Première partie

Techniques algébriques, à travers l'histoire

Chapitre 1

Calculs polynomiaux

Résumé -

Nous commençons l'année en revisitant l'histoire de la résolution des équations polynomiales. On donne/rapelle ainsi quelques bases essentielles pour la suite : existe-t-il un moyen pour résoudre toute équation polynomiale, on verra que l'important est de pouvoir factoriser; que nous dit la géométrie des équations polynomiales; que nous dit l'analyse pour une résolution (approchée) d'une équation polynomiale.
Ce sera l'occasion de mettre en place quelques définitions et de nombreux savoir-faire. Comme chaque chapitre, celui-ci commence par des problèmes ouverts.

Enfin, voici une liste de petites vidéo ou conférence visionnable sur internet et en lien avec le sujet. A visionner à loisir :

- *Mathieu Bautista - L'histoire du x. https ://www.youtube.com/watch ?v=AW7uNg9RLCs*
- *MicMath - Conique à la plage. https ://www.youtube.com/watch ?v=eFPhYYKCyFc*
- *El Jj - Différence équations et fonctions. https ://www.youtube.com/watch ?v=sJKjFgtIBKY*

Sommaire

1. Quelques problèmes

1.1. Problèmes

? Problème 1 - Kwarizmi

Résoudre, comme El Kwarizmi, l'équation $x^2 + 10x = 39$ en n'exploitant que des méthodes géométriques (calcul d'aire, nombres positifs...).
On commence par considérer un carré de côté x et deux rectangles de côtés x et 5. On complète...
Adapter la méthode pour résoudre $x^2 + 21 = 10x$.
Que pensez-vous de la nature des nombres a, b ou c si l'on souhaite générer la méthode pour résoudre l'équation associée au trinôme du second degré : $ax^2 + bx + c = 0$

? Problème 2 - Mauvais vers italiens

« *Tartalea exposa sa solution en mauvais vers italiens* » (Lagrange, Oeuvres, 1795).
Il existe une méthode (Tartglia-Cardan) pour résoudre les équations de degré 3 (voir cours). Comment pouvait-on l'écrire alors qu'il n'existait ni le symbole =, ni les notations x^2...?

? Problème 3 - Tartaglia et Cardan

Trouver toutes les solutions de l'équation $x^3 + 6x = 20$, avec la méthode de Tartaglia et Cardan, en posant $x = u - v$ et en exploitant les symétries du problème (on peut résoudre : $u^3 - v^3 = \ldots$ et $u^3 v^3 = \ldots$).
Et pour une équation de type $ax^3 + px + q = 0$?
De même donner les solutions de l'équations $x^4 + px^2 + 1 = 0$.

? Problème 4 - Polynôme de degré 2 et représentation graphique

Représenter la courbe d'équation $y = ax^2 + bx + c$.
On fera la différence entre $a > 0$ et $a < 0$. On notera en particulier les coordonnées du sommet

? Problème 5 - Chercher à côté...

Si on sait que $f(x_0) = 0,001$, quelle stratégie mettre en place pour trouver une racine de f ? On peut chercher du côté de x_0, pas très loin...

? Problème 6 - Formule de Taylor et développement limité

Soit $f : x \mapsto a_0 + a_1 x + \cdots + a_n x^n$, une fonction polynomiale de degré n.
Pour tout $k \in \mathbb{N}$, exprimer $f^{(k)}(0)$ en fonction des nombres (a_i) qui définissent la fonction polynomiale.
$f^{(k)}(0)$ *est la valeur en* 0 *de la* k^e *dérivée de la fonction* f.

1.2. Vocabulaires et contextes

Principe : On appelle équation, une relation calculatoire avec des objets inconnues i.e. à expliciter, définies à partir de cette (ces) relation(s) calculatoire(s).

Définition - Résoudre une équation
Soit E un ensemble (souvent $E=\mathbb{R}$ ou $\mathbb{C}$ ou $\mathbb{R}^p\ldots$, mais pas uniquement).
Soit une application $f:E\to F$ et $b\in F$(= $\mathbb{R}$ ou $\mathbb{C}$ ou autre).
On dit qu'on résout l'équation $f(x)=b$, d'inconnue $x\in E$, lorsqu'on trouve tous les « nombres » $x\in E$ tel que $f(x)=b$.

On parle d'antécédent de b par f

Remarque - Il n'existe qu'une méthode infaillible pour résoudre une équation : Faire l'essai.
Soit $x_0\in E$. Alors (calcul) : $f(x_0)=\ldots$.
Si la réponse est b, on a trouvé une solution. Sinon, on essaye à nouveau.
Mais cela est souvent long, surtout si l'ensemble E est infini…
Autre question : pourquoi se concentrer (uniquement?) sur des équations polynomiales?

Pour aller plus loin - Des tortues à l'infini
Les mots application (ou fonction), $\mathbb{R}$, $\mathbb{C}$… seront définis (légalisés) plus tard dans l'année ou plus loin dans le polycopié.

2. Equation polynomiale. Algèbre et géométrie

La motivation historiquement première est la motivation ludique. La plupart des résultats ici a été obtenu dans le cadre de joutes mathématiques, de défis si l'on préfère.
Les révolutions successives qui marquent la mathématiques européennes consistent souvent en l'installation de nouvelles notations. Cela crée des ponts!

2.1. Révolution 1 : Viète

Algebra Nova

Histoire - Algèbre spécieuse
MONTUCLA *dans Histoire des mathématiques : « Il est peu de mathématiciens à qui l'algèbre doive plus qu'à cet homme célèbre… On doit d'abord à M. Viète d'avoir établi l'usage des lettres pour désigner, non seulement les quantités inconnues, mais même celles qui sont connues, ce qui fit donne à son algèbre le nom de spécieuse, nom qu'elle a gardé longtemps, à cause que tout y est représenté par des symboles…*
C'est nous osons le dire à ce changement que l'algèbre est redevable d'une grande partie de ces progrès. »
Comment énoncer aisément la règle du discriminant sans les lettres a, b et c ?

Heuristique - Généraliser avec des lettres
FRANÇOIS VIÈTE a l'idée fondamentale d'écrire des lettres $A,B,C,\ldots X$ pour les inconnues et les données d'un problème (ie. les connus!) et de faire les calculs algébriques. Dès lors, aucun problème des anciens Grecs ne semble résister au schéma :

$$\text{Problème géométrique} \xrightarrow{\text{mêmes lettres}} \text{Problème algèbrique} \xrightarrow{\text{calculs}} \text{Solution}$$

et Viète écrit en majuscules : « *NVLLVM NON PROBLEMA SOLVERE* ».

Saut historique et définition

Définition - Fonction polynomiale (d'une variable)
On dit que $f:\mathbb{R}\to\mathbb{R}$ est une application polynomiale s'il existe $n\in\mathbb{N}$ et $a_0,a_1,\ldots a_n\in\mathbb{R}$ (voire $\mathbb{C}$) tels que :

pour tout réel x, $f(x)$ est obtenu par le calcul $\sum_{k=0}^{n} a_k x^k$.

On note $f:x\mapsto\sum_{k=0}^{n} a_k x^k$.

Remarque - Un unique type de calcul, pour tout x
C'est toujours le même calcul : une combinaison linéaire des puissances de x qui donne le valeur de $f(x)$.
Cette définition s'étend au cas de plusieurs variables (à avoir en tête en particulier pour les manipulations)

Définition - Polynôme de plusieurs variables
On appelle fonction polynomiale de p variables (abrégé ici en « polynôme de p variables ») une fonction de la forme :

$$f_p : (x_1, x_2, \dots x_p) \mapsto \sum_{(k_1,\dots k_p)\in\mathbb{N}^p} a_{k_1,k_2,\dots k_p} \prod_{i=1}^{p} x_i^{k_i}$$

la somme étant finie (nombre de termes est fini) où $\forall\ (k_1,\dots k_p) \in \mathbb{N}^p$, $a_{k_1,k_2,\dots k_p} \in \mathbb{R}$ (ou $\mathbb{C}$) ;
Il s'agit d'une combinaison linéaire finie de puissances entières des inconnues réelles (ou complexes) $x_1,\dots x_p$.
La somme et le produit de deux fonctions polynomiale est une fonction polynomiale.
On appelle degré de f_p, le nombre $\max\{k_1+k_2+\dots+k_p \mid a_{k_1,k_2,\dots k_p} \neq 0\}$. Si f n'a qu'une variable, $\deg(f) = \max\{n \in \mathbb{N}, \text{ tel que } a_n \neq 0\}$.

Exemple - $f : (x,y,z) \mapsto 3x^2y - 2xyz - xz^2$
Il s'agit d'une fonction polynomiale de trois variables, de degré 3.

Histoire - François Viète

François Viète (1540-1603), n'est pas un mathématicien professionnel, mais un avocat. Néanmoins, il sera le représentant français des joutes mathématiques (cryptanalyse) avec les italiens ou les anglais. Il est célèbre pour son algèbre nouvelle où il est le premier à exploiter les lettres pour décrire les nombres dans des équations. Ce point de vue est révolutionnaire!

Vocabulaire

On donne un peu de vocabulaire et une notation bien pratique, même si historiquement cela s'est fait beaucoup plus tardivement...

Analyse - Unicité d'écriture
Considérons deux fonctions polynomiales $f : x \mapsto \sum_{k=0}^{n} a_k x^k$ et $g : x \mapsto \sum_{k=0}^{m} b_k x^k$.
Supposons qu'il s'agisse bien de deux écritures différentes, donc $\{h \in [\![0, \max(n,m)]\!] \mid a_h \neq b_h\}$ est non vide.
Comme tout ensemble inclus dans $\mathbb{N}$, fini (ou majoré) et non vide, il admet un plus grand élément, notons le h.
On a donc pour tout $x \in \mathbb{R}$:

$$f(x) - g(x) = (a_h - b_h)x^h + \sum_{k=0}^{h-1} (a_k - b_k)x^k$$

On veut montrer que $f \neq g$, ie : $\exists\ x_0 \in \mathbb{R}$ tel que $f(x_0) \neq g(x_0)$.

- Méthode 1 (réflexes de terminale) :
  $\lim\limits_{x\to+\infty} f(x)-g(x) = \text{signe}(a_h-b_h)\infty$ et $f-g$ est continue, donc nécessairement ie : $\exists\ x_0 \in \mathbb{R}$ tel que $f(x_0) - g(x_0) \neq 0$.
- Méthode 2 (même chose mais redémontré). On note $c_k = a_k - b_k$.

  $$\frac{f(x)-g(x)}{x^h} = a_h - b_h + \sum_{k=0}^{h-1} \frac{c_k}{x^{h-k}}$$

  Pour tout $k \in [\![0, h-1]\!]$, $\dfrac{c_k}{x^{h-k}} \underset{x\to+\infty}{\longrightarrow} 0$.
  Il existe donc $X_k \in \mathbb{R}_+^*$ tel que $\forall\ x \geqslant X_k$, $\left|\dfrac{c_k}{x^{h-k}}\right| \leqslant \dfrac{|a_h-b_h|}{2h}$.
  Il suffit de prendre $X_k = \sqrt[h-k]{2h\frac{|c_k|}{|a_h-b_h|}}$ Et donc pour $x \geqslant X := \max(X_0, X_1, \cdots X_{h-1})$ (inégalité triangulaire) :

  $$\left|\frac{f(x)-g(x)}{x^h} - (a_h-b_h)\right| = \left|\sum_{k=0}^{h-1} \frac{c_k}{x^{h-k}}\right| \leqslant \sum_{k=0}^{h-1} \frac{|a_h-b_h|}{2h} = \frac{|a_h-b_h|}{2}$$

Ainsi, comme $|\alpha| \leqslant \beta \iff -\beta \leqslant \alpha \leqslant \beta$, on a donc pour $x \geqslant X$:

$$-\frac{|a_h - b_h|}{2} \leqslant \frac{f(x)-g(x)}{x^h} - (a_h - b_h) \leqslant \frac{|a_h - b_h|}{2}$$

$$\text{Et donc si } a_h - b_h > 0 : \frac{f(x)-g(x)}{x^h} \geqslant -\frac{a_h - b_h}{2} + (a_h - b_h) = +\frac{a_h - b_h}{2} > 0$$

$$\text{Et si } a_h - b_h < 0 : \frac{f(x)-g(x)}{x^h} \leqslant \frac{b_h - a_h}{2} + (a_h - b_h) = -\frac{b_h - a_h}{2} < 0$$

Ainsi, par contraposée : si deux écritures polynomiales donne une même fonction, c'est nécessairement la même écriture!

Définition - Degré et $[f]_k$
Soit f une application polynomiale. Alors son écriture est unique.
Ainsi, il existe un unique $n \in \mathbb{N}$ et un unique $(n+1)$-uplet de réels $(a_0, a_1, \dots, a_n)$ tels que $f : x \mapsto \sum_{k=0}^{n} a_k x^k$ avec $a_n \neq 0$.
Le nombre entier n s'appelle le degré de la fonction polynomiale f.
$a_n x^n$ s'appelle le terme dominant de la fonction polynomiale f.
Et on aura pour habitude de noter $[f]_k$, le k coefficient a_k (devant x^k) dans l'écriture de f.

2.2. Révolution 2 : Descartes

Heuristique - Algèbre et géométrie
La motivation est ici GEOMETRIQUE. Une nouvelle façon de concevoir le problème est de maintenant lui donné un sens géométrique et en particulier de donner à $\mathbb{R}$ le sens du continu (comme une droite) et à $f(\mathbb{R})$ une déformation continue de $\mathbb{R}$.

Pour aller plus loin - Notations
Les notations x^5... a été popularisé par Descartes (avant on écrivait simplement xxxxx).
Le symbole = a été popularisé par Leibniz (avant on écrit adeq).

Représentation graphique

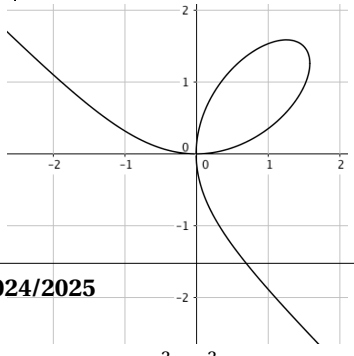
En mathématique, l'apport principal de Descartes a consisté à donner une vision géométrique à l'algèbre et une approche calculatoire à la géométrie (de l'ordre de la précision du langage). C'est une véritable révolution.
Ce qui suit reste vrai même si H n'est pas poiynomiale.

Définition - Graphe dans $\mathbb{R}^2$
Soit $H : \mathbb{R}^2 \to \mathbb{R}$.
On note $\Gamma_H = \{(x,y) \in \mathbb{R}^2 \mid H(x,y) = 0\}$, une sous-partie de $\mathbb{R}^2$.
On dit que $H(x,y) = 0$ est une équation du graphe Γ.

Sans condition supplémentaire sur H, Γ peut être très variés.

Exemple - Folium de Descartes
DESCARTES et ROBERTVAL ont étudié dans leur correspondance la courbe $\mathscr{F}$ d'équation $x^3 + y^3 - 3xy = 0$ (cubique).
Sa représentation graphique est dans la marge.

Représentation - Folium



$\mathscr{F}$ d'équation $x^3 + y^3 - 3xy = 0$

Définition - Exemple. Graphe d'une fonction
Si $f : \mathbb{R} \to \mathbb{R}$, on appelle representation graphique de f (ou graphe de f), la courbe $\mathscr{C}$ d'équation $y = f(x)$

Dans ce cas $H : (x, y) \mapsto y - f(x)$. Une telle courbe ne peut « revenir en arrière ». En effet, cela signifierait qu'un nombre x_0 aurait deux images.

Représentation des fonctions polynomiales...

Heuristique - Réprésentation

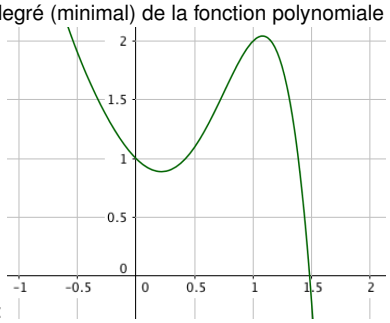
La représentation d'une fonction polynomiale $x \mapsto \sum_{k=0}^{n} a_k x^k$ est continue.
Cette fonction polynomiale est de degré n, donc admet au plus n racines (de l'équation polynomiale).
Par dérivation (que l'on expliquera plus loin), il y a également au plus n sens de variation différents...

Pour aller plu
Nous reverrons
nuité.
Vous pouvez déj
nition formalisé

Exercice
Donner l'exemple d'une fonction polynomiale de degré 4 n'ayant que deux sens de variations différentes

Correction
$x \mapsto x^4$, tout simplement

Exercice
Quel est le degré (minimal) de la fonction polynomiale dont la représentation graphique est



donnée par :

Correction
C'est un polynôme de degré impair.
Le terme dominant est nécessairement négatif (dominant en $\pm\infty$).
Il y a 3 variations, donc le degré est au moins 3.
(Ici, il s'agit de $x \mapsto x^5 + x^3 + 2x^2 - x + 1$.)

Calculs algébriques et symétries géométriques

Proposition - Translation
La courbe Γ présente une invariance par translation de vecteur $\vec{u} = T\vec{i}$, si et seulement si, on a l'équivalence : $H(x, y) = 0 \iff H(x + T, y) = 0$

Remarque - Qui sont x et y ?

Est-ce pour tout x, y ou avec l'existence d'un x et d'un y ?
Cela doit être vrai pour x et y dans l'ensemble de définition, puis $x + T$ doit y appartenir également...

Savoir faire - Graphe d'une fonction $y = f(x)$ (translation)

Dans le cas particulier d'une fonction, $\mathscr{C}_f$ présente d'une invariance par translation de vecteur $\vec{u} = T\vec{i}$ si et seulement si $\forall\, x \in \mathscr{D}_f,\ f(x + T) = f(x)$. On dit que f est T-périodique.
En particulier cos ou sin sont 2π-périodique.

Analyse - Symétrie axiale, centrale

Si la courbe présente une symétrie axiale autour de l'axe $x = x_0$.
Alors on a $H(x_0 + x, y) = 0 \iff H(x_0 - x, y) = 0$.
En posant $x' = x_0 + x$, on trouve $x_0 - x = x_0 - (x' - x_0) = 2x_0 - x'$.
Ce qui donne l'équivalence suivante : $H(x, y) = 0 \iff H(2x_0 - x, y) = 0$.
La réciproque est vérifiée.

Proposition - Symétrie axiale
La courbe Γ présente une symétrie axiale d'axe $x = x_0$,
si et seulement si, on a l'équivalence : $H(x, y) = 0 \iff H(2x_0 - x, y) = 0$

Savoir faire - Graphe d'une fonction $y = f(x)$ (symétrie axiale)

Dans le cas particulier d'une fonction, $\mathscr{C}_f$ présente d'un symétrie axiale d'axe $x = x_0$ si et seulement si $\forall\ x \in \mathscr{D}_f$, $f(2x_0 - x) = f(x)$.
En particulier pour $x_0 = 0$, on a la caractérisation $f(-x) = f(x)$. On dit alors que f est paire (comme $x \mapsto x^2$).

Une symétrie centrale de centre $M_0(x_0, y_0)$ est la composition d'une symétrie d'axe $x = x_0$ et d'axe $y = y_0$.

Proposition - Symétrie centrale
La courbe Γ présente une symétrie centrale de centre $M(x_0, y_0)$,
si et seulement si, on a l'équivalence : $H(x, y) = 0 \iff H(2x_0 - x, 2y_0 - y) = 0$

Savoir faire - Graphe d'une fonction $y = f(x)$ (symétrie centrale)

Dans le cas particulier d'une fonction, $\mathscr{C}_f$ présente d'un symétrie centrale centrée en $M_0(x_0, y_0)$ si et seulement si $\forall\ x \in \mathscr{D}_f$, $f(2x_0 - x) = 2y_0 - f(x)$.
En particulier pour $x_0 = y_0 = 0$, on a la caractérisation $f(-x) = -f(x)$. On dit alors que f est impaire (comme $x \mapsto x^3$).

Une astuce pour le calcul

De manière générale, on ne démontre pas un résultat de calcul par une exploitation graphique, mais cela permet largement de vérifier une série de calculs.

Truc & Astuce pour le calcul - Transformer le résultat d'un calcul en une représentation visuelle

Depuis Descartes, l'algèbre et la géométrie sont totalement liés et il est possible de passer « du diable de l'algèbre à l'ange de la géométrie » (Hermann Weyl).
Il est donc important de savoir faire cette transformation : donner du sens géométrique à un calcul algébrique. On peut ainsi anticiper ou vérifier un résultat.

Exercice

Montrer que le système $\begin{cases} x^2 + y^2 + 2x - 2y &= 0 \\ x^2 + y^2 - 4y + 3 &= 0 \end{cases}$ admet exactement deux solutions.

Correction

Ces équation s'écrivent : $(x+1)^2 + (y-1)^2 = 2$ et $x^2 + (y-2)^2 = 1$ Il s'agit de l'intersection du cercle de centre $(-1, 1)$ de rayon $\sqrt{2}$ (qui passe par O) et du cercle de centre $(0, 2)$ et de rayon 1.
Le changement de registre ici est une force si on sait bien l'employer, mais aussi un gros problème pour les élèves bloqués dans leur registre.

Truc & Astuce pour le calcul - Avec des fonctions

Le calcul sur les fonctions est à l'intersection du calcul algébrique, du calcul graphique, du calcul différentiel et intégrale, du calcul de limites...
Les problèmes où interviennent ces fonctions sont aussi très variés, et il n'est pas rare de voir des changements de registre, d'un domaine à l'autre dans une même problème! Il faut avoir un esprit bien souple...
C'est tout particulièrement le cas de l'étude des polynômes (où l'on bascule facilement d'un domaine à l'autre).

Exercice

Démontrer que le polynôme x^3+3x-1 admet une unique racine sur $\mathbb{R}$

Correction

Il ne faut surtout pas factoriser, mais faire l'étude de la fonction $f : x \mapsto x^3+3x-1$, dérivable et finalement strictement croissante. Elle admet une unique racine sur $\mathbb{R}$ (théorème de la bijection).

3. Opérer (=calculer) avec des polynômes

3.1. Développer

Stabilité

On commence par le développement, c'est a priori plus simple car c'est une opération *mécanique*.
A ce stade, c'est surtout la pratique qui justifie le théorème suivant :

Théorème - Stabilité

L'addition, la soustraction et la multiplication de deux (ou plus) fonctions polynomiales donne une nouvelle fonction polynomiale.

Exercice

Soient $f : x \mapsto 2+x-x^2+3x^3$ et $g : x \mapsto 1+x-x^2+x^4$ définies sur $\mathbb{R}$.
Evaluer, pour tout $x \in \mathbb{R}$, $(f+g)(x)$ et $(f \times g)(x)$

Correction

$f+g : x \mapsto 3+2x-2x^2+3x^3+x^4$.
$f \times g : x \mapsto 2+3x-2x^2+x^3+6x^4$.

Un peu plus théorique :

Exercice

Soient $f : x \mapsto a_0+a_1x+\cdots a_nx^n$ et $g : x \mapsto b_0+b_1x+\ldots b_mx^m$ définies sur $\mathbb{R}$.

1. Evaluer, pour tout $x \in \mathbb{R}$, $(f+g)(x)$ et $(f \times g)(x)$.
2. Exprimer alors, pour tout $k \in \mathbb{N}$, $[f+g]_k$ en fonction des $[f]_i$ et $[g]_j$.
3. Exprimer alors, pour tout $k \in \mathbb{N}$, $[f \times g]_k$ en fonction des $[f]_i$ et $[g]_j$.

Correction

1.
2. On trouve $[f+g]_k = [f]_k + [g]_k$, et de manière générale, cette opération est linéaire.
3. On a le produit de Cauchy : $[f \times g]_k = \sum_{i=0}^{k} [f]_i [g]_{k-i} = \sum_{i+j=k} [f]_i [g]_j$, valable même si les coefficients sont nuls.

Développer malin

Mais il existe plusieurs façons de développer un produit polynomial. Certaines sont plus intelligentes que d'autres...

Exercice

Développer $(a+b+c)^3$

Correction

La solution sera de la forme $\sum A_{i,j,k} a^i b^j c^k$ avec $i+j+k=3$. Reste à trouver la valeur de $A_{i,j,k}$.

Pour obtenir, par exemple, le nombre $A_{1,1,1}$, il faut prendre un a, un b et un c, il y a $3\times2\times1=3!=6$ possibilités.
Pour obtenir, par exemple, le nombre $A_{1,2,0}$, il faut prendre un a, deux b et 0 c, il y a $3\times1\times1=3=3$ possibilités.
Ensuite il y a une symétrie entre a, b et c qui commutent : $A_{1,2,0}=A_{0,2,1}=\cdots=A_{0,1,2}$.
Finalement : $(a+b+c)^3=6abc+3(a^2b+a^2c+b^2c+b^2a+c^2a+c^2b)$

Exercice

On admet que $(a+b)^4=a^4+4a^3b+6a^2b^2+4ab^3+b^4$.
En organisant convenablement votre calcul $(a+b)^5$, trouver comment passer des coefficients (1,4,6,4,1) de $(a+b)^4$ à ceux de $(a+b)^5$.
Cela vous rappelle-t-il quelque chose ? En déduire la formule générale qui donne une expression de $(a+b)^n$.

Correction

$$\begin{array}{rccccccc}(a+b)^5=(a+b)^4(a+b)=a(a+b)^4+b(a+b)^4&&&&&&&\\ =&a^5&+4a^4b&+6a^3b^2&+4a^2b^3&+ab^4&&\\ &&+a^4b&+4a^3b^2&+6a^2b^3&+4ab^4&+b^5&\\ =&(1+0)a^5&+(4+1)a^4b&+(6+4)a^3b^2&+(4+6)a^2b^3&+(1+4)ab^4&+(0+1)b^5&\\ =&a^5&+5a^4b&+10a^3b^2&+10a^2b^3&+5ab^4&+b^5&\end{array}$$

On retrouve le triangle de Pascal : de $(1,4,6,4,1)$ à $(1,5,10,10,5,1)$... Cela ne fait pas une démonstration mais permet d'anticiper, comprendre, retenir la formule du binôme de Newton :

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

Truc & Astuce pour le calcul - Anticipation

Il s'agit de ne plus être à chaque instant derrière son calcul, mais bien en avant !. Il s'agit bien là aussi de voir quelque chose... Lorsque le calcul demandé est ouvert (on ne donne pas une forme fermée : montrer que $A=B$), il faut savoir vers où l'on va.
Par exemple, lorsqu'on dérive une fonction, ce qui nous intéresse souvent c'est de connaître son signe. Il faut donc donner une forme factorisée...

Exercice

On peut exprimer de 3 façons différentes $A=f(x)=(3x^2+8x-1)-(x^2+3x-4)$. Associer chacune de ces expressions à l'exploitation qu'on peut en faire.

$A=2x^2+5x-3$	étude du minimum de f
$A=-6+3(x+1)+2(x+1)^2$	développement limité de f en -1
$A=2\left(x+\frac{5}{4}\right)^2-\frac{49}{8}$	étude du signe de f
$A=(x+3)(2x-1)$	étude du polynôme f

Correction

$A=2x^2+5x-3$	$\longrightarrow$	étude du polynôme f
$A=-6+3(x+1)+2(x+1)^2$	$\longrightarrow$	développement limité de f en -1
$A=2\left(x+\frac{5}{4}\right)^2-\frac{49}{8}$	$\longrightarrow$	étude du minimum de f
$A=(x+3)(2x-1)$	$\longrightarrow$	étude du signe de f

Truc & Astuce pour le calcul - Reconnaissance de formes

En visualisant les formes dans les formules, il est plus aisé de garder en mémoire le calcul effectué (pour le retro-contrôle) et surtout, il est plus aisé de savoir dans quel ordre faire le calcul de manière à être efficace : ne pas se tromper et agir rapidement.

Exercice

Démontrer :
$4[(a^2-b^2)cd+(c^2-d^2)ab]^2+[(a^2-b^2)(c^2-d^2)-4abcd]^2=(a^2+b^2)^2(c^2+d^2)^2$
On pourra y voir la forme $A=B^2-C^2$...

Correction

Notons $B=(a^2+b^2)(c^2+d^2)$ et $C=(a^2-b^2)(c^2-d^2)-4abcd$. Alors
$B^2-C^2=(B-C)(B+C)$
$=\left[(a^2+b^2)(c^2+d^2)-(a^2-b^2)(c^2-d^2)+4abcd\right]\left[(a^2+b^2)(c^2+d^2)+(a^2-b^2)(c^2-d^2)-4abcd\right]$
$=\left[2a^2d^2+2b^2c^2+4abcd\right]\left[2a^2c^2+2b^2d^2+4abcd\right]=4\left[(ad+bc)^2\right]\left[(ac-bd)^2\right]$
On reconnait une fois E^2+F^2+2EF et une fois F^2+G^2-2FG. Donc
$B^2-C^2=4\big((ad+bc)(ac-bd)\big)^2=4(a^2dc+c^2ab-b^2cd-d^2ab)^2=4[(a^2-b^2)cd+(c^2-d^2)ab]^2$

Attention - Ne pas trop écrire

Pour apprendre à se projeter vers l'avant, il faut ne pas écrire trop de calculs intermédiaires. De nombreuses petites réécritures doivent être simplement pensées, sans être écrites.
Le prix à payer : une insécurité forte pour l'élève.
Le prix à gagner : une plus grande concentration et une meilleure vitesse d'exécution !
Deux ans avant les concours, il faut investir dans cette stratégie.

Truc & Astuce pour le calcul - Garder en mémoire « vive » le calcul

Si on garde en mémoire immédiate le calcul, il est possible de déceler les erreurs plusieurs lignes de calcul plus loin. On évite des erreurs bêtes, comme des signes + et − qui se mélangent, une page qui se tourne et qui conduit à des nombres qu'on oublie…
Pour apprendre à exploiter une mémoire globalisante du calcul, on peut essayer après chaque calcul à réécrire le résultat obtenu sur une page blanche. Evidemment, un tel résultat doit rester en mémoire immédiate, il n'est pas nécessaire de le placer en mémoire de travail plus profonde.

Exercice
Quel est le résultat obtenu lors du dernier calcul que vous avez effectué ?

Correction

3.2. Factoriser

Avec les petits Bernoullis

Analyse - $x^k - a^k$

Considérons une inconnue réelle x et une connue a.
$(x-a)(x^4+x^3a+x^2a^2+xa^3+a^4) = x^5+x^4a+x^3a^2+x^2a^3+xa^4-x^4a-x^3a^2-x^2a^3-xa^4-a^5 = x^5-a^5$.
Soit $k \in \mathbb{N}^*$.
Calculons

$$\begin{aligned}(x-a)\sum_{i=0}^{k-1} x^i a^{k-1-i} &= (x-a)\times\left(x^{k-1}+x^{k-2}a+x^{k-3}a^2+\ldots xa^{k-2}+a^{k-1}\right)\\ &= x^k - x^{k-1}a + x^{k-1}a - x^{k-2}a^2 + \cdots + x^2a^{k-2} - xa^{k-1} + xa^{k-1} - a^k = x^k - a^k\end{aligned}$$

Exercice
Comment rendre ce calcul rigoureux ?

Correction
Avec des suites géométriques, ou par récurrence (voir plus bas) ou par télescopage (cf chapitre sur les sommes)

Pour aller plus loin - Polynômes de Bernoulli

On appelle (ici en MPSI3) cette famille de polynômes, les « petits Bernoullis », pour ne pas les confondre.

Définition - Les petits Bernoullis

Pour tout $n \in \mathbb{N}$ et $a \in \mathbb{K}$, on note $b_a^n : x \mapsto x^n + ax^{n-1} + \ldots a^{n-1}x + a^n = \sum_{i=0}^{n} x^i a^{n-i}$.
On appelle cette famille de fonctions polynomiales $(b_a^n)_{n\in\mathbb{N}}$, les petits Bernoullis.
On a alors :

$$\forall\, x \in \mathbb{K}, \qquad (x-a)\times b_a^n(x) = x^{n+1} - a^{n+1}$$

Il faut faire une démonstration. On exploite le résultat sur les sommes des termes consécutifs d'une suite géométrique, vu en terminale.

Démonstration
Notons que pour $x \neq a$:

$$b_a^n(x) = \sum_{i=0}^{n} x^i a^{n-i} = a^n \sum_{i=0}^{n} \left(\frac{x}{a}\right)^i = a^n \frac{1-\left(\frac{x}{a}\right)^{n+1}}{1-\frac{x}{a}} = a\frac{a^n - \frac{x^{n+1}}{a}}{a-x} = \frac{x^{n+1}-a^{n+1}}{x-a}$$

Donc pour $x \neq a$: $(x-a)b_a^n(x) = x^{n+1} - a^{n+1}$.
Et si $x = a$: on a aussi $(x-a)b_a^n(x) = 0 = x^{n+1} - a^{n+1}$. □

Exercice
En notant que $b_a^{n+1} = ab_a^n + x^{n+1}$ (à démontrer), montrer par récurrence la factorisation de Bernoulli

---

Correction
Notons, pour tout $n \in \mathbb{N}$, $\mathscr{P}_n$: « $\forall\, x \in \mathbb{K}$, $(x-a)b_a^n(x) = x^{n+1} - a^{n+1}$. »
- $(x-a)b_a^0(x) = (x-a)1 = x-a$. Donc $\mathscr{P}_0$ est vraie.
- Soit $n \in \mathbb{N}$. On suppose $\mathscr{P}_n$ vraie : $\forall\, x \in \mathbb{K}$, $(x-a)b_a^n(x) = x^{n+1} - a^{n+1}$.
  Pour tout $x \in \mathbb{K}$, $(x-a)b_a^{n+1}(x) = (x-a)(x^{n+1} + ab_a^n) = x^{n+2} - ax^{n+1} + a(x^{n+1} - a^{n+1}) = x^{n+2} - a^{n+2}$.
  Donc $\mathscr{P}_{n+1}$ est vraie.

Grâce à une racine

On a les corollaires important suivant :

Proposition - Factorisation (1)
Soit f une fonction polynomiale de degré n.
Pour tout $a \in \mathbb{K}$, il existe g_a, fonction polynomiale de degré $n-1$ tel que

$$\forall\, x \in \mathbb{K}, \qquad f(x) - f(a) = (x-a)g_a(x)$$

Démonstration
Supposons que $f(x) = \sum_{k=0}^{n} c_k x^k$ avec $c_n \neq 0$.
Alors

$$\begin{aligned} f(x) - f(a) &= \sum_{k=0}^{n} c_k x^k - \sum_{k=0}^{n} c_k a^k = \sum_{k=1}^{n} c_k (x^k - a^k) + c_0 - c_0 \\ &= \sum_{k=1}^{n} c_k (x-a) b_a^{k-1}(x) = (x-a) \times \underbrace{\sum_{k=1}^{n} c_k b_a^{k-1}(x)}_{=g_a(x)} \end{aligned}$$

On note bien que g_a est une fonction polynomiale. Le terme x^{n-1}, apparait avec le coefficient $c_n \neq 0$ dans b_a^{n-1} uniquement. Aucun monôme de degré plus élevé n'apparait dans la combinaison linéaire. □

Corollaire - Factorisation (2)
Soit f une application polynomiale sur $\mathbb{K}$(=$\mathbb{R}$ ou $\mathbb{C}$) de degré n.
Soit $x_0 \in \mathbb{K}$ tel que $f(x_0) = 0$.
Alors il existe g, application polynomiale de degré $n-1$ telle que

$$\underbrace{\text{pour tout } x \in \mathbb{K},}_{\forall\, x \in \mathbb{K}} \qquad f(x) = (x - x_0) \times g(x)$$

Savoir que g existe est une excellente chose.
Mais savoir comment obtenir g connaissant x_0 et f est encore mieux (sans développer tous les petits Bernoullis) !
On peut donc décrire un algorithme pour obtenir g :

Pour aller plus loin - Division euclidienne
En fait on obtient ici un algorithme de division euclidienne. Mais uniquement par un polynôme de degré 1.
On reprendra cela plus tard.

Savoir faire - Factoriser

Notons $f : x \mapsto x^3+2x^2-x-2$.
Alors $f(1) = 1+2-1-2 = 0$. Donc $f(x) = (x-1) \times g(x)$.
Pour appliquer l'algorithme, on prend l'habitude d'écrire le plus à gauche le terme sur lequel on agit en premier (comme pour une division euclidienne entière), on écrit donc les polynômes dans le sens des puissances décroissantes :

$$\begin{array}{llll|l} x^3 & +2x^2 & -x & -2 & x-1 \\ \hline & (2+1)x^2 & & & 1x^2+3x+2 \\ & & (-1+3)x & & \\ & & & -2+2 & \end{array}$$

Donc $x^3+2x^2-x-2 = (x-1)(x^2+3x+2)$.
On oublie jamais de **vérifier le calcul réciproque** s'il est beaucoup plus simple !

Trouver toutes les racines

La factorisation permet de décomposer un problème en plusieurs sous-problèmes.

Proposition - Factorisation
Soient f et g deux fonctions polynomiales à valeurs dans $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$. Soit $x \in \mathbb{K}$.
$f(x) \times g(x) = 0$ si et seulement si $f(x) = 0$ ou $g(x) = 0$

Pour ce genre de proposition, on fait un raisonnement en deux temps (double implication) :

Pour aller plus loin - Anneaux intègres
Un ensemble d'élément neutre 0 *pour une première loi* + *et vérifiant :*

$$a \times b = 0 \Longrightarrow a = 0 \text{ ou } b = 0$$

est appelé ensemble intègre.
$\mathbb{R}$ *ou* $\mathbb{C}$ *sont intègres. Ce n'est pas le cas de* $\mathcal{M}_n(\mathbb{R})$.

Démonstration
Si $f(x) = 0$ ou $g(x) = 0$, alors $f(x) \times g(x) = 0$, car l'un des deux termes est nuls.
Réciproquement. Si $f(x) \times g(x) = 0$, alors car $\mathbb{K}$ est intègre : $f(x) = 0$ ou $g(x) = 0$ □

Exercice
Trouver toutes les racines réelles de l'équation $x^4 - 5x^2 + 4 = 0$

Correction
On note $f(x) = x^4 - 5x^2 + 4$.
On a $f(1) = 1-5+4 = 0$, et comme f est paire (bi-carrée), $f(-1) = 0$.
$f(x) = (x-1)(x^3+x^2-4x-4) = (x-1)(x+1)(x^2-4) = (x-1)(x+1)(x-2)(x+2)$.
Un produit de nombres réels est nul si et seulement si l'un des facteurs est nul ($\mathbb{R}$ intègre).

$$f(x) = 0 \Longleftrightarrow x = 1 \text{ ou } x = -1 \text{ ou } x = 2 \text{ ou } x = -2$$

Théorème - Factorisation multiple
Soit f une fonction polynomiale de degré n. Soit $p \leqslant n$
Si $x_1, x_2, \dots x_p$, p solutions différentes de l'équation $f(x) = 0$ (racines de f).
Alors il existe g_p, fonction polynomiale de degré $n-p$ tel que

$$\forall\, x \in \mathbb{K}, \qquad f(x) = (x-x_1)(x-x_2)\dots(x-x_p) \times g_p(x)$$

Démonstration
On fait une récurrence finie sur $p \leqslant n$.
— $\mathscr{P}_0$ (et $\mathscr{P}_1$) est vraie.
— Soit $p < n$. Supposons que $\mathscr{P}_p$ est vraie.
Soient $x_1, x_2, \dots x_p, x_{p+1}$ $p+1$ solutions de $f(x) = 0$.
Alors d'après $\mathscr{P}_p$. Il existe g_p de degré $n-p$ tel que

$$\forall\, x \in \mathbb{K}, \qquad f(x) = (x-x_1)(x-x_2)\dots(x-x_p) \times g_p(x)$$

Puis $f(x_{p+1}) = 0 = (x_{p+1}-x_1)(x_{p+1}-x_2)\dots(x_{p+1}-x_p)\times g_p(x_{p+1})$.
Comme $x_{p+1} \neq x_i$, alors $x_{p+1}-x_i \neq 0$ et nécessairement $g_p(x_{p+1}) = 0$.
Et donc il existe g_{p+1} de degré $n-p-1$ tel que

$$\forall\, x \in \mathbb{K}, \quad g_p(x) = (x-x_{p+1})\times g_{p+1}(x) \Rightarrow \forall\, x \in \mathbb{K}, \quad f(x) = (x-x_1)\dots(x-x_{p+1})\times g_{p+1}(x)$$

Et donc $\mathscr{P}_{p+1}$ est vraie.
La récurrence est démontrée □

On arrive à un résultat énoncé par Descartes, mais pas vraiment démontré…

Corollaire - Nombre maximal de solution
Une équation polynomiale de degré n admet au plus n solutions différentes

Démonstration
Il suffit de raisonner par l'absurde □

Identification

Il n'existe qu'une seule fonction polynomiale nulle :

Proposition - Une seule fonction polynomiale nulle
Si $f : x \mapsto \sum_{i=0}^{n} a_i x^i$ est une fonction polynomiale nulle
alors pour tout $i \in \mathbb{N}$, $a_i = 0$.

Démonstration
On fait un raisonnement par contraposée.
Si f n'est pas le polynôme nulle, elle est de degré n avec $n \geqslant 0$ (rappelons que le polynôme nul est de degré $-\infty$).
Donc f admet au plus n racines. Donc f n'est pas identiquement nul.
Bilan : $\exists\, i \in \mathbb{N}$ tel que $a_i \neq 0 \Longrightarrow f \neq 0$. □

Avec la même démonstration améliorée et adaptée à $f-g$, on trouve :

Proposition - Identification des coefficients
Soient f et g deux fonctions polynomiales et $E \subset \mathbb{C}$ tel que $\forall\, x \in E$, $f(x) = g(x)$.
Si $\text{card}(E) > \deg(f-g)$, alors $f = g$.
Plus précisément : pour tout $k \in \mathbb{N}$, $[f-g]_k = 0$ donc $[f]_k = [g]_k$

Remarque - Essentiel
On aura noté que :
- $[p]_k$ est le coefficient du polynôme p devant le monôme x^k. C'est une application linéaire.
- L'addition, la multiplication et la composition de deux polynômes donnent toujours un polynôme.

3.3. Expliciter formellement les racines

Avec les notations de Viète, il est facile de donner toutes les racines (complexes) d'un polynôme de degré 2 à coefficients dans $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$, à condition de savoir calculer la racine carrée d'un nombre complexe (cas où $\Delta \in \mathbb{C} \setminus \mathbb{R}$).

Proposition - Discriminant
On considère l'équation $ax^2+bx+c = 0$, où $a, b, c \in \mathbb{R}$ ou $\mathbb{C}$.
On note $\Delta = b^2 - 4ac$.
Si δ vérifie $\delta^2 = \Delta$, alors les racines de cette équation sont $\dfrac{-b+\delta}{2a}$ et $\dfrac{-b-\delta}{2a}$

On commence par une petite remarque :

Remarque - Théorème de Viète

Soient $(S,P)\in\mathbb{C}^2$. Les solutions du système

$$\begin{cases} z_1+z_2=S \\ z_1\times z_2=P \end{cases}$$

sont exactement (à permutation près) les solutions de $x^2-Sx+P=0$.

En effet :

$$(x-z_1)(x-z_2)=x^2-(z_1+z_2)x+z_1z_2=x^2-Sx+P$$

Démonstration

Il suffit de calculer

$$(x-\frac{-b+\delta}{2a})(x-\frac{-b-\delta}{2a}=x^2-\left(\frac{-b+\delta}{2a}+\frac{-b-\delta}{2a}\right)x+\frac{-b+\delta}{2a}\times\frac{-b-\delta}{2a}=x^2+\frac{b}{a}x+\frac{b^2-\delta^2}{4a^2}=\frac{1}{a}(ax^2+bx+c)=0$$

□

Remarque - Autres idées de démonstration ?

Avec la forme canonique.

En réfléchissant sur la symétrie et la moyenne des racines… On ne démontre pas ces résultats qui datent du XVI siècle :

Proposition - Formule de Tartaglia-Cardan (1545)

Les solutions complexes z_k $(k\in\{0,1,2\})$ de l'équation du troisième degré $x^3+px+q=0$ où p et $q\in\mathbb{R}$ sont donnés par

$$z_k=u_k+v_k$$

avec $u_k=j^k\sqrt[3]{\frac{1}{2}(-q+\sqrt{\frac{-\Delta}{27}})}$, $v_k=j^{-k}\sqrt[3]{\frac{1}{2}(-q-\sqrt{\frac{-\Delta}{27}})}$

et où $j=e^{\frac{2i\pi}{3}}$ et $\Delta=-(4p^3+27q^2)$ est le discriminant de l'équation.

On peut vérifier que $\Delta=(z_0-z_1)^2(z_1-z_2)^2(z_2-z_0)^2$ et $3u_kv_k=-p$

Exercice

Faire la démonstration

Correction

On calcule. Intelligemment…

Exemple - Racines de l'équation $x^3-2x=4$

Ici $p=-2$ et $q=-4$. Donc $\Delta=-(4\times(-8)+27\times16)=-25\times16$.

Ainsi $u_k=j^k\sqrt[3]{\frac{1}{2}(4+\frac{20}{\sqrt{27}}}=\frac{1}{\sqrt{3}}j^k\sqrt[3]{(6\sqrt{3}+10}$ et de même $v_k=\frac{1}{\sqrt{3}}j^k\sqrt[3]{(6\sqrt{3}-10}$

Ainsi $z_0=\frac{1}{\sqrt{3}}\left(\sqrt[3]{(6\sqrt{3}+10}+\sqrt[3]{(6\sqrt{3}-10}\right)=\cdots=2$.

On peut par exemple chercher des nombres entiers a,b tels que $(a\sqrt{3}+b)^3=6\sqrt{3}+10\ldots$ par développement, divisibilité, essai/erreur, on trouve : $(\sqrt{3}+1)^3=6\sqrt{3}+10$ et $(\sqrt{3}-1)^3=6\sqrt{3}-10\ldots$

On peut vérifier qu'il est plus simple ici de voir que 2 *est racine évidente, factoriser par* $(x-2)$*, puis trouver les deux autres racines conjuguées…*

Remarque - Degré 4

Il existe également une formule pour l'équation de degré 4. Vous la trouverez sans problème sur internet.

Vous y trouverez également l'histoire de la découverte des ces formules. C'est assez intéressant Nous ne démontrons pas :

Théorème - Ruffini-Abel (1824)

Pour tout entier $n\geqslant5$, il n'existe pas de formule générale exprimant « par radicaux » les racines d'un polynôme quelconque de degré n.

C'est-à-dire de formule n'utilisant que les coefficients, la valeur 1, les

quatre opérations et l'extraction des racines n-ièmes

4. Equation polynomiale et analyse

4.1. La meilleure méthode : l'essai/erreur

Heuristique - La meilleure solution

Si l'on veut résoudre un problème, dont on ne sait rien excepté ses réalisations pour certains réalisations des variables. Alors la solution naturelle consiste à faire des essais/erreurs.
Concrètement, pour résoudre $f = 0$, on prend une première valeur pour x. On essaye $f(x_1)$. Est-il égal à 0 ?
Si non, on essaye une autre valeur...

Est-il possible d'apprendre de nos essais/erreurs ?

Histoire - Niels Henrik Abel

Niels Henrik ABEL *(1802-1829) est un mathématicien norvégien génial, mort très jeune. Une météorite dans le ciel septentrional.*

4.2. Retro-contrôle

Une méthode classique en ingénierie (mais aussi en biologie) est d'exploiter le retro-contrôle ou une retro-action positive.

Truc & Astuce pour le calcul - Exploiter le retro-contrôle

Pour du calcul raisonné, il s'agit d'abord de réinjecter les résultats obtenus dans la formule initiale pour voir si le résultat est juste.
Mais si le résultat n'est pas juste, alors tout n'est pas perdu : il ne faut pas repartir de 0. Avec le calcul de vérification, il est parfois possible de voir où est l'erreur (erreur de signe, oublie d'un puissance...).
Et le second résultat ne doit pas être trop éloigné du premier résultat (plus grand si la fonction est croissante...).
Principe : on crée une suite (x_n) qui converge vers x, la vraie valeur que l'on cherche.
A chaque étape, on prend $x_{n+1} = x_n + y_n$, meilleure approximation de x que x_n,
et donc y_n est une suite telle que :
— y_n est beaucoup plus petit que x_n, donc négligeable face à x_n
— $(y_n) \to 0$
— pour $k \in N$, $k \geqslant 2$, y_n^k est toujours beaucoup plus petit que y_n, donc négligeable face à y_n^k.

Nous allons expliquer la méthode à partir d'un exercice.

Exercice

On cherche à donner une valeur approchée de $\sqrt{8}$. Donner une valeur approchée (à deux chiffres), par rétro-contrôle

---

Correction

$\sqrt{8} \approx \sqrt{9} = 3$. On a donc $\sqrt{8} = 2,\ldots$. On pose donc $x_1 = 3$. On considère $x_2 = x_1 + y_1 = 3 + y_1$. On a alors $x_2^2 = x_1^2 + 2x_1y_1 + y_1^2 = 9 + 6y_1 + y_1^2$.

On aimerait être proche de 8 pour x_2^2, et donc si y_1^2 est négligeable par rapport à y_1, on a $6y_1 = -1$, donc $y_1 = -\frac{1}{6}$.

On considère donc $x_2 = 3 - \frac{1}{6} = \frac{17}{6}$.

On considère ensuite $x_3 = x_2 + y_2$, donc $x_3^2 = x_2^2 + 2x_2y_2 + y_2^2 = \frac{289}{36} + \frac{17}{3}y_2 + y_2^2$.

On aimerait être proche de 8 pour x_3^2, et donc si y_2^2 est négligeable par rapport à y_2, on a $\frac{17}{3}y_2 = -\frac{288-289}{36} = \frac{-1}{36}$, donc $y_2 = -\frac{-1}{204}$.

On considère donc $x_3 = \frac{17}{6} - \frac{1}{204} = \frac{17^2 \times 2 - 1}{204} = \frac{577}{204}$.

On a (calculatrice) : $\sqrt{8} = 2,8284\ldots$

4.3. Méthode de la sécantes

Analyse - Principe

On cherche à résoudre une équation polynomiale $f(x) = 0$.
On essaye : $f(x_1) = y_1 \neq 0$ et $f(x_2) = y_2 \neq 0$.
Que valeur x_3 choisir pour obtenir une bonne approximation de la solution à $f(x) = 0$.
On peut faire une représentation graphique.

- Si $y_1 = -y_2$, on a envie d'essayer $x_3 = \frac{1}{2}(x_1 + x_2)$.
- si y_1 est proche de 0 et y_2 est loin. On va prendre x_3 proche de x_1.
  Plus grand, plus petit? Selon les signes de y_1 et y_2, on peut imaginer que x_3 est entre x_1 et x_2, ou non.
- ...

Est-il possible de réponde quantitativement à cette question?
On connait deux points de la courbe $y = f(x)$, on peut imaginer que la droite qui passe par ces deux points est la meilleure approximation de la fonction polynomiale.
On cherche alors la racine du polynôme, de degré 1, qui passe par ces deux points.
C'est la fonction polynomiale $d(x) = \frac{y_2-y_1}{x_2-x_1}(x - x_1) + y_1$.
En effet, par construction : $f(x_1) = y_1$ et $f(x_2) = y_2$.
On prend alors x_3, tel que $d(x_3) = 0$, i.e. : $x_3 = x_1 - \dfrac{x_2 - x_1}{y_2 - y_1} y_1$

Définition - Algorithme de la sécante

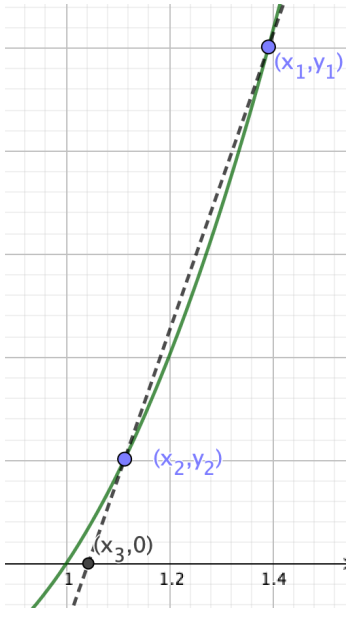
Considérons la fonction polynomiale $f(x) = \sum_{k=0}^{n} a_k x^k$.
On considère deux nombres $a, b \in \mathbb{R}$, puis la suite (u_n) définie par :

$$u_0 = a,\ u_1 = b \forall\ n \in \mathbb{N}, u_{n+2} = u_{n+1} - \frac{u_{n+1} - u_n}{f(u_{n+1}) - f(u_n)} f(u_{n+1})$$

La suite ainsi définie suit l'algorithme de la sécante

Sous certaines conditions, relativement robuste, la suite (u_n) converge vers une racine de f

Représentation - Méthode de la sécante



Pour aller plus loin - Convergence

Nous verrons plus loin ce que signifie la convergence pour une suite.
Vous pouvez déjà chercher à donner une définition formalisée...

4.4. Vers la dérivation. Méthode de la tangente

Analyse - Quand 0 s'invite

Au moment de la convergence vers une limite notée x, on trouve que u_n comme u_{n+1} est proche de x. Et donc $u_{n+1} - u_n$ est proche de 0.
Mais c'est aussi le cas de $f(u_n) \approx f(x) \approx f(u_{n+1})$.
On se trouve donc naturellement en présence d'une forme indéterminée $\frac{0}{0}$.
Comment gérer cette forme?

Heuristique - Toute forme indéterminée $\frac{0}{0}$...

| ... peut se voir comme un calcul de dérivée, avec la formule de L'HOSPITAL.

Rappelons ce qu'est un nombre dérivée :

Définition - Nombre dérivée

Soit f une fonction (polynomiale), on appelle dérivée de f en x_0, le nombre :

$$\lim_{h\to 0} \frac{f(x_0 + h) - f(x_0)}{h} = \lim_{x\to x_0} \frac{f(x) - f(x_0)}{x - x_0}$$

On note ce nombre $f'(x_0)$, selon la notation de Lagrange (1797).

Exemple - Monôme et polynôme. Dérivation

Considérons h petit (en valeur absolue), que pensez de $(x + h)^n$?
C'est un polynôme en x de degré n. Il vaut presque x^n (surtout si h petit). On peut appliquer les petits Bernoullis :

$$k > 0, \qquad (x+h)^k - x^k = -(x^k - (x+h)^k) = -(x - (x+h)) \times b^k_{x+h} = h \times \sum_{i=0}^{k-1} (x+h)^{k-1-i} x^i$$

Notons que pour $k=0$, $(x+h)^k - x^k = 1-1=0$. Donc si $f(x)=\sum_{k=0}^{n} a_k x^k$, par linéarité :

$$f(x+h)-f(x) = h \times \sum_{k=1}^{n} a_k \sum_{i=0}^{k-1} (x+h)^{k-1-i} x^i$$

Par conséquent (par continuité polynomiale)

$$\frac{f(x+h)-f(x)}{h} = \sum_{k=1}^{n} \sum_{i=0}^{k-1} (x+h)^{k-1-i} x^i \underset{h\to 0}{\Longrightarrow} \sum_{k=1}^{n} a_k \sum_{i=0}^{k-1} x^k = \sum_{k=1}^{n} k a_k x^{k-1}$$

Finalement : la dérivée de $x \mapsto b_0 + b_1 x + \dots b_n x^n$ est $x \mapsto b_1 + 2b_2 x + \dots n b_n x^{n-1}$.

Analyse - Méthode de la tangente

On a donc pour $u_n = x + h_n$, avec h_n proche de 0, $f(u_n) = f(x+h_n) = f(x) + h_n f'(x)$:

$$\frac{u_{n+1}-u_n}{f(u_{n+1})-f(u_n)} = \frac{(x+h_{n+1})-(x+h_n)}{[f(x)+h_{n+1}f'(x)]-[f(x)+h_n f'(x)]} = \frac{1}{f'(x)}$$

On a l'aglorithme suivant qu'on associe à NEWTON, pour son utilisation en toute généralité :

Définition - Algorithme de la tangente

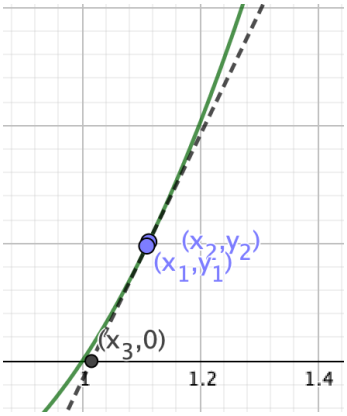
Considérons la fonction polynomiale $f(x)=\sum_{k=0}^{n} a_k x^k$.
On considère deux nombres $a \in \mathbb{R}$, puis la suite (u_n) définie par :

$$v_0 = a, : \forall\ n \in \mathbb{N}, v_{n+1} = v_n - \frac{1}{f'(v_n)} f(v_n)$$

La suite ainsi définie suit l'algorithme de la tangente

Sous certaines conditions, relativement robuste, la suite (v_n) converge vers une racine de f

Représentation - Méthode de la tangente (Newton)



Pour aller plus loin - Méthode de la Newton

En informatique, au second semestre, nous justifierons plus précisément l'algorithme de Newton. Nous donnerons des conditions de convergence.

5. Bilan

Synthèse

- ⇝ En science, les problèmes se traduisent sous forme d'équations, souvent polynomiales. Ces fonctions polynomiales sont très présentes car elles sont stables par addition, multiplication et composition. La multiplication s'appelle développement. On développe avec intelligence !
- ⇝ Le plus important pour résoudre une équation est de savoir factoriser. Le théorème de factorisation est très important : il permet de séparer les problèmes de résolution.
  Dans quelques rares cas (degré faible), il existe des formules explicites qui donnent les expressions des racines d'une fonction polynomiale.
- ⇝ On peut aussi regarder les fonctions polynômiales comme des transformations géométriques de la droite réelle. Faire le lien : fonction polynomiale/représentation géométrique permet d'enrichir chacun des deux points de vue. On peut penser à l'exemple des coniques.
- ⇝ Une autre idée est d'exploiter ce lien pour chercher les racines : localement une branche de courbe polynomiale ressemble à un segment de droite. Les algorithmes de la sécante ou de la tangente exploitent cette idée pour trouver une valeur approchée d'une racine d'une fonction polynomiale.

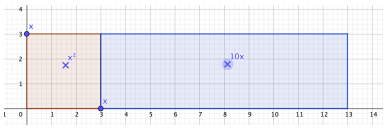
Savoir-faire et Truc & Astuce du chapitre

- Savoir-faire - Graphe d'une fonction $y = f(x)$ (translation)
- Savoir-faire - Graphe d'une fonction $y = f(x)$ (symétrie axiale)
- Savoir-faire - Graphe d'une fonction $y = f(x)$ (symétrie centrale)
- Truc & Astuce pour le calcul - Transformer le résultat d'un calcul en une représentation visuelle
- Truc & Astuce pour le calcul - Avec des fonctions
- Truc & Astuce pour le calcul - Anticipation
- Truc & Astuce pour le calcul - Reconnaissance de formes
- Truc & Astuce pour le calcul - Garder en mémoire « vive » le calcul
- Savoir-faire - Factoriser
- Truc & Astuce pour le calcul - Exploiter le retro-contrôle

Notations

Notations	*Définitions*	*Propriétés*	*Remarques*
$[f]_k$	coefficient de la fonction polynomiale f devant le monôme x^k	Linéarité : $[\lambda f+\mu g]_k = \lambda[f]_k + \mu[g]_k$ Convolution : $[f\times g]_k = \sum_{i=0}^{k}[f]_i[g]_{k-i}$	$\deg f = n \iff \forall\, k > n, [f]_k = 0$
$b_a^n(x)$	Petit Bernoulli en a d'indice n	$\forall\, x\in\mathbb{K}\ b_a^n(x) = \sum_{i=0}^{n} x^i a^{n-i}$ (fonction polynomiale)	$x^{n+1}-a^{n+1} = (x-a)\times b_a^n(x)$

Retour sur les problèmes

1. 
2. Il faut transformer en phrase les opérations $\times$, $\sqrt{\cdot}\ldots$
3. $x = u - v$, $x^3 = (u-v)^3 = u^3 - 3u^2v + 3uv^2 - v^3$.
   On a donc : $x^3+6x-20 = u^3 - v^3 - 3uv(u-v)+6x-20 = u^3 - v^3 + 3x(2-uv) - 20$.
   Ajoutons la condition $uv = 2$, on a donc $u^3 - v^3 = 20$ et $u^3v^3 = 8$.
   Ainsi, u^3 et $-v^3$ sont racines de $x^2 - Sx + P = x^2 - 20x - 8 = 0$.
   $\delta = 432$, puis $u^3 = 10+\sqrt{108}$ et $v^3 = 10-\sqrt{108}$.
   Et enfin, $x = \sqrt[3]{10+\sqrt{108}} - \sqrt[3]{10-\sqrt{108}}$.
   Autre méthode : $x^3+6x = 20$? Par essai : $x = 2$ fonctionne.

   $$x^3+6x-20 = (x-2)(x^2+2x+10) = (x-2)(x+1-3i)(x+1+3i)$$

   Puis, on exploite la formule de Cardan et enfin on reconnait une bicarré :

   $$x^4+px^2+1 = (x^2 - \frac{-p+\sqrt{p^2-4}}{2})(x^2 - \frac{-p-\sqrt{p^2-4}}{2})$$

   $$= (x-\sqrt{\frac{-p+\sqrt{p^2-4}}{2}})(x+\sqrt{\frac{-p+\sqrt{p^2-4}}{2}})(x-\sqrt{\frac{-p-\sqrt{p^2-4}}{2}})(x+\sqrt{\frac{-p-\sqrt{p^2-4}}{2}})$$

4. Voir cours de terminale (ou de première)
5. C'est la méthode de la retro-action.
6. Par récurrence : $f^{(k)}(x) = \sum_{i=k}^{n} i(i-1)\ldots(i-k)a_i x^{i-k} = k!a_k + \frac{(k+1)!}{1!}a_{k+1}x + \cdots + \frac{n!}{(n-k)!}a_n x^{n-k}$.
   Et donc $f^{(k)}(0) = k!a_k$. Ainsi, $a_k = \dfrac{f^{(k)}(0)}{k!}$.

Chapitre 2

Calculs et opérations avec des sommes ou des produits

Résumé -

Pour commencer, nous apprenons à manipuler les symboles $\sum$ et $\prod$.
Nous prendrons également le temps d'étudier les méthodes d'étude des doubles sommes (finies).
Les résultats se basent sur la commutativité de l'addition. Tant que les sommes sont finies, il n'y a pas de surprise dans les résultats obtenus. Le cas infini sera étudié plus tard.
Nous obtenons alors un premier résultat intéressant : les coefficients binomiaux (point de vue complémentaire à celui vu au lycée) et leur application à la formule du binôme de Newton. Nous faisons un petit plongeon dans l'histoire des mathématiques au XVII : les coefficients binomiaux agitaient la communauté des mathématiciens de l'époque.

Enfin, voici une liste de petites vidéo ou conférence visionnable sur internet et en lien avec le sujet (de pénibilités variées) :

- *Yvan Monka - Symbole Sigma. https ://www.youtube.com/watch ?v=0zspJuzo7L8*
- *El Jj - Nombres de Catalan. http ://eljjdx.canalblog.com/archives/2017/02/20/34959863.html*
- *Maths moi ça - L'étonnant triangle de Pascal. https ://www.youtube.com/watch ?v=IzkfjbWffpc*

Sommaire

1. Quelques problèmes

Problème 7 - Nombres triangulaires

Lorsqu'on fait la somme 1+1+1+... 1, on peut décrire tous les nombres entiers.
Si on somme ensuite les résultats obtenus : $1+2+\cdots+n$*, quel nombre obtient-on ?*
Ce nombre est appelé nombre triangulaire d'ordre n, noté T_n. Par exemple : $T_4 = 1+2+3+4 = 10$.
Si on somme ensuite les résultats obtenus : $T_1+T_2+\cdots+T_n = 1+3+6+10+\cdots+T_n$*, quel nombre obtient-on ? Et si on continue, toujours ?*

Problème 8 - Développement

Donner, pour tout entier n, la forme développée des applications polynomiales $x \mapsto (x-1)(x-2)(x-3)\dots(x-n)$ et $x \mapsto (1+x)(1+2x)(1+3x)\dots(1+nx)$

Problème 9 - Suite de nombres. Et le suivant ?

Prenons la suite obtenue à la question précédente : $1,4,10,20,35$. Quel est le terme suivant ?
Et de manière générale étant donnée une suite quelconque de n termes donnés explicitement, comment trouver le terme suivant ?

Problème 10 - Interpolation à pas constant

Quelle est la fonction f de degré minimal tel $f(0) = 1$, $f(1) = -1$, $f(2) = 1$, $f(3) = 4$?
Généraliser la question et donner une réponse...

Problème 11 - Développement de puissance

Considérons le calcul typiquement algébrique : $(a+b)^n$ où a et b sont deux nombres quelconques et n en entier.
L'expérience montre que pour $n = 4$ (par exemple), on trouve en développant cette expression : $a^4+4a^3b+6a^2b^2+4ab^3+b^4$.
Est-il possible de décrire simplement cette expression : quels sont les facteurs a et b obtenus, est-il facile d'exprimer/calculer le nombre situé devant a^ib^j ?
Et que se passe-t-il si l'on considère plutôt $(a+b+c+\dots)^m$?

2. Symboles $\sum$ et $\prod$

2.1. Définition

Remarque - Terme sans ambiguïté

La façon naturelle d'écrire une somme longue est de :

1. donner les premiers termes

2. s'assurer que la suite logique des termes est comprises, puis de remplacer ces termes (en grands nombres) par des points de suspension
3. donner la valeur du dernier terme.

Mais comment s'assurer qu'il n'y ait pas ambiguïté? On préfère souvent une formulation explicite :

5) qui utilise Σ pour désient utile pour pitre, Euler a portants aux

Définition - Notation $\sum$ et $\prod$
Soient $a_1,\dots,a_n$ n nombres réels ou complexes. L'addition dans $\mathbb{R}$ ou $\mathbb{C}$ étant commutative (on somme dans l'ordre que l'on souhaite) et associative (les parenthèses ne sont pas nécessaires), on note

$$\sum_{i=1}^{n} a_i = a_1 + \cdots + a_n$$

De même on note

$$\prod_{i=1}^{n} a_i = a_1 a_2 \dots a_n.$$

Le terme a_i s'écrit sous forme d'une formule dépendant de i.

Remarque - Bons usages et généralisation de notation
Bien évidemment, l'indice i (qui peut aussi s'appeler k,l,m,p ...) de la somme peut démarrer à une valeur entière autre que 1, toutefois la valeur de départ (sous le signe $\sum$) doit être inférieure à la valeur d'arrivée.

Exercice

Ecrire avec le symbole $\sum$, le calcul $1+2+4+\cdots+128$

Correction

On reconnait des puissances successives de 2. $1+2+4+\cdots+128 = \sum_{k=0}^{7} 2^k$

Plus généralement,

Définition - Extension de notation $\sum$ et $\prod$
Si I est un sous-ensemble fini de $\mathbb{N}$, $I = \{i_1, i_2, \dots, i_p\}$, on note

$$\sum_{i\in I} a_i = a_{i_1} + a_{i_2} + \cdots + a_{i_p} = \sum_i a_i \mathbb{1}_I(i)$$

où $\mathbb{1}_I : i \mapsto \begin{cases} 1 \text{ si } i \in I \\ 0 \text{ si } i \notin I \end{cases}$ est l'indicatrice de I.
On peut également noter, si $\mathscr{P}(i)$ désigne une propriété sur les entiers (parité, imparité...),

$$\sum_{i|\mathscr{P}(i)} a_i = \sum_{i\in\{j\in\mathbb{N}|\mathscr{P}(j) \text{ vraie }\}} a_i = \sum_i a_i [\mathscr{P}(i)]$$

où $[\mathscr{P}(i)]$ est *la notation d'Iverson* qui vaut 1 ssi $\mathscr{P}(i)$ est vraie et 0 sinon.
Ces notations se généralisant au produit.

Pour aller plus loin - Descriptions des ensembles
Nous verrons qu'un ensemble se définit en règle générale, soit par extension : $I = \{i_1, i_2, \dots, i_p\}$ soit par compréhension : $I = \{i \mid \mathscr{P}(i) \text{ (vraie)}\}$

Exemple - $\sum_{i=1}^{n} a_i$
Par exemple $\sum_{i=1}^{n} a_i = \sum_{i\in[\![1,n]\!]} a_i = \sum_{1\leqslant i\leqslant n} a_i$.
Cette dernière est un abus de $\sum_{i\in\mathbb{N} \mid 1\leqslant i\leqslant n} a_i$.

Exercice

Sans utiliser la notation $\sum$, donner l'expression développée de

$$\sum_{i|0\leqslant i\leqslant 6} \frac{1}{2i+1}, \quad \sum_{i|0\leqslant 2i\leqslant 7} \left(3i + \frac{1}{i+1}\right), \quad \sum_{i|0\leqslant i^3\leqslant 10} \frac{1}{i^2+i+1}.$$

Pour aller plus loin - Notation $\mathbb{N}_n$
On note, tout au long de l'année :

$$\mathbb{N}_n = [\![1, n]\!]$$

Cet ensemble commence bien à 1.
Cette notation n'est pas standardisée

Correction

$\sum_{i|0\leqslant i\leqslant 6} \frac{1}{2i+1} = \frac{1}{1}+\frac{1}{3}+\frac{1}{5}+\frac{1}{7}+\frac{1}{9}+\frac{1}{11}+\frac{1}{13}$ - $\sum_{i|0\leqslant 2i\leqslant 7} 3i+\frac{1}{i+1} = 3+6+9+1+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}$ - $\sum_{i|0\leqslant i^3\leqslant 10} \frac{1}{i^2+i+1} = \frac{1}{1}+\frac{1}{3}+\frac{1}{7}$

Attention - Attention à ne pas donner une existence à une variable muette!

Que vaut $\left(\sum_{k=1}^{n} 2^k\right)+k$? Rien!

Exercice

Exprimer la somme des inverses de tous les nombres premiers inférieurs à N

Correction

Par exemple $\sum_p \frac{1}{p}[p\leqslant N][p,\text{ premier}]$ ou encore $\sum_{p\leqslant N}\frac{1}{p}[p,\text{ premier}]$ ou $\sum_{p\in\mathscr{P}\cap\mathbb{N}_N}\frac{1}{p}\ldots$

2.2. Quatre règles opératoires

Nouvelle description de l'ensemble

Si on a un doute, on exploite **au brouillon** des points de suspension.

Analyse - Changement d'indice

Supposons que l'on ait à calculer $S=\sum_{a\in A} a$.

On peut supposons arbitrairement que $A=\{a_1,a_2,\ldots a_n\}$.

Alors S peut s'écrire sous la forme $\sum_{i\in M} a_i$ où M est une description quelconque de l'ensemble $[\![1,n]\!]$.

Deux cas sont classiques : $M=\{1,2,\ldots n\}$ ou encore $M=\{n,n-1,n-2,\ldots n-(n-1)\}$.

Ainsi, on a $S=\sum_{i=1}^{n} a_i=\sum_{h=0}^{n-1} a_{n-h}$, par exemple.

On peut imaginer d'autres écritures...

Savoir faire - Exemple de changement

Faire le changement d'indice $i=n-k$ pour $\sum_{k=1}^{n}(2k+1)$

On a donc $k=n-i$ et $2k+1=2n-2i+1=(2n+1)-2i$.

Et le tableau de correspondance :

k	i
1	$n-1$
n	0

Donc $\sum_{k=1}^{10}(2k+1)=\sum_{i=0}^{n-1}[(2n+1)-2i]$

On notera que les termes i sont notés dans l'ordre croissant (ce qui inverse l'ordre du calcul effectivement réalisé).

Exercice

Compléter les expressions qui suivent :

$$\sum_{i=1}^{n} a_i=\sum_{\ldots}^{\ldots} a_{k-1}=\sum_{\ldots}^{\ldots} a_{n-h}$$

Correction

$\sum_{i=1}^{n} a_i=\sum_{k=2}^{n+1} a_{k-1}=\sum_{h=0}^{n-1} a_{n-h}$

Somme, par récurrence

Analyse - $\sum$ et récurrence

Le symbole somme est particulièrement liée au raisonnement par récurrence.

En effet celui-ci a également pour vocation de formaliser le raisonnement avec des points de suspension.

Plus précisément, si on note, pour tout $n \in \mathbb{N}$, $S_n = \sum_{k=0}^{n} a_k$, alors (S_n) est exactement la suite définie par $S_0 = a_0$ et par la relation de *récurrence* : pour tout $n \in \mathbb{N}$, $S_{n+1} = S_n + a_{n+1}$

Proposition - A savoir!
On a pour $(a_i)_i, (b_i)_i, \lambda \in \mathbb{R}$ ou $\mathbb{C}$, :

$$\sum_{i=0}^{n}(a_i+b_i) = \sum_{i=0}^{n} a_i + \sum_{i=0}^{n} b_i \qquad \sum_{i=0}^{n} \lambda a_i = \lambda \sum_{i=0}^{n} a_i$$

$$\prod_{i=0}^{n} a_i b_i = \prod_{i=0}^{n} a_i \times \prod_{i=0}^{n} b_i \qquad \prod_{i=0}^{n} \lambda a_i = \lambda^{n+1} \prod_{i=0}^{n} a_i$$

Démonstration
On note $R_n = \sum_{i=1}^{n}(a_i+b_i)$, $S_n = \sum_{i=1}^{n} a_i$ et $T_n = \sum_{i=1}^{n} b_i$,
puis pour tout entier n, $\mathscr{P}_n$: $R_n = S_n + T_n$.
— $R_0 = (a_0+b_0) = a_0+b_0 = S_0+T_0$, donc $\mathscr{P}_0$ est vraie.
— Soit $n \in \mathbb{N}$, supposons que $\mathscr{P}_n$ est vraie.

$$R_{n+1} = R_n + (a_{n+1}+b_{n+1}) = S_n + T_n + a_{n+1} + b_{n+1} = S_{n+1} + R_{n+1}$$

Donc $\mathscr{P}_{n+1}$ est alors vraie.
La récurrence est démontrée.
Pour le second résultat, on va exploiter une méthode : l'utilisation d'un invariant de boucle.
Notons, pour tout $n \in \mathbb{N}$, $A_n = \sum_{i=1}^{n} \lambda a_i - \lambda \sum_{i=1}^{n} a_i$.

$A_{n+1} = \sum_{i=1}^{n} \lambda a_i + \lambda a_{n+1} - \lambda(\sum_{i=1}^{n} a_i + a_{n+1}) = A_n$.
Donc (A_n) est une suite constante, égale à $A_0 = \lambda a_0 - \lambda a_0 = 0$.
Donc pour tout $n \in \mathbb{N}$, $\sum_{i=1}^{n} \lambda a_i = \lambda \sum_{i=1}^{n} a_i$.
De même :
D'après la commutativité de la multiplication :

$$\prod_{i=1}^{n} a_i b_i = (a_1 b_1) \times (a_2 b_2) \cdots (a_n b_n) = (a_1 a_2 \cdots a_n) \times (b_1 b_2 \cdots b_n) = \prod_{i=1}^{n} a_i \times \prod_{i=1}^{n} b_i$$

En conséquence du cas précédent ($b_i = \lambda$, pour tout $i \in \mathbb{N}_n$) $\prod_{i=1}^{n} \lambda a_i = (\lambda a_1 \times \lambda a_2 \cdots \times \lambda a_n) =$

$\lambda^n (a_1 a_2 \cdots a_n) = \lambda^n \prod_{i=1}^{n} a_i$

□

Remarque - le $n+1$ de λ^{n+1} dans le dernier produit
… est donc en fait le cardinal de I, ensemble sur lequel on fait le produit.

Sommation par paquets

Exercice
Si A et B sont deux ensembles disjoints de E, montrer que pour tout $i \in E$, $\mathbb{1}_{A\cup B}(i) - \mathbb{1}_A(i) - \mathbb{1}_B(i) = 0$.
En déduire que $\sum_{i \in A\cup B} a_i = \sum_{i\in A} a_i + \sum_{i \in B} a_i$.

Correction
Comme A et B sont disjoints, ou bien $i \in A$ et $i \notin B$, ou bien $i \notin A$ et $i \in B$ ou bien $i \notin A$ et $i \notin B$. Pour chacun de ces trois cas, on a respectivement :

$$\mathbb{1}_{A\cup B} - \mathbb{1}_A - \mathbb{1}_B = \begin{cases} 1-1-0=0 \\ 1-0-1=0 \\ 0-0-0=0 \end{cases}$$

Dans tous les cas : $\mathbb{1}_{A\cup B} = \mathbb{1}_A + \mathbb{1}_B$.
Puis, on multiplie par a_i et on additionne :

$$\sum_{i\in A\cup B} a_i = \sum_i a_i \mathbb{1}_{A\cup B}(i) = \sum_i a_i \mathbb{1}_A(i) + \sum_i a_i \mathbb{1}_B(i) = \sum_{i\in A} a_i + \sum_{i\in B} a_i$$

Généralisons ce résultat. Il faut commencer par définir une notation :

Définition - Réunion disjointe
On note $C = A \uplus B$, par signifier la double information :
— $C = A \cup B$
— $A \cap B = \emptyset$
Autrement écrit : $x \in C$ si et seulement si $x \in A$ ou (exclusif) $x \in B$.
On peut généraliser cette notation à plusieurs ensembles :
$C = \biguplus_{i=1}^{n} A_i$ signifie : $x \in C$ si et seulement si $\exists\, !i \in \mathbb{N}_n$ tel que $x \in A_i$.

On a alors la relation de Chasles pour les sommes :

Pour aller plus loin - Autre notation
On trouve parfois également la notation : $C = A \sqcup B$ ou encore $C = A \cup\!\!\!\cdot\, B$.
Ce + est là pour pourvoir signifier un résultat que l'on verra plus loin, si $C = \biguplus_{i=1}^{n} A_i$, alors $\text{card}(C) = \sum_{i=1}^{n} \text{card}(A_i)$.

Proposition - Sommation par paquets
Soit une famille $(E_r)_{r\in S}$ une famille d'ensembles indexés par S.
On suppose qu'il s'agit d'une famille d'ensembles disjoints 2 à 2 :
$\forall\, r \neq r' \in S,\ E_r \cap E_{r'} = \emptyset$.
Alors
$$\sum_{r\in S}\left(\sum_{k\in E_r} a_k\right) = \sum_{k\in \biguplus_{r\in S} E_r} a_k$$

On voit ici apparaître une double somme. On en reparlera plus loin.
Exemple - $\mathbb{N}$
On peut vouloir couper la somme en deux parties : les indices pairs et les indices impairs.
Dans ce cas $S = \{1,2\}$, $E_1 = \mathscr{P} \cap \mathbb{N}_r$ et $E_2 = \mathscr{I} \cap N_r$, par exemple.

Démonstration
On a là encore : $E = \biguplus_{r\in S} E_r$ et donc

$$\begin{array}{rcl} k \in E & \iff & \exists\, !r \in S \text{ tel que } k \in E_r \\ \mathbb{1}_E(k) = 1 & \iff & \exists\, r \in S \text{ tel que } \mathbb{1}_{E_r}(k) = 1 \text{ et } \forall\, s \in S\setminus\{r\}, \mathbb{1}_{E_s}(k) = 0 \\ & \iff & \sum_{r\in S} \mathbb{1}_{E_r}(k) = 1 \end{array}$$

Donc, ceci étant vrai pour tout k, on a l'égalité de fonctions

$$\mathbb{1}_E = \sum_{r\in S} \mathbb{1}_{E_r}$$

Alors, comme précédemment :

$$\begin{aligned} \sum_{k\in E} a_k &= \sum_k a_k \mathbb{1}_E(k) = \sum_k a_k \left(\sum_{r\in S} \mathbb{1}_{E_r}(k)\right) \\ &= \sum_k \sum_{r\in S} a_k \mathbb{1}_{E_r}(k) = \sum_{r\in S}\sum_k a_k \mathbb{1}_{E_r}(k) \\ &= \sum_{r\in S}\sum_{k\in E_r} a_k \end{aligned}$$

□

Remarque - Interversion des symboles Σ
Dans la démonstration, on a inversé deux symboles Σ. Cela sera justifié par la suite, car il s'agit d'une **somme finie**.
C'est comme si on sommait les termes dans un tableau : d'abord en ligne, puis en colonne ; ou d'abord en colonne, puis en ligne. Dans tous les cas, on obtient le même résultat car il s'agit d'additionner une et une seule fois tous les termes du tableau.

Savoir faire - Exploiter une sommation par paquets
On a parfois intérêt à découper l'ensemble E en (réunion de m) sous-ensembles disjoints $E = E_1 \uplus E_2 \cdots \uplus E_m$.

On calcule alors la somme par paquets :

$$\sum_{r=1}^{m}\left(\sum_{k\in E_r} a_k\right) = \sum_{k\in E} a_k$$

Télescopage

C'est la seule méthode ui donne une formule explicite.

Savoir faire - Méthode du télescopage (ou dominos)

Soit (u_n) une suite. Soient $p, n \in \mathbb{N}$ tels que $p \leqslant n$

Alors $\sum_{k=p}^{n}(u_{k+1}-u_k) = u_{n+1}-u_p$

$\left(=(u_{p+1}-u_p)+(u_{p+2}-u_{p+1})+(u_{p+3}-u_{p+2})+\cdots+(u_{n+1}-u_n)\right)$

Remarque - « Voir » le téléscopage

Deux remarques :

- Dans la plupart des situations, la somme ne se présente pas directement sous la forme $\sum_{k=p}^{n}(u_{k+1}-u_k)$, il faut commencer par faire "apparaître" cette forme.
- On peut aussi avoir intérêt à écrire la somme avec des points de suspension pour confirmer le télescopage aperçu

Exercice

Calculer $\sum_{k=1}^{n} \ln\frac{k+1}{k}$.

Correction

$\sum_{k=1}^{n} \ln\frac{k+1}{k} = \sum_{k=1}^{n}[\ln(k+1)-\ln(\mathbb{K})] = \ln(n+1)-\ln(1) = \ln(n+1)$

Pour aller plus loin - Théorème fondamentale de l'analyse entière?

Intégrer↔ Dériver sont les deux problèmes inverses l'un de l'autre. Ils sont définis pour des fonctions de la variable réelle.
Pour la variable entière, les fonctions sont des suites. Et intégrer consiste simplement à faire le calcul $\sum_{k=0}^{n} a_k$.
Alors à quoi correspond la dérivation de la variable entière? Au calcul $a_{n+1}-a_n$! (c'est pas exemple comme cela qu'on voit si une suite est croissante...).
La formule du télescopage présente ce lien : $\sum \leftrightarrow \delta$

2.3. Avec Python

Il arrive souvent que l'on rencontre des calculs de sommes à effectuer sous Python. La méthode est simple, on emploie des boucles :

- `for`, si l'on connait bien l'ensemble sur lequel est definie i
- `while`, si i est définie par une propriété

Informatique - $\sum_{k=n}^{m} a(k)$

```
def Somme1(n,m):
    S=0
    for k in range(n,m+1):
        S=S+a(k)
    return(S)
```

Informatique - $\sum_{i\,|\,f(i)\leqslant n} a(i)$

```
def Somme2(n):
    S,i=0,0
    while f(i)<=n:
        S=S+a(i)
        i=i+1
    return(S)
```

Remarque - L'ordinateur (calculatrice) comme un obstacle?

L'une des principales raisons de la faiblesse des élèves pour le calcul est l'usage trop tôt de la calculatrice.
Il faut laisser le temps pour comprendre, maitriser, puis ne pas oublier le sens du

calcul avant de passer à la calculatrice.
Néanmoins, il ne faut pas systématiquement tout jeter!

Truc & Astuce pour le calcul - Écrire un programme pour mieux comprendre

Ecrire un programme permet souvent de mieux comprendre la nature du calcul.
C'est le cas en particulier :

— pour le calcul de somme.
— pour le calcul de probabilités.
Dans ce cas, ce n'est pas le calcul mais la modélisation elle-même du problème qui est mieux comprise.

Exercice
Ecrire une boucle (double ?) pour faire le calcul :

$$\sum_{i=1}^{100}\sum_{j=i}^{200-i} i\times j$$

Correction

```
S=0
for i in range (101):
   for j in range(i,200-i+1):
      S=S+i*j
return(S)
```

Remarque - Varier les paramètres et informatique

L'informatique permet aussi de facilement faire varier les paramètres. On a vu la force de l'usage des paramètres dans le calcul. Mais c'est aussi, de manière générale, la force de l'excellent mathématicien. Python nous aidera largement : ce n'est que légèrement une boite noire pour nous, nous aurons donc pour ambition de bien maîtriser tous nos faits et gestes informatiques (pas de clicothérapie!).

2.4. Des sommes connues

Proposition - Sommes de puissances d'entiers consécutifs
Soit $n\in\mathbb{N}^*$. Alors :

$$\sum_{k=1}^{n} k=\frac{n(n+1)}{2};\quad \sum_{k=1}^{n} k^2=\frac{n(n+1)(2n+1)}{6};\quad \sum_{k=1}^{n} k^3=\left(\frac{n(n+1)}{2}\right)^2$$

On peut faire une récurrence, ou bien chercher à démontrer le résultat directement. Une méthode classique : user le télescopage.

Démonstration
Nous démontrons le second résultat en admettant le premier.
Notons d'abord que pour tout entier k : $(k+1)^3-k^3=3k^2+3k+1$.
Donc, en voyant un téléscopage :

$$(n+1)^3-1=\sum_{k=1}^{n}[(k+1)^3-k^3]=\sum_{k=1}^{n}[3k^2+3k+1]=3\sum_{k=1}^{n}k^2+3\sum_{k=1}^{n}k+\sum_{k=1}^{n}1$$

Ainsi (toujours d'abord factoriser!) :

$$3\sum_{k=1}^{n}k^2=(n+1)^3-1-\frac{3n(n+1)}{2}-n=\frac{(n+1)}{2}[2(n+1)^2-2-3n]=\frac{(n+1)(2n^2+n)}{2}=\frac{n(n+1)(2n+1)}{2}$$

□

Exercice
Démontrer par récurrence les résultats précédents

Correction

Posons pour tout $n\in\mathbb{N}^*$, $\mathscr{P}_n$:« $\sum_{k=1}^{n}k^3=\left(\frac{n(n+1)}{2}\right)^2$ ».

— $\sum_{k=1}^{1} k^3 = 1$ et $\left(\frac{1(1+1)}{2}\right)^2 = 1$

— Soit $n \in \mathbb{N}$ et supposons $\mathscr{P}_n$ vérifiée.

$$\sum_{k=1}^{n+1} k^3 = \sum_{k=1}^{n} k^3 + (n+1)^3 = \left(\frac{n(n+1)}{2}\right)^2 + (n+1)^3$$
$$(n+1)^2\left[\tfrac{1}{4}(n^2+4n+4)\right] = \frac{(n+1)^2(n+2)^2}{4}$$

Donc $\mathscr{P}_{n+1}$ est vraie.

Savoir faire - Se passer du formalisme d'une récurrence ou invariant de boucle

On peut souvent se passer de la formalisation de la récurrence (mais avec les mêmes calculs).

Ici on considère la suite $u_n = \sum_{k=1}^{n} k^3 - \left(\frac{n(n+1)}{2}\right)^2$.

On note que $u_{n+1} = u_n$ (même calcul que $\mathscr{P}_n \Rightarrow \mathscr{P}_{n+1}$) et que $u_1 = 0$.

Donc pour tout $n \in \mathbb{N}^*$, $u_n = 0$. CQFD.

Proposition - Calcul d'une somme arithmétique

Pour une suite arithmétique :

$$\forall\, n \in \mathbb{N}, u_{n+1} = u_n + r$$

on a :

$$\sum_{k=n}^{m} u_k = (m-n+1) \times \frac{u_m + u_n}{2}$$

C'est-à-dire :

somme de termes succ. = (nb de termes) × (moyenne des termes extrêmes)

Démonstration

Pour tout k, $u_k = u_n + (k-n)r$.

$$\sum_{k=n}^{m} u_k = \sum_{k=n}^{m} u_n + r\sum_{k=n}^{m}(k-n) = (m-n+1)u_n + r\sum_{i=0}^{m-n} i$$

$$= (m-n+1)u_n + r\frac{1}{2}(m-n)(m-n+1) = \frac{m-n+1}{2}(u_n + r(m-n) + u_n) = (m-n+1) \times \frac{u_m+u_n}{2}$$

□

Exercice

Démontrer le résultat en suivant la « méthode de Gauss » (double somme inversée...)

Correction

On note S, la somme en question : $2S = S + S = \sum_{k=n}^{m} u_k + \sum_{k=n}^{m} u_k$. Dans la première somme on note $h = k$ et dans la seconde on fait le changement d'indice $h = m - k + n$:

$$2S = \sum_{h=n}^{m} u_h + \sum_{h=n}^{m} u_{m+n-h} = \sum_{h=n}^{m} (u_n + (h-n)r + u_n + (m+n-h-n)r)$$
$$= \sum_{h=n}^{m} (u_n + u_n + (m-n)r) = \sum_{h=n}^{m} (u_n + u_m) = (m-n+1)(u_n + u_m)$$

Proposition - Calcul d'une somme géométrique

Soit $x \in \mathbb{R}$ (ou $x \in \mathbb{C}$) et $n \in \mathbb{N}$. On a alors :

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

Plus généralement pour une suite géométrique de raison $q \neq 1$, on a :

$$\text{somme de termes successifs} = \text{1er terme} \times \frac{1-q^{\text{nb de termes}}}{1-q}$$

Démonstration
Si $x = 1$, $\sum_{k=0}^{n} x^k = \sum_{k=0}^{n} 1 = n+1$
Si $x \neq 1$, $(1-x)\sum_{k=0}^{n} nx^k = \sum_{k=0}^{n} (x^k - x^{k+1}) = x^0 - x^{n+1}$, télescopage. □

Exercice
Calculer $1+2+4+8+16+32+64+128$

Correction
C'est la somme de 8 termes consécutifs d'une suite géométrique de raison 2.
Donc $1+2+4+8+16+32+64+128 = 1 \times \frac{1-2^8}{1-2} = 2^8 - 1 = 255$.
Quel lien avec l'écriture en binaire des nombres entiers sur 8 octets ?

On se rappelle, avec les petits Bernoulli :

Proposition - Une factorisation à connaître
Soient a et b deux réels (ou deux complexes), alors :

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

Application - $3^n - 2^n$, sous forme d'une addition de n termes
$3^n - 2^n = (3-2)(3^{n-1} + 2\ldots 3^{n-2} + 4\ldots 3^{n-3} + \cdots + 2^{n-2}\ldots 3 + 2^{n-1}$.

Application - Factoriser $a^5 - b^5$
Vu au chapitre précédent.
$a^5 - b^5 = (a-b)(a^4 + a^3b + a^2b^2 + ab^3 + b^4)$

Démonstration
On développe et utilisons la méthode du téléscopage :

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

Puis en prenant $a = 1$, $1 - x^{n+1} = (1-x)\sum_{k=0}^{n} x^k \ldots$ □

Exercice
Peut-on factoriser $a^n + b^n$? Si oui, factoriser le.

Correction
Si $b = -a$, $a^n + b^n = a^n(1 + (-1)^n) = 0$ si n est impair.
On peut factoriser par $a+b$ dans ce cas :

$$a^n + b^n = (a+b)(a^{n-1} - a^{n-2}b + a^{n-3}b^2 + \cdots + (-1)^k a^{n-1-k}b^k + \cdots + b^{n-1})$$

car $n-1$ est alors pair

2.5. Sommes doubles (multiples...)

Heuristique - Somme à multiples indices
On peut faire une somme d'éléments pris dans un ensemble fini. Mais la description de ces éléments n'est pas toujours naturellement données sous la forme $x_i, i \in [\![0,n]\!]$.
Parfois les éléments apparaissent comme les éléments d'un tableau (*matrice*) et sont donc doublement (ou plus) indexés : $x_{i,j}, i \in \mathbb{N}_n, j \in \mathbb{N}_m$.
Les choses se présentent différemment selon que i et j sont « indépendants » entre eux ou non.

Cas i et j indépendants. Produit cartésien d'ensembles

Analyse - Du sens des formules

Comment décrire avec des symboles $\sum$ la somme suivante :

$$S = a_{1,1} + a_{1,2} + \cdots + a_{1,m} + a_{2,1} + \cdots + a_{2,m} + \cdots + a_{n,1} + \cdots + a_{n,m}$$

On peut voir cette somme de la façon suivante :

$$S = \left(a_{1,1} + a_{1,2} + \cdots + a_{1,m}\right) + \left(a_{2,1} + \cdots + a_{2,m}\right) + \cdots + \left(a_{n,1} + \cdots + a_{n,m}\right)$$

On y voit n sommes, chacune composée d'une somme de m termes :

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

Evidemment, la somme étant finie, l'addition commutative, on a également :

$$S = a_{1,1} + a_{2,1} + \cdots + a_{n,1} + a_{1,2} + \cdots + a_{n,2} + \cdots + a_{1,m} + \cdots + a_{n,m} = \sum_{j=1}^{m} \left(\sum_{i=1}^{n} a_{i,j} \right)$$

Et finalement on peut écrire :

$$S = \sum_{i=1}^{n} \left(\sum_{j=1}^{m} a_{i,j} \right) = \sum_{j=1}^{m} \left(\sum_{i=1}^{n} a_{i,j} \right) = \sum_{(i,j) \in \mathbb{N}_n \times \mathbb{N}_m} a_{i,j}$$

Définition - Somme double

On considère une famille de nombres réels ou complexes $(a_{i,j})$ indexée par deux indices i et j, i compris entre 1 et n, j compris entre 1 et m où n et m sont deux entiers non nuls donnés. On peut représenter ces nombres par un tableau où l'indice i désigne la ligne et l'indice j la colonne :

$$\begin{array}{ccccccc} a_{1,1} & \cdots & a_{1,j} & \cdots & \cdots & a_{1,m} \\ \vdots & & \vdots & & & \vdots \\ a_{i,1} & \cdots & a_{i,j} & \cdots & \cdots & a_{i,m} \\ \vdots & & \vdots & & & \vdots \\ a_{n,1} & \cdots & a_{n,j} & \cdots & \cdots & a_{n,m} \end{array}$$

La somme de tous les éléments de ce tableau est notée

$$\sum_{1 \leqslant i \leqslant n, 1 \leqslant j \leqslant m} a_{i,j} = \sum_{(i,j) \in [\![1,n]\!] \times [\![1,m]\!]} a_{i,j}.$$

Savoir faire - Somme multiple (indépendance)

Pour la calculer on peut procéder d'au moins deux façons, la première consiste à faire d'abord la somme des termes ligne par ligne, puis d'additionner les résultats :

$$\sum_{1 \leqslant i \leqslant n, 1 \leqslant j \leqslant m} a_{i,j} = \sum_{i=1}^{n} \underbrace{\left(\sum_{j=1}^{m} a_{i,j} \right)}_{\text{somme de la ligne } i} = \sum_{i=1}^{n} \sum_{j=1}^{m} a_{i,j},$$

une seconde étant de sommer d'abord les termes colonne par colonne puis d'additionner les résultats :

$$\sum_{1 \leqslant i \leqslant n, 1 \leqslant j \leqslant m} a_{i,j} = \sum_{j=1}^{m} \underbrace{\left(\sum_{i=1}^{n} a_{i,j} \right)}_{\text{somme de la colonne } j} = \sum_{j=1}^{m} \sum_{i=1}^{n} a_{i,j}.$$

Remarque - Diagonale
Il y a d'autres possibilités, par exemple en sommant suivant des lignes diagonales lorsque $n = m$.
On reverra cela plus loin.

Exercice
Calculer $\sum_{1\leqslant i\leqslant n, 1\leqslant j\leqslant p} ij$

Correction
$S = \sum_{i=1}^{n}\left(\sum_{j=1}^{p} ij\right) = \sum_{i=1}^{n} i\left(\sum_{j=1}^{p} j\right) = \sum_{i=1}^{n} i\frac{p(p+1)}{2} = \frac{n(n+1)p(p+1)}{4}$

Comme le montre l'exercice précédent :

Proposition - Produit de deux sommes (développement ou factorisation)
Soient des réels (ou des complexes) a_i et b_j, $1 \leqslant i \leqslant n$, $1 \leqslant j \leqslant p$. Alors :

$$\left(\sum_{i=1}^{n} a_i\right)\left(\sum_{j=1}^{p} b_j\right) = \sum_{(i,j)\in[\![1,n]\!]\times[\![1,p]\!]} a_i b_j$$

Démonstration
C'est assez simple, par développement :

$$\left(\sum_{i=1}^{n} a_i\right)\left(\sum_{j=1}^{p} b_j\right) = \left(\sum_{i=1}^{n} a_i\left(\sum_{j=1}^{p} b_j\right)\right) = \sum_{(i,j)\in[\![1,n]\!]\times[\![1,p]\!]} a_i b_j$$

□

Cas i et j dépendants

Heuristique - Cas : i et j dépendants
Ici on somme seulement certains termes du tableau rectangulaire $a_{i,j}$.
Et donc les indices sont dépendants l'un de l'autre!
Par exemple, dans cette situation, les valeurs prises par j (à l'intérieur de la somme) dépendent de celles prises par i (à l'extérieur de la somme). Il y a, en revanche, souvent liberté dans le choix de l'ordre de sommation (d'abord i ou d'abord j).

Analyse - $G = \{a_{i,j}, i \in \mathbb{N}_n, j \in A_i\}$
On suppose que G est décrit parfaitement ainsi : $\{a_{i,j}, i \in A, j \in A_i\}$.
Donc

$$\sum_{a\in G} a = \sum_{i,j} a_{i,j}\mathbb{1}_G(a_{i,j}) = \sum_{i\in A}\left(\sum_{j\in A_i} a_{i,j}\right)$$

Evidemment, la somme $\sum_{j\in A_i}\left(\sum_{i\in A} a_{i,j}\right)$ ne signifie rien (A_i ne peut exister car i n'existe pas) et donc les choses ne commutent pas facilement ici.

Proposition - Somme double classique $\sum_{1\leqslant j\leqslant i\leqslant n} a_{i,j}$
Soit $(a_{i,j})_{(i,j)\in\mathbb{N}^2}$ une famillle de nombres réels ou complexes :

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

Démonstration
On a les calculs suivants :

$$\begin{aligned}\sum_{1\leqslant j\leqslant i\leqslant n} a_{i,j} &= \sum_{i,j} a_{i,j}[1\leqslant j\leqslant i\leqslant n] = \sum_i \left(\sum_j a_{i,j}[1\leqslant j\leqslant i][1\leqslant i\leqslant n]\right)\\ &= \sum_i \left(\sum_j a_{i,j}[1\leqslant j\leqslant i]\right)[1\leqslant i\leqslant n] = \sum_{i=1}^{n}\left(\sum_{j=1}^{i} a_{i,j}\right)\\ &= \sum_j \left(\sum_i a_{i,j}[j\leqslant i\leqslant n][1\leqslant j\leqslant n]\right) = \sum_{j=1}^{n}\left(\sum_{i=j}^{i} a_{i,j}\right)\end{aligned}$$

□

Exercice
Calculer $\sum_{1\leqslant i\leqslant j\leqslant n} ij$

Correction

$$S = \sum_{j=1}^{n} j\left(\sum_{i=1}^{j} i\right) = \sum_{j=1}^{n} \frac{j^2(j+1)}{2} = \frac{1}{2}\left(\frac{n^2(n+1)^2}{4} + \frac{n(n+1)(2n+1)}{6}\right)$$
$$= \frac{n(n+1)(3n^2+7n+2)}{24} = \frac{n(n+1)(3n+1)(n+2)}{24}$$

Pour aller plus loin - Formalisme/sens
*On notera ici (mais c'est souvent le cas en mathématiques) que le formalisme est plus **efficace** que la recherche de sens : il permet une sorte d'automat(h)isation.*
Néanmoins, il est bon de toujours accompagné l'un par l'autre.

Savoir faire - Somme multiple (dépendance)
Pour la calculer on ordonne les indices de sommation :

1. on choisit celle qui sera le plus à l'extérieur (donc à gauche) des symboles $\sum$. Elle ne dépend que de paramètres fixés et d'aucun indice.
2. on choisit ensuite la suivante, la seconde dans l'ordre des sommes. Elle dépend des paramètres fixés et de l'indice précédent

…

Exemple : $\sum_{1\leqslant i<j\leqslant n} a_{i,j} = \sum_{i=1}^{n-1}\left(\sum_{j=i+1}^{n} a_{i,j}\right) = \sum_{j=2}^{n}\left(\sum_{i=1}^{j-1} a_{i,j}\right)$

Analyse - Sens de ces modes de sommations
On somme seulement certains termes du tableau (et pas tous).
Si E désigne l'ensemble des indices (i,j) des nombres que l'on désire sommer, on notera $\sum_{(i,j)\in E} a_{i,j}$ cette somme.
Pour le cas présent $E = \{(i,j)\in\mathbb{N}^2 | 1\leqslant j\leqslant i\leqslant n\}$ (termes sous ou sur la diagonale principale du tableau carré à n lignes et n colonnes), on a alors

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

Il est recommandé de comprendre les tableau suivants :
Première sommation :

$$\underbrace{\sum_{i=1}^{n}}_{\downarrow_2}\underbrace{\sum_{j=1}^{i}}_{\rightarrow_1} a_{i,j}:\quad \begin{array}{ccccc} a_{1,1} & & & & \\ \downarrow_2 & \ddots & & & \\ a_{h,1} & \rightarrow_1 & a_{h,h} & & \\ \downarrow_2 & & \downarrow_2 & \ddots & \\ a_{n,1} & \rightarrow_1 & a_{n,h} & \rightarrow_1 & a_{n,n}\end{array}$$

Deuxième sommation :

$$\underbrace{\sum_{j=1}^{n}}_{\rightarrow_2}\underbrace{\sum_{i=j}^{n}}_{\downarrow_1} a_{i,j}:\quad \begin{array}{ccccc} a_{1,1} & & & & \\ \downarrow_1 & \ddots & & & \\ a_{h,1} & \rightarrow_2 & a_{h,h} & & \\ \downarrow_1 & & \downarrow_1 & \ddots & \\ a_{n,1} & \rightarrow_2 & a_{n,h} & \rightarrow_2 & a_{n,n}\end{array}$$

2.6. Exercice d'applications

Exercice

Retrouver la valeur de $S=\sum_{k=1}^{n} k$, en notant que $S=\sum_{k=1}^{n}\sum_{i=1}^{k} 1$

Correction

$S=\sum_{k=1}^{n} k=\sum_{k=1}^{n}\sum_{i=1}^{k} 1=\sum_{i=1}^{n}\sum_{k=i}^{n} 1=\sum_{i=1}^{n}(n-i+1)=\sum_{i=1}^{n}(n+1)-\sum_{i=1}^{n} i.$
Donc $2S=n(n+1)$ et donc $S=\frac{n(n+1)}{2}$.

Exercice

Montrer que $\left(\sum_{k=1}^{n} d_k\right)^2=\sum_{k=1}^{n} d_k^2+2\sum_{1\leqslant i<j\leqslant n} d_i d_j.$

Correction

$\left(\sum_{k=1}^{n} d_k\right)^2=\left(\sum_{k=1}^{n} d_k\right)\times\left(\sum_{k=1}^{n} d_k\right)^2=\sum_{(k,h)\in\mathbb{N}_n^2} d_k d_h.$
Or $\mathbb{N}_n^2=\{(k,k),k\in\mathbb{N}_n\}\uplus\{(k,h),k<h\in\mathbb{N}_n\}\uplus\{(k,h),k>h\in\mathbb{N}_n\}$.
On peut aussi exploiter la notation d'Iverson $\mathbb{N}^2=\{(k,h)\mid[k=h]\}\uplus\{(k,h)\mid[k<h]\}\uplus\{(k,h)\mid[k>h]\}$.
Par sommation par paquets :
$\left(\sum_{k=1}^{n} d_k\right)^2=\sum_{k=1}^{n} d_k^2+\sum_{1\leqslant k<h\leqslant n} d_k d_h+\sum_{1\leqslant k>h\leqslant n} d_k d_h=\sum_{k=1}^{n} d_k^2+2\sum_{1\leqslant k<h\leqslant n} d_k d_h,$
en faisant $k\leftrightarrow h$ dans la dernière somme.

On peut aussi démontrer l'inégalité célèbre de Cauchy-Schwarz :

Exercice

1. En développant $\sum_{1\leqslant i<j\leqslant 3}(a_i b_j-a_j b_i)^2$, montrer que $\left(\sum_{k=1}^{3} a_k b_k\right)^2\leqslant\left(\sum_{k=1}^{3} a_k^2\right)\left(\sum_{k=1}^{3} b_k^2\right).$
2. De même, montrer l'inégalité de Cauchy-Schwarz :
$$\left(\sum_{k=1}^{n} a_k b_k\right)^2\leqslant\left(\sum_{k=1}^{n} a_k^2\right)\left(\sum_{k=1}^{n} b_k^2\right)$$
3. A quelle condition a-t-on : $\sum_{k=1}^{n}(a_i b_i)^2=\left(\sum_{k=1}^{n} a_k^2\right)\left(\sum_{k=1}^{n} b_k^2\right)$

Correction

1. On a
$$\begin{aligned}
0\leqslant & \sum_{1\leqslant i<j\leqslant 3}(a_i b_j-a_j b_i)^2=(a_1b_2-a_2b_1)^2+(a_1b_3-a_3b_1)^2+(a_2b_3-a_3b_2)^2\\
0\leqslant & (a_1b_2)^2+(a_1b_3)^2+(a_2b_1)^2+(a_2b_3)^2+(a_3b_1)^2+(a_3b_2)^2\\
& -2(a_1a_2b_1b_2)-2(a_1a_3b_1b_3)-2(a_2a_3b_2b_3)\\
0\leqslant & \sum_{i\neq j}(a_i b_j)^2+\sum_i(a_i b_i)^2-\sum_i(a_i b_i)^2-2\sum_{i<j}(a_i a_j b_i b_j)\\
0\leqslant & \sum_{i,j} a_i^2 b_j^2-\sum_{i,j}(a_i b_i a_j b_j)=\sum_{\mathbb{K}} a_k^2\sum_{\mathbb{K}} b_k^2-\left(\sum_{\mathbb{K}}(a_k b_k)\right)^2
\end{aligned}$$
Donc $\left(\sum_{k=1}^{3} a_k b_k\right)^2\leqslant\left(\sum_{k=1}^{3} a_k^2\right)\left(\sum_{k=1}^{3} b_k^2\right).$
2. De même en remplaçant 3 par n :
$$\begin{aligned}
0\leqslant & \sum_{1\leqslant i<j\leqslant n}(a_i b_j-a_j b_i)^2\\
0\leqslant & \sum_{i=1}^{n}\sum_{j\neq i} a_i^2 b_j^2+\sum_{i=1}^{n}(a_i b_i)^2-\sum_{i=1}^{n}(a_i b_i)^2-\sum_{i<j}(a_i a_j b_i b_j)-\sum_{i<j}(a_j a_i b_j b_i)\\
0\leqslant & \sum_{k=1}^{n} a_k^2\sum_{k=1}^{n} b_k^2-\left(\sum_{k=1}^{n}(a_k b_k)\right)^2
\end{aligned}$$
Donc $\left(\sum_{k=1}^{n} a_k b_k\right)^2\leqslant\left(\sum_{k=1}^{n} a_k^2\right)\left(\sum_{k=1}^{n} b_k^2\right)$

3. Il y a égalité, si et seulement si (il faut bien être attentif à chaque signe !) (si $b_n \neq 0$) :

$$\forall\, i<j, \qquad a_i b_j - a_j b_i = 0 \Longleftrightarrow \forall\, i \leqslant n a_i = \frac{a_n}{b_n} b_i$$

Exercice

Ecrire le développement polynomiale de

$$\prod_{i=1}^{n}(x-a_i)$$

Correction

Il s'agit d'une fonction polynomiale. Plus largement (en notant $|I|$, le cardinal de I) :

$$\prod_{i=1}^{n}(x_i-a_i) = \sum_{I\subset\mathbb{N}_n}\left(\prod_{i\in I} x_i \times \prod_{j\in\mathbb{N}_n\setminus I}(-a_j)\right)$$

$$\prod_{i=1}^{n}(x-a_i) = \sum_{I\subset\mathbb{N}_n}\left(\prod_{i\in I} x \times \prod_{j\in\mathbb{N}_n\setminus I}(-a_j)\right) = \sum_{I\subset\mathbb{N}_n}\left(x^{|I|}\times(-1)^{n-|I|}\prod_{j\in\mathbb{N}_n\setminus I} a_j\right)$$

On range alors en fonction des valeurs du cardinal de I, afin de trouver des monômes de même degré.

Il s'agit d'une sommation par paquets $\{I\subset\mathbb{N}_n\} = \biguplus_{k=0}^{n}\{I\subset\mathbb{N}_n, |I|=k\}$:

$$\prod_{i=1}^{n}(x-a_i) = \sum_{k=0}^{n}\left[\sum_{\substack{I\subset\mathbb{N}_n\\|I|=k}}\left(x^{|I|}\times(-1)^{n-|I|}\prod_{j\in\mathbb{N}_n\setminus I} a_j\right)\right] = \sum_{k=0}^{n}(-1)^{n-k}x^k\left(\sum_{\substack{I\uplus J=\mathbb{N}_n\\|I|=k}}\prod_{j\in J} a_j\right)$$

Cela peut s'écrire aussi :

$$\prod_{i=1}^{n}(x-a_i) = \sum_{k=0}^{n}(-1)^{n-k}x^k\left(\sum_{\substack{J\subset=\mathbb{N}_n\\|J|=n-k}}\prod_{j\in J} a_j\right)$$

3. Coefficients binomiaux et formule du binôme

3.1. Factorielles et coefficients binomiaux

Définitions

Les remarques en marge permettent de s'intéresser aux nombres :

Définition - Factorielle et coefficient binomial

Pour n et p éléments de $\mathbb{N}$, $p\leqslant n$, on pose :

$0! = 1$ et pour $n\geqslant 1$, $n! = n\times(n-1)\times\cdots\times 1$ qui se lit "factorielle n"

$\binom{n}{p} = \frac{n(n-1)\ldots(n-p+1)}{p!} = \frac{n!}{p!(n-p)!}$ qui se lit « p parmi n »

On généralise la notation à tout $p\in\mathbb{Z}$: si $p<0$ ou $p>n$ (si on n'a pas $0\leqslant p\leqslant n$), alors $\binom{n}{p}=0$.

Remarque - Plus tard...

• Il n'est pas évident que pour $p\leqslant n$, $\binom{n}{p}$ est un nombre entier.
Si tel est le cas (on le verra plus loin) cela signifie que $p!$ divise tout nombre de la forme $n(n-1)\ldots(n-p+1)\ldots$.

• Nos reprendrons la notion de coefficient binomial lorsque nous ferons du dénombrement. Cela expliquera véritablement l'origine de ce calcul.

Informatique - Calcul de la factorielle avec une boucle

```
def factorielle(n):
    f=1
    for k in range(1,n):
        f=f*(k+1)
    return(f)
```

Pour aller plus loin - Coefficients binomiaux réels

On sent qu'on pourrait s'intéresser plutôt au nombres

$$\binom{x}{p} = \frac{x(x-1)\ldots(x-p+1)}{p!}$$

avec $x\in\mathbb{C}$ et $p\in\mathbb{N}$.
Dans ce cas, la méthode du triangle de Pascal doit s'adapter, alors que la formule du binôme de Newton reste vraie!

Propriétés immédiates

Proposition - Propriétés
Pour tout nombres entiers $n \in \mathbb{N}$ (naturels) et $p \in \mathbb{Z}$ (relatifs) :

$$\binom{n}{0} = \binom{n}{n} = 1; \quad \binom{n}{1} = n; \quad \binom{n}{2} = \frac{n(n-1)}{2};$$
$$\binom{n}{p} = \binom{n}{n-p}$$
$$\binom{n}{p} = \frac{n}{p}\binom{n-1}{p-1} \text{ (fausse pour } p = 0\,!)$$
$$\binom{n}{p} = \binom{n-1}{p-1} + \binom{n-1}{p} \text{ (Relation de Pascal)}$$

Démonstration
Les premiers résultats sont immédiats (simple jeu d'écriture).
Le second est assez simple, mais il faut différencier : $p \in [\![0, n]\!]$ ou $p < 0 \iff n - p > n$ ou $p > n \iff n - p > 0$.
Démontrons les deux derniers. Si $1 \leqslant p \leqslant n$,

$$\binom{n}{p} = \frac{n!}{p!(n-p)!} = \frac{n(n-1)!}{p(p-1)!((n-1)-(p-1))!} = \frac{n}{p}\binom{n-1}{p-1}$$

si $p < 0$, alors $\binom{n}{p} = 0$ et $\binom{n-1}{p-1} = 0$; si $p > n$, alors $\binom{n}{p} = 0$ et $\binom{n-1}{p-1} = 0$.
Plus intéressant, si $1 \leqslant p \leqslant n-1$,

$$\binom{n-1}{p-1} + \binom{n-1}{p} = \frac{(n-1)!}{(p-1)!(n-p)!} + \frac{(n-1)!}{p!(n-1-p)!} = \frac{(n-1)![p+(n-p)]}{p!(n-p)!} = \binom{n}{p}$$

Si $p = n$: $\binom{n-1}{p-1} + \binom{n-1}{p} = 1 + 0 = 1 = \binom{n}{n}$.
Si $p > n$: $\binom{n-1}{p-1} + \binom{n-1}{p} = 0 + 0 = 0 = \binom{n}{n}$.
De même si $p < 0$. □

Exercice

Pour n, p, simplifier $\sum_{k=p}^{n} \binom{k}{p}$. *On pourra y « voir » un télescopage*

Correction

$$\sum_{k=p}^{n} \binom{k}{p} = \binom{p}{p} + \sum_{k=p+1}^{n} \left(\binom{k+1}{p+1} - \binom{\mathbb{K}}{p+1} \right) = 1 + \binom{n+1}{p+1} - \binom{p+1}{p+1} = \binom{n+1}{p+1}$$

Histoire - Blaise Pascal

Blaise PASCAL *(1623-1662) est un français, génie des mathématiques (mais pas uniquement). Il redémontre tout Euclide, seul à 12 ans. Il fonde la géométrie projective (avec Desargues) et la « géométrie du hasard » avec Fermat.*
Mais c'est aussi l'inventeur de la brouette, de la première machine à calculer ou du premier transport en commun parisien…
Il présente le triangle de Pascal dans le traité du triangle arithmétique (mais ce n'est pas lui le premier à le découvrir).

3.2. Triangle de Pascal

De ces propriétés on déduit un moyen simple de calculer les coefficients binomiaux :

Savoir faire - Triangle de Pascal
On peut alors construire le triangle de Pascal pour pouvoir calculer facilement (addition et non multiplication) les coefficients binomiaux. On écrit ainsi dans un tableau :

$$\begin{array}{ccccccccc}
\binom{0}{0} & & & & & & 1 & & \\
\binom{1}{0} & \binom{1}{1} & & & & & 1 & 1 & \\
\binom{2}{0} & \binom{2}{1} & \binom{2}{2} & & & & 1 & 2^{=1+1} & 1 \\
\binom{3}{0} & \binom{3}{1} & \binom{3}{2} & \binom{3}{3} & & \longrightarrow & 1 & \underline{3} & \underline{3} & 1 \\
\binom{4}{0} & \binom{4}{1} & \binom{4}{2} & \binom{4}{3} & \binom{4}{4} & & 1 & 4 & \underline{6^{=3+3}} & 4 & 1 \\
\binom{5}{0} & \binom{5}{1} & \binom{5}{2} & \binom{5}{3} & \binom{5}{4} & \binom{5}{5} & 1 & 5 & 10 & 10 & 5 & 1 \\
 & \vdots & & & \vdots & & & \vdots & & & \vdots &
\end{array}$$

On l'a déjà dit :

Proposition - Nombres entiers

Pour n et p éléments de $\mathbb{N}$, $p \leqslant n$, $n!$ et $\binom{n}{p}$ sont des entiers naturels.

Démonstration

Pour la factorielle : il s'agit d'un produit d'entiers naturels.
Pour le coefficient binomial, on réalise une récurrence sur le niveau n.
On pose, pour tout $n \in \mathbb{N}$, $\mathscr{P}_n$: « $\forall\, p \in \mathbb{Z}$, $\binom{n}{p} \in \mathbb{N}$.

- $\mathscr{P}_0$ est vraie car $\binom{0}{0} = 1$ et pour tout $p \neq 0$, $\binom{n}{p} = 0$.
- Soit $n \in \mathbb{N}$. Supposons que $\mathscr{P}_n$ est vraie.
  Soit $p \in \mathbb{Z}$, alors $\binom{n+1}{p} = \binom{n}{p} + \binom{n}{p-1}$, addition de deux entiers d'après $\mathscr{P}_n$.
  Donc $\mathscr{P}_{n+1}$ est vraie.

□

Informatique - Triangle de Pascal

En exploitant des listes (de listes) en informatique, il est possible de créer la n^{e} ligne du triangle de Pascal.

```
def Pascal(n):
    L=[0]*(n+1)
    for h in range(n+1):
        L[h]=[1]+[0]*n
    print(L)
    for h in range(n):
        for k in range(h+1):
            L[h+1][k+1]=L[h][k]+L[h][k+1]
    return(L)
```

Histoire - Isaac Newton

On attribue couramment à Isaac NEWTON *(1642-1707) la révolution scientifique de la science moderne, grâce à son travail mathématique sur les calculs différentiel et intégral et son travail en mécanique (relation fondamentale de la dynamique, loi des forces…).*
La formule du binôme qui apparait ici est en fait du à Pascal (1654), et généralisée une première fois par Newton (avec son travail d'interpolation polynomiale - 1671) puis par Abel avec tout exposant $n \in \mathbb{R}$ au début du XIX siècle.

3.3. Formule du binôme

Proposition - Formule du binôme (de Newton)

Soient $n \in \mathbb{N}$ et a, b deux réels ou deux complexes (ou éléments de tout anneau et tels que $a \times b = b \times a$). Alors :

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

Avec $a=b=1$:

Corollaire -

$$\sum_{p=0}^{n}\binom{n}{p}=2^n$$

Démonstration
Démontrons ce résultat par récurrence.

Notons, pour tout $n\in\mathbb{N}$, $\mathscr{P}_n$: « $(a+b)^n=\sum_{k=0}^{n}\binom{n}{\mathbb{K}}a^{\mathbb{K}}b^{n-k}$ ».

— $(a+b)^0=1$ et $\sum_{k=0}^{0}\binom{n}{\mathbb{K}}a^{\mathbb{K}}b^{n-k}=\binom{0}{0}a^0b^0=1$.
donc $\mathscr{P}_0$ est vraie.

— Soit $n\in\mathbb{N}$, supposons que $\mathscr{P}_n$ est vraie.
On a alors

$$(a+b)^{n+1}=(a+b)\times\sum_{k=0}^{n}\binom{n}{\mathbb{K}}a^{\mathbb{K}}b^{n-k}=\binom{n}{n}a^{n+1}+\underbrace{\sum_{k=0}^{n-1}\binom{n}{\mathbb{K}}a^{k+1}b^{n-k}}_{h=k+1}+\underbrace{\sum_{k=1}^{n}\binom{n}{\mathbb{K}}a^{\mathbb{K}}b^{n-k+1}}_{h=k}+\binom{n}{0}b^{n+1}$$

$$=a^{n+1}+\sum_{h=1}^{n}\left(\binom{n}{h-1}+\binom{n}{h}\right)a^hb^{n+1-h}+b^{n+1}=\sum_{k=0}^{n+1}\binom{n+1}{\mathbb{K}}a^{\mathbb{K}}b^{n+1-k}$$

d'après la formule de Pascal et comme $\binom{n+1}{n+1}=1=\binom{n+1}{0}$.
Donc $\mathscr{P}_{n+1}$ est alors vérifiée.

On a ainsi démontré par récurrence le résultat attendu. □

Exercice

Calculer $\sum_{0\leqslant p\leqslant n;p\text{ pair}}\binom{n}{p}$ et $\sum_{0\leqslant p\leqslant n;p\text{ impair}}\binom{n}{p}$.

Correction

Notons P_n et I_n ces deux sommes. On a vu que $P_n+I_n=\sum_{k=0}^{n}\binom{n}{\mathbb{K}}=2^n$.

Puis, $I_n=\sum_{0\leqslant p\leqslant n;p\text{impair}}\binom{n}{p}=\sum_{0\leqslant p\leqslant n;p\text{impair}}\binom{n-1}{p}+\sum_{0\leqslant p\leqslant n;p\text{impair}}\binom{n-1}{p-1}=I_{n-1}+P_{n-1}=2^{n-1}$.

Et ainsi $P_n=2^n-I_n=2^n-2^{n-1}=2^{n-1}(2-1)=2^{n-1}=I_n$

Pour aller plus loin - De qui parle-t-on ?

« Il est de bonne famille et il a reçu une excellente éducation. Prodigieusement doué pour les mathématiques, à vingt et un ans il publiait une étude sur le binôme de Newton, qui fit sensation dans toute l'Europe et lui valut de devenir titulaire de la chaire de mathématiques dans une de nos petites universités. Tout donnait à penser qu'il allait faire une carrière extrêmement brillante. Mais l'homme avait une hérédité chargée, qui faisait de lui une sorte de monstre, avec des instincts criminels d'autant plus redoutables qu'ils étaient servis par une intelligence exceptionnelle. Des bruits fâcheux coururent bientôt sur lui dans l'Université, qui l'obligèrent à se démettre. Il vint à Londres où il se mit à donner des cours destinés aux officiers de l'armée. »

4. Bilan

Synthèse

⇝ En mathématique, la sélection (naturelle) opère également et un langage se crée. Des définitions et concepts sont sélectionnés, comme des mots de la langue ; et le formalisme comme l'écriture de ce langage.
Le formalisme doit être compact (souvent), ressemblant à sa notion attachée (rarement) et efficace calculatoirement (absolument). C'est le cas de la notation $\sum$ (ou $\prod$) qu'on utilise comme un « auto-mathisme » avec un peu d'habitude : changement de variable, sommation par paquets, télescopage, somme multiple…

⇝ Pour bien comprendre cette utilisation, on se rend compte que l'enjeu est de maîtriser la manipulation de l'ensemble des indices. Cette notion d'ensemble est au coeur des mathématiques, nous reviendrons dessus aux chapitres 9 et 10. (Il faut aussi que l'ensemble des nombres soit commutatif).

⇝ Le nombres de sous-ensemble à k éléments à partir d'un ensemble à n élément est également un calcul qui se présente dans de très nombreuses branches des mathématiques. Il est formalisé par le coefficient binomial et une expression de langage : « k parmi n ».

⇝ De nombreuses propriétés (issues de domaines variés) peuvent lui être attachés : triangle de Pascal, binôme de Newton, nombre de chemins dans des arbres binaires fusionnés, inversion de Pascal…

Savoir-faire et Truc & **Astuce du chapitre**

- Savoir-faire - Changement d'indice
- Savoir-faire - Sommation par paquets
- Savoir-faire - Méthode de téléscopage
- Truc & Astuce pour le calcul - Ecrire un programme pour mieux comprendre
- Truc & Astuce pour le calcul - Invariant de boucle
- Savoir-faire - Somme multiple (indépendance)
- Savoir-faire - Somme multiple (dépendance)
- Savoir-faire - Triangle de Pascal

Notations

	Propriétés	*Remarques*
Iverson pour la propriété $\mathscr{P}$	$\forall\, i$, $[\mathscr{P}(i)] \in \{0,1\}$ avec $[\mathscr{P}(i)] = 1$ ssi $\mathscr{P}(i)$ vraie	
de n (lire dans ce sens !)	$n! = \prod_{k=1}^{n} k$	$0! = 1$ et par récurrence $n! = n \times (n-1)!$
binomial de k parmi n	$\binom{n}{k} = \frac{n!}{k!(n-k)!} = \frac{(n-k+1)_k}{k!}$ (si n n'est pas entier)	C'est le nombre de sous-ensemble à k éléments pris dans un ensemble E à n éléments.
	$(1+x)^r = \sum_{k=0}^{+\infty} \binom{k}{r} x^k$ (même si $r \notin \mathbb{N}\ldots$)	

Retour sur les problèmes

7. Nombres triangulaires. $T_n = \frac{n(n+1)}{2} = \binom{n+1}{2}$. Si on note T_n^k, telle que $T_n^k = \sum_{h=1}^{n} T_h^{k-1}$ et $T_n^0 = n$.
   On trouve alors par récurrence : $T_n^k = \binom{n+k}{k+1}$.
8. A calculer... Il n'y a pas de résultats intéressants, sauf en petite taille ou en expression formelle.
9. Suite de nombres. Et le suivant ?
   Une stratégie : apprendre des évolutions sur les premiers nombres, donc calculer la suite $\delta(u)_n = u_{n+1} - u_n$. Pour en déduire $u_m = u_{m-1} + \delta(u)_{m-1}$.
   Si ce n'est pas suffisant, évaluer (par récurrence) : $\delta^k(u)_n = \delta^{k-1}(u)_{n+1} - \delta^{k-1}(u)_n \ldots$
10. Interpolation à pas constant. Voir activités
11. Développement de puissance.
    C'est le coefficient multinomial de Newton : $(a_1 + a_2 + \cdots + a_k)^n = \sum_{m_1+m_2+\cdots+m_k=m} \frac{n!}{m_1!m_2!\ldots m_k!} \prod_{i=1}^{k} a_i^{m_i}$.
    Vous comprenez ?

Chapitre 3

Résolution de systèmes linéaires.

Résumé -

Il n'est pas rare également d'avoir à résoudre des équations algébriques sous forme de systèmes d'équations linéaires.
Le but de ce (tout petit) chapitre est de mettre en place les définitions efficaces, les méthodes (algorithmes) de résolution adéquates (ce que n'est pas la substitutiuon…) et de voir apparaître par le calcul bien mené, la structure sous-jacente. On voit ici l'exemple typique du calcul linéaire, ou matriciel et la structure clé d'espaces vectoriels.

Enfin, voici une liste de petites vidéo ou conférence visionnable sur internet et en lien avec le sujet. A visionner à loisir :

- *Exo7Math - Systèmes linéaires - partie 1 : introduction. https ://www.youtube.com/watch ?v=0uYJ3RNL5SU*
- *Mathéma-TIC - Règle de Cramer. https ://www.youtube.com/watch ?v=CNzVk1evqac*

Sommaire

1. Quelques problèmes

? Problème 12 - Résolution d'un système linéaire

Il n'est pas rare que l'on rencontre en SI ou en physique (ou en mathématiques) un système d'équation de la forme suivante à résoudre :

$$\begin{cases} 2x & + & 3y & - & 4z & = & 1 \\ x & - & 3y & + & z & = & -1 \\ x & + & y & - & z & = & 1 \end{cases}$$

Est-il possible de trouver la/les solution(s) en (x,y,z) de ce système directement ?
On peut croire et anticiper que ces solutions s'expriment sous la forme d'un calcul différent (mais équivalent, d'une certaine façon) pour x, y et z, à partir des nombres $2;3;-4;1\ldots$
Comment expliciter ce calcul ?

? Problème 13 - Résolution « au petit bonheur »

Lorsque le système est gros (plus de quatre équations), les méthodes aléatoires de résolution ne fonctionnent pas bien.
Il faut s'organiser. Qu'avons-nous le droit de faire ? Existe-t-il une méthode, un algorithme (programmable informatiquement, par exemple) qui donne à coup sûr l'ensemble des solutions d'un système d'équations linéaires ?

? Problème 14 - Forme de l'ensemble des solutions

Dans la pratique, lorsqu'on rencontre un système de 3 équations à 3 inconnues, on trouve une unique solution.
Est-ce toujours vrai ? Sinon, qu'est-ce qui fait que cela peut être faux ?
Et, plus généralement, s'il y a n équations et p inconnues, combien y a-t-il de solutions ?

? Problème 15 - Impact des paramètres

Il arrive aussi en science appliquée (et en mathématiques (en interne) également) que l'on trouve un système à paramètre. Par exemple :

$$\begin{cases} ax & + & 3y & - & 4z & = & 1 \\ (a-1)x & - & 3y & + & z & = & -1 \\ x & + & y & - & z & = & 1 \end{cases}$$

Dans ce cas là, à quoi ressemble l'ensemble des solutions ?

2. Systèmes linéaires. Equivalence

Dans cette partie, même si nous énonçons des résultats (sous forme de définitions et propositions à apprendre), nous ne faisons pas (encore) de démonstration comme annoncé au cours inaugural.

2.1. Vocabulaire

Remarque - Le formalisme de LEIBNIZ

On commence par **paramétrer notre problème** : on l'élargit en donnant une écriture symbolique (paramétre lettré) aux nombres. Puis on élargit la problème : pourquoi seulement trois équations et trois inconnues?
Pour permettre l'étude systématique des systèmes linéaires, on a besoin d'une suite de nombre doublement indexé les $(a_{i,j})$. C'est Leibniz qui en a eu l'idée (1678). Et comme souvent, à partir du moment où le formalisme et les définitions associées sont bons, les théorèmes tombent comme des fruits mûrs...

Définition - Système linéaire de n équations à p inconnues

Un **système linéaire** à n équations et p inconnues à coefficient dans $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$ est un système S d'équations de la forme :

$$\begin{cases} a_{1,1}x_1 + a_{1,2}x_2 + \cdots + a_{1,p}x_p &= b_1 \\ a_{2,1}x_1 + a_{2,2}x_2 + \cdots + a_{2,p}x_p &= b_2 \\ \vdots & \vdots \\ a_{n,1}x_1 + a_{n,2}x_2 + \cdots + a_{n,p}x_p &= b_n \end{cases}$$

où
- les $a_{i,j} \in \mathbb{K}$;
- $(b_1, b_2, \dots b_n) \in \mathbb{K}^n$ est appelé le second membre de l'équation ;
- $(x_1, x_2 \dots x_p) \in \mathbb{K}^p$ est appelé l'inconnue du système.

On appelle **système homogène** le système obtenu en remplaçant chaque b_i par 0 (second membre nul).

On appelle **solution** du système S, l'ensemble $\mathscr{S} \subset \mathbb{R}^p$ des p-uplets $(\overline{x}_1, \overline{x}_2, \dots \overline{x}_p)$ qui vérifient les n équations :
$\forall\, i \in \mathbb{N}_n,\ a_{i,1}\overline{x}_1 + a_{i,2}\overline{x}_2 + \cdots + a_{i,p}\overline{x}_p = b_i$

Histoire - Mathématiques des neuf livres

Le plus vieux traité mathématique proposant l'étude de système d'équations linéaires est le Chiu chang suan shu, que l'on appelle les mathématiques des neufs livres, écrit en 160 av. JC. Bien que de nature différentes (moins géométrique et plus algorithmique), ce livre est au mathématiques chinoise, ce que les éléments d'Euclide sont aux mathématiques européennes.

Remarque - Notations

On a pour habitude de noter S les systèmes et de manière plus stylisé l'ensemble des solutions $\mathscr{S}$.
On associe parfois des indices.

Remarque - Solution(s) du système homogène

Un système linéaire homogène admet toujours au moins la solution nulle : $\mathscr{S}_0 = (x_1, x_2 \dots x_p) = (0, 0 \dots, 0) = O_p$

Heuristique - Résolution

Il y a en gros deux méthodes pour résoudre un tel système.
La méthode de Leibniz (fin du XVII) fonctionne très bien pour les petites dimensions $n, p \leqslant 3$ et la formule de Cramer que l'on donnera ensuite. Elle marche bien également *en théorie* pour les grandes dimensions.
La méthode (dite) de Gauss est algorithmique, nous la privilégierons pour les plus grandes dimensions.

2.2. Systèmes équivalents

Analyse - Sens des flèches $\Rightarrow$, $\Leftarrow$

« Naturellement », lorsqu'on écrit $S_1 \Rightarrow S_2$, cela signifie qu'on peut passer du système S_1 au système S_2.
Ainsi, tous les éléments qui vérifient les équations de S_1 vérifient aussi celles de S_2.
En terme d'ensemble solution : toutes les solutions de S_1 sont soolutions de S_2.
Donc $\mathscr{S}_1 \subset \mathscr{S}_2$.
On peut donc résumer : $S_1 \Rightarrow S_2$ si et seulement si $\mathscr{S}_1 \subset \mathscr{S}_2$.

Définition - Systèmes équivalents
Deux systèmes S_1 et S_2 sont dits équivalents, relation notée

$$S_1 \iff S_2$$

si et seulement si les ensembles de solutions sont les mêmes : $\mathscr{S}_1 = \mathscr{S}_2$.

Exercice

Les systèmes $S_1 : \begin{cases} 2x+y=1 \\ -x-y=0 \end{cases}$ et $S_2 : \begin{cases} 2x+y=1 \end{cases}$ sont-ils équivalents ?

Correction
Non.
Le couple $(0,1) \in \mathscr{S}_2$, c'est une solution du second système d'équations mais pas du premier car $0-1 \neq 0$

Attention - Pas d'abus
Typiquement ici, il ne faut pas abuser du symbole d'équivalence !
Dans l'exercice, on écrit donc jamais :

$$S_1 : \begin{cases} 2x+y=1 \\ -x-y=0 \end{cases} \iff 2x+y=1$$

Il n'y a ici qu'une implication.

3. Résolution explicite. cas des petits systèmes $n = p = 2$ ou $n = p = 3$

3.1. Vers la formule de Cramer

Etude formelle du système simple $n = p = 2$

Comme Leibniz, commençons par étudier le système

$$S : \quad \begin{cases} ax & +by & = \alpha \\ cx & +dy & = \beta \end{cases}$$

où a, b, c, d et α, β sont des paramètres, alors que x, y sont les inconnues.

Analyse - Etude « à la lycéenne »
On substitue (évidemment : **si** $a \neq 0$) :

$$S \Longrightarrow x = \frac{\alpha - by}{a}$$

Donc (on a définitivement perdu l'équivalence) (**si** $ad-bc \neq 0$) :

$$S \Longrightarrow \frac{c\alpha - cby}{a} + dy = \beta \Longrightarrow \frac{ad-bc}{a} y = \frac{a\beta - c\alpha}{a} \Longrightarrow y = \frac{a\beta - c\alpha}{ad-bc}$$

On trouve alors en réinjectant : $S \Longrightarrow x = \dfrac{\alpha(ad-bc) - b(a\beta - c\alpha)}{a(ad-bc)} = \dfrac{\alpha d - \beta b}{ad-bc}$.

Remarque - Avons-nous trouver les solutions ?
Non !!
Ici, on a $\mathscr{S} \subset \left\{ \left(\dfrac{\alpha d - \beta b}{ad-bc}, \dfrac{-\alpha c + \beta a}{ad-bc} \right) \right\}$.
Mais peut-être que $\mathscr{S}$ est l'ensemble vide !
Il faut donc faire une réciproque. Sinon, il y a une faute (courante) de logique.

Analyse - Réciproque
Toujours dans le cas $ad-bc$, on vérifie que

$$a\frac{\alpha d - \beta b}{ad-bc} + b\frac{-\alpha c + \beta a}{ad-bc} = \alpha \frac{ad-bc}{ad-bc} = \alpha$$
$$c\frac{\alpha d - \beta b}{ad-bc} + d\frac{-\alpha c + \beta a}{ad-bc} = \beta \frac{ad-bc}{ad-bc} = \beta$$

Donc la réciproque est vérifiée. On a bien l'égalité : $\mathscr{S} = \left\{\left(\frac{\alpha d - \beta b}{ad - bc}, \frac{-\alpha c + \beta a}{ad - bc}\right)\right\}$.

Remarque - $a \neq 0$?

Dans l'analyse, il a fallut faire séparer le cas $a \neq 0$.
Mais le résultat obtenu ne dépend pas de $a = 0$ ou $a \neq 0$. Donc on aurait sûrement pu s'en passer.
D'ailleurs, la réciproque montre que ce n'est pas un cas important.
Seule la situation $ad - bc \neq 0$ ou $ad - bc = 0$ est importante !

Définition - Déterminant d'un système 2×2
On appelle **déterminant du système** 2×2 (i.e. $n = 2$ et $p = 2$)

$$S \quad \begin{cases} a_{1,1}x_1 + a_{1,2}x_2 &= b_1 \\ a_{2,1}x_1 + a_{2,2}x_2 &= b_2 \end{cases}$$

le nombre $\delta_S = a_{1,1}a_{2,2} - a_{1,2}a_{2,1}$ souvent noté $\begin{vmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{vmatrix}$.

Cas d'un système 3×3

Leibniz propose d'étendre sa méthode pour $n > 2$.

Définition - Déterminant d'un système 3×3
On appelle **déterminant du système** 3×3 (i.e. $n = 3$ et $p = 3$)

$$S \quad \begin{cases} a_{1,1}x_1 + a_{1,2}x_2 + a_{1,3}x_3 &= b_1 \\ a_{2,1}x_1 + a_{2,2}x_2 + a_{2,3}x_3 &= b_2 \\ a_{3,1}x_1 + a_{3,2}x_2 + a_{3,3}x_3 &= b_3 \end{cases}$$

le nombre $\delta_S = a_{1,1}a_{2,2}a_{3,3} + a_{1,2}a_{2,3}a_{3,1} + a_{1,3}a_{2,1}a_{3,2} - a_{1,2}a_{2,1}a_{3,3} - a_{3,1}a_{2,2}a_{1,3} - a_{1,1}a_{3,2}a_{2,3}$ souvent noté $\begin{vmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{vmatrix}$.

Pour aller plus loin - Extension de taille n
Lorsque l'on souhaite généraliser la notion de déterminant, on comprend qu'il faut prendre une combinaison linéaire des termes constitués exactement de nombre tous pris dans des lignes et des colonnes différentes.
Mais il y a une règle de signe à respecter. Il n'est pas clair que Leibniz ait bien compris cette règle.
Cramer, 50 ans plus tard l'a retrouvé (sans avoir connaissance des travaux de Leibniz).
Un mathématicien japonais SEKI KOWA *(1642-1708) a effleuré toute cette découverte.*

Formule de Cramer

Nous verrons dans le cours sur les déterminants, au second semestre, une petite notice sur Gabriel Cramer (1704-1752)

Savoir faire - Formule de CRAMER

Si le déterminant du système S (n=p=2) est non nul, la solution de $\mathscr{S}$ de

$$(S) \quad \begin{cases} a_{1,1}x_1 + a_{1,2}x_2 &= b_1 \\ a_{2,1}x_1 + a_{2,2}x_2 &= b_2 \end{cases}$$

est $\mathscr{S} = \{(\overline{x}_1, \overline{x}_2)\} = \left\{\left(\frac{\begin{vmatrix} b_1 & a_{1,2} \\ b_2 & a_{2,2} \end{vmatrix}}{\begin{vmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{vmatrix}}, \frac{\begin{vmatrix} a_{1,1} & b_1 \\ a_{2,1} & b_2 \end{vmatrix}}{\begin{vmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{vmatrix}}\right)\right\}$.

Si le déterminant du système S (n=p=3) est non nul, la solution de $\mathscr{S}$ de

$$(S) \quad \begin{cases} a_{1,1}x_1 + a_{1,2}x_2 + a_{1,3}x_3 &= b_1 \\ a_{2,1}x_1 + a_{2,2}x_2 + a_{2,3}x_3 &= b_2 \\ a_{3,1}x_1 + a_{3,2}x_2 + a_{3,3}x_3 &= b_3 \end{cases}$$

Pour aller plus loin - Sytème de taille $n \times n$
Dans le calcul du déterminant qui donne x_i, on remplace la i-ème colonne par la colonne du second membre.
La formule de Cramer se généralise à tout système carré, avec une bonne définition du déterminant généralisée.

$$\text{est } \mathscr{S} = \{(\overline{x}_1, \overline{x}_2, \overline{x}_3)\} = \left\{ \left(\frac{\begin{vmatrix} b_1 & a_{1,2} & a_{1,3} \\ b_2 & a_{2,2} & a_{2,3} \\ b_3 & a_{3,2} & a_{3,3} \end{vmatrix}}{\begin{vmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{vmatrix}}, \frac{\begin{vmatrix} a_{1,1} & b_1 & a_{1,3} \\ a_{2,1} & b_2 & a_{2,3} \\ a_{3,1} & b_3 & a_{3,3} \end{vmatrix}}{\begin{vmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{vmatrix}}, \frac{\begin{vmatrix} a_{1,1} & a_{1,2} & b_1 \\ a_{2,1} & a_{2,2} & b_2 \\ a_{3,1} & a_{3,2} & b_3 \end{vmatrix}}{\begin{vmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{vmatrix}} \right) \right\}$$

4. Algorithme su pivot de GAUSS

Heuristique - Il y a une bonne méthode et une mauvaise méthode…

La **mauvaise méthode** est la SUBSTITUTION.
On perd des équations et cela est **plus long**.
La **bonne méthode** consiste à faire des OPERATIONS ELEMENTAIRES.
Il faut les connaître, mais ce n'est pas sufffisant.
Il faut aussi savoir **bien** les mener. C'est la partie difficile de l'algorithme.

4.1. Systèmes équivalents : opérations élémentaires

On commence par trouver des invariants de l'ensemble des solutions. Cela permet de raisonner avec des systèmes équivalents

Définition - Opérations élémentaires
On appelle **opérations élémentaires** sur le système dont la ligne i est noté L_i, les opérations suivantes :
- pour tout $i, j \leqslant n$, échange des lignes L_i et L_j codé : $L_i \leftrightarrow L_j$
- pour tout $i \leqslant n$, $\lambda \neq 0$, la multiplication de la ligne L_i par λ codé : $L_i \leftarrow \lambda L_i$
- pour tout $i, j \leqslant n$, $\alpha \in \mathbb{K}$, l'ajout à la ligne L_i de α fois la ligne L_j codé : $L_i \leftarrow L_i + \alpha L_j$

Proposition - Invariant des solutions
Les opérations élémentaires conservent exactement les solutions du système

Démonstration
Il est évident que chaque opération conduit à une implication $\Rightarrow$ sur les systèmes.
Et par ailleurs, pour chacune de ces opérations, il est possible de « revenir en arrière ».
Pour cela, il suffit de voir que

1. l'opération $L_i \leftrightarrow L_j$ répétée deux fois redonne le système initiale.
2. l'opération $L_i \leftarrow \frac{1}{\lambda} L_i$ ($\lambda \neq 0$), après l'opération $L_i \leftarrow \lambda L_i$, redonne le système initiale.
3. l'opération $L_i \leftarrow L_i - \alpha L_j$, après l'opération $L_i \leftarrow L_i + \alpha L_j$, redonne le système initiale.

On trouve donc $\mathscr{S} \Rightarrow \mathscr{S}' \Rightarrow \mathscr{S}$.
Cela signifie bien : $\mathscr{S} \Longleftrightarrow \mathscr{S}'$. □

4.2. Algorithme du pivot de GAUSS

Il reste à bien exploiter ces opérations élémentaires afin de **toujours** trouver la solution d'un système linéaire.

Histoire - Algorithme de Fang-Cheng
On ne prête qu'au riche : en occident, nous associons l'algorithme ici étudié (central en mathématiques de l'algèbre linéaire) au grand Carl-Freidrich Gauss, mais il semble qu'il soit bien connu et expliqué dans les mathématiques des neuf livres (première impression en 1084, 5 siècle avant Gutemberg…).
L'auteur de cet ouvrage est un certain CHANG TS'ANG, *et sa méthode s'appelle le modèle rectangulaire écrit Fang-Cheng.*

Savoir faire - Algorithme du pivot de GAUSS
Pour résoudre un système linéaire, on applique des opérations élémentaires pour le rendre triangulaire.

1. On cherche un coefficient non nul dans la première colonne (devant la première inconnue) le plus simple possible (on va devoir

diviser par ce nombre).
Si tous les coefficients sont nuls, on passe à l'inconnue suivante.

2. On échange la ligne où l'on a trouvé ce coefficient avec la ligne 1.
3. Pour $j \in [\![2,n]\!]$, on effectue l'opération $L_j \leftrightarrow L_j - \dfrac{a_{1,j}}{a_{1,1}} L_1$ (ce qui permet d'annuler le coefficient devant x_1 pour la ligne j puis pour toute la colonne.
4. On recommence la première étape, en oubliant la première équation et en s'intéressant à l'inconnue suivante, tant qu'il reste des inconnues.

Après avoir appliqué cet algorithme, le système obtenu est triangulaire, et l'on peut déterminer ses solutions en partant de la dernière équation et en remontant à la première (sans substitution : cela demande un « mouvement » supplémentaire…).
Il peut arriver que l'on retrouve en dernière équation plus d'une inconnue. Dans ce cas, on garde une seule inconnue, les autres deviennent des variables libres que l'on considère alors en second membre.

4.3. Applications. Différents formes de l'ensemble des solutions

Application - Trois systèmes à résoudre : $S_1 : \begin{cases} x & -y & +z & =1 \\ -x & -y & +z & =3 \\ 2x & +y & +z & =0 \end{cases}$,

$S_2 : \begin{cases} x & +y & +2z & =1 \\ 2x & -y & +z & =-1 \\ x & -2y & -z & =-2 \end{cases}$ **et** $S_3 : \begin{cases} x & +y & +2z & =1 \\ 2x & -y & +z & =-1 \\ x & -2y & -z & =2 \end{cases}$

On applique la méthode du pivot de Gauss, on trouve :

$$S_1 \Leftrightarrow \left\{ \begin{array}{rrrl} \underline{x} & -y & +z & =1 \\ & -2y & +2z & =4 \\ & 3y & -z & =-2 \end{array} \right| \begin{array}{l} \\ L_2 \leftarrow L_2 + L_1 \\ L_3 \leftarrow L_3 - 2L_1 \end{array}$$

$$\Leftrightarrow \left\{ \begin{array}{rrrl} x & -y & +z & =1 \\ & \underline{-2y} & +2z & =4 \\ & & 2z & =4 \end{array} \right| \begin{array}{l} \\ \\ L_3 \leftarrow L_3 + \frac{3}{2} L_2 \end{array} \Leftrightarrow \left\{ \begin{array}{rl} x & =-1 \\ y & =0 \\ z & =2 \end{array} \right.$$

Donc $\mathscr{S}_1 = \{(-1,0,2)\}$
Pour le second système

$$S_2 \Leftrightarrow \left\{ \begin{array}{rrrl} \underline{x} & +y & +2z & =1 \\ & -3y & -3z & =-3 \\ & -3y & -3z & =-3 \end{array} \right| \begin{array}{l} \\ L_2 \leftarrow L_2 - 2L_1 \\ L_3 \leftarrow L_3 - L_1 \end{array}$$

$$\Leftrightarrow \left\{ \begin{array}{rrl} x & +y & =1-2z \\ & y & =3-z \end{array} \right. \Leftrightarrow \left\{ \begin{array}{rl} x & =-2-z \\ y & =3-z \end{array} \right.$$

Donc $\mathscr{S}_2 = \{(-2-z, 3-z, z), z \in \mathbb{R}\}$ que l'on prend l'habitude de noter : $\{(-2,3,0) + z(-1,-1,1), z \in \mathbb{R}\}$.
Pour le troisième système

$$S_3 \Leftrightarrow \left\{ \begin{array}{rrrl} \underline{x} & +y & +2z & =1 \\ & -3y & -3z & =-3 \\ & -3y & -3z & =1 \end{array} \right| \begin{array}{l} \\ L_2 \leftarrow L_2 - 2L_1 \\ L_3 \leftarrow L_3 - L_1 \end{array}$$

$$\Leftrightarrow \left\{ \begin{array}{rrrl} x & +y & 2z & =1 \\ & y & z & =3 \\ & & 0 & =4 \end{array} \right.$$

Donc $\mathscr{S}_3 = \emptyset$.

Remarque - Nombre de solution
Notons de cette application qu'un système carré (ici 3×3) peut avoir :

Pour aller plus loin - Début du second semestre
Ceci est le début historique de l'algèbre linéaire que nous verrons longuement en début de second semestre (avec la question des déterminants). N'anticipons pas trop pour le moment…

— aucune solution
— une unique solution
— une infinité de solution

⚠ **Attention - Lorsqu'il y a infinité de solution...**
↯ il n'y a pas unicité d'écriture de cet ensemble

5. Bilan

Synthèse

⇝ Pour la résolution d'un système linéaire il existe une méthode infaillible : l'algorithme du pivot de Gauss.
Plus la taille du système est grande, plus il est nécessaire d'exploiter cette méthode.

⇝ Pour des systèmes plus petits, on peut exploiter directement la formule de Cramer

⇝ L'ensemble des solutions est soit un ensemble vide, soit l'addition d'une solution particulière et de l'ensemble (*espace vectoriel*) des solutions de l'équation homogène. L'écriture de cet ensemble n'est pas unique.

⇝ Il faut être très attentif au symbole $\Longleftrightarrow$ parfois employé abusivement...

Savoir-faire et Truc & Astuce du chapitre

— Savoir-faire - Formule de CRAMER
— Savoir-faire - Algorithme du pivot de GAUSS

Notations

Notations	*Définitions*	*Propriétés*	*Remarq*
$S_1 \Longleftrightarrow S_2$	Les systèmes S_1 et S_2 ont les mêmes solutions	$\mathscr{S}_1 = \mathscr{S}_2$	On pass rations
$S_1 \Longrightarrow S_2$	Les solutions de S_1 sont des solutions de S_2	$\mathscr{S}_1 \subset \mathscr{S}_2$	On pass « non lé
$\delta_S = a_{1,1}a_{2,2} - a_{1,2}a_{2,1} = \begin{vmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{vmatrix}$	Déterminant du système S (de taille 2) ou de la matrice associée		On exp formule

Retour sur les problèmes

12. Résolution d'un système linéaire. $\mathscr{S} = \{(1,1,1)\}$.
13. Résolution « au petit bonheur ». Voir l'algorithme de Gauss.
14. Forme de l'ensemble des solutions.
    Il peut y avoir :
    — aucune solution (dans ce cas le système n'est pas homogène)
    — une solution de la forme $\mathscr{S} = \{(\overline{x}_1, \dots \overline{x}_p) + \sum_{k=1}^{r} \lambda_k \overrightarrow{u_k}\}$ où r est le rang du système, $(\overline{x}_1, \dots \overline{x}_p)$ est une solution particulière et pour tout k, $\overrightarrow{u_k}$ est un vecteur de $\mathbb{K}^p$, solution non nulle de l'équation homogène.
    (Ces vecteurs sont linéairement indépendants).
15. Impact des paramètres.
    — Si $a = 8$, alors $\mathscr{S} = \emptyset$,
    — Si $a \neq 8$, alors $\mathscr{S} = \left\{\left(\frac{2}{8-a}, \frac{a-16}{a-8}, \frac{10}{8-a}\right)\right\}$

Chapitre 4

Calculs trigonométriques

Résumé -

Il s'agit, pour commencer, de revoir les propriétés des fonctions trigonométriques. Nous choisissons une présentation constructive, selon le sens de l'histoire et de la formation du lycéen. Nous nous appuyons sur les formules de base : $\cos(a+b) = \cos a\cos b - \sin a\sin b$ *et* $\sin(a+b) = \sin a\cos b + \cos a\sin b$ *pour développer toute la trigonométrie.*

Avec l'exponentielle complexe (chapitre 7), les formules seront revues plus efficacement.

Pour résoudre $f(x) = y$*, il faut pouvoir écrire* $x = f^{-1}(y)$*, i.e. trouver la fonction* f^{-1} *réciproque de* f*. On s'intéresse donc à la trigonométrie réciproque (*arcsin, arccos *et* arctan*)*

Vidéos :

- *Kitoumath. Les mathématiques fantastiques - Les formules de trigo à apprendre sans peine. https ://www.youtube.com/watch ?v=lKj1zQpToxA*
- *Micmath - La conjugaison complexe est un automorphisme de corps. https ://www.youtube.com/watch ?v=AVDMpnwsztg*
- *Les maths en finesse - Racines nieme de l'unité. https ://www.youtube.com/watch ?v=aZLGdnktO8k*

Sommaire

1. Problèmes

Pour aller plus loin - Triangles semblables
On dit que deux triangles sont semblables si deux (et donc trois) angles sont de mêmes mesures

Problème 16 - Fonctions définies sur des angles nulles et droits. Et plus loin ?

Pour tout angle θ, les triangles rectangles dont l'un des angles vaut θ sont tous semblables.
Il y a donc un coefficient de proportionnel entre les mesures des côtés de ces triangles, qui dépend uniquement de θ. Prenons un triangle rectangle de référence d'hypothènuse égale à 1.
On note $\cos\theta$ la mesure du côté adjacent et $\sin\theta$ le côté opposé.
Que se passe-t-il si l'angle dépasse 90° ?

Problème 17 - Unité de mesure d'angles

Au début du lycée, on vous a fait changer l'unité de mesure des angles : des degrés à des radians ?
Pourquoi ? Qu'est-ce qu'on y gagne, pour chaque unité ?

Problème 18 - Relation entre les formules de trigonométrie

Quelles relations entre $\cos(a+b)$ et $\cos a$, $\cos b$. Et d'autres ?
De même peut-on linéariser $\cos(a)\sin(b)$ (c'est-à-dire, l'écrire sous forme d'une somme !).
Beaucoup de formules !
Une autre question, non négligeable : comment apprendre toutes ces formules ?

Histoire - Claude Ptolémé
Ptolémée (Ptolémaïs de Thébaïde (Haute-Égypte) vers 90 - Canope vers 168) est un astronome et astrologue grec qui vécut à Alexandrie (Égypte).

Il est connu pour ses apports en géographie et en mathématique (géométrie et trigonométrie).
Même s'il n'exploitait pas les fonctions cos *et* sin *mais plutôt la fonction corde (cord), il donna le premier élan (après Hipparque ?) à la trigonométrique que nous connaissons. Ces travaux ont été repris par les mathématiques indiennes (III à VI siècle) puis les mathématiques arabes (VIII à XIII siècle)*

Problème 19 - Equation polynomiale et trigonométrie

Montrer que pour tout entier n, et pour tout $\theta \in \mathbb{R}$,
$\cos(n\theta)$ s'exprimer comme une fonction polynomiale en $\cos\theta$.
Ce polynôme s'appelle le polynôme de Tchebychev d'ordre n.

Problème 20 - Fonction réciproque

Si souvent, nous aurons besoin d'inverser les relations $\sin\theta = x$ en $\theta = \sin^{-1}(x)$.
Mais, sin ou cos ne sont pas des fonctions bijectives. Comment faire ?

2. Fonctions trigonométriques

2.1. Construction historique

Analyse - Triangles rectangles semblables
D'après le théorème de Thalès, le rapport des longueurs de deux côtés similaires de deux triangles semblables est constant.
Dans l'ensemble des triangles rectangles, deux triangles sont semblables dès qu'ils ont chacun un angle de même mesure.
Ainsi les rapports de deux côtés consécutifs dans un triangle rectangle ne dépendent que de la mesure d'un angle (non droit).

Ce qui donne une première définition, *non encore totalement satisfaisante* :
Soit θ un angle compris entre 0 et 90 degrés.

Soit ABC un triangle rectangle en A et tel que $\widehat{ABC} = \theta$.

Alors le rapport $\dfrac{AB}{BC}$ est indépendant du triangle considéré, noté $\cos^\circ \theta$.

De même $\dfrac{AC}{BC}$ est indépendant du triangle considéré, noté $\sin^\circ \theta$.

Et $\dfrac{AC}{AB}$ est indépendant du triangle considéré, noté $\tan^\circ \theta$.

La notation choisie ici n'est pas standardisée, et est sûrement contradictoire avec celle vue en fin de collège. Elle permet de différencier de la vraie fonction cosinus (définie avec des angles en radians).

Analyse - Simplification : hypoténuse égale à 1

Pour éviter tout problème de division, on peut se concentrer sur les triangles rectangles dont l'hypothènuse a une longueur unité ($BC = 1$).
Ainsi, pour chaque angle θ compris entre 0 et 90 degrés, il existe un « unique » triangle rectangle en A, tel que $\widehat{BAC} = \theta$ et $BC = 1$. Dans ce cas $\cos^\circ \theta = AB$ et $\sin^\circ \theta = AC$.

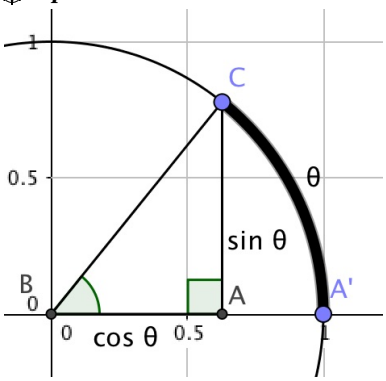
Analyse - Mesure naturelle d'angle

Les cultures (babyloniennes...) choisirent des mesures d'angles différentes. Toutes étaient proportionnelles les unes aux autres, une est-elle plus naturelle que les autres ? Oui !
Après l'étape précédente, nous considérons des triangles rectangles d'hypothènuse de longueur 1. D'une certaine façon, tous ces triangles se trouvent dans le cercle de rayon 1 et centrée en B. L'angle en B est proportionnel à la longueur de la corde CA' (voir dessin).
Ainsi l'angle habituellement (jusqu'en seconde) noté 90ř a une longueur : quart de périmetre de cercle de rayon 1 i.e. $\frac{1}{4}(2\pi \times 1) = \frac{\pi}{2}$. On a la correspondance :

$$\theta^\circ = \frac{180}{\pi}\theta^r \iff \theta^r = \frac{\pi}{180}\theta^\circ$$

Le calcul devient simple : il suffit de mesurer une longueur (avec une corde qu'on superpose, puis qu'on détend).

Représentation - Radian



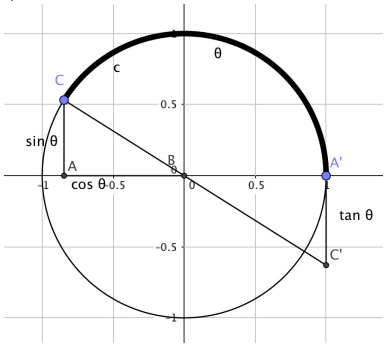
2.2. Fonctions sinus et cosinus

Représentation

D'après ce que l'on a vu, par construction, on retrouve $\cos\theta$ et $\sin\theta$ sur la figure comme indiqué en marge.
Par définition : $\tan\theta = \frac{\sin\theta}{\cos\theta} = \frac{\tan\theta}{1}$.
Donc, par théorème de Thalès, en plaçant la parallèle au sinus dans le triangle prolongé tel que $BA' = 1$, on retrouve le sinus en $A'C'$.

Reste une dernière étape : franchir les angles frontières de 0 et $\frac{\pi}{2}$ radians (ou 0° et 90°). Avec la dernière représentation, ce n'est pas compliqué : on continue la projection sur l'axe BA' et sur l'axe orthogonal.

Représentation - Cercle trigonométrique



Périodicité et symétrie

On a alors les résultats suivants, qu'il faut surtout savoir retrouver :
Exercice
Compléter les résultats suivants :

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

$\sin(\pi-\theta) =$ $\quad \cos(\pi-\theta) =$ $\quad \sin(\frac{\pi}{2}-\theta) =$ $\quad \cos(\frac{\pi}{2}-\theta) =$

Correction

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

$\sin(\pi-\theta) = \sin\theta$ $\quad \cos(\pi-\theta) = -\cos\theta$ $\quad \sin(\frac{\pi}{2}-\theta) = \cos\theta$ $\quad \cos(\frac{\pi}{2}-\theta) = \sin\theta$

Avec la définition suivante :

Définition - Congruence modulo α
Soient θ, θ' et α trois réels.
On dit que θ est congru à θ' modulo α
s'il existe $k \in \mathbb{Z}$ tel que $\theta = \theta' + k\alpha$:

$$\theta \equiv \theta'[\alpha] \iff \exists k \in \mathbb{Z} \,|\, \theta = \theta' + k\alpha$$

on a la proposition :

Proposition - Propriétés des congruences
Soit $\alpha \in \mathbb{R}$. On a pour tout $(\theta, \theta', \theta'') \in \mathbb{R}^3$:
- $\theta \equiv \theta[\alpha]$ (reflexivité)
- $\theta \equiv \theta'[\alpha] \Rightarrow \theta' \equiv \theta[\alpha]$ (symétrie)
- $(\theta \equiv \theta'[\alpha]$ et $\theta' \equiv \theta''[\alpha]) \Rightarrow \theta \equiv \theta''[\alpha]$ (transitivité)

On dit que la relation de congruence modulo α est une relation d'équivalence.

Démonstration
Reflexivité : $\theta \equiv \theta[\alpha]$, il suffit de prendre $k = 0$.
Symétrie : Si $\theta \equiv \theta'[\alpha]$, alors il existe k tel que $\theta = \theta' + k\alpha$, donc $\theta' = \theta + (-k)\alpha$ et donc $\theta' \equiv \theta[\alpha]$.
Transitivité : Si $(\theta \equiv \theta'[\alpha]$ et $\theta' \equiv \theta''[\alpha]$
alors il existe k et k' tel que $\theta = \theta' + k\alpha$, $\theta' = \theta'' + k\prime\alpha$.
donc $\theta = \theta'' + k'\alpha + k\alpha = \theta'' + (k + k')\alpha$, donc $\theta \equiv \theta''[\alpha]$ □

A savoir parfaitement retrouver :

Savoir faire - Cas d'égalité de sinus ou de cosinus
Pour tout $(\theta, \theta') \in \mathbb{R}^2$ on a :

$$\sin\theta = \sin\theta' \iff$$

$$\cos\theta = \cos\theta' \iff$$

Démonstration
En exploitant les représentations graphiques, on se place par 2π-périodicité sur un intervalle simple.
Puis, on trouve : $\sin\theta = \sin\theta' \iff \begin{cases} \theta \equiv & \theta' & [2\pi] \\ \theta \equiv & \pi - \theta' & [2\pi] \end{cases}$
De même : $\cos\theta = \cos\theta' \iff \begin{cases} \theta \equiv & \theta' & [2\pi] \\ \theta \equiv & -\theta' & [2\pi] \end{cases}$ □

2.3. Fonction tangente

Définition - Tangente d'un angle
Soit $\theta \in \mathbb{R}$, $\theta \not\equiv \frac{\pi}{2}[\pi]$. On appelle *tangente* de θ le réel, noté $\tan\theta$, défini par :

$$\tan\theta = \frac{\sin\theta}{\cos\theta}$$

Remarque - fonction cotangente
On définit de même la fonction cotangente sur $\mathbb{R} \setminus \pi\mathbb{Z}$ par $\text{cotan}\theta = \dfrac{\cos\theta}{\sin\theta}$.
Sur le cercle trigonométrique, on la trouve sur la tangente au cercle au point $(0,1)$.
Si $\theta \not\equiv 0\left[\dfrac{\pi}{2}\right]$ on a $\text{cotan}\theta = \dfrac{1}{\tan\theta}$.

Proposition - (Im)parité et périodicité
Soit $\theta \in \mathbb{R}$, $\theta \not\equiv \frac{\pi}{2}[\pi]$. On a

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

Démonstration
On applique, directement la définition, par exemple : $\tan(\theta+\pi) = \dfrac{\sin(\theta+\pi)}{\cos(\theta+\pi)} = \dfrac{-\sin\theta}{-\cos\theta} = \tan\theta$ □

Exercice
Etudier et représenter la fonction tan

Correction
La fonction tan est définie sur $\mathscr{D} = \{x \in \mathbb{R} \mid \cos\theta \neq 0\} = \mathbb{R} \setminus \{\frac{\pi}{2} + k\pi; k \in \mathbb{Z}\}$.
La fonction tan est π-périodique d'après la formule trouvée plus haut.
On peut l'étudier sur $]-\frac{\pi}{2}, \frac{\pi}{2}[$, puis exploiter des translations de vecteurs $\pi\vec{i}$.
Par division de deux fonctions dérivables, tan est dérivable sur son ensemble de définition.

$$\forall\, x \in \mathscr{D}, \qquad \tan'(x) = \frac{\cos^2(x)+\sin^2(x)}{\cos^2(x)} = 1+\tan^2(x) = \frac{1}{\cos^2(x)}$$

La fonction tan est donc strictement croissante sur les intervalles de la forme $]\dfrac{-\pi}{2}, \dfrac{-\pi}{2}[$ et à valeurs dans $]-\infty, \infty[$.
$\tan(0) = 0$ et la pente de la tangente au point 0a pour pente : $\tan'(k\pi) = \dfrac{1}{\cos(k\pi)} = 1$.
Enfin, comme $\dfrac{1}{\cos^2}$ est croissante sur $[0, \frac{\pi}{2}[$, tan est convexe sur $[0, \frac{\pi}{2}[$ et $\dfrac{1}{\cos^2}$ est décroissante sur $]-\frac{\pi}{2}, 0]$, tan est concave sur $]-\frac{\pi}{2}, 0]$.
On obtient la représentation graphique suivante :

Proposition - Cas d'égalité de tangentes
Pour tout $(\theta, \theta') \in \mathbb{R}^2$ on a :

$$\tan\theta = \tan\theta' \Leftrightarrow \theta \equiv \theta'[\pi]$$

3. Formules trigonométriques

Nous démontrons la plupart des relations avec des angles inférieurs à $\frac{\pi}{2}$, puis nous étendons les résultats par périodicité/symétrie.

3.1. Formules de Regiomontanus

Très important! A connaître par coeur, absolument! Il peut être bon d'avoir un moyen mnémotechnique auprès de soi...

Histoire - Trigonométrie : une vieille discipline
Ces formules apparaissent pour la première fois chez Ptolémée, 150 après J-C. On les retrouve chez Regiomontanus en 1464

Proposition - Formules fondamentales

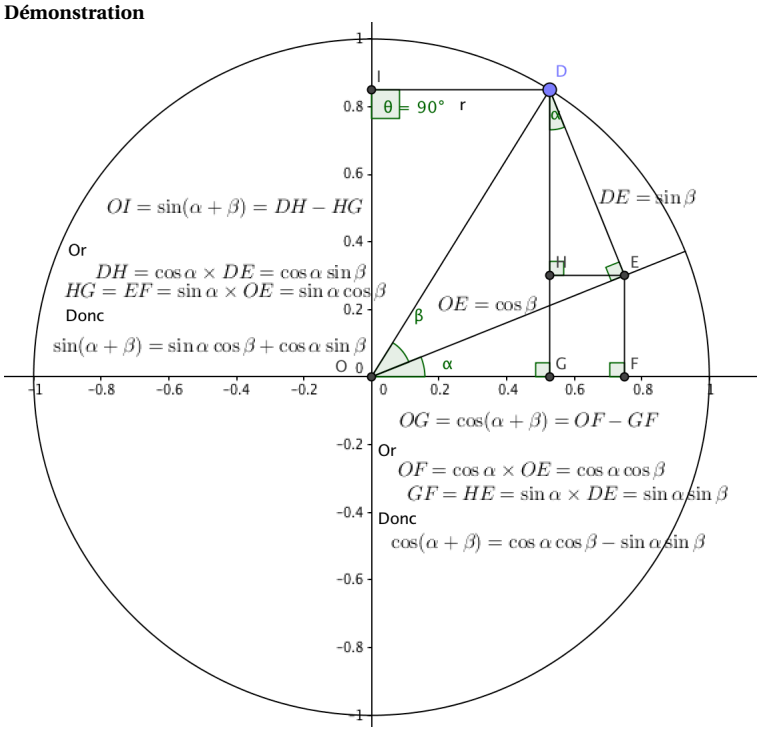
$$\cos^2\theta + \sin^2\theta = 1 \qquad 1+\tan^2\theta = \frac{1}{\cos^2\theta} \text{ où } \cos^2\theta = (\cos\theta)^2$$

$$\cos(a+b) = \cos a\cos b - \sin a\sin b \qquad \cos(a-b) = \cos a\cos b + \sin a\sin b$$
$$\sin(a+b) = \sin a\cos b + \sin b\cos a \qquad \sin(a-b) = \sin a\cos b - \sin b\cos a$$

Truc & Astuce pour le calcul - Exploiter les symétries du calcul
Une piste pour retrouver la formule $\cos(a+b)$.
Nous savons qu'il existe une relation, mais laquelle. Notons $\varphi(a,b) = \cos(a+b)$.
La relation doit vérifier :
— $\varphi(b,a) = \varphi(a,b)$, cela ne peut donc pas être $\varphi(a,b) = \sin a\cos b - \sin b\cos a$.

— $\varphi(-a,-b) = \varphi(a,b)$, cela ne peut donc pas être $\varphi(a,b) = \sin a\cos b + \sin b\cos a$.
— $\varphi(a,-a) = \cos(0) = 1$, cela ne peut donc pas être $\varphi(a,b) = \cos a\cos b + \sin a\sin b$, dans ce cas $\varphi(a,-a) = \cos^2 a - \sin^2 a \neq 1$ (pour la plupart des a)

Démonstration



La première formule dérive directement de la relation de Pythagore.
Avec $A(\cos\theta,\sin\theta)$, $M\cos\theta,0)$ et $O(0,0)$, on a OMA est rectangle.
D'après le théorème de Pythagore : $OA^2 = OM^2 + MA^2$ ie $1 = \cos^2\theta + \sin^2\theta$.
Concernant les formules $\cos(a+b)\ldots$, on pourra trouver une démonstration plus efficace avec les notations exponentielle…
Pour obtenir $\cos(a-b)$, on remplace β par $-b$ et donc $\cos\beta = \cos b$ alors que $\sin\beta = -\sin b\ldots$ □

Exercice
On peut aussi exploiter les équations différentielles.
On note $f : x \mapsto \cos(a+x)$. Montrer que f est solution du problème de Cauchy :
$$\begin{cases} y''+y &= 0 \\ y(0) &= \cos a \\ y'(0) &= -\sin(a) \end{cases}.$$
En déduire une expression de f.

Correction
Il suffit de faire le calcul. La solution du problème de Cauchy est unique, c'est une combinaison linéaire de cos et sin.
Donc il existe A et B tel que $\cos(a+x) = A\cos x + B\sin x$.
Puis avec les valeurs en $x = 0$, de f et f', on a $A = \cos a$ et $B = -\sin a$

De cet exercice, on déduit un nouveau moyen mnemotechnique pour retenir les formules de Regiomontanus.

Truc & Astuce pour le calcul - Combinaison linéaire en $\cos x$ et $\sin x$
$x \mapsto \cos(a+x)$ est une fonction, combinaison linéaire de $\cos x$ et $\sin x$.
Il existe A,B **indépendant de** x tel que $\cos(a+x) = A\cos x + B\sin x$.
En particulier pour $x = 0$ et $x = \frac{\pi}{2}$: $\cos a = A\times 1 + B\times 0$ et $\cos(a+\frac{\pi}{2}) = -\sin a = A\times 0 + B\times 1$.
Donc pour tout $a, x \in \mathbb{R}$: $\cos(a+x) = \cos a\cos x - \sin a\sin x$.
$x \mapsto \sin(a+x)$ est une fonction, combinaison linéaire de $\cos x$ et $\sin x$.
Il existe C,D **indépendant de** x tel que $\sin(a+x) = C\cos x + D\sin x$.
En particulier pour $x = 0$ et $x = \frac{\pi}{2}$: $\sin a = C\times 1 + D\times 0$ et $\sin(a+\frac{\pi}{2}) =$

$+\cos a = C\times 0 + D\times 1$.

Donc pour tout $a, x \in \mathbb{R}$: $\sin(a+x) = \sin a\cos x + \cos a \sin x$.

Savoir les déduire ou les retrouver.

Proposition - Formules fondamentales (bis)

$$\tan(a+b) = \frac{\tan a + \tan b}{1-\tan a\tan b} \qquad \tan(a-b) = \frac{\tan a - \tan b}{1+\tan a\tan b}$$

$$\sin 2a = 2\sin a\cos a \qquad \cos 2a = \cos^2 a - \sin^2 a = 2\cos^2 a - 1 = 1 - 2\sin^2 a$$

$$\tan 2a = \frac{2\tan a}{1-\tan^2 a}$$

$$\cos^2 a = \frac{1+\cos 2a}{2} \qquad \sin^2 a = \frac{1-\cos 2a}{2}$$

Exercice

Démontrer ces formules

Correction

Quelques unes :

$$\tan(a-b) = \frac{\sin(a-b)}{\cos(a-b)} = \frac{\sin a\cos b - \cos a\sin b}{\cos a\cos b + \sin a\sin b} = \frac{\frac{\sin a\cos b - \cos a\sin b}{\cos a\cos b}}{\frac{\cos a\cos b + \sin a\sin b}{\cos a\cos b}} = \frac{\tan a - \tan b}{1+\tan a\tan b}$$

$\cos(2a) = \cos(a+a) = \cos^2 a - \sin^2 a = \cos^2 a - (1-\cos^2 a) = 2\cos^2 a - 1 = (1-\sin^2 a) - \sin^2 a = 1 - 2\sin^2 a$...

3.2. Produit en somme et réciproquement

Savoir les déduire ou les retrouver.

Proposition - Transformation de produit en somme

$$\cos a\cos b = \frac{1}{2}\big(\cos(a+b) + \cos(a-b)\big)$$

$$\sin a\sin b = \frac{1}{2}\big(\cos(a-b) - \cos(a+b)\big)$$

$$\sin a\cos b = \frac{1}{2}\big(\sin(a+b) + \sin(a-b)\big)$$

Exercice

Comment exploiter les symétries du calcul pour « deviner » les égalités

Correction

On note (par exemple) $\varphi(a,b) = \sin a\sin b$. Donc $\varphi(-a,-b) = \varphi(a,b)$, la formule contient des cos (fonctions paires).

$\varphi(a,0) = 0$, donc il y a une soustraction de $\cos(a+b) = \cos a$ et $\cos(a-b) = \cos a$. $\varphi(a,a) = \sin^2 a = \frac{1}{2}(1-\cos 2a)$ ce qui donne la première formule...

Exercice

Démontrer ces formules

Correction

La démonstration se fait à l'envers. Il faut donc bien intuiter d'où partir....

Par exemple, pour obtenir $\sin a\sin b$, on se souvient que cela s'obtient avec $\cos(a+b)$.

Donc on note : $\cos(a+b) = \cos a\cos b - \sin a\sin b$ et $\cos(a-b) = \cos a\cos b + \sin a\sin b$.

Il faut supprimer les $\cos a\cos b$, on retranche donc :

$$\cos(a+b) - \cos(a-b) = -\sin a\sin b - \sin a\sin b \Longrightarrow \sin a\sin b = \frac{1}{2}(\cos(a-b) - \cos(a+b))$$

Exercice

Comment exploiter les symétries du calcul pour « deviner » les égalités

Correction

Remarque - Notation exponentielle

Avec les notations exponentielles, le résultats sera plus immédiat (on pourra le trouver dans le sens direct).

Pour aller plus loin - Trisection de l'angle

Un problème antique consistait à trouver comment couper un angle en trois parts égales.

La réponse de Viète (1593 - incomplète car elle ne donne pas de construction), consiste à remarquer que $\sin(3\alpha) = 3\sin\alpha - 4\sin^3\alpha$.

Si on connait 3α *et donc* $\sin 3\alpha = S$. *Il s'agit de résoudre :* $4x^3 - 3x + S = 0$.

Or la formule d'Al-Khwarismi donne pour solution à cette équation :

$$x = \sqrt[3]{-\frac{S}{8} + \sqrt{-\frac{S^2}{64} - \frac{1}{4}}} + \sqrt[3]{-\frac{S}{8} - \sqrt{-\frac{S^2}{64} - \frac{1}{4}}}$$

Proposition - Transformation de somme en produit

$$\cos p + \cos q = 2\cos\left(\frac{p+q}{2}\right)\cos\left(\frac{p-q}{2}\right)$$
$$\cos p - \cos q = -2\sin\left(\frac{p+q}{2}\right)\sin\left(\frac{p-q}{2}\right)$$
$$\sin p + \sin q = 2\sin\left(\frac{p+q}{2}\right)\cos\left(\frac{p-q}{2}\right)$$
$$\sin p - \sin q = 2\cos\left(\frac{p+q}{2}\right)\sin\left(\frac{p-q}{2}\right)$$

Exercice
Démontrer ces formules

Correction
Par exemple :

$$\cos p - \cos q = \cos(\frac{p+q}{2}+\frac{p-q}{2}) - \cos(\frac{p+q}{2}-\frac{p-q}{2})$$
$$= \cos(\frac{p+q}{2})\cos(\frac{p-q}{2}) - \sin(\frac{p+q}{2})\sin(\frac{p-q}{2}) - \cos(\frac{p+q}{2})\cos(\frac{p-q}{2}) - \sin(\frac{p+q}{2})\sin(\frac{p-q}{2})$$
$$= -2\sin(\frac{p+q}{2})\sin(\frac{p-q}{2})$$

Attention - Remarque
Il n'y a pas de formule générale pour transformer $\cos p \pm \sin q$.
Sauf à exploiter $\sin q = \cos(\frac{\pi}{2} - q)\ldots$

Exemple - Calcul de $\sum_{k=0}^{n} \cos kt$

Soit $t \in]0,\pi]$. On pose $S_n = \sum_{k=0}^{n} \cos kt$.

Calculer $2(\sin\frac{t}{2})S_n$ puis déduire S_n (que l'on exprimera comme produit ou quotient de trois termes sinus ou cosinus).

$$2(\sin\frac{t}{2})S_n = 2\sum_{k=0}^{n}\sin\frac{t}{2}\cos(kt) = \sum_{k=0}^{n}\left(\sin(\frac{2k+1}{2}t) + \sin(\frac{-2k+1}{2}t)\right) = \sum_{k=0}^{n}\left(\sin(\frac{2k+1}{2}t) - \sin(\frac{2k-1}{2}t)\right)$$

Puis par télescopage :

$$2(\sin\frac{t}{2})S_n = \sin(\frac{2n+1}{2}t) - \sin(-\frac{1}{2}t) = \sin(\frac{2n+1}{2}t) + \sin(\frac{1}{2}t) = 2\sin(\frac{2n+2}{4}t)\cos(\frac{2n}{4}t) = 2\sin(\frac{n+1}{2}t)\cos(\frac{n}{2}t)$$

Ainsi

$$S_n = \frac{\cos\frac{n}{2}t\sin\frac{n+1}{2}t}{\sin\frac{1}{2}t}$$

3.3. Angle moitié

Proposition - Utilisation de la tangente de l'angle moitié
On note $t = \tan\dfrac{\theta}{2}$. Alors :

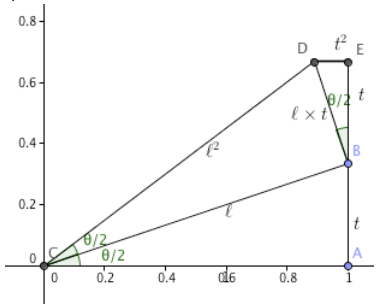
$$\sin\theta = \frac{2t}{1+t^2} \qquad \cos\theta = \frac{1-t^2}{1+t^2} \qquad \tan\theta = \frac{2t}{1-t^2}$$

Remarque - Calcul intégral
Ces résultats servent souvent dans le calcul d'intégrale avec des fonctions trigonométriques

Démonstration
Un graphique nous aide là aussi. Mais on peut directement, faire le calcul après avoir rappeler : $1 + t^2 = 1 + \tan^2\frac{\theta}{2} = \dfrac{1}{\cos^2\frac{\theta}{2}}$:

$$\sin\theta = \sin(2\frac{\theta}{2}) = 2\sin\frac{\theta}{2}\cos\frac{\theta}{2} = 2\tan\frac{\theta}{2}\cos^2\frac{\theta}{2} = \frac{2t}{1+t^2}$$

Représentation - Formules d'addition



$$\cos\theta = \cos(2\frac{\theta}{2}) = 2\cos^2\frac{\theta}{2} - 1 = \frac{2}{1+t^2} - 1 = \frac{1-t^2}{1+t^2}$$

□

Remarque - Trucs pour ne pas écrire de bêtises...

On peut vérifier que si $\theta = 0$: $t = 0$, $\sin\theta = 0$,

mais aussi $\cos\theta = 1$ et $\sin^2 + \cos^2 = 1$,

ou encore, si $\theta = \frac{\pi}{2}$, $t = 1$ et $\tan\theta = \infty$..., donc le dénominateur de tan s'annule en 1 et -1.

ou toujours, tan = $\frac{\sin}{\cos}$...

Exercice

Exemple d'emploi des notations exponentielles.

Notons α l'argument du complexe $z = 1 + it$.

Calculer z^2, quel est l'argument du complexe z^2 ? En déduire les relations recherchées ?

Sauriez-vous en déduire l'expression de $\cos\theta$ en fonction de $r = \tan\frac{\theta}{3}$?

Correction

$z^2 = (1-t^2) + 2it$, c'est le complexe de module $\sqrt{(1-t^2)^2 + 4t^2} = \sqrt{1+2t^2+t^4} = (1+t^2)$ et d'argument 2α.

Donc $\cos(2\alpha) = \frac{\mathrm{Re}(z^2)}{1+t^2} = \frac{1-t^2}{1+t^2}$ et $\sin(2\alpha) = \frac{\mathrm{Im}(z^2)}{1+t^2} = \frac{2t}{1+t^2}$.

De même $z^3 = [1-3t^2] + i[3t-t^3]$, de module $\sqrt{(1-3t^2)^2 + (3t-t^3)^2} = \sqrt{1+3t^2+3t^4+t^6} = (1+t^2)^{3/2}$ et d'argument 3α.

Donc $\cos(3\alpha) = \frac{1-3t^2}{(1+t^2)^{3/2}}$ et $\sin(3\alpha) = \frac{3t-t^3}{(1+t^2)^{3/2}}$...

Savoir faire - Méthode pour transformer $a\cos t + b\sin t$ en $A\cos(t-\phi)$

On écrit

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

Comme $\left(\frac{a}{\sqrt{a^2+b^2}}\right)^2 + \left(\frac{b}{\sqrt{a^2+b^2}}\right)^2 = 1$, il existe $\phi \in \mathbb{R}$ tel que

$$\cos\phi = \frac{a}{\sqrt{a^2+b^2}} \text{ et } \sin\phi = \frac{b}{\sqrt{a^2+b^2}}$$

d'où en posant $A = \sqrt{a^2+b^2}$, on a

$$a\cos t + b\sin t = A\left(\cos\phi\cos t + \sin\phi\sin t\right) = A\cos(t-\phi).$$

La fonction $s : t \mapsto a\cos t + b\sin t$ représente donc un signal sinusoïdal d'amplitude A de phase initiale $-\phi$(instant $t = 0$) .

Exercice

Factoriser $\sin\theta + \cos\theta$, $\sqrt{3}\cos x - \sin x$.

Correction

$\sin\theta + \cos\theta = \sqrt{2}\left(\sin\theta\sin\frac{\pi}{4} + \sin\theta\cos\frac{\pi}{4}\right) = \sqrt{2}\cos(\theta - \frac{\pi}{4})$.

Comme $\sqrt{3}^2 + 1 = 4 = 2^2$, on a,

$$\sqrt{3}\cos x - \sin x = 2\left(\sqrt{3}2\cos x - \frac{1}{2}\sin x\right) = 2\left(\cos\frac{-\pi}{6}\cos x + \sin\frac{-\pi}{6}\sin x\right) = 2\cos(x+\frac{\pi}{6})$$

4. Trigonométrie réciproque

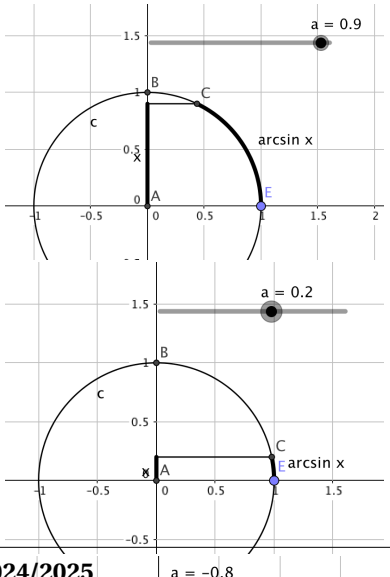
4.1. Arcsinus

Définition - Arcsinus

Pour tout $x \in [-1,1]$ il existe un unique $\theta \in [-\frac{\pi}{2}, \frac{\pi}{2}]$ vérifiant $x = \sin\theta$. Ce réel θ est appelé arcsinus de x et noté $\arcsin x$ On a donc :

$$\theta = \arcsin x \Leftrightarrow \left(\sin\theta = x \text{ et } \theta \in [-\frac{\pi}{2}, \frac{\pi}{2}]\right)$$

Représentation - Quelques valeurs de $\arcsin x$



Attention - Intervalle d'arrivée
De même qu'il a été choisi de prendre l'unique racine positive de a, lorsqu'on écrit $\sqrt{a}$ (et non $-\sqrt{a}$ qui vérifie également $(-\sqrt{a})^2 = a$) ; on choisit ici un résultat dans $[-\frac{\pi}{2}, \frac{\pi}{2}]$, il faut donc penser à ajouter un angle...

$$\sin\theta = x \Longleftrightarrow \theta \equiv \arcsin(x)[2\pi] \text{ ou } \theta \equiv \pi - \arcsin(x)[2\pi]$$

Il faut connaître les valeurs remarquables suivantes :

x	0	$\frac{1}{2}$	$\frac{1}{\sqrt{2}}$	$\frac{\sqrt{3}}{2}$	1
$\arcsin x$					

Exercice
Calculer $\arcsin(\sin\frac{2\pi}{3}), \arcsin(\sin\frac{23\pi}{6})$.

Correction
Où est le piège ? dans l'intervalle d'arrivée.
Comme $\frac{2\pi}{3} > \frac{\pi}{2}$, $\arcsin(\sin\frac{2\pi}{3}) = \pi - \frac{2\pi}{3} = \frac{\pi}{3}$. De même $\frac{23\pi}{6} = 4\pi - \frac{\pi}{6}$, $\arcsin(\sin\frac{23\pi}{6}) = -\frac{\pi}{6}$

Histoire - Série arcsin
On trouve chez WALLIS *et les mathématiciens anglais pré-newtonien :*
$\arcsin(x) = x + \frac{1}{2}\frac{x^3}{3} + \frac{1\times3}{2\times4}\frac{x^5}{5} + \frac{1\times3\times5}{2\times4\times6}\frac{x^7}{7} + \ldots$

Proposition - Composition de fonctions trigonométriques et arcsin
On a :

$$\forall\theta \in [-\frac{\pi}{2}, \frac{\pi}{2}], \quad \arcsin(\sin\theta) = \theta$$

$$\forall x \in [-1,1], \quad \sin(\arcsin x) = x$$

$$\forall x \in [-1,1], \quad \cos(\arcsin x) = \sqrt{1-x^2}$$

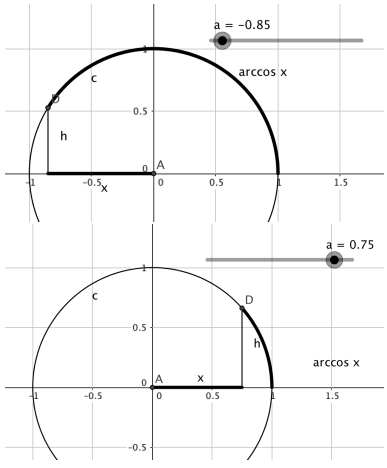
Démonstration
Les deux premiers résultats s'obtiennent en appliquant tout simplement la définition.
Comme $\cos^2 = 1 - \sin^2$, on a donc pour $x \in [-1,1]$,

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

donc $\cos(\arcsin(x)) = \pm\sqrt{1-x^2}$.
Or $\arcsin(x) \in [-\frac{\pi}{2}, \frac{\pi}{2}]$, donc $\cos(\arcsin(x)) \geqslant 0$ et donc $\cos(\arcsin(x)) = +\sqrt{1-x^2}$ □

4.2. Arccosinus

Représentation - Quelques valeurs de arccos x



Définition - Arccosinus
Pour tout $x \in [-1,1]$ il existe un unique $\theta \in [0,\pi]$ vérifiant $x = \cos\theta$. Ce réel θ est appelé arccosinus de x et noté $\arccos x$ On a donc :

$$\theta = \arccos x \Leftrightarrow \Big(\cos\theta = x \text{ et } \theta \in [0,\pi]\Big)$$

Attention - Intervalle d'arrivée
Comme précédemment :

$$\cos\theta = x \Longleftrightarrow \theta \equiv \arccos(x)[2\pi] \text{ ou } \theta \equiv -\arccos(x)[2\pi]$$

Il faut connaître les valeurs remarquables suivantes :

x	0	$\frac{1}{2}$	$\frac{1}{\sqrt{2}}$	$\frac{\sqrt{3}}{2}$	1
$\arccos x$					

Exercice
Calculer $\arccos(\cos\frac{4\pi}{3}), \arccos(\cos\frac{25\pi}{6})$.

Correction
Comme $\frac{4\pi}{3} \in]\pi, 2\pi[$, $\arccos(\cos\frac{2\pi}{3}) = -\frac{4\pi}{3} + 2\pi = \frac{2\pi}{3}$. De même $\frac{25\pi}{6} = 4\pi + \frac{\pi}{6}$, $\arccos(\cos\frac{25\pi}{6}) = \frac{\pi}{6}$

Proposition - Composition de fonctions trigonométriques et arccos
On a :

$$\forall \theta \in [0,\pi], \quad \arccos(\cos\theta) = \theta$$

$$\forall x \in [-1,1], \quad \cos(\arccos x) = x$$

$$\forall x \in [-1,1], \quad \sin(\arccos x) = \sqrt{1-x^2}$$

Exercice
Faire la démonstration

Correction
Comme $\sin^2 = 1 - \cos^2$, on a donc pour $x \in [-1,1]$,

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

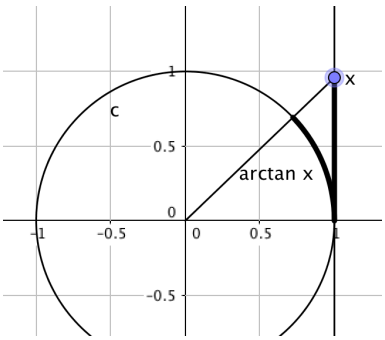
donc $\sin(\arccos(x)) = \pm\sqrt{1-x^2}$.
Or $\arccos(x) \in [0,\pi]$, donc $\sin(\arccos(x)) \geqslant 0$ et donc $\sin(\arccos(x)) = +\sqrt{1-x^2}$

4.3. Arctangente

Définition - Arctangente
Pour tout $x \in \mathbb{R}$ il existe un unique $\theta \in]-\frac{\pi}{2}, \frac{\pi}{2}[$ vérifiant $x = \tan\theta$. Ce réel θ est appelé arctangente de x et noté $\arctan x$ On a donc :

$$\theta = \arctan x \Leftrightarrow \left(\tan\theta = x \text{ et } \theta \in]-\frac{\pi}{2}, \frac{\pi}{2}[\right)$$

Représentation - Quelques valeurs de $\arctan x$



Il faut connaître les valeurs remarquables suivantes :

x	0	$\frac{1}{\sqrt{3}}$	1	$\sqrt{3}$
$\arctan x$				

Proposition - Composition de fonctions trigonométriques et arctan
On a :

$$\forall \theta \in]-\frac{\pi}{2}, \frac{\pi}{2}[, \quad \arctan(\tan\theta) = \theta$$

$$\forall x \in \mathbb{R}, \quad \tan(\arctan x) = x$$

$$\forall x \in \mathbb{R}, \quad \cos(\arctan x) = \frac{1}{\sqrt{1+x^2}}$$

$$\forall x \in \mathbb{R}, \quad \sin(\arctan x) = \frac{x}{\sqrt{1+x^2}}$$

Histoire - Série arctan
On trouve chez GREGORY *et les mathématiciens anglais pré-newtonien :*
$\arcsin(x) = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \ldots$.
Puis, avec, une formule d'approximation de π *:*
$\pi = 16\arctan\frac{1}{5} - 4\arctan\frac{1}{239}$.

Démonstration
On exploite $\cos^2 = \frac{1}{1+\tan^2}$, puis la positivité.
Et $\sin^2 = 1 - \cos^2 = 1 - \frac{1}{1+\tan^2} = \frac{\tan^2}{1+\tan^2} \ldots$ □

Proposition - Relation complexe
Soit $x \in \mathbb{R}$, alors $\arg(1+ix) = \arctan(x)$

Démonstration
Soit $z = 1+ix$, note $\rho = \sqrt{1+x^2}$ son module et θ sont argument.

$$z = 1+ix = \rho(\cos\theta + i\sin\theta)$$

Donc a donc $\tan\theta = \dfrac{\sin\theta}{\cos\theta} = \dfrac{\rho\sin\theta}{\rho\cos\theta} = \dfrac{x}{1} = x$.
Ainsi $\arctan x = \theta = \arg(1+ix)$. □

Remarque - Aspect analytique
On étudiera dans un prochain chapitre les aspects analytiques de ces fonctions (dérivées, développement limités...)

Pour aller plus loin - Formule d'approximation de π : Formule de Machin (1706)
$4\arctan\frac{1}{5} - \arctan\frac{1}{239} = \cdots = \arg\left((1+\frac{i}{5})^4(1-\frac{i}{239})\right) = \arg(\frac{114244}{149375}(1+i)) = \frac{\pi}{4}$.
Comme $\arctan(x) \approx x + \frac{1}{6}x^3 + \frac{3}{8}x^5 + \ldots$, *on trouve une excellente approximation de* π.
Cette formule donna la meilleure approximation de π *connue durant tout le XVIII ième siècle.*

5. Bilan

Synthèse

⇝ En géométrie (et physique), nous pratiquons la projection orthogonale, cela consiste à multiplier par $\cos\theta$ (ou $\sin\theta$) la longueur de l'hypoténuse. Différents calculs se présentent à nous : $\cos(a+b)$, $\cos(a)\cos(b)$ ou $\cos a + \cos b$ (et tout ce que l'on peut imaginer de manière équivalente avec sin ou tan). Il existe alors de nombreuses relations calculatoires **à apprendre**!

⇝ Très souvent la question se pose de manière réciproque : étant donné une longueur de quel angle en est-elle le cos? Le problème, la fonction n'est pas injective : on peut avoir $\theta \neq \theta'$ et $\cos\theta = \cos\theta'$. On restreint donc l'intervalle image. On crée ainsi une fonction réciproque arccos à la fonction $\cos_{|[0,\pi]} : [0,\pi] \to [-1,1]$. De même pour les fonctions $\sin_{|[-\frac{\pi}{2},\frac{\pi}{2}]}$ et $\tan_{[-\frac{\pi}{2},\frac{\pi}{2}]}$. Au passage, on trouve une méthode complémentaire (algébrique) dans le simple cas de la racine carrée d'un nombre complexe.

Savoir-faire et Truc & Astuce du chapitre

— Savoir-faire - Cas d'égalité de sinus ou de cosinus
— Truc & Astuce pour le calcul - Exploiter les symétries du calcul
— Truc & Astuce pour le calcul - Combinaison linéaire en $\cos x$ et $\sin x$
— Savoir-faire - Méthode pour transformer $a\cos t + b\sin t$ en $A\cos(t-\varphi)$

Notations

Notations	*Définitions*	*Propriétés*	*Remarques*
cos, sin, tan	Fonctions cosinus, sinus et tangentes.	Dans un triangle ABC rectangle en A, $\cos B = \dfrac{AB}{BC}$, $\sin B = \dfrac{AC}{BC}$ et $\tan B = \dfrac{AC}{AB} = \dfrac{\cos B}{\sin B}$	Tout un chapitre à connaître!
arccos, arcsin, arctan	Fonctions réciproques de $\cos_{|[0,\pi]}^{|[-1,1]}$, $\sin_{|[-\pi/2,\pi/2]}^{|[-1,1]}$, $\tan_{|[-\pi/2,\pi/2]}^{|[-1,1]}$ respectivement.	$\cos(\arccos(x)) = x$ pour tout $x \in [-1,1]$	A savoir maîtriser

Retour sur les problèmes

16. Voir le cours
17. Les formules d'additions de cos et sin restent vraies si les angles sont en degré.
Pourquoi changer d'unité au lycée?

L'inégalité $\sin x \leqslant x \leqslant \tan x$ est vraie pour x en radian. Pour x en degré, on aurait plutôt : $\sin x \leqslant \frac{\pi}{180} x \leqslant \tan x$.

On trouve alors, en radian : $\cos x \leqslant \dfrac{\sin x}{x} \leqslant 1$ en faisant $x \to 0$: on trouve $\sin'(x) = 1$, puis avec les formules d'addition : $\sin' = -\cos$.

Si les angles sont en degré : il faut un coefficient multiplicatif. C'est donc une relation pénible.

Bilan : si on passe en radian, c'est parce qu'on s'intéresse à propriétés analytiques des fonctions trigonométriques...

18. A apprendre. Mais l'apprendre, c'est toujours plus compliqué.

19. $\cos(n\theta) = \mathrm{Re}((e^{i\theta})^n) = \mathrm{Re}((\cos\theta + i\sin\theta)^n) = \displaystyle\sum_{0\leqslant k\leqslant \frac{n}{2}} \binom{n}{2k} (-1)^k \cos^{n-2k}\theta (1-\cos^2\theta)^k = T_n(\cos\theta)$

20. arc sin et arc cos...

Chapitre 5

Ensemble des nombres complexes

Résumé -

Dans ce chapitre, nous reprenons des résultats de lycée sur les nombres complexes et leur lien avec la géométrie du plan. Ce sont des bons outils pour reprendre les propriétés trigonométriques !
Comme, ces nombres ont été inventés/découverts pour résoudre tout type d'équation polynomiale, il est normal que ce chapitre soit associé à la recherche des solutions de $x^n = 1$, i.e. la recherche des racines de l'unité.
Puis nous nous intéressons aux transformations géométriques du plan. Lorsque l'espace géométrique étudié est de dimension 2 (plan $\mathbb{R}^2$), les nombres complexes sont de parfaits outils pour faire cette étude. En effet, ces nombres ont un lien fort avec la géométrie (revue au chapitre suivant) : addition de complexes = translation et multiplication de complexes = homothétie et rotation (similitude)...

- *Micmath - La conjugaison complexe est un automorphisme de corps. https ://www.youtube.com/watch ?v=AVDMpnwsztg*
- *Les maths en finesse - Racines nieme de l'unité. https ://www.youtube.com/watch ?v=aZLGdnktO8k*
- *Exo7Math - Nombres complexes part.4 : géométrie .https ://www.youtube.com/watch ?v=ej9zpQYsQs8*
- *AEV - Nombres complexes / Applications à la géométrie. https ://www.youtube.com/watch ?v=HfxqAQ1SiGo*

Sommaire

1. Problèmes

? Problème 21 - Multiplication de nombres

Que représente la multiplication complexe $Z = z \times z'$ pour des points $M(z)$ et $M'(z')$.
En particulier, existe-t-il un algorithme géométrique pour tracer cette multiplication ?
On pourra dans un premier temps, considérer que $z \in \mathbb{R}$, $z \in \mathbb{U}$

? Problème 22 - Théorème de Napoléon

En exploitant les nombres complexes, démontrer le théorème de Napoléon :
Si nous construisons trois triangles équilatéraux à partir des côtés d'un triangle quelconque, tous à l'extérieur ou tous à l'intérieur, les centres de ces triangles équilatéraux forment eux-mêmes un triangle équilatéral.
Beaucoup de problèmes (théorème de Ptolémée, théorème de Cotes...), du plan se démontre par *calculs* avec des nombres complexes.

? Problème 23 - Transformation du plan

A l'aide des nombres complexes, comment trouver toutes les transformations du plan qui conserve les longueurs et/ou les angles ?

? Problème 24 - Application en physique

On connait beaucoup d'applications en physique des nombres complexes (notés j), en particulier en électricité ou pour l'étude de Fourier.
La loi de Snell-Descartes donne pour la réfraction entre deux milieux d'indices n_1 et n_2 : $n_1 \sin\theta_1 = n_2 \sin\theta_2$.
Lorsque $n_1 > n_2$ et θ_1 proche de 0 (donc $\sin\theta_1$ proche de 1), on trouve $\sin\theta_2 = \dfrac{n_1}{n_2}\sin\theta_1 > 1$, et donc $\cos\theta_2 \in \mathbb{C}$.
Avec ce nombre complexe, Augustin FRESNEL unifiait en une seule formule ce qui se présentait jusqu'alors sous deux formes. Quelle est cette formule ?

2. EULER : manipulateur des nombres du diable

2.1. Racine de polynômes

Histoire - Citation
« Au reste tant les vrayes racines que les fausses ne sont pas tousiours reelles ; mais quelques seulement imaginaires ; c'est à dire qu'on peut bien touiours en imaginer autant que iay dit en chasque Equation ; mais qu'il n'y a quelquefois aucune quantité, qui corresponde a celles qu'on imagine. » DESCARTES *(1637)*

Analyse - Problème de Cardan (1545)
Un segment de longueur 10 est coupé en deux parties de longueur a et b. De quelle manière le faire, de sorte que l'aire du rectangle dont chaque côté vaut a et b puisse valoir 40 ?
On a les équations $a+b=10$ et $a\times b=40$, donc $a^2-10a=a(a-10)=a\times b=40$.
Les racines de cette équation $x^2-10x-40=0$ est $5+\sqrt{-15}$ et $5-\sqrt{-15}$.
Evidemment cela est problématique (racine d'un nombre négatif) sauf que
— $(5+\sqrt{-15})+(5-\sqrt{-15})=10$

— $(5+\sqrt{-15})\times(5-\sqrt{-15}) = 5^2-(-15) = 25+15 = 40$

La solution fonctionne avec ces nouveaux nombres, à condition de garder les même règles de calcul qu'avec les nombres classiques…

Les règles de calcul sont donnés par Raphaël Bombelli (1572), dans son algebra. Pendant deux siècles les mathématiciens se querellent quant à leur existence et à leurs emplois.

Exercice

On reprend un exercice historique de Bombelli.

En reprenant les règles classiques de calcul, évaluer $(2+\sqrt{-1})^3$.

En employant les formules de Cardan, trouver les racines de $x^3 = 15x+4$.

Correction

Binôme de Newton : $(2+\sqrt{-1})^3 = 2^3+3\sqrt{-1}\times 2^2+3(\sqrt{-1})^2\times 2+(\sqrt{-1})^3 = 8+12\sqrt{-1}-6-\sqrt{-1} = 2+11\sqrt{-1}$.

2.2. Calcul algébrique

Euler invente la notation i bien pratique et les manipule avec précision. Il écrit à Diderot : « $e^{i\pi} = -1$ donc Dieu existe ».

Remarque - Unicité

Un complexe est un « nombre » z qui s'écrit $z = a+ib$ où a et b sont des réels et i vérifie $i^2 = -1$. Cette écriture est unique et s'appelle la forme algébrique de z.

Pour aller plus loin - Construction de $\mathbb{C}$

Rassurons nous : nous construirons bien $\mathbb{C}$ par la suite, selon la méthode proposée par Cauchy (avec des classes d'équivalence).

Comme la construction de $\mathbb{C}$ n'est pas donnée, il n'est pas possible de démontrer l'unicité de l'écriture de z, ni la justification des règles de calcul. Nous les admettons alors. Ce n'est pas rien…

Définition - Notation de nombre complexe (1748)

Soit $z = a+ib$ un complexe (a et b sont des réels).

$a = \mathrm{Re}\, z$ s'appelle la **partie réelle** de z.

$b = \mathrm{Im}\, z$ s'appelle la **partie imaginaire** de z;

z est dit **imaginaire pur** *($z \in i\mathbb{R}$) si sa partie réelle est nulle.*

$\overline{z} = a-ib$ s'appelle le **conjugué** de $z = a+ib$.

$|z| = \sqrt{a^2+b^2}$ s'appelle le **module** de z.

Notons tout de suite comment se comportent les calculs habituelles addition et multiplication sur $\mathbb{C}$ (il s'agit presque de définition ici…)

Définition - (et proposition ?) $\mathbb{C}$ est un corps

Pour tout $(z,z') \in \mathbb{C}^2$, $\lambda, \lambda' \in \mathbb{R}$,

— $\mathrm{Re}(\lambda z+\lambda' z') = \lambda \mathrm{Re}(z)+\lambda' \mathrm{Re}(z')$ (*la partie réelle est $\mathbb{R}$-linéaire sur $\mathbb{C}$*)

— $\mathrm{Im}(\lambda z+\lambda' z') = \lambda \mathrm{Im}(z)+\lambda' \mathrm{Im}(z')$ (*la partie imaginaire est $\mathbb{R}$-linéaire sur $\mathbb{C}$*)

— si $z = a+ib$ et $z' = a'+ib'$, alors $z\times z' = (aa'-bb')+i(ab'+a'b)$

En particulier $z\times\overline{z} = a^2+b^2 = |z|^2 = |\overline{z}|^2$, donc $\dfrac{1}{z} = \dfrac{\overline{z}}{|z|^2}$.

Remarque - Importance du conjugué

Le fait que $z\overline{z}$ est un nombre réel (pur) explique l'importance de la notion de conjugué.

Démonstration

Comme

$$\lambda z+\lambda' z' = (\lambda a+i\lambda b)+(\lambda' a'+i\lambda' b') = (\lambda a+\lambda' a')+i(\lambda b+\lambda' b')$$

Donc

$$\mathrm{Re}(\lambda z+\lambda' z') = \lambda \mathrm{Re}(z)+\lambda' \mathrm{Re}(z') \quad \text{et} \quad \mathrm{Im}(\lambda z+\lambda' z') = \lambda \mathrm{Im}(z)+\lambda' \mathrm{Im}(z')$$

Puis :

$$z\times z' = (a+ib)\times(a'+ib') = (aa'+iab'+iba'+i^2bb' = (aa'+bb')+i(ab'+a'b)$$

□

On a alors

Proposition - Conjugaison
On a les propriétés du conjugué :

$$\forall (z,z')\in\mathbb{C}^2,\forall a\in\mathbb{R},\quad \overline{\overline{z}}=z \qquad \overline{z+z'}=\overline{z}+\overline{z'}$$
$$\overline{zz'}=\overline{z}\,\overline{z'} \qquad \overline{\left(\frac{1}{z}\right)}=\frac{1}{\overline{z}}$$
$$\mathrm{Re}\,z=\frac{z+\overline{z}}{2} \qquad \mathrm{Im}\,z=\frac{z-\overline{z}}{2i}$$

Démonstration
Supposons que $z=a+ib$ et $z'=a'+ib'$. Alors

$$\overline{\overline{z}}=\overline{a-ib}=a+ib=z$$
$$\overline{z+z'}=\overline{(a+ib)+(a'+ib')}=\overline{(a+a')+i(b+b')}=(a+a')-i(b+b')=(a-ib)+(a'-ib')=\overline{z}+\overline{z'}$$
$$z\times z'=\overline{(a+ib)\times(a'+ib')}=\overline{(aa'-bb')+i(ab'+a'b)}=(aa'-bb')-i(ab'+a'b)=(a-ib)\times(a'-ib')=\overline{z}\times\overline{z'}$$
$$\overline{\frac{1}{z}}=\frac{1}{|z|^2}\overline{z}=\frac{1}{|\overline{z}|^2}\overline{z}=\frac{1}{\overline{z}}\left(=\frac{z}{|z|^2}\right)$$
$$\frac{z+\overline{z}}{2}=\frac{2a+i0}{2}=a=\mathrm{Re}(z) \qquad \frac{z-\overline{z}}{2i}=\frac{0+2ib}{2i}=b=\mathrm{Im}(z)$$

□

Proposition - Puissance et conjugaison
On définit les puissances d'un nombre complexe par

$$\begin{cases} z^0=1 \\ \forall n\in\mathbb{N},\ z^{n+1}=z^n z \end{cases}$$

On a alors $\forall n\in\mathbb{N},\ \overline{z^n}=\overline{z}^n$.
Pour $z\neq 0$ et $n\in\mathbb{N}$, on pose $z^{-n}=\dfrac{1}{z^n}=(z^n)^{-1}$, on a alors $\forall n\in\mathbb{Z},\ \overline{z^n}=\overline{z}^n$.

Exercice
Faire la démonstration

Correction
Il faut faire une récurrence

Exercice
Démontrer la formule de Moivre (1707) : $(\cos a+i\sin a)^n=\cos(na)+i\sin(na)$.
(Il est compliqué de comprendre comment il a cette idée, au. hasard ?)

Correction
On le fait par récurrence.
Pour $n=0$, on a bien $(\cos a+i\sin a)^0=1=\cos 0+i\sin 0$.
Supposons que le résultat est vrai au rang n quelconque.
$(\cos a+i\sin a)^{n+1}=(\cos(na)+i\sin(na))(\cos a+i\sin a)=(\cos(na)\cos a-\sin(na)\sin a)+i(\cos(na)\sin a+\sin(na)\cos a))=\cos((n+1)a)+i\sin((n+1)a)$

Proposition - Propriétés du module
On a les propriétés du module :

$$\forall (z,z')\in\mathbb{C}^2,a\in\mathbb{R}_+,\quad |z|=\sqrt{z\overline{z}} \qquad |az|=a|z| \qquad |zz'|=|z||z'|$$
$$|z|=|\overline{z}|=|-z| \qquad \left|\frac{z}{z'}\right|=\frac{|z|}{|z'|}\ (\text{si } z'\neq 0)$$

Démonstration
Les deux premiers résultats ont été démontrés. Puis

$$|zz'|=\sqrt{(zz')\overline{(zz')}}=\sqrt{zz'\overline{z}\,\overline{z'}}=\sqrt{z\overline{z}}\sqrt{z'\overline{z'}}=|z||z'|$$
$$|z|=\left|z'\times\frac{z}{z'}\right|=|z'|\times\left|\frac{z}{z'}\right|$$

Il ne reste plus qu'à diviser par $|z'|$. □

Remarque - Valeur absolue et module
Pour un réel, valeur absolue et module coïncident !

2.3. Représentation graphique (addition et longueur)

On munit le plan d'un repère orthonormé direct $(0,\vec{u},\vec{v})$. Le point M de coordonnées (a,b), caractérisé par $\overrightarrow{OM} = a\vec{u} + b\vec{v}$, peut alors être représenté par le complexe $z = a + ib$.

Définition - Affixe d'un point. Affixe d'un vecteur
$z = a+ib$ est alors appelé **affixe** du point $M(a,b)$, on peut noter $z = \text{Aff}(M)$.
Réciproquement, le point M est appelé (point) image de z.
De même, si $\vec{w}$ est un vecteur de coordonnées (a,b), $a+ib$ est appelé affixe de $\vec{w}$ (noté $\text{Aff}(\vec{w})$), lui-même appelé (vecteur) image du complexe $a+ib$.

Remarque - Axes
Les points de l'axe des abscisses correspondent aux points d'affixe réelle.
Les points de l'axe des ordonnés correspondent aux points d'affixe imaginaire pure.

Proposition - Opération complexe et correspondance sur le plan géométrique
Si z est l'affixe de M alors $\overline{z}$ est l'affixe du symétrique de M par rapport à l'axe des abscisses.
Si $z = \text{Aff}(M)$ alors $|z|$ est égal à la distance OM.
Si $z = \text{Aff}(M)$ et $z_0 = \text{Aff}(M_0)$, alors $\text{Aff}(\overrightarrow{M_0M}) = z - z_0$ et $|z - z_0| = M_0M$

Histoire - John WALLIS (1616-1703)

Il semble que John Wallis ait imaginé représenter les nombres complexes dans un plan. Certainement, avait-il compris comment additionner les nombres complexes ; mais l'essentiel : il n'avait pas compris le rôle géométrique de la multiplication.
Wallis est par ailleurs un grand manipulateur de l'infini.

2.4. Inégalités

Théorème - Inégalités
Pour $(z,z') \in \mathbb{C}^2$, on a les inégalités suivantes :

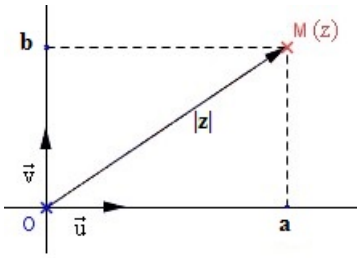
$$\text{Re}\, z \leqslant |z| \text{ avec égalité si et seulement si } z \in \mathbb{R}^+$$

$$\text{Im}\, z \leqslant |z| \text{ avec égalité si et seulement si } z \in i\mathbb{R}^+$$

$$\Big||z| - |z'|\Big| \leqslant |z+z'| \leqslant |z| + |z'| \textbf{ (Inégalité triangulaire)}$$

avec égalité dans l'inégalité de droite
si et seulement si $z' = 0$ ou il existe $\lambda \in \mathbb{R}^+$ tel que $z = \lambda z'$ (z, z' positivement liés).

Représentation - Affixe d'un point (ou d'un vecteur)



Attention - Module ou valeur absolue ?
Il y a des modules et des valeurs absolues partout ici

Analyse - Interprétation de l'inégalité triangulaire
Si A, B, C sont trois points tels que $z = \text{Aff}(\overrightarrow{AB})$ et $z' = \text{Aff}(\overrightarrow{BC})$, alors

$$|z+z'| \leqslant |z| + |z'| \text{ avec égalité si et seulement si } z, z' \text{ positivement liés}$$

signifie que $AC \leqslant AB + BC$ avec égalité si et seulement si A, B, C sont alignés et B entre A et C.

Démonstration
Avec les notations habituelles,

$$|z|^2 = a^2 + b^2 \geqslant a^2 \Longrightarrow |z| \geqslant |a| = |\text{Re}(z)|$$

Il y a égalité si et seulement si $b = 0$, i.e. $\text{Im}(z) = 0$.
Pour l'inégalité triangulaire, réécrivons les calculs (au carrés) pour mieux voir comment comparer :

$$|z+z'|^2 = (a+a')^2 + (b+b')^2 = a^2 + b^2 + a'^2 + b'^2 + 2aa' + 2bb' = |z|^2 + |z'|^2 + 2\text{Re}(z\overline{z'})$$

$$(|z|+|z'|)^2 = |z|^2+|z'|^2+2|zz'| = |z|^2+|z'|^2+2|z\overline{z'}|$$

Or comme on a vu que $|z\overline{z'}| \geqslant \mathrm{Re}(z\overline{z'})$, on a l'égalité attendue ;
la condition équivalente à l'égalité est $\mathrm{Im}(z\overline{z'}) = 0$ et $\mathrm{Re}(z\overline{z'}) > 0$
i.e. $ab'-a'b = 0$ et $aa'+bb' > 0$ i.e. $z' = \lambda z$ avec $\lambda = \frac{a'}{a} = \frac{b'}{b} > 0$
Enfin, en supposant $|z| \geqslant |z'|$, en considérant $Z = z+z'$ et $Z' = -z'$, on a

$$|z| = |Z+Z'| \leqslant |Z|+|Z'| = |z+z'|+|z'| \quad \Longrightarrow \quad \Big||z|-|z'|\Big| = |z|-|z'| \leqslant |z+z'|$$

□

Par récurrence :

Proposition - Inégalités
Pour n complexes $z_1, \dots, z_n$ on a

$$|z_1+\cdots+z_n| \leqslant |z_1|+\cdots+|z_n|.$$

Proposition - Caractérisation des complexes remarquables

$$z = 0 \Leftrightarrow |z| = 0 \Leftrightarrow \mathrm{Re}\, z = \mathrm{Im}\, z = 0$$

$$z \in \mathbb{R} \Leftrightarrow \mathrm{Im}\, z = 0 \Leftrightarrow \overline{z} = z \Leftrightarrow |z|^2 = (\mathrm{Re}\, z)^2$$

$$z \in i\mathbb{R} \Leftrightarrow \mathrm{Re}\, z = 0 \Leftrightarrow \overline{z} = -z \Leftrightarrow |z|^2 = (\mathrm{Im}\, z)^2$$

3. Le visionnaire : GAUSS et la multiplication complexe

3.1. Les complexes de module 1

En 1800, les mathématiciens manipulent les nombres complexes, mais ces nombres manquent de légitimité.
C'est Gauss qui les justifie géométriquement sur $\mathbb{R}^2$ (Argand et Wessel semblent, chacun de leur côté, avoir eu la même idée).

Le groupe unitaire $\mathbb{U}$

Définition - Groupe unitaire
On note $\mathbb{U}$ l'ensemble des complexes de module 1, c'est aussi le cercle unité de $\mathbb{C}$, ensemble des affixes des points du cercle trigonométrique

$$\mathbb{U} = \{z \in \mathbb{C} \mid |z| = 1\}.$$

Proposition - Conjugaison sur $\mathbb{U}$
$\forall (z,z') \in \mathbb{U}^2,\ zz' \in \mathbb{U}, \qquad \forall z \in \mathbb{U},\ \overline{z} = \dfrac{1}{z} \in \mathbb{U}.$
On dit que l'ensemble $\mathbb{U}$ muni de l'opération multiplication est un groupe commutatif.

Démonstration
$|zz'| = |z| \times |z'| = 1 \times 1 = 1$, donc $zz' \in \mathbb{U}$.
On a vu $\dfrac{1}{z} = \dfrac{\overline{z}}{|z|^2} = \overline{z}$ car $|z| = 1$. □

Interprétation géométrique du calcul $u \times z$ pour $u \in \mathbb{U}$ et $z \in \mathbb{C}$

Analyse - Géométrique

Soit $u \in \mathbb{U} \setminus \{1\}$ et $z \in \mathbb{C}$. Considérons les 4 points du plan $A(1)$, $B(u) \neq A$, $C(z)$ et $D(u \times z)$.

Alors $AC = |z-1|$ et $BD = |uz-u| = |u| \times |z-1| = AC$.

On a vu (raisonnement analyse-synthèse) alors que :

si (AB) et (CD) ne sont pas parallèles, il existe une unique rotation r du plan tel que $r(A) = B$ et $r(C) = D$.

C'est la rotation dont le centre Ω est à l'intersection des médiatrices de $[AB]$ et $[CD]$ et d'angle $\widehat{\overrightarrow{\Omega A}, \overrightarrow{\Omega B}}$. Or $OA = OB = 1$ et $OC = |z| = |u||z| = |uz| = OD$. Donc O est à l'intersection des médiatrices, par unicité de ce point : $O = \Omega$.

Ainsi, $r(C) = D$, où r est la rotation de centre $O(0)$ et d'angle $\widehat{\overrightarrow{OA}, \overrightarrow{OB}}$.

On montrera plus loin que $(AB)//(CD) \iff \dfrac{z_D - z_C}{z_B - z_A} \in \mathbb{R} \iff z \in \mathbb{R}$.

Or dans ce cas, on peut appliquer directement le théorème de Thalès : pour les triangles OAB et OCD semblables.

On retrouve donc le résultat précédent.

Définition - Argument de $u \in \mathbb{U}$, de $z \in \mathbb{C}$

Soit $u \in \mathbb{U}$. On note I, le point du plan d'affixe 1 et M celui d'affixe u.

On appelle argument de $u \in \mathbb{U}$ noté $\text{Arg}(u)$, l'angle (principal) $(\overrightarrow{OI}, \overrightarrow{OM}) \in]-\pi, \pi]$.

Dans un premier temps, on note $\measuredangle\theta$ ce nombre complexe de module 1 et d'argument θ.

On a alors $u = \cos\theta + i\sin\theta$, pour $\theta = \text{Arg}(u)$.

Remarque - Angle aigu

Le résultat se conçoit bien pour $\theta \in [0, \frac{\pi}{2}]$, mais il reste vrai pour toute mesure d'angle (dans $\mathbb{R}$), par propriété de parité de parité (cos), imparité (sin) et 2π-périodicité.

Démonstration

Si on note X, le point d'affixe $\text{Re}(u)$, alors le triangle OXM est rectangle en X.

Son hypothénuse a une longueur égale à $OM = |u| = 1$, donc comme $\theta \equiv (\overrightarrow{OI}, \overrightarrow{OM})[2\pi]$,

on a $OX = \cos\theta = \text{Re}(u)$ et $XM = \sin\theta = \text{Im}(u)$.

Ainsi $u = \cos\theta + i\sin\theta$. □

Pour aller plus loin - Le problème est en fait lié à la relation d'égalité...

Nous allons essayer de faire avec, au mieux. Mais lorsqu'on aura proprement défini ce que sont des relations d'équivalence, cela sera plus simple...

Proposition - Multiplication par $u \in \mathbb{U}$

Soit $z \in \mathbb{C}$ et $u \in \mathbb{U} \setminus \{1\}$.

Notons $\theta = \arg(u)$.

Alors $u \times z$ est l'affixe du point obtenu par rotation de centre 0 et d'angle θ, à partir du point d'affixe z

La démonstration a été faite plus haut (dans l'analyse).

Attention - Mauvaise définition de l'argument...

La définition donnée ici est peu satisfaisante; avec un argument principal, on n'a pas nécessairement $\text{Arg}zz' = \text{Arg}z + \text{Arg}z'$, mais seulement des congruences.

3.2. Notation exponentielle

Argument ou argument (classe d'équivalence)

Corollaire - Propriété de $\angle\cdot$
Pour tout $\theta,\theta' \in \mathbb{R}$, $\angle\theta \times \angle\theta' = \angle(\theta+\theta')$

Démonstration
$\angle\theta \times \angle\theta'$ est l'affixe du point M' obtenu à partir de M d'affixe $\angle\theta'$ par rotation d'angle θ.
Il s'agit donc du point de $\mathbb{U}$ d'argument $\theta+\theta'$.
A noter qu'on peut « dépasser » π… □

Analyse - Non injectivité
Mais pour revenir à la question de l'argument, i.e. de $z = \angle\theta$ à $\theta = \text{Arg}z$, on a un soucis de non bijectivité (en fait de non injectivité) de $\angle$: $\angle\theta = \angle\theta' \not\Rightarrow \theta = \theta'$.
On doit considérer la classe d'équivalences de θ :

$$\overline{\theta} = \{\alpha \mid \exists\, k \in \mathbb{Z}, \alpha = \theta + 2k\pi\}$$

Notons donc $\theta \equiv \theta'[2\pi]$ pour exprimer : $\exists\, k \in \mathbb{Z}$ tel que $\theta' - \theta = 2k\pi$.

Définition - et Propriété. Argument de z
Pour tout $z \in \mathbb{U}$, on note $\arg z = \{\text{Arg}z + 2k\pi, k \in \mathbb{Z}\}$, on le considère *comme* un nombre.
Pour $z, z' \in \mathbb{U}$: $\arg(z \times z') = \arg z + \arg z'$ ou (de manière équivalente) : $\text{Arg}(z \times z') \equiv \text{Arg}z + \text{Arg}z'\ [2\pi]$.

Démonstration
Notons $\theta = \text{Arg}z$ et $\theta' = \text{Arg}z'$. On a $\theta, \theta' \in]-\pi,\pi]$, donc $\theta+\theta' \in]-2\pi, 2\pi]$
Alors $zz' = \angle\theta \times \angle\theta' = \angle\theta + \theta'$.
• Si $\theta+\theta' \in]-\pi,\pi]$, alors $\text{Arg}zz' = \theta + \theta^{p}rime = \text{Arg}z + \text{Arg}z'$.
• Si $\theta+\theta' \in]\pi, 2\pi]$, alors $\theta+\theta'-2\pi \in]-\pi, 0] \subset]-\pi,\pi]$ et $\text{Arg}zz' = \theta+\theta'-2\pi$.
• Si $\theta+\theta' \in]-2\pi, -\pi]$, alors $\theta+\theta'+2\pi \in]0,\pi] \subset]-\pi,\pi]$ et $\text{Arg}zz' = \theta+\theta'+2\pi$.
Par conséquent, dans tous les cas : $\text{Arg}zz' \equiv \text{Arg}z + \text{Arg}z'[2\pi]$. □

Notation d'Euler

Définition - Notation d'Euler
Nous verrons que dans le cas réel, on appelle exponentielle les fonctions qui vérifient $f(a+b) = f(a) \times f(b)$.
Elles s'écrivent (dans le cas réel) sous la forme $x \mapsto A^x$ où $A = f(1)$.
Par uniformité de notation, suivant L. Euler, on notera maintenant $e^{i\theta} = \angle\theta = \cos\theta + i\sin\theta$

Pour aller plus loin - Une vraie démonstration
Euler a bien démontré cette démonstration ; il ne s'agit pas d'une simple notation. Il faut donc voir qu'ici, :
— *il s'agit bien du nombre $e = 2,718281\ldots$*
— *il s'agit bien d'une puissance complexe*

On a alors, plus globalement :

Théorème - Propriétés
Soient $(\theta,\theta') \in \mathbb{R}^2$. On a :

$$e^{i(\theta+\theta')} = e^{i\theta}e^{i\theta'} \qquad \overline{e^{i\theta}} = e^{-i\theta} = \frac{1}{e^{i\theta}}$$

$$e^{i\frac{\pi}{2}} = i \qquad e^{i\pi} = -1$$

$$e^{i\theta} = 1 \Leftrightarrow \theta \equiv 0[2\pi] \Leftrightarrow \theta \in 2\pi\mathbb{Z} \qquad e^{i\theta} = e^{i\theta'} \Leftrightarrow \theta \equiv \theta'[2\pi]$$

Histoire - D'où vient la notation e ?
Ce n'est pas le e d'exponentielle, mais bien le e du suisse Leonard Euler
La formule d'Euler $e^{i\theta} = \cos\theta + i\sin\theta$ a été obtenue à partir du développement sous forme infinie de exp, cos *et* sin *(chapitre précédent), mais Euler n'a pas compris les propriétés géométriques du produit de complexes.*

Démonstration
On a alors $1 = e^{i0} = e^{i(\theta-\theta)} = e^{i\theta}e^{-i\theta}$, donc $e^{-i\theta} = \frac{1}{e^{i\theta}} = \overline{e^{i\theta}}$, car $e^{i\theta} \in \mathbb{U}$.
Le reste s'obtient facilement… □

Corollaire - Formule d'additions trigonométriques
Soient $a, b \in \mathbb{R}$,
$\cos(a+b) = \cos a \cos b - \sin a \sin b$ et $\sin(a+b) = \sin a \cos b + \cos a \sin b$

Démonstration
$\cos(a+b) + i\sin(a+b) = e^{i(a+b)} = e^{ia} \times e^{ib} = (\cos a + i \sin a)(\cos b + i \sin b)$
$= (\cos a \cos b - \sin a \sin b) + i(\sin a \cos b + \cos a \sin b)$
Reste à identifier les parties réelles et imaginaires. □

Exercice
En déduire les formules donnant $\cos(a-b)$ et $\sin(a-b)$.

Correction
Par parité de cos et imparité de sin, en faisant $b \mapsto -b$:
$\cos(a-b) = \cos a \cos b + \sin a \sin b$ et $\sin(a-b) = \sin a \cos b - \cos a \sin b$

3.3. Formules d'Euler et de de Moivre

Formules

Proposition - Formules d'Euler
$$\cos\theta = \mathrm{Re}(e^{i\theta}) = \frac{e^{i\theta} + e^{-i\theta}}{2} \quad , \quad \sin\theta = \mathrm{Im}(e^{i\theta}) = \frac{e^{i\theta} - e^{-i\theta}}{2i}.$$

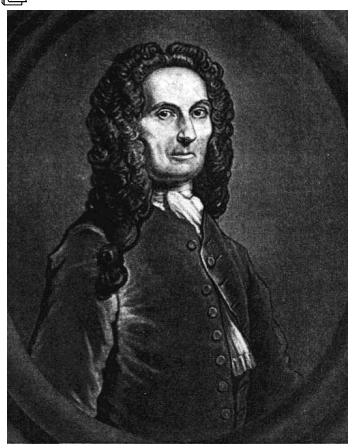
Exercice
Calculer $\frac{1}{3} + \frac{1}{4}$, en déduire une expression de $\cos\frac{7\pi}{12}$ et $\sin\frac{7\pi}{12}$.
On exploitera ces résultats plus tard.

Correction
$\frac{1}{3} + \frac{1}{4} = \frac{7}{12}$, donc
$$e^{i\frac{7\pi}{12}} = e^{i\frac{1}{3}\pi + i\frac{1}{4}\pi} = (\cos\frac{\pi}{3} + i\sin\frac{\pi}{3}) \times (\cos\frac{\pi}{4} + i\sin\frac{\pi}{4})$$
$$= (\frac{1}{2} + i\frac{\sqrt{3}}{2}) \times (\frac{\sqrt{2}}{2} + i\frac{\sqrt{2}}{2}) = \frac{\sqrt{2}-\sqrt{6}}{4} + i\frac{\sqrt{2}+\sqrt{6}}{4}$$
Donc : $\cos\frac{7\pi}{12} = \frac{\sqrt{2}-\sqrt{6}}{4}$ et $\sin\frac{7\pi}{12} = \frac{\sqrt{2}+\sqrt{6}}{4}$

Proposition - Formule de Moivre
$\forall\, n \in \mathbb{Z}$, $(e^{i\theta})^n = e^{in\theta}$
$\forall\, n \in \mathbb{Z}$, $(\cos\theta + i\sin\theta)^n = \cos n\theta + i\sin n\theta$

Démonstration
Par récurrence. On note pour tout $n \in \mathbb{N}$, $\mathscr{P}_n$: « $(e^{i\theta})^n = e^{in\theta}$ ».
— $(e^{i\theta})^0 = 1 = e^{i0\theta}$, donc $\mathscr{P}_0$ est vraie.
— Soit $n \in \mathbb{N}$, supposons que $\mathscr{P}_n$ est vraie.
On a donc $(e^{i\theta})^{n+1} = (e^{i\theta})^n \times e^{i\theta} = e^{in\theta}e^{i\theta}$ d'après $\mathscr{P}_n$.
puis $e^{i\theta})^{n+1} = e^{i(n+1)\theta}$, donc $\mathscr{P}_{n+1}$ est alors vérifiée.
La récurrence fonctionne bien et le résultat est démontrée pour tout n entier naturel.
Puis si $m \in \mathbb{Z}$ avec $m < 0$, alors en notant $n = -m$
$$(e^{i\theta})^m = (e^{i\theta})^{-n} = \frac{1}{(e^{i\theta})^n} = \frac{1}{e^{in\theta}} = e^{-in\theta} = e^{im\theta}$$
La seconde formule de Moivre est la notation algébrique de la précédente. □

Histoire - De Moivre



Abraham De MOIVRE *(1667, 1754), mathématicien d'origine française, mais qui du vivre en Angleterre. Hormis la formule de de Moivre (1707), il est connu pour son ouvrage sur les probabilités : the Doctrine of chances*

Angle moitié (pour factoriser)

Truc & Astuce pour le calcul - Factorisation de l'angle moitié
Lorsqu'on rencontre un expression de la forme $e^{ia} \pm e^{ib}$ (a, b réels), il faut toujours penser à factoriser par la moitié :
$$a = \frac{a+b}{2} + \frac{a-b}{2} \quad , \quad b = \frac{a+b}{2} - \frac{a-b}{2}$$

Cela donne :

$$e^{ia} \pm e^{ib} = e^{i\frac{a+b}{2}}\left(e^{i\frac{a-b}{2}} \pm e^{-i\frac{a-b}{2}}\right)$$

Et on applique les formules d'Euler

Exercice

Factoriser $1+e^{i\theta}$ et $1-e^{i\theta}$. (Re)trouver les formules donnant $1 \pm \cos\theta$

Correction

$$1+e^{i\theta} = e^{i\theta/2}(e^{-i\theta/2}+e^{i\theta/2}) = 2\cos\frac{\theta}{2}e^{i\theta/2} \quad , \quad 1-e^{i\theta} = e^{i\theta/2}(e^{-i\theta/2}-e^{i\theta/2}) = -2i\sin\frac{\theta}{2}e^{i\theta/2}$$

En prenant les parties réelles :

$$1+\cos\theta = 2\cos^2\frac{\theta}{2} \qquad 1-\cos\theta = 2\sin^2\frac{\theta}{2}$$

Exercice

Nouveau calcul de $\sum_{k=0}^{n}\cos kt$ et $\sum_{k=0}^{n}\sin kt$

Correction

On fait un seul calcul, la partie réelle donnera le premier résultat, la partie imaginaire le second.

$$S_n = \sum_{k=0}^{n} e^{ikt} = \sum_{k=0}^{n}(e^{it})^k = 1\frac{1-(e^{it})^{n+1}}{1-e^{it}} = \frac{1-e^{i(n+1)t}}{1-e^{it}} = \frac{e^{i(n+1)t/2}(-2i\sin\frac{(n+1)t}{2})}{e^{it/2}(-2i\sin\frac{t}{2})} = e^{int/2}\frac{\sin(n+1)t/2}{\sin t/2}$$

Formule d'Euler+ de Moivre + série géométrique + angle moitié +...

Donc

$$\sum_{k=0}^{n}\cos kt = \mathrm{Re}(S_n) = \frac{\cos(nt/2)\sin(n+1)t/2}{\sin t/2} \quad , \quad \sum_{k=0}^{n}\sin kt = \mathrm{Im}(S_n) = \frac{\sin(nt/2)\sin(n+1)t/2}{\sin t/2}$$

Linéarisation

Savoir faire - Linéarisation

Il s'agit d'exprimer $\cos^n\theta$ ou $\sin^n\theta$ sous forme d'une somme de $\cos k\theta$ ou $\sin k\theta$

(il ne doit plus y avoir de puissances ni de produits de cosinus ou sinus).

- Ecrire $\cos^n\theta = \left(\frac{e^{i\theta}+e^{-i\theta}}{2}\right)^n$.
- Développer avec la formule du binôme.
- Regrouper les termes conjugués pour faire apparaître des cosinus ou des sinus.

Si il n'y a pas de faute de calculs, vous devez obtenir un nombre réel : **donc simplification des i...**

Exercice

Linéariser $\cos^3\theta$, $\sin^4\theta$.

Correction

$$\cos^3\theta = \left(\frac{e^{i\theta}+e^{-i\theta}}{2}\right)^3 = \frac{e^{3i\theta}+3e^{i\theta}+3e^{-i\theta}+e^{3i\theta}}{8} = \frac{1}{4}\cos(3t)+\frac{3}{4}\cos(t)$$

$$\sin^4\theta = \left(\frac{e^{i\theta}-e^{-i\theta}}{2i}\right)^4 = \frac{e^{4i\theta}-4e^{2i\theta}+6-4e^{-2i\theta}+e^{4i\theta}}{16} = \frac{1}{8}\cos(4t)-\frac{1}{2}\cos(2t)+\frac{3}{8}$$

Pour aller plus loin - Polynôme de Tchebychev

Les polynômes de Tchebychev (env. 1860) sont définis par récurrence par :

$$T_0(x) = 1, T_1(x) = x$$

$$T_{n+2}(x) = 2xT_{n+1}(x) - T_n(x)$$

Ils vérifient alors $\forall\ n \in \mathbb{N}, \forall\ \theta \in \mathbb{R}$,

$$T_n(\cos\theta) = \cos(n\theta)$$

Savoir faire - Expressions de $\cos(nt)$ et $\sin(nt)$ en fonction de $\cos t$ et $\sin t$

- Ecrire $\cos(nt) = \mathrm{Re}(e^{int}) = \mathrm{Re}[(e^{it})^n]$ ou $\sin(nt) = \mathrm{Im}[(e^{it})^n]$.
- Utiliser la formule du binôme pour calculer $(e^{it})^n = (\cos t + i\sin t)^n$.
- Récupérer la partie réelle (ou imaginaire) en séparant les indices pairs des indices impairs.

Exercice

Ecrire $\cos 3t$ en fonction des puissances de $\cos t$, $\sin 3t$ comme le produit de $\sin t$ et d'une expression contenant des puissances de $\cos t$. Faire de même avec $\cos 5t$ et $\sin 5t$.

Correction

Là encore, on fait les deux questions en même temps, puis on prend la partie réelle et imaginaire.

$$\cos(3t)+i\sin(3t)=(e^{it})^3=(\cos t+i\sin t)^3=\cos^3 t+3i\cos^2 t\sin t-3\cos t\sin^2 t-i\sin^3 t$$

$$\cos(3t)=\cos^3 t-3\cos t\sin^2 t=\cos^3 t-3\cos t(1-\cos^2 t)=4\cos^3 t-3\cos t$$

$$\sin(3t)=3\cos^2 t\sin t-\sin^3 t=3(1-\sin^2 t)\sin t-\sin^3 t=3\sin t-4\sin^3 t$$

$$\cos(5t)+i\sin(5t)=(e^{it})^5=(\cos t+i\sin t)^5=\cos^5 t+5i\cos^4 t\sin t-10\cos^3 t\sin^2 t-10i\cos^2 t\sin^3 t+5\cos t\sin^4 t+i\sin^5 t$$

$$\cos(5t)=\cos^5 t-10\cos^3 t\sin^2 t+5\cos t\sin^4 t=16\cos^5 t-20\cos^3 t+5\cos t$$

$$\sin(5t)=5\cos^4 t\sin t-10\cos^2 t\sin^3 t+\sin^5 t=16\sin^5 t-20\sin^3 t+5\sin t$$

3.4. Argument, forme trigonométrique

Définition - Argument

Soit $z\in\mathbb{C}$, $z\neq 0$, on a $\dfrac{z}{|z|}\in\mathbb{U}$ donc il existe $\theta\in\mathbb{R}$ tel que $\dfrac{z}{|z|}=e^{i\theta}$.

On dit que θ est un argument de z. On note $\theta=\arg z$.

L'écriture $z=re^{i\theta}$ où $r=|z|$ est appelée forme trigonométrique de z.

Pour aller plus loin - Moins loin...

Rappelons qu'on a dit précédemment que l'argument était plutôt à considérer comme un ensemble de valeurs ; une classe d'équivalence.

L'argument est une fonction logarithmique (qui vérifie $f(ab)=f(a)+f(b)$), mais multivariée (à plusieurs valeurs).

Proposition - Arithmétique de la congruence

Si $(z,z')\in(\mathbb{C}^*)^2$, on a

$$\arg\overline{z}\equiv-\arg z\ [2\pi]$$
$$\arg\frac{1}{z}\equiv-\arg z\ [2\pi]$$
$$\arg(zz')\equiv(\arg z+\arg z')\ [2\pi]$$
$$\arg\left(\frac{z}{z'}\right)\equiv(\arg z-\arg z')\ [2\pi]$$

Attention - Pas d'unicité

Il n'y a pas unicité de l'argument, il est défini à 2π près. On peut imposer l'unicité de l'argument en le choisissant dans un intervalle de longueur 2π (en géneral $]-\pi,\pi]$ ou $[0,2\pi[$).

Remarque - Géométriquement

Soit z un complexe non nul et M le point d'affixe z. Toute mesure de l'angle orienté $(\vec{u},\overrightarrow{OM})$ est un argument de z.

Il faut savoir caractériser les complexes non nuls réels (resp. réels positifs, resp. réels négatifs, resp. imaginaires purs) par leur argument.

Proposition - Relation arg et arctan

Soit $x\in\mathbb{R}$, alors $\arg(1+ix)\equiv\arctan x\ [2\pi]$.

Soit $z\in\mathbb{C}$, alors $\arg z\equiv\arctan\frac{\mathrm{Im}(z)}{\mathrm{Re}(z)}\ [\pi]$. Précisément :

$$\arg z\equiv\begin{cases}\arctan\frac{\mathrm{Im}(z)}{\mathrm{Re}(z)}\ [2\pi] & \text{si } \mathrm{Re}(z)>0\\ \pi+\arctan\frac{\mathrm{Im}(z)}{\mathrm{Re}(z)}\ [2\pi] & \text{si } \mathrm{Re}(z)>0\end{cases}$$

Démonstration
Soit $z = 1 + ix$, on note $\rho = \sqrt{1+x^2}$ son module et θ son argument principal.

$$z = 1 + ix = \rho\cos\theta + i\rho\sin\theta$$

On peut identifier :

$$x = \frac{x}{1} = \frac{\rho\sin\theta}{\rho\cos\theta} = \tan\theta \Longleftrightarrow \theta = \arctan x$$

□

4. Racines d'un nombre complexe

4.1. Recherche de racines carrées

On dit que $Z \in \mathbb{C}$ est une racine carrée de $z \in \mathbb{C}$ si $Z^2 = z$. On dispose de deux méthodes pour chercher les racines carrées de z.

Résolution trigonométrique (la meilleure!)

Truc & Astuce pour le calcul - Racine carrée. Exploitation de la forme trigonométrique

On considère un complexe non nul z écrit sous forme trigonométrique $z = |z|e^{i\alpha}$, et on cherche Z sous forme trigonométrique $Z = \rho e^{i\theta}$ où $\rho > 0$.
On a alors $Z^2 = \rho^2 e^{2i\theta}$, on fait ensuite une sorte d'identification entre les modules et les arguments (mais attention...).

Exercice

Trouver les racines carrées de $z = \dfrac{1-i}{\sqrt{3}-i}$.

On rappelle que $\cos\frac{7\pi}{12} = \frac{\sqrt{2}-\sqrt{6}}{4}$ *et* $\sin\frac{7\pi}{12} = \frac{\sqrt{2}+\sqrt{6}}{4}$

Correction

Il faut écrire z sous forme trigonométrique, on multiplie le dénominateur par la quantité conjuguée (non nécessaire) :

$$z = \frac{(1-i)(\sqrt{3}+i)}{3+1} = \frac{(1+\sqrt{3})+i(1-\sqrt{3})}{4}$$

$$|z|^2 = \frac{1}{16}((1+\sqrt{3})^2 + (1-\sqrt{3})^2) = \frac{1}{2} \quad \Longrightarrow |z| = \frac{1}{\sqrt{2}}$$

Si $\arg(z) = \theta$, alors $\cos\theta = \dfrac{\sqrt{2}}{4}(1+\sqrt{3}) = \sin\frac{7\pi}{12}$ et $\sin\theta = \dfrac{\sqrt{2}}{4}(1-\sqrt{3}) = \cos\frac{7\pi}{12}$.
Donc $\theta = \dfrac{\pi}{2} - \dfrac{7\pi}{12} = \dfrac{-\pi}{12}$ Les racines carrées de z sont alors : $Z_1 = \frac{1}{2^{1/4}}e^{-i\theta/24}$ et $Z_2 = -\frac{1}{2^{1/4}}e^{-i\theta/24}$.

La méthode-algorithmique précédente nous permet d'affirmer :

Proposition - Deux racines complexes
Tout complexe non nul possède exactement deux racines carrées complexes (opposées).

Résolution algébrique

Truc & Astuce pour le calcul - Racine carrée. Exploitation de la forme algébrique

On considère donc un complexe non nul z écrit sous forme algébrique $z = x + iy$, et on cherche Z sous forme algébrique $Z = X + iY$.
Le principe est d'écrire l'égalité des modules, des parties réelles et imaginaires de z et Z^2 pour se ramener à une résolution simple de système

donnant X^2, Y^2 et le signe de XY.

$$Z^2 = z \Leftrightarrow \begin{cases} X^2+Y^2 = \sqrt{x^2+y^2} \\ X^2-Y^2 = x \\ 2XY = y \end{cases}$$

On résout le système formée par les deux premières équations, la troisième donne le signe de XY.

Exercice
Déterminer les racines carrées de $2-3i$.

Correction
On considère $Z = X+iY$ avec $Z^2 = 2-3i$.
Donc $|Z|^2 = X^2+Y^2 = |Z^2| = \sqrt{4+9} = \sqrt{13}$.
Puis $\mathrm{Re}(Z^2) = X^2-Y^2 = 2$.
On a donc $X^2 = \frac{2+\sqrt{13}}{2}$ et $Y^2 = -1+\frac{\sqrt{13}}{2}$.
Et comme $\mathrm{Im}(Z^2) = 2XY = -3 < 0$, X et Y sont de signe opposé.
Ainsi $Z = \pm\left(\sqrt{1+\frac{\sqrt{13}}{2}} - i\sqrt{-1+\frac{\sqrt{13}}{2}}\right)$

Equation du second degré

Le théorème suivant a déjà été vu. Mais ici, on insiste sur le fait que les coefficients peuvent être des nombres complexes.

Proposition - Nombre de racines et degré
L'équation $az^2+bz+c=0$, avec $(a,b,c) \in \mathbb{C}^3$, $a \neq 0$, admet deux solutions complexes (éventuellement confondues) $z_1 = \dfrac{-b-\delta}{2a}$ et $z_2 = \dfrac{-b+\delta}{2a}$ où δ est une racine carrée complexe de b^2-4ac.

Remarque - Bien connu...
On retrouve le résolution déjà connue d'une équation du second degré à coefficients réels.

Proposition - Théorème de Viète
Soient $(S,P) \in \mathbb{C}^2$. Les solutions du système

$$\begin{cases} z_1+z_2 = S \\ z_1 \times z_2 = P \end{cases}$$

sont exactement (à permutation près) les solutions de $z^2-Sz+P=0$

Exercice
Résoudre dans $\mathbb{C}$ le système d'équation $\begin{cases} z_1+z_2 = 3 \\ z_1 \times z_2 = 1-3i \end{cases}$.

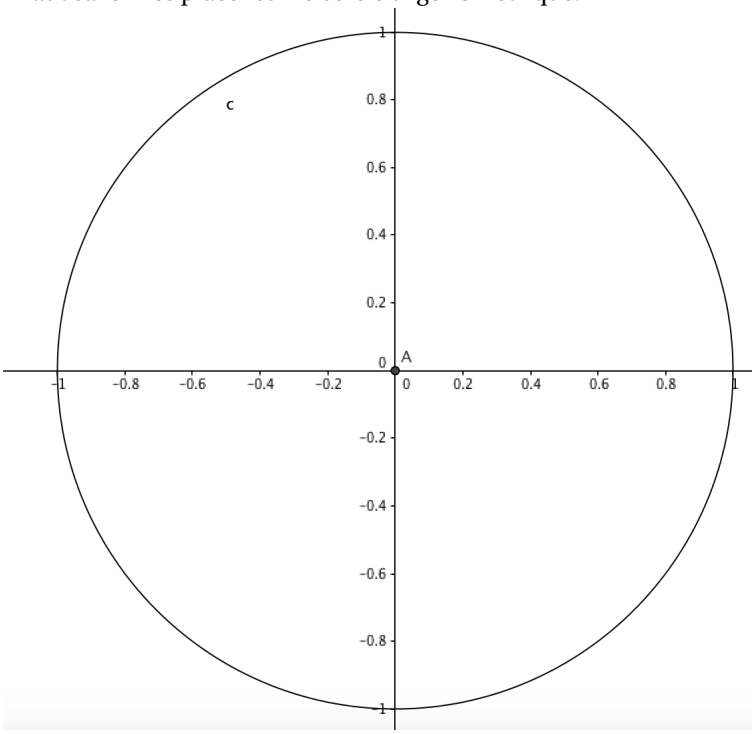
Correction
Il faut trouver les solutions de l'équation $z^2-3z+1-3i=0$.
Le discriminant est $\Delta = 9-4+12i = 5+12i$.
On cherche $\delta = a+ib$ tel que $\delta^2 = \Delta$, donc
$|\delta|^2 = a^2+b^2 = |\delta^2| = |\Delta| = \sqrt{25+144} = \sqrt{169} = 13$.
$\Re(\delta^2) = a^2-b^2 = \Re(\Delta) = 5$.
Donc $a^2 = \dfrac{13+5}{2} = 9$ et $b^2 = 4$.
Enfin, comme $2ab = \Im(\delta^2) = \Im(\Delta) = 12 > 0$, on a donc $\delta = 3+2i$ ou $\delta = -3-2i$.
Et par conséquent, les racines de l'équations sont (à permutation près) :
$z_1 = \dfrac{3+3+2i}{2} = 3+i$ et $z_2 = \dfrac{3-3-2i}{2} = -i$.

4.2. Racines n-ièmes de l'unité

Pour aller plus loin - Groupes $(\mathbb{U}_n, \times)$
L'ensemble $\mathbb{U}_n$ des racines n-ieme de l'unité, avec la loi de multiplication $\times$ en fait un groupe.
Nous reprendrons son étude plus précisément, dans quelques semaines.

Théorème - Les n solutions de $z^n = 1$
Soit $n \in \mathbb{N}^*$. Les n racines n-ièmes de l'unité, c'est à dire les solutions de l'équation $z^n = 1$, sont les n nombres $e^{\frac{2ik\pi}{n}}$ avec $k \in \{0, 1, \dots, n-1\}$.
On note $\mathbb{U}_n = \left\{ e^{\frac{2ik\pi}{n}} ; k \in \{0, 1, \dots, n-1\} \right\}$

On obtient donc pour
$n = 2$: 1 et -1 ;
$n = 3$: 1, $j = e^{\frac{2i\pi}{3}} = \exp\frac{2i\pi}{3}$ et $j^2 = \overline{j} = e^{\frac{4i\pi}{3}} = \exp\frac{4i\pi}{3}$;
$n = 4$: $1, i, -1$ et $-i$.
Il faut savoir les placer sur le cercle trigonométrique.



Proposition - Somme des racines n-ième
Soit $n \in \mathbb{N}$, $n \geqslant 2$. La somme des racines n-ièmes de l'unité est nulle.
En particulier $1 + j + j^2 = 0$.

Démonstration
On peut identifier dans le développement polynomiale ou faire le calcul :
$\sum_{k=0}^{n-1} e^{\frac{2ik}{n}\pi} = \sum_{k=0}^{n-1} (e^{\frac{2i\pi}{n}})^k = 1 \dfrac{1 - e^{2in\pi/n}}{1 - e^{2i\pi/n}} = 0$ □

4.3. Racines n-ièmes d'un nombre complexe

Théorème - Racines n-ièmes de z_0
Soient $z_0 \in \mathbb{C}^*$ et $n \in \mathbb{N}^*$. Alors z_0 a exactement n racines n-ièmes (solutions de $z^n = z_0$).
Si $z_0 = |z_0| e^{i\alpha}$, alors ce sont les

$$z_k = |z_0|^{1/n} e^{i(\frac{\alpha}{n} + \frac{2k\pi}{n})} \text{ où } k \in \{0, 1, \dots, n-1\}.$$

Exercice

Déterminer les racines n-ièmes de $\frac{1+\sqrt{3}i}{1-i}$.

On rappelle que $\cos\frac{7\pi}{12} = \frac{\sqrt{2}-\sqrt{6}}{4}$ *et* $\sin\frac{7\pi}{12} = \frac{\sqrt{2}+\sqrt{6}}{4}$

Correction

On note $z = \frac{1+\sqrt{3}i}{1-i} = \frac{(1+\sqrt{3}i)(1+i)}{(1-i)(1+i)} = \frac{(1-\sqrt{3})+i(1+\sqrt{3})}{2} = \sqrt{2}\times(\frac{\sqrt{2}-\sqrt{6}}{4}) + i\frac{\sqrt{2}+\sqrt{6}}{4}) = \sqrt{2}e^{i\frac{7\pi}{12}}$.

Ainsi, les racines n-ièmes de z sont les nombres $\rho\epsilon_k$, pour $k \in \{0,1,\dots n-1\}$, avec $\rho = 2^{1/2n}$ et $\epsilon_k = e^{i\frac{7\pi}{12n}+\frac{2k\pi}{n}}$.

Exercice

Résoudre dans $\mathbb{C}$ l'équation $(z-1)^6 + (z+1)^6 = 0$.

Correction

On a donc $\left(\frac{z-1}{z+1}\right)^6 = -1$. Puis $\frac{z-1}{z+1}$, racine 6-ième de $-1 = e^{i\pi}$.

Donc $\frac{z-1}{z+1} = \epsilon_k$, avec $k \in \{0,1,\dots,5\}$ et $\epsilon_k = e^{i(2k+1)\pi/6}$.

Et ensuite $z-1 = \epsilon_k(z+1)$ donc $(1-\epsilon_k)z = 1+\epsilon_k$, donc $z_k = \frac{1+\epsilon_k}{1-\epsilon_k}$.

En factorisant par l'angle moitié : $z_k = \frac{\cos(\frac{2k+1}{12}\pi)}{-i\sin(\frac{2k+1}{12}\pi)} = \frac{i}{\tan(\frac{2k+1}{12}\pi)}$

On voit bien sur ces exemples qu'il est préférable d'exploiter la forme géométrique lorsqu'on cherche des racines de nombres complexes…

5. $\mathbb{R}^2 = \mathbb{C} = \mathscr{P}$

5.1. Regard géométrique sur le plan complexe

Proposition - Identification

On munit le plan $\mathscr{P}$ d'un repère orthonormé $(O,\vec{i},\vec{j})$.

Soient A,B,A',B' quatre points du plan. On a alors (*mesure des angles orientés de vecteurs*) :

$|z_A| = OA$

$AB = |z_B - z_A|$

$\arg z_A \equiv (\vec{i},\overrightarrow{OA})[2\pi]$ et plus généralement : $\arg(z_B - z_A) \equiv (\vec{i},\overrightarrow{AB})[2\pi]$

$\arg\left(\frac{z_{B'}-z_{A'}}{z_B-z_A}\right) \equiv (\overrightarrow{AB},\overrightarrow{A'B'})[2\pi]$

Le dernier résultat, essentiel, mérite une démonstration

Démonstration

Il s'agit d'angle orienté :

$$(\overrightarrow{AB},\overrightarrow{A'B'}) = (\vec{i},\overrightarrow{A'B'}) - (\vec{i},\overrightarrow{AB}) = \arg(z_{B'}-z_{A'}) - \arg(z_B - z_A) = \arg\left(\frac{z_{B'}-z_{A'}}{z_B-z_A}\right)$$

□

Analyse - Le nombre $Z = z' \times \overline{z}$

Etant donné deux vecteurs $\overrightarrow{AB}$ d'affixe $z(= z_B - z_A)$ et $\overrightarrow{A'B'}$ d'affixe $z'(= z_{B'} - z_{A'})$, alors le dernier calcul pour trouver l'angle entre les deux vecteurs conduit à nous intéresser au nombre complexe :

$$Z = z' \times \overline{z}$$

Son argument donne l'angle (orienté) entre ces deux vecteurs.

Définition - Le complexe Z

Soient A,B,A',B' quatre points du plan d'affixe z_A, z_B, $z_{A'}$ et $z_{B'}$ respectivement.

On note $Z = (z_{B'} - z'_A) \times \overline{(z_B - z_A)}$.

Alors
$$\arg Z \equiv (\overrightarrow{AB}, \overrightarrow{A'B'})[2\pi] \quad \text{et} \quad |Z| = \|\overrightarrow{AB}\| \times \|\overrightarrow{A'B'}\|$$
où $\|\overrightarrow{u}\| = \sqrt{x^2+y^2}$ est par définition la norme (longueur) du vecteur $\overrightarrow{u}(x,y)$.
Et donc
$$Z = \|\overrightarrow{AB}\|\|\vec{A'B'}\| \left[\cos(\overrightarrow{AB}, \overrightarrow{A'B'}) + i\sin(\overrightarrow{AB}, \overrightarrow{A'B'})\right]$$

Le calcul donne :

Proposition - Partie réelle et imaginaire de Z
Avec les mêmes notations et en notant $\overrightarrow{u} = \overrightarrow{AB}$ d'affixe $z = z_B - z_A = x + iy$ et $\overrightarrow{u'} = \overrightarrow{A'B'}$ d'affixe $z' = z_{B'} - z_{A'} = x' + iy'$.
On rappelle que $Z = z' \times \overline{z}$. On a alors :
$$\mathrm{Re}(Z) = \|\overrightarrow{AB}\|\|\vec{A'B'}\| \cos(\overrightarrow{AB}, \overrightarrow{A'B'}) = xx' + yy'$$
$$\mathrm{Im}(Z) = \|\overrightarrow{AB}\|\|\vec{A'B'}\| \sin(\overrightarrow{AB}, \overrightarrow{A'B'}) = xy' - x'y$$

Démonstration
Il s'agit juste de vérifier l'égalité avec les parties réelles et imaginaires de z et z'.
$$Z = z \times \overline{z'} = (x+iy) \times (x'-iy') = (xx'+yy') + i(xy'-x'y)$$
□

5.2. Lignes de niveau

Une droite est un ensemble de points alignés. Pour définir une (équation de) droite, on exploite donc les deux points de vue sur l'alignement.

Théorème - Utilisation des complexes - Droite
1. La droite (AB) privée des points A et B (d'affixes respectives a et b) est l'ensemble des points M d'affixe z vérifiant
$$\arg\left(\frac{z-b}{z-a}\right) \equiv 0[\pi].$$
2. La droite (AB) privée du point A est l'ensemble des points M d'affixe z vérifiant
$$\frac{z-b}{z-a} = \overline{\left(\frac{z-b}{z-a}\right)}$$
3. La droite (AB) est l'ensemble des points M d'affixe z vérifiant
$$(z-b)(\overline{z-a}) = (\overline{z-b})(z-a).$$

Démonstration
Notons $\mathscr{D}$, la droite (AB) privées des points A et B (pour avoir un dénominateur non nul dans le calcul).
$$M(z) \in \mathscr{D} \iff (\overrightarrow{AM}, \overrightarrow{BM}) \equiv 0[\pi] \iff \arg\left(\frac{z-b}{z-a}\right) \equiv 0[\pi]$$
Et
$$M(z) \in \mathscr{D} \iff \mathrm{Im}((b-z)\overline{(a-z)}) = 0 \iff (b-z)\overline{(a-z)} \in \mathbb{R} \iff (z-b)(\overline{z-a}) = \overline{(z-b)(\overline{z-a})}$$
$$\iff (z-b)(\overline{z-a}) = (\overline{z-b})(z-a) \iff \frac{z-b}{z-a} = \overline{\left(\frac{z-b}{z-a}\right)}$$
□

niveau circu-

l'affixe z véri-

infinie si $\theta \equiv$
privé de ces

suivant les
émonstration
information
ce cercle se
trice de $[AB]$
gle au centre
e...).

Exercice

Donner l'équation de la droite (complexe) qui passe par les points $A(1+i)$ et $B(2-i)$.

Correction

$$M(z) \in (AB) \iff (z-(1+i))(\overline{z-(2-i)}) = (\overline{z-(1+i)})(z-(2-i)) \iff z\overline{z}-(2+i)z-(1+i)\overline{z}+(1+3i) = z\overline{z}-(1+i)z-(2-i)\overline{z}+1-3i$$
$$\iff (1+2i)z-(1-2i)\overline{z}-6i = 0$$

On peut vérifier que A et B appartiennent bien à la droite :
$(1+2i)(1+i)-(1-2i)(1-i) = 2i\text{Im}((1+2i)(1+i)) = 6i$ et $2i\text{Im}((1+2i)(2-i)) = 6i$.

Théorème - Ligne de niveau - Cercle

Soient A, B, C trois points distincts d'affixes respectives a, b et c.

1. $(AB) \perp (AC) \Leftrightarrow \arg\left(\dfrac{c-a}{b-a}\right) \equiv \dfrac{\pi}{2}[\pi]$
2. L'ensemble des points M d'affixe z vérifiant
$$\arg\left(\frac{z-b}{z-a}\right) \equiv \frac{\pi}{2}[\pi]$$
est le cercle de diamètre $[AB]$ privé des points A et B.
3. L'ensemble des points M d'affixe z vérifiant
$$\frac{z-b}{z-a} = -\overline{\left(\frac{z-b}{z-a}\right)}$$
est le cercle de diamètre $[AB]$ privé du point A.
4. L'ensemble des points M d'affixe z vérifiant $(z-b)\overline{(z-a)} = -\overline{(z-b)}(z-a)$ est le cercle de diamètre $[AB]$.

Démonstration

Le premier point découle d'une proposition précédente.
Nous allons procéder en deux temps (double inclusion d'ensemble).
On note $\mathscr{E} = \left\{M(z) \in \mathbb{C} \mid \arg\dfrac{z-b}{z-c} \equiv \dfrac{\pi}{2}[\pi]\right\}$ et $\mathscr{C}$ le cercle de diamètre $[AB]$.
Il allons de montrer, par double inclusion, que $\mathscr{E} = \mathscr{C} \setminus \{A, B\}$.

— Le centre du cercle $\mathscr{C}$ est K d'affixe $k = \dfrac{a+b}{2}$ et de rayon $r = \dfrac{1}{2}|AB| = \dfrac{|b-a|}{2}$

$$z \in \mathscr{C} \setminus \{A, B\} \iff z = k + re^{i\theta} \text{ avec } \theta \neq \theta_0[\pi]$$

où $\theta_0 = \arg(a-b) = (\vec{i}, \overrightarrow{BA})$, donc $a - b = 2re^{i\theta_0}$.
Or

$$\frac{z-b}{z-a} = \frac{k-b+re^{i\theta}}{k-a+re^{i\theta}} = \frac{(a-b)+2re^{it\theta}}{(b-a)+2re^{i\theta}} = \frac{e^{i\theta}+e^{i\theta_0}}{e^{i\theta}-e^{i\theta_0}} = \frac{2e^{i\frac{\theta-\theta_0}{2}}\cos\frac{\theta-\theta_0}{2}}{2e^{i\frac{\theta-\theta_0}{2}}\,i\sin\frac{\theta-\theta_0}{2}} = -\frac{\cos\frac{\theta-\theta_0}{2}}{\sin\frac{\theta-\theta_0}{2}}\,i$$

Donc $\arg\left(\dfrac{z-b}{z-a}\right) \equiv \arg(-i) \equiv \dfrac{\pi}{2}[\pi]$. Donc $\mathscr{C} \setminus \{A, B\} \subset \mathscr{E}$.

— Réciproquement, si $z \in \mathscr{E}$, il existe $R \in \mathbb{R}$ avec $z \neq a, b$

$$\frac{z-b}{z-a} = Ri \iff z-b = Ri(z-a) \iff (1-Ri)z = b-Ria \iff z = \frac{b-Ria}{1-Ri} = \frac{(b+R^2a)+Ri(b-a)}{1+R^2}$$

Donc, toujours avec $k = \frac{a+b}{2}$, affixe de K, milieu de $[AB]$,

$$z-k = z-\frac{a+b}{2} = \frac{(b-a)[1-R^2+2Ri]}{2(1+R^2)} \Longrightarrow |z-k| = \frac{|b-a|\times|1-R^2+2Ri|}{2(1+R^2)} = \frac{|b-a|}{2}$$

Ce nombre est contant (indépendant de z), c'est le rayon du cercle représenté par $\mathscr{E}$.
Et on a donc $\mathscr{E} \subset \mathscr{C} \setminus \{A, B\}$

Pour finir, nous savons que si $Z \neq 0$,

$$\arg(Z) \equiv \frac{\pi}{2}[\pi] \longleftrightarrow Z \in i\mathbb{R} \longleftrightarrow Z = -\overline{Z}$$

□

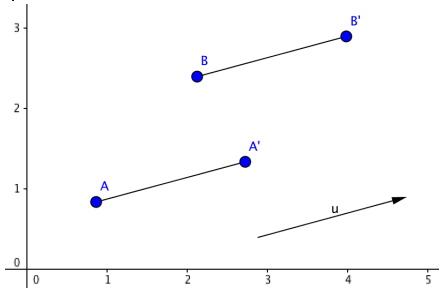
5.3. Transformations du plan (point de vue complexe)

Transformations « élémentaires »

Définition - Transformation du plan
On appelle transformation du plan toute bijection du plan dans lui-même.

Attention - Projection
Une projection sur une droite n'est pas une transformation du plan.

Représentation - Translation de vecteur $\vec{u}$



Définition - Translation
On appelle translation de vecteur $\overrightarrow{u}$ l'application

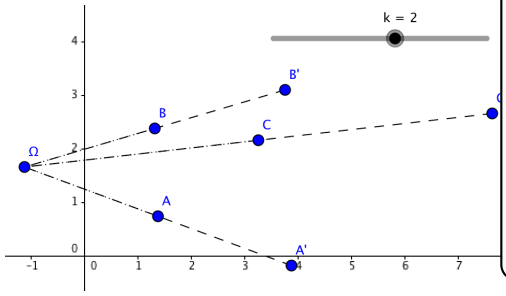
$$\begin{array}{rccc} t_{\overrightarrow{u}} : & \mathscr{P} & \to & \mathscr{P} \\ & M & \mapsto & M' \text{ tel que } \overrightarrow{MM'} = \overrightarrow{u} \end{array}$$

En complexes, $t_{\overrightarrow{u}}$ est représentée par l'application de $\mathbb{C}$ dans $\mathbb{C}$ définie par

$$z \mapsto z + z_0$$

où z_0 est l'affixe du vecteur $\overrightarrow{u}$.

Représentation - Homothétie de centre Ω et de rapport $k = 2$



Définition - Homothétie
On appelle homothétie de centre Ω et de rapport le réel $k \neq 0$ l'application

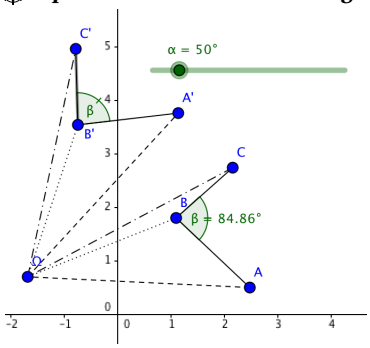
$$\begin{array}{rccc} h_{\Omega,k} : & \mathscr{P} & \to & \mathscr{P} \\ & M & \mapsto & M' \text{ tel que } \overrightarrow{\Omega M'} = k\overrightarrow{\Omega M} \end{array}$$

En complexes, $h_{\Omega,k}$ est représentée par l'application de $\mathbb{C}$ dans $\mathbb{C}$ définie par

$$z \mapsto z_0 + k(z - z_0)$$

où z_0 est l'affixe du point Ω.

Représentation - Rotation d'angle α



Définition - Rotation
On appelle rotation de centre Ω et d'angle θ l'application

$$\begin{array}{rccc} R_{\Omega,\theta} : & \mathscr{P} & \to & \mathscr{P} \\ & M & \mapsto & M' \text{ tel que } \begin{cases} \Omega M' = \Omega M \\ (\overrightarrow{\Omega M}, \overrightarrow{\Omega M'}) \equiv \theta\,[2\pi] \end{cases} \text{ si } M \neq \Omega \\ & \Omega & \mapsto & \Omega \end{array}$$

En complexes, $R_{\Omega,\theta}$ est représentée par l'application de $\mathbb{C}$ dans $\mathbb{C}$ définie par

$$z \mapsto z_0 + e^{i\theta}(z - z_0)$$

où z_0 est l'affixe du point Ω.

Similitudes (directes)

Il s'agit de composition de rotation et d'homothétie...

Définition - Similitude directe
On appelle similitude directe du plan toute transformation représentée dans le plan complexe par une application de la forme

$$z \mapsto az + b$$

avec $(a, b) \in \mathbb{C}^* \times \mathbb{C}$.

Remarque - Les transformations élémentaires
Translations, homothéties et rotations sont des similitudes directes, il suffit de regarder leur expression en complexe.

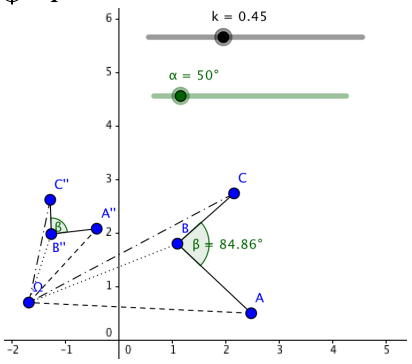
Analyse - Résultats caractéristiques
Si A, B sont deux points distincts d'images respectives A', B' par la transformation associée à $z \mapsto az + b$, alors

$$(\overrightarrow{AB}, \overrightarrow{A'B'}) \equiv \arg a\,[2\pi] \text{ et } \frac{A'B'}{AB} = |a|.$$

Proposition - Similitude directe
Une similitude directe conserve les angles et les rapports des distances.
Si A, B, C sont trois points distincts d'images respectives A', B', C' par Alors

$$(\overrightarrow{AB}, \overrightarrow{AC}) \equiv (\overrightarrow{A'B'}, \overrightarrow{A'C'})\,[2\pi] \qquad \text{et} \qquad \frac{AB}{AC} = \frac{A'B'}{A'C'}$$

Représentation - Similitude (directe)



Démonstration
Supposons que la transformation est $z \mapsto az + b$. Et notons z_M, l'affixe du point M, générique.

$$\frac{z_{A'} - z_{B'}}{z_{A'} - z_{C'}} = \frac{az_A + b - az_B - b}{az_A + b - az_C - b} = \frac{z_A - z_B}{z_A - z_C}$$

En prenant l'argument et le module, on retrouve respectivement la conservation des angles et des longueurs. □

Théorème - Caractérisation
Soit f la transformation du plan représentée par $z \mapsto az + b$ avec $a \in \mathbb{C}^*, b \in \mathbb{C}$.

- Si $a = 1$, f est la translation de vecteur d'affixe b.
- Si $a \neq 1$, f admet un unique point fixe (point invariant) $\Omega(\frac{b}{1-a})$ appelé centre de la similitude.
  f s'écrit alors : $f = h \circ r = r \circ h$ avec
  - r rotation de centre Ω d'angle de mesure $\arg a$
  - h homothétie de centre Ω et de rapport $|a|$.

  On dit que $|a|$ est le rapport de la similitude, et $\arg a$ est (la mesure de) l'angle de la similitude.

Démonstration
Soit $\Omega(z)$, un point fixe de f i.e. : $z = f(z)$ i.e. $z = az + b$, donc $z = \dfrac{b}{1-a}$ $(a \neq 1)$.
On a donc un unique point fixe. Notons $z_0 = \frac{b}{1-a}$.
Ainsi, pour tout $z \in \mathbb{C}$,

$$f(z) - z_0 = f(z) - f(z_0) = (az + b) - (az_0 + b) = a(z - z_0) \qquad \Longrightarrow \qquad f(z) = z_0 + a(z - z_0)$$

En notant $a = |a|e^{i\arg(a)}$, on retrouve le produit complexe ou la composition de transformation de la rotation $(\times e^{i\theta})$ et de l'homothétie $(\times |a|)$. □

Remarque - Cas particuliers avec $a \neq 1$:
- Si $a \in \mathbb{R}^*$, f est l'homothétie de centre Ω de rapport a.
- Si $|a| = 1$, f est la rotation de centre Ω, d'angle $\theta = \arg a$.

Corollaire - Caractérisation de la similitude
La similitude de centre d'affixe z_0, de rapport k et d'angle θ est représentée par

$$z \mapsto z_0 + ke^{i\theta}(z - z_0).$$

Savoir faire - Reconnaître une similitude

Etant donnée une transformation, pour reconnaître une similitude il faut :

1. chercher le point fixe : la solution de $f(z) = z$.
   On le note z_0, c'est le centre de la similitude.
2. chercher le complexe a tel que $f(z) - z_0 = a(z - z_0$.
   Ce complexe a donne le rapport et l'angle de la similitude

Proposition - Composée de similitudes
La composée de deux similitudes est une similitude dont le rapport est le produit des rapports et l'angle, la somme des angles.
On a les cas particuliers suivants :

- la composée de deux translations est une translation de vecteur la somme des deux vecteurs,
- la composée de deux homothéties est soit une homothétie soit une translation,
- la composée de deux rotations de même centre est une rotation de même centre et d'angle la somme des deux angles (éventuellement d'angle nul, i.e. l'identité du plan)
- la composée de deux rotations de centres distincts, d'angles respectifs θ et θ' est :
  - une rotation si $\theta + \theta' \not\equiv 0[2\pi]$
  - une translation (éventuellement de vecteur nul, i.e. l'identité) sinon.

Démonstration
Considérons les deux similitudes $z \mapsto az + b$ et $z \mapsto a'z + b'$.
La composée des deux similitudes est

$$z \mapsto a(a'z + b') + b = aa'z + (ab' + b)$$

On reconnait, à cas particuliers près, une similitude. Pour les cas particuliers, on reprend chacun dans le même ordre :

- la composée de deux translations est une translation de vecteur la somme des deux vecteurs.
  Dans ce cas $a = a' = 1$, on a bien une translation.
- la composée de deux homothéties est soit une homothétie soit une translation.
  Dans ce cas $a, a' \in \mathbb{R}$, donc $aa' \in \mathbb{R}$; on a donc une homothétie ou une translation (si $aa' = 1$)
- la composée de deux rotations de même centre est une rotation de même centre et d'angle la somme des deux angles (éventuellement d'angle nul, i.e. l'identité du plan).
  Dans ce cas, en notant Ω ce centre : $r \circ r'(\Omega) = r(\Omega) = \Omega$. Donc Ω est centre de la similitude.
  Comme $a, a' \in \mathbb{U}$, alors $aa' \in \mathbb{U}$ et donc cette similitude est une rotation.
  Son angle est $\arg(aa') = \arg(a) + \arg(a')$.
- la composée de deux rotations de centres distincts, d'angles respectifs θ et θ'.
  On a $a = e^{i\theta}$ et $a' = e^{i\theta'}$, donc il s'agit de l'application $z \mapsto e^{i(\theta+\theta')}z + (ab' + b)$.
  Et donc il s'agit d'une
  - rotation si $\theta + \theta' \not\equiv 0[2\pi]$
  - translation (éventuellement de vecteur nul, i.e. l'identité) sinon.

□

Corollaire - Transformation réciproque (ou inverse)
On a les transformations réciproques suivantes :

$$t_{\vec{u}}^{-1} = t_{-\vec{u}}; \quad h_{\Omega,k}^{-1} = h_{\Omega,\frac{1}{k}}; \quad R_{\Omega,\theta}^{-1} = R_{\Omega,-\theta}.$$

Proposition - Transformation du plan
Etant donné deux segments $[MN]$ et $[M'N']$ de longueurs non nulles, il existe une et une seule similitude directe transformant M en M' et N en N'.

Démonstration
On peut typiquement faire une démonstration en analyse-synthèse.
Supposons que cette transformation existe, notée $z \mapsto az + b$.
On a donc $z_{M'} = az_M + b$ et $z_{N'} = az_N + b$. Et donc $a = \dfrac{z_{M'} - z_{N'}}{z_M - z_N}$.

a est donc unique et de même $b = z_{M'} - \dfrac{z_{M'} - z_{N'}}{z_M - z_N} z_M$.
Mais pour s'assurer de l'existence, il faut qu'on retrouve bien $az_N + b = z_{N'}$:

$$az_N + b = \frac{z_{M'} - z_{N'}}{z_M - z_N} z_N + z_{M'} - \frac{z_{M'} - z_{N'}}{z_M - z_N} z_M = \frac{(z_{M'} - z_{N'})(z_N - z_M)}{z_M - z_N} + z_{M'} = -z_{M'} + z_{N'} + z_{M'} = z_{N'}$$

□

Symétries

Définition - Symétries
On appelle :

— Symétrie centrale de centre Ω l'application

$$\begin{array}{rccc} s_\Omega : & \mathscr{P} & \to & \mathscr{P} \\ & M & \mapsto & M' \text{ tel que } \overrightarrow{\Omega M'} = -\overrightarrow{\Omega M} \end{array}$$

En complexes, s_Ω est représentée par l'application de $\mathbb{C}$ dans $\mathbb{C}$ définie par $z \mapsto 2z_0 - z$ où z_0 est l'affixe du point Ω.

— Symétrie orthogonale d'axe la droite $\mathscr{D}$ (ou réflexion d'axe $\mathscr{D}$) l'application

$$\begin{array}{rccc} s_\mathscr{D} : & \mathscr{P} & \to & \mathscr{P} \\ & M & \mapsto & M' \text{ tel que } \begin{cases} M' = M \text{ si } M \in \mathscr{D} \\ \mathscr{D} \text{ est la médiatrice de } [MM'] \text{ sinon} \end{cases} \end{array}$$

La symétrie orthogonale par rapport à Ox est représentée par l'application de $\mathbb{C}$ dans $\mathbb{C}$ définie par $z \mapsto \overline{z}$.

Remarque - Symétrie centrale
Une symétrie centrale peut être considérée comme une homothétie de rapport -1 ou une rotation d'angle π, c'est une similitude directe.

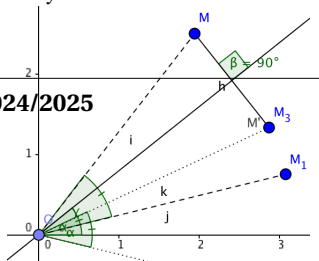
Attention - Réflexion
Une réflexion n'est pas une similitude directe

Proposition - Involution
Une symétrie est une transformation du plan : $s^{-1} = s$.

Pour aller plus loin - Symétrie axiale (orthogonale) et composition
Une réflexion peut s'écrire comme une composée de similitude directe et de conj : $z \mapsto \overline{z}$. La figure suivant donne une décomposition possible :

$$r(z) = \left[R_{0,\theta} \circ conj \circ R_{O,\theta}^{-1}\right](z)$$

où θ est l'angle entre l'axe de abscisse et l'axe de symétrie.



Démonstration
Toute involution : (s^2 = id) est nécessairement bijective car $\forall\ y, \exists\ x = s(y)$ tel que $y = s(x)$.
En effet $s(x) = s(s(y)) = s^2(y) = y$.
Toute symétrie est une involution :

- si $s_\Omega(M) = M'$ et $s_\Omega(M') = M''$, alors $\overrightarrow{\Omega M} = -\overrightarrow{\Omega M'} = \overrightarrow{\Omega M''}$, et donc $M = M'' = s_\Omega^2(M)$.
- si $s_{\mathscr{D}}(M) = M'$ et $s_{\mathscr{D}}(M') = M''$, alors $\mathscr{D}$ est la médiatrice de $[MM']$ et de $[M'M'']$,
  donc (MM') et $(M'M'')$ sont parallèles (perpendiculaire à une même troisième)
  et comme elles ont un point commun : M', ces droites sont confondues,
  donc M, M', M'' sont alignés.
  $\mathscr{D}$ étant médiatrice de $[MM']$ et de $[M'M'']$, les milieux de $[MM']$ et de $[M'M'']$ sont confondus.
  Et finalement : $M = M'' = s_{\mathscr{D}}^2(M)$

□

Pour aller plu
L'inversion est u
en géométrie pr
L'inversion par r
Ω(ω) et de rayon
M′(z′) tel que (z
Pour que Ω puis
nit sur ℂ ∪ {∞}

6. Bilan

Synthèse

- ⇝ Les relations trigonométriques permettent d'obtenir une relation algébrique simple lorsqu'on considère la fonction d'EULER à valeurs complexes $t \mapsto \cos t + i \sin t$, notée e^{it} : $e^{i(a+b)} = e^{ia} \times e^{ib}$. A partir de celle-ci, on retrouve toutes les formules (en exploitant la relation de DE MOIVRE, le binôme de NEWTON ou les fonctions linéaires Re ou Im). **A savoir-faire manipuler absolument!**
- ⇝ Très souvent la question se pose de manière réciproque : étant donné une longueur de quel angle en est-elle le cos ? Le problème, la fonction n'est pas injective : on peut avoir $\theta \neq \theta'$ et $\cos\theta = \cos\theta'$. On restreint donc l'intervalle image. On crée ainsi une fonction réciproque arc cos à la fonction $\cos_{|[0,\pi]} : [0,\pi] \to [-1,1]$. De même pour les fonctions $\sin_{|[-\frac{\pi}{2},\frac{\pi}{2}]}$ et $\tan_{|[-\frac{\pi}{2},\frac{\pi}{2}]}$.
- ⇝ On se pose la même question pour $z \mapsto z^2$ et plus largement $z \mapsto z^n$. Etant donné un nombre complexe $Z = \rho e^{i\alpha}$, il y a exactement n nombres complexes différents $z_1, \dots z_n$ tels que pour tout $k \in \mathbb{N}_n$, $(z_k)^n = Z$. Ce sont les racines n-ième de Z. Pour les obtenir, on se ramène **aux classiques racines $n$ième de l'unité** $e^{2ik\pi/n}$n en divisant par $\sqrt[n]{\rho}e^{i\alpha/n}$.
  Au passage, on trouve une méthode complémentaire (algébrique) dans le simple cas de la racine carrée d'un nombre complexe.
- ⇝ La géométrie plane (de $\mathbb{R}^2$) se code parfaitement par du calcul, sur $\mathbb{C}$. Les concepts naturels de géométrie (longueur, orthogonalité, parallélisme) se réduise exactement par du calcul, selon le rêve de Descartes ou de Leibniz.
- ⇝ En retour, le calcul complexe simple : addition, multiplication, division devient une opération géométrique : translation, similitude sans ou avec conjugaison respectivement voire inversion…

Savoir-faire et Truc & Astuce du chapitre

- Truc & Astuce pour le calcul - Factorisation de l'angle moitié
- Savoir-faire - Linéarisation
- Savoir-faire - Expression de $\cos(nt)$ et $\sin(nt)$ en fonction de $\cos t$ et $\sin t$
- Truc & Astuce pour le calcul - Racine carrée. Exploitation de la forme trigonométrique
- Truc & Astuce pour le calcul - Racine carrée. Exploitation de la forme algébrique
- Savoir-faire - Obtenir l'équation d'une droite (avec un point et un vecteur directeur ou avec deux points)

— Savoir-faire - Interprétation en terme de projection
— Savoir-faire - Obtenir l'équation d'une droite (avec un point et un vecteur normal)
— Savoir-faire - Reconnaître une similitude

Notations

Notations	*Définitions*	*Propriétés*	*Remarques*
Re, Im	Fonctions partie réelles et parties imaginaires, appliquées à un nombre complexe	Elles sont $\mathbb{R}$-linéaires ($\forall\ a_1, a_2 \in \mathbb{R}, z_1, z_2 \in \mathbb{C}$, $\mathrm{Re}(a_1 z_1 + a_2 z_2) = a\mathrm{Re}(z_1) + a_2\mathrm{Re}(z_2)\ldots$)	
$e^{i\theta}$	$e^{i\theta} := \cos\theta + i\sin\theta$ (Euler)	$e^{i(\theta+\theta')} = e^{i\theta} \times e^{i\theta'}$	Relations de de Moivre : $\cos\theta = \mathrm{Re}(e^{i\theta}) = \frac{1}{2}(e^{i\theta} + e^{-i\theta})$, $\sin\theta = \mathrm{Im}(e^{i\theta}) = \frac{1}{2i}(e^{i\theta} - e^{-i\theta})$.
j	$j := e^{2i\pi/3} = \dfrac{-1+i\sqrt{3}}{2}$	$j^3 = 1$, $1 + j + j^2 = 0$	Racine primitive troisième de l'unité (avec $\overline{j} = j^2$)
$\mathbb{U}_n$	Ensemble des racines n-ième de l'unité $\mathbb{U}_n := \{z \in \mathbb{C} \mid z^n = 1\}$	$z \in \mathbb{U}_n$ si et seulement si $\exists\,!k \in \mathbb{N}_n$ (ou $\exists\,!k \in [\![0, n-1]\!]$) tq $z = e^{2ik\pi/n}$	Reviendra souvent dans l'année.

Retour sur les problèmes

21. Multiplication des nombres.
$z \times z'$ donne (géométriquement) le nombre complexe obtenu par similitude de centre O, de longueur $|z|$, et d'angle $\arg(z)$ à partir du point $Z'(z')$.
On peut l'obtenir en appliquant le théorème de Thalès. On note I, l'intersection du cercle unité et de la droite OZ'. Puis on trace la parallèle à IZ passant par Z'. L'intersection de cette droite avec OZ donne le point $Z''(z \times z')$.
22. Le centre du triangle équilatéral sur le côté $[AB]$ a pour affixe $z_1 = \dfrac{b - e^{i\pi/6}a}{1 - e^{i\pi/6}} = \dfrac{ibe^{-i\pi/12} - iae^{i\pi/12}}{2\sin\frac{\pi}{12}}$ car $(b - z_1) = e^{i\pi/6}(a - z_1)\ldots$.
23. Transformation du plan
Tout le chapitre répond à cet exercice
24. $\cos\theta_2 = i\dfrac{\sqrt{n_1^2\sin^2\theta_1 - n_2^2}}{n_2}$. Pour le reste, voir avec M. Lagoute.

Chapitre 6

Calcul matriciel

Résumé -

Dans un premier temps nous (re)définissons les opérations matricielles : addition, multiplication par un nombre (ce qui donne un espace vectoriel) et multiplication (ce qui donne avec l'addition un anneau).
Puis nous nous concentrons sur les matrices carrées; elles jouent un rôle très important, puisque très fréquemment, on est amené à multiplier une matrice par elle-même (pour cela elle doit être carrée). C'est l'occasion de définir et de voir les propriétés des puissances de matrice, ainsi que de la trace.
On cherche également à inverser de telle matrice (si possible). Le meilleur moyen est d'utiliser un algorithme de Gauss (pivot de Gauss ou algorithme de Gauss-Jordan).

Sommaire

Dans tout le chapitre, m, n, p, q sont des entiers naturels non nuls et $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$ (on pourrait plus généralement considérer que $\mathbb{K}$ est un corps (commutatif)).

1. Problèmes

? Problème 25 - Pourquoi des matrices ?

Pour comprendre le monde qui nous entoure, on mesure des objets puis des dépendances entre ces objets.
Ces dépendances peuvent être multidimensionnelles, elles sont souvent enregistrées dans un tableau. C'est ce que l'on voit en particulier en économétrie, ou des domaines encore plus large : biologie, géographie...
Additionner des nombres a du sens lorsqu'on a des mesures de même unité d'objets de même nature. C'est seulement au XV-ième siècle que les additions se sont généralisées et se sont dégagées de leur signifiants.
Peut-on définir une addition des tableaux, qui ait du sens ? Peut-on généraliser cette addition à tous tableaux ?
Quelles sont les propriétés naturelles qui découlent : associativité, commutativité, élément neutre (Groupe) ? Voire : espace vectoriel, avec une multiplication par un scalaire.

? Problème 26 - Pourquoi un tel produit ?

On peut addition des tableaux, peut-on les multiplier ?
Donner une incarnation naturelle dans un problème physique qui justifie la règle de multiplication matricielle :

$$[AB]_{i,j} = \sum_{k=0}^{n} [A]_{i,k}[B]_{k,j}$$

? Problème 27 - Racines carrées

Puisque le produit de deux matrices existe, la puissance entière en découle : $A^2 = A \times A$ et par récurrence : $A^n = A \times A^{n-1}$.
Peut-on définir des puissances entières négatives : A^{-4} ?
Et plus largement des puissances non entières. Par exemple, la racine carrée de A serait une matrice B tel que $B^2 = A$.
Concrètement, la matrice $A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ admet-elle une racine carrée ? Plusieurs (combien ?) ?

? Problème 28 - Anneau non commutatif des matrices

L'anneau des matrices carrées de taille n : $(\mathcal{M}_n(\mathbb{K}), +, \times)$ est un exemple d'anneau non commutatif : $A \times B \neq B \times A$. Cela suggère quelques questions :

- Parmi les propriétés classiques d'anneaux, lesquelles ne sont plus vraies dans cet anneau ?
- Et inversement, existe-t-il des matrices A tel que pour tout B : $A \times B = B \times A$
- Etant donnée A, existe-t-il une condition simple, nécessaire et/ou suffisante sur B pour que $A \times B = B \times A$?

— On sait que les éléments du groupe $\mathcal{M}_n(\mathbb{K})^\times := GL_n(\mathbb{K})$ sont les matrices inversibles et ne sont nécessairement pas des diviseurs de 0.

? Problème 29 - Matrices inversibles

Existe-t-il un moyen simple, algorithmique, pour étudier l'inversibilité d'une matrice ? Peut-on l'exploiter pour obtenir également A^{-1} (dans le cas où A est inversible) ?

2. Esnsemble $\mathcal{M}_{n,p}(\mathbb{K})$

Cet ensemble est considéré comme un espace vectoriel ici.

2.1. Ensemble des matrices

Définition - Matrices

Une matrice à n lignes et p colonnes à coefficients dans $\mathbb{K}$ est une famille $(a_{ij})_{\substack{1\leqslant i\leqslant n\\1\leqslant j\leqslant p}}$ d'éléments de $\mathbb{K}$ indexée par $[\![1,n]\!]\times[\![1,p]\!]$. On parle aussi de matrice de type (n,p) ou de matrice $n\times p$. On note $\mathcal{M}_{n,p}(\mathbb{K})$ l'ensemble des matrices à n lignes et p colonnes à coefficients dans $\mathbb{K}$. Si $A\in\mathcal{M}_{n,p}(\mathbb{K})$,

$$A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1p}\\ a_{21} & a_{22} & \cdots & a_{2p}\\ \vdots & \vdots & \vdots & \vdots\\ a_{n1} & a_{n2} & \cdots & a_{np}\end{pmatrix}=(a_{ij})_{1\leqslant i\leqslant n;1\leqslant j\leqslant p}=(a_{i,j})_{\substack{1\leqslant i\leqslant n\\1\leqslant j\leqslant p}}$$

a_{ij} est le coefficient de la i-ième ligne, j-ième colonne.
Deux matrices A et B sont donc égales si elles ont même nombre de lignes, même nombre de colonnes et mêmes coefficients.
Si $n=p$ on note $\mathcal{M}_n(\mathbb{K})$ l'ensemble des matrices carrées d'ordre n.

Exemple - Premier exemple

$$(i-j)_{\substack{1\leqslant i\leqslant 2\\1\leqslant j\leqslant 3}}=\begin{pmatrix}0 & -1 & -2\\ 1 & 0 & -1\end{pmatrix}$$

Définition - Quelques cas particuliers

Quelques matrices « de référence » sont à connaître :

— La matrice nulle de $\mathcal{M}_{n,p}(\mathbb{K})$ est la matrice qui ne contient que des coefficients nuls :

$$(0)_{\substack{1\leqslant i\leqslant n\\1\leqslant j\leqslant p}}=(0)=O_{n,p}$$

— si $n=1$, on dit que $A=\begin{pmatrix}a_1 & \dots & a_p\end{pmatrix}$ est une matrice ligne.

— si $p=1$, $B=\begin{pmatrix}b_1\\ \vdots\\ b_n\end{pmatrix}$ est appelée matrice colonne.

Parmi les matrices carrées d'ordre n, quelques matrices joueront un rôle particulier :

— La matrice identité de $\mathcal{M}_n(\mathbb{K})$ est la matrice I_n qui possède des 1 sur

la diagonale et des 0 en dehors de la diagonale :

$$I_n = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} = (a_{ij})_{\substack{1\leqslant i\leqslant n \\ 1\leqslant j\leqslant n}} = (a_{ij})_{1\leqslant i,j\leqslant n}$$

où $a_{ij} = 0$ si $i \neq j$ et $a_{ii} = 1$ pour $i = 1$ à n.

— une matrice diagonale est une matrice carrée dont seuls les éléments diagonaux sont non nuls :

$$\text{diag}(a_{11},\ldots,a_{nn}) = \begin{pmatrix} a_{11} & 0 & \cdots & \cdots & 0 \\ 0 & a_{22} & 0 & \cdots & 0 \\ 0 & 0 & a_{33} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & a_{nn} \end{pmatrix} \quad (i \neq j \Rightarrow a_{ij} = 0)$$

Ainsi $I_n = \text{diag}(\underbrace{1,1,\ldots 1}_{n})$.

— une matrice scalaire est une matrice diagonale dont tous les éléments sont identiques :

$$\begin{pmatrix} \lambda & 0 & \cdots & \cdots & 0 \\ 0 & \lambda & 0 & \cdots & 0 \\ 0 & 0 & \lambda & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda \end{pmatrix} \quad (i \neq j \Rightarrow a_{ij} = 0 \text{ et } a_{ii} = a_{jj} = \lambda)$$

— une matrice triangulaire supérieure est une matrice carrée dont les éléments au-dessous de la diagonale sont nuls :

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ 0 & a_{22} & a_{23} & \cdots & a_{2n} \\ 0 & 0 & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & a_{nn} \end{pmatrix} \quad (i > j \Rightarrow a_{ij} = 0);$$

— une matrice triangulaire inférieure est une matrice carrée dont les éléments au-dessus de la diagonale sont nuls.

$$\begin{pmatrix} a_{11} & 0 & \cdots & \cdots & 0 \\ a_{21} & a_{22} & 0 & \cdots & 0 \\ a_{31} & a_{32} & a_{33} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nn} \end{pmatrix} \quad (i < j \Rightarrow a_{ij} = 0);$$

2.2. Opérations (vectorielles) sur les matrices

Addition

Définition - Addition de deux matrices de même taille
La somme de deux matrices $A = (a_{ij})_{\substack{1\leqslant i\leqslant n \\ 1\leqslant j\leqslant p}}$ et $B = (b_{ij})_{\substack{1\leqslant i\leqslant n \\ 1\leqslant j\leqslant p}}$ de $\mathscr{M}_{n,p}(\mathbb{K})$

est la matrice définie par la formule suivante :

$$A+B=(a_{ij}+b_{ij})_{\substack{1\leqslant i\leqslant n\\1\leqslant j\leqslant p}},$$

on ajoute les coefficients qui ont la même position.
Il s'agit d'une loi interne sur $\mathcal{M}_{n,p}(\mathbb{K})$.

Application - Exemple

$$\begin{pmatrix}1&2&3\\4&5&6\end{pmatrix}+\begin{pmatrix}2&3&4\\5&6&7\end{pmatrix}=\begin{pmatrix}3&5&7\\9&11&13\end{pmatrix}.$$

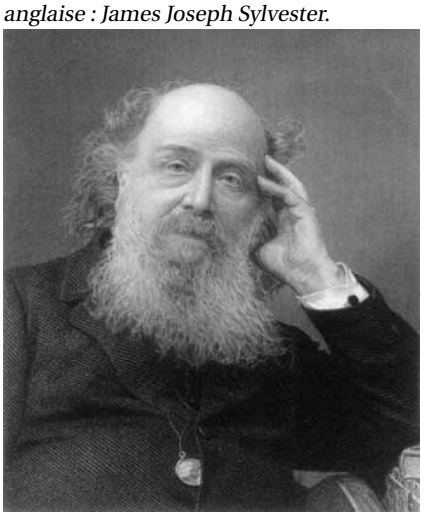
Analyse - Groupe $(\mathcal{M}_{n,p}(\mathbb{K}),+)$
On remarque que pour A, B, $C\in\mathcal{M}_{n,p}(\mathbb{K})$,

- $A+(0)_{\substack{1\leqslant i\leqslant n\\1\leqslant j\leqslant p}}=(0)_{\substack{1\leqslant i\leqslant n\\1\leqslant j\leqslant p}}+A=A.$
- $(A+B)+C=A+(B+C)$;
- $A+B=B+A.$
- $(a_{ij})_{\substack{1\leqslant i\leqslant n\\1\leqslant j\leqslant p}}+(-a_{ij})_{\substack{1\leqslant i\leqslant n\\1\leqslant j\leqslant p}}=(0)_{\substack{1\leqslant i\leqslant n\\1\leqslant j\leqslant p}}$;

Le théorème suivant en découle :

Théorème - Le groupe $(\mathcal{M}_{n,p}(\mathbb{K}),+)$
L'ensemble $\mathcal{M}_{n,p}(\mathbb{K})$ muni de l'addition + est donc un groupe commutatif, d'élément neutre la matrice nulle de $\mathcal{M}_{n,p}(\mathbb{K})$.

Histoire - Le terme de matrice
Le terme de matrice pour désigner les tableaux rectangulaires considérées ici fut introduit en 1850 par le mathématicien américain d'origine anglaise : James Joseph Sylvester.



James Joseph Sylvester est né 1814 et mourut en 1897. Son origine juive l'obligea à aller enseigner en Amérique.

Multiplication par un scalaire

Définition - Multiplication par un scalaire
Le produit d'une matrice A de $\mathcal{M}_{n,p}(\mathbb{K})$ par $\alpha\in\mathbb{K}$ est la matrice notée αA définie par :

$$\alpha(a_{ij})_{\substack{1\leqslant i\leqslant n\\1\leqslant j\leqslant p}}=(\alpha a_{ij})_{\substack{1\leqslant i\leqslant n\\1\leqslant j\leqslant p}}.$$

On définit ainsi une loi externe sur $\mathcal{M}_{n,p}(\mathbb{K})$ à domaine d'opérateur $\mathbb{K}$

Et on vérifie facilement les propriétés suivantes :

Proposition - Propriétés de la multiplication scalaire

- $1A=A$
- $\alpha(\beta A)=(\alpha\beta)A$
- $(\alpha+\beta)A=\alpha A+\beta A,$
- $\alpha(A+B)=\alpha A+\alpha B.$

Espace vectoriel de dimension finie

Analyse - Pour définir explicitement, sans quiproquo, une matrice, il faut…
$n\times p$ coefficients de $\mathbb{K}$ qu'il faut placer aux différentes positions de la matrice.
Il s'agit donc d'un espace vectoriel de dimension $n\times p$, et dont la base *canonique* est donnée par la référence de position.
Chaque élément de la base correspond à une position. Chaque élément de la base a donc un double indice : k et ℓ.

Théorème - L'espace vectoriel $(\mathcal{M}_{n,p}(\mathbb{K}),+,.)$
$(\mathcal{M}_{n,p}(\mathbb{K}),+,.)$ est un $\mathbb{K}$-e.v de dimension np.
La base canonique est formée par les $n\times p$ matrices $E_{k\ell}$ $(1\leqslant k\leqslant n; 1\leqslant \ell\leqslant p)$ où $E_{k\ell}$ est la matrice ne contenant que des 0 sauf l'élément d'indices k,ℓ qui vaut 1, soit

$$E_{k\ell}=(\delta_{ki}\delta_{j\ell})_{\substack{1\leqslant i\leqslant n\\ 1\leqslant j\leqslant p}}$$

On a donc $\dim_{\mathbb{K}}\mathcal{M}_{n,p}(\mathbb{K})=n\times p$.

Savoir faire - Notations

Par la suite, on notera ${}^{i}[A]_{j}$ ou $\text{Coef}_{i,j}(A)$, le coefficient en ligne i et colonne j de la matrice A.
On a donc

$${}^{i}[\lambda A+\mu B]_{j}=\lambda\,{}^{i}[A]_{j}+\mu\,{}^{i}[B]_{j}\qquad \text{Coef}_{i,j}(\lambda A+\mu B)=\lambda\text{Coef}_{i,j}(A)+\mu\text{Coef}_{i,j}(B)$$

$$\forall\, i,j,\quad {}^{i}[\cdot]_{j}\text{ ou Coef}_{i,j}\text{ est une application linéaire de }\mathcal{M}_{n,p}(\mathbb{K})$$

On notera également $L_i(A)$ (respectivement $C_j(A)$), la ligne i (respectivement colonne j de A).
On note que ${}^{i}[AB]_{j}=L_i(A)\times C_j(B)$, c'est un nombre.

2.3. Transposition

Définition - Matrice transposée
Soit $A=(a_{ij})_{1\leqslant i\leqslant n,1\leqslant j\leqslant p}\in\mathcal{M}_{n,p}(\mathbb{K})$,
on définit la **transposée** de A, notée ${}^{t}A$ ou A^{T} par

$$\forall\, i\leqslant p, j\leqslant n:\qquad {}^{i}[A^{T}]_{j}={}^{i}[{}^{t}A]_{j}={}^{j}[A]_{i}$$

On a $A^{T}\in\mathcal{M}_{p,n}(\mathbb{K})$.

La transposée d'une matrice s'obtient en "échangeant" lignes et colonnes :

Exemple - Matrice 3×3

$$\text{Si } A=\begin{pmatrix}1&-1&-1&-3\\0&1&3&1\\-1&1&1&3\end{pmatrix}\text{ alors }{}^{t}A=\begin{pmatrix}1&0&-1\\-1&1&1\\-1&3&1\\-3&1&3\end{pmatrix}$$

Théorème - Isomorphisme
Sous réserve que la taille des matrices permette d'effectuer les différentes opérations, on a :

$$(A+B)^{T}=A^{T}+B^{T};\quad (\lambda A)^{T}=\lambda A^{T};\quad (A^{T})^{T}=A$$

La transposition est donc un isomorphisme entre les espaces vectoriels $\mathcal{M}_{n,p}(\mathbb{K})$ et $\mathcal{M}_{p,n}(\mathbb{K})$.

Exercice
Faire la démonstration

Correction
Jeu d'écriture

3. Multiplication matricielle

3.1. Définition

Définition - Produit de deux matrices
Le produit d'une matrice $A = (a_{ij})_{\substack{1\leqslant i\leqslant n\\1\leqslant j\leqslant p}}$ de $\mathcal{M}_{n,p}(\mathbb{K})$ par une matrice $B = (b_{ij})_{\substack{1\leqslant i\leqslant p\\1\leqslant j\leqslant q}}$ de $\mathcal{M}_{p,q}(\mathbb{K})$ est une matrice de $\mathcal{M}_{n,q}(\mathbb{K})$ définie par

$$C = AB = (c_{ij})_{\substack{1\leqslant i\leqslant n\\1\leqslant j\leqslant q}}$$

où

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \ldots + a_{ip}b_{pj} = \sum_{k=1}^{p} a_{ik}b_{kj}.$$

Savoir faire - Notation et multiplication matricielle
Par la suite, on notera $\text{Coef}_{i,j}(A)$, le coefficient en ligne i et colonne j de la matrice A.
On a donc

$$^{i}[AB]_{j} = \sum_{k=1}^{p} {}^{i}[A]_{k}\,{}^{k}[B]_{j}$$

Il faut savoir passer d'un sens vers l'autre : $\text{Coef}_{i,j}(AB) = \sum_{k=1}^{p} \text{Coef}_{i,k}(A)\text{Coef}_{k,j}(B)$

et aussi $\sum_{k=1}^{p} \text{Coef}_{i,k}(A)\text{Coef}_{k,j}(B) = \text{Coef}_{i,j}(AB)$.

Pour aller plus loin - La convention d'Einstein
La convention d'Einstein en physique consiste à voir dans toute répétition de deux lettres muettes une somme. Ainsi le symbole $\sum$ peut être enlevé.
Le fait qu'une telle convention existe signifie la fréquence importante des opérations du type

$$\sum_{k=1}^{n} a_{i,k}b_{k,j}$$

écrit par Einstein : $a_{i,k}b_{k,j}\ldots$

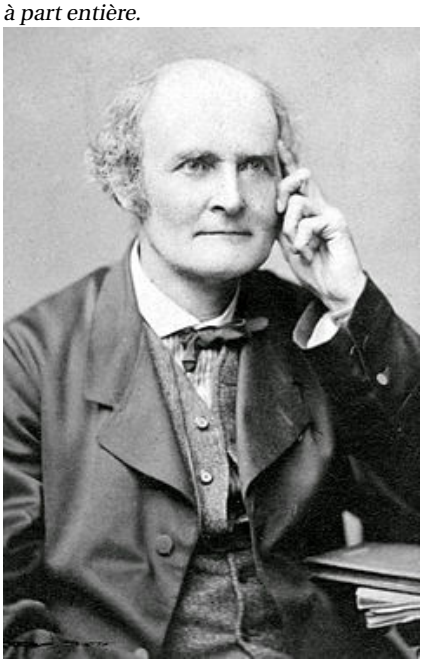
Attention - Taille des matrices
On ne peut pas multiplier une matrice de $\mathcal{M}_{3,4}(\mathbb{K})$ avec une matrice de $\mathcal{M}_{5,6}(\mathbb{K})$! Il faut que le nombre de colonnes de la première matrice soit égal au nombre de lignes de la seconde.

Savoir faire - Présentation des calculs
Une méthode pratique de présentation des calculs :

$$\begin{pmatrix} \ldots & \ldots & b_{1j} & \ldots \\ \ldots & \ldots & b_{2j} & \ldots \\ \vdots & \vdots & \vdots & \vdots \\ \ldots & \ldots & b_{pj} & \ldots \end{pmatrix}$$

$$\begin{pmatrix} \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots \\ a_{i1} & a_{i2} & \ldots & a_{ip} \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots \end{pmatrix} \begin{pmatrix} \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots \\ \ldots & \ldots & c_{ij} & \ldots \\ \vdots & \vdots & & \vdots \\ \vdots & \vdots & \vdots & \vdots \end{pmatrix}$$

Histoire - Arthur Cayley
C'est Arthur Cayley qui définit le produit matriciel, dans le premier article (1855) qui étudie les matrices comme objets mathématiques à part entière.



Arthur Cayley est un brillant avocat puis mathématicien anglais né en 1821 et mort en 1895. C'est un génie très hétéroclite.

Application - Produit de deux matrices

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \quad \begin{pmatrix} 1+6+15 & 2+8+18 \\ 4+15+30 & 8+20+36 \end{pmatrix} \quad = \quad \begin{pmatrix} 22 & 28 \\ 49 & 64 \end{pmatrix}$$

Exemple - Petits calculs

Soient $A = \begin{pmatrix} 2 & -3 & -4 \\ 3 & 1 & 5 \end{pmatrix}$ et $B = \begin{pmatrix} 3 & -3 & 2 \\ -1 & 5 & -2 \\ -1 & 3 & 0 \end{pmatrix}$. Calculer, si cela est possible, AB, BA, A^2, B^2.

Exercice

Simplifier le produit :

$$\sum_{h=1}^{n}\sum_{\ell=1}^{n}\sum_{j=1}^{n} a_{\ell,j}b_{i,h}c_{h,\ell}d_{j,m}$$

Correction

Il est évident que le résultat dépend de i et de m et d'aucun autre nombre.

Les nombres $_\ell[A]^j\ldots$ sont des nombres réels, on peut donc les faire commuter.

$$\sum_{h=1}^{n}\sum_{\ell=1}^{n}\sum_{j=1}^{n} a_{\ell,j}b_{i,h}c_{h,\ell}d_{j,m} = \sum_{h=1}^{n}\sum_{\ell=1}^{n}\sum_{j=1}^{n} b_{i,h}c_{h,\ell}a_{\ell,j}d_{j,m} = \text{Coef}_{i,m}(BCAD)$$

Il s'agit du nombre en ligne i et colonne m de la matrice $B \times C \times A \times D$.

3.2. Interprétation en terme de systèmes linéaires

Analyse - Multiplication par une matrice colonne

Si $A = (a_{ij})_{1\leqslant i\leqslant n,\, 1\leqslant j\leqslant p}$, et $B = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$, $X = \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix}$, alors

$$AX = \begin{pmatrix} a_{11} & \ldots & \ldots & a_{1p} \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & \ldots & \ldots & a_{np} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ \vdots \\ x_p \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + \ldots + a_{1p}x_p \\ \vdots \\ \vdots \\ a_{n1}x_1 + \ldots + a_{np}x_p \end{pmatrix}$$

Proposition - $(S) \Leftrightarrow AX = B$

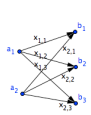
L'équation $AX = B$ pour des matrices est une manière compacte d'écrire un système linéaire général avec n équations, p inconnues et un second

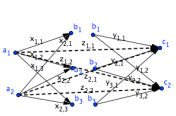
Pour aller plus loin - Interprétation en terme de graphes

Un dessin permettra d'expliquer au mieux ce à quoi peut servir une matrice. Il faut d'abord considérer :

1. *un ensemble de départ; par exemple : $A = \{a_1, a_2\}$*
2. *un ensemble d'arrivée; par exemple : $B = \{b_1, b_2, b_3\}$*
3. *un jeu de flèches entre les deux ensembles, associés à des nombres.*
   Par exemple, la flèche $a_i \to b_j = {}_iX^j$ indique la relation entre a_i et b_j.

C'est cette flèche qui est abstraite, la plus ouverte possible. Elle peut être par exemple un temps de trajet, un coefficient de proportionalité... Ainsi la figure 1 suivante est représentée par la matrice $\begin{pmatrix} x_{1,1} & x_{1,2} & x_{1,3} \\ x_{2,1} & x_{2,2} & x_{2,3} \end{pmatrix}$.





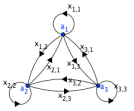


FIG.1-GRAPHE - FIG.2-PRODUIT -FIG.3- CARRÉE MATRICIEL

La figure 3 représente un matrice d'un ensemble dans lui-même (comme pour les endomorphismes), elle est nécessairement carrée.

La figure 2 représente alors le produit matriciel. En effet, il s'agit de savoir comment aller alors de $A = \{a_1, a_2\}$ à $C = \{c_1, c_2\}$. La réponse est obtenue sous forme de matrice $z_{i,j}$, où $z_{i,j}$ indique les chemins de a_i à c_j. Il y a en fait trois possibilités : passer par b_1, b_2 ou b_3, cela donne donc exactement :

$$\begin{aligned} z_{i,j} &= x_{i,1}y_{1,j} + x_{i,2}y_{2,j} + x_{i,3}y_{3,j} \\ &= \sum_{k=1}^{3} x_{i,k}y_{k,j} \\ &= \sum_{k=1}^{3} Coef_{i,k}(X)Coef_{k,j}(Y) \end{aligned}$$

membre $B = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$

$$(S)\begin{cases} a_{11}x_1 & +\ldots & +a_{1p}x_p & = b_1 \\ a_{21}x_1 & +\ldots & +a_{2p}x_p & = b_2 \\ \vdots & + & +\vdots & = \vdots \\ a_{n1}x_1 & +\ldots & +a_{np}x_p & = b_n. \end{cases}$$

Nous reviendrons sur ce parallèle lorsque nous prendrons le temps de résoudre des systèmes linéaires.

3.3. Propriété du produit

Proposition - Associativité du produit
Le produit de matrices est associatif.
Plus précisément si $A \in \mathcal{M}_{n,p}(\mathbb{K})$, $B \in \mathcal{M}_{p,q}(\mathbb{K})$, $C \in \mathcal{M}_{q,m}(\mathbb{K})$ alors on a

$$(AB)C = A(BC)$$

qui est une matrice de $\mathcal{M}_{n,m}(\mathbb{K})$.

Démonstration
Le seul problème réside dans la manipulation des indices.
Appelons $D = (d_{ih})$ la matrice AB; le coefficient de la $i$ème ligne et de la $h$ème colonne est donné par

$$d_{ih} = \sum_{k=1}^{n} a_{ik}b_{kh},$$

donc le coefficient $i,\ j$ de $(AB)C = E$ est donné par

$$\sum_{h=1}^{p} d_{ih}c_{hj} = \sum_{h=1}^{p}\left(\sum_{k=1}^{n} a_{ik}b_{kh}\right)c_{hj}$$

$$= \sum_{h=1}^{p}\sum_{k=1}^{n} a_{ik}b_{kh}c_{hj} = e_{ij}.$$

Calculons maintenant le coefficient $k,\ j$ de $D' = BC = (d'_{kj})$

$$d'_{kj} = \sum_{h=1}^{p} b_{kh}c_{hj},$$

le coefficient $i,\ j$ de $E' = A(BC) = (e'_{ij})$ est obtenu en faisant :

$$e'_{ij} = \sum_{k=1}^{n} a_{ik}d'_{kj} = \sum_{k=1}^{n} a_{ik}\left(\sum_{h=1}^{p} b_{kh}c_{hj}\right)$$

$$= \sum_{k=1}^{n}\sum_{h=1}^{p} a_{ik}b_{kh}c_{hj}$$

d'où l'égalité que l'on appelle associativité. □

Remarque - En terme de Coef$_{i,j}$
Ce que l'on a démontré c'est :

$$^{i}[(AB)C]_{j} = \sum_{h,k}^{i} [A]_{h}^{h}[B]_{k}^{k}[C]_{j} = {}^{i}[A(BC)]_{j}$$

Proposition - Bilinéarité
Si A et B sont des matrices de $\mathcal{M}_{n,p}(\mathbb{K})$ et $C,\ D$ de $\mathcal{M}_{p,q}(\mathbb{K})$, $\lambda, \mu \in \mathbb{K}$ alors

$$A(\lambda C + \mu D) = \lambda AC + \mu AD \quad \text{et} \quad (\lambda A + \mu B)C = \lambda AC + \mu BC$$

En résumé l'application $(A, C) \in \mathcal{M}_{n,p}(\mathbb{K}) \times \mathcal{M}_{p,q}(\mathbb{K}) \mapsto AC \in \mathcal{M}_{n,q}(\mathbb{K})$ est bilinéaire.

Pour aller plus loin - Application aux probabilités
Ecrire : « $z_{i,j}$ indique les chemins de a_i à c_j. Il y a en fait trois possibilités : passer par b_1, b_2 ou b_3 » nous rappelle la formule des probabilités totales.
En effet, on se souvient que si (B_1, B_2, B_3) forme un système complet d'événements, alors :

$$\mathbf{P}(A_i) = \mathbf{P}(B_1) \times \mathbf{P}_{B_1}(A_i) + \mathbf{P}(B_2) \times \mathbf{P}_{B_2}(A_i)$$

$$+\mathbf{P}(B_3) \times \mathbf{P}_{B_3}(A_i)$$

Ce se résume en

$$(A+B)(C+D) = AC + AD + BC + BD$$

Démonstration
On ne fera qu'un calcul.

$$\mathrm{Coef}_{i,j}\big(A(\lambda C + \mu D)\big) = \lambda \sum_{h=1}^{p} \mathrm{Coef}_{i,h}(A)\mathrm{Coef}_{h,j}(C) + \mu \sum_{h=1}^{p} \mathrm{Coef}_{i,h}(A)\mathrm{Coef}_{h,j}(D)$$

(...) □

Proposition - Cas à connaître!
$(E_{i,j})_{\substack{1\leqslant i\leqslant n\\ 1\leqslant j\leqslant p}}$ désigne la base canonique de $\mathscr{M}_{n,p}(\mathbb{K})$ i.e. ${}^{a}[E_{i,j}]_b = \delta_{a,i}\delta_{b,j}$
et $(F_{i,j})_{\substack{1\leqslant i\leqslant p\\ 1\leqslant j\leqslant q}}$ celle de $\mathscr{M}_{p,q}(\mathbb{K})$.
Alors $E_{i,j} \times F_{k,\ell} = \delta_{k,j} G_{i,\ell}$, avec $(G_{s,t})_{\substack{1\leqslant s\leqslant n\\ 1\leqslant r\leqslant q}}$ base canonique de $\mathscr{M}_{n,q}(\mathbb{K})$

Démonstration
Il s'agit de calculer $E_{i,j}F_{k,\ell}$ pour les différents quadruplets possibles (i,j,k,ℓ).
Résultat à retenir (ou à savoir retrouver rapidement). Il faut quatre lettres muettes (2 pour E et 2 pour F), et deux lettres (provisoires) pour préciser la position :

$$\mathrm{Coef}_{a,b}(E_{i,j}F_{k,\ell}) = \sum_{h=1}^{p} \mathrm{Coef}_{a,h}(E_{i,j})\mathrm{Coef}_{h,b}(F_{k,\ell}) = \sum_{h=1}^{p} \delta_{a,i}\delta_{h,j}\delta_{h,k}\delta_{b,\ell} = \begin{cases} 0 & \text{si } k \neq j \\ \delta_{a,i}\delta_{b,\ell} & \text{si } k = j \end{cases}$$

Donc si $a \neq i$ ou $b \neq \ell$, $\mathrm{Coef}_{a,b}(E_{i,j}F_{k,\ell}) = 0$.
Et si $k = j$, alors pour $a = i$, $b = \ell$, on a $\mathrm{Coef}_{a,b}(E_{i,j}F_{k,\ell}) = 1$.
Donc $E_{i,j} \times F_{k,\ell} = \delta_{k,j} G_{i,\ell}$, avec $(G_{s,t})_{\substack{1\leqslant s\leqslant n\\ 1\leqslant r\leqslant q}}$ base canonique de $\mathscr{M}_{n,q}(\mathbb{K})$. □

Pour aller plus loin - Transposition et graphe
La matrice transposée est la matrice associée au graphe où le sens des flèches s'inverse par rapport à la situation initiale

Exercice
Comment écrire la matrice $AE_{i,j}$ à partir de la matrice A ?
De même pour $E_{i,j}A$?

Correction
Pour tout k, h,

$${}^{k}[AE_{i,j}]_h = \sum_{s=1}^{n} {}^{k}[A]_s \delta_{i,s}\delta_{j,h} = \delta_{j,h}\,{}^{k}[A]_i$$

Donc si $h \neq 0$, on trouve le nombre nul et si $h = j$, on obtient le nombre ${}^{k}[A]_i$, situé précédemment en colonne i de A.
On obtient donc une matrice formée d'une unique colonne non nulle, l'ancienne colonne i de A, située en colonne j.
De même $E_{i,j}A$ est la matrice formée d'une unique ligne non nulle, l'ancienne ligne j de A, située en colonne i.
On peut aussi exploiter la formule précédente et la linéarité du produit (et distribution)

Proposition - Transposition d'un produit
Pour $A \in \mathscr{M}_{n,p}(\mathbb{K})$ et $B \in \mathscr{M}_{p,q}(\mathbb{K})$, on a

$${}^{t}(A \times B) = {}^{t}B \times {}^{t}A$$

Attention - Eviter d'écrire des bêtises
Notons bien que ${}^{t}B \in \mathscr{M}_{q,p}(\mathbb{K})$ et ${}^{t}A \in \mathscr{M}_{p,n}(\mathbb{K})$.
Donc le produit ${}^{t}A \times {}^{t}B$ n'aurait aucun sens (aucune raison que $n = q$.)

Démonstration
Pour tout $(i,j) \in \mathbb{N}_n \times N_q$,

$$\begin{aligned}\text{Coef}_{j,i}\left({}^t(A\times B)\right) = \text{Coef}_{i,j}\left((A\times B)\right) &= \sum_{h=1}^{p} \text{Coef}_{i,h}(A)\times \text{Coef}_{h,j}(B)\\ = \sum_{h=1}^{p} \text{Coef}_{h,i}({}^tA)\times \text{Coef}_{j,h}({}^tB) &= \sum_{h=1}^{p} \text{Coef}_{j,h}({}^tB)\times \text{Coef}_{h,i}({}^tA)\\ &= \text{Coef}_{j,i}({}^tB\times {}^tA)\end{aligned}$$

□

3.4. Produit par blocs

Proposition - Produit par blocs
Soient deux matrices $M \in \mathcal{M}_{n,p}(\mathbb{K})$, $M' \in \mathcal{M}_{p,q}(\mathbb{K})$.
Considérons des entiers $k \leqslant n, \ell \leqslant p, m \leqslant q$ et des matrices $A \in \mathcal{M}_{k,\ell}(\mathbb{K})$, $B \in \mathcal{M}_{k,p-\ell}(\mathbb{K})$, $C \in \mathcal{M}_{n-k,\ell}(\mathbb{K})$, $D \in \mathcal{M}_{n-k,p-\ell}(\mathbb{K})$, $A' \in \mathcal{M}_{\ell,m}(\mathbb{K})$, $B' \in \mathcal{M}_{\ell,q-m}(\mathbb{K})$, $C' \in \mathcal{M}_{p-\ell,m}(\mathbb{K})$, $D' \in \mathcal{M}_{p-\ell,q-m}(\mathbb{K})$ telles que M et N s'écrivent par blocs

$$M = \begin{pmatrix} A & B \\ C & D \end{pmatrix}, \quad M' = \begin{pmatrix} A' & B' \\ C' & D' \end{pmatrix}$$

Alors, on peut calculer le produit MM' par blocs de la manière suivante :

$$MM' = \begin{pmatrix} AA'+BC' & AB'+BD' \\ CA'+DC' & CB'+DD' \end{pmatrix}$$

Attention - Bien faire attention aux dimensions
Il est nécessaire que les dimensions correspondent bien.
Sinon, le calcul écrit n'aurait pas de sens

Démonstration
On calcule $\text{Coef}_{i,j}(MM')$ suivant la position de i par rapport à k et celle de j par rapport à m.
On va faire un cas : $i \leqslant k$ et $j \geqslant m+1$. Donc $\text{Coef}_{i,j}(M\times M')$ se trouve en haut à droite.

$$\begin{aligned}\text{Coef}_{i,j}(MM') &= \sum_{h=1}^{p} \text{Coef}_{i,h}(M)\text{Coef}_{h,j}(M')\\ &= \sum_{h=1}^{\ell} \text{Coef}_{i,h}(M)\text{Coef}_{h,j}(M') + \sum_{h=\ell+1}^{p} \text{Coef}_{i,h}(M)\text{Coef}_{h,j}(M')\\ &= \sum_{h=1}^{\ell} \text{Coef}_{i,h}(A)\text{Coef}_{h,j-m}(B') + \sum_{h=\ell+1}^{p} \text{Coef}_{i,h-\ell}(B)\text{Coef}_{h-\ell,j-m}(D')\end{aligned}$$

Car pour
- $i \leqslant k$, $h \leqslant \ell$, $\text{Coef}_{i,h}(M) = \text{Coef}_{i,h}(A)$,
- $i \leqslant k$, $h \geqslant \ell+1$, $\text{Coef}_{i,h}(M) = \text{Coef}_{i,h-\ell}(B)$,
- $h \leqslant \ell$, $j \geqslant m+1$, $\text{Coef}_{h,j}(M') = \text{Coef}_{h,j-m}(B')$,
- $h \leqslant \ell$, $j \leqslant m+1$, $\text{Coef}_{h,j}(M') = \text{Coef}_{h-\ell,j-m}(D')$,

Donc

$$\text{Coef}_{i,j}(MM') = \text{Coef}_{i,j-m}(AB') + \text{Coef}_{i,j-m}(BD') = \text{Coef}_{i,j-m}(AB'+BD')$$

□

On peut avoir intérêt à considérer les matrices sous forme d'une association de colonnes ou de lignes.
On peut voir ces opérations, comme une forme de calculs parallèles, à la physicienne.

Proposition - Matrice et association de colonnes ou de lignes
Soient $A, B \in \mathcal{M}_n(\mathbb{K})$.
Il nous arrivera de noter $A = \begin{pmatrix} \frac{L_1}{L_2} \\ \hline \vdots \\ \hline L_n \end{pmatrix}$
comme une association de n (matrices ou vecteurs) lignes, avec $L_i(A) = L_i$
Il nous arrivera de noter $B = (C_1|C_2|\dots|C_n)$
comme une association de n (matrices ou vecteurs) colonnes, avec $C_j(B) = C_j$
On a alors $AB = (AC_1|AC_2|\dots|AC_n) = \begin{pmatrix} \frac{L_1B}{L_2B} \\ \hline \vdots \\ \hline L_nB \end{pmatrix}$ mais aussi ${}^i[A \times B]_j = L_i \times C_j$.

Exercice
Comment écrire la matrice $E_{i,j}A$ à partir de la matrice A et $AE_{i,j}$, en raisonnant par blocs ?

Correction
$L_k(E_{i,j}A) = L_k(E_{i,j})A = \delta_{k,i}L_i(E_{i,j})A = \delta_{k,i}(L_i(E_{i,j}C_1(A)|\cdots|L_n(E_{i,j}C_1(A)) = \delta_{k,i}({}^j[A]_1|{}^j[A]_2|\cdots{}^j[A]_n) = \delta_{k,i}L_j(A)$

4. Les matrices carrées

4.1. L'anneau $(\mathcal{M}_n(\mathbb{K}), +, \times)$

Pour aller plus loin - Algèbre ?
On appelle $\mathbb{K}$-algèbre un ensemble $\mathcal{A}$ muni de deux opérations interne (notées ici + et $\times$) et une opération externe (notée ici $\cdot$) telle que $(\mathcal{A}, +, \cdot)$ est un $\mathbb{K}$-espace vectoriel et $(\mathcal{A}, +, \times)$ est un anneau.

Pour aller plus loin - Algèbre
$(\mathcal{M}_n(\mathbb{K}), +, \times, .)$ est une $\mathbb{K}$-algèbre de dimension n^2.

Heuristique - Pourquoi les matrices carrées ?
Si on multiplie deux matrices de $\mathcal{M}_n(\mathbb{K})$ on trouve un élément de $\mathcal{M}_n(\mathbb{K})$, la multiplication est donc interne dans $\mathcal{M}_n(\mathbb{K})$.
On peut ainsi effectuer les calculs $A \times B$ et $B \times A$, mais aussi A^k pour tout entier k…
Nous verrons qu'ains $(\mathcal{M}_n(\mathbb{K}), +, \times)$ est un anneau.

Théorème - La $\mathbb{K}$-algèbre $(\mathcal{M}_n(\mathbb{K}), +, \times, .)$
$(\mathcal{M}_n(\mathbb{K}), +, \times)$ est un anneau non commutatif et non intègre dès que $n \geqslant 2$, d'élément unité I_n.

Exemple - Non commutativité et non intégrité
Vérifiez que
$$\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \neq \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$$
puis que
$$\begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 2 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Exemple - L'inverse d'une matrice d'ordre 2
Soit $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$.
Alors en notant $A' = \begin{pmatrix} d & -b \\ c & a \end{pmatrix}$, on a :
$$A \times A' = A' \times A = (ad - bc)I_2$$
Donc si $ad - bc \neq 0$, A est inversible et $A^{-1} = \frac{1}{ad-bc}A'$.

4.2. Puissance de matrices

Définition - Puissance d'une matrice
Soit $A \in \mathcal{M}_n(\mathbb{K})$, on définit par récurrence :

$$A^0 = I_n, \qquad \forall\, k \in \mathbb{N}, A^{k+1} = A \times A^k$$

On a alors, par commutation :

$$A^k = \underbrace{A \times \ldots \times A}_{k \text{ fois}} = A^m \times A^{k-m} \qquad \text{(pour tout } m \leqslant k\text{)}$$

Les règles de calcul dans un anneau (comme dans $\mathbb{R}$ ou $\mathbb{C}$) s'appliquent d'où :

Proposition - Formules matricielles
Pour $A, B \in \mathcal{M}_n(\mathbb{K})$, si $AB = BA$ alors

$$(A+B)^p = \sum_{k=0}^{p} \binom{p}{k} A^k B^{p-k} \text{ (Formule du binôme de Newton)}$$

$$A^p - B^p = (A-B)(A^{p-1} + A^{p-2}B + \cdots + AB^{p-2} + B^{p-1})$$

$$I_n - A^p = (I_n - A)(I_n + A + A^2 + \cdots + A^{p-1})$$

4.3. Inversibilité d'une matrice

Définition - Inversibilité de A
On dit qu'une **matrice carré d'ordre** n A est inversible, si elle admet un inverse pour la loi $\times$, c'est-à-dire s'il existe une matrice carrée d'ordre n B telle que

$$BA = AB = I_n$$

(où I_n est la matrice identité). B est alors notée A^{-1} et appelée inverse de A.

Remarque - Des matrices non inversibles
On ne peut pas inverser de matrices de $\mathcal{M}_{n,p}(\mathbb{K})$ si $n \neq p$.
Une matrice carrée n'est pas nécessairement inversible :
par exemple, la matrice nulle O_n n'a pas d'inverse car

$$\forall A \in \mathcal{M}_n(\mathbb{K}),\ AO_n = O_n \neq I.$$

Exemple - Matrice non inversible, moins triviale
De plus il y a des matrices non nulles qui n'ont pas d'inverse.
Par exemple $N = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$ n'a pas d'inverse.
En effet, pour $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$,

$$A \times N = \begin{pmatrix} b & 0 \\ d & 0 \end{pmatrix}$$

Elle ne peut pas être égale à I_2.

Remarque - Rappels
Nous avons vu que l'ensemble des inverses d'un anneau forme un groupe.
On l'avait noté $A^\times$.

Définition - Le groupe $GL_n(\mathbb{K})$
L'ensemble $(GL_n(\mathbb{K}), \times)$ des inversibles de l'anneau $\mathcal{M}_n(\mathbb{K})$ est un groupe, non commutatif.
On l'appelle le groupe linéaire.
On a donc pour $A, B \in GL_n(\mathbb{K})$, $(AB)^{-1} = B^{-1}A^{-1}$.

Proposition - Inverse de la transposé
Si A est une matrice carrée inversible alors ${}^tA$ est aussi inversible et $({}^tA)^{-1} = {}^t(A^{-1})$.

Démonstration
« Le coup des chaussettes dans le tiroir… »
$(AB)\times(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AA^{-1} = I_n$ et $(B^{-1}A^{-1})\times(AB) = B^{-1}(A^{-1}A)B = B^{-1}B = I_n$
Donc $(AB)^{-1} = B^{-1}A^{-1}$.
De même, en transposant un produit :

$$I_n = {}^t(I_n) = {}^t(A\times A^{-1}) = {}^t(A^{-1})\times {}^tA \qquad I_n = {}^t(I_n) = {}^t(A^{-1}\times A) = {}^tA\times {}^t(A^{-1})$$

Donc $({}^tA)^{-1} = {}^t(A^{-1})$ □

Pour aller plus loin - Polynôme annulateur
On montrera en seconde année que pour tout matrice $M \in \mathcal{M}_n(\mathbb{K})$, l'ensemble des polynômes annulateurs de M

$$\{P \in \mathbb{K}[X] \mid P(M) = 0\}$$

est non vide. On sait assurément qu'il existe au moins un polynôme de degré n dans cet ensemble (le polynôme caractéristique de M - d'après le théorème de Cayley-Hamilton).
Et mieux, comme $\mathbb{K}[X]$ est un anneau euclidien, cet ensemble est forcément de la forme

$$\{P \in \mathbb{K}[X] \mid P(M) = 0\} = \mu_P\mathbb{K}[X]$$

où μ_P est un polynôme particulier : minimal (en degré), unitaire et propre à P. On l'appelle le polynôme minimal de P

Exercice
Soit $J \in \mathcal{M}_n(\mathbb{R})$ la matrice dont tous les coefficients sont des 1.

1. On suppose $n = 2$. Calculer J^2, J^3, J^k pour $k \in \mathbb{N}$.
2. Mêmes questions avec $n \geqslant 2$ quelconque. J est-elle inversible ?
3. Calculer A^p où $A = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}$.

Correction

1. Pour $n = 2$, on montre aisément que $J^2 = 2J$, $J^3 = 4J$ et pour tout $k \in \mathbb{N}$, $J^k = 2^{k-1}J$ (récurrence).
2. Maintenant, on suppose que $n \geqslant 2$.

$$\text{Coef}_{i,j}(J^2) = \sum_{h=1}^{n} \text{Coef}_{i,h}(J)\text{Coef}_{h,j}(J) = \sum_{h=1}^{n} 1 = n = n\text{Coef}_{i,j}(J)$$

Donc $J^2 = nJ$ Par récurrence, supposons $J^k = n^{k-1}J$.

$$J^{k+1} = J^kJ = n^{k-1}JJ = n^{k-1}nJ = n^kJ = n^{(k+1)-1}J$$

Si J était inversible, on aurait les égalités :

$$J = J^{-1}\times J^2 = nJ^{-1}J = nI_n$$

Ce qui est faux.

3. $A = I_3 + J_3$, I_3 commute avec J_3 et donc on peut appliquer la formule de Newton :

$$A^p = \sum_{k=0}^{p}\binom{p}{\mathbb{K}}I_3^{p-k}J^k = I_3 + \sum_{k=1}^{p}\binom{p}{\mathbb{K}}I_3^{p-k}3^{k-1}J$$

$$= I_3 + \frac{1}{3}\left(\sum_{k=1}^{p}\binom{p}{\mathbb{K}}3^k\right)J = I_3 + \frac{1}{3}((3+1)^p - 1)J$$

$$= \frac{1}{3}\begin{pmatrix} 4^p+2 & 4^p-1 & 4^p-1 \\ 4^p-1 & 4^p+2 & 4^p-1 \\ 4^p-1 & 4^p-1 & 4^p+2 \end{pmatrix}$$

(On peut vérifier pour $p = 1$ que cela fonction bien…

Savoir faire - Exploiter un polynôme annulateur pour trouver M^{-1}
Soit $M \in \mathcal{M}_n(\mathbb{K})$. Supposons que le polynôme $P = \sum_{k=0}^{d} a_kX^k$ annule M, c'est-à-dire que $P(M) = \sum_{k=0}^{d} a_kM^k = 0$. Alors

— si $a_0 \neq 0$, on a alors $I_n = \frac{-1}{a_0}\left(\sum_{k=1}^{d} a_kM^k\right) = M\times\frac{-1}{a_0}\left(\sum_{k=1}^{d} a_kM^{k-1}\right)$.
Et donc nécessairement M est inversible et $M^{-1} = \frac{-1}{a_0}\left(\sum_{k=0}^{d-1} a_{k+1}M^{\mathbb{K}}\right)$

— si $a_0 = 0$, alors il faut faire un raisonnement par l'absurde :
si M était inversible alors en multipliant par M^{-1}, on a

$$M^{-1}\times P(M)=\sum_{k=1}^{d}a_kM^{k-1}=0 \text{ et donc } \sum_{k=0}^{d-1}a_{k+1}M^{\mathbb{K}}=0.$$

Il y a alors deux options,

ou bien cette somme n'est pas nulle, et par l'absurde, M n'est pas inversible,

ou bien cette somme vaut bien 0 et donc on recommence au point initial.

Savoir faire - Exploiter un polynôme annulateur pour trouver M^n

Soit $M\in\mathcal{M}_n(\mathbb{K})$. Supposons que le polynôme $P=\sum_{k=0}^{d}a_kX^k$ annule M, c'est-à-dire que $P(M)=\sum_{k=0}^{d}a_kM^k=0$.

Alors, on peut faire la division euclidienne de X^n par P :

il existe $Q_n,R_n\in\mathbb{K}[X]$ tels que $X^n=Q_n(X)\times P(X)+R_n(X)$ avec $\deg(R_n)<\deg(P)=d$.

On a alors, puisque $M^n=0+R_n(M)$ car $P(M)=0$.

Cela permet de

— démontrer que $\{M^n,\in\mathbb{Z}\}\subset \text{vect}(I_n,M,M^2,\dots M^{d-1})$ (résultat théorique classique)

— calculer explicitement M^n, si l'on sait faire explicitement cette division euclidienne. Pour faire celle-ci, il arrive souvent qu'on utilise les racines de P…

Application - Inverse et puissance de $M=\begin{pmatrix}1&1&0\\0&1&1\\0&0&1\end{pmatrix}$

On remarque que $(M-I_3)^3=0$, donc on a un polynôme annulateur de M :

$$P(X)=(X-1)^3=X^3-3X^2+3X-1$$

Ainsi : $M(M^2-3M+3I_3)=M^3-3M^2+3M=I_3$.

Donc M est inversible d'inverse $M^{-1}=M^2-3M+3I_3=\begin{pmatrix}1&-1&1\\0&1&-1\\0&0&1\end{pmatrix}$ De même, si l'on fait la division euclidienne, on a :

$$X^n=Q_n(X-1)^3+R_n \qquad \text{où}\deg(R_n)\leqslant 2$$

On peut exploiter la formule de Taylor, en 1 (en notant $P(X)=X^n$) :

$$X^n=\sum_{k=0}^{n}\frac{P^{(\mathbb{K})}(1)}{k!}(X-1)^k=\underbrace{P(1)+P'(1)(X-1)+\frac{P''(1)}{2}(X-1)^2}_{=R_n}+(X-1)^3\underbrace{\sum_{k=3}^{n}\frac{P^{(\mathbb{K})}(1)}{k!}(X-1)^{k-3}}_{=Q_n}$$

Or $P(1)=1$, $P'(1)=n$ et $P''(1)=n(n-1)$. Ainsi

$$R_n(X)=1+n(X-1)+\frac{1}{2}n(n-1)(X-1)^2$$

Ainsi $M^n=\begin{pmatrix}1&n&\frac{n(n-1)}{2}\\0&1&n\\0&0&1\end{pmatrix}$

Exercice

Soit $A=(a_{pq})\in\mathcal{M}_n(\mathbb{C})$ définie par $a_{pq}=\exp(\frac{2i\pi pq}{n})$ et $\overline{A}=(\overline{a_{pq}})$.

Calculer $A\overline{A}$ et en déduire que A est inversible.

Correction

(On ne prendra pas i, à cause du i imaginaire pure).

Pour tout $k,\ell\in\mathbb{N}_n$,

$$\text{Coef}_{k,\ell}(A\overline{A}=\sum_{h=1}^{n}a_{k,h}\overline{a_{h,\ell}}=\sum_{h=1}^{n}\exp\big(\frac{2i\pi}{n}(kh-h\ell)\big)$$

$$= \sum_{h=1}^{n} \exp\Big(\frac{2hi\pi}{n}(k-\ell)\Big)$$

$$= \begin{cases} \displaystyle\sum_{h=1}^{n} 1 = n & \text{si } k=\ell \\ \displaystyle\sum_{h=1}^{n} \Big(\exp\Big(\frac{2i\pi}{n}(k-\ell)\Big)\Big)^h = \frac{1-\exp\Big(\frac{2ni\pi}{n}(k-\ell)\Big)}{1-\exp\Big(\frac{2i\pi}{n}(k-\ell)\Big)} = 0 & \text{si } k\neq\ell \end{cases}$$

Donc $A\overline{A} = nI_n$.
Notons alors que $\overline{A}A = \overline{A\overline{A}} = nI_n$.
Ainsi A est inversible et $A^{-1} = \frac{1}{n}\overline{A}$.

4.4. Quelques sous-ensembles remarquables

Matrices diagonales

Proposition - Matrices scalaires
L'ensemble des matrices scalaires d'ordre n, à coefficients dans $\mathbb{K}$, est un sous-espace vectoriel de $\mathcal{M}_n(\mathbb{K})$ de dimension 1, contenant I_n et stable par la multiplication (c'est un sous-anneau commutatif et aussi une sous-algèbre commutative de $\mathcal{M}_n(\mathbb{K})$).

Pour aller plus loin - Base 1
Donc l'espace des matrices scalaires est $\text{vect}(I_n)$*, de dimension 1.*

Proposition - Espace des matrices diagonales
L'ensemble des matrices diagonales d'ordre n, à coefficients dans $\mathbb{K}$, est un sous-espace vectoriel de $\mathcal{M}_n(\mathbb{K})$ de dimension n, contenant I_n et stable par la multiplication (c'est un sous-anneau commutatif et une sous-algèbre commutative de $\mathcal{M}_n(\mathbb{K})$).

Pour aller plus loin - Base 2
Donc l'espace des matrices diagonales est $\text{vect}\big((E_{i,i})_{i\in\mathbb{N}_n}\big)$*, de dimension n.*

Proposition - Inverse de matrices diagonales
Si D est diagonale, $D = \text{diag}(d_1,\dots,d_n)$, alors D est inversible si et seulement pour tout $i\in\{1,\dots,n\}$, $d_i\neq 0$ et alors $D^{-1} = \text{diag}(\frac{1}{d_1},\dots,\frac{1}{d_n})$.
Donc D^{-1} est elle-même une matrice diagonale.

Démonstration
Pour tout $i,j\in\mathbb{N}_n$, par définition de D

$$\text{Coef}_{i,j}(A\times D) = \sum_{h=1}^{n} \text{Coef}_{i,h}(A)\text{Coef}_{h,j}(D) = d_j\text{Coef}_{i,j}(A)$$

Donc pour que $A\times D = I_n$,
il faut pour tout $i\in\mathbb{N}_n$, $\text{Coef}_{i,i}(AD) = d_i\text{Coef}_{i,i}(A) = 1$, il est nécessaire que $d_i\neq 0$.
En prenant alors $\text{Coef}_{i,i}(A) = \frac{1}{d_i}$, on a donc $\text{Coef}_{i,i}(AD) = 1$.
il faut également pour tout $j\neq i$, $\text{Coef}_{i,j}(AD) = 0 = d_i\text{Coef}_{i,j}(A)$,
donc nécessairement $\text{Coef}_{i,j}(A) = 0$.
Par conséquent, pour que D soit inversible, il faut $d_i\neq 0$, pour tout i.
Et la seule matrice qui conviendrait pour inverse est $A = \text{diag}(\frac{1}{d_1},\dots\frac{1}{d_n})$ (car $AD = I_n$).
Un calcul simple confirme : $D\times A = I_n$ également.
On a trouvé alors une condition nécessaire et suffisante avec, alors, une expression de D^{-1}. □

Matrices symétriques et antisymétriques

Pour aller plus loin - Bases 3
Donc $\mathcal{S}_n(\mathbb{K}) = \text{vect}\big((E_{i,j}+E_{j,i})_{i\leqslant j\in\mathbb{N}_n}\big)$*, espace de dimension* $\frac{1}{2}n(n+1)$.
Et $\mathcal{A}_n(\mathbb{K}) = \text{vect}\big((E_{i,j}-E_{j,i})_{i<j\in\mathbb{N}_n}\big)$*, espace de dimension* $\frac{1}{2}n(n-1)$.

Définition - Matrices symétriques et antisymétriques
Soit $A\in\mathcal{M}_n(\mathbb{K})$.
A est dite **symétrique** si ${}^tA = A$, soit si pour tout (i,j), $a_{ij} = a_{ji}$ (ou ${}^i[A]_j = {}^j[A]_i$) ;
on note $\mathcal{S}_n(\mathbb{K})$ l'ensemble des matrices symétriques d'ordre n à coefficients dans $\mathbb{K}$.
A est dite **antisymétrique** si ${}^tA = -A$, soit si pour tout (i,j), $a_{ij} = -a_{ji}$ (ou ${}^i[A]_j = -{}^j[A]_i$) ;

on note $\mathcal{A}_n(\mathbb{K})$ (ou $\mathcal{AS}_n(\mathbb{K})$) l'ensemble des matrices antisymétriques d'ordre n à coefficients dans $\mathbb{K}$.

En l'absence d'ambiguité, on peut noter $\mathcal{S}_n$ et $\mathcal{A}_n$.

Proposition - Diagonale d'une matrice antisymétrique
Les éléments diagonaux d'une matrice antisymétrique sont nuls.

Pour aller plus loin - Matrices et graphes
Que signifie pour un graphe que sa matrice associée est symétrique? Antisymétrique?

Démonstration
A, antisymétrique. Pour tout $i \in \mathbb{N}_n$,

$$\text{Coef}_{i,i}(A) = -\text{Coef}_{i,i}(A) \Longrightarrow 2\text{Coef}_{i,i}(A) = 0$$

□

Ensemble des matrices triangulaires (supérieures)

Théorème - Espace des matrices triangulaires
L'ensemble des matrices triangulaires supérieures (respectivement inférieures) d'ordre n, à coefficients dans $\mathbb{K}$, est un sous-espace vectoriel de $\mathcal{M}_n(\mathbb{K})$ de dimension $\frac{n(n+1)}{2}$, contenant I_n.
Il est stable pour la multiplication (donc est un sous-anneau et une sous-algèbre de $\mathcal{M}_n(\mathbb{K})$).

Pour aller plus loin - Base 4
Donc il s'agit de $\text{vect}\left((E_{i,j})_{i\leqslant j\in\mathbb{N}_n}\right)$*, espace de dimension* $\frac{1}{2}n(n+1)$.

Attention - Pas trop vite
L'ensemble des matrices triangulaires d'ordre n, à coefficients dans $\mathbb{K}$, n'est pas un sous-espace vectoriel de $\mathcal{M}_n(\mathbb{K})$.
L'addition d'une matrice triangulaire supérieure et d'une matrice triangulaire inférieure ne donne pas une matrice triangulaire, la plupart du temps

Pour aller plus loin - Inverse d'une matrice triangulaire supérieure
Nous verrons que si T une matrice triangulaire supérieure, carrée.
Alors T est inversible ssi $\forall\ i \in \mathbb{N}_n$, $Coef_{i,i}(T) \neq 0$.
Et dans ce cas, T^{-1} est également triangulaire supérieure

Remarque - Mais notons :
Une matrice à la fois triangulaire inférieure et supérieure est diagonale.

Savoir faire - Montrer qu'une matrice est triangulaire supérieure
Il faut montrer (condition nécessaire et suffisante) :

$$\forall\ i,j \in \mathbb{N}_n, \quad i > j \quad \Longrightarrow \quad \text{Coef}_{i,j}(T) = 0$$

Il faut démontrer la stabilité par somme (facile) et par multiplication de matrices triangulaires supérieures.

Démonstration
Soit T_1 et T_2 deux matrices triangulaires supérieures.
Pour tout $i > j$:

$$\begin{aligned}\text{Coef}_{i,j}(T_1\,T_2) &= \sum_{k=1}^{n} \text{Coef}_{i,k}(T_1)\text{Coef}_{k,j}(T_2)\\ &= \sum_{k=1}^{i-1} \underbrace{\text{Coef}_{i,k}(T_1)}_{=0\,:\,i>k}\text{Coef}_{k,j}(T_2) + \sum_{k=i}^{n} \text{Coef}_{i,k}(T_1)\underbrace{\text{Coef}_{k,j}(T_2)}_{=0\,:\,k\leqslant i>j} = 0\end{aligned}$$

Donc $T_1\,T_2$ est triangulaire supérieure. □

4.5. Trace d'une matrice carrée

Définition - Trace d'une matrice
Soit $A \in \mathcal{M}_n(\mathbb{K})$ une matrice carrée. On appelle trace de A le scalaire égal à la somme de ses coefficients diagonaux :

$$\mathrm{Tr}\, A = \sum_{i=1}^{n} a_{ii} = \sum_{i=1}^{n} {}^{i}[A]_{i} = \sum_{i=1}^{n} \mathrm{Coef}_{i,i}(A)$$

Proposition - Propriété de la trace
L'application

$$\begin{aligned} \mathrm{Tr} : \quad & \mathcal{M}_n(\mathbb{K}) \to \mathbb{K} \\ & A \mapsto \mathrm{Tr}\, A \end{aligned}$$

est une forme linéaire sur $\mathcal{M}_n(\mathbb{K})$: $\mathrm{Tr}(\lambda A + \mu B) = \lambda \mathrm{Tr}\, A + \mu \mathrm{Tr}\, B$ et

$$\forall (A,B) \in \mathcal{M}_n(\mathbb{K})^2,\ \mathrm{Tr}(AB) = \mathrm{Tr}(BA).$$

Démonstration
Par linéarité de $\mathrm{Coef}_{i,i}$:

$$\mathrm{Tr}(\lambda A + \mu B) = \sum_{i=1}^{n} \mathrm{Coef}_{i,i}(\lambda A + \mu B) = \lambda \sum_{i=1}^{n} \mathrm{Coef}_{i,i}(A) + \mu \sum_{i=1}^{n} \mathrm{Coef}_{i,i}(B) = \lambda \mathrm{Tr}(A) + \mu \mathrm{Tr}(B)$$

$$\mathrm{Tr}(AB) = \sum_{i=1}^{n} {}^{i}[AB]_{i} = \sum_{i=1}^{n} \sum_{h=1}^{n} {}^{i}[A]_{h}^{h}[B]_{i} = \sum_{h=1}^{n} \sum_{i=1}^{n} {}^{h}[B]_{i}^{i}[A]_{h} = \mathrm{Tr}(BA)$$

□

5. Opérations élémentaires sur les matrices

5.1. Opérations élémentaires sur les lignes d'une matrice

Heuristique - Lien résolution de système/inversion de matrice
Le calcul $A \times X = b$ où X et b sont des matrices colonnes est exactement l'écriture d'un système linéaire.
La résolution $X = A^{-1}b$ (si A est inversible donc carrée) exploite les opérations élémentaires sur les lignes du système pour appliquer l'algorithme de Gauss.
Essayons de transférer directement la même idée sur les colonnes de A.

Définition - Opérations élémentaires sur les lignes
On appelle opération élémentaire sur les lignes L_i de la matrice A l'une des transformations suivantes effectuée sur A :

- **Permutation (ou échange)de deux lignes**
  Pour $i \neq j$, $L_i \leftrightarrow L_j$ signifie que l'on permute la i-ième et la j-ième lignes de la matrice.
- **Addition d'un multiple d'une ligne à une autre ligne**
  Pour $i \neq j$, $L_j \leftarrow L_j + \lambda L_i$ signifie que l'on remplace la j-ième ligne L_j de la matrice par $L_j + \lambda L_i$, où $\lambda \in K$.
- **Multiplication d'une ligne par un scalaire NON NUL**
  Pour tout $\alpha \in K$, $\alpha \neq 0$, $L_i \leftarrow \alpha L_i$ signifie que l'on remplace la i-ième ligne par αL_i.

Chacune des manipulations précédentes correspond à un produit matriciel :

Proposition - Transformation élémentaire sur les lignes comme un produit

Effectuer une transformation élémentaire sur les lignes d'une matrice $A \in \mathcal{M}_{n,p}(\mathbb{K})$ revient à calculer le produit matriciel à gauche de A : EA où E est l'une des matrices suivantes :

— Pour $L_i \leftrightarrow L_j$,

$$E = \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 0 & & 1 & & \\ & & & \ddots & & & \\ & & 1 & & 0 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix} = I_n - E_{ii} - E_{jj} + E_{ij} + E_{ji}$$

C'est une matrice de transposition, notée habituellement $P_{i,j} (= P_{j,i})$

— Pour $L_i \leftarrow L_i + \lambda L_j$

$$E = \begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & & \\ & \lambda & & \ddots & \\ & & & & 1 \end{pmatrix} = I_n + \lambda E_{ij} \quad \lambda \text{ en ligne } i, \text{ colonne } j$$

C'est une matrice de transvection, notée habituellement $T_{i,j}(\lambda)$

— Pour $L_i \leftarrow \alpha L_i$

$$E = \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & 0 & \\ & & 1 & & & & \\ & & & 1 & & & \\ & & & & \alpha & & \\ & & 0 & & & \ddots & \\ & & & & & & 1 \end{pmatrix} = I_n + (\alpha - 1)E_{ii}$$

C'est une matrice de dilatation, notée habituellement $D_i(\alpha)$

Démonstration

Soit $A \in \mathcal{M}_{n,p}(\mathbb{K})$ et $E_{ij} \in \mathcal{M}_n(\mathbb{K})$ une des matrices de la base canonique.

On rappelle que ${}^h[BA]_k = \sum_{r=1}^{n} {}^h[B]_r \, {}^r[A]_k$.

Donc ${}^h[E_{i,j}A]_k = \sum_{r=1}^{n} {}^h[E_{i,j}]_r \, {}^r[A]_k = {}^h[E_{i,j}]_j \, {}^j[A]_k = \begin{cases} 0 & \text{si } h \neq i \\ {}^j[A]_k & \text{si } h = i \end{cases}$

On a donc $L_h(E_{i,j}A) = \begin{cases} 0 & \text{si } h \neq i \\ L_j(A) & \text{si } h = i \end{cases} = \delta_{h=i} L_j(A)$ Alors $E_{ij}A$ est la matrice de $\mathcal{M}_{n,p}(\mathbb{K})$ dont toutes les lignes sont nulles sauf la i-ième , qui est la j-ième ligne de A :

$$\forall\, h \neq i \, L_h(E_{i,j}A) = 0 \qquad L_i(E_{i,j}A) = L_j(A)$$

$$\forall\, h \neq i, \ell \in \mathbb{N}_n \text{Coef}_{h,\ell}(E_{i,j}A) = 0 \qquad \text{Coef}_{i,\ell}(E_{i,j}A) = \text{Coef}_{j,\ell}(A)$$

Puis comme $I_n A = A$ et en examinant chacune des opérations élémentaires on obtient le résultat.

• $L_k\big((I_n - E_{i,i} - E_{j,j} + E_{i,j} + E_{j,i})A\big) = L_k(A) - \delta_{k,i}L_i(A) - \delta_{k,j}L_j(A) + \delta_{k,i}L_j(A) + \delta_{k,j}L_i(A)$.

$$L_k(EA) = \begin{cases} L_i(A) - L_i(A) + L_j(A) = L_j(A) & \text{si } k = i \\ L_j(A) - L_j(A) + L_i(A) = L_i(A) & \text{si } k = i \\ L_k(A) & \text{sinon} \end{cases}$$

• $L_k\big((I_n + \lambda E_{i,j})A\big) = L_k(A) + \delta_{k,i}\lambda L_j(A)$.

$$L_k(EA) = \begin{cases} L_i(A) + \lambda L_j(A) & \text{si } k = i \\ L_k(A) & \text{sinon} \end{cases}$$

- $L_k\big((I_n + (\alpha-1)E_{i,i})A\big) = L_k(A) + (\alpha-1)\delta_{k,i}L_i(A)$.

$$L_k(EA) = \begin{cases} L_i(A) + (\alpha-1)L_i(A) = \alpha L_i(A) & \text{si } k = i \\ L_k(A) & \text{sinon} \end{cases}$$ □

Proposition - Opération élémentaire en I_n
On considère une opération élémentaire, notée φ, qui transforme une matrice $A \in \mathscr{M}_{n,p}(\mathbb{K})$ en la matrice $\varphi(A) \in \mathscr{M}_{n,p}(\mathbb{K})$ et la matrice I_n en $\varphi(I_n)$.

Alors $\varphi(A) = \varphi(I_n) \times A$.

Et par récurrence, si $\varphi_1, \varphi_2, \dots \varphi_k$ sont k transformations élémentaires (sur les lignes) qui s'appliquent à des matrices possèdant n lignes, alors, pour tout $A \in \mathscr{M}_{n,p}(\mathbb{K})$:

$[\varphi_k \circ \cdots \circ \varphi_1](A) = \varphi_k(I_n) \times \cdots \times \varphi_1(I_n) \times A$.

Démonstration
Cela a bien un sens de considérer que φ peut s'appliquer à A et à I_n, car il s'agit d'opération sur les lignes et A et I_n ont le même nombre de lignes.
Pour prouver ce résultat il suffit de vérifier que pour les trois transformations possibles on a bien $E = \varphi(I_n)$, puisque $\varphi(A) = EA$.
On démontre ensuite l'hérédité de la récurrence : Supposons que $\varphi_k \circ \cdots \circ \varphi_1(A) = \varphi_k(I_n) \times \cdots \times \varphi_1(I_n) \times A$.
Notons $A' = \varphi_k \circ \cdots \circ \varphi_1(A)$.
Alors

$$[\varphi_{k+1} \circ \varphi_k \circ \cdots \circ \varphi_1(A) = \varphi_{k+1}(\varphi_k \circ \cdots \circ \varphi_1(A)) = \varphi_{k+1}(A') = \varphi_{k+1}(I_n) \times A' = \varphi_{k+1}(I_n) \times \varphi_k(I_n) \times \cdots \times \varphi_1(I_n) \times A$$

□

Cela se concrétise dans le savoir faire suivant

Savoir faire - Retenir les opérations matricielles codant les opérations élémentaires

Pour une opération sur les lignes de A, il s'agit toujours de produit à gauche de A.
(On verra pour les colonnes, il s'agit de produits à droites de A)
Par quelle matrice ?
C'est toujours par la matrice qu'on obtient lorsqu'on applique la transformation élémentaire en question à I_n.
Ainsi, par exemple, si l'on veut faire $L_3 \leftarrow L_3 - 2L_2$, pour $n = 3$, on multiplie par la matrice

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow[L_3 \leftarrow L_3 - 2L_2]{} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{pmatrix}$$

Proposition - Inversibilité des opérations élémentaires
Si une matrice B est déduite de A par une opération élémentaire φ, alors A peut se déduire de B par l'opération inverse φ^{-1} suivant le tableau suivant :

φ	φ^{-1}
$L_i \leftrightarrow L_j$	$L_i \leftrightarrow L_j$
$L_j \leftarrow L_j + \lambda L_i$	$L_j \leftarrow L_j - \lambda L_i$
$L_i \leftarrow \alpha L_i$	$L_i \leftarrow \frac{1}{\alpha} L_i$

Proposition - Inversibilité des matrices élémentaires
Si φ est une opération élémentaire sur les lignes alors la matrice carrée $\varphi(I_n)$ est inversible et $\varphi(I_n)^{-1} = \varphi^{-1}(I_n)$.
On a alors $P_{i,j}^{-1} = P_{i,j}$ (symétrique, involutif),
$\big(T_{i,j}(\lambda)\big)^{-1} = T_{i,j}(-\lambda)$, et $\big(D_i(\alpha)\big)^{-1} = D_i(\frac{1}{\alpha})$.

Démonstration
D'après la proposition précédente $\varphi^{-1}(I_n)\varphi(I_n) = \varphi^{-1}(\varphi(I_n)) = I_n$ □

5.2. Opérations élémentaires sur les colonnes d'une matrice

On définit les mêmes opérations élémentaires sur les colonnes que sur les lignes. On obtient alors les résultats suivants :

Proposition - Transformation élémentaire sur les colonnes comme un produit
Effectuer une transformation élémentaire sur les colonnes d'une matrice $A \in \mathcal{M}_{n,p}(\mathbb{K})$ revient à calculer le produit matriciel AF où F est l'une des matrices suivantes :
— Pour $C_i \leftrightarrow C_j$,
$$F = I_p - E_{ii} - E_{jj} + E_{ij} + E_{ji}$$
— Pour $C_i \leftarrow C_i + \lambda C_j$
$$F = I_p + \lambda E_{ji}$$
— Pour $C_i \leftarrow \alpha C_i$
$$F = I_p + (\alpha - 1)E_{ii}$$

Démonstration
Soit $A \in \mathcal{M}_{n,p}(\mathbb{K})$ et $E_{ij} \in \mathcal{M}_p(\mathbb{K})$ une des matrices de la base canonique. Alors AE_{ij} est la matrice de $\mathcal{M}_{n,p}(\mathbb{K})$ dont toutes les colonnes sont nulles sauf la i-ième , qui est la j-ième colonne de A. $AI_p = A$ et en examinant chacune des opérations élémentaires on obtient le résultat. □

Proposition - Opération élémentaire en I_p
On considère une opération élémentaire, notée ψ, qui transforme une matrice $A \in \mathcal{M}_{n,p}(\mathbb{K})$ en la matrice $\psi(A) \in \mathcal{M}_{n,p}(\mathbb{K})$ et la matrice I_p en $\psi(I_p)$.
Alors $\psi(A) = A\psi(I_p)$.
Et par récurrence, si $\psi_1, \psi_2, \dots \psi_k$ sont k transformations élémentaires (sur les colonnes) qui s'appliquent à des matrices possédant p lignes, alors, pour tout $A \in \mathcal{M}_{n,p}(\mathbb{K})$:
$\psi_k \circ \cdots \circ \psi_1(A) = A \times \psi_1(I_p) \times \cdots \times \psi_k(I_p)$.

Démonstration
On démontre le résultat par récurrence :

$$\psi_{k+1}(\psi_k \circ \cdots \circ \psi_1)(A) = \psi_{k+1}(A') = A'\psi_{k+1}(I_p) = A \times \psi_1(I_p) \times \cdots \times \psi_k(I_p) \times \psi_{k+1}(A)$$

par hypothèse de récurrence. □

5.3. Transformation par opérations élémentaires (matrices inversibles)

Conservation de l'inversibilité

Histoire - Gauss et le calcul matriciel linéaire
Encore Gauss… Voir le cours d'arithmétique.

Pour aller plus loin - Mieux!
Plus que la conservation du caractère inversible, c'est le rang d'une matrice qui est exac-

Proposition - Conservation de l'inversibilité
Soient $A, B \in \mathcal{M}_n(\mathbb{K})$. On suppose que A est inversible.
Alors AB est inversible si et seulement si B est inversible

Démonstration
Si B est inversible, alors AB est inversible (d'inverse $B^{-1}A^{-1}$).
Si AB est inversible, alors $B = A^{-1}(AB)$ est inversible (d'inverse $(AB)^{-1}A$) □

Corollaire - Conservation d'inversibilité par les opérations élémentaires
Les opérations élémentaires sur les lignes (ou sur les colonnes) d'une matrice carrée conservent le caractère inversible/non inversible d'une matrice.

Démonstration
Une opération élémentaire sur une matrice est un produit à gauche par une matrice inversible.
D'après le théorème précédent, on en déduit la conservation du critère d'inversibilité □

Algorithme

Théorème - Transformation de Gauss-Jordan
Soit $A \in \mathcal{M}_n(\mathbb{K})$.
A est inversible si et seulement s'il est possible de transformer A en une matrice triangulaire supérieure, sans 0 sur la diagonale, à l'aide uniquement d'opérations élémentaires portant sur les lignes.
Dans ce cas, on peut terminer la décomposition de A vers I_n par suite d'opérations élémentaires.
Si on applique alors à la matrice I_n les mêmes opérations élémentaires, dans le même ordre, on obtient A^{-1}.

Histoire - Fang-cheng
« Cette méthode est connue des Chinois depuis au moins le Ier siècle de notre ère. Elle est référencée dans le livre chinois Jiuzhang suanshu (Les Neuf Chapitres sur l'art mathématique), dont elle constitue le huitième chapitre, sous le titre « Fang cheng » (la disposition rectangulaire). La méthode est présentée au moyen de dix-huit exercices. Dans son commentaire daté de 263, Liu Hui en attribue la paternité à Chang Ts'ang, chancelier de l'empereur de Chine au IIe siècle avant notre ère. » (pompage sans scrupule de Wikipedia...)

On commence par un lemme, que l'on démontre, puis que l'on applique aux matrices triangulaires supérieures avant de faire la démonstration globale.

Lemme - Trigonalisation
Si $M = \left(\begin{array}{c|c} A & B \\ \hline 0_{n-s,s} & C \end{array}\right) \in \mathcal{M}_n(\mathbb{K})$, où $A \in GL_s(\mathbb{K})$ ($s \in \mathbb{N}_n$, quelconque) et $C \in \mathcal{M}_{n-s}(\mathbb{K})$, alors :
• ou bien $\forall\, i \in \mathbb{N}_{n-s}$, ${}^{i}[C]_1 = 0$ et alors M n'est pas inversible
• ou bien $\exists\, i \in \mathbb{N}_{n-s}$ tel que ${}^{i}[C]_1 \neq 0$, alors $\exists\, T$, produit de matrices élémentaires telle que $T \times M = \left(\begin{array}{c|ccc} A & & B & \\ \hline & x & & \\ & 0 & & \\ 0_{n-s,s} & \vdots & & C' \\ & 0 & & \end{array}\right)$ avec $x \neq 0$.

Pour aller plus loin - Réussite de l'algorithme de Gauss - Conditionnement
Théoriquement, la question ne se pose pas : l'algorithme réussit toujours sa mission.
Mais dans la pratique, en particulier informatiquement, il peut y avoir des approximations numériques (par exemple à 2^{-52} près en Python).
Supposons donc la donnée de A avec une erreur à δA près.
On note (par exemple) $\|B\|_2 = \sqrt{(Tr(B^T \times B)}$, il s'agit d'une norme. Elle est de plus multiplicative : $\|AB\|_2 \leqslant \|A\|_2 \times \|B\|_2$.
On définit alors le conditionnement de A

$$cond(A) = \|A\|_2 \times \|A^{-1}\|_2$$

On a alors l'erreur relative à l'arrivée proportionnelle à l'erreur relative du départ par le conditionnement.

$$\frac{\|\delta A^{-1}\|_2}{\|A^{-1}\|_2} \leqslant cond(A)\frac{\|\delta A\|_2}{\|A\|_2}$$

On dit qu'une matrice est mal conditionnée si $cond(A)$ est important. Plus une matrice est mal conditionnée, plus l'impact d'une petite erreur au départ est important sur l'erreur finale.

Démonstration
Les cas d'étude sont bien complémentaires.
• Supposons qu'il existe $i \in \mathbb{N}_{n-s}$ tel que ${}^{i}[C]_1 \neq 0$,
Transformons C en C' par $L_1 \leftrightarrow L_i$, codée par le produit à gauche par $\left(\begin{array}{c|c} I_s & 0 \\ \hline 0 & P_{1,i} \end{array}\right)$,
puis pour k de 2 à $n-s$, $L_k \leftarrow L_k - \underbrace{\frac{{}^{k}[C']_1}{{}^{1}[C']_1}}_{\neq 0} L_1$. On obtient bien le résultat attendu.
• Supposons que pour tout $i \in \mathbb{N}_{n-s}$, ${}^{i}[C]_1 = 0$.

A étant inversible, on considère $M' = \left(\begin{array}{c|c} A^{-1} & 0_{s,n-s} \\ \hline 0_{n-s,s} & I_{n-s} \end{array}\right)$, d'inverse $(M')^{-1} = \left(\begin{array}{c|c} A & 0_{s,n-s} \\ \hline 0_{n-s,s} & I_{n-s} \end{array}\right)$.

Par produit par blocs :

$$M' \times M = \left(\begin{array}{c|c} I_s & A^{-1}B \\ \hline 0_{n-s,s} & \begin{array}{c} 0 \\ 0 \\ \vdots \\ 0 \end{array} \quad C' \end{array}\right)$$

Puis on réalise : $C_{s+1}(M'M) \leftarrow C_{s+1}(M'M) - \sum_{j=1}^{s} {}^{j}[M'M]_{s+1} \cdot C_j(M'M)$ pour obtenir $\tilde{M}$.

Cela est codée par un produit matricielle à droite : $\tilde{M} = M'M \times \overline{T}$, avec $\overline{T}$, produit de matrices élémentaires donc inversible.

Or, $\tilde{M}$ est caractérisé par le fait qu'elle possède une colonne nulle (la $s+1$-ième).

Ainsi $\tilde{M}$ est diviseur de 0 ($\tilde{M} \times E_{s+1,1} = 0$), donc non inversible.

Comme $\overline{T}$ et M' sont inversible, nécessairement M ne l'est pas. □

Corollaire - Matrices triangulaires supérieures

Si T est une matrice (carrée) triangulaire supérieure, avec des coefficients non nuls sur la diagonale.

Alors il existe T', triangulaire supérieure et inversible telle que $T' \times T = I_n$.

Réciproquement, si T est une matrice (carrée) triangulaire supérieure, avec (au moins) un coefficients nuls sur la diagonale, alors T n'est pas inversible.

Démonstration

• Sens direct.

On va démontrer le résultat par récurrence sur n, l'ordre de la matrice T.

— Si $n = 1$, alors T' est la matrice à un élément : $\frac{1}{{}_1[T]_1}$.

— Soit $n \in \mathbb{N}$. Supposons le résultat vraie pour toute matrice triangulaire supérieure d'ordre n.

Soit T, triangulaire supérieure d'ordre $n+1$.

D'après l'analyse précédente, il existe T_1, triangulaire supérieure inversible telle que $T_1 \times T = \left(\begin{array}{cc} \tilde{T} & 0_{n-1,1} \\ 0_{1,n-1} & {}^{n}[T]_n \end{array}\right)$.

$\tilde{T}$ est l'extraction de T sur les n premières lignes et colonnes, elle est donc triangulaire supérieure, sans coefficient nul sur la diagonale.

On peut appliquer l'hypothèse de récurrence.

Il existe $\tilde{T}'$ triangulaire supérieure et inversible telle que $\tilde{T}' \times \tilde{T} = I_n$.

Considérons T_2, définie par blocs par : $\left(\begin{array}{cc} \tilde{T}' & 0_{n-1,1} \\ 0_{1,n-1} & \frac{1}{{}^{n}[T]_n} \end{array}\right)$, triangulaire supérieure.

On a alors (par blocs) :

$$T_2 \times T_1 \times T = \left(\begin{array}{cc} \tilde{T}' & 0_{n-1,1} \\ 0_{1,n-1} & \frac{1}{{}^{n}[T]_n} \end{array}\right) \times \left(\begin{array}{cc} \tilde{T} & 0_{n-1,1} \\ 0_{1,n-1} & {}^{n}[T]_n \end{array}\right) = \left(\begin{array}{cc} \tilde{T}'\tilde{T} & 0_{n-1,1} \\ 0_{1,n-1} & 1 \end{array}\right) = I_{n+1}$$

Et comme le produit de deux matrices triangulaires supérieures et inversible est triangulaire supérieure et inversible, l'hypothèse de récurrence est héréditaire.

Enfin, comme T' est inversible, en multipliant à gauche par T'^{-1}, on trouve $T = T'^{-1}$ et donc $T \times T' = I_n$.

• Réciproquement, s'il y a un 0 sur la diagonale.

En se trouve dans la situation du lemme en considérant le nombre $s := \min\{h \mid {}^{h}[T]_h = 0\} - 1$. □

On peut enfin démontrer l'algorithme de Gauss.

Démonstration

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

On applique le lemme colonne après colonne (pour s de 1 à n), tant que la matrice M reste visiblement inversible.

(Comme le statut de M ne peut pas changer elle était inversible ou non dès le début, mais nous l'ignorions).

La matrice A_s qui apparait à l'étape s est alors triangulaire supérieure, inversible dans qu'il n'apparait pas de 0 sur la diagonale. • ou bien, à une étape, un zéro est apparu sur la diagonale et alors M n'est pas inversible. • ou bien, on peut aller au bout du lemme de trigonalisation, à chaque étape A_s est bien triangulaire supérieure, sans 0 sur la diagonale donc inversible.

Ainsi, par opérations élémentaires sur les lignes de M on obtienne une matrice triangulaire T, cela signifie

$$E_k E_{k-1} \dots E_1 M = T$$

On termine d'inverser par opérations élémentaires sur les lignes de T : $E_m E_{m-1} \times E_{k+1} T = I_n$. Cela signifie

$$E_m E_{m-1} \dots E_1 A = I_n$$

En posant $B = E_m E_{m-1} \dots E_1 \in \mathcal{M}_n(\mathbb{K})$, on obtient $BA = I_n$, A étant inversible, on a $BAA^{-1} = I_n A^{-1}$ d'où $B = A^{-1}$. Or $B = E_m E_{m-1} \dots E_1 = E_m E_{m-1} \dots E_1 I_n$ est la matrice déduite de I_n par la même succession des m opérations. □

Remarque - Et les colonnes

On peut également procéder à l'aide d'opérations élémentaires sur les colonnes puisque l'on aura alors $AF_1 \dots F_p = I_n$ d'où $A^{-1} = I_n F_1 \dots F_m$,

Attention - Surtout pas!

Mais en revanche il ne faut surtout **pas mélanger les deux types d'opérations** car l'on aboutit à $E_m E_{m-1} \dots E_1 A F_1 \dots F_p = I_n$ qui n'est pas de la forme $AB = I_n$ et ne permet pas de conclure à l'inversibilité de A.
Mais néanmoins, comme souvent, rien n'est perdu dans ce cas là : on montrera que A est équivalente à I_n, donc de même rang : n, donc elle est inversible…

Corollaire - Inversion à droite suffisante

Soit $A \in \mathcal{M}_n(\mathbb{K})$. Supposons qu'il existe $B \in \mathcal{M}_n(\mathbb{K})$ tel que $AB = I_n$.
Alors A est inversible et $A^{-1} = B$.
(De même, B est inversible et $B^{-1} = A$)

Démonstration

On applique une série de transformations élémentaires à A pour l'échelonner.

$$\exists\, E_1, E_2, \dots E_k, T \text{ telles que } E_k E_{k-1} \cdots E_1 A = T$$

On a donc $TB = E_k E_{k-1} \cdots E_1$.
Puis en multipliant à droite par $E_1^{-1} \dots E_k^{-1}$, on a $T(B(E_k E_{k-1} \cdots E_1)^{-1}) = I_n$.
On a vu que si T avait un coefficient nul sur la diagonale, il est impossible d'obtenir I_n.
Donc T n'a aucun coefficient nul sur la diagonale, il est inversible.
Par produit de matrices inversibles : $A = (E_k E_{k-1} \cdots E_1)^{-1} T$ est inversible.
Et donc $B = A^{-1} AB = A^{-1}$. □

Pour aller plus loin - Autres algorithmes

Il existe d'autre méthode (raffinement de Gauss) pour inverser une matrice. Elles ont une complexité algorithmique plus faible (on gagne sur la constante dans le O - ou mieux encore dans certains cas tout particulier de matrices). Il s'agit des méthodes LU, QR, Choleski. On peut aussi employer des méthodes itératives assez différentes dans leur philosophie comme les méthodes de Jacobi… On crée une suite de matrice (A_n), avec $A_0 = A$ et $\lim(A_n) = A^{-1}$. On s'arrête quand on est « assez proche » de A^{-1}.

Pour aller plus loin - Coût de l'algorithme de Gauss

Si l'on suit parfaitement la démarche de l'algorithme de Gauss, la complexité calculatoire est en $O(n^3)$ où n est l'ordre de la matrice considérée.
En effet, il faut mettre en place trois boucles imbriquées, indexée par n (pas tout à fait l'une est triangulaire).

Savoir faire - Inversibilité et inverse d'une matrice par l'algorithme de Gauss

Calculer l'inverse de $A = \begin{pmatrix} 1 & 0 & -1 \\ 1 & -1 & 0 \\ -1 & 0 & 0 \end{pmatrix}$.

On applique l'algorithme de Gauss-Jordan en considérant $(A|I_3)$:

$$\left(\begin{array}{ccc|ccc} 1 & 0 & -1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 & 1 & 0 \\ -1 & 0 & 0 & 0 & 0 & 1 \end{array}\right) \longrightarrow \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & 0 & -1 \\ 1 & 0 & -1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 & 1 & 0 \end{array}\right) \quad \begin{cases} L_1 \leftarrow -L_3 \\ L_2 \leftarrow L_1 \\ L_3 \leftarrow L_2 \end{cases}$$

$$\longrightarrow \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 1 & 0 & 1 \\ 0 & -1 & 0 & 0 & 1 & 1 \end{array}\right) \quad \begin{cases} L_2 \leftarrow L_2 - L_1 \\ L_3 \leftarrow L_3 - L_1 \end{cases} \longrightarrow \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & 0 & -1 \\ 0 & 1 & 0 & 0 & -1 & -1 \\ 0 & 0 & 1 & -1 & 0 & -1 \end{array}\right) \quad \begin{cases} L_2 \leftarrow \\ L_3 \leftarrow \end{cases}$$

Par suite d'opérations élémentaires, on a transformé A en I_3 donc A est inversible.
Par suite des mêmes opérations élémentaires, on a transformé I_3 en A^{-1},
donc $A^{-1} = \begin{pmatrix} 0 & 0 & -1 \\ 0 & -1 & -1 \\ -1 & 0 & -1 \end{pmatrix}$.

A retenir : **dans une opération élémentaire, il ne faut pas perdre l'information d'une ligne**. Autrement écrit, il faut toujours pouvoir revenir en arrière!!

Exercice

Calculer l'inverse de $A=\begin{pmatrix}1&2&3\\-1&-1&-2\\2&2&5\end{pmatrix}$.

Correction

On applique l'algorithme de Gauss-Jordan en considérant $(A|I_3)$:

$$\left(\begin{array}{ccc|ccc}1&2&3&1&0&0\\-1&-1&-2&0&1&0\\2&2&5&0&0&1\end{array}\right)\longrightarrow\left(\begin{array}{ccc|ccc}1&2&3&1&0&0\\0&1&1&1&1&0\\0&-2&-1&-2&0&1\end{array}\right)\quad\begin{cases}L_2\leftarrow L_1+L_2\\L_3\leftarrow L_3-2L_1\end{cases}$$

$$\longrightarrow\left(\begin{array}{ccc|ccc}1&2&3&1&0&0\\0&1&1&1&1&0\\0&0&1&0&2&1\end{array}\right)\quad\{\ L_3\leftarrow L_3+2L_2\quad\longrightarrow\left(\begin{array}{ccc|ccc}1&0&0&-1&-4&-1\\0&1&0&1&-1&-1\\0&0&1&0&2&1\end{array}\right)\quad\begin{cases}L_1\leftarrow L_1-2L_2-L_3\\L_2\leftarrow L_2-L_3\end{cases}$$

Par suite d'opérations élémentaires, on a transformé A en I_3 donc A est inversible.

Par suite des mêmes opérations élémentaires, on a transformé I_3 en A^{-1}, donc $A^{-1}=\begin{pmatrix}-1&-4&-1\\1&-1&-1\\0&2&1\end{pmatrix}$.

Remarque - Et si A n'est pas inversible?

Si A n'est pas inversible, alors l'algorithme conduit à une matrice échelonné (triangulaire supérieure) dont la diagonale possède (au moins) un zéro. Il est alors totalement impossible d'obtenir I_n.

Nous verrons plus loin que l'algorithme donne alors une nouvelle information à propos de la matrice A : son rang!

Exercice

Déterminer les inverses de

$$A=\begin{pmatrix}2&1&-2\\0&3&1\\-2&-3&5\end{pmatrix},\ B=\begin{pmatrix}1&-1&0\\1&2&1\\1&1&0\end{pmatrix},\ C=\begin{pmatrix}10&9&1\\9&10&5\\1&5&9\end{pmatrix},\ D=\begin{pmatrix}1&2&1&-1\\1&-1&1&1\\2&1&3&-1\\1&1&2&1\end{pmatrix}$$

Correction

On applique l'algorithme de Gauss-Jordan en considérant $(A|I_3)$:

$$\left(\begin{array}{ccc|ccc}2&1&-2&1&0&0\\0&3&1&0&1&0\\-2&-3&5&0&0&1\end{array}\right)\longrightarrow\left(\begin{array}{ccc|ccc}2&1&-2&1&0&0\\0&3&1&0&1&0\\0&-2&3&1&0&1\end{array}\right)\quad\{\ L_3\leftarrow L_3+L_1$$

$$\longrightarrow\left(\begin{array}{ccc|ccc}1&\frac12&-1&\frac12&0&0\\0&3&1&0&1&0\\0&0&\frac{11}{3}&1&\frac23&1\end{array}\right)\quad\begin{cases}L_1\leftarrow\frac12L_1\\L_3\leftarrow L_3+\frac23L_2\end{cases}$$

$$\longrightarrow\left(\begin{array}{ccc|ccc}1&\frac12&-1&\frac12&0&0\\0&1&\frac13&0&\frac13&0\\0&0&1&\frac{3}{11}&\frac{2}{11}&\frac{3}{11}\end{array}\right)\quad\begin{cases}L_2\leftarrow\frac13L_2\\L_3\leftarrow-\frac{3}{11}L_3\end{cases}$$

$$\longrightarrow\left(\begin{array}{ccc|ccc}1&\frac12&-1&\frac12&0&0\\0&1&0&\frac{-1}{11}&\frac{3}{11}&\frac{-1}{11}\\0&0&1&-\frac{3}{11}&\frac{2}{11}&\frac{3}{11}\end{array}\right)\quad\{\ L_2\leftarrow L_2-\frac13L_3$$

$$\longrightarrow\left(\begin{array}{ccc|ccc}1&0&0&\frac{9}{11}&\frac{1}{22}&\frac{7}{22}\\0&1&0&\frac{-1}{11}&\frac{3}{11}&\frac{-1}{11}\\0&0&1&-\frac{3}{11}&\frac{2}{11}&\frac{3}{11}\end{array}\right)\quad\{\ L_1\leftarrow L_1-\frac12L_2+L_3$$

Par suite d'opérations élémentaires, on a transformé A en I_3 donc A est inversible.

Par suite des mêmes opérations élémentaires, on a transformé I_3 en A^{-1}, donc $A^{-1}=\frac{1}{22}\begin{pmatrix}18&1&7\\-2&6&-2\\6&4&6\end{pmatrix}$.

$C^{-1}=\begin{pmatrix}65&-76&35\\-76&89&-41\\35&-41&19\end{pmatrix}$ et $D^{-1}=\frac17\begin{pmatrix}8&9&-3&-4\\3&-1&-2&2\\-6&-5&4&3\\1&2&-3&3\end{pmatrix}$

6. Bilan

Synthèse

⇝ Les problèmes à double entrées se codent dans des tableaux. Ces problèmes s'articulent lorsqu'ils ont une entrée commune ; cela se code par la multiplication des tableaux. Nous avons ainsi défini des opérations sur les tableaux (à taille fixée), donnant une structure d'anneau et d'espace vectoriel à cet ensemble (lorsque les dimensions correspondent bien). Attention c'est un anneau non commutatif, avec des diviseurs de zéros…

⇝ Un espace est particulièrement intéressant, celui des matrices de taille carrée, car dans cette situation, elles codifient les transformations d'un espace sur lui-même. Cas très fréquent.
Pour ces matrices, on peut être amené à étudier leurs puissances A^n (cf. Sciences physiques). Mais ici on se concentre surtout sur l'étude de l'inversibilité de A et le calcul de A^{-1} si possible.

⇝ On met en place un algorithme pour répondre à ces questions. C'est un algorithme sur les matrices, mais qui se code aussi par du calcul matriciel. Ainsi on voit les matrices (ou plutôt des ensembles de matrices) agir sur les matrices (ou plutôt d'autres ensembles de matrices). Cela nous donne l'idée de considérer des actions de groupes $GL_n(\mathbb{K})$ (groupe des matrices carrées de taille n, inversibles) sur tout plein d'ensembles.
En exploitant cette idée, on dégage la notion de rang d'une matrice ; il est invariant par produit par matrices inversibles à droite et à gauche et on trouve une forme normalisée adaptée (pour la relation d'équivalence qui conserve le rang). Cette méthode est exploitée similairement dans plein d'autres parties des mathématiques (matrices congruentes, matrices semblables…)

Savoir-faire et Truc & Astuce du chapitre

— Savoir-faire - Notations
— Savoir-faire - Notation et multiplication matricielle
— Savoir-faire - Présentation des calculs
— Savoir-faire - Exploiter un polynôme annulateur pour trouver M^{-1}
— Savoir-faire - Exploiter un polynôme annulateur pour trouver M^n
— Savoir-faire - Montrer qu'une matrice est triangulaire supérieure
— Savoir-faire - Retenir les opérations matricielles codant les transformations élémentaires.
— Savoir-faire - Inversibilité et inverse d'une matrice par algorithme de Gauss

Notations

Notations	*Définitions*	*Propriétés*	*Remarques*
$\mathcal{M}_{n,k}(\mathbb{K})$	Ensemble (espace) des matrices avec n lignes et p colonnes à coefficients dans $\mathbb{K}$		On note $\mathcal{M}_n(\mathbb{K})$ l'espace $\mathcal{M}_{n,n}(\mathbb{K})$
$\mathcal{S}_n(\mathbb{K})$, resp. $\mathcal{A}_n(\mathbb{K})$	Ensemble (espace) des matrices symétrique (nécessairement carrées), resp. antisymétrique	$A^T = A$, resp. $A^T = -A$	
$GL_n(\mathbb{K})$	Groupe (linéaire) des matrices carrées d'ordre n inversible		
${}^i[A]_j$ ou $[A]_{i,j}$ ou $\text{Coeff}_{i,j}(A)$	Coefficient en ligne i et colonne j de la matrice A	C'est une application linéaire	${}^i[AB]_j = \sum_{k=1}^{n} {}^i[A]_k {}^k[B]_j$
${}^tA$ ou A^T	Transposée de la matrice A	${}^i[A^T]_j = {}^j[A]_i$	Application linéaire involutive
A^n, A^{-1}	Puissance n^e de A (par récurrence). Inverse de A (si inversible)		
$\text{Tr}(A)$	Trace de A (somme des coeff. sur la diagonale)	$\text{Tr}(A) = \sum_{i=1}^{n} {}^i[A]_i$	$\text{Tr}(AB) = \text{Tr}(BA)$
$(E_{i,j})_{i,j}$	Base canonique de $\mathcal{M}_{n,p}(\mathbb{K})$	${}^h[E_{i,j}]_k = \delta_{i,h}\delta_{j,k}$	$E_{i,j} \times E_{h,k} = \delta_{j,h}E_{i,k}$
$P_{i,j} = I_n + E_{i,j} + E_{j,i} - E_{i,i} - E_{j,j}$	Matrice de transposition	Inversion des lignes (resp. colonnes) i et j si multiplication par la gauche (resp. la droite)	$P_{i,j}^{-1} = P_{i,j}$
$T_{i,j}(\lambda) = I_n + \lambda E_{i,j}$	Matrice de transvection	$L_i \leftarrow L_i + \lambda L_j$ à gauche (resp. $C_j \rightarrow C_j + \lambda C_i$ - à droite)	$T_{i,j}(\lambda)^{-1} = T_{i,j}(-\lambda)$
$D_i(\alpha) = I_n + (\alpha - 1)E_{i,i}$	Matrice de dilatation	$L_i \leftarrow \alpha L_i$ à gauche (resp. $C_i \rightarrow \alpha C_i$ - à droite)	$D_i(\alpha)^{-1} = D_i(\frac{1}{\alpha})$
$J_r(n,p) = \begin{pmatrix} I_r & O_{r,p-r} \\ O_{n-r,r} & O_{n-r,p-r} \end{pmatrix}$	Matrice J_r	$\text{rg}(A) = r \iff \exists P, Q \in GL_n(\mathbb{K}) \times GL_p(\mathbb{K})$ tel que $A = P \times J_r \times Q$	

Retour sur les problèmes

25. Addition naturelle des éléments de mêmes caractéristiques (donc situés dans une même case)

26. L'interprétation en terme de graphes qui figure dans la marge donne des éléments de réponse à ce problème.
$[A]_{ij}$ est le coefficient de dépendance de la moyenne de l'élève i par rapport à la note j.
$[B]_{j,k}$ est le poids de la partie k dans la note j.
Alors $[AB]_{i,k}$ est le coefficient de dépendance de la moyenne de l'élève i par rapport aux parties k.

27. La matrice de la question n'admet aucune racine. On démontre qu'on devrait la chercher parmi les matrices triangulaires avec des zéros sur la diagonale : $a = \begin{pmatrix} 0 & \alpha & \beta \\ 0 & 0 & \gamma \\ 0 & 0 & 0 \end{pmatrix}$.
On a alors $a^2 = \begin{pmatrix} 0 & 0 & \alpha\gamma \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$. Il n'y a pas de racine.
En revanche, le même calcul montre que pour tout $\alpha \in \mathbb{R}^*$ et $\beta \in \mathbb{R}$, $a = \begin{pmatrix} 0 & \alpha & \beta \\ 0 & 0 & \frac{1}{\alpha} \\ 0 & 0 & 0 \end{pmatrix}$ est une racine de $a = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$, qui admet donc une double infinité de racines…

28. L'anneau des matrices n'est pas commutatif, les éléments ne sont pas tous inversibles ni régulier. Il y a des diviseurs de 0.
Les seules matrices qui commutent avec toutes les matrices sont les λI_n.
($AE_{i,j} = \sum_k [A]_{k,i}E_{k,j}$ et $E_{i,j}A = \sum_{\mathbb{K}} [A]_{j,k}E_{i,k} \Rightarrow [A]_{i,j} = 0$ si i $\neq j$ et $[A]_{i,i} = [A]_{j,j}$)
Essentiellement (à peu de choses près…) si B commute avec A, alors il existe $P \in \mathbb{K}[X]$ tel que $B = P(A)$…

29. Oui si A n'est pas inversibles, alors $\text{Ker}(A) \neq \emptyset$. Il existe une colonne X tel que $AX = 0$.
    Si $B = (X|X|\cdots|X)$ alors $AB = 0$.
    Pour étudier l'inversibilité de A, on peut étudier son noyau. On peut aussi exploiter la méthode de Gauss-Jordan.
30. Voir cours.

Deuxième partie

Techniques analytiques, à travers l'histoire

Chapitre 7

Fonctions *à la Euler*

Résumé -

Dans ce chapitre, on s'intéresse aux principales fonctions usuelles des mathématiques pré-eulerien (donc, jusqu'à la fonction Γ*, exclue).*
Nous reprenons les constructions historiques qui conduisent à ces fonctions, en espérant qu'ainsi, les propriétés caractéristiques de chacune seront mieux mémorisées.
Pour préparer le cours sur la dérivation des fonctions usuelles, nous donnerons une série d'inégalités localisées suffisantes pour calculer les dérivées.
Enfin, nous terminons ce chapitre en donnant une liste d'évaluations numériques des fonctions usuelles sous forme de somme infinies (convergentes). Nous en reparlerons (beaucoup) plus tard.

— *Kahn Académy - Qu'est-ce qu'une fonction exponentielle ?. https ://www.youtube.com/watch ?v=pBeGfLoId4I*
— *Micmath - Merveilleux logarithmes. https ://www.youtube.com/watch ?v=rWfl7Pw8YVE*

Sommaire

1. Problèmes

Histoire - Evolution de la notion de fonction

Pour Leibniz puis Euler, qui a été le premier à utiliser le mot de « fonction » et pour les mathématiciens du XVIII-ième siècle, l'idée de relation fonctionnelle était plus ou moins assimilée à l'existence d'une formule mathématique simple exprimant la nature exacte de cette relation. Cette conception s'est révélée trop étroite pour les exigences de la physique mathématique, et l'idée de fonction, ainsi que la notion de limite qui lui est associée, ont subi un long processus de clarification et de généralisation.

Par exemple : $x \mapsto \begin{cases} 1 & si\ x \leqslant 1 \\ x^2 & si\ x > 1 \end{cases}$ *n'est pas une fonction pour Euler; mais pour nous (et vous?) c'est bien une fonction!*

? Problème 30 - Fonction par morceaux

L'application $x \mapsto \begin{cases} 1 & \text{si } x \leqslant 1 \\ x^2 & \text{si } x > 1 \end{cases}$ est-elle bien une fonction (à la « Euler ») ?

? Problème 31 - Comment calculer π^e ?

Pour l'étude des fonctions usuelles, il est très important pour chacun d'être en mesure de savoir comment faire le calcul des valeurs, sans raccourci avec la calculatrice.
Il faudra néanmoins faire ces calculs pour les nombres rationnels (et pour les nombres réels, on verra plus loin...).
Les nombres π et e sont des nombres réels bien définis.
Comment faire ce calcul? Les nombres réels sont approchés par des nombres rationnels.

On va donc commencer par essayer de calculer $\left(\frac{22}{7}\right)^{\frac{8}{3}}$, puis de manière générale de r^z où $r, z \in \mathbb{Q}$.
Deux temps d'analyse :

1. On fixe $z \in \mathbb{Q}$ et on étudie $t \mapsto t^z$, pour $t \in \mathbb{R}$. Ce sont les fonctions puissances.
2. On fixe $r \in \mathbb{Q}$ puis $r \in \mathbb{R}$ et on étudie $t \mapsto r^t$, pour $t \in \mathbb{R}$. Ce sont les fonctions exponentielles.

? Problème 32 - Interpolation de la suite géométrique

Existe-t-il une fonction simple (polynomiale) qui interpole la suite géométrique de raison r (= 2 par exemple) ?
Comment l'étudier?

? Problème 33 - De la multiplication à l'addition

Additionner deux nombres de tailles n se fait en gros en $2n$ calculs. Pour les multiplier l'algorithme classique nécessite n^2 multiplications de chiffres, puis une addition de n nombres...
Peut-on trouver un moyen simple qui transmute les multiplications en additions? Une fonction telle que : $f(a \times b) = f(a) + f(b)$?

2. Généralités sur les fonctions

2.1. Définition et représentation d'une fonction

Définition - Fonction

Une *fonction* d'une variable réelle, à valeurs réelles, est un "procédé" qui à chaque élément x d'un sous-ensemble $\mathscr{D}$ de $\mathbb{R}$ associe un réel $f(x)$ parfaitement déterminé. Une telle fonction est notée

$$\begin{array}{rl} f: & \mathscr{D} \to \mathbb{R} \\ & x \mapsto f(x) \end{array}$$

$f(x)$ est appelé *image* de x par f.
Si $y \in \mathbb{R}$, tout élément x de $\mathscr{D}$ vérifiant $f(x) = y$ est appelé un *antécédent* de y.

Remarque - Ensemble de définition
Le fait d'écrire $f : \mathscr{D} \to \mathbb{R}$ sous-entend que $f(x)$ est parfaitement défini si $x \in \mathscr{D}$.
Il est toutefois fréquent en pratique que l'ensemble $\mathscr{D}$ ne soit pas connu a priori. A partir de l'expression de $f(x)$, il faut déterminer l'ensemble des réels x pour lesquels elle a un sens. Cet ensemble est appelé *domaine de définition* de f (généralement noté $\mathscr{D}_f$). Il est souvent constitué d'une réunion d'intervalles.

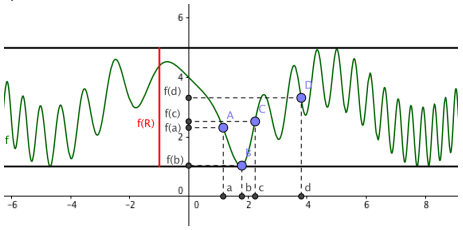
Définition - Graphe d'une fonction
$\mathscr{C} = \{(x, f(x)); x \in \mathscr{D}_f\}$ s'appelle le *graphe* de f, le plan étant muni d'un repère $(O, \vec{i}, \vec{j})$, on appelle *représentation graphique* ou *courbe représentative de f* la représentation de cet ensemble dans le plan. La courbe représentative a pour équation $y = f(x)$.

Définition - Image de $\mathscr{D}$ par une fonction. Fonction restreinte
Si $\mathscr{D}' \subset \mathscr{D}_f$ ($\mathscr{D}'$ sous ensemble de $\mathscr{D}_f$), on note $f(\mathscr{D}')$ l'ensemble des images des éléments de $\mathscr{D}'$: $f(\mathscr{D}') = \{f(x); x \in \mathscr{D}'\}$. $f(\mathscr{D}')$ s'appelle *l'ensemble image (ou image directe)* de $\mathscr{D}'$ par f.
Soit $f : \mathscr{D}_f \to \mathbb{R}$ une fonction et I un intervalle inclus dans $\mathscr{D}_f$. On appelle *restriction* de f à I la fonction g définie sur I par : $\forall x \in I, g(x) = f(x)$. On la note $f_{|I}$.

Définition - Ensemble des applications
Soit I un intervalle de $\mathbb{R}$, non vide et non réduit à un point.
On note $\mathscr{F}(I, \mathbb{R})$ l'ensemble des fonctions définies sur I à valeurs dans $\mathbb{R}$ (on parle aussi d'applications de I dans $\mathbb{R}$)

Savoir faire - Transformation sur le graphe
Soient $f : \mathscr{D}_f \to \mathbb{R}$ et $a \in \mathbb{R}$.
Comment obtient-on les domaines de définition ainsi que les graphes (ou représentations graphiques) des fonctions $x \mapsto f(x) + a$, $x \mapsto f(x + a)$, $x \mapsto f(a - x)$, $x \mapsto af(x)$?
$x \mapsto f(x) + a : \mathscr{D}_1 = \mathscr{D}$ et translation de $a\vec{j}$,
$x \mapsto f(x + a) : \mathscr{D}_2 = \{x \mid x + a \in \mathscr{D}\} = \mathscr{D} - a$ et translation de $-a\vec{i}$,
$x \mapsto f(a - x) : \mathscr{D}_3 = \{x \mid a - x \in \mathscr{D}\}$ (ex : si $\mathscr{D} = [c, d[$, alors $\mathscr{D}_3 =]a - d, a - c]$) et reflexion (symétrie) d'axe d'équation $x = \frac{a}{2}$ $x \mapsto af(x) : \mathscr{D}_4 = \mathscr{D}$ et homotétie-axiale de rapport a et de « centre » l'axe des abscisses.

Représentation - Fonction



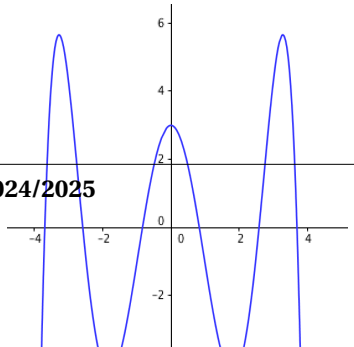
Pour tracer on commence bien par tous les points en noir : $x_i, f(x_i)$, on en déduit $A_i(x_i, f(x_i))$. Puis on relie tous ces points A_i, une courbe apparait !
L'image de f est ici donné en rouge : c'est le segment $[1, 5]$.

Exercice
Montrer qu'une suite (u_n) est également une fonction

Correction
Il s'agit d'une application de $\mathbb{N}$ dans $\mathbb{R}$, avec $u : n \mapsto u_n$

2.2. Opérations sur les fonctions

Représentation - Fonction paire



Définition - Opérations classiques
Dans $\mathscr{F}(I,\mathbb{R})$ on définit les opérations suivantes :
— Addition : si $(f,g)\in\mathscr{F}(I,\mathbb{R})^2$, la fonction $f+g$ est définie sur I par

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

— Multiplication par un réel : si $f\in\mathscr{F}(I,\mathbb{R})$ et $\lambda\in\mathbb{R}$, la fonction λf est définie sur I par

$$\forall x\in I,\ (\lambda f)(x)=\lambda f(x)$$

— Produit de deux fonctions : si $(f,g)\in\mathscr{F}(I,\mathbb{R})^2$, la fonction fg est définie sur I par

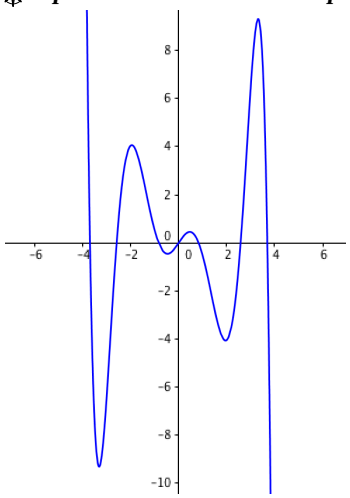
$$\forall x\in I,\ (fg)(x)=f(x)\times g(x)$$

— Valeur absolue d'une fonction : si $f\in\mathscr{F}(I,\mathbb{R})$ la fonction $|f|$ est définie sur I par

$$\forall x\in I,\ |f|(x)=|f(x)|$$

Remarque - Lois internes...
Addition et multiplication de deux applications sont des lois internes sur $\mathscr{F}(I,\mathbb{R})$, commutatives, associatives, l'application nulle est élément neutre pour l'addition et toute application f admet un opposé : l'application $-f$, l'application constante égale à 1 est élément neutre pour la multiplication, la multiplication est distributive par rapport à l'addition. (On parle d'anneau pour décrire un ensemble ayant deux telles lois internes).

Représentation - Fonction impaire



Définition - Composée
Soient $f:I\to\mathbb{R}$, $g:J\to\mathbb{R}$ vérifiant $f(I)\subset J$. On définit la composée $g\circ f:I\to\mathbb{R}$ par

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

En général, même lorsque les deux fonctions $g\circ f$ et $f\circ g$ ont un sens, elles sont différentes.

Exercice

Définir proprement les deux composées (si cela est possible) des fonctions suivantes (ou de leurs restrictions) :

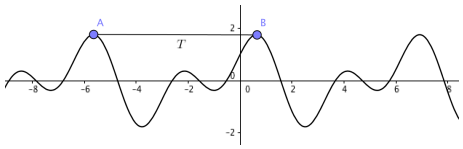
$$f:\begin{array}{l}\mathbb{R}\to\mathbb{R}\\ x\mapsto x^2-1\end{array}\quad\text{et}\quad g:\begin{array}{l}\mathbb{R}\setminus\{0\}\to\mathbb{R}\\ x\mapsto\frac{1}{x}\end{array}$$

Correction

$$f\circ g:\begin{array}{l}\mathbb{R}^*\to\mathbb{R}\\ x\mapsto\frac{1}{x^2}-1\end{array}\quad\text{et}\quad g\circ f:\begin{array}{l}\mathbb{R}\setminus\{-1,1\}\to\mathbb{R}\\ x\mapsto\frac{1}{x^2-1}\end{array}$$

2.3. Vocabulaire d'analyse de fonctions

Représentation - Fonction périodique de période T



Parité, périodicité

Définition - Vocabulaire (et propriétés)
Soit $f:\mathscr{D}_f\to\mathbb{R}$ une fonction
• On dit que f est *paire* si

$$\forall x\in\mathscr{D}_f,\ -x\in\mathscr{D}_f\text{ et }f(-x)=f(x)$$

$\mathscr{C}_f$ est alors symétrique par rapport à Oy
• On dit que f est *impaire* si

$$\forall x\in\mathscr{D}_f,\ -x\in\mathscr{D}_f\text{ et }f(-x)=-f(x)$$

$\mathscr{C}_f$ est alors symétrique par rapport à O.
• $f \in \mathscr{F}(\mathbb{R},\mathbb{R})$ est dite périodique s'il existe $T>0$ tel que

$$\forall x \in D_f,\ x+T \in D_f, x-T \in D_f \text{ et } f(x+T)=f(x).$$

T est une période de f. $\mathscr{C}_f$ est alors invariante par translation de vecteur $T\vec{i}$.

On laisse les démonstrations au lecteur.

Fonctions monotones

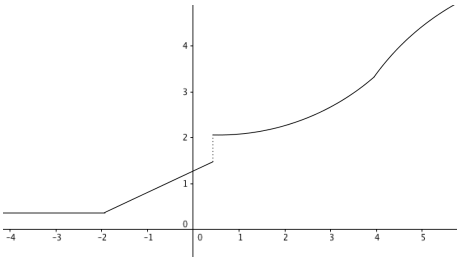
Définition - Fonction croissante, décroissante et monotone
On dit qu'une fonction f
- est croissante sur I si $\forall (x,y) \in I^2,\ x \leqslant y \Rightarrow f(x) \leqslant f(y)$.
- est décroissante sur I si $\forall (x,y) \in I^2,\ x \leqslant y \Rightarrow f(x) \geqslant f(y)$.
- est monotone sur I si elle est croissante sur I ou décroissante sur I.
- est *strictement* croissante sur I si $\forall (x,y) \in I^2,\ x < y \Rightarrow f(x) < f(y)$.
- est *strictement* décroissante sur I si $\forall (x,y) \in I^2,\ x < y \Rightarrow f(x) > f(y)$.

Proposition - Composition et monotonie
La composée de deux applications monotones de même sens de variation (respectivement de sens contraire) est une application croissante (respectivement décroissante).

La démonstration est simple...

Démonstration
Supposons que f et g sont décroissants.
Soit $x \leqslant y \in \mathscr{D}_{g\circ f}$, alors :
$f(x) \geqslant f(y)$
$(g\circ f)(x) \leqslant (g\circ f)(y)$.
Donc $g\circ f$ est décroissante...□

Représentation - Fonction strictement croissante



Fonction (a priori) croissante (et non dérivable!).

Fonctions convexes/concaves

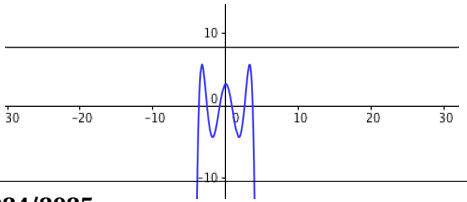
Définition - Fonction convexe, concave
On dit qu'une fonction f
- est convexe sur I si
  $\forall (x,y) \in I^2, \forall\ \lambda \in [0,1]\ f(\lambda x+(1-\lambda)y) \leqslant \lambda f(x)+(1-\lambda)f(y)$.
- est strictement convexe sur I si
  $\forall (x,y) \in I^2, \forall\ \lambda \in]0,1[\ f(\lambda x+(1-\lambda)y) < \lambda f(x)+(1-\lambda)f(y)$.
- est concave sur I si
  $\forall (x,y) \in I^2, \forall\ \lambda \in [0,1]\ f(\lambda x+(1-\lambda)y) \geqslant \lambda f(x)+(1-\lambda)f(y)$.
- est strictement concave sur I si
  $\forall (x,y) \in I^2, \forall\ \lambda \in]0,1[\ f(\lambda x+(1-\lambda)y) > \lambda f(x)+(1-\lambda)f(y)$.

Pour aller plus loin - Fonctions convexes
Nous étudierons tout particulièrement les fonctions convexes dans un prochain chapitre

Fonctions (ou applications) bornées

Définition - Fonctions majorées, minorées, bornées
On dit qu'une fonction f :
- est majorée s'il existe $M \in \mathbb{R}$ tel que $\forall x \in \mathscr{D}_f,\ f(x) \leqslant M$.
- est minorée s'il existe $m \in \mathbb{R}$ tel que $\forall x \in \mathscr{D}_f,\ m \leqslant f(x)$.
- est bornée si elle est majorée et minorée.

Représentation - Fonction majorée, minorée?



Fonction (a priori) majorée et non minorée, donc non bornée.

Proposition - Bilan
Une application $f : I \to \mathbb{R}$ est bornée
si et seulement si il existe $M > 0$ tel que $\forall x \in I, |f(x)| \leqslant M$.

Extremums

Définition - Maximum, minimum et extremum
On dit que $M \in \mathbb{R}$ est un maximum sur I de f si
- $\forall x \in I, f(x) \leqslant M$
- et il existe $a \in I$ tel que $f(a) = M$
  on dit que f présente un maximum en a.

On dit que $m \in \mathbb{R}$ est un minimum sur I de f si
- $\forall x \in I, f(x) \geqslant m$
- et il existe $a \in I$ tel que $f(a) = m$
  on dit que f présente un minimum en a.

On parle d'extremum lorsque l'on a un maximum ou un minimum. On note

$$M = \max_{x \in I} f(x) \text{ et } m = \min_{x \in I} f(x)$$

Définition - Maximum local
On dit que $M = f(a)$ est un *maximum local* (maximum ou minimum)
s'il existe $\epsilon > 0$ tel que pour tout $x \in]a-\epsilon, a+\epsilon[\cap D_f$, $f(x) \leqslant f(a)$.
On parle de *maximum strict*, lorsque pour $x \in]a-\epsilon, a+\epsilon[\cap D_f$ et $x \neq a$, $f(x) < f(a)$.

Remarque - Voisinage de a (dans $\mathbb{R}$
Pour décrire un tel ensemble V, on parle souvent de voisinage de a.
Ainsi la restriction de f à $V =]a-\epsilon, a+\epsilon[\cap D_f$ présente en a un extremum (stricte) respectivement).

Remarque - Extension de définition
La définition s'étend à minimum local et minimum local strict et à maximum local et maximum local strict.

Représentation - Exemple
Sur la représentation de la fonction majorée non minorée, on peut voir trois maximums (maxima) locaux, tous stricts et deux sont aussi maximum global...

2.4. Bijections et réciproques

On commence par rappeler ce qu'est une fonction bijective avant de voir le rôle joué ici par la dérivée.

Bijection

Définition - Bijection
Soit f une fonction définie sur un sous-ensemble D de $\mathbb{R}$ à valeurs dans $\mathbb{R}$ et $J \subset \mathbb{R}$.
On dit que f est bijective de D sur J (ou réalise une bijection de D sur J)
— si tout élément de D a son image dans J
— et si tout élément de J admet un unique antécédent par f dans D
Formellement :

$$\forall x \in D, f(x) \in J \quad \text{et} \quad \forall y \in J, \exists! x \in D, y = f(x).$$

Remarque - Si, par convention $J = f(D)$
Si l'on pose $J = f(D) = \{f(x); x \in D\} = \{y \in \mathbb{R} \mid \exists x \in D, y = f(x)\}$,
alors f est une bijection de D sur $J = f(D)$ si et seulement si $\forall y \in J, \exists! x \in D,\ y = f(x)$.

Exemple - Application exponentielle
L'application exponentielle est une bijection de $\mathbb{R}$ sur $\mathbb{R}_+^*$

Définition - Application (bijection) réciproque
Si f est bijective de D sur J, on définit une fonction g par

$$\begin{array}{rll} g: & J \to D & \\ & y \mapsto x & \mid y = f(x) \text{ (unique antécédent de } y \text{ par } f) \end{array}$$

Cette fonction g est elle-même bijective et appelée bijection réciproque de f, et notée f^{-1}.

Exemple - Application logarithmique
L'application logarithmique (naturelle) est une bijection de $\mathbb{R}_+^*$ sur $\mathbb{R}$. C'est la bijection réciproque de la fonction exponentielle.

Proposition - Application directe
Si f est une bijection de D sur J, on a

$$\forall x \in D, (f^{-1} \circ f)(x) = x;$$

$$\forall y \in J, (f \circ f^{-1})(y) = y;$$

$$y = f(x) \Leftrightarrow x = f^{-1}(y).$$

Théorème - Réciproque et représentation graphique
Soit $f : I \to J$ une bijection et $f^{-1} : J \to I$ sa bijection réciproque. Alors
- dans un repère orthonormé, $\mathscr{C}_f$ et $\mathscr{C}_{f^{-1}}$ sont symétriques par rapport à la première bissectrice (droite d'équation $y = x$).
- Si f est monotone sur I alors f^{-1} est monotone sur J, de même sens de variations.

On verra des exemples plus loin.

Démonstration
Pour montrer la symétrie dans le graphique, il nous faut alors montrer l'équivalence entre :
$M(x, y) \in \mathscr{C}_f$ et $M'(y, x) \in \mathscr{C}_{f^{-1}}$.
Or
$$M' \in \mathscr{C}_{f^{-1}} \Leftrightarrow f^{-1}(y) = x \Leftrightarrow y = f(x) \Leftrightarrow M \in \mathscr{C}_f$$
Concernant la monotonie, supposons que f soit croissante, :
considérons $x \leqslant y \in J$,
si $f^{-1}(x) > f^{-1}(y)$, alors par croissance de f : $x > y$. Absurde.
donc, nécessairement : $f^{-1}(x) \leqslant f^{-1}(y)$ et f^{-1} est croissante.
De même si f^{-1} est croissante, alors sa réciproque, i.e. f est également croissante.
Le cas décroissant est laissé en exercice.□

On admet les théorèmes qui suivent et qui seront démontrés ultérieurement :

Théorème - Théorème de la bijection
Soit f une fonction définie sur I, continue, strictement monotone sur I (intervalle de $\mathbb{R}$),
alors f est bijective de I sur $J = f(I)$.
Sa bijection réciproque f^{-1} est continue sur J.

2.5. Etude des branches infinies en ∞

Le but est de préciser l'allure de la courbe représentative de f au voisinage de (+ ou -) l'infini.

Savoir faire - Etude des branches infinies (et définition)

Soit f une fonction définie au voisinage de $\pm\infty$.

1. si $\lim\limits_{x\to\infty} f(x) = \ell$ avec $\ell \in \mathbb{R}$, la courbe admet une **asymptote horizontale** d'équation $y = \ell$.
2. si $\lim\limits_{x\to\infty} f(x) = \infty$, on calcule $\dfrac{f(x)}{x}$.
   (a) si $\lim\limits_{x\to\infty} \dfrac{f(x)}{x} = 0$, il y a une **branche parabolique horizontale** (ou de direction Ox).
   (b) si $\lim\limits_{x\to\infty} \dfrac{f(x)}{x} = \infty$, il y a une **branche parabolique verticale** (ou de direction Oy).
   (c) si $\lim\limits_{x\to\infty} \dfrac{f(x)}{x} = a$, $a \in \mathbb{R}^*$, il faut calculer $f(x) - ax$.
      i. si $\lim\limits_{x\to\infty} f(x) - ax = b \in \mathbb{R}$, la courbe admet une **asymptote oblique** d'équation $y = ax + b$.
      ii. si $\lim\limits_{x\to\infty} f(x) - ax = \infty$ il y a une **branche parabolique oblique de direction** $y = ax$.

Exercice

Énoncer des fonctions présentant :

1. une asymptote horizontale (on donnera l'équation de cette asymptote)
2. une branche parabolique horizontale
3. une branche parabolique verticale
4. une asymptote oblique (on donnera l'équation de cette asymptote)
5. une branche parabolique oblique (on donnera la direction), mais pas d'asymptote oblique

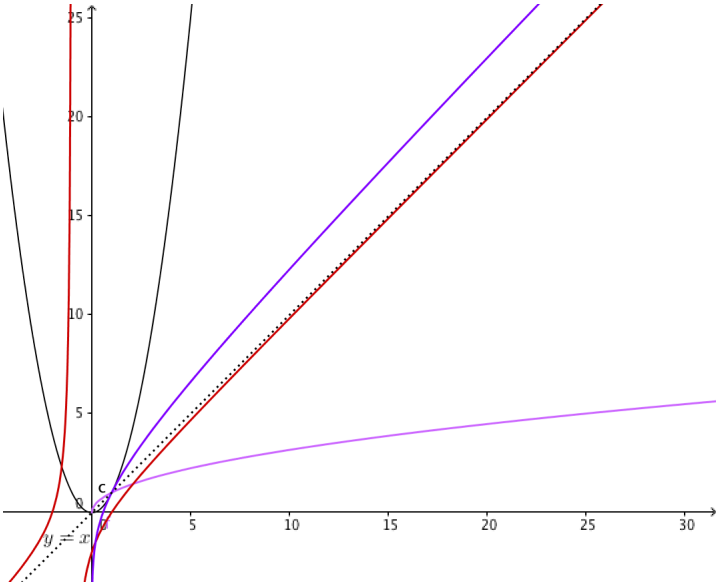
---

Correction

1. $y = \arctan(x)$ admet une asymptote horizontale : $y = \frac{\pi}{2}$
2. $y = \sqrt{x}$ admet une branche parabolique horizontale
3. $y = x^2$ admet une branche parabolique verticale
4. $y = \dfrac{2x^2+x-2}{x+1} (= 2x - 1 - \frac{1}{x+1})$ admet une asymptote oblique d'équation $y = 2x - 1$
5. $y = 3x + \ln x$ admet une branche parabolique oblique d'équation $y = 3x$.

Exercice

Indiquer sous chacun des quatre graphiques suivants lesquels présentent des branches paraboliques, des asymptotes...



Correction

En noir : branche parabolique verticale, en bleu une branche parabolique oblique d'équation $y = x$, en rouge un courbe avec asymptote d'équation $y = x$ et en violet une courbe avec une direction parabolique horizontale.

Remarque - Courbe asymptote

Plus généralement on dit que la courbe représentative de f admet au voisinage de $+\infty$ (ou $-\infty$ ou un réel a) admet pour courbe asymptote la courbe d'équation $y = g(x)$ si $\lim\limits_{x\to+\infty} f(x) - g(x) = 0$.

3. Fonctions trigonométriques

On reprend, de manière analytique (où le paramètre x devient une variable) les fonctions trigonométriques vues précédemment.

3.1. Fonctions circulaires

Proposition - Aspect analytique

Les fonctions sinus et cosinus sont 2π-périodiques.
sin est une fonction impaire alors que cos est une fonction paire.

Savoir faire - Transférer un problème trigonométrique « en a », vers « en 0 »

Il faut exploiter les formules trigonométriques : $\sin(a+h) = \sin a \cos h + \cos a \sin h$ et $\cos(a+h) = \cos a \cos h - \sin a \sin h$, à connaître par coeur.

Analyse - Inégalité fondamentale

On termine cette partie par une observation essentielle :

Rappelons que θ est la longueur de l'angle

(on peut aussi comparer l'aire du triangle BCA', égale à $\dfrac{1\times CA}{2} = \dfrac{1}{2}\sin\theta$

à l'aire (plus grande) de la portion du disque BCA égale à $\dfrac{\theta}{2\pi}\times\pi 1^2 = \dfrac{1}{2}\theta$.)

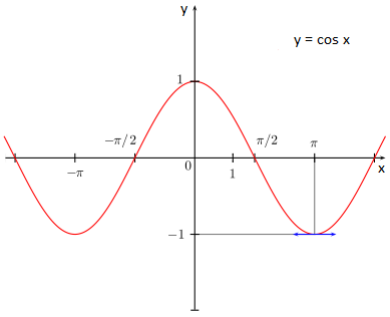
Pour la seconde inégalité, on calcule deux aire :

— l'aire de la portion du disque BAC : elle vaut $\frac{\theta}{2}R^2 = \dfrac{\theta}{2}$ car $R = 1$

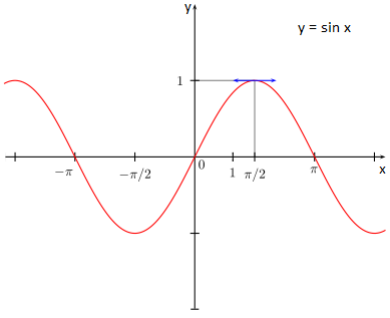
— l'aire du triangle rectangle $BA'C'$: elle vaut $\frac{1}{2}BA'A'C' = \dfrac{\tan\theta}{2}$.

« Clairement » la seconde aire est plus petite que la première.

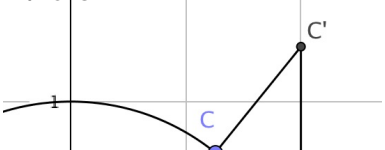
Représentation - Fonction cosinus



Représentation - Fonction sinus



Représentation - Cercle trigonométrique



Proposition - Inégalité
On a pour tout $x \in]-\frac{\pi}{2}, \frac{\pi}{2}[$,

$$|\sin x| \leqslant |x| \leqslant |\tan x| = \frac{|\sin x|}{\cos x}$$

Et en particulier $\lim\limits_{x\to 0} \dfrac{\sin x}{x} = 1$

Démonstration
On a vu (dans l'analyse) que pour $x \in]0, \frac{\pi}{2}[$, $\sin x \leqslant x \leqslant \tan x$.
Puis en multipliant par $-1 < 0$: $-\sin x \geqslant -x \geqslant -\tan x$.
et par imparité si $x_- \in]-\frac{\pi}{2}, 0[$: $\sin(x_-) = -\sin(-x_-) \geqslant -(-x_-) \geqslant -\tan(-x_-) = \tan(x_-)$.
Ainsi pour $x \in]-\frac{\pi}{2}, \frac{\pi}{2}[$: $|\sin(x)| \leqslant |x| \leqslant |\tan(x)| = \dfrac{|\sin x|}{\cos x}$
Donc en divisant par $|x|$ et en séparant en deux inégalités :

$$\left|\frac{\sin x}{x}\right| \leqslant 1 \qquad \text{et} \qquad \cos x \leqslant \left|\frac{\sin x}{x}\right|$$

Par encadrement : $\dfrac{\sin x}{x} \underset{x\to 0}{\longrightarrow} 1$.
Ainsi sin est dérivable en 0 de dérivée égale à 1. □

Exemple - Calculatrice. Calculer $\sin(0,01234)$
On obtient $\sin(0,01234) = 0,01233968682$

Exercice
En majorant le module de $e^{ix} - 1$, montrer que pour tout $x \in \mathbb{R}$, $\sin^2 x + (\cos x + 1)^2 \leqslant x^2$.

Correction
Soit $x \in \mathbb{R}$, la factorisation par l'angle moitié donne : $e^{ix} - 1 = e^{i\frac{x}{2}}(2i \sin \frac{x}{2})$, donc

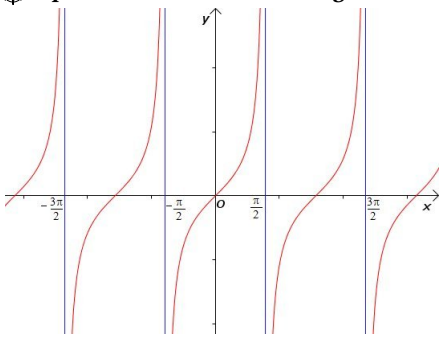
$$|e^{ix} - 1| = 1 \times 2|\sin \frac{x}{2}| \leqslant |x|$$

Si on élève au carré :

$$(\cos x - 1)^2 + \sin^2 x \leqslant x^2$$

Proposition - Fonction tangente - Aspect analytique
La fonction tangente est ainsi définie sur $\mathbb{R} \setminus \{\frac{\pi}{2} + k\pi; k \in \mathbb{Z}\}$ ($\mathbb{R}$ privé des points de la forme $\frac{\pi}{2} + k\pi$ avec $k \in \mathbb{Z}$).
tan est impaire, π-périodique

Représentation - Fonction tangente



Démonstration
Il s'agit de la division de deux fonctions dérivables. Sur $\mathbb{R} \setminus \{\frac{\pi}{2} + k\pi; k \in \mathbb{Z}\}$ (le cosinus ne s'annule pas) : □

Exercice
Etudier et représenter la fonction tan

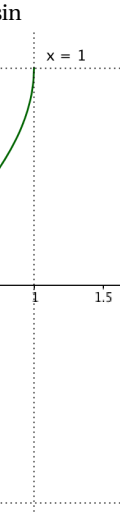
Correction
La fonction tan est définie sur $\mathscr{D} = \{x \in \mathbb{R} \mid \cos\theta \neq 0\} = \mathbb{R} \setminus \{\frac{\pi}{2} + k\pi; k \in \mathbb{Z}\}$.
La fonction tan est π-périodique d'après la formule trouvée plus haut.
On peut l'étudier sur $]-\frac{\pi}{2}, \frac{\pi}{2}[$, puis exploiter des translations de vecteurs $\pi\vec{i}$.
Par division de deux fonctions dérivables, tan est dérivable sur son ensemble de définition.

$$\forall\, x \in \mathscr{D}, \qquad \tan'(x) = \frac{\cos^2(x) + \sin^2(x)}{\cos^2(x)} = 1 + \tan^2(x) = \frac{1}{\cos^2(x)}$$

La fonction tan est donc strictement croissante sur les intervalles de la forme $]\dfrac{-\pi}{2}, \dfrac{-\pi}{2}[$ et à valeurs dans $]-\infty, \infty[$.
$\tan(0) = 0$ et la pente de la tangente au point 0a pour pente : $\tan'(k\pi) = \dfrac{1}{\cos(k\pi)} = 1$.
Enfin, comme $\dfrac{1}{\cos^2}$ est croissante sur $[0, \frac{\pi}{2}[$, tan est convexe sur $[0, \frac{\pi}{2}[$ et $\dfrac{1}{\cos^2}$ est décroissante sur $]-\frac{\pi}{2}, 0]$, tan est concave sur $]-\frac{\pi}{2}, 0]$.
On obtient la représentation graphique suivante :

3.2. Fonctions circulaires réciproques

Fonction arcsin



Définition - Arcsinus
La fonction sinus est continue et strictement croissante sur $[-\frac{\pi}{2},\frac{\pi}{2}]$ donc réalise une bijection de $[-\frac{\pi}{2},\frac{\pi}{2}]$ sur $[-1,1]$.
La bijection réciproque s'appelle arcsinus, $\arcsin : [-1,1] \to [-\frac{\pi}{2},\frac{\pi}{2}]$.
Elle est impaire, strictement croissante. On a donc :

$$t = \arcsin x \Leftrightarrow \left(\sin t = x \text{ et } t \in [-\frac{\pi}{2},\frac{\pi}{2}]\right)$$

Proposition - Rappels
On a :
- $\forall x \in [-\frac{\pi}{2},\frac{\pi}{2}]$, $\arcsin(\sin x) = x$
- $\forall x \in [-1,1]$, $\sin(\arcsin x) = x$
- $\forall x \in [-1,1]$, $\cos(\arcsin x) = \sqrt{1-x^2}$
- $\forall x \in]-1,1[$, $\tan(\arcsin x) = \frac{x}{\sqrt{1-x^2}}$

Analyse - Questions
Comment démontrer tout cela?
Il faut connaître les valeurs remarquables suivantes :

x	0	$\frac{1}{2}$	$\frac{1}{\sqrt{2}} = \frac{\sqrt{2}}{2}$	$\frac{\sqrt{3}}{2}$	1
$\arcsin x$	0	$\frac{\pi}{6}$	$\frac{\pi}{4}$	$\frac{\pi}{3}$	$\frac{\pi}{2}$

Exercice
Comparer $\arcsin x$ et x, à partir de la double inégalité fondamentale du sinus.

Correction
On a vu pour tout $x \in]-\frac{\pi}{2},\frac{\pi}{2}[$, $|\sin x| \leqslant |x| \leqslant |\tan x|$.
Posons $x = \arcsin u$, pour $u \in]-1,1[$; donc $x \in]-\frac{\pi}{2},\frac{\pi}{2}[$.

$$|u| \leqslant |\arcsin u| \leqslant |\tan(\arcsin u)| = \frac{|u|}{\sqrt{1-u^2}}.$$

Fonction arccos

Définition - Arccosinus
La fonction cosinus est continue et strictement décroissante sur $[0,\pi]$ donc réalise une bijection de $[0,\pi]$ sur $[-1,1]$.
La bijection réciproque s'appelle arccosinus, $\arccos : [-1,1] \to [0,\pi]$.
Elle est strictement décroissante et on a donc :

$$t = \arccos x \Leftrightarrow \left(\cos t = x \text{ et } t \in [0,\pi]\right)$$

Proposition - Rappels
On a :
- $\forall x \in [0,\pi]$, $\arccos(\cos x) = x$
- $\forall x \in [-1,1]$, $\cos(\arccos x) = x$
- $\forall x \in [-1,1]$, $\sin(\arccos x) = \sqrt{1-x^2}$
- $\forall x \in [-1,1], x \neq 0$, $\tan(\arccos x) = \frac{\sqrt{1-x^2}}{x}$
- $\forall x \in [-1,1]$, $\arcsin x + \arccos x = \frac{\pi}{2}$

Il faut connaître les valeurs remarquables suivantes :

x	0	$\frac{1}{2}$	$\frac{1}{\sqrt{2}} = \frac{\sqrt{2}}{2}$	$\frac{\sqrt{3}}{2}$	1
$\arccos x$	$\frac{\pi}{2}$	$\frac{\pi}{3}$	$\frac{\pi}{4}$	$\frac{\pi}{6}$	0

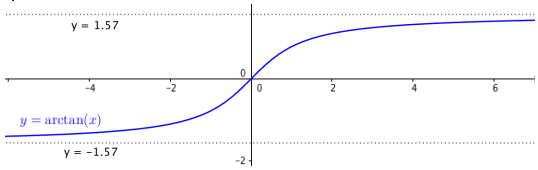
Fonction arctan

Définition - Arctangente
La fonction tangente est continue et strictement croissante sur $]-\frac{\pi}{2},\frac{\pi}{2}[$ donc réalise une bijection de $]-\frac{\pi}{2},\frac{\pi}{2}[$ sur $\mathbb{R}$.
La bijection réciproque s'appelle arctangente, $\arctan : \mathbb{R} \to]-\frac{\pi}{2},\frac{\pi}{2}[$.
Elle est impaire, strictement croissante et on a donc :

$$t = \arctan x \Leftrightarrow \left(\tan t = x \text{ et } t \in]-\frac{\pi}{2},\frac{\pi}{2}[\right)$$

Proposition - Rappels
On a :
- $\forall x \in]-\frac{\pi}{2},\frac{\pi}{2}[$, $\arctan(\tan x) = x$
- $\forall x \in \mathbb{R}$, $\tan(\arctan x) = x$
- $\forall x \in \mathbb{R}$, $\cos(\arctan x) = \dfrac{1}{\sqrt{1+x^2}}$
- $\forall x \in \mathbb{R}$, $\sin(\arctan x) = \dfrac{x}{\sqrt{1+x^2}}$
- $\forall x \in \mathbb{R}$, $\arctan x + \arctan \frac{1}{x} = \begin{cases} \frac{\pi}{2} & \text{si } x > 0 \\ -\frac{\pi}{2} & \text{si } x < 0 \end{cases}$

Représentation - Fonction arctan



Il faut connaître les valeurs remarquables suivantes :

x	0	$\frac{1}{\sqrt{3}} = \frac{\sqrt{3}}{3}$	1	$\sqrt{3}$
$\arctan x$	0	$\frac{\pi}{6}$	$\frac{\pi}{4}$	$\frac{\pi}{3}$

4. Fonctions polynomiales et puissances rationnelles

On reprend, de manière analytique (où le paramètre x devient une variable) les fonctions puissances et polynomiales vues précédemment.

Pour aller plus loin - Besoin d'inégalités
Pour étudier les limites variées (à l'infini, ou calculer des dérivées), on a ***toujours*** *besoin d'encadrement.*
Dans chaque partie, on trouvera des inégalités importantes

4.1. Fonction puissance entière relative

Définition

Définition - Puissance entière > 0
Soit $n \in \mathbb{N}^*$, on qualifie de fonction puissance entière l'application $x \mapsto x^n$, i.e. définie par récurrence par $x \mapsto x \times x^{n-1}$ et $x^0 = 1$.
Son ensemble de définition est $\mathbb{R}$. C'est une fonction continue (nous le verrons plus tard).
Cette application est paire si n est pair, et impaire si n est impair.

nces et puis-

it chez Bom-
5), Descartes

Il faut savoir représenter ces fonctions.

Définition - Puissance entière < 0
Soit $m \in \mathbb{Z}_-$, on qualifie de fonction puissance entière négative l'application $x \mapsto x^m = \dfrac{1}{x^{-m}}$, i.e. définie par récurrence par $x \mapsto \dfrac{1}{x} \times x^{m+1}$ et $x^0 = 1$.
Son ensemble de définition est $\mathbb{R}^*$. C'est une fonction continue sur $\mathbb{R}_+^*$ et sur $\mathbb{R}_-^*$.
Cette application est paire si n est pair, et impaire si n est impair.

Il faut savoir représenter ces fonctions, en particulier l'hyperbole $\mathscr{C}$ associé à $x \mapsto \frac{1}{x}$ $(m = -1)$.

Proposition - Morphisme
On a pour tout $m, n \in \mathbb{Z}$ et $x \in \mathbb{R}^*$, $x^m \times x^n = x^{m+n}$.
Vrai également en $x = 0$, si $n, m > 0$.

Exercice
A démontrer

Correction
On fixe $m \in \mathbb{Z}$, et on fait une récurrence sur $n \in \mathbb{N}$.
Puis, on étudie le cas $n < 0$, avec $-m \in \mathbb{Z}$: $x^{-m} \times x^{-n} = x^{-m+-n} \dots$

Inégalités de croissance

Fixons $n \in \mathbb{N}$.
Pour la croissance de $x \mapsto x^n$, on exploite le binôme de Newton, par exemple :

Proposition - Binôme de Newton. Croissance
Soit $n \in \mathbb{N}$. Pour tout $a \in \mathbb{R}$, $(a+x)^n = \displaystyle\sum_{k=0}^{n} \binom{n}{k} a^{n-k} x^k$.
Avec $a, x > 0$, on a donc $x < x' \Rightarrow x^n < x'^n$, soit la croissance de $x \mapsto x^n$ sur $\mathbb{R}_+$.

Démonstration
$a = x' - x > 0$, $x'^n = x^n + \underbrace{\displaystyle\sum_{k=0}^{n-1} \binom{n}{k} a^{n-k} x^k}_{>0} > x^n$. □

Exercice
Faire la démonstration par récurrence.
On fera bien attention aux variables fixées et celles qui sont libres

Correction
On peut fixer x et x' dans $\mathbb{R}_+$, par exemple : $0 \leqslant x < x'$.
Posons, pour tout entier $n \in \mathbb{N}^*$, $\mathscr{P}_n$: « $0 \leqslant x^n < x'^n$. »
- $\mathscr{P}_1$ est vraie par hypothèse.
- Soit $n \in \mathbb{N}^*$. Supposons que $\mathscr{P}_n$ est vraie.
  Comme $x \geqslant 0$: $x^{n+1} = x^n \times x \leqslant x'^n \times x$, d'après $\mathscr{P}_n$.
  Puis comme $x < x'$ et $x'^n > 0$: $x'^n \times x < x'^n \times x' = x'^{n+1}$.
  Donc $\mathscr{P}_{n+1}$ est vraie.

La récurrence est démontrée : $\forall\ n \in \mathbb{N}^*$, $0 \leqslant x^n < x'^n$.
Ceci est vraie pour tout $0 \leqslant x < x'$, on a donc :
$\forall\ (x, x', n) \in \mathbb{R}_+ \times \mathbb{R}_+^* \times \mathbb{N}^*$ tel que $x < x\prime$: $x^n < x'^n$,
On a donc encore : $\forall\ n \in \mathbb{N}^*$, $\forall\ 0 \leqslant x < x'$, $x^n < x'^n$, donc $t \mapsto t^n$ est strictement croissante sur $\mathbb{R}_+$.
Peut-on se passer de $0 \leqslant$ *dans l'écriture de* $\mathscr{P}_n$ *?*

Inégalités de Bernoulli

Proposition - Inégalités de Bernoulli
Pour tout $n \in \mathbb{N}^*$, pour tout $x \in]-1, +\infty[$, $(1+x)^n \geqslant 1+nx$.
Pour tout $n \in \mathbb{N}^*$, pour tout $x \in]-1, \frac{1}{n}[$, $(1+x)^n \leqslant \dfrac{1}{1-nx}$

Démonstration
Par récurrence, on note $\mathscr{P}_n$: « $\forall\, x > -1$, $(1+x)^n \geqslant 1+nx$ et $\forall\, x \in [-1, \frac{1}{n}]$, $(1+x)^n \leqslant \dfrac{1}{1-nx}$ ».

— Le résultat est vrai pour $n=0$ (et $n=1$ (immédiat)).
Donc $\mathscr{P}_0$ est vraie.

— Soit $n \in \mathbb{N}$. Supposons que $\mathscr{P}_n$ est vraie.
Soit $x > -1$. On a donc $(1+x)^{n+1} = (1+x)^n \times (1+x) \leqslant (1+nx)(1+x) = 1+(n+1)x+nx^2$,
car : $1+x>0$ et $nx^2 \geqslant 0$.
De même si $x \in]-1, \frac{1}{n+1}[\subset]-1, \frac{1}{n}[$, $(1+x)^{(n+1)} = (1+x)\times(1+x)^n \leqslant \dfrac{1+x}{1-nx}$ car $1+x>0$.
Or $(1+x)(1-(n+1)x) = 1-nx-(n+1)x^2 \leqslant 1-nx$ car $(n+1)x^2 \geqslant 0$.
Ainsi, en divisant par $(1-(n+1)x)(1-nx) \geqslant 0$: $(1+x)^{(n+1)} \leqslant \dfrac{1+x}{1-nx} \leqslant \dfrac{1}{1-(n+1)x}$
Donc $\mathscr{P}_{n+1}$ vraie.

On obtient ainsi la première et la troisième inégalité (en composant par $t \mapsto \frac{1}{t}$ décroissante sur $\mathbb{R}+$).
□

L'exercice suivant permet de montrer la continuité des fonctions puissances.
Exercice
On fixe $n \in \mathbb{N}$.Montrer que $\lim_{x\to 0+}(1+x)^n = 1$, puis la continuité à droite de $t \mapsto t^n$ en 1 et en tout $x \in \mathbb{R}$.

Correction
Soit n fixé. Pour tout $x \in [0, \frac{1}{n}]$,

$$1+nx \leqslant (1+x)^n \leqslant \frac{1}{1-nx}$$

Par encadrement : même limite à gauche et à droite, pour $x \to 0^+$, on a donc $(1+x)^n \underset{x\to 0^+}{\longrightarrow} 1$. Soit $a \in \mathbb{R}$ et $h>0$, alors $(a+h)^n = a^n(1+\frac{h}{a})^n$, ainsi si $0 \leqslant \frac{h}{a} \leqslant \frac{1}{2}$,
on a : $a^n(1+n\frac{h}{a}) \leqslant (a+h)^n = a^n(1+\frac{h}{a})^n \leqslant a^n(1+2n\frac{h}{a})$.
Là encore, pour $h \to 0^+$, on a donc $(a+h)^n \to a^n$, ce qui conduit à la continuité à droite de $t \mapsto t^n$.

4.2. Fonctions polynomiales

Définition - Fonction polynomiale
On appelle fonction polynomiale une fonction de la forme :

$$f : x \mapsto a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n = \sum_{k=0}^{n} a_k x^k$$

où $\forall\, k \in [\![0, n]\!]$, $a_k \in \mathbb{R}$ (ou $\mathbb{C}$) ;
Il s'agit d'une combinaison linéaire finie de puissances entières de la variable x. On parle de polynôme simple ou de polynôme à une variable.
On dit que $f(x) = b$ est une équation polynomiale si f est une fonction polynomiale.

Remarque - Notations
On remarque que les lettres de fin d'alphabet sont en générale associées à des inconnues. Les lettres de début d'alphabet aux variables connues.
On comprend pour les inconnues : on ne peut pas faire autrement. Mais pourquoi associer des lettres à des nombres connus?
Par propriétés calculatoires sur $\mathbb{R}$ ou $\mathbb{C}$:

Proposition - Propriété de l'ensemble des fonctions polynomiales
Si f et g sont deux fonctions polynomiales, alors :
— $f+g$ est une fonction polynomiale
— $f\times g$ est une fonction polynomiale
— $f\circ g$ est une fonction polynomiale.

Démonstration
Il suffit d'écrire les calculs. Cela est lié aux propriétés des anneaux $\mathbb{R}$ et $\mathbb{C}$.
Pour la composition, c'est plus subtile. On note que, par récurrence, grâce au résultat sur les produits :
Pour tout g, polynomiale, tout $k\in\mathbb{N}^*$: $g^k = g^{k-1}\times g$ est également polynomiale.
puis d'après le résultat sur la somme, par récurrence : $\sum_{k=0}^{n} a_k g^k(x) = f\circ g$ est polynomiale.
□

Rappelons :

Théorème - Factorisation multiple
Soit f une fonction polynomiale de degré n. Soit $p\leqslant n$
Si $x_1, x_2, \dots x_p$, p solutions différentes de l'équation $f(x)=0$ (racines de f).
Alors il existe g_p, fonction polynomiale de degré $n-p$ tel que

$$\forall\, x\in\mathbb{K},\qquad f(x) = (x-x_1)(x-x_2)\dots(x-x_p)\times g_p(x)$$

Corollaire - Nombre maximal de solution
Une équation polynomiale de degré n admet au plus n solutions différentes

4.3. Fonction puissance rationnelle

Définition

Analyse - Bijection de $x\mapsto x^n$ sur $\mathbb{R}_+$
Pour tout $n\in\mathbb{N}^*$, $x\mapsto x^n$ est strictement croissante sur $I=\mathbb{R}_+$ et à valeurs dans $J=\mathbb{R}_+$.
Elle est continue et donc admet une application réciproque de $J=\mathbb{R}_+$ sur $I=\mathbb{R}_+$.
Dans le cas où n est impair, on peut élargir la définition de $I=\mathbb{R}$ sur $J=\mathbb{R}$.

Définition - Racine n-ième
On note $\sqrt[n]{\cdot}$ la bijection réciproque de $x\mapsto x^n$.
On a donc :

$$\text{si } n \text{ est pair}\quad \sqrt[n]{x} = x^{1/n} \quad \text{pour } x\geqslant 0$$
$$\text{si } n \text{ est impair}\quad \sqrt[n]{x} = \begin{cases} x^{1/n} & \text{pour } x\geqslant 0 \\ -|x|^{1/n} = -(-x)^{1/n} & \text{pour } x\leqslant 0 \end{cases}$$

Pour n impair on a donc $\sqrt[n]{-x} = -\sqrt[n]{x}$.

Définition - Puissance rationnelle
Soit $r=\frac{p}{q}\in\mathbb{Q}^*$ (avec $q\in\mathbb{N}^*$, $p\in\mathbb{Z}$), on qualifie de fonction puissance rationnelle l'application $x\mapsto x^r$ où x^r vérifie $(x^r)^q = x^p$.
Son ensemble de définition est $\mathbb{R}_+^*$. C'est une fonction croissante et continue (nous l'admettons à ce stade).

Exemple - $5^{2/3}$

Il s'agit du nombre y tel que $y^3 = 5^2 = 25$. (Existe-t-il un et un seul nombre y qui vérifie cette équation $y^3 = 25$?).

Par dichotomie (et croissance) : $2^3 = 8$ et $3^3 = 27$, donc $y \in]2,3[$. Et l'on peut chercher à préciser...

$2,9^3 = 24,389$. Donc $y \in]2,9;3[$...

Inégalités complémentaires

Analyse - Image de $[0,1[$ **et image de** $]1,+\infty[$ **par** $x \mapsto x^r$

Si $x > 1$, alors pour $n, m \in \mathbb{N}$: $x^n > 1$.

Et pour tout $y < 1$, $y^m < 1$.

Et nécessairement y tel que $y^m = x^n > 1$ vérifie $y > 1$.

Proposition - Comparaison des fonctions puissances

Soient $r < r' \in \mathbb{Q}$.

Alors pour tout $x > 1$, $x^r < x^{r'}$. Et si $x < 1$, $x^r > x^{r'}$

Démonstration

La soustraction $r' - r = r''$ donne un nombre rationnel, positif.

On a alors, par positivité de x^r :

$x^r < x^{r'} \iff x^{r''} = x^{r'-r} > 1$.

Or $x > 1$ et $r'' \in \mathbb{Q}^+$, donc $x^{r''} > 1$.

Si $x < 1$, on raisonne avec $\frac{1}{x} > 1$.

□

Attention - Ne pas confondre les variables x **et** n

Dès maintenant, on fait bien attention lorsqu'on compare à x fixé x^r et x^s,

et lorsqu'on compare à r fixé x^r et x'^r.

5. Exponentielles et logarithmes

Heuristique - Histoire

Les mathématiciens ont d'abord rencontrer les fonctions logarithmiques (STEVIN, BRIGGS, NEPER) au XVIeme siècle.

Ils cherchaient un processus pour transformer multiplication (complexe) en addition (plus simple).

Il s'agissait d'interpoler la réciproque des suites géométriques : $n \mapsto a^n$, vérifiant $a^{n+m} = a^n \times a^m$.

Pour faciliter les démonstrations du cours, nous remonterons l'histoire (d'abord exponentielles avant logarithmes).

5.1. ExponentielleS

Heuristique - Equation fonctionnelle

Soit $a \in \mathbb{R}$, si $n, m \in \mathbb{N}$, $a^{n+m} = a^n \times a^m$.

Cette relation est centrale si l'on s'intéresse à $x \mapsto a^x$.

Considérons donc $f : \mathbb{R} \to \mathbb{R}$ vérifiant pour tous nombres réels $x, y \in \mathbb{R}$: $f(x+y) = f(x) \times f(y)$.

Est-ce qu'une telle relation est suffisante pour définir parfaitement aucune (non car $t \mapsto 2^t$ semble bien aller, au moins pour $t \in \mathbb{Q}$), une fonction f, ou plusieurs? Et dans ce cas, que

Pour aller plus loin - Notation

Dans Quadrature n°137, on trouve les notations $\cdot$ pour $+$, $\cdot$ pour $\times$ et $^{\cdot}$ pour une puissance. On trouve alors $a^{\cdot p} \cdot a^{\cdot q} = a^{\cdot(p.q)}$ ce qui s'interprète comme $a^p \times a^q = a^{p+q}$ et autres relations automatisées du même genre.

, rajouter pour différencier ces différentes fonctions?

On notera le pluriel :

Définition - Fonctions exponentielles
On qualifie de fonctions exponentielles les applications $f : \mathbb{R} \to \mathbb{R}$, non nulles vérifiant :
$\forall\, x, y \in \mathbb{R} : f(x+y) = f(x) \times f(y)$.

Proposition - Propriété commune à toutes les exponentielles
Si f est une fonction vérifiant : $f(x+y) = f(x) \times f(y)$, alors f est à valeurs dans $\mathbb{R}_+$.
Puis : ou bien $f : x \mapsto 0$, ou bien $f(0) = 1$.
Donc une fonction exponentielle est à valeurs dans $\mathbb{R}_+$ et $f(0) = 1$.

Démonstration
Pour tout $x \in \mathbb{R}$, $f(x) = f(\frac{x}{2} + \frac{x}{2}) = \left(f(\frac{x}{2})\right)^2 \geqslant 0$.
$f(x) = f(0+x) = f(0) \times f(x)$.
Si il existe x tel que $f(x) \neq 0$, alors en divisant par $f(x) : f(0) = 1$.
Sinon, pour tout $x \in \mathbb{R}$, $f(x) = 0$ et on a bien (réciproquement) $f(x+y) = 0 = f(x) \times f(y)$. □

Proposition - Base
Si f est une fonction exponentielle non nulle, alors il existe un nombre $a \in \mathbb{R}_+$, tel que pour tout $x \in \mathbb{Q}$, $f(x) = a^x$

Savoir faire - Etude d'une équation fonctionnelle de $\mathbb{N}$ à $\mathbb{R}$
L'étude d'une équation fonctionnelle se fait souvent de la façon suivante :

1. Par récurrence, en étudiant $f(sn)$ pour $n \in \mathbb{N}$ et s quelconque.
2. Par imparité/parité (ou autre symétrie), on étudie $f(sm)$ pour $m \in \mathbb{Z}$.
3. On retrouve ensuite le résultat pour $m \in \mathbb{Q}$.
4. Ensuite, on exploitera (plus tard) un argument de continuité ou bien un argument de croissance

Démonstration
Considérons une fonction exponentielle, non nulle.
Soit $s \in \mathbb{R}$, fixé.
Posons, pour tout $n \in \mathbb{N}$, $\mathscr{P}_n$: « $f(sn) = (f(s))^n$.
— Le résultat est vraie pour $n = 0$ (car $f(s)^0 = 1 = f(0)$) et $n = 1$.
— Soit $n \in \mathbb{N}$. Supposons que $\mathscr{P}_n$ est vraie.
Alors $f(s(n+1)) = f(sn+s) = f(sn) \times f(s) = (f(s))^{n+1}$ d'après $\mathscr{P}_n$.
Soit $t \in \mathbb{R}$, $f(t+(-t)) = f(0) = 1 = f(t) \times f(-t)$, donc $f(-t) = \dfrac{1}{f(t)}$.
Ainsi, pour $t \leftarrow sn$, alors $f(s \times (-n)) = \dfrac{1}{f(sn)} = \dfrac{1}{(f(s))^n} = (f(s))^{-n}$ Ainsi, $s \leftarrow 1$ et $a = f(1) > 0$,
on a pour tout $m \in \mathbb{Z}$, $f(m) = f(1m) = f(1)^m = a^m$.
Considérons, alors $r = \frac{p}{q}$ avec $p \in \mathbb{Z}$ et $q \in \mathbb{N}^*$.
on a $f(qr) = f(p) = a^p$ et $f(qr) = f(r)^q$ avec $s \leftarrow r$ et $n \leftarrow q$.
on a donc, en prenant la racine q-ième : $f(r) = a^{\frac{p}{q}} = a^r$ □

Proposition - Variations
Soi f une fonction exponentielle est non nulle,
f est strictement croissante sur $\mathbb{Q}$ si $f(1)(= a) > 1$
et elle est décroissante sur $\mathbb{Q}$ si $f(1)(= a) < 1$

Démonstration
Supposons que $f(1) = a > 1$. Soient $r_1, r_2 \in \mathbb{Q}$, avec $r_1 \leqslant r_2$. Notons $r = r_2 - r_1 \in Q \cap]0, +\infty[$
$f(r_2) = f(r_1 + r) = f(r_1) \times f(r)$. Or $f(r) = a^r > 1$ car $a > 1$.
Donc $f(r_2) > f(r_1)$, ainsi f est croissante sur $\mathbb{Q}$.
De même si $f(1) < 1$.
□

Pour aller plus loin - Suite décimale
Prenons un exemple pour mieux comprendre ici.
Considérons le nombre π, on a alors $\pi = 3,1415\ldots$.
Dans ce cas : $d_0 = 3$, $d_1 = 3,1$, $d_2 = 3,14$, $d_3 = 3,141\ldots$

Analyse - Comment définir $f(x)$ pour $x \in \mathbb{R}$

Il reste à étendre la définition pour des nombres $x \in \mathbb{R}$.
Considérons le développement décimal de x, il existe (on le démontrera plus tard) une suite d_n tel que pour tout $n \in \mathbb{N}$, $d_n \leqslant x < d_n + 10^{-n}$ avec $(d_n)_n$ croissante et $(d_n + 10^{-n})_n$ décroissante de limite x.
Supposons que $a > 1$.
La suite $(a^{d_n})_{n\in\mathbb{N}}$ est une suite croissante, majorée par a^{d_0+1}, elle converge.
On note $f(x) = a^x$, cette limite.

Exercice
On note $a = f(1)$. Montrer que pour tout $x \in \mathbb{R}$ et $r \in \mathbb{Q}$, $f(rx) = f(x)^r$.
Quelle formule obtient concernant les puissances de a ?

Correction
En reprenant pour $m \in \mathbb{Z}$, $f(mx) = (f(x))^m$, puis pour $r \in \mathbb{Q}$, $f(rx) = f(x)^r$.
On a donc $a^{rx} = (a^x)^r$. Puis en passant à la limite pour $(r_n) \to x'$: $a^{xx'} = (a^x)^{x'}$.

Pour aller plus loin - Définition de a^x
Il faudrait démontrer que cette valeur ne dépend pas de la suite (d_n) choisie, convergente vers x.

On résume et admet les derniers résultats (il nous manque la continuité) :

Théorème - Fonctions exponentielles. Bilan
Les fonctions exponentielles non nulles sont continues.
Elles vérifient : $f(0) = 1$ et pour tout $x, y \in \mathbb{R}$, $f(x + y) = f(x) \times f(y)$ et $f(x \times y) = f(x)^y$ $(y \in \mathbb{Q})$.
Il existe $a(= f(1)) \in \mathbb{R}$ tel que $f(x) = a^x$ (par définition de a^x si $x \in \mathbb{R} \setminus \mathbb{Q}$).
Si $a > 1$, alors f est strictement croissante sur $\mathbb{R}$, avec $\lim\limits_{-\infty} f = 0$ et $\lim\limits_{+\infty} f = +\infty$.
Si $a < 1$, alors f est strictement décroissante sur $\mathbb{R}$, avec $\lim\limits_{-\infty} f = +\infty$ et $\lim\limits_{+\infty} f = 0$.

5.2. LA fonction exponentielle

Nous verrons que ces fonctions sont dérivables. L'une a la propriété essentielle de vérifier $f'(0) = 1$. C'est LA fonction exponentielle avec $a = e$ (LE « e »). Prenons un autre définition.

Pour aller plus loin - Critère de convergence pour une suite réelle
Nous verrons, sans cercle vicieux, que toute suite de nombres réelles, majorée et croissante (à partir d'un certain rang) est convergente.
C'est une propriété caractéristique de $\mathbb{R}$.

Définition - La fonction exponentielle naturelle
Soit $x \in \mathbb{R}$. La suite $(x_n) = ((1 + \frac{x}{n})^n)$ est majorée, croissante à partir d'un certain rang, donc convergente.
Notons $\exp(x)$ la limite de (x_n).

Il faut démontrer la convergence de la suite :

Démonstration
• Positivité à partir d'un certain rang.
Pour tout réel x, notons que $\frac{x}{n} \to 0$ pour $n \to +\infty$.
Donc à partir d'un certain rang N_x tel que $\forall\, n \geqslant N_x$, $-\frac{1}{2} \leqslant \frac{x}{n} \leqslant \frac{1}{2}$.
Ainsi, pour tout $n \geqslant N_x$, $x_n > 0$. On notera plus simplement que ce rang N_x est $\max(0, \lceil -x \rceil)$.
• Croissance (à partir d'un certain rang)
On peut utiliser le critère d'une suite positive (à partir d'un certain rang) : par comparaison relative à 1).

$$\frac{x_{n+1}}{x_n} = \left(1 + \frac{x}{n}\right)\left(\frac{\frac{n+1+x}{n+1}}{\frac{n+x}{n}}\right)^{n+1} = \left(\frac{n+x}{n}\right)\left(\frac{n^2+nx+n}{n^2+nx+n+x}\right)^{n+1}$$

$$\frac{x_{n+1}}{x_n} = \left(\frac{n+x}{n}\right)\left(1 - \frac{x}{(n+1)(n+x)}\right)^{n+1}$$

Rappelons l'inégalité de Bernoulli : $(1+a)^m \leqslant 1+ma$ si $a > -1$ et $m \in \mathbb{N}^*$.

Prenons alors $m \leftarrow n+1 (\in \mathbb{N}^*)$ et $a \leftarrow -\dfrac{x}{(n+1)(n+x)}$.

On a bien : $a > -1$, car :

- si $x \leqslant 0$, $x < n+x$ donc $x < (n+1)(n+x)$ dès que $n > 0$.
- si $x < 0$, $-x > 0$ et $(n+1)(n+x) > 0$ dès que $n > -x$, i.e. $n \leqslant N_x$.

Donc :

$$\left(1 - \frac{x}{(n+1)(n+x)}\right)^{n+1} \leqslant 1 + (n+1)\frac{-x}{(n+1)(n+x)} = \frac{n}{n+x}$$

Enfin, comme $\dfrac{n+x}{n} \leqslant 0$, le sens de l'inégalité est inchangé :

$$\frac{x_{n+1}}{x_n} \leqslant \frac{n+x}{n} \times \frac{n}{n+x} = 1$$

• Majoration.

Soit $s = \lfloor |x| \rfloor + 1 \in \mathbb{N}$, donc $\frac{x}{sn} \in]-\frac{1}{n}, \frac{1}{n}[$.

$$x_{sn} = \left(\left(1 + \frac{x}{sn}\right)^n\right)^s \leqslant \left(\frac{1}{1 - n\frac{x}{sn}}\right)^s = \left(\frac{s}{s-x}\right)^s \text{ (constante en } n\text{).}$$

(La fonction $t \mapsto t^s$ est croissante sur $\mathbb{R}_+$.)

Soit $m \geqslant \max(N_x, N'_x)$, et $n = \lfloor \frac{m}{s} \rfloor + 1$

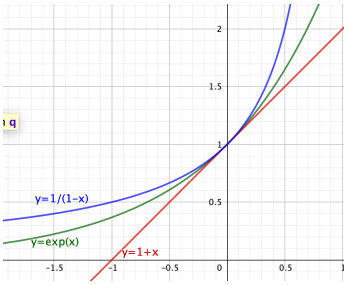
on a $n \geqslant \frac{m}{s}$, donc $N'_x \leqslant m \leqslant ns$.

puis par croissance de (x_n) pour tout $m \geqslant N'_x$: $x_m \leqslant x_{ns} \leqslant \left(\dfrac{s}{s-x}\right)^s$ □

Proposition - Inégalités

On a pour tout $x \in]-1, 1[$, $1 + x \leqslant \exp x \leqslant \dfrac{1}{1-x}$.

Représentation - Inégalités



Remarque - Elargissement de l'intervalle

En fait, le résultat est vrai pour tout $x \in \mathbb{R}$ pour la première inégalité et sur $]-\infty, 1[$ pour la seconde.

Mais en réalité, elles nous serviront surtout pour x proche de 0, où les trois termes valent 1…

Démonstration

On applique les inégalités de Bernoulli, à partir d'un certain rang, puisque $\frac{x}{n} \to 0$ ($\frac{x}{n} < \frac{1}{n}$) :

$$1 + x = 1 + n\frac{x}{n} \leqslant \left(1 + \frac{x}{n}\right)^n = x_n \leqslant \frac{1}{1 - n\frac{x}{n}} = \frac{1}{1-x}$$

On passe ensuite à la limite : à gauche et à droite les termes sont constants. Au centre, la suite converge vers e^x. □

Théorème - exp est une fonction exponentielle.

La fonction $x \mapsto \exp x$ est une fonction exponentielle appellée LA fonction exponentielle.

On a donc pour tout $x \in \mathbb{R}$, $\exp(x) = e^x$ où $e = \exp(1) = \lim\limits_{n\to+\infty}\left(1 + \dfrac{1}{n}\right)^n$

Elle vérifie donc : $\forall\, x, y \in \mathbb{R}$, $\exp(x+y) = e^{x+y} = e^x e^y = \exp(x) \times \exp(y)$ et $\exp(x \times y) = e^{xy} = (e^x)^y = (\exp(x))^y$.

Démonstration

Plus compliqué. Soient $x, y \in \mathbb{R}$.

Notons $x_n = \left(1 + \frac{x}{n}\right)^n$, $y_n = \left(1 + \frac{y}{n}\right)^n$ et $z_n = \left(1 + \frac{x+y}{n}\right)^n$.

Alors $\dfrac{x_n y_n}{z_n} = \dfrac{\left(1 + \frac{x}{n} + \frac{y}{n} + \frac{xy}{n^2}\right)^n}{\left(1 + \frac{x+y}{n}\right)^n} = \left(1 + \dfrac{t_n}{n}\right)^n$,

avec $t_n = \dfrac{xy}{n+x+y} \to 0$ pour $n \to +\infty$.

Il existe donc une valeur N tel que pour tout $n \geqslant N$, $t_n \in [-\frac{1}{2}, \frac{1}{2}]$, donc $\frac{t_n}{n} \in \subset [-\frac{1}{n}, \frac{1}{n}[\subset [-1, \frac{1}{n}[$

On a alors $1 + t_n \leqslant \dfrac{x_n y_n}{z_n} \leqslant \dfrac{1}{1-t_n}$ et donc par unicité des limites : $\dfrac{\exp(x)\exp(y)}{\exp(x+y)} = \lim \dfrac{x_n y_n}{z_n} = 1$.

Donc $\exp(x+y) = \exp(x)\exp(y)$.

exp est une fonction exponentielle. □

Application - Evaluation approchée de $(1+\frac{1}{30})^{100}$
Si l'on considère $n = 100$ proche de l'infini, on a donc

$$\left(1+\frac{1}{30}\right)^{100} = \left(1+\frac{\frac{10}{3}}{100}\right)^{100} \approx \exp\left(\frac{10}{3}\right)$$

On vérifie à la calculatrice : $\exp\frac{10}{3} = 28,03162$ et $(1+\frac{1}{30})^{100} = 26,548739$.

Analyse - Binôme de Newton pour $(1+\frac{1}{n})^n$ et une approximation d'Euler
L'argument d'Euler, à partir du binôme de Newton donne :

$$\begin{aligned}(1+\frac{1}{n})^n &= \sum_{k=0}^{n}\binom{n}{k}\frac{1}{n^k} = 1+\frac{n}{n}+\frac{n(n-1)}{2!n^2}+\frac{n(n-1)(n-2)}{3!n^3}+\dots\\ &= 1+1+\frac{1(1-\frac{1}{n})}{2!}+\frac{1(1-\frac{1}{n})(1-\frac{2}{n})}{3!}+\dots\end{aligned}$$

Euler ajoute : « *si n est un nombre plus grand qu'aucune quantité assignable, la fraction $\frac{n-1}{n}$ égalera l'unité* » et conclue par $e = 1+1+\frac{1}{2!}+\frac{1}{3!}+\dots$.
La méthode laisse à désirer, mais le résultat est juste :

Proposition - Formules d'Euler (1746)

$$e = \lim_{n\to+\infty}\left(1+\frac{1}{n}\right)^n = \sum_{k=0}^{+\infty}\frac{1}{k!}$$

Et pour tout $x \in \mathbb{R}$,

$$e^x = \lim_{n\to+\infty}\left(1+\frac{x}{n}\right)^n = \sum_{k=0}^{+\infty}\frac{x^k}{k!}$$

Histoire - D'où vient la notation e ?
Leonard EULER *(1707-1783) est le mathématicien le plus prolixe de l'histoire (avec Cauchy ?).*
C'est un calculateur de génie, doté d'une mémoire prodigieuse (hypemnésique).

5.3. LogarithmeS

Les fonctions exponentielles non constante égale à 0 ou 1 sont continues et strictement monotone, elles admettent donc une fonction récipropque.

Histoire - Neper
L'écossais John Napier (1550-1617) (ou Neper) cherche au XV siècle une fonction qui faciliterait les calculs : elle transformerait le produit (compliqué car beaucoup de calculs) en addition (moins de calculs).
Il trouve le logarithme.
On a donc $\ln(ab) = \ln a + \ln b$.
L'histoire peut être un moyen mnémotechnique.

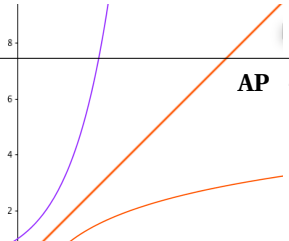
Définition - Fonctions logarithmes
On appelle fonctions logarithmes toute fonctions g réciproques de fonctions exponentielles f non constantes.
Si cette dernière est $x \mapsto a^x$ (avec $a \notin \{0,1\}$), alors la fonction logarithme est qualifiée « de base a ». On note souvent $g = \log_a$ ou $\ln_a$.
Elle est définie sur $\mathbb{R}_+$, à valeurs dans $\mathbb{R}$ et vérifie alors
$\forall\, x, y \in \mathbb{R}_+,\ g(x\times y) = g(x)+g(y)$.

Exemple - Différents logarithmes bien connus
Le logarithme en base 10 est un classique de la physique.
Il indique en gros la taille en nombre de chiffres.
Le logarithme en base 2 est un classique de l'informatique.

Démonstration
L'existence de g est liée à la bijectivité de f.
f étant à valeurs dans $\mathbb{R}_+$, on a pour tout $X = f(x) \in \mathbb{R}_+$ (x existe bien, c'est $g(X)$) et $Y = f(y) \in \mathbb{R}_+$: $g(X\times Y) = g(f(x)\times f(y)) = g(f(x+y)) = x+y = g(X)+g(Y)$. □

Théorème - Fonctions logarithmes. Bilan
Les fonctions logaritmes sont continues, définies sur $\mathbb{R}_+^*$, à valeurs dans $\mathbb{R}$.
Elles vérifient : $g(1) = 0$ et pour tout $x, y, t \in \mathbb{R}_+$, $g(x\times y) = g(x)+g(y)$ et $g(x^t) = t\times g(x)$.
Il existe $a(= f(1)) \in \mathbb{R}$ tel que $g(a) = 1$.
Si $a > 1$, alors g est strictement croissante sur $\mathbb{R}$, avec $\lim\limits_{0} g = -\infty$ et $\lim\limits_{+\infty} g = +\infty$.

Représentation - Exponentielles et logarithmes selon la valeur de a



Si $a < 1$, alors g est strictement décroissante sur $\mathbb{R}$, avec $\lim\limits_{0} g = +\infty$ et $\lim\limits_{+\infty} f = -\infty$.

Démonstration
Faisons juste la démonstration de $g(x^t)$:
On sait, pour $X = g(x)$ que : $f(t \times X) = f(X)^t$, i.e. $f(t \times g(x)) = x^t$. Puis en composant par g : $t \times g(x) = g(x^t)$. □

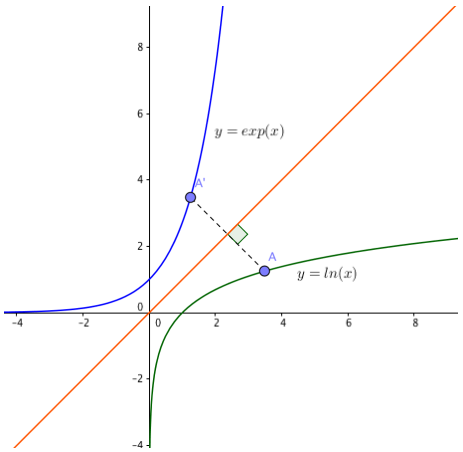
Savoir faire - Calculer « à la main » le logarithme en base a de x ?

On n'a pas d'autre solution (pour le moment) mais cela est suffisant (compte-tenu de la contrainte que l'on s'est donné) de procéder par dichotomie.
On regarde la suite (a^n) et l'on trouve $n \in \mathbb{N}$ tel que $a^n \leqslant x < a^{n+1}$.
Puis, on sélectionne $a_0 = n$ et $b_0 = n+1$. Puis on peut (par exemple), considérer $c = \frac{a_0+b_0}{2} \in \mathbb{Q}$, évaluer a^c.

Si $a^c < x$, on prend $a_1 = c$ et $b_1 = b_0$, sinon on prend $a_1 = a_0$ et $b_1 = c\ldots$

On a deux suites adjacentes (a_n) et (b_n) convergente vers y avec $a^y = x$, donc $y = \log_a(x)$.

Représentation - Exponentielle et logarithme naturels



Définition - Le logarithme naturel
La fonction logarithme réciproque de la fonction exponentielle, donc le logarithme en base $e = \exp 1$ est appelé logarithme naturel.
On le note ln.
ln est définie, continue et croissante sur $\mathbb{R}_+$, à valeurs dans $\mathbb{R}$.
Numériquement : $\ln(1) = 0$, $\ln e = 1$, $\lim\limits_{0} \ln = -\infty$ et $\lim\limits_{+\infty} \ln = +\infty$.
On a les propriétés algébrique : $\ln(ab) = \ln a + \ln b$, $\ln(a^b) = b \ln a\ldots$
On a pour tout $u \in \mathbb{R}_+^*$, $\ln u \leqslant u - 1$.

Histoire - Logarithme naturel
Il y a un abus historique ici.
On a démontré historiquement qu'il s'agit bien du logarithme naturel en le définissant comme primitive de $x \mapsto \frac{1}{x}$ (primitive dont on a démontré qu'il s'agissait d'un logarithme (Fermat - 1636)). C'est le naturel car $\ln'(1) = 1$.
La définition donnée ici vient de Euler (un siècle plus tard)

Tout est immédiat, sauf l'inégalité qu'on démontre :

Démonstration
On sait que pour tout $x \in]-1, +\infty[\subset \mathbb{R}$, $1 + x \leqslant \exp(x)$.
En composant par ln, croissante : $\ln(1+x) \leqslant x$.
Posons $u = 1 + x \in \mathbb{R}_+^*$: $\ln u \leqslant u - 1$
□

Proposition - Ecriture en fonction du logarithme naturel
Soit g la fonction logarithme de base a, réciproque de $f : x \mapsto a^x$.
Alors on a :

$$\forall\, x \in \mathbb{R},\ f(x) = a^x = \exp(x \times \ln a) \qquad \forall\, x \in \mathbb{R}_+,\ g(x) = \frac{\ln x}{\ln a}$$

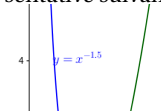
Démonstration
Comme ln est un logarithme : $\exp(x \ln a) = \exp(\ln a^x) = a^x$, car ln est réciproque de exp.
Puis, $G : x \mapsto \frac{\ln x}{\ln a}$ est une fonction logarithme

$$G(x \times y) = \frac{\ln(xy)}{\ln a} = \frac{\ln x + \ln y}{\ln a} = G(x) + G(y)$$

de base a, car $G(a) = \frac{\ln a}{\ln a} = 1$. Donc $G = g$ □

5.4. Retour sur les fonctions puissances, avec un exposant non rationnel

Représentation - Fonctions puissances réelles (différentes valeurs de a)
On en déduit les variations et la courbe représentative suivant les valeurs de a :



Définition - Fonction puissance réelle
Pour $\alpha \in \mathbb{R}$, on définit sur $]0,+\infty[$ la fonction puissance α par :
$$\begin{array}{rccc} g_\alpha : &]0,+\infty[& \to & \mathbb{R} \\ & x & \mapsto & x^\alpha = \exp(\alpha \ln x) \end{array}$$

Proposition - Fonction puissance réelle
Elle vérifie les propriétés suivantes :
$\forall (x,y) \in]0,+\infty[^2$, $\forall (\alpha,\beta) \in \mathbb{R}^2$,
- $(xy)^\alpha = x^\alpha y^\alpha$
- $x^{\alpha+\beta} = x^\alpha x^\beta$
- $(x^\alpha)^\beta = x^{\alpha\beta}$
- Si $\alpha = 0$, g_α est constante égale à 1.
- Si $\alpha > 0$, g_α est croissante et $\lim\limits_{x\to 0^+} g_\alpha(x) = 0$
- Si $\alpha < 0$, g_α est décroissante et $\lim\limits_{x\to 0^+} g_\alpha(x) = +\infty$

Pour $\alpha > 0$, on étudie l'existence d'une demi-tangente en 0 :
- Si $\alpha < 1$, la courbe $y = g_\alpha(x)$ admet une tangente verticale en 0.
- Si $\alpha = 1$, la courbe $y = g_\alpha(x)$ admet une tangente de pente 1 en 0.
- Si $\alpha > 1$, la courbe $y = g_\alpha(x)$ admet une tangente horizontale en 0.

5.5. Croissances comparées

Heuristique - Logarithme comme nombre de chiffres
Commençons par une remarque, avec $x = 10^n$, on a $\dfrac{\ln(x)}{x} = n10^{-n}\ln 10 \underset{n\to+\infty}{\longrightarrow} 0$.
Notons que $\ln 10 \approx 2{,}302$, donc $\ln x = \ln(10) \times \log_{10}(x)$ et que $\log_{10}(x)$ est en première approximation le nombre de chiffres de x, alors pour x grand, $\ln x$ est le nombre de chiffres de x multiplié par un peu plus de 2.
Exemple : $\ln(123456789) \approx 2{,}3 \times \log(1{,}23 \times 10^9) = 2{,}3 \times (9 + \ln(1{,}23)) \approx 20$, ce qui est petit

Il s'agit, a priori, de formes indéterminées. Elles sont levées :

Théorème - Croissance comparée
Soient a et b deux réels strictement positifs. On a
$$\lim_{x\to+\infty} \frac{(\ln x)^b}{x^a} = 0; \quad \lim_{x\to 0^+} x^a |\ln x|^b = 0;$$
$$\lim_{x\to+\infty} \frac{e^{ax}}{x^b} = +\infty; \quad \lim_{x\to-\infty} |x|^b e^{ax} = 0.$$

Démonstration
On commence par lever l'indétermination de $\dfrac{\ln(x)}{x}$ en $+\infty$. Tout en découlera…
Commençons par une remarque, avec $x = 10^n$, on a $\dfrac{\ln(x)}{x} = n10^{-n}\ln 10 \underset{n\to+\infty}{\longrightarrow} 0$.
On sait que pour tout $x > 0$, $\ln x \leqslant x - 1 < x$,
en particulier $\frac{1}{2}\ln x = \ln(\sqrt{x}) < \sqrt{x}$.
Donc, pour $x > 1$: $0 < \dfrac{\ln x}{x} < \dfrac{2\sqrt{x}}{x} = \dfrac{2}{\sqrt{x}}$. Par encadrement : $\lim\limits_{x\to+\infty} \dfrac{\ln x}{x} = 0$.
Pour $a, b > 0$,
$$\frac{(\ln x)^b}{x^a} = \left(\frac{\frac{b}{a}\ln(x^{a/b})}{x^{a/b}}\right)^b \underset{x\to+\infty}{\longrightarrow} 0$$
par composition de limites : avec $u = x^{a/b}$, puis $v \mapsto v^b$.
Avec cette nouvelle limite, en composant avec $x : t \mapsto \frac{1}{t}$,
$$x^a |\ln x|^b = \frac{|-\ln t|^b}{t^a} \underset{\substack{t\to+\infty \\ x\to 0}}{\longrightarrow} 0$$

Avec la première limite, en composant avec $x : t \mapsto \ln(t)$ i.e. $t = e^x$,

$$\frac{e^{ax}}{x^b} = \frac{t^a}{(\ln t)^b} \underset{\substack{t\to+\infty\\ x\to+\infty}}{\longrightarrow} +\infty$$

Avec la deuxième limite, en composant avec $x : t \mapsto \ln(t)$ i.e. $t = e^x$,

$$|x|^b e^{ax} = |\ln t|^b t^a \underset{\substack{t\to 0\\ x\to-\infty}}{\longrightarrow} 0$$

□

Exercice

Fixons $a, b \in \mathbb{R}_+^*$.

Montrer que pour x grand : $e^{\frac{a}{b+1}x} \geqslant \frac{a}{b+1}x$, en déduire $\frac{e^{ax}}{x^b} \geqslant \left(\frac{a}{b+1}\right)^{b+1} x$.

Conclure sur la valeur de $\lim\limits_{x\to+\infty} \frac{e^{ax}}{x^b}$.

Correction

5.6. Fonctions hyperboliques directes

Définition - Fonctions hyperboliques

Les fonctions ch (cosinus hyperbolique), sh (sinus hyperbolique) et th (tangente hyperbolique) sont définies sur $\mathbb{R}$ par :

$$\operatorname{ch} x = \frac{e^x + e^{-x}}{2}, \quad \operatorname{sh} x = \frac{e^x - e^{-x}}{2}, \quad \operatorname{th} x = \frac{\operatorname{sh} x}{\operatorname{ch} x} = \frac{e^x - e^{-x}}{e^x + e^{-x}}.$$

Proposition - Fonctions hyperboliques

La fonction ch est paire, les fonctions sh et th sont impaires.

$\operatorname{ch}(-x) = \operatorname{ch} x$ $\quad$ $\operatorname{sh}(-x) = -\operatorname{sh} x$

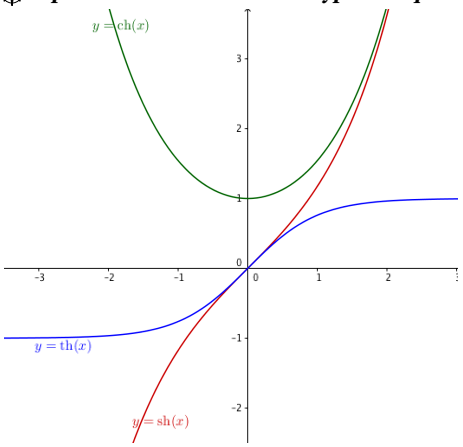
$\operatorname{ch}^2 x - \operatorname{sh}^2 x = 1$

$\operatorname{ch} x + \operatorname{sh} x = e^x$ $\quad$ $\operatorname{ch} x - \operatorname{sh} x = e^{-x}$

$\operatorname{ch}(a+b) = \operatorname{ch} a \operatorname{ch} b + \operatorname{sh} a \operatorname{sh} b$ $\quad$ $\operatorname{sh}(a+b) = \operatorname{sh} a \operatorname{ch} b + \operatorname{ch} a \operatorname{sh} b$

$\operatorname{th}(a+b) = \frac{\operatorname{th} a + \operatorname{th} b}{1 + \operatorname{th} a \operatorname{th} b}$

Représentation - Fonctions hyperboliques



Démonstration

$\operatorname{ch}^2(x) - \operatorname{sh}^2(x) = \frac{1}{4}(e^{2x} + 2 + e^{-2x} - e^{2x} + 2 - e^{-2x}) = 1.$

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

$\operatorname{ch} a \operatorname{ch} b + \operatorname{sh} a \operatorname{sh} b = \frac{1}{4}((e^a + e^{-a})(e^b + e^{-b}) + (e^a - e^{-a})(e^b - e^{-b})) = \frac{1}{4}(e^{a+b} + e^{a-b} + e^{-a+b} +$

$e^{-a-b} + e^{a+b} - e^{a-b} - e^{-a+b} + e^{-a-b}) = \frac{1}{2}\left(e^{a+b} + e^{-(a+b)}\right) = \operatorname{ch}(a+b).$ □

6. Sommes numériques infinies

On commence par une extension de notation :

Définition - Extension du coefficient binomial

Pour $\alpha \in \mathbb{R}$ et $k \in \mathbb{N}$, on note :

$$\binom{\alpha}{k} = \frac{\alpha(\alpha-1)\dots(\alpha-k+1)}{k!}$$

La section suivante donne des résultats connus pour l'essentiel depuis le XVIII siècle, au moins.
Mais les démonstrations satisfaisantes sont plus tardives (ABEL et successeurs du XIX).
Pour vous, elles auront officiellement lieu l'année prochaine lorsque vous verrez le cours sur les séries entières.

Proposition - Egalités sommatoires
On a les égalités suivantes :

$x \in ?$	fonction	somme	Auteur (Année)
$\mathbb{R}$	$\sin(x)$	$= \sum_{k=0}^{+\infty} \frac{(-1)^k}{(2k+1)!} x^{2k+1}$	NEWTON(1669), LEIBNIZ(1691)
$\mathbb{R}$	$\cos(x)$	$= \sum_{k=0}^{+\infty} \frac{(-1)^k}{(2k)!} x^{2k}$	NEWTON(1669), LEIBNIZ(1691)
$\mathbb{R}$	$\tan(x)$	$= x + \frac{1}{3}x^3 + \frac{2}{5}x^5 + \frac{17}{315}x^7 + \dots$	JAC. BERNOULLI(1702)
$[-1,1]$	$\arctan(x)$	$= \sum_{k=0}^{+\infty} \frac{(-1)^k}{2k+1} x^{2k+1}$	GREGORY(1671), LEIBNIZ(1674)
$[-1,1]$	$\arcsin(x)$	$= x + \frac{1}{2}\frac{x^3}{3} + \frac{1\cdot 3}{2\cdot 4}\frac{x^5}{5} + \dots$	NEWTON(1669)
$\mathbb{R}$	$(1+x)^n$	$= \sum_{k=0}^{n} \binom{n}{k} x^k$	PASCAL(1654)
$]-1,1[$	$(1+x)^\alpha$	$= \sum_{k=0}^{+\infty} \binom{\alpha}{k} x^k$	NEWTON(1666)
$\mathbb{R}$	$\exp(x)$	$= \sum_{k=0}^{+\infty} \frac{1}{k!} x^k$	EULER(1748)
$]-1,1[$	$\ln(1+x)$	$= \sum_{k=1}^{+\infty} \frac{(-1)^{k-1}}{k} x^k$	MERCATOR(1668)
$\mathbb{R}$	$\mathrm{sh}(x)$	$= \sum_{k=0}^{+\infty} \frac{1}{(2k+1)!} x^{2k+1}$	EULER(1748)
$\mathbb{R}$	$\mathrm{ch}(x)$	$= \sum_{k=0}^{+\infty} \frac{1}{(2k)!} x^{2k}$	EULER(1748)

Pour aller plus loin - Produit infini
Euler démontre également (1748) :

$$\forall\, x \in \mathbb{R}, \sin x = x \prod_{k=1}^{+\infty} \left(1 - \frac{x^2}{k^2\pi^2}\right)$$

Exercice
Donner l'expression formalisée de $\arcsin(x)$.

Correction
$\arcsin(x) = \sum_{k=0}^{+\infty} \frac{(2k)!}{2^{2k}k!^2} \frac{x^{2k+1}}{2k+1}$

Exercice
On rappelle la formule de MACHIN : $\frac{\pi}{4} = 4\arctan\frac{1}{5} - \arctan\frac{1}{239}$.
Combien de calcul pour obtenir 10 décimales de π satisfaisantes ? Faites les à la calculatrice

Correction
$\frac{1}{5} = 0.2$, et donc $\left(\frac{1}{5}\right)^n = 2^n \times 10^{-n}$.
Avec $n = 10$, on a $2^10 \approx 10^3$ et donc $\left(\frac{1}{5}\right)^n \approx 10^{3-10} = 10^{-7}$.
Avec $n = 14$, cela doit être suffisant...

7. Bilan

Synthèse

⇝ On s'intéresse aux fonctions dont le domaine est dans $\mathbb{R}$. Beaucoup de définitions : images, restrictions, ou additions, multiplications et com-

positions de fonctions. Des adjectifs pour les fonction : périodiques, paires ou impaires, majorées, minorées, bornées, monotone, strictement croissante…

⇝ Plusieurs fonctions de référence sont à connaître : exponentielle(s) et logarithme(s) en toute base ; les fonctions puissances ; les fonctions circulaires (sin, arccos…) et hyperboliques directes. On doit savoir comment comparer ces fonctions lorsqu'elles sont en compétition au voisinage de point problématique.

⇝ Les fonctions complexes de la variables réelles s'étudient de la même façon (même si la représentation est plus complexe). En fait, ce qui compte, c'est la nature de la variable « de départ ».

Plusieurs inégalités sont apparues dans ce chapitre. Il est bon de s'en souvenir :

Inégalité de fondamentale de trogonométrie : $\forall\, x \in \left]-\frac{\pi}{2},\frac{\pi}{2}\right[, \quad |\sin x| \leqslant |x| \leqslant |\tan x| = \frac{|\sin x|}{\cos x}$

Inégalité de Bernoulli : $\begin{cases} \forall\, n \in \mathbb{N}^*, \forall\, x \in]-1,+\infty[, & 1+nx \leqslant (1+x)^n \\ \forall\, n \in \mathbb{N}^*, \forall\, x \in \left]-1,\frac{1}{n}\right[& (1+x)^n \leqslant \dfrac{1}{1-nx} \end{cases}$

Inégalité de l'exponentielle/logarithme : $\begin{cases} \forall\, x \in]-1,1[& 1+x \leqslant \exp(x) \leqslant \dfrac{1}{1-x} \end{cases} \forall\, x > 0 \qquad \ln x \leqslant x-1$

Il est bon de connaître les stricts croissances des fonctions puissances ($\alpha > 0$) : $x < y \Rightarrow x^\alpha < y^\alpha$ et des fonctions exponentielles ($x > 1$) $\alpha < \beta \Rightarrow x^\alpha < x^\beta$.
Par ailleurs :
• les fonctions convexes comme exp, $x \in \mathbb{R}_+^* \mapsto \frac{1}{x}$ $x \mapsto x^\alpha$ ($\alpha > 1$), $x \in [0,\pi] \mapsto \sin x$
• ou concaves comme ln, $x \mapsto x^\alpha$ ($\alpha \in]0,1[$), $x \in [-\pi,\pi] \mapsto \cos x$
offrent de nombreuses inégalités utiles :

$\forall\, x,y \in \mathscr{D}_f \forall\, \lambda \in [0,1], \quad f(\lambda x+(1-\lambda)y) \leqslant \lambda f(x)+(1-\lambda)f(y)$ – cas convexe

$\forall\, x,y \in \mathscr{D}_f \forall\, \lambda \in [0,1], \quad \lambda f(x)+(1-\lambda)f(y) \leqslant f(\lambda x+(1-\lambda)y)$ – cas concave

Savoir-faire et Truc & **Astuce du chapitre**

— Savoir-faire - Transformation sur le graphe
— Savoir-faire - Etude des branches infinies (et définition)
— Savoir-faire - Transférer un problème trigonométrique « en a », vers « en 0 »
— Savoir-faire - Etude d'une équation fonctionnelle de $\mathbb{N}$ à $\mathbb{R}$
— Savoir-faire - Calculer « à la main » le logarithme en base a de x ?

Retour sur les problèmes

12. Elle est bien continue, mais elle n'est pas de classe $\mathscr{C}^1$. Euler ne l'aurait probablement pas considérée comme une fonction.
13. C'est l'application $x \mapsto 2^x u_0 = u_0 e^{x\ln 2}$
14. Les seules applications de cette forme sont les applications $x \mapsto A\ln(x)$ (où A est constante).
    Elles sont définie sur $\mathbb{R}_+$. Peut-on les étendre sur $\mathbb{R}$ en entier ?
    Si $x > 0$, $\ln(-x) = \ln(e^{i\pi}x) = \ln x + i\pi[2\pi]$…
15. On vient de terminer ce chapitre en répondant à cette question.

Chapitre 8

Utilisation de la dérivation

Résumé -

Ce chapitre se présente comme un large résumé du cours de première S avec ajouts de fonctions usuelles (vues en terminale ou en MPSI). On reprend tous les résultats en admettant la notion fine de limite, ils nous serviront d'outils pour des études plus poussés par la suite ou bien pour leur application dans d'autres domaines scientifiques. Le but est donc de raffiner notre technique (calculs)....
Il ne faut pas oublier que la notion de fonctions est une notion fine des mathématiques modernes, elle a mis plusieurs siècle à émerger : Leibniz, puis Euler, puis Cantor et Weierstrass... pour arriver jusqu'à nous. Et encore, le joyau n'est pas encore définitivement ciselé (distributions de Schwartz...). Sur internet :

- *OptimalSup-Spé - Cours Fonctions usuelles. https ://www.youtube.com/watch ?v=xTbtre9dgmQ*
- *Micmaths - Merveilleux logarithmes - https ://www.youtube.com/watch ?v=rWfl7Pw8YVE*
- *El Jj - Différences équations/fonctions (pas de questions stupides#01)*

Sommaire

1. Problèmes

? Problème 34 - Optimisation par FERMAT. Raisonnement prédérivatoire

Reprenons l'exemple de Fermat qui cherche à trouver la position E sur un segment $[AC]$ de manière à ce que $AE \times EC$ soit maximum.
Notons $a = AE$ et $b = EC$. On a donc $a + b = d (= AC)$ fixé. On cherche la valeur maximal de $AE \times EC = a(d-a) = ad - a^2$.
La méthode de Fermat consiste donc à ajouter une valeur e à AE, on a alors une nouvelle valeur qui vaut : $(a+e)d - (a+e)^2 = ad - a^2 + e(d - 2a) - e^2$.
Ajouter une valeur e à AE consiste donc à faire évaluer $AE \times EC$ d'une valeur $ad - a^2 + e(d-2a) - e^2 - ad + a^2 = e(d-2a) + e^2$.
Fermat dit qu'il faut « adégaler » les deux valeurs cela donne $e(d-2a) + e^2 = 0$. On peut simplifier par $e \neq 0$.
On trouve donc $d - 2a + e = 0$, et il prend $e = 0$ et donc $a = \frac{d}{2}$. E est au milieur de $[AC]$. Qu'en pensez-vous ?

? Problème 35 - Optimisation

On passe notre vie à optimiser nos actions, nos décisions (la plus efficace, la moins longue...).
Comment assurément trouver une stratégie qui permet à tous les coups d'optimiser (au moins localement) nos décisions.
Si elle se mesure sous la forme Résultat(Action), on doit trouver : Résultat(Action+δ)<Résultat(Action) et Résultat(Action-δ)<Résultat(Action) également...

? Problème 36 - Dérivation : concept local ou global ?

A priori l'optimisation précédente de la fonction Résultat donne des valeurs différentes selon la valeur Action où l'on se place. Ainsi, la notion de dérivation est un concept local.
Et donc, dans un second temps, une fois que pour tout x, $f'(x)$ est bien défini, on peut seulement réunir toutes ces valeurs en une seule fonction f.
D'où vient le miracle que dans la pratique, on peut directement agir de f à f' (de fonction à fonction) ?

? Problème 37 - Dérivations de fonctions usuelles et des opérations

Et ainsi, en particulier, quelles sont les transformations de dérivation pour les fonctions usuelles ? Et également, que deviennent les opérations classiques $(+, \times, \circ)$ pour la dérivation ?

? Problème 38 - Se concentrer sur un problème

D'une certaine façon, regarder la dérivation de f en x_0, c'est faire l'approximation affine (donc relativement simple) mais juste des valeurs $f(x)$ pour x proche de x_0.
Ces valeurs approchées peuvent se concentrer autour de 0. Et si l'on se

trouve autour d'un cas de limite problématique du type : $\frac{\to 0}{\to 0}$, ne peut-on pas exploiter les approximations affines (et donc les dérivées) pour obtenir une bonne évaluation de cette forme indéterminée ?

? Problème 39 - Etude de fonctions complexes

Comment étudier des fonctions à valeurs dans $\mathbb{C}$: $f : \mathbb{R} \to \mathbb{C}$ (comme $\theta \mapsto e^{i\theta}$) ?
Nous les avons définies lors du chapitre précédent ; comment les dériver maintenant ?
Mieux : comment étudier des fonctions de la variables complexes ? Peut-on les dériver, comment cela s'adapte ?

2. Dérivation

2.1. Approche historique

Remarque - Connaissance a priori

L a dérivation est une limite, il faudrait donc commencer par définir proprement (=formellement) la définition de la limite (d'une suite ou d'une fonction en un point). Or on suit (pour commencer l'année) le parcours historique des mathématiques, cette définition n'a été formalisée qu'au XIXème siècle (Bolzano, Cauchy, Weierstrass), après les théorèmes fondamentaux de Newton, Leibniz (fin du XVIIème siècle), Euler (XVIIIème siècle)… Nous nous contenterons de savoir que :

— toute suite de réels, croissante et majorée, converge (dans $\mathbb{R}$)
— si les limites existent bien et que les calculs sont possibles, alors la limite d'une somme, respectivement d'un produit, respectivement d'une composition est la somme des limites, respectivement le produit des limites, respectivement la composition des limites.

Pour aller plus loin - Analyse non standard

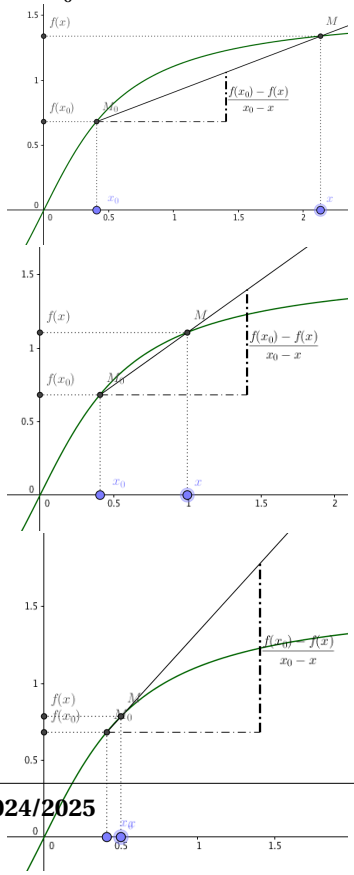
En 1961, Abraham Robinson crée une nouvelle logique mathématique en s'appuyant sur les nombres hyperréels.
Edward NELSON, *en 1977, renouvelle l'analyse non standatd (IST) en donnant un nouveau sens aux infinitésimaux et définit un nouveau calcul…*

Heuristique - Infinitésimaux

Le calcul différentiel et le calcul intégral ont été finalisés indépendamment et associés par Leibniz (1684) et Newton (1671, publié en 1736).
Leurs raisonnement repose sur une notion floue d'infinitésimaux (ou de fluente), notion non convaincante à l'époque.
Johan Bernoulli (1692) popularise auprès de toute sa dynastie, du marquis de L'Hospital et surtout d'Euler, la découvert de Leibniz (« une énigme plutôt qu'un explication ») et pense que des explications trop abondantes au sujet de l'infiniment petit pourrait troubler l'entendement de ceux qui ne sont pas « accoutumés à de longues explications ».
En fait, il manque la notion claire de limite, dont d'Alembert (1754) s'approche le plus. La bonne notion de limite est définie par Bolzano (1817), Cauchy (1823) ou Weierstrass (1861) et donnent ainsi une définition bien formalisée et solide mathématiquement (permettant de démontrer des résultats), abandonnant la notion d'Infinitésimaux.
Ce chapitre s'appuie sur des mathématiques du XVII. Nous ferons les démonstrations après avoir défini la notion de limite.

Représentation - Nombre dérivée

Il s'agit d'une notion « dynamique » (limite). Il faut le voir comme un film, en plusieurs temps $x \mapsto x_0$:



On retrouve ce type de schémas dans le cours de Leibniz ou de Joh. Bernoulli

Remarque - Cours de physique

En physique, en MPSI, le cours est proche des idées de Leibniz et Newton, et c'est très bien ainsi.
La notation $\frac{\partial f}{\partial x}$ est celle de Leibniz. La notation $\dot{x}$ est due à Newton !
C'est Lagrange, après un détour par Euler, qui impose la notation f'.

Remarque - Limite

On va faire comme si on connaissait la notion de limite, application linéaire ($\lim \lambda f + g = \lambda \lim f + \lim g$)…

2.2. Dérivabilité

Définition - Nombre dérivée
Soient I un intervalle de $\mathbb{R}$, f une fonction définie sur I et x_0 un point de I.
f est dérivable en x_0 si la fonction $x \mapsto \frac{f(x)-f(x_0)}{x-x_0}$, définie sur $I-\{x_0\}$, admet une limite finie en x_0.
On a alors

$$f'(x_0) = \lim_{x\to x_0} \frac{f(x)-f(x_0)}{x-x_0} = \lim_{h\to 0} \frac{f(x_0+h)-f(x_0)}{h}$$

f est dérivable sur I si elle est dérivable en tout point de I.
$f'(x)$ s'appelle le nombre dérivé en x et la fonction qui à $x \in I$ associe $f'(x)$ s'appelle la (fonction) dérivée de f.

2.3. Approximation linéaire

On donne parfois une définition équivalente de la dérivation de f en x_0.
Son avantage : elle cache la question de la limite dans l'hypothèse de continuité de ϵ en x_0. .

Pour aller plus loin - Continuité = limite
Nous verrons que la notion de continuité de f en un point x_0 est équivalente à la notion d'existence de limite de f en x_0.

Proposition - Définition de Weierstrass
Soit f définie sur un intervalle I. Soit $x_0 \in I$
f est dérivable en x_0 si et seulement si
il existe $A \in \mathbb{R}$, ϵ continue en x_0 avec $\epsilon(x_0) = 0$ tels que

$$f(x) = f(x_0) + (x-x_0)(A+\epsilon(x))$$

On a alors $A = f'(x_0)$

Démonstration
Si f est dérivable en x_0, on considère $A = f'(x_0)$ et $\epsilon : x \mapsto \frac{f(x)-f(x_0)}{x-x_0} - f'(x_0)$.
Ces fonctions répondent aux conditions.
Réciproquement, si $f(x) = f(x_0) + (x-x_0)(A+\epsilon(x))$, alors $\frac{f(x)-f(x_0)}{x-x_0}$ admet une limite égale à A en x_0. □

Application - Pour quelques fonctions usuelles (connaissant les dérivées en 0
On trouve avec $\epsilon_i(x) \to 0$ pour $x \to 0$:

$$\exp(x) = 1 + x + x\epsilon_1(x)$$

$$\ln(1+x) = x + x\epsilon_2(x)$$

$$(1+x)^\alpha = 1 + \alpha x + x\epsilon_3(x)$$

C'est le début du calcul des équivalents.

Pour aller plus loin - Linéarisation
Autre idée essentielle pour l'exploitation de la dérivation : transformer un problème paramétré par une fonction plus ou moins compliquée en un autre problème linéaire. Pour cela on remplace f par une droite, la plus proche c'est-à-dire la dérivée de f (en un x_0 bien choisi).
C'est en particulier la méthode de Newton que nous reverrons en informatique. Il s'agit de resoudre $f(x) = 0$.

Définition - Équation de la tangente
Soit f définie sur I, dérivable en $x_0 \in I$.
La droite d'équation

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

s'appelle la tangente à la courbe représentative de f en $M(x_0, f(x_0))$

Truc & Astuce pour le calcul - A propos de la tangente
On notera :
— qu'il s'agit bien de l'équation d'une droite ($y = ax+b$)
— qu'elle passe bien par le point M : $f(x_0) + f'(x_0)(x_0-x_0) = f(x_0)$.

. — que sa pente vaut $f'(x_0)$.

2.4. Règles de dérivation

Les résultats suivants seront démontrées plus tard dans l'année (linéarité de la limite). Ils ont été énoncés pour la première fois par LEIBNIZ en 1684. Le but pour le moment est de se former pour le calcul!

Proposition - Résultats
Soient u et v dérivable en x_0
- $u+v, uv, u^n, \exp u$ sont dérivables en x_0 ;
- si, en outre, u ne s'annule pas sur un intervalle contenant x_0 et est dérivable en x_0 alors $\frac{1}{u}, \ln|u|$ sont dérivables en x_0 ;
- si, en outre, v ne s'annule pas sur un intervalle contenant x_0 et est dérivable en x_0 alors $\frac{u}{v}$ est dérivable en x_0 ;
- si, en outre, $\alpha \in \mathbb{R}^*$ et si u est strictement positive sur un intervalle contenant x_0, alors u^α est dérivable en x_0

Si $u\circ\phi$ est définie, si φ est dérivable en x_0 et u dérivable en $\varphi(x_0)$ alors $u\circ\varphi$ est dérivable en x_0 et $(u\circ\varphi)'(x_0) = \varphi'(x_0) \times u'(\varphi(x_0))$.

Résumé :

Proposition - Résumé des formules usuelles de dérivation

fonction f de la forme	fonction dérivée f'
$u+v$	$u'+v'$
uv	$u'v+uv'$
$u_1u_2\dots u_n$	$u_1'u_2\dots u_n + u_1u_2'u_3\dots u_n + \dots + u_1u_2\dots u_{n-1}u_n'$
u^n où $n\in\mathbb{N}^*$	$nu'u^{n-1}$
$\frac{1}{u}$	$\frac{-u'}{u^2}$
$\frac{u}{v}$	$\frac{u'v-uv'}{v^2}$
$u\circ\phi$	$\phi'\times u'\circ\phi = u'\circ\phi\times\phi'$
$u_1\circ\dots\circ u_n$	$(u_1'\circ u_2\circ\dots\circ u_n)\times(u_2'\circ u_3\circ\dots\circ u_n)\times\dots\times u_n'$
$\exp u$	$u'\times\exp u$
$\ln\lvert u\rvert$	$\frac{u'}{u}$
u^α où $\alpha\in\mathbb{R}^*$	$\alpha u'u^{\alpha-1}$

Ces formules sont vraies, sous réserve, évidente, de dérivabilité des fonctions qui interviennent. Exercice
Avec les résultats admis en début de chapitre, démontrer la formule de dérivation de $u\times v$.

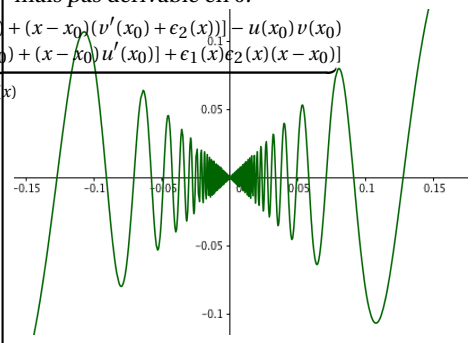
Correction
On exploite la méthode de Weierstrass

$$\begin{aligned}[uv](x)-[uv](x_0) &= u(x)v(x)-u(x_0)v(x_0) = [u(x_0)+(x-x_0)(u'(x_0)+\epsilon_1(x))][v(x_0)+(x-x_0)(v'(x_0)+\epsilon_2(x))]-u(x_0)v(x_0)\\ &= (x-x_0)[v(x_0)u'(x_0)+u(x_0)v'(x_0)+\underbrace{\epsilon_1(x)[v(x_0)+(x-x_0)v'(x_0)]+\epsilon_2(x)[u(x_0)+(x-x_0)u'(x_0)]+\epsilon_1(x)\epsilon_2(x)(x-x_0)}_{:=\epsilon(x)}]\end{aligned}$$

Comme $\epsilon(x)\to 0$ (par addition et produit des limites), on trouve bien le résultat attendu.
Nous reprendrons toutes ces démonstrations par la suite.

Pour aller plus loin - Toutes les fonctions sont dérivables?

La fonction $x\mapsto x\cos\frac{1}{x}$ est continue sur $\mathbb{R}$, mais pas dérivable en 0.



Mais pire, il existe des fonctions partout continues et nulle part dérivable. La construction est subtile. Par exemple :

$$x\mapsto\sum_{n=1}^{+\infty}\frac{\sin(10^n\pi x)}{2^n}$$

2.5. Dérivation de fonctions usuelles

Fonctions exponentielles et logarithmes

Savoir faire - Encadrement pour le calcul de limite

Pour obtenir des résultats sur les limites (indéterminées), la meilleure stratégie est l'encadrement :
si $a(x) \leqslant b(x) \leqslant c(x)$ et que a et c admettent la même limite en x_0 égale à ℓ,
alors b admet également une limite en x_0, égale à ℓ.

Analyse - Exponentielle (naturelle) en 0

On sait que pour tout $x \in]-1,1[$, $1+x \leqslant e^x \leqslant \dfrac{1}{1-x}$, donc $x \leqslant e^x - 1 \leqslant \dfrac{x}{1-x}$.
Puis en divisant par $x \neq 0$:

$$1 \geqslant \frac{e^x-1}{x} \geqslant \frac{1}{1-x} \text{ (cas } x<0 \qquad 1 \leqslant \frac{e^x-1}{x} \leqslant \frac{1}{1-x} \text{ (cas } x>0$$

Par encadrement : exp est dérivable en 0, de valeur égale à 1.

Proposition - Dérivation des fonctions exponentielles

exp est dérivable sur $\mathbb{R}$ et pour tout $x \in \mathbb{R}$, $\exp'(x) = \exp(x)$.
Pour $a > 0$ et $a \neq 1$. Si $f : x \mapsto a^x$, alors f est dérivable sur $\mathbb{R}$ et pour tout $x \in \mathbb{R}$, $f'(x) = \ln a \times a^x$.

Savoir faire - Transférer un problème exponentiel « en a », vers « en 0 »

Si on étudie exp au voisinage de a, donc des points de la forme $a+h$ avec h proche de 0,
on exploite $\exp(a+h) = \exp a \times \exp(h)$.
Donc on factorise (à l'extérieur) par $\exp a$ et on se concentre sur une étude en ϵ, ϵ proche de 0

On appliquera souvent cette méthode dans le calcul de développement limité.

Démonstration

Soit $x_0 \in \mathbb{R}$, pour tout $h = x - x_0 \in \mathbb{R}$,

$$\frac{\exp(x)-\exp x_0}{x-x_0} = \frac{\exp(x_0+h)-\exp x_0}{h} = e^{x_0}\frac{e^h-1}{h} \to e^{x_0}$$

On rappelle que $a^x = \exp(x \times \ln a)$. En posant $u = \ln a(x-x_0)$:

$$\frac{a^x - a^{x_0}}{x-x_0} = \frac{\exp(x_0 \ln a + h \ln a) - \exp(x_0 \ln a)}{h \ln a} \times \ln a = e^{x_0 \ln a} \times \ln a \frac{e^u - 1}{u} \to \ln a a^{x_0}$$

car pour $x \to x_0$, on a $u \to 0$ □

Par addition et composition (démontrée plus loin)

Proposition - Fonction trigonométrique hyperbolique

Les fonctions ch, sh et th sont dérivables sur $\mathbb{R}$ et

$$\forall\, x \in \mathbb{R},\ \mathrm{ch}'(x) = \mathrm{sh}\, x,\ \mathrm{sh}'(x) = \mathrm{ch}\, x \text{ et } \mathrm{th}'(x) = 1 - \mathrm{th}^2(x) = \frac{1}{\mathrm{ch}\, x}.$$

Démonstration

$\mathrm{ch}'(x) = e^x + (-1)e^{-x} = \mathrm{sh}(x)$, $\mathrm{sh}'(x) = e^x - (-1)e^{-x} = \mathrm{ch}(x)$.
$\mathrm{th}'(x) = \dfrac{\mathrm{ch}^2(x) - \mathrm{sh}^2(x)}{\mathrm{ch}^2(x)} = 1 - \mathrm{th}^2(x) = \dfrac{1}{\mathrm{ch}^2(x)}$. □

Remarque - $\frac{\partial}{\partial x} e^{-x}$?

On peut noter que $e^{-x} = \left(e^{-1}\right)^x$, c'est une fonction exponentielle de base e^{-1}.
Donc de dérivée : $\ln(e^{-1}) \times (e^{-1})^x = -e^{-x}$.

Analyse - Logarithme (naturel) en 1

On vient de voir que $\lim\limits_{x\to 0}\dfrac{e^x-1}{x}=1$.
En composant par $X=e^x$ et donc $x=\ln X$, avec l'équivalence $x\to 0 \Leftrightarrow X\to 1$, on a

$$\lim_{X\to 1}\frac{X-1}{\ln X-\underbrace{\ln 1}_{=0}}=1$$

En prenant l'inverse, on a donc ln est dérivable en 1, de valeur égale à $\frac{1}{1}=1$.

Proposition - Dérivation des fonctions logarithmes
ln est dérivable sur $\mathbb{R}_+$ et pour tout $x\in\mathbb{R}_+$, $\ln'(x)=\dfrac{1}{x}$.
Si $g: x\mapsto \ln_a(x)$, alors g est dérivable sur $\mathbb{R}_+$ et pour tout $x\in\mathbb{R}_+$, $g'(x)=\dfrac{1}{(\ln a)\times x}$.

Savoir faire - Transférer un problème logarithme « en a », vers « en 1 »

Si on étudie ln au voisinage de a, donc des points de la forme $a+h$ avec h proche de 0,
on factorise (à l'intérieur) par a : $a+h=a(1+\frac{h}{a})$ et exploite $\ln(a+h)=\ln a+\ln(1+\frac{h}{a})$.
Et on se concentre sur une étude en $1+\epsilon$, ϵ proche de 0

On appliquera souvent cette méthode dans le calcul de développement limité.

Démonstration
Soit $x_0\in\mathbb{R}_+$, pour tout $h=x-x_0\in\mathbb{R}$ (et $u=\dfrac{x-x_0}{x_0}$),

$$\frac{\ln(x)-\ln x_0}{x-x_0}=\frac{\ln(x_0(1+\frac{h}{x_0})-\ln x_0}{h}=\frac{1}{x_0}\times\frac{\ln(1+u)}{u}\to\frac{1}{x_0}$$

On rappelle que $\ln_a(x)=\dfrac{\ln x}{\ln a}$. Par produit (démontré plus loin), $\ln'_a(x_0)=\dfrac{1}{\ln a}\ln'(x_0)=\dfrac{1}{(\ln a)\times x_0}$ □

Fonctions puissances

Proposition - Fonctions puissances
La fonction puissance $x\mapsto x^\alpha$ est dérivable sur son ensemble de définition (privé de 0, si besoin : $\alpha-1<0$ alors que $\alpha>0$), de dérivée : $x\mapsto \alpha x^{\alpha-1}$.

Démonstration
Il faudrait étudier tous les cas $\alpha\in\mathbb{N}$, $\alpha\in\mathbb{Z}$, $\alpha\in\mathbb{Q}$.
Notons que $x\mapsto x^\alpha=\exp(\alpha\ln x)$, de dérivée (par composition) : $x\mapsto \alpha\dfrac{1}{x}\exp(\alpha\ln x)=\alpha\dfrac{x^\alpha}{x}=\alpha x^{\alpha-1}$. □

Par sommation, on retrouve ce qu'on a déjà démontré :

Proposition - Fonctions polynomiales
Les fonctions polynomiales sont dérivables sur $\mathbb{R}$ et précisément, si $f: x\mapsto\sum\limits_{k=0}^{n}a_kx^k$, on a pour tout $x\in\mathbb{R}$, $f'(x)=\sum\limits_{k=0}^{n}ka_kx^{k-1}=\sum\limits_{k=1}^{n}ka_kx^{k-1}=\sum\limits_{k=0}^{n-1}(k+1)a_{k+1}x^k$.
On notera qu'il s'agit encore d'un polynôme, de degré $\deg f-1$.

L'ensemble de définition des fonctions puissances entières étant $\mathbb{R}$ en entier, les résultats précédents ne sont pas suffisants. Il faut refaire une démonstration.
Il serait tout de même étonnant de trouver des résultats différents...

Fonctions circulaires

Proposition - Fonction trigonométrique
sin, cos et tan sont dérivables sur leur ensemble de définition. Précisément :
$\forall\, x \in \mathbb{R},\ \sin'(x) = \cos x,\ \forall\, x \in \mathbb{R},\ \cos'(x) = -\sin x.$ et pour tout $x \in \mathbb{R} \setminus \{\frac{\pi}{2} + k\pi, k \in \mathbb{Z}\}$, $\tan'(x) = 1 + \tan^2 x = \dfrac{1}{\cos^2 x}$.

On applique une méthode déjà connue.

Démonstration
Par encadrement (chapitre précédent) : $\dfrac{\sin x}{x} \underset{x\to 0}{\longrightarrow} 1$.
Ainsi sin est dérivable en 0 de dérivée égale à 1.
De même étudions la dérivabilité de cos en 0 :
$\cos x - 1 = \mathrm{Re}(e^{ix} - 1) = \mathrm{Re}(e^{ix/2}(2i \sin \frac{x}{2})) = -2\sin^2 \frac{x}{2}$.
Donc $\dfrac{\cos x - 1}{x} = -\sin \frac{x}{2} \dfrac{\sin \frac{x}{2}}{\frac{x}{2}} \to 0$ (par produit 0×1).
Ainsi cos est dérivable en 0 de dérivée égale à 0.
Puis $\dfrac{\sin(x_0) - \sin x}{x_0 - x} = \dfrac{\sin(x_0 + h) - \sin x_0}{h} = \dfrac{\sin x_0(\cos h - 1) + \sin h \cos x_0}{h} = \dfrac{\cos h - 1}{h} \sin x_0 + \dfrac{\sin h}{h} \cos x_0 \to \cos x_0$ pour $h \to 0$.
Puis $\dfrac{\cos(x_0) - \cos x}{x_0 - x} = \dfrac{\cos(x_0 + h) - \cos x_0}{h} = \dfrac{\cos x_0(\cos h - 1) - \sin h \sin x_0}{h} = \dfrac{\cos h - 1}{h} \cos x_0 - \dfrac{\sin h}{h} \sin x_0 \to -\sin x_0$ pour $h \to 0$.
Pour la fonction tangente, on se souvient que $\tan(x_0 + h) - \tan x_0 = \dfrac{\tan x_0 + \tan h}{1 - \tan x_0 \tan h} - \tan x_0 = \dfrac{\tan h(1 + \tan^2 x_0)}{1 + \tan x_0 \tan h}$.
Donc $\dfrac{\tan(x_0 + h) - \tan x_0}{h} = \dfrac{\sin h}{h}(1 + \tan^2 x_0) \dfrac{1}{\cos h(1 + \tan x_0 \tan h)} \underset{h\to 0}{\longrightarrow} 1 + \tan^2(x_0) = \dfrac{1}{\cos^2 x_0}$
□

Les résultats sur les dérivées des fonctions réciproques arc sin... seront vus loin lorsqu'on étudiera la dérivation de fonctions réciproques.

2.6. Dérivées seconde, troisième...

Définition - Dérivables
Soient I un intervalle et f une fonction définie sur I.
On dit que f est deux fois dérivable sur I si f est dérivable sur I et si f' est elle-même dérivable sur I, on note f'' la dérivée seconde de f (c'est-à-dire la dérivée de f').
Si f'' est encore dérivable sur I, on dit que f est trois fois dérivable sur I et on note $f''' = f^{(3)} = (f'')'$ sa dérivée troisième, et ainsi de suite.

Définition - Fonction de classe $\mathscr{C}^k$
On dit que f est de classe $\mathscr{C}^1$ ou continûment dérivable si elle est dérivable et que f' est continue, de classe $\mathscr{C}^2$ ou 2 fois continûment dérivable si elle est deux fois dérivable et que f'' est continue...

Comme les dérivées des fonctions usuelles vues plus haut sont de la forme de fonctions usuelles, par récurrence :

Proposition - Fonctions usuelles
Les fonctions usuelles vues précédemment sont de classe $\mathscr{C}^\infty$ sur leur ensemble de définition (sauf les fonctions puissances $x \mapsto x^\alpha$ qui peuvent perdre le 0 dans l'ensemble à la dérivée n-ième si $n > \alpha$...).

3. Quelques utilisations de la dérivation

3.1. Variations

En utilisant les théorèmes suivants (qui seront démontrés ultérieurement), on peut étudier les variations d'une fonction que l'on résume **systématiquement** dans un tableau de variations, complété par les limites aux bornes (et non par des phrases!)

Théorème - Monotonie et fonction dérivée
Soient I un **intervalle** de $\mathbb{R}$ et f une fonction dérivable sur I. Alors :
f est constante sur I si et seulement si $\forall x \in I$, $f'(x) = 0$;
f est croissante sur I si et seulement si $\forall x \in I$, $f'(x) \geqslant 0$;
f est décroissante sur I si et seulement si $\forall x \in I$, $f'(x) \leqslant 0$.
Si $f' > 0$ (resp. $f' < 0$) sauf en un nombre fini de points de I, alors f est strictement croissante (resp. décroissante) sur I.

Attention - Intervalle!
Il est indispensable que I soit un intervalle. On peut en effet trouver f définie sur D, dérivable sur D et vérifiant $\forall x \in D$, $f'(x) < 0$ mais qui n'est pas décroissante sur D.

Exercice
Donner un tel contre-exemple

Correction

3.2. Bijections et réciproques

Théorème - Dérivation de la bijection réciproque
Soient I, J deux intervalles de $\mathbb{R}$ et $f : I \to J$ bijective de I sur J. On suppose que f est dérivable en $x_0 \in I$.
Alors f^{-1} est dérivable en $f(x_0)$ **si et seulement si** $f'(x_0) \neq 0$
et on a alors

$$(f^{-1})'(f(x_0)) = \frac{1}{f'(x_0)}$$

ou encore, avec $y_0 = f(x_0)$,

$$(f^{-1})'(y_0) = \frac{1}{f' \circ f^{-1}(y_0)}$$

Remarque - Se souvenir

On peut retrouver la formule en dérivant l'égalité $f^{-1}\circ f = \mathrm{Id}$, en effet :

$$(f^{-1}\circ f)' = f' \times (f^{-1})'\circ f = (\mathrm{Id})' = (x \mapsto 1)$$

mais attention à ne pas oublier la condition de dérivabilité de f^{-1} !

Si $f'(x) \neq 0$: $(f^{-1})'(f(x)) = \dfrac{1}{f'(x)}$ si $y = f(x)$ et $f'(x) = f'(f^{-1}(y)) \neq 0$:

$$(f^{-1})'(y) = \frac{1}{f'(f^{-1}(y))}$$

Proposition - Fonctions trigonométriques réciproques

Les fonctions arcsin et arccos sont de classe $\mathscr{C}^\infty$ sur $]-1,1[$,

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

La fonction arctan est de classe $\mathscr{C}^\infty$ sur $\mathbb{R}$,

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

Démonstration

$x \in \mathscr{D}_{(\sin^{-1})'} \iff \sin'(\arcsin(x)) \neq 0 \iff \cos(\arcsin(x) \neq 0 \iff \arcsin x \notin \{-\frac{\pi}{2}, \frac{\pi}{2}\} \iff x \notin \{-1,1\}$.

Donc $\mathscr{D}_{\arcsin'} =]-1,1[$ et

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

$x \in \mathscr{D}_{(\cos^{-1})'} \iff \cos'(\arccos(x)) \neq 0 \iff \sin(\arccos(x) \neq 0 \iff \arccos x \notin \{0,\pi\} \iff x \notin \{-1,1\}$.

Donc $\mathscr{D}_{\arccos'} =]-1,1[$ et

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

$x \in \mathscr{D}_{(\tan^{-1})'} \iff \tan'(\arctan(x)) \neq 0 \iff x \notin \{\frac{\pi}{2} + 2k\pi, k \in \mathbb{Z}\}$.

Or cette dernière affirmation est toujours vraie. Donc $\mathscr{D}_{\arctan'} = \mathbb{R}$ et

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

Pour aller plus loin - Création de fonctions

Dans un autre sens, on peut considérer :

1. *f strictement monotone et continue de I sur J.*
2. *Donc f admet une fonction réciproque f^{-1} de J sur I.*
3. *Notons Φ, une primitive de f^{-1}.*

Que peut-on dire de Φ ?
Montrer que la plupart des fonctions transcendantes rencontrées dans ce cours peuvent être obtenus de la sorte…

Corollaire -

Si $f : I \to \mathbb{R}$ est :

- continue strictement monotone sur l'intervalle I,
- dérivable sur I et
- $\forall x \in I$, $f'(x) \neq 0$

alors f est bijective de I sur $J = f(I)$, donc $f^{-1} : J \to I$ existe (et est bijective).

Mieux : f^{-1} est dérivable sur J et $(f^{-1})' = \dfrac{1}{f'\circ f^{-1}}$.

Graphiquement : $\mathscr{C}_{f^{-1}}$ est la symétrique de $\mathscr{C}_f$ par rapport à l'axe $y = x$.

Savoir faire - Etudier si f est un $\mathscr{C}^1$-difféomorphisme

On considère f, de classe $\mathscr{C}^1$ sur un intervalle I.

1. On coupe I, en réunion d'intervalles $I_1, I_2, \dots I_n$ sur lesquels f est strictement monotone
   (A noter que cela n'est pas toujours possible…. Donnez des contre-exemple).
   Cela conduit à étudier le signe strictement non nul de f'. Les frontières sont en dans $\{x \mid f'(x) = 0\}$.
2. Sur chaque intervalle I_k, on note $J_k = \{y = f(x), x \in I_k\}$ et donc $f_k = f^{|J_k}_{|I_k} : I_k \to J_k$ est bijective.
3. Puis, on considère $J'_k = \{y = f(x), x \in I_k \mid f'(x) \neq 0\} \subset J_k$.
   (Souvent les problèmes sont sur les bords des J_k, lié à la jonction de I_k et I_{k+1} ou I_{k-1}, mais cela n'est pas nécessaire).

Alors $f_k^{-1} : J_k \to I_k$ est de classe $\mathscr{C}^1$ sur J'_k et pour tout $y \in J'_k$:

$$(f^{-1})'(y) = (f_k^{-1})'(y) = \frac{1}{f'(f_k^{-1}(y)} = \frac{1}{f'(f^{-1}(y)}$$

Gardons en mémoire que graphiquement : $\mathscr{C}_{f^{-1}}$ est la symétrique de $\mathscr{C}_f$ par rapport à l'axe $y = x$.
Les cas problématique correspond à une pente nulle pour $\mathscr{C}_f$ et donc infinie pour $\mathscr{C}_{f^{-1}}$...

Exercice

La fonction sh réalise une bijection de $\mathbb{R}$ sur $\mathbb{R}$. La bijection réciproque est notée argsh (argument sinus hyperbolique).
La fonction ch réalise une bijection de $[0, +\infty[$ sur $[1, +\infty[$. La bijection réciproque est notée argch (argument cosinus hyperbolique).
La fonction th réalise une bijection de $\mathbb{R}$ sur $]-1, 1[$. La bijection réciproque est notée argth (argument tangente hyperbolique).

1. Dérivées.
   (a) Montrer que les fonctions argsh, argch, argth sont dérivables respectivement sur $\mathbb{R}$, $]1, +\infty[$ et $]-1, 1[$.
   (b) Montrer que $(\text{argsh})'(x) = \frac{1}{\sqrt{x^2+1}}$ sur $\mathbb{R}$
   (c) Montrer que $(\text{argch})'(x) = \frac{1}{\sqrt{x^2-1}}$ sur $]1, +\infty[$
   (d) Montrer que $(\text{argth})'(x) = \frac{1}{1-x^2}$ sur $]-1, 1[$
2. Expressions logarithmiques.
   (a) Montrer que $\forall x \in \mathbb{R}, \quad \text{argsh}\, x = \ln(x + \sqrt{x^2+1})$
   (b) Montrer que $\forall x \in]1, +\infty[, \quad \text{argch}\, x = \ln(x + \sqrt{x^2-1})$
   (c) Montrer que $\forall x \in]-1, 1[, \quad \text{argth}\, x = \frac{1}{2}\ln\frac{1+x}{1-x}$

Correction

3.3. Inégalités

Savoir faire - Obtenir une inégalité

Pour démontrer une inégalité, une méthode est d'étudier la fonction formée par la différence des deux membres, d'établir son tableau de variations pour obtenir son signe.

Exercice

Montrer les inégalités :

$$\forall x \in \mathbb{R}, 1 - \frac{x^2}{2} \leqslant \cos x \leqslant 1;$$
$$\forall x \in \mathbb{R}^+, x - \frac{x^3}{6} \leqslant \sin x \leqslant x;$$
$$\forall x \in \mathbb{R}^-, x - \frac{x^3}{6} \geqslant \sin x \geqslant x.$$

Correction

Commençons par noter que $\forall\, x \in \mathbb{R}, \cos x \leqslant 1$.
Les fonctions sin, cos et polynomiales sont dérivables sur $\mathbb{R}$, de même de leurs additions :

$$f_1 : x \mapsto \cos x - 1 + \frac{x^2}{2} \qquad f_2 : x \mapsto \sin x - x \qquad f_3 : x \mapsto \sin x - x + \frac{x^3}{6}$$

Les dérivées sont respectivement :

$$f_1' : x \mapsto -\sin x + x \qquad f_2' : x \mapsto \cos x - 1 \qquad f_3' : x \mapsto \cos x - 1 + \frac{x^2}{2}$$

On en déduit les tableaux de variations dans l'ordre déductif suivant :

Pour aller plus loin - Querelle entre Fermat et Descartes

Fermat et Descartes se sont querellés sur les lois de refraction (loi de Snell-Descartes). Pour Fermat, la démonstration de Descartes était incomplète. Le Toulousain proposait de trouver

x	$-\infty$		0		$+\infty$
f_2'		$-$		$-$	
f_2		↘	0	↘	
$f_1'(=-f_2)$		$-$	0	$+$	
f_1		↘	0	↗	
$f_3'(=f_1)$		$+$	0	$+$	
f_3		↗	0	↗	

Donc $\cos x \leqslant 1$ et pour $x \in \mathbb{R}_+$, $f_1(x) \geqslant 0$, donc $\cos x \geqslant 1 - x$.
Et pour tout $x \in \mathbb{R}_+$, $f_3(x) \geqslant 0$, donc $\sin x \geqslant x - \frac{x^3}{6}$ et $f_2(x) \leqslant 0$ donc $\sin x \leqslant x$.
Et pour tout $x \in \mathbb{R}_-$, $f_3(x) \leqslant 0$, donc $\sin x \leqslant x - \frac{x^3}{6}$ et $f_2(x) \geqslant 0$ donc $\sin x \geqslant x$.

Savoir faire - Obtenir un maximum (avec une dérivée)

Pour obtenir un extremum de f (dérivable deux fois) sur I, on résout $f'(x) = 0$.
On note x_0 la solution de cette équation (donc $f'(x_0) = 0$).
- si $f''(x_0) > 0$, alors f présente un minimum (local) en x_0.
- si $f''(x_0) < 0$, alors f présente un maximum (local) en x_0.
- si $f''(x_0) = 0$, alors on ne peut rien dire.

Remarque - Convexité

On rappelle que si pour tout $x \in I$, $f''(x) \geqslant 0$, on dit que f est convexe sur I.

3.4. Calculs de limites (lever les indéterminations)

On peut calculer certaines limites en reconnaissant le taux de variations d'une fonction dérivable connue.

Exercice

Rappeler les valeurs des limites suivantes :

$$\lim_{x\to 0} \frac{\sin x}{x}; \quad \lim_{x\to 0} \frac{\tan x}{x}; \quad \lim_{x\to 0} \frac{\cos x - 1}{x}.$$

En déduire

$$\lim_{x\to 0} \frac{\cos x - 1}{x^2}$$

Correction

Il s'agit de limite de taux de variations donc d'un calcul de dérivée :
$\lim_{x\to 0} \frac{\sin x}{x} = \sin'(0) = \cos(0) = 1$, $\frac{\tan x}{x} = \tan'(0) = 1 + \tan^2(0) = 1$ et $\lim_{x\to 0} \frac{\cos x - 1}{x} = \cos'(0) = \sin(0) = 0$.
Enfin, en posant $u = \frac{x}{2}$:

$$\frac{\cos x - 1}{x^2} = -2\frac{\sin^2(\frac{x}{2})}{x^2} = -\frac{\sin^2(u)}{2u^2} = -\frac{\sin^2(u)}{2u^2} = \frac{-1}{2}\left(\frac{\sin u}{u}\right)^2$$

En faisant tendre $x \to 0$, donc $u \to 0$, on a par produit de limite : $\lim_{x\to 0} \frac{\cos x - 1}{x^2} = -\frac{1}{2} \times 1^2 = -\frac{1}{2}$

Proposition - Règle de L'Hospital (1696)

Soient f et g deux fonctions définies sur I, dérivables en $x_0 \in I$.
On suppose que $f(x_0) = g(x_0) = 0$ et que $\dfrac{f'}{g'}$ admet une limite en x_0.

Alors $\displaystyle\lim_{x\to x_0} \frac{f(x)}{g(x)} = \lim_{x\to x_0} \frac{f'(x)}{g'(x)}$

On fait la démonstration dans le cas où $g'(x_0) \neq 0$. Sinon, on exploitera l'inégalité des accroissements finis (dans un chapitre prochain).

Démonstration
Avec les notations de Weierstrass :

$$f(x) = (x-x0)(f'(x_0)+\epsilon_1(x)) \qquad g(x) = (x-x0)(g'(x_0)+\epsilon_2(x))$$

$$\frac{f(x)}{g(x)} = \frac{f'(x_0)+\epsilon_1(x)}{g'(x_0)+\epsilon_2(x)} \underset{x\to x_0}{\longrightarrow} \frac{f'(x_0)}{g'(x_0)}$$

□

ation en $+\infty$
règle de L'Hô-
u type $\frac{\pm\infty}{\pm\infty}$.
$\lim_a \frac{f'}{g'} = \ell$,

Attention - Pas d'abus
La règle ne s'applique qu'en cas d'indetermination.

$$-4 = \lim_{x\to 1} \frac{3x^2+1}{2x-3} \neq \lim_{x\to 1} \frac{6x}{2} = 3$$

Savoir faire - Lever une indétermination par la règle de L'Hospital
1. D'abord on vérifie bien qu'on peut appliquer la règle de L'Hospital :

$$f = \frac{n}{d} \text{ avec } n(x) \underset{x\to a}{\longrightarrow} 0 \text{ et } d(x) \underset{x\to a}{\longrightarrow} 0$$

2. On calcule $n'(x)$, $n''(x) \dots n^{(k)}(x) \dots$ jusque à ce que $n^{(k)}(x) \underset{x\to a}{\not\longrightarrow} 0$
Et de même $d'(x)$, $d''(x) \dots d^{(\ell)}(x) \dots$ jusqu'à ce que $d^{(\ell)}(x) \underset{x\to a}{\not\longrightarrow} 0$.
3. On note $m = \min\{k \in \mathbb{N} \mid n^{(k)}(x) \not\to 0 \text{ ou } d^{(k)}(x) \not\to 0\}$.
Et on connait donc $\lim_a \frac{n^{(k)}}{d^{(k)}}$, et par la règle de L'Hospital : $\lim_a \frac{n^{(k-1)}}{d^{(k-1)}}$ a la même valeur.
4. Et on remonte ainsi à $\lim_a \frac{n}{d}$, dont
on sait seulement maintenant qu'elle existe et qu'elle a cette même valeur.

Exercice
Calculer $\lim_{x\to 0} \frac{\cos 2x - 1}{x^3+5x^2}$.
Evidemment, on exploitera les résultats non encore démontrés sur les fonctions usuelles

Correction
On note $f : x \mapsto \cos 2x - 1$, dérivable, $f'(x) = -2\sin(2x)$, donc $f'(0) = 0$.
On note $g : x \mapsto x^3 + 5x^2$, dérivable, $g'(x) = 3x^2 + 10x$, donc $g'(0) = 0$.
On a toujours une forme indéterminée.
On reprend avec $f''(x) = -4\cos(2x)$, donc $f''(0) = -4$.
Et $g''(x) = 6x + 10$ donc $g''(0) = 10$.
Ainsi

$$\lim_{x\to 0} \frac{f(x)}{g(x)} = \lim_{x\to 0} \frac{f'(x)}{g'(x)} = \frac{-4}{10} = \frac{-2}{5}$$

3.5. Approximation polynomiale (développement limité)

Heuristique - Approximation polynomiale LOCALE
Nos fonctions usuelles ne sont pas des polynômes, on ne cherche par à remplacer une fonction usuelle par une fonction polynomiale sur un intervalle de $\mathbb{R}$ contenant une infinité de points.
Mais on peut faire des comparaison au voisinage d'un point.
La définition de Weierstrass de la dérivabilité de φ en x_0 avec pour valeur $A(=\varphi'(x_0))$ exprime :

$$\exists\, A \in \mathbb{R}, \exists\, \epsilon : I \to \mathbb{R} \text{ continue en } x_0 \text{ avec } \epsilon(x_0) = 0 \mid \forall\, x \in I, \varphi(x) = \varphi(x_0) + A(x-x_0) + (x-x_0)\epsilon(x)$$

On va donc chercher une fonction polynomiale f de degré n tel que

$$\varphi(x) = f(x) + (x-x_0)^n \epsilon(x) \qquad \text{avec } \epsilon(x) \underset{x\to x_0}{\longrightarrow} 0$$

, *On commence par étudier les cas $x_0 = 0$.*

Proposition - Développement polynomiale en 0
Soit $n \in \mathbb{N}^*$.
On considère φ de classe $\mathscr{C}^{n+1}$ sur I, contenant 0.
On note $T : x \mapsto \sum_{k=0}^{n} \frac{\varphi^{(k)}(0)}{k!} x^k$. Alors

$$\lim_{x\to 0} \frac{\varphi(x) - T(x)}{x^n} = 0$$

Autrement écrit : il existe $\epsilon : I \to \mathbb{R}$ tel que $\epsilon(x) \underset{x\to 0}{\longrightarrow} 0$ et $f(x) = T(x) + x^n\epsilon(x)$.

On commence par un lemme (qui sera enrichit dans la suite du cours)

Lemme - Dérivation monôme
Notons $m_s : x \mapsto x^s$.
Alors pour tout $k \in \mathbb{N}$, $m_s^{(k)}(0) = 0$ si $k \neq s$ et $m_s^{(s)}(0) = s!$.

Commençons par démontrer le lemme

Démonstration
Pour le lemme, on remarque que m_s est une fonction polynomiale donc de classe $\mathscr{C}^\infty$.
Par récurrence (finie), on note pour $k \leqslant s$, $\mathscr{P}_k : m_s^{(k)} : x \mapsto \frac{s!}{(s-k)!} x^{s-k}$.

— $\mathscr{P}_0$ est vraie
— Soit $k \leqslant s-1$. Supposons que $\mathscr{P}_k$ est vraie.

$$m_s^{(k+1)}(x) = (m_s^{(k)})'(x) = \left(x \mapsto \frac{s!}{(s-k)!} x^{s-k}\right)'(x) = \frac{(s-k)\times s!}{(s-k)!} x^{s-k-1} = \frac{s!}{(s-k-1)!} x^{s-k-1}$$

Donc $\mathscr{P}_{k+1}$ est vraie.

On a donc, ensuite, $m_s^{(s)}(x) = s!$ (constante) et pour tout $k > s$, $m_s^{(k)}(x) = 0$, en particulier $m_s^{(k)}(0) = 0$.
□

Démonstration
On applique la règle de L'hopsital.
1. Notons $den : x \mapsto x^n$. Donc $den = s_n$.
En application directe du lemme : $den^{(k)}(0) = 0$ si $k \leqslant n-1$ et $den^{(n)}(0) = n! (\neq 0)$.
2. Notons $num : x \mapsto f(x) - T(x)$.
Par linéarité (addition ici) n est de classe $\mathscr{C}^n$ et pour tout $k \leqslant n$, $(f-T)^{(k)}(0) = f^{(k)}(0) - T^{(k)}(0)$.
Puis par linéarité de la dérivation d'un polynôme (pour $k \leqslant n$) :

$$T^{(k)} = \left(\sum_{i=0}^{n} [T]_i s_i\right)^{(k)} = \sum_{i=0}^{n} [T]_i s_i^{(k)} = 0 + \cdots + 0 + [T]_k k! + 0 + \cdots + 0 = f^{(k)}(0)$$

On a donc $num^{(k)}(0) = 0$, pour tout $k \leqslant n$. On peut enfin appliquer la règle de L'Hospital en remontant : Pour tout k de 1 à n, $\frac{num^{n-k}}{den^{n-k}} \underset{x\to 0}{\longrightarrow} 0$ □

Exemple - DL de $x \mapsto (1+x)^\alpha$ ($\alpha \notin \mathbb{N}$)
Notons $f_\alpha : x \mapsto (1+x)^\alpha$.
Alors $f'_\alpha(x) = \alpha(1+x)^{\alpha-1}$, et par récurrence :

$$f_\alpha^{(k)}(x) = \prod_{i=0}^{k-1} (\alpha - i) \times (1+x)^{\alpha-k}$$

Et donc $f(x) = \sum_{k=0}^{n} \frac{\alpha(\alpha-1)\cdots(\alpha-k+1)}{k!} x^k + x^n\epsilon(x)$.

Exercice
Quelle formule obtient-on pour $\alpha = -1$?

Correction
La formule de Lagoute $\frac{1}{1+x} = 1 - x + x^2 - x^3 + \cdots + (-1)^n x^n + x^n\epsilon(x)$ ou suite géomtérique
Exercice

1. Donner une équation différentielle simple vérifiée par exp
2. En déduire le $DL_n(0)$ de exp.
3. En déduire le $DL_n(0)$ de ch et sh
4. Exprimer, pour tout $n \in \mathbb{N}$, la dérivée n-ième de $x \mapsto e^{ix}$.
5. En déduire le $DL_n(0)$ de cos et sin

Correction

1. On note $y = \exp$, on a $y' = y$.
2. Par récurrence immédiate : $y^{(n)} = y = \exp$ et donc $\exp^{(n)}(0) = e^0 = 1$.
   Donc pour x proche de 0, $\exp(x) = \sum_{k=0}^{n} \frac{1}{k!} x^k + x^n \epsilon(x)$.
3. Par linéarité de la dérivation (ou du calcul polynomial) ou par périodicité des dérivées :
   Pour x proche de 0, $\mathrm{ch}(x) = \sum_{k=0}^{\lfloor n/2 \rfloor} \frac{1}{(2k)!} x^{2k} + x^n \epsilon(x)$
   Pour x proche de 0, $\mathrm{sh}(x) = \sum_{k=0}^{\lfloor n/2 \rfloor} \frac{1}{(2k+1)!} x^{2k+1} + x^n \epsilon(x)$
4. Notons $E : x \mapsto e^{ix}$, on a $E^{(n)} = i^n e^{ix} = e^{ix + n\frac{\pi}{2}} = \cos(x + n\frac{\pi}{2}) + i \sin(x + n\frac{\pi}{2})$.
   En prenant les parties réelles et imaginaires :
   Pour x proche de 0, $\cos(x) = \sum_{k=0}^{\lfloor n/2 \rfloor} (-1)^k \frac{1}{(2k)!} x^{2k} + x^n \epsilon(x)$
   Pour x proche de 0, $\sin(x) = \sum_{k=0}^{\lfloor n/2 \rfloor} (-1)^k \frac{1}{(2k+1)!} x^{2k+1} + x^n \epsilon(x)$

Remarque - Développement limité
L'exercice qui suit, prépare à la théorie des DL (second semestre). On a un stock de DL connue (fonction usuelle), puis on combine le résultat avec des calculs polynomiaux

Exercice

On rappelle que $e^x = 1 + x + \frac{x^2}{2} + o(x^2)$ au voisinage de 0 et $\cos(x) = 1 - \frac{x^2}{2} + \text{ø}(x^2)$ au voisinage de 0.
Donner le DL_2 au voisinage de 0 de $x \mapsto \cos(x)e^x$ et de $x \mapsto \cos(e^x - 1)$

Correction
Le produit polynomial :

$$e^x \cos x = (1 + x + \frac{1}{2}x^2)(1 - \frac{1}{2}x^2) + o(x^2) = 1 + x - \frac{1}{2}x^3 - \frac{1}{4}x^4 + o(x^2) = 1 + x + o(x^2)$$

$$\cos(\underbrace{e^x - 1}_{:=u \to 0}) = 1 - \frac{1}{2}u^2 + o(u^2) = 1 - (x + \frac{x^2}{2})^2 + o(x^2) = 1 - x^2 + o(x^2)$$

4. Dérivation de fonctions réelles à valeurs complexes

4.1. Fonctions à valeurs complexes

Aparté. L'exponentielle complexe

Définition - Exponentielle complexe
Soit $z \in \mathbb{C}$. On définit l'exponentielle complexe de z par

$$\exp z = e^z = e^{\mathrm{Re}(z)} e^{i\mathrm{Im}(z)}.$$

Son module est $e^{\mathrm{Re}(z)}$ et un argument est $\mathrm{Im}(z)$.

Proposition - Elargissement
On retrouve les propriétés classiques de l'exponentielle :
- L'exponentielle complexe coïncide bien avec l'exponentielle sur $\mathbb{R}$ pour les réels ou avec l'exponentielle des imaginaires purs.
- Pour $(z, z') \in \mathbb{C}^2$, on a : $e^{z+z'} = e^z e^{z'}$ et $\dfrac{1}{e^z} = e^{-z}$.

Démonstration
Si $z = a + 0i \in \mathbb{R}$, alors $\exp(z) = e^a$.
Si $z = a + ib$ et $z' = a' + ib'$ (notation standard), alors

$$e^{z+z'} = e^{(a+a')+i(b+b')} = e^{a+a'}\,e^{i(b+b')} = e^a e^{a'} e^{ib} e^{ib'} = e^{a+ib} e^{a'+ib'} = e^z e^{z'}$$

En particulier :

$$e^{-z}e^z = e^{-z+z} = e^0 = 1 \qquad \Longrightarrow \qquad e^{-z} = \frac{1}{e^z}$$

□

Exercice

1. Résoudre $e^z = 1$.
2. Résoudre $e^z = 1 - i\sqrt{3}$.
3. Résoudre plus généralement $e^z = z_0$
4. Déterminer l'image d'une droite d'équation $x = a$ par l'application $\exp : \mathbb{C} \to \mathbb{C}^*$.
5. Déterminer l'image d'une droite d'équation $y = b$ par l'application $\exp : \mathbb{C} \to \mathbb{C}^*$.

Correction

1. Soit $z \in \mathbb{C}$.
$$e^z = 1 \iff \mathrm{Re}(z) = 0, \mathrm{Im}(z) \equiv 0[2\pi] \iff z \in \{2ik\pi, k \in \mathbb{Z}\}$$
2. Soit $z \in \mathbb{C}$
$$e^z = 1 - i\sqrt{3} \iff e^{\mathrm{Re}(z)} = \sqrt{1+3} = 2, \mathrm{Im}(z) \equiv \frac{\pi}{3}[2\pi] \iff z \in \{\ln 2 + i(\frac{\pi}{3} + 2k\pi), k \in \mathbb{Z}\}$$
3. Résoudre plus généralement $e^z = z_0$.
$$e^z = z_0 \iff e^{\mathrm{Re}(z)} = |z_0|, \mathrm{Im}(z) \equiv \arg(z_0)[2\pi] \iff z \in \{\ln|z_0| + i(\arg(z_0) + 2k\pi), k \in \mathbb{Z}\}$$
4. Déterminer l'image d'une droite $\mathscr{D}_a$ d'équation $x = a$ par l'application $\exp : \mathbb{C} \to \mathbb{C}^*$.
$z \in \mathscr{D}_a \iff \exists\, y \in \mathbb{R} \mid z = a + iy$. Et $\exp(z) = e^a e^{iy}$. Donc :
$$\exp(\mathscr{D}_a) = \mathscr{C}_{e^a},$$
cercle de centre O et de rayon e^a.
Il faudrait vérifier (trivialement) l'inclusion réciproque...
5. Déterminer l'image d'une droite $\mathscr{D}_b$ d'équation $y = b$ par l'application $\exp : \mathbb{C} \to \mathbb{C}^*$.
$z \in \mathscr{D}_b \iff \exists\, x \in \mathbb{R} \mid z = x + ib$. Et $\exp(z) = e^x e^{ib}$. Donc :
$$\exp(\mathscr{D}_a) = \mathscr{D}_{\theta_b},$$
demi-droite d'origine O et faisant un angle b avec l'axe des abscisses.
Il faudrait vérifier (trivialement) l'inclusion réciproque...

Proposition - Résolution d'équation
— Pour $(z, z') \in \mathbb{C}^2$, $\exp z = \exp z'$ si et seulement si $z - z' \in 2i\pi\mathbb{Z}$.
— Tout complexe $z_0 \neq 0$ peut s'écrire sous la forme $\exp z$.

Démonstration
Notons $z = a + ib$ et $z' = a' + ib'$.

$$e^z = e^{z'} \iff e^a e^{ib} = e^{a'} e^{ib'} \iff \begin{cases} e^a = e^{a'} \\ b \equiv b'[2\pi] \end{cases} \iff z - z' \in 2i\pi\mathbb{Z}$$

Par construction, si $z_0 \neq 0$, en prenant $z = \ln(|z_0|) + i\arg(z_0)$ (possible car $|z_0| \neq 0$) :

$$e^z = e^{\ln(|z_0|)} e^{i\arg(z_0)} = |z_0| e^{i\arg(z_0)} = z_0$$

□

Fonctions d'une variable réelle, à valeurs complexes

I désigne un intervalle de $\mathbb{R}$.

Définition - Découpage d'une fonction à valeurs dans $\mathbb{C}$
Soient $f_1 : I \to \mathbb{R}$ et $f_2 : I \to \mathbb{R}$. On pose : $\forall x \in I,\ f(x) = f_1(x) + i f_2(x)$.
$f : I \to \mathbb{C}$ est une fonction définie sur I à valeurs dans $\mathbb{C}$ *(si I n'est pas précisé au départ, son domaine de définition est l'intersection de ceux de f_1 et f_2).*
On définit les fonctions $\mathrm{Re}(f)$ (partie réelle de f) et $\mathrm{Im}(f)$ (partie imaginaire de f) à valeurs dans $\mathbb{R}$ par

$$\mathrm{Re}(f) = f_1 : \begin{array}{ccc} I & \to & \mathbb{R} \\ x & \mapsto & \mathrm{Re}(f(x)) \end{array} \qquad \mathrm{Im}(f) = f_2 : \begin{array}{ccc} I & \to & \mathbb{R} \\ x & \mapsto & \mathrm{Im}(f(x)) \end{array}$$

On dit que f est continue sur I si f_1 et f_2 sont continues sur I.

Exemple - $t \mapsto e^{it}$

L'exemple classique (simple) est la fonction circulaire exponentielle : $\exp_i : t \mapsto \cos t + i \sin t$.
Sa partie réelle est la fonction cosinus et sa partie imaginaire est la fonction sinus.
On a vu qu'il s'agissait bien d'une fonction exponentielle car vérifiant :

$$\exp_i(a+b) = \exp_i(a) \times \exp_i(b)$$

mais à valeurs complexes et non réels (donc on perd $\exp_i(a) > 0$, les limites à l'infini...).

Analyse - Justifions que sa base est bien e^i

On a défini $\exp(x) = e^x$ comme la limite de $x_n = \left(1 + \frac{x}{n}\right)^n$.
Que se passe-t-il si l'on considère $x = it$?
x_n est alors une suite à valeurs complexes dont le module est. :

$$\left|1 + \frac{ix}{n}\right|^n = \sqrt{\left(1 - \frac{x^2}{n^2}\right)^n}$$

Comme $\frac{x^2}{n}$ tend vers 0, à partir d'un certain rang :

$$1 - n\frac{x^2}{n^2} \leqslant \left(1 - \frac{x^2}{n^2}\right)^n \leqslant \frac{1}{1 + n\frac{x^2}{n^2}}$$

Le deux termes à gauche et à droite tendent vers 1, par encadrement (et continuité de $\sqrt{\cdot}$), $|x_n| \to \sqrt{1} = 1$.
Et si θ_n est l'argument de x_n, on a

$$\theta_n = \arg\left(1 + \frac{ix}{n}\right)^n = n \arg(1 + i\frac{x}{n}) = n \arctan\frac{x}{n}$$

On démontre que pour $u \to 0$, $\frac{\arctan u}{u} \to 1$, donc $\dfrac{\arctan\frac{x}{n}}{\frac{x}{n}} \to 1$ et donc $\theta_n \to x$.
Ainsi $x_n \to \cos x + i \sin x$, ce qui justifie officiellement le signe égal de $e^{ix} = \cos x + i \sin x$.

4.2. Dérivation d'une fonctions d'une variable réelle, à valeurs complexes

I désigne un intervalle de $\mathbb{R}$.

Définition - Dérivation d'une fonction à valeurs dans $\mathbb{C}$
Soit $f : \mathbb{R} \to \mathbb{C}$, une fonction à valeurs complexes.
Soient $f_1 : I \to \mathbb{R}$ et $f_2 : I \to \mathbb{R}$ telles que $\forall x \in I$, $f(x) = f_1(x) + i f_2(x)$.
On dit que f est dérivable, respectivement de classe $\mathscr{C}^1$ sur I si f_1 et f_2 sont dérivables, respectivement de classe $\mathscr{C}^1$ sur I.
Si f est dérivable sur I, on définit sa dérivée par :

$$\forall x \in I, \quad f'(x) = f_1'(x) + i f_2'(x).$$

On note $\mathscr{C}(I, \mathbb{C})$ (resp. $\mathscr{C}^1(I, \mathbb{C})$) l'ensemble des fonctions continues (resp. de classe $\mathscr{C}^1$ de I dans $\mathbb{C}$).

Remarque - Commutativité de la dérivation complexe et de la partie réelle ou imaginaire

Si f à valeurs complexes est dérivable on a donc $(\mathrm{Re}(f))' = \mathrm{Re}(f')$ et $(\mathrm{Im}(f))' = \mathrm{Im}(f')$.

Proposition - Constance et dérivation
Soit $f : I \to \mathbb{C}$ dérivable sur I. Alors f est constante sur I si et seulement si f' est nulle sur I.

Pour la démonstration, on admet ces résultats pour des fonctions de $\mathbb{R}$ dans $\mathbb{R}$.

Démonstration
Sur I :

$$f' = 0 \iff \begin{cases} f_1' = 0 \\ f_2' = 0 \end{cases} \iff \begin{cases} f_1 \text{ constante} \\ f_2 \text{ constante} \end{cases} \iff f \text{ constante}$$

□

4.3. Propriétés

Attention - Croissance sur $\mathbb{C}$?
Parler de croissance (ou de décroissance) d'une fonction à valeurs dans $\mathbb{C}$ n'a pas de sens.

Proposition - Opérations
Soient f et g deux fonctions dérivables sur I à valeurs dans $\mathbb{C}$ et $\alpha \in \mathbb{C}$. Alors $f+g$, fg et αf sont dérivables sur I et

$$\begin{aligned} \forall x \in I, \quad (f+g)'(x) &= f'(x) + g'(x) \\ (fg)'(x) &= f'(x)\,g(x) + f(x)\,g'(x) \\ (\alpha f)' &= \alpha f' \end{aligned}$$

Si g ne s'annule pas, alors $\dfrac{1}{g}$ et $\dfrac{f}{g}$ sont dérivables sur I et

$$\begin{aligned} \forall x \in I, \quad \left(\frac{1}{g}\right)'(x) &= \frac{-g'(x)}{g(x)^2} \\ \left(\frac{f}{g}\right)'(x) &= \frac{f'(x)g(x) - f(x)g'(x)}{g(x)^2} \end{aligned}$$

Exercice
Faire les démonstration.
Il suffit d'exploiter la commutativité entre dérivation et partie réelle ou imaginaire.

Correction

4.4. Composition avec l'exponentielle complexe

Théorème - Dérivée de l'exponentielle complexe
Soit $\phi : I \to \mathbb{C}$ dérivable sur I. Alors

$$\begin{aligned} \psi : \quad I &\to \mathbb{C} \\ x &\mapsto e^{\phi(x)} \end{aligned}$$

est dérivable sur I et : $\forall x \in I, \quad \psi'(x) = \phi'(x) e^{\phi(x)}$

Démonstration
On note $\phi = \phi_1 + i\phi_2$.
On a alors

$$\forall\, x \in I, \qquad \psi(x) = e^{\phi_1(x)} \times e^{i\phi_2(x)} = e^{\phi_1(x)} \times (\cos(\phi_2(x)) + i\sin(\phi_2(x)))$$

dérivable par produit :

$$\begin{array}{rcl}\forall\, x\in I,\quad \psi'(x) &=& \phi_1'(x)e^{\phi 1(x)}e^{i\phi_2(x)}+e^{\phi_1(x)}\phi_2'(x)(-\sin(\phi_2(x))+i\cos(\phi_2(x)))\\ &=& \phi_1'(x)e^{\phi 1(x)}e^{i\phi_2(x)}+e^{\phi_1(x)}\phi_2'(x)i(i\sin(\phi_2(x))+\cos(\phi_2(x)))\\ &=& (\phi_1'(x)+i\phi_2'(x))e^{\phi(x)}=\phi'(x)e^{\phi(x)}\end{array}$$

□

⚠ Attention - i joue un rôle comparable à celle d'un nombre réel

↝ On fera bien attention à ne pas oublier le i dans la dérivation !

Exercice

Calculer la dérivée $f : x \mapsto (\arcsin(x))e^{2x+ix^2}$.

Correction

f est dérivable sur $\mathbb{C}$, comme produit et composition de fonctions dérivables sur $]-1,1[$.

$$\forall\, x\in]-1,1[,\qquad f'(x)=\left(\frac{1}{\sqrt{1-x^2}}+(2+2ix)\arcsin(x)\right)e^{2x+ix^2}$$

Application - Fonction polaire

On considère $g : \mathbb{R}\to\mathbb{C}$, $\theta\mapsto\rho(\theta)e^{i\theta}$. On dit que l'on définie une fonction polaire.
On suppose que $\rho : \mathbb{R}\to\mathbb{R}_+$ est dérivable sur I.

• Montrons que g est dérivable et calculons g'.
Par produit de fonctions dérivables sur I, g est dérivable sur I.
Et pour tout $x\in I$, $g'(\theta)=\rho'(\theta)e^{i\theta}+i\rho(\theta)e^{i\theta}$.

• Il arrive que pour ce genre de fonction, on ait besoin de calculer $|g'(\theta)|$.
On utilise la quantité conjuguée :

$$\begin{array}{rcl}\forall\,\theta\in I,\quad |g'(\theta)|^2 &=& g'(\theta)\overline{g'(\theta)}=(\rho'(\theta)e^{i\theta}+i\rho(\theta)e^{i\theta})\overline{(\rho'(\theta)e^{i\theta}+i\rho(\theta)e^{i\theta})}\\ &=& (\rho'(\theta)e^{i\theta}+i\rho(\theta)e^{i\theta})(\rho'(\theta)e^{-i\theta}-i\rho(\theta)e^{-i\theta})\\ &=& [\rho'(\theta)]^2+[\rho(\theta)]^2\end{array}$$

Donc pour tout $\theta\in I$, $|g'(\theta)|=\sqrt{[\rho'(\theta)]^2+[\rho(\theta)]^2}$

5. Bilan

Synthèse

- ⇝ Au voisinage d'un point, on peut regarder comment la fonction évolue : c'est la valeur de la dérivée en ce point qui nous l'indique. On associe alors (si possible) f une nouvelle fonction : celle qui donne la valeur de dérivée en tout point. MIRACLE. Le passage $f\to f'$ se décrit très bien en terme d'algorithme (sans avoir à passer par le nombre dérivée). Il faut donc apprendre des tas de règles de calculs (à démontrer, par ailleurs...).
- ⇝ Ainsi plusieurs fonctions de référence sont à connaître : exponentielle(s) et logarithme(s) en toute base ; les fonctions puissances ; les fonctions circulaires (sin, arc cos...) et hyperboliques directes. On doit savoir comment comparer ces fonctions lorsqu'elles sont en compétition au voisinage de point problématique.
- ⇝ Le théorème de la bijection est un résultat important. Il faut le connaître, même si pour l'heure la démonstration nous échappe sans une définition précise de la continuité (et donc de $\mathbb{R}$).
- ⇝ En retour, la fonction dérivée permet de mieux comprendre localement la fonction. On peut alors étudier des variations, faire de l'optimisation (locale) ou lever des indéterminations (pour du calcul de limite).
- ⇝ Les fonctions complexes de la variables réelles s'étudient de la même façon (même si la représentation est plus complexe). En fait, ce qui compte, c'est la nature de la variable « de départ ».

Savoir-faire et Truc & Astuce du chapitre

- Truc & Astuce pour le calcul - A propos de la tangente
- Savoir-faire - Encadrement pour le calcul de limite
- Savoir-faire - Transférer un problème exponentiel de « en a » vers « en 0 ».
- Savoir-faire - Transférer un problème logarithme « en a » vers « en 1 ».
- Savoir-faire - Transférer un problème trigonométrique « en a » vers « en 0 ».
- Savoir-faire - Etudier si f est un $\mathscr{C}^1$-difféomorphisme
- Savoir-faire - Obtenir une inégalité
- Savoir-faire - Obtenir un maximum (dérivation)
- Savoir-faire - Lever une indétermination par la règle de L'Hospital
- Savoir-faire - Etude des branches infinies (et définition)

Retour sur les problèmes

19. On a une fonction $A(a_0+e)$ qui donne l'aire en fonction de a_0+e. Elle est optimale en a_0 lorsque $A(a_0-\epsilon) < A(a_0)$ et $A(a_0+\epsilon) < A(a_0)$ avec $\epsilon > 0$.
    Le calcul $A(a_0+e)$ donne exactement $A(a_0)+eA'(a_0)+\ldots$. On trouve, avec la méthode de Fermat : $A'(a_0)=0$, ce qui donne bien l'optimal de A.
20. Il faut nécessairement que Résultat$'$(Action) = 0, ce qui conduit à une équation donc la valeur recherchée Action_0 est solution. Malheureusement, la condition bien que nécessaire n'est pas suffisante.
21. C'est un vrai miracle que cela soit aussi simple (et que le tableau ne soit pas plus compliqué...).
    A moins que cela soit l'inverse : c'est parce que $y'=y$ se rencontre partout que l'exponentielle est si importante...
22. Cours
23. C'est le but de la formule de L'Hospital.
24. On vient de terminer ce chapitre en répondant à cette question.

Chapitre 9

Fonctions primitives et équations différentielles

Résumé -

Dans ce chapitre, le point de vue adopté est celui de Newton et Leibniz puis Euler, qui font de l'intégration l'opération inverse de la dérivation, c'est-à-dire que si f est une fonction définie sur I intervalle de $\mathbb{R}$ contenant a et b, admettant une primitive F, on définit l'intégrale de a à b de f par $\int_a^b f(t)\,dt = F(b)-F(a) = \left[F(t)\right]_a^b$.
Par propriétés de la dérivation (linéarité et primitivation d'inégalités), on obtient la linéarité et la croissance de l'intégrale ainsi définie. On s'exercera aussi techniquement pour maitriser ce calcul intégral, le fameux calculus.
Hypothèse forte : on admet toujours qu'une fonction continue sur un intervalle I admet des primitives sur I. La démonstration aura lieu plus tard.
Dans la très grande famille des équations différentielles, nous nous concentrons uniquement sur les équations différentielles linéaires d'ordre 1 et les équations différentielles d'ordre 2 à coefficients constants. Cela réduit largement les équations différentielles rencontrées, mais la raison est bonne : ce sont celles que l'on rencontre le plus souvent (à cause de la linéarisation des problèmes physiques) et ce sont celles que l'on sait résoudre (la fonction exponentielle a été crée pour cela).
Quelques liens youtube :

- *Hedacademy - Les primitives. https ://www.youtube.com/watch ?v=05ikcAFcPZ0*
- *Exo7 - Equation différentielle - https ://www.youtube.com/watch ?v=dkjXofPNMDo*
- *5min Lebesgue - forme idéale - https ://www.youtube.com/watch ?v=9x91d0UnBTw*

Sommaire

1. Problèmes

? Problème 40 - Lien primitve/intégrale

Calculer l'aire d'une surface est un « vieux » problème. Par exemple : quelle est l'aire d'une ellipse, d'une lunule ?
Optimiser, trouver un maximum ou un minimum est un autre « vieux » problème des mathématicien.
Existe-t-il un lien (profond, donc) entre les deux ?

? Problème 41 - Problèmes historiques

Galilée affirme en 1638 que la forme d'une chaîne suspendue entre deux clous est presqu'une parabole. Huygens, démontre 20 ans plus tard que ceci est faux. La solution, appelée caténaire (ou chainette) est donnée une vingtaine d'année plus tard par Leibniz et Johann Bernoulli.
Lors du séjour de Leibniz à Paris (1672-1673) durant lequel il suit des cours d'Huygens, Claude Perrault lui pose le problème suivant : *quelle est la courbe qui a la propriété qu'en chacun de ses points P, le segment de la tangente entre P et l'axe x et de longueur constante a ?* Pour concrétiser cette question, Perrault tire de son gousset une "hotlogio portabili suae thecae argenteae" et la fait glisser sur la table. Il précise qu'aucun mathématicien parisien ni toulousain (Fermat) n'a été capable d'en trouver l'équation.

? Problème 42 - Primitivisation des opérations algébriques

Pour le calcul de dérivation, le cours est grossièrement constitué d'un tableau de dérivation usuelle et de trois savoir-faire : comment dériver une addition, comment dériver une multiplication, comment dériver une composition ?
Avec ça, on arrive à tout (ou presque)...
Qu'en est-il de la primitivisation (opération inverse de la dérivation) ? Un tableau semble tout à fait possible, la règle de l'addition semble également bien s'inverser. Mais pour la multiplication ? Et la composition ?
Quelle est la primitive de $f \times g$? Quelle est la primitive de $f \circ g$?

? Problème 43 - Fraction rationnelle

Il est facile d'intégrer (ici, trouver la primitive) de toutes fonctions polynomiales.
Est-il possible d'intégrer toute division de polynôme, c'est-à-dire toute fraction rationnelle ?
Dans le cas d'une forme factorisée, cela donne :
Quelle est une primitive de $x \mapsto \dfrac{P(x)}{(x-a)(x-b)^n(x^2+cx+d)^m}$ avec $c^2 - 4d < 0$?

? Problème 44 - Implication

Si une fonction est dérivable, alors elle est continue. La réciproque est bien entendu fausse.
Existe-t-il un lien entre « admettre une primitive » et « être continue » ?
Comment gère-t-on une fonction non continue ?
Par exemple, existe-t-il une primitive à tan ? Existe-t-il une primitive

continue à tan ?

? Problème 45 - Problème de M. Lagoute

Que dire de $y'' + y = 0$? Tout dire !
C'est le fameux oscillateur harmonique, vu et revu en cours de sciences physiques.

? Problème 46 - Existence. Unicité

Est-ce qu'une équation différentielle admet toujours une, une seule, solution ?
Ce n'est pas le cas de $y'' + y = 0$ qui admet au moins comme solution $x \mapsto \cos x$ et $x \mapsto \sin x$. En admette-t-elle d'autres ? Peut-on lister toutes les solutions ?
Et alors, quel type de contraintes ajoutées pour obtenir une équation différentielle avec une et une seule solution ?
Et par ailleurs, existe-t-il des équations différentielles sans solution ?

? Problème 47 - Linéarisation

Dans ce cas, nous étudions que des équations différentielles linéaires.
Comment faire lorsque l'équation différentille n'est pas linéaire ? Peut-on revenir au cas précédent ? Par ailleurs, existe-t-il des moyens complémentaires (informatiques, par exemple) pour étudier des équations différentielles variées ?

2. Primitives

I désigne un intervalle de $\mathbb{R}$.
On commence par se concentrer sur les primitives, bien qu'il ne s'agit que d'un cas particulier d'une équation différentielle. Mais c'est la méthode essentielle pour résoudre les équations différentielles linéaires d'ordre 1.

2.1. Définitions

Définition - Primitives
Soit f une fonction définie sur un intervalle I de $\mathbb{R}$ à valeurs dans $\mathbb{R}$ (resp. à valeurs dans $\mathbb{C}$). On appelle primitive de f sur I toute fonction F **dérivable** sur I à valeurs dans $\mathbb{R}$ (resp. à valeurs dans $\mathbb{C}$) telle que, sur I, $F' = f$.

Proposition - CNS de primitive sur $\mathbb{C}$
$F : I \to \mathbb{C}$ est une primitive de $f : I \to \mathbb{C}$
si et seulement si $\mathrm{Re}F$ et $\mathrm{Im}F$ sont des primitives de $\mathrm{Re}f$ et $\mathrm{Im}f$ (resp.).

Proposition - Définition à constante additive près
Deux primitives de f sur l'intervalle I diffèrent d'une constante.
C'est-à-dire que si F est une primitive de f sur I alors l'ensemble des primitives de f sur I est

$$\left\{x \mapsto F(x) + C ; C \in \mathbb{K}\right\}$$

Pour aller plus loin - Classe d'équivalence
On a déjà vu, comme il n'y a pas unicité de la primitive, primitiver devrait plutôt signifier : donner un ensemble, une classe d'équivalence pour la relation d'équivalence : $f \equiv g$ ssi $f - g \in \mathbb{R}$ (est une constante).

où $\mathbb{K} = \mathbb{R}$ (resp. $\mathbb{K} = \mathbb{C}$) si f est à valeurs dans $\mathbb{R}$ (resp. dans $\mathbb{C}$).

Démonstration
Soit F, une primitive de f. On a les équivalences :

$$\begin{aligned}\varphi \text{ est primitive de } f &\iff \varphi' = f = F' \iff (\varphi - F)' = 0\\ &\iff \exists\, C \in \mathbb{K} \text{ tel que } \varphi - F = C\end{aligned}$$

(ce dernier résultat bien connu, sera démontré un peu plus tard. □

2.2. Primitives usuelles

En reprenant simplement le tableau de dérivations des fonctions usuelles, on trouve :

Pour aller plus loin - Tableau fini ?
Pour être utile, une table de primitives doit comporter plusieurs centaines de pages. Mentionnons ici celles de Gröbner et Hofreiter (1949) et de Gradshteyn et Ryzhik (1980). Aujourd'hui de nombreux logiciels de calcul symbolique contiennent de telles tables

Proposition - Tableau des primitives usuelles des fonctions à valeurs réelles (1)

fonction	primitives (C est une constante réelle)
x^α où $\alpha \in \mathbb{R}\setminus\{-1\}$	$\dfrac{x^{\alpha+1}}{\alpha+1} + C$
$\dfrac{1}{x}$	$\ln\lvert x\rvert + C$
e^x	$e^x + C$
$e^{\beta x}$ où $\beta \in \mathbb{C}\setminus\{0\}$	$\dfrac{e^{\beta x}}{\beta} + C$
$\operatorname{sh}\beta x$ où $\beta \neq 0$	$\dfrac{\operatorname{ch}\beta x}{\beta} + C$
$\operatorname{ch}\beta x$ où $\beta \neq 0$	$\dfrac{\operatorname{sh}\beta x}{\beta} + C$
$\sin x$	$-\cos x + C$
$\cos x$	$\sin x + C$
$\sin\beta x$ où $\beta \neq 0$	$-\dfrac{\cos\beta x}{\beta} + C$
$\cos\beta x$ où $\beta \neq 0$	$\dfrac{\sin\beta x}{\beta} + C$
$1 + \tan^2 x$	$\tan x + C$
$\ln x$	$x\ln x - x + C$
$\dfrac{1}{1+x^2}$	$\arctan x + C$
$\dfrac{1}{\sqrt{1-x^2}}$	$\arcsin x + C$ ou $-\arccos x + C'$

Proposition - Tableau des primitives usuelles des fonctions à valeurs réelles (2)

fonction f de la forme :	primitive F
$u'(x)u(x)^\alpha$ où $\alpha \in \mathbb{R}\setminus\{-1\}$	$\dfrac{u(x)^{\alpha+1}}{\alpha+1} + C$
$\dfrac{u'(x)}{u(x)}$	$\ln\lvert u(x)\rvert + C$
$u'(x)e^{u(x)}$	$e^{u(x)} + C$

Ces formules sont valables sur tout **intervalle** I où f (ou u) est continue.

Attention - Primitive sur deux intervalles
Si f admet des primitives sur la réunion de deux intervalles disjoints, on peut avoir des constantes différentes sur chacun des deux intervalles.
On rencontrera particulièrement cette situation dans le chapitre sur les

, équations différentielles.

2.3. Quelques cas particuliers

Exponentielles et trigonométrie

Savoir faire - $e^{\alpha x}\cos\beta x$ ou $e^{\alpha x}\sin\beta x$

Pour $f(x)=e^{\alpha x}\cos\beta x$ ou $f(x)=e^{\alpha x}\sin\beta x$ $(\alpha,\beta\in\mathbb{R})$, il suffit de primitiver $e^{(\alpha+i\beta)x}$ et récupérer partie réelle ou imaginaire.

Savoir faire - $\sin^n x\cos^m x$

Pour $f(x)=\sin^n x\cos^m x$, on peut linéariser.
Rappelons que pour cela on utilise les formules de de Moivre : $\sin^n x=\left(\frac{e^{ix}-e^{-ix}}{2i}\right)^n\dots$

Exercice

Déterminer des primitives des fonctions suivantes :

$$f_1(x)=e^{3x}\sin(2x);\quad f_2(x)=\sin^3(x)\cos^4(x)$$

Correction

$f_1=\mathrm{Im}(e^{3x+2ix})=\mathrm{Im}(e^{(3+2i)x})$.

$$F_1(x)=\mathrm{Im}\left(\frac{1}{3+2i}e^{(3+2i)x}\right)=\mathrm{Im}\left(\frac{3-2i}{13}e^{(3+2i)x}\right)=\frac{e^{3x}}{13}\big(3\sin(2x)-2\cos(2x)\big)$$

$$f_2(x)=\left(\frac{e^{ix}-e^{-ix}}{2i}\right)^3\times\left(\frac{e^{ix}+e^{-ix}}{2}\right)^4=\frac{i}{128}\left(e^{3ix}-3e^{ix}+3e^{-ix}-e^{-3ix}\right)\left(e^{4ix}+4e^{2ix}+6+4e^{-2ix}+e^{-4ix}\right)$$

$$=\frac{i}{128}\left((e^{7ix}-e^{-7ix})+(e^{5ix}-e^{-5ix})-3(e^{3ix}-e^{-3ix})-3(e^{ix}-e^{-ix})\right)$$

$$=\frac{-1}{64}\left(\sin(7x)+\sin(5x)-3\sin(3x)-3\sin(x)\right)$$

Donc

$$F_2(x)=\frac{1}{448}\cos(7x)+\frac{1}{320}\cos(5x)-\frac{1}{64}\cos(3x)-\frac{3}{64}\cos(x)$$

Idée pour « vérifier » le calcul : faire un DL en 0 :

$$f_2(x)\sim x^3\qquad et\qquad \frac{-1}{64}\left(\sin(7x)+\sin(5x)-3\sin(3x)-3\sin(x)\right)=\frac{-1}{64}\left[(7+5-9-3)x-\frac{1}{6}(343+125-81-3)x^3+o(x^3)\right]=x^3+o(x^3)$$

Remarque - Si on a un doute…

On peut toujours vérifier en calculant la dérivée de la primitive obtenue.

Ici la dérivée de $x\mapsto\frac{e^{3x}}{13}\big(3\sin(2x)-2\cos(2x)\big)$ est

$$x\mapsto\frac{e^{3x}}{13}\big((9\sin(2x)-6\cos(2x))+(6\cos(2x)+4\sin(2x))\big)=e^{3x}\sin(2x)$$

Fractions rationnelles

Des résultats et un savoir-faire à connaître :

Proposition - Tableau de primitives de certaines fonctions rationnelles (fonctions à valeurs dans $\mathbb{R}$)

fonction	primitives (C est une constante réelle)
$\frac{1}{x-a}$	$\ln\lvert x-a\rvert + C$
$\frac{1}{(x-a)^n}$ où $n \in \mathbb{N}^* \setminus \{1\}$	$\frac{-1}{n-1}\frac{1}{(x-a)^{n-1}} + C$
$\frac{1}{x^2+a^2}$	$\frac{1}{a}\arctan\frac{x}{a} + C$
$\frac{2x+p}{x^2+px+q}$	$\ln\lvert x^2+px+q\rvert + C$
$\frac{2x+p}{(x^2+px+q)^n}$ où $n \in \mathbb{N}^* \setminus \{1\}$	$\frac{-1}{n-1}\frac{1}{(x^2+px+q)^{n-1}} + C$

Exercice
A démontrer

Correction
Il suffit de dériver les cases de droites

Remarque - Division euclidienne
Lorsque le polynôme au numérateur a un degré plus important que celui du dénominateur, on commence par effectuer une division euclidienne pour se trouver en présence des cas précédent.
Nous reverrons les divisions euclidiennes en fin de premier semestre

Pour aller plus loin - Avec l'arithmétique complexe (à manipuler avec précaution...)
On rappelle que $\arg(1+ix) = \arctan(x)$, *donc si* $z = e^{i\theta} = 1+ix$, *on a donc*

$$\ln(1+ix) = \ln(z) = i\theta = i\arg(1+ix) = i\arctan(x)$$

Donc

$$\int \frac{1}{x^2+1} = \frac{i}{2}\int \frac{1}{x+i} - \frac{1}{x-i}$$

$$= \frac{i}{2}(\ln(x+i) - \ln(x-i))$$

$$= \frac{i}{2}\left(\ln(1+\frac{i}{x}) - \ln(1-\frac{i}{x})\right)$$

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

En fait le logarithme complexe est définie à une constante $2i\pi$ *près...*

Savoir faire - Fractions rationnelles $\frac{1}{ax^2+bx+c}$

Pour les fractions rationnelles de la forme $f(x) = \frac{1}{ax^2+bx+c}$ $((a,b,c) \in \mathbb{R}^3, a \neq 0)$:

— si $\Delta > 0$, le trinôme a deux racines réelles distinctes α et β. On cherche alors $\lambda, \mu \in \mathbb{R}$ tels que

$$f(x) = \frac{\lambda}{x-\alpha} + \frac{\mu}{x-\beta}$$

et on primitive avec $\ln|.|$

$$F(x) = \lambda\ln\lvert x-\alpha\rvert + \mu\ln\lvert x-\beta\rvert = \ln\lvert(x-\alpha)^\lambda(x-\beta)^\mu\rvert$$

— si $\Delta = 0$, on a alors

$$f(x) = \frac{1}{a(x-\alpha)^2}$$

on primitive directement avec la fraction rationnelle

$$F(x) = \frac{-1}{a(x-\alpha)}$$

— si $\Delta < 0$, on met le dénominateur sous forme canonique

$$f(x) = \frac{1}{(x+\alpha)^2+\beta^2}$$

on reconnaît une fonction composée qui se primitive avec arctan

$$F(x) = \frac{1}{\beta}\arctan\left(\frac{x+\alpha}{\beta}\right)$$

Exercice
Déterminer des primitives des fonctions suivantes :

$$f_1(x) = \frac{1}{(x^2-x-2)}; \quad f_2(x) = \frac{1}{x^2+x+1}; \quad f_3(x) = \frac{2x-1}{x^2+x+1}; \quad f_4(x) = \frac{2x+1}{x^2+4x+4}.$$

Correction

$x^2-x-2=(x+1)(x-2)$, donc il existe $\lambda,\mu\in\mathbb{R}$ tel que $\forall\, x\in\mathbb{R}$: $\dfrac{1}{x^2-x-2}=\dfrac{\lambda}{x+1}+\dfrac{\mu}{x-2}=\dfrac{(\lambda+\mu)x+(-2\lambda+\mu)}{x^2-x-2}$.

On peut (et on doit...) prendre $\lambda=-\mu=\frac{-1}{3}$, on a donc $\dfrac{1}{x^2-x-2}=\dfrac{1}{3}\left(\dfrac{1}{x+2}-\dfrac{1}{x-1}\right)$ Donc une primitive de f_1 est $F_1:x\mapsto\dfrac{1}{3}\ln\left|\dfrac{x+2}{x-1}\right|+C$.

Le discriminant de x^2+x+1 est $\Delta=1-4=-3<0$, on a $x^2+x+1=(x+\frac{1}{2})^2+\frac{3}{4}$.

Une primitive de f_2 est donc $F_2:x\mapsto\dfrac{2}{\sqrt{3}}\arctan\dfrac{2x+1}{\sqrt{3}}+C$

$f_3(x)=\dfrac{2x+1}{x^2+x+1}-2\dfrac{1}{x^2+x+1}$, on reconnait $\frac{u'}{u}-2f_2$.

Donc une primitive de f_3 est $F_3:x\mapsto\ln(x^2+x+1)-2\dfrac{2}{\sqrt{3}}\arctan\dfrac{2x+1}{\sqrt{3}}+C$

Pourquoi n'est-il pas besoin de valeur absolue ici ? Enfin, $f_4(x)=\dfrac{2x+1}{(x+2)^2}=2\dfrac{x+2-\frac{3}{2}}{(x+2)^2}=\dfrac{2}{x+2}-3\dfrac{1}{(x+2)^2}$.

Donc une primitive de f_4 est $F_4:x\mapsto\ln|x+2|+3\dfrac{1}{x+2}+C$.

3. Intégrales

3.1. Théorème fondamental et conséquences

Extension de l'intégrale sur $\mathbb{C}$

Définition - Notation
Soit $f:I\to\mathbb{R}$, pour tout $a<b\in I$, on note

$$\int_a^b f(t)\mathrm{d}t$$

l'aire (algébrique) comprise entre les segments de droites $x=a$, $y=0$ et $x=b$, et la courbe $y=f(x)$.
Nous admettons son existence si f est continue sur $[a,b]$.

Définition - Intégrale de f sur $[a,b]$
Si $f:[a,b]\to\mathbb{C}$ est continue sur $[a,b]$, on appelle intégrale de f sur $[a,b]$ le nombre complexe

$$\int_a^b f(t)\,dt=\int_a^b \mathrm{Re}\,f(t)\,dt+i\int_a^b \mathrm{Im}\,f(t)\,\mathrm{d}t.$$

⚠ Attention - Définition ?
Est-ce vraiment une définition ? Non, car on ne sait pas bien ce qu'est ce calcul. Préciser qu'il s'agit du nombre obtenu à partir d'une primitive de f, c'est tourner en rond par rapport au théorème et au corollaire qui suivent.
A ce stade, on est obligé de prendre ce nombre comme construit à partir de f, a et b. Au second semestre, nous prendrons le temps de bien montrer l'existence et donner un algorithme de calcul de ce nombre. Nous verrons qu'il n'est pas nécessaire que f soit continue (f pourrait être moins régulière)

Exercice

Soit P une application polynomiale. On suppose que pour tout $m \in \mathbb{N}$, $\int_0^{2\pi} e^{-imt} P(e^{it}) \mathrm{d}t = 0$.
Montrer que P est nul

Correction

Pour fixer les notations, on suppose que $P(x) = \sum_{k=0}^{d} a_k x^k$.
Alors $\int_0^{2\pi} e^{-imt} P(e^{it}) \mathrm{d}t = \sum_{k=0}^{d} \int_0^{2\pi} a_k e^{i(k-m)t} \mathrm{d}t = 2\pi a_m + \sum_{k \neq m} \left[\frac{a_k}{i(k-m)\pi} e^{i(k-m)t} \right]_0^{2\pi} = 2\pi a_m$.
Donc pour tout $m \in \mathbb{N}$, $a_m = 0$.

Fonction : intégrale de sa borne supérieure

Théorème - Théorème fondamental du calcul différentiel
Soit f continue sur I, intervalle de $\mathbb{R}$, à valeurs dans $\mathbb{R}$ ou $\mathbb{C}$. Soit $a \in I$.
Alors la fonction

$$\begin{array}{rccc} F : & I & \to & K \\ & x & \mapsto & \displaystyle\int_a^x f(t)\,\mathrm{d}t \end{array}$$

est de classe C^1 (c'est-à-dire dérivable de dérivée continue) sur I et $F' = f$.
C'est de plus l'unique primitive de f nulle en $a \in I$.

Corollaire - Existence de primitive
Toute fonction continue sur I admet une primitive sur I.

Attention - Pas toutes les primitives avec cette notation
On n'obtient pas toutes les primitives ainsi.
Ainsi, pour $f(x) = \cos x$, la primitive $F(x) = \sin x + 2$ ne peut s'obtenir ainsi. En effet, on aurait alors $F(x) = \sin x - \sin a$, or il n'existe pas de $a \in \mathbb{R}$ tel que $\sin a = -2$.

Remarque - Démonstration ?
Ce théorème est admis, pour le démontrer il faudrait une meilleure définition de $\int_a^b f(t)\mathrm{d}t$

Fusion : primitive et intégrale (de sa borne supérieure)

Théorème - Calcul fondamental
Soit f continue sur I intervalle de $\mathbb{R}$ contenant a et b. Soit F une primitive de f. Alors

$$\int_a^b f(t)\,\mathrm{d}t = F(b) - F(a) = \Big[F(t)\Big]_a^b.$$

Définition - Notation par extension
On notera, par extension des notations précédentes :
- $\int^x f(t)\mathrm{d}t$, une primitive quelconque de f.
  On pourra même considérer avec cette notation l'ensemble de toutes les primitives de f
- $\big[F(t)\big]_{\cdot}^{x} = F(x)$

Corollaire - Avec f'
Soit f de classe C^1 sur I. Alors

$$\forall (a,x)\in I^2,\ f(x)-f(a)=\int_a^x f'(t)\,\mathrm{d}t.$$

3.2. Quelques propriétés de l'intégrale

Proposition - Linéarité, croissance, Chasles...
Pour des fonctions f et g continues sur un intervalle I à valeurs dans $\mathbb{R}$, on a, pour $a,b\in I$, les propriétés suivantes :

— **linéarité** : si λ et μ sont deux réels,

$$\int_a^b (\lambda f+\mu g)(t)\,\mathrm{d}t=\lambda\int_a^b f(t)\,dt+\mu\int_a^b g(t)\,\mathrm{d}t$$

— **relation de Chasles** : pour $c\in]a,b[$,

$$\int_a^b f(t)\,\mathrm{d}t=\int_a^c f(t)\,dt+\int_c^b f(t)\,\mathrm{d}t$$

— **positivité** : si $a\leqslant b$ **et** $\forall x\in[a,b]$, $f(x)\geqslant 0$ alors

$$\int_a^b f(t)\,\mathrm{d}t\geqslant 0$$

— **croissance** : si $a\leqslant b$ **et** $\forall x\in[a,b]$, $f(x)\geqslant g(x)$ alors

$$\int_a^b f(t)\,dt\geqslant\int_a^b g(t)\,\mathrm{d}t$$

Truc & Astuce pour le calcul - Encadrer une intégrale
Pour encadrer une intégrale, on encadre la fonction à intégrer

Savoir faire - Inégalité des accroissements finis (avec f de classe $\mathscr{C}^1$)
Si f est de classe $\mathscr{C}^1$ sur $[a,b]\subset I$.
Notons $M=\sup_{[a,b]} f'$ et $m=\inf_{[a,b]} f'$, on a donc, pour tout $t\in[a,b]$:

$$f'(t)-m\geqslant 0\qquad\text{et}\qquad M-f'(t)\geqslant 0$$

On intègre sur $[a,b]$:

$$f(b)-f(a)-m(b-a)=\int_a^b f'(t)\mathrm{d}t-m\int_a^b m\mathrm{d}t\geqslant 0\Longrightarrow f(b)-f(a)\geqslant m(b-a)$$

$$f(b)-f(a)-M(b-a)=\int_a^b f'(t)\mathrm{d}t-M\int_a^b m\mathrm{d}t\leqslant 0\Longrightarrow f(b)-f(a)\leqslant M(b-a)$$

Pour aller plus loin - Inégalité des accroissements finis
Ce savoir-faire de l'inégalité des accroissements finis sera améliore plus loin : il suffira que f soir dérivable (et non nécessairement de classe $\mathscr{C}^1$.

Remarque - Combinaison linéaire
Si f et g sont deux fonctions, λ et μ deux réels $\lambda f+\mu g$ s'appelle une combinaison linéaire à coefficients réels de f et g.

Proposition - Fonctions à valeurs complexes
Soient f, g deux fonctions continues sur $[a,b]$ à valeurs dans $\mathbb{C}$ et λ, μ deux complexes. Alors on a les propriétés suivantes :

— **linéarité** :

$$\int_a^b (\lambda f + \mu g)(t)\,dt = \lambda \int_a^b f(t)\,\mathrm{d}t + \mu \int_a^b g(t)\,\mathrm{d}t$$

— **relation de Chasles** : pour $c \in]a,b[$,

$$\int_a^b f(t)\,\mathrm{d}t = \int_a^c f(t)\,\mathrm{d}t + \int_c^b f(t)\,\mathrm{d}t$$

Attention - Sur $\mathbb{C}$, pas de relation d'ordre…
Donc la croissance de l'intégrale sur $\mathbb{C}$ n'a pas de sens. Mais on peut exploiter les modules, si l'on souhaite faire des encadrements de la partie réelle et la partie imaginaire…

Exemple - Cela reste vraie, même si $c \notin]a,b[$
Si, par exemple, on a $c < b < a$ et F une primitive de f sur $[c,a]$:

$$\int_c^a f(t)\mathrm{d}t = \int_c^b f(t)\mathrm{d}t + \int_b^a f(t)\mathrm{d}t$$

Donc $\int_a^b f(t)\mathrm{d}t = F(b) - F(a) = -\int_b^a f(t)\mathrm{d}t = \int_c^b f(t)\mathrm{d}t - \int_c^a f(t)\mathrm{d}t = \int_a^c f(t)\mathrm{d}t + \int_c^b f(t)\mathrm{d}t$.

3.3. Technique 1 : Intégration par parties

Enoncé

Théorème - Intégration par parties
Si u et v sont deux fonctions de classe C^1 sur un intervalle I de $\mathbb{R}$, à valeurs dans $\mathbb{R}$ ou $\mathbb{C}$, alors

$$\forall (a,b) \in I^2, \qquad \int_a^b u'(t)v(t)\,\mathrm{d}t = \Big[u(t)v(t)\Big]_a^b - \int_a^b u(t)v'(t)\,\mathrm{d}t$$

Démonstration
$u(a)v(a) - u(b)v(b) = [uv]_a^b = \int_a^b (uv)' = \int_a^b uv' + \int_a^b u'v$.
Donc $\int_a^b u'v = [uv]_a^b - \int_a^b uv'$ □

Obtenir une primitive

Savoir faire - Obtenir une primitive avec une IPP cachée
Pour calculer une primitive par IPP de $f = u'v$, notée

$$\int_{\cdot}^{x} f(t)\,\mathrm{d}t$$

(attention, il s'agit d'**une fonction** et non d'un scalaire), on peut écrire

$$\int_{\cdot}^{x} u'(t)v(t)\,\mathrm{d}t = u(x)v(x) - \int_{\cdot}^{x} u(t)v'(t)\,\mathrm{d}t + C$$

Exemple - Primitive de arctan

On cherche une primitive de arctan : $\int_.^x \arctan(t)\mathrm{d}t$. Les fonctions $u : t \mapsto t$ et $v : t \mapsto \arctan t$ sont de classe $\mathscr{C}^1$ sur $\mathbb{R}$.

$$\int_.^x \arctan(t)\mathrm{d}t = \int_.^x u'(t)v(t)\mathrm{d}t = [t\arctan t]_.^x - \int_.^x \frac{t}{1+t^2}\mathrm{d}t = x\arctan x - \frac{1}{2}\ln(1+x^2)+C$$

Exercice

Avec une intégration par parties trouver une primitive de $x \mapsto \frac{x^2}{(x^2+1)^2}$ puis de $x \mapsto \frac{1}{(x^2+1)^2}$.
On verra une autre méthode plus loin.

Correction

Les applications $u : x \mapsto \frac{-1}{x^2+1}$ et $v : x \mapsto \frac{1}{2}x$ sont de classe $\mathscr{C}^1$ sur $\mathbb{R}$.

$$\int \frac{x^2}{(x^2+1)^2}\mathrm{d}x = \int \frac{1}{2}x \times \frac{2x}{(x^2-1)^2}\mathrm{d}x = \int v(t)u'(t)\mathrm{d}t = u(x)v(x) - \int u(x)v'(x)\mathrm{d}x + C$$

$$= \frac{-x}{2(x^2+1)} + \int \frac{1}{2(x^2+1)}\mathrm{d}x + C = \frac{-1}{2}\left(\frac{x}{x^2+1} + \arctan(x)\right) + C$$

Puis

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

Par linéarité :

$$\int \frac{1}{(x^2+1)^2}\mathrm{d}x = \arctan x - \frac{1}{2}\left(\frac{-x}{x^2+1} + \arctan(x)\right) + C' = \frac{1}{2}\left(\arctan(x) + \frac{x}{x^2+1}\right) + C'$$

Primitives de $P(x)e^{\alpha x}$

Savoir faire - $f(x) = P(x)e^{\alpha x}$

Pour $f : t \mapsto P(t)e^{\alpha t}$ où P est une fonction polynomiale, on peut faire $\deg(P)$ intégrations par parties (IPP) en dérivant $v : t \mapsto P(t)$ et en intégrant $u' : t \mapsto e^{\alpha t}$.
On peut appliquer la même méthode pour $f : t \mapsto P(t)\sin(\alpha t)$ ou $f : t \mapsto P(t)\cos(\alpha t)$.

Remarque - Autre méthode

On peut aussi directement chercher une primitive de la forme $Q(t)e^{\alpha t}$ avec $\deg Q = \deg P$.
On dérive cette fonction et identifie les coefficients de Q.

Exercice

Calculer

$$I = \int_0^{\pi/2} t\sin t\,\mathrm{d}t, \quad J = \int_0^1 e^{2x}(6x^2+2x-4)\,\mathrm{d}x$$

Correction

Les fonctions $u : t \mapsto -\cos t$ et $v : t \mapsto t$ sont de classe $\mathscr{C}^1$ sur $\mathbb{R}$,

$$I = [-t\cos t]_0^{\pi/2} + \int_0^{\pi/2} \cos t = 0 + [\sin(t)]_0^{\pi/2} = 1$$

Les fonctions $u : x \mapsto \frac{1}{2}e^{2x}$, $v : x \mapsto 6x^2+2x-4$ et $w : x \mapsto 12x+2$ sont de classe $\mathscr{C}^1$ sur $\mathbb{R}$,

$$J = \left[\frac{1}{2}e^{2x}(6x^2+2x-4)\right]_0^1 - \frac{1}{2}\int_0^1 e^{2x}(12x+2)\mathrm{d}x = 2e^2+2-\frac{1}{4}\left[e^{2x}(12x+2)\right]_0^1 + \frac{1}{4}\int_0^1 12e^{2x}\mathrm{d}x$$

$$= 2e^2+2-\frac{14}{4}e^2+\frac{1}{2}+\frac{3}{2}e^2-\frac{3}{2} = 1$$

Exercice

Pour aller plus loin : trouver une formule générale de $\int e^{\alpha x}P(x)$ qui exploite les puissances de α et les dérivées de P

Correction

$$\int_a^b e^{\alpha x}P(x) = \sum_{k=0}^{n} \frac{(-1)^k}{\alpha^{k+1}}\left(e^{\alpha a}P^{(k)}(a) - e^{\alpha b}P^{(k)}(b)\right)$$

Pour aller plus loin - Fonction beta

On note $B(x,y) = \int_0^1 t^{x-1}(1-t)^{y-1}\mathrm{d}t$.
On montre par IPP que
$B(x,y+1) = \frac{y}{x+y}B(x,y)$.
Puis si n et m sont des entiers :

$$B(n+1,m+1) = \frac{n!m!}{(n+m+1)!} = \frac{1}{(n+m+1)\binom{n+m}{n}}$$

Cela peut donner un sens à $\binom{x+y}{x}$, *avec* $x, y \in \mathbb{R}\ldots$

Primitives de $P(x)\ln\big(Q(x)\big)$

Savoir faire - $f(x) = P(x)\ln\big(Q(x)\big)$

Pour $f : t \mapsto P(t)\ln\big(Q(t)\big)$ où P est une fonction polynomiale, on peut faire une intégrations par parties (IPP) en dérivant $v : t \mapsto \ln\big(Q(t)\big)$ et en intégrant $u' : t \mapsto P(t)$.
On se retrouve alors en présence d'une fraction rationnelle, que l'on sait intégrer, en principe...

Exercice
Calculer

$$I_b = \int_0^b x\ln(x^2+1)\,\mathrm{d}x.$$

On pourra remarquer que $x^3 = x(x^2+1) - x$...

Correction
Les fonctions $u : x \mapsto \frac{1}{2}x^2$ et $v : x \mapsto \ln(x^2+1)$ sont de classe $\mathscr{C}^1$ sur $\mathbb{R}$,

$$I_b = \left[\frac{1}{2}x^2\ln(x^2+1)\right]_0^b - \int_0^b \frac{x^3}{x^2+1}\,\mathrm{d}x = \frac{1}{2}b^2\ln(b^2+1) + \int_0^b \frac{x}{x^2+1} - x\mathrm{d}x$$
$$= \frac{1}{2}b^2\ln(b^2+1) + \frac{1}{2}\ln(b^2+1) - \frac{1}{2}b^2$$

3.4. Technique 2 : Changement de variables

Enoncé

Théorème - Changement de variable
Soient I, J des intervalles de $\mathbb{R}$, $\alpha, \beta \in I$,
Soient $\phi : I \to \mathbb{R}$ de classe C^1 telle que $\phi(I) \subset J$ et $f : J \to \mathbb{R}$ (ou $\mathbb{C}$) continue.
Alors

$$\int_{\phi(\alpha)}^{\phi(\beta)} f(t)\,\mathrm{d}t = \int_\alpha^\beta f(\varphi(x))\varphi'(x)\,\mathrm{d}x.$$

Démonstration
Notons F, une primitive de f, on a alors :

$$\int_{\phi(\alpha)}^{\phi(\beta)} f(t)\,\mathrm{d}t = F(\phi(\alpha)) - F(\phi(\beta)) = [F\circ\phi]_\alpha^\beta = \int_\alpha^\beta f(\varphi(x))\varphi'(x)\,\mathrm{d}x.$$

car $(F\circ\phi)' = \phi' \times F'\circ\phi = \phi' \times f\circ\phi$. □

Analyse - Deux cas possibles
Deux cas se produise : une application directe de ce théorème, ou bien une application avec la fonction réciproque de ϕ.

1. Cas direct (non bijectif) :
   On doit calculer $\displaystyle\int_a^b f(u)\mathrm{d}u$.
   On pose $u = \varphi(t)$, où φ est de classe $\mathscr{C}^1$ sur $[\alpha,\beta]$, avec $\varphi(\alpha) = a$ et $\varphi(\beta) = b$
   et donc $\mathrm{d}u = \varphi'(t)\mathrm{d}t$.
   Ainsi : $\displaystyle\int_a^b f(u)\,\mathrm{d}u = \int_\alpha^\beta f(\varphi(t))\varphi'(t)\,\mathrm{d}t$
2. Cas indirect (bijectif) :
   On doit calculer $\displaystyle\int_a^b f(u)\mathrm{d}u$.
   On pose $t = \psi(u)$, où ψ est de classe $\mathscr{C}^1$ sur $[a,b]$, et bijective de $[a,b]$ sur $[\alpha,\beta]$.
   avec $\varphi = \psi^{-1}$, on a donc $u = \varphi(t)$
   et donc $\mathrm{d}u = \varphi'(t)\mathrm{d}t = \dfrac{1}{\psi'(\varphi(t))}\mathrm{d}t \iff \psi'(u)\mathrm{d}u = \mathrm{d}t$.
   Ainsi : $\displaystyle\int_a^b f(u)\,\mathrm{d}u = \int_a^b f(\psi^{-1}(t))\frac{\mathrm{d}t}{\psi'(\psi^{-1}(t))}$

Mais, c'est la pratique qui compte, pas d'apprendre ces formulations!

Pour aller plus loin - Aire d'un cercle
Donner l'aire d'un cercle de rayon r. **AP**

Savoir faire - Changement de variable - dans la pratique

on veut calculer $\int_{\phi(\alpha)}^{\phi(\beta)} f(t)\,\mathrm{d}t$. On pose $t=\phi(x)$ (changement de variable), on remplace alors

- t par $\phi(x)$
- $\mathrm{d}t$ par $\phi'(x)\,\mathrm{d}x$ $\qquad (\phi'(x)=\dfrac{d\phi(x)}{dx}=\dfrac{dt}{dx})$
- t varie de $\phi(\alpha)$ à $\phi(\beta)$ par x varie de α à β (et inversement)

On peut faire un tableau

Exercice

Calculer par changement de variables les intégrales suivantes

$$I=\int_0^{\pi/2} x\sin(x^2)\,\mathrm{d}x,\quad J=\int_0^1 \sqrt{1-t^2}\,\mathrm{d}t$$

Correction

On considère $\phi : t\mapsto\sqrt{t}$, de classe $\mathscr{C}^1$ sur $\left[0,\frac{\pi^2}{4}\right]$,
On pose alors $x=\phi(t)$, et donc, comme $\phi'(t)=\frac{1}{2\sqrt{t}}$

$$I=\int_0^{\pi^2/4}\sqrt{t}\sin t\times\frac{1}{2\sqrt{t}}\mathrm{d}t=\left[\frac{-1}{2}\cos t\right]_0^{\pi^2/4}=\frac{1}{2}(1-\cos\frac{\pi^2}{4})$$

On considère $\psi : t\mapsto\sin(t)$, de classe $\mathscr{C}^1$ sur $\left[0,\frac{\pi}{2}\right]$,
On pose alors $x=\psi(t)$, et donc, comme $\psi'(t)=\cos(t)$

$$I=\int_0^{\pi/2}\sqrt{1-\sin^2 t}\times\cos(t)\mathrm{d}t=\frac{1}{2}\int_0^{\pi/2}(1+\cos(2t))\,\mathrm{d}t=\frac{1}{2}\left[t+\frac{1}{2}\sin(2t)\right]_0^{\pi/2}=\frac{\pi}{4}$$

(C'est l'aire d'un quart de disque...)

Application : calcul de primitive

Soulignons l'importance que ϕ soit bijective pour exploiter ϕ^{-1}...

Savoir faire - Calculer une primitive par changement de variable

On cherche une primitive F de f sur I,

- on pose $t=\phi(x)$ et donc $\mathrm{d}t=\phi'(x)\mathrm{d}x$ où ϕ est une **bijection de classe** C^1 de J sur I,
- on cherche une primitive $G(x)=\int f(\phi(x))\phi'(x)\,\mathrm{d}x$ et on prend $F(t)=G(\phi^{-1}(t))$.

Exemple - Primitive de $t\mapsto\frac{1}{t^2+a^2}$

Par exemple en posant $x=\frac{t}{a}$, on a

$$\int\frac{dt}{t^2+a^2}=\int\frac{1}{a}\frac{dx}{x^2+1}=\frac{1}{a}\arctan x+C=\frac{1}{a}\arctan\frac{t}{a}+C$$

Attention - Ne pas oublier de revenir à la variable de départ

Pour éviter les erreurs (oubli de revenir à la variable de départ...) on aurait intérêt à écrire

$$F(x)=\int_{\cdot}^{x}\frac{dt}{t^2+a^2}$$

Exercice

Soit $a \in \mathbb{R}$. Donner une primitive de $x \mapsto \dfrac{1}{(x^2+a^2)^2}$ en faisant le changement de variable $\tan t = \frac{x}{a}$.
On rappelle que $1+\tan^2 u = \dfrac{1}{\cos^2 u}$

Correction

On pose $\tan\theta = \frac{t}{a}$, $\frac{1}{a}\mathrm{d}t = (1+\tan^2\theta)\mathrm{d}\theta$

$$\int^x \frac{\mathrm{d}t}{(t^2+a^2)^2} = \frac{1}{a^4}\int^x \frac{\mathrm{d}t}{(1+(\frac{t}{a})^2)^2} = \frac{1}{a^3}\int^{\arctan(x/a)} \cos^2\theta \mathrm{d}\theta = \frac{1}{2a^3}\int^{\arctan(x/a)} (\cos 2\theta + 1)\mathrm{d}\theta$$

$$= \frac{1}{2a^3}\left[\frac{1}{2}\sin\left(2\arctan\frac{x}{a}\right) + \arctan\frac{x}{a}\right] + C = \frac{1}{2a^3}\left[\tan(\arctan\tfrac{x}{a})\cos^2(\arctan\tfrac{x}{a}) + \arctan\frac{x}{a}\right] + C$$

$$= \frac{1}{2a^3}\frac{x}{a} \times \frac{1}{1+\frac{x^2}{a^2}} + \frac{1}{2a^3}\arctan\frac{x}{a} + C = \frac{ax + \arctan\frac{x}{a}}{2a^3(a^2+x^2)} + C$$

Exercice

Aller plus loin : Donner l'expression des primitives de $x \mapsto \dfrac{1}{(x^2+a^2)^n}$

Correction

Savoir faire - Bijection par morceaux

Lorsque le changement de variable doit être bijectif, mais ne l'est que par morceaux, alors

1. on cherche une primitive sur chaque morceaux d'intervalle
2. on « recolle » chaque morceaux en ajustant les constante de manière à ce que la primitive soit bien continue.

L'exercice suivant illustre ce savoir-faire.

Exercice

Donner l'ensemble de définition et calculer la primitive de $h : x \mapsto \dfrac{1}{2+\cos x}$
On posera $t = \tan\frac{x}{2}$

Correction

h est définie et continue sur $\{x \in \mathbb{R} \mid \cos x \neq -2\} = \mathbb{R}$, elle admet donc des primitives sur cet ensemble.
D'abord, pour tout $x \in \mathbb{R}$, $h(x) = \dfrac{1}{2+\frac{1-\tan^2\frac{x}{2}}{1+\tan^2\frac{x}{2}}} = \dfrac{1+\tan\frac{x}{2}}{3+\tan\frac{x}{2}}$,

$$H(x) = \int^x \frac{1+\tan^2\frac{t}{2}}{3+\tan^2\frac{t}{2}}\mathrm{d}t$$

On réalise le changement de variable $u = \tan\frac{t}{2}$, donc $\mathrm{d}u = \frac{1}{2}(1+\tan^2\frac{t}{2})\mathrm{d}t$.

$$H(x) = \int^{\tan\frac{x}{2}} \frac{2}{3+u^2}\mathrm{d}u = \frac{2}{3}\int^{\tan\frac{x}{2}} \frac{\mathrm{d}u}{1+[\frac{u}{\sqrt{3}}]^2} = \frac{2}{\sqrt{3}}\left[\arctan\frac{u}{\sqrt{3}}\right]^{\tan\frac{x}{2}}$$

$$\exists K \in \mathbb{R}, \qquad H(x) = \frac{2}{\sqrt{3}}\arctan\left(\frac{\tan\frac{x}{2}}{\sqrt{3}}\right) + K$$

MAIS le changement de variable posée : $u = \tan\frac{t}{2}$ n'est pas bijective sur $\mathbb{R}$.
Il s'agit de bijection d'intervalle de la forme $]-\pi+2j\pi;\pi+2j\pi[$ sur $\mathbb{R}$.
Donc les solutions obtenues sont plutôt :

$$\exists (K_j)_{j\in\mathbb{Z}} \in \mathbb{R}^{\mathbb{Z}}, \qquad \forall\, x \in](2j-1)\pi;(2j+1)\pi[, H(x) = \frac{2}{\sqrt{3}}\arctan\left(\frac{\tan\frac{x}{2}}{\sqrt{3}}\right) + K_j$$

Puis comme H est dérivable, elle est nécessairement continue donc continue en $\frac{\pi}{2} + j\pi$.

$$\lim_{x\to((2j+1)\pi)^+} H(x) = \lim_{x\to((2(j+1)\pi)^-} J(x)$$

Et comme, par composition des limites :

$$\lim_{x\to((2j+1)\pi)^+} H(x) = \lim_{x\to((2j+1)\pi)^+} \frac{2}{\sqrt{3}}\arctan\left(\frac{\tan\frac{x}{2}}{\sqrt{3}}\right) + K_{j+1} = \lim_{y\to-\infty}\frac{2}{\sqrt{3}}\arctan(y) + K_{j+1} = -\frac{\pi}{\sqrt{3}} + K_{j+1}$$

$$\lim_{x\to((2j+1)\pi)^-} \frac{2}{\sqrt{3}}\arctan\left(\frac{\tan\frac{x}{2}}{\sqrt{3}}\right) + K_j = \lim_{y\to+\infty}\frac{2}{\sqrt{3}}\arctan(y) + K_j = \frac{\pi}{\sqrt{3}} + K_j$$

On a donc, pour tout $j \in \mathbb{Z}$,

$$-\frac{\pi}{\sqrt{3}}K_{j+1} = \frac{\pi}{\sqrt{3}} + K_j$$

Donc $K_{j+1} = K_j + \frac{2\pi}{\sqrt{3}}$. On reconnait une suite arithmétique : $K_j = K_0 + \frac{2\pi}{\sqrt{3}}j$.
Et finalement, comme

$$x \in](2j-1)\pi;(2j+1)\pi[\iff \frac{x}{\pi}+1 \in]2j,2j+2[\iff \frac{x+\pi}{2\pi} \in]j,2j+1[\iff j = \left\lfloor \frac{x+\pi}{2\pi} \right\rfloor$$

on peut affirmer :

$$\exists\, K(=K_0) \in \mathbb{R} \text{ tel que :} \qquad \forall\, x \in \mathbb{R}, \qquad H(x) = \frac{2}{\sqrt{3}}\arctan\left(\frac{\tan\frac{x}{2}}{\sqrt{3}}\right) + \frac{2\pi}{\sqrt{3}} \times \left\lfloor \frac{x+\pi}{2\pi} \right\rfloor + K$$

Fonctions définies à partir de fonctions trigonométrique

Savoir faire - Calcul pour $f(t) = \sin^n t \cos^m t$ avec n ou m impair

Pour $f(t) = \sin^n t \cos^m t$, on peut linéariser, ou,

- si n est impair, effectuer le changement de variables $u = \cos t$ (ou isoler un $\sin t$ et dans $\sin^{n-1} t$ remplacer $\sin^2 t$ par $1-\cos^2 t$ puis reconnaître des primitives),
- si m est impair, effectuer le changement de variables $u = \sin t$ (ou remplacer $\cos^2 t$ par $1-\sin^2 t$).

Exercice

Calculer par changement de variables l'intégrale suivante

$$\int_0^{\pi/2} \sin^2 u \cos^3 u \,\mathrm{d}u$$

Correction

On pose $t = \sin u$, de classe $\mathscr{C}^1$ sur $[0, \frac{\pi}{2}]$. On a donc $\mathrm{d}t = \cos u \mathrm{d}u$

$$\int_0^{\pi/2} \sin^2 u \cos^3 u \,\mathrm{d}u = \int_0^1 t^2(1-t^2)\mathrm{d}t = \left[\frac{1}{3}t^3 - \frac{1}{5}t^5\right]_0^1 = \frac{2}{15}$$

Savoir faire - Cas général ($\tan\frac{x}{2}$)

D'une manière générale, les changements de variables utiles pour les fonctions construites avec de fonctions trigonométriques sont $t = \cos x$, $t = \sin x$, $t = \tan x$, $t = \tan\frac{x}{2}$.

On rappelle que si $t = \tan\frac{x}{2}$, alors $\cos x = \dfrac{1-t^2}{1+t^2}$, $\sin x = \dfrac{2t}{1+t^2}$ et $\tan x = \dfrac{2t}{1-t^2}$.

Exercice

Calculer par changement de variables l'intégrale suivante

$$\int_{\pi/3}^{\pi/2} \frac{\mathrm{d}t}{\sin t}$$

Correction

On pose $t = \tan\frac{x}{2}$ de classe $\mathscr{C}^1$ sur $[\frac{\pi}{3}, \frac{\pi}{2}]$.
On a $\mathrm{d}t = \frac{1}{2}(1+\tan^2\frac{x}{2})\mathrm{d}x$, donc $\mathrm{d}x = \dfrac{2\mathrm{d}t}{1+t^2}$.

$$\int_{\pi/3}^{\pi/2} \frac{\mathrm{d}t}{\sin t} = \int_{1/\sqrt{3}}^1 \frac{t^2+1}{2t} \times \frac{2}{1+t^2}\mathrm{d}t = \int_{1/\sqrt{3}}^1 \frac{\mathrm{d}t}{t} = \frac{1}{2}\ln 3$$

Simplification des calculs

Théorème - Simplification des calculs
Soit $f : [-a, a] \to \mathbb{R}(\mathbb{C})$ une fonction continue :
- si f est paire, $\displaystyle\int_{-a}^{a} f(t)\mathrm{d}t = 2\int_{0}^{a} f(t)\mathrm{d}t$;
- si f est impaire, $\displaystyle\int_{-a}^{a} f(t)\mathrm{d}t = 0$;
- si $f : \mathbb{R} \to \mathbb{R}$ (ou $\mathbb{C}$) est une fonction continue T-périodique,
$$\int_{a+T}^{b+T} f(t)\,dt = \int_{a}^{b} f(t)\mathrm{d}t \text{ et } \int_{a}^{a+T} f(t)\mathrm{d}t = \int_{0}^{T} f(t)\mathrm{d}t.$$

Démonstration
Si f est paire,
$$\int_{-a}^{a} f(t)\mathrm{d}t = \int_{-a}^{0} f(t)\mathrm{d}t + \int_{0}^{a} f(t)\mathrm{d}t = \underbrace{\int_{a}^{0} f(-u)(-\mathrm{d}u)}_{u=-t} + \underbrace{\int_{0}^{a} f(u)(\mathrm{d}u)}_{u=t} = 2\int_{0}^{a} f(u)\mathrm{d}u$$

Si f est impaire,
$$\int_{-a}^{a} f(t)\mathrm{d}t = \int_{-a}^{0} f(t)\mathrm{d}t + \int_{0}^{a} f(t)\mathrm{d}t = \underbrace{\int_{a}^{0} f(-u)(-\mathrm{d}u)}_{u=-t} + \underbrace{\int_{0}^{a} f(u)(\mathrm{d}u)}_{u=t} = \int_{0}^{a} f(u)(\mathrm{d}u) - \int_{0}^{a} f(u)(\mathrm{d}u) = 0$$

Si f est T-périodique
$$\int_{a+T}^{b+T} f(t)\,dt = \underbrace{\int_{a}^{b} f(u+T)\mathrm{d}u}_{u=t-T} = \int_{a}^{b} f(u)\mathrm{d}u$$

Notons $n = \left\lfloor \frac{a}{T} \right\rfloor$, donc $nT \leqslant a < (n+1)T$:
$$\int_{a}^{a+T} f(t)\mathrm{d}t = \int_{a}^{(n+1)T} f(t)\mathrm{d}t + \int_{(n+1)T}^{a+T} f(t)\mathrm{d}t = \int_{a-nT}^{(n+1)T-nT} f(t)\mathrm{d}t + \int_{(n+1)T-(n+1)T}^{a+T-(n+1)T} f(t)\mathrm{d}t$$
$$= \int_{a-nT}^{T} f(t)\mathrm{d}t + \int_{0}^{a-nT} f(t)\mathrm{d}t = \int_{0}^{a-nT} f(t)\mathrm{d}t + \int_{a-nT}^{T} f(t)\mathrm{d}t = \int_{0}^{T} f(t)\mathrm{d}t$$
d'après la formule de Chasles □

Exemple - Réduction d'intervalle
Soit i, une fonction T périodique.
$$\int_{0}^{nT} i(t)\mathrm{d}t = n\int_{0}^{T} i(t)\mathrm{d}t$$

Variable dans les bornes de l'intégrale (composition)

Savoir faire - Variable dans le bornes de l'intégrale
Il arrive qu'on doit étudier des fonctions de la forme
$$g : x \mapsto \int_{f_1(x)}^{f_2(x)} h(t)\mathrm{d}t$$
, sans pouvoir exprimer explicitement H, une primitive de h.
Néanmoins, la simple existence de H, permet d'écrire :
$$g(x) = H(f_2(x)) - H(f_1(x))$$
Dont on déduit de nombreuses informations. Par exemple : g est dérivable si f_1 et f_2 le sont. Et dans ce cas :
$$\forall\, x \in \mathscr{D}, \qquad g'(x) = f_2'(x) \times h(f_2(x)) - f_1'(x) \times h(f_1(x))$$

Exercice

Soit $H : x \mapsto \displaystyle\int_x^{2x} \frac{\mathrm{d}t}{\sqrt{1+t^4}}$. Etudier la parité de H. Montrer que H est dérivable sur $\mathbb{R}$, calculer H' et dresser le tableau de variations de H. Déterminer les limites de H en $+\infty$ et $-\infty$.

Correction

$H(-x) = \displaystyle\int_{-x}^{-2x} \frac{\mathrm{d}t}{\sqrt{1+t^4}} = \int_x^{2x} \frac{-\mathrm{d}u}{\sqrt{1+u^4}} = -H(x)$ en faisant le changement de variable $u = -t$.
Donc H est impaire. On étudie sur $\mathbb{R}_+$, on appliquera la symétrie ensuite.
L'application $g : t \mapsto \dfrac{1}{\sqrt{1+t^4}}$ est continue sur $\mathbb{R}$, donc elle admet une primitive G sur $\mathbb{R}$.
On a alors $H(x) = G(2x) - G(x)$. Et par composition, H est dérivable sur $\mathbb{R}$ et

$$\forall\, x \in \mathbb{R}, \qquad H'(x) = 2G'(2x) - G'(x) = 2g(2x) - g(x) = \frac{2}{\sqrt{1+16x^4}} - \frac{1}{\sqrt{1+x^4}}$$

Pour tout $x > 0$,

$$H'(x) \leqslant 0 \iff \frac{4}{1+16x^4} \leqslant \frac{1}{1+x^4} \iff 3 - 12x^4 \leqslant 0 \iff 1 - 4x^4 \leqslant 0$$

$$H'(x) \leqslant 0 \iff (1-2x^2)(1+2x^2) \leqslant 0 \iff (1-\sqrt{2}x)(1+\sqrt{2}x) \leqslant 0 \iff x \in \left[0, \frac{1}{\sqrt{2}}\right]$$

donc H est croissante sur $\left[0, \frac{1}{\sqrt{2}}\right]$, puis décroissante sur $\left[\frac{1}{\sqrt{2}}, +\infty\right[$.
Par ailleurs $H(0) = 0$ et par décroissance de $t \mapsto \dfrac{1}{\sqrt{1+t^4}}$ sur $[x, 2x]$,

$$0 \leqslant H(x) \leqslant \int_x^{2x} \frac{1}{\sqrt{1+x^4}}\mathrm{d}t = \frac{2x-x}{\sqrt{1+x^4}} \leqslant \frac{x}{\sqrt{x^4}} = \frac{1}{x}$$

Et par encadrement : $\lim\limits_{x\to+\infty} H(x) = 0$.

4. Equation différentielle (dérivation/primitivation tordue)

4.1. Vocabulaire

D'une manière générale on appelle équation différentielle une équation faisant intervenir les dérivées successives d'une même fonction, elle est du premier ordre si elle porte sur la fonction et sa dérivée première, du second ordre si elle porte sur la fonction et ses dérivées première et seconde...
La résolution d'un problème de Cauchy est la résolution d'une équation différentielle avec des conditions initiales.
Plus précisément :

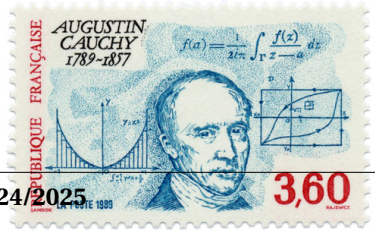
Définition - Equation différentielle linéaire du premier ordre
Une équation différentielle (E) est dite *linéaire et du premier ordre* si elle s'écrit $\alpha(t)y' + \beta(t)y = \gamma(t)$ où α, β, γ sont trois fonctions définies sur un intervalle I de $\mathbb{R}$ (à valeurs dans $\mathbb{R}$ ou $\mathbb{C}$).
Elle est dite *normalisée* si elle s'écrit $y' + a(t)y = b(t)$ où a, b sont deux fonctions définies sur un intervalle I de $\mathbb{R}$ (à valeurs dans $\mathbb{R}$ ou $\mathbb{C}$).
$b(t)$ (ou $\gamma(t)$) est le *second membre*, l'équation est dite *sans second membre ou homogène* si la fonction b est nulle.

Remarque - Mise sous forme normale
Si α ne s'annule pas sur I l'équation différentielle $\alpha(t)y' + \beta(t)y = \gamma(t)$ se ramène à une équation normalisée.

Définition - Problème de Cauchy du premier ordre
On appelle problème de Cauchy du premier ordre la donnée d'une équation différentielle du premier ordre et d'une condition initiale $y(t_0) = y_0$ où $t_0 \in \mathbb{R}$ et $y_0 \in \mathbb{C}$ (ou $\mathbb{R}$).

Histoire - Révolution newtonienne
La double découverte, par Newton, des lois de physique qui s'expriment sous forme d'équations différentielles et des méthodes mathématiques pour les résoudre a permis la révolution scientifique en Europe. Pendant deux siècles, les techniques se sont affinées et la maîtrise de tous les phénomènes de sciences physiques s'est élargies. Au milieu du XIX-ième siècle, les scientifiques croyaient au déterminisme totalement compris (l'avenir totalement écrit sous nos yeux)...

Histoire - Louis-Augustin Cauchy



Louis-Augustin Cauchy (1789-1857) est un mathématicien français aux intérêts très large en mathématiques. Son attitude face aux événements politiques contemporains l'a beaucoup

Définition - Solutions
Soit f une fonction de I dans $\mathbb{C}$.
f est solution de (E) $\quad \alpha(t)y' + \beta(t)y = \gamma(t)$ si

(1) f est dérivable sur I,
(2) $\forall t \in I, \quad \alpha(t)f'(t) + \beta(t)f(t) = \gamma(t)$.

0n pourra noter S_E l'ensemble des solutions de (E).
Résoudre l'équation différentielle (E) c'est donc déterminer l'ensemble S_E, c'est-à-dire trouver toutes les solutions sur I.
On appelle *courbe intégrale* de (E) la courbe représentative d'une solution de (E).

Remarque - Convention réelle
En l'absence d'indications contraires, si les fonctions α, β, γ sont à valeurs dans $\mathbb{R}$, on notera S_E l'ensemble des fonctions de I dans $\mathbb{R}$ qui sont solutions de (E).

Définition - Solution d'un problème de Cauchy
Résoudre le problème de Cauchy défini par (E) et $y(t_0) = y_0$, c'est déterminer toutes les solutions f de (E) vérifiant $f(t_0) = y_0$.

Remarque - Une/la solution ?
On parle souvent de solution particulière de (E), il s'agit en fait d'UNE fonction qui est solution, par opposition à solution générale qui est la forme générale des solutions.
Par exemple, $t \mapsto e^{3t}$, $t \mapsto 4e^{3t}$ sont des solutions particulières de $y' = 3y$ alors que $t \mapsto \lambda e^{3t}$ est la solution générale.

Savoir faire - Découper I pour avoir des équations normalisées
Si α s'annule sur I, on cherchera à découper I en plusieurs intervalles ouverts sur lesquels elle ne s'annule pas pour se ramener à des équations normalisées.
Ensuite on cherchera les solutions sur I par « recollement », c'est-à-dire que l'on regardera, parmi les fonctions définies par morceaux sur chacun des intervalles, celles qui sont dérivables sur I (problème aux points de recollement, c'est-à-dire ceux où α s'annulait) et vérifient (E) sur I.

Le principe suivant est d'usage fréquent en physique : il permet de s'intéresser à des seconds membres simples.

Proposition - Principe de superposition des solutions
Si le second membre de l'équation (E) est de la forme $b(t) = b_1(t) + \cdots + b_n(t)$
et si l'on connaît des solutions particulières $\widetilde{y}_1, \ldots, \widetilde{y}_n$ des équations avec les seconds membres $b_1(t), \ldots, b_n(t)$,
alors une solution particulière de (E) est $\widetilde{y} = \widetilde{y}_1 + \cdots + \widetilde{y}_n$.

Démonstration
On exploite la linéarité de l'équation différentielle linéaire.
On a

$$\widetilde{y}' + a\widetilde{y} = (\widetilde{y}_1 + \cdots + \widetilde{y}_n)' + a(\widetilde{y}_1 + \cdots + \widetilde{y}_n)$$
$$= (\widetilde{y}_1' + \widetilde{y}_1) + \cdots + (\widetilde{y}_n' + a\widetilde{y}_n) = b_1 + \cdots + b_n = b$$

□

4.2. Equation différentielle linéaire d'ordre 1

Principe

On considère désormais l'équation différentielle linéaire normalisée

$$(E) \qquad y' + a(t)y = b(t)$$

où a et b sont continues sur I, intervalle de $\mathbb{R}$.
Dans la suite $\mathbb{K}$ désigne $\mathbb{R}$ ou $\mathbb{C}$.

Heuristique - Démonstration et savoir-faire. Que retenir ?

Pour les démonstrations, nous allons décomposer l'application $F_a : \mathscr{D}(\mathbb{R}) \to \mathscr{D}(\mathbb{R})$, $y \mapsto y' + ay$ en applications, plus ou moins inversible. Nous verrons alors que l'équation différentielle est une dérivation « tordue ».
A la fin du cours, nous donnerons une méthode qu'on appliquera directement lors des exercices.
Sauf pour les exercices théoriques (du type inégalités différentielles).

Analyse

Analyse - Décomposition de F_a

Notons $D : y \mapsto y'$ et pour une application h quelconque $\varphi_h : y \mapsto \exp(h) \times y$.
Alors :

1. $\varphi_{-h} \circ \varphi_h : y \mapsto \exp(-h) \times (\exp(h) \times y) = \exp(-h+h) \times y = y = \mathrm{id}(y)$.
   Ainsi $\varphi_{-h} \circ \varphi_h = \mathrm{id} = \varphi_h \circ \varphi_{-h}$. Donc φ_h est inversible et $\varphi_h^{-1} = \varphi_{-h}$.
2. $\varphi_{-h} \circ D \circ \varphi_h : y \mapsto \exp(-h)[\exp(h)y]' = \exp(-h)[h'y + y']\exp(h) = y' + h'y = F_{h'}(y)$.
   Ainsi : $\varphi_{-h} \circ D \circ \varphi_h = F_{h'}$.

Ainsi, si $a \leftarrow h'$ donc $A \leftarrow h$ UNE primitive de a :

$$F_a(y) = b \iff \varphi_{-A} \circ D \circ \varphi_A(y) = b \iff D \circ \varphi_A(y) = \varphi_A(b)$$

Le problème : D n'est pas bijective… Il faudra gérer des constantes…
On continuera cette résolution plus loin.

Remarque - Non commutativité de D et φ_A

Malheureusement, on n'a pas $F_a = \varphi_{-A} \circ D \circ \varphi_A = D \circ \varphi_{-A} \circ \varphi_A = D$.
C'est là où on voit que l'équation différentielle est une dérivation tordue…

Exercice

Résoudre (E) $\quad (1+t^2)y' + 4ty = 0$ sur $I = \mathbb{R}$.

Correction

Sous forme normalisée (et homogène) : $y' + \dfrac{4t}{1+t^2}y = 0$, l'équation est définie sur $\mathbb{R}$ $(1+t^2 \neq 0)$.
On note alors $a(t) = \dfrac{4t}{1+t^2}$, de primitive $A(t) = 2\ln(1+t^2)$.
On a donc $D \circ \varphi_A(y) = \exp(A(t)) \times 0 = 0$, donc il existe $C \in \mathbb{R}$ tel que $y = \varphi_{-A}(C) = C\exp(-A)$.
Les solutions de (E) sont donc $t \mapsto Ce^{-2\ln(1+t^2)} = \frac{C}{(1+t^2)^2}$.

Remarque - Comment retenir l'ensemble des solutions ?

On dispose d'une méthode intuitive (mais qui n'est pas une démonstration car elle suppose que l'on sait que les solutions y ne s'annulent en aucun point) pour retrouver le résultat dans le cas des fonctions à valeurs réelles (c'est-à-dire lorsque $\mathbb{K} = \mathbb{R}$) en écrivant $\dfrac{y'(t)}{y(t)} = -a(t)$ et en primitivant.

$$\frac{y'(t)}{y(t)} = -a(t) \quad \Longrightarrow \ln(y(t)) = -A(t) + K$$
$$\Longrightarrow y(t) = \exp(-A(t)+K) = Ce^{-A(t)} \text{ avec } C = e^K$$

On écrirait plutôt, heuristiquement :

$$(ye^{\int a})' = [y' + ay]e^{\int a} = 0 \Longrightarrow ye^{\int a} = K \Longrightarrow y = Ke^{-\int a}$$

Pour aller plus loin - Simplification de problème

On rencontre toujours en mathématique le diagramme suivant où un problème M (une matrice par exemple) est transformé en un problème D (matrice diagonale par exemple) plus simple à résoudre.

$$\begin{array}{ccc} x & \xrightarrow{M} & y \\ P\downarrow & & \downarrow P \\ \cdot & \xrightarrow{D} & \cdot \end{array}$$

Il y a deux chemins : celui par M, et celui par P, D, puis P (à l'envers). La composition s'écrit de droite à gauche. Cela donne

$$M = P^{-1} \times D \times P$$

Attention - Variable x, variable t ?
Nous avons noté y la fonction de la variable t, mais l'on peut bien évidemment avoir d'autres notations, par exemple y fonction de la variable x (équations différentielles donnant l'ordonnée en fonction de l'abscisse) ou x en fonction de t (équations différentielles donnant l'abscisse en fonction du temps) ou encore z en fonction de t (équations différentielles donnant l'affixe en fonction du temps).

Structures des solutions et méthodes

Théorème - Structure de l'ensemble S_E
La solution générale de l'équation (E) est la somme d'une solution particulière et de la solution générale de l'équation homogène associée (H), ce qui peut aussi s'écrire :
Si $\tilde{y}$ (à lire « y tilde ») est une solution particulière de l'équation (E) alors

$$S_E = \left\{ t \mapsto Ce^{-A(t)} + \tilde{y}(t);\ C \in \mathbb{K} \right\}.$$

Remarque - Existence
Insistons : L'existence d'(au moins) une solution est donnée par la forme de l'équation (E) :
- normalisée,
- avec a continue (admettant une primitive - plus exactement)
- avec b continue ($t \mapsto b(t)e^{A(t)}$ admettant une primitive - plus exactement)

Démonstration
Soit $\tilde{y}$ une solution particulière. On a donc $\varphi_{-A} \circ D \circ \varphi_A(\tilde{y}) = b$.
Donc l'équation devient :

$$D \circ \varphi_A(y) = D \circ \varphi_A(\tilde{y}) \iff \exists\, C \in \mathbb{R} \text{ tel que } \varphi_A(y) = C + \varphi_A(\tilde{y})$$

En exploitant l'inverse de φ_A, on trouve S_E (et il y a équivalence donc égalité d'ensembles). □

Pour aller plus loin - Equation à variables séparables
La méthode présentée dans le remarque (comment retenir l'ensemble des solutions ?)est celle proposée dans le cas des équations à variables séparables : $y' \times f(y) = g(t)$. Avec un changement de variables, il s'agit d'un simple calcul d'intégrale (présenté ainsi, cela fonctionne mieux)...
En 1840, le mathématicien belge Pierre Verhulst propose un modèle de dynamique de population qui conduit à l'équation

$$y' = ay\left(1 - \frac{y}{K}\right)$$

où a et K sont des constantes associées à la population considérée.
La résolution donne :

$$y(t) = K\frac{1}{1 + \left(\frac{K}{y_0} - 1\right)e^{-at}}$$

On trouve en particulier : $y(0) = y_0$ et $y \underset{t\to+\infty}{\longrightarrow} K$. Ce qui est contraire au modèle de Maltus

Trouver une solution particulière $\tilde{y}$

Analyse - Résoudre l'équation
Continuons notre résolution. Nous en étions à $\varphi_{-A} \circ D \circ \varphi_A = b$.
Donc en remplaçant par leur valeur : $D \circ \varphi_A(y) = be^A$, qu'il faut intégrer à une constante près.
Notons $e^{A(t)}y(t) = \int_{\cdot}^{t} b(u)e^{A(u)}\mathrm{d}u + K$, donc

$$y(t) = \underbrace{Ke^{-A(t)}}_{\text{solution homogène}} + \underbrace{e^{-A(t)}\int_{\cdot}^{t} b(u)e^{A(u)}\mathrm{d}u}_{\text{solution particulière}}$$

.
Si $b = 0$, on a bien une solution homogène avec le premier terme de cette somme-solution.
L'enjeu est donc maintenant de trouver une solution particulière, peut-être plus simplement...
On doit à Lagrange la méthode de la variation de la constante qui répond explicitement à cette question (parmi d'autres méthodes).

Savoir faire - Comment trouver une équation particulière ? Méthode de « variation de la constante »
D'après le théorème précédent, ce qui reste est de trouver une solution particulière.
Sans indication donnée par l'énoncé, la méthode classique à suivre est la suivante :

1. On normalise l'équation différentielle. Cela peut nécessiter une étude sur plusieurs intervalles.
2. On résout l'équation différentielle homogène : $y = Ce^{-A(t)}$
3. On cherche une solution particulière :
   — en cherchant une solution évidente,
   — en utilisant le principe de superposition des solutions,
   — en essayant des fonctions simples (polynomiales lorsque a et b le sont, trigonométriques lorsque a et b le sont...),
   — en faisant varier la constante C, c'est-à-dire sous la forme

   $$\widetilde{y} : t \mapsto C(t)e^{-A(t)}$$

   (La constante C devient variable : « méthode de la variation de la constante »).
   Le calcul (à refaire à chaque fois - il permet de vérifier la bonne résolution de l'équation homogène) conduit à :

   $$C'(t) = b(t)e^{A(t)}$$

   C'est un « simple » calcul de primitive
4. Les solutions générales sont alors de la forme

   $$y : t \mapsto (K + C(t))e^{-A(t)}$$

   avec C définie au point précédent

Problème de Cauchy

Théorème - Problème de Cauchy
Soit $t_0 \in I$ et $y_0 \in \mathbb{K}$.
Il existe une unique solution sur I de l'équation linéaire normalisée (E) $y' + a(t)y = b(t)$ vérifiant la condition initiale $y(t_0) = y_0$.

Démonstration
Les solutions de (E) sont de la forme $y = \widetilde{y} + Ce^{-A(t)}$ avec C constante « à déterminer ».
On a alors $y(t_0) = y_0 = \widetilde{y}(t_0) + Ce^{-A(t_0)}$, donc $C = (y_0 - \widetilde{y}(t_0))e^{A(t_0)}$.
Ce calcul donne une et une seule solution car $e^{-A(t_0)} > 0$ □

Remarque - Résoudre un problème de Cauchy
Pour résoudre un problème de Cauchy, on résout (E) puis on cherche la solution vérifiant $y(t_0) = y_0$.

Pour aller plus loin - Courbe intégrale
Cela signifie également qu'il existe une unique courbe intégrale passant par le point (t_0, y_0).

Applications. Cas classiques

Exercice
Résoudre (E) $y' + ty = t$.

Correction

1. L'équation est sous forme normalisée.
2. L'équation homogène est $y' + ty = 0$ de solution $t \mapsto Ce^{-\frac{1}{2}t^2}$.
3. Une solution particulière de E est $y = 1$

L'ensemble des solutions est donc $\left\{t \mapsto 1 + Ce^{-\frac{1}{2}t^2}, C \in \mathbb{R}\right\}$.

Exercice
Résoudre (E) $z' = (1+i)z - 2it^2 + 2$.

Correction

1. L'équation est sous forme normalisée. A noter qu'ici $z : \mathbb{R} \mapsto \mathbb{C}$
2. L'équation homogène est $z' - (1+i)z = 0$ de solution $t \mapsto Ce^{(1+i)t} = Ce^t e^{it}$.
3. Le second membre est une fonction polynômiale de degré 2 : $t \mapsto -2it^2 + 2$.
   On cherche une solution particulière de la même forme : $\widetilde{z}(t) = At^2 + Bt + C$
   $$\widetilde{z}'(t) - (1+i)\widetilde{z}(t) = -A(1+i)t^2 + (2A - B(1+i))t + (B - C(1+i)) = -2it^2 + 2$$
   En identifiant : $A = \frac{2i}{1+i} = 1+i$, puis $B = 2$ et $C = 0$, une solution particulière : $\widetilde{z} : t \mapsto (1+i)t^2 + 2t$
   (On peut/doit vérifier)

L'ensemble des solutions est donc $\left\{t \mapsto 2t + (1+i)t^2 + Ce^{(1+i)t}, C \in \mathbb{C}\right\}$.

⚠ Attention - Ensemble des solutions particulières selon le second membre

Attention, si le second membre est colinéaire à la solution homogène, il faut alors chercher une solution particulière de « degré » plus élevé…

Exercice

Résoudre (E) $y' + y = 2e^x + 4\sin x + 3\cos x$.

Correction

1. L'équation est sous forme normalisée.
2. L'équation homogène est $y' + y = 0$ de solution $t \mapsto Ce^{-x}$.
3. On applique la méthode de la variation de la constante pour trouver une solution particulière $\widetilde{y}(x) = C(x)e^{-x}$.
   $$\widetilde{y}'(x) + \widetilde{y}(x) = C'(x)e^{-x} = 2e^x + 4\sin x + 3\cos x$$
   Donc $C'(x) = 2e^{2x} + 4\sin x e^x + 3\cos x e^x$, fonction à primitiver. On commence par primitiver $(\cos x + i\sin x)e^x = e^{(1+i)x}$ en $\frac{1}{1+i}e^{(1+i)x} = \frac{e^x}{2}(1-i)(\cos x + i\sin x)$.
   En prenant les parties réelles et imaginaires :
   $$\int \cos x e^x \mathrm{d}x = \frac{e^x}{2}(\cos x + \sin x) \qquad \int \sin x e^x \mathrm{d}x = \frac{e^x}{2}(-\cos x + \sin x)$$
   On a donc $C(x) = e^{2x} + 2e^x(-\cos x + \sin x) + \frac{3}{2}e^x(\cos x + \sin x) = e^{2x} + e^x(\frac{7}{2}\sin x - \frac{1}{2}\cos x)$.

L'ensemble des solutions est donc $\left\{x \mapsto e^x + \frac{1}{2}(7\sin x - \cos x) + Ce^{-x}, C \in \mathbb{R}\right\}$.

Savoir faire - Etudier une inéquation différentielle

Supposons qu'on ait l'inéquation $y' + ay \leqslant b$.
On reprend les notations précédentes, avec $\varphi_h : y \mapsto (t \mapsto e^{h(t)} \times y(t))$.
On a $\varphi_{-A} \circ D \circ \varphi_A \circ y \leqslant b$ ou encore $\forall\, u \in I : (D \circ \varphi_A)(y)(u) \leqslant e^{A(u)}b(u)$ car $e^z > 0$.
Puis par croissance de l'intégration pour $t \geqslant t_0$:
$$\varphi_A(y)(t) - \varphi_A(y)(t_0) \leqslant \int_{t_0}^{t} b(u)e^{A(u)}\mathrm{d}u.$$
Et donc $y(t) \leqslant e^{A(t_0)-A(t)}y(t_0) + e^{-A(t)}\int_{t_0}^{t} b(u)e^{A(u)}\mathrm{d}u$.
(La constante t_0 est arbitraire, il faut nécessairement en fixer une!).
On reconnait une solution de l'équation différentielle pour le cas frontière.

Pour aller plus loin - Inéquations

Souvent les inéquations se transforment en de nouvelles inégalités, où les cas frontières sont exactement obtenus avec les solutions de l'équation identique (⩽ transformé en =).

Cas d'une équation non résolue. Problème du recollement

Savoir faire - Cas non résolue

A résoudre une équation $a(t)y' + b(t)y = c(t)$ sur I avec a qui s'annule sur I.

1. On étudie l'équation sur des sous-intervalles de I où a ne s'annule pas.
   Par exemple si $a(t) = t$ et $I = \mathbb{R}$; on étudie sur $\mathbb{R}_+^*$ et sur $\mathbb{R}_-^*$.
2. On obtient une famille de solutions, paramétrée sur chacun des sous-intervalles par une variable α (par exemple).

3. On essaye de « **recoller** » les solutions. Pour cela, on regarde les limites de y et de y' au voisinage du point t_0 qui annule a de manière à étudier la continuité et la dérivabilité de y en t_0.
   Souvent ces limites dépendent de la valeur paramètre α.

Dans la suite du cours : le théorème de prolongement de classe $\mathscr{C}^1$ et les méthodes de calcul asymptotiques seront très utiles pour résoudre ce problème

Exercice

Résoudre l'équation $t^2y' + y = 1$ sur un intervalle I de $\mathbb{R}$ (on discutera suivant la position de 0 par rapport à I).

Correction

1. L'équation n'est pas sous forme normalisée. Il faut diviser par t^2, cela n'est possible que sur $\mathbb{R}_+^*$ et sur $\mathbb{R}_-^*$.
   L'équation normalisée est alors $y' + \frac{1}{t^2}y = \frac{1}{t^2}$.
   On considère donc, pour la suite, I un intervalle de $\mathbb{R}_+^*$ ou de $\mathbb{R}_-^*$.
2. L'équation homogène normalisée est $y' + \frac{1}{t^2}y = 0$ de solution $t \mapsto Ce^{\frac{1}{t}}$.
3. Une solution particulière de E est $y = 1$
4. L'ensemble des solutions est donc sur $\mathbb{R}_+^*$: $\left\{t \mapsto 1 + Ce^{\frac{1}{t}}, C \in \mathbb{R}\right\}$ et $\mathbb{R}_-^*$: $\left\{t \mapsto 1 + C'e^{\frac{1}{t}}, C' \in \mathbb{R}\right\}$

On a trouvé une solution sur tout intervalle de $\mathbb{R}_+^*$, sur tout intervalle de $\mathbb{R}_-^*$.
Est-il possible d'avoir une solution sur un intervalle de $\mathbb{R}$ (contenant 0) ? Pour cela on pratique un recollement :

Soit y une telle solution, alors il existe $C, C' \in \mathbb{R}$ tel que : $\forall\, t > 0$, $y(t) = 1 + Ce^{\frac{1}{t}}$ et $\forall\, t < 0$, $y(t) = 1 + C'e^{\frac{1}{t}}$.

Cette fonction y est nécessairement de classe $\mathscr{C}^1$ sur $I \subset \mathbb{R}$.

Elle est donc continue en 0. Or $\lim\limits_{t\to 0^+} y(t) = \begin{cases} 1 & \text{si } C = 0 \\ +\infty & \text{si } C > 0 \\ -\infty & \text{si } C < 0 \end{cases}$.

Il est donc nécessaire que $C = 0$ et dans ce cas $y(t) = 1$, pour tout $t \geqslant 0$.

De même $\lim\limits_{t\to 0^-} y(t) = 1$ (pour tout C'). Donc y est bien continue en 0

Il faut aussi que y soit dérivable sur I, donc dérivable en 0. Or $y'(t) = 0$, pour tout $t > 0$,

alors que pour tout $t < 0$, $y'(t) = \frac{-1}{t^2}Ce^{\frac{1}{t}} \underset{t\to 0^-}{\longrightarrow} 0$ (comment lever l'indétermination ?).

Bilan : l'ensemble des solutions y (fonctions de classe $\mathscr{C}^1$) solutions de (E) sur I contenant 0 est

$$\left\{t \mapsto \begin{cases} 1 & \text{si } t \geqslant 0 \\ 1 + Ce^{\frac{1}{t}} & \text{si } t < 0 \end{cases}, C \in \mathbb{R}\right\}$$

Notez bien la méthode de recollement!

Remarque - Règle de L'Hospital

Il n'est pas rare que pour faire le recollement, nous ayons besoin de la règle de L'Hospital pour lever l'indétermination que l'on obtient en calculant, pour s frontière de I $\lim_{t\to s} y'(t) = \lim_{t\to s} \dfrac{c(t) - b(t)y(t)}{a(t)}$.
A utiliser, sans modération!

4.3. Equation différentielle linéaire d'ordre 2 à coefficients constants

Enoncé

$\mathbb{K}$ désigne toujours $\mathbb{R}$ ou $\mathbb{C}$. Insistons : ici a, b, c **sont constants**, $a \neq 0$ (sinon on revient au cas précédent).
L'année prochaine, vous élargirez ce point de vue.

Définition - Equations différentielles linéaires du second ordre à coefficients constants
Soient $(a, b, c) \in \mathbb{K}^3$, $a \neq 0$ et $u : I \to \mathbb{K}$ une fonction continue sur l'intervalle I de $\mathbb{R}$.
On dit qu'une fonction $f : I \to \mathbb{K}$ est solution de l'équation différentielle du

second ordre à coefficients constants

$$(E) \quad ay'' + by' + cy = u(t)$$

si

(1) f est deux fois dérivable sur I

(2) $\forall t \in I, af''(t) + bf'(t) + cf(t) = u(t)$

L'équation $ar^2 + br + c = 0$ est appelée *équation caractéristique associée.*

Résolution de l'équation homogène associée

On considère donc l'équation homogène $(H) \quad ay'' + by' + cy = 0 \quad a \neq 0$.

Analyse - Composition de $F_{-\alpha} \circ F_{-\beta}$

Conservons les mêmes notations et comme a, b et c sont constants, prenons des primitives affines.

On a donc $(F_{-\alpha} \circ F_{-\beta})(y) = (y' - \beta y)' - \alpha(y' - \beta) = y'' - (\alpha + \beta)y' + \alpha\beta y$.

Et donc, en reprenant les formules de Viète : α et β sont les racines $x^2 - (\alpha + \beta)x + \alpha\beta$.

Considérons donc les racines α et β du polynôme $ax^2 + bx + c = a(x - \alpha)(x - \beta)$.

Alors, $aF_{-\alpha} \circ F_{-\beta}(y) = ay'' + by' + cy$.

Cette analyse ne marche pas lorsque les racines sont les mêmes : $\alpha = \beta$.

On étudie les cas particuliers selon la nature des solutions (double ou simples, complexes ou réelles) de l'équation caractéristique associée à l'équation homogène.

Théorème - Cas complexe

— Si $ar^2 + br + c = 0$ possède deux racines distinctes r_1 et r_2 alors l'ensemble des solutions à valeurs dans $\mathbb{C}$ est :

$$S_H = \left\{ t \mapsto \lambda e^{r_1 t} + \mu e^{r_2 t}; (\lambda, \mu) \in \mathbb{C}^2 \right\}$$

— Si $ar^2 + br + c = 0$ possède une racine double $r_0 \in \mathbb{C}$ alors l'ensemble des solutions à valeurs dans $\mathbb{C}$ est :

$$S_H = \left\{ t \mapsto (\lambda + \mu t) e^{r_0 t}; (\lambda, \mu) \in \mathbb{C}^2 \right\}$$

Démonstration

Notons r_1 et r_2 les racines de $ax^2 + bx + c = a(x - r_1)(x - r_2)$. (On peut avoir $r_1 = r_2$).

Alors y solution de $ay'' + by' + cy = 0 \iff F_{-r_1 \cdot t} \circ F_{-r_2 \cdot t}(y) = 0$.

Si l'on compose (comme précédemment), on a :

$$\begin{aligned} ay'' + by' + cy = 0 \quad &\iff \varphi_{r_1 \cdot t} \circ D \circ \varphi_{-r_1 \cdot t} \circ \varphi_{r_2 \cdot t} \circ D \circ \varphi_{-r_2 \cdot t}(y) = 0 \\ &\iff \exists\, C_1 \in \mathbb{C} \text{ tel que } \circ D \circ \varphi_{-r_2 \cdot t}(y) = C_1 \exp((r_1 - r_2)t) \\ &\iff \exists\, C_1 \in \mathbb{C} \text{ tel que } \circ D \circ \varphi_{-r_2 \cdot t}(y) = C_1 \exp((r_1 - r_2)t) \end{aligned}$$

— Si $r_1 = r_2$, $ay'' + by' + cy = 0 \iff \exists\, C_1, C_2 \in \mathbb{C}$ tels que $y = (C_1 t + C_2)\exp(r_2 t)$.

— Si $r_1 \neq r_2$, $ay'' + by' + cy = 0 \iff \exists\, C_1, C_2 \in \mathbb{C}$ tels que $y = \dfrac{C_1}{r_1 - r_2} \exp(r_1 \cdot t) + C_2 \exp(r_2 \cdot t)$.

□

Pour le cas réel, la partie correspondant aux racines complexes ($\Delta < 0$) est plus subtile.

Théorème - Cas réel

On suppose ici a, b, c réels, $a \neq 0$.

— Si $ar^2 + br + c = 0$ possède deux racines réelles distinctes r_1 et r_2 alors l'ensemble des solutions à valeurs dans $\mathbb{R}$ est :

$$S_H = \left\{ t \mapsto \lambda e^{r_1 t} + \mu e^{r_2 t}; (\lambda, \mu) \in \mathbb{R}^2 \right\}$$

— Si $ar^2+br+c=0$ possède une racine double $r_0 \in \mathbb{R}$ alors l'ensemble des solutions à valeurs dans $\mathbb{R}$ est :

$$S_H = \left\{ t \mapsto (\lambda + \mu t) e^{r_0 t} ; (\lambda, \mu) \in \mathbb{R}^2 \right\}$$

— Si $ar^2+br+c=0$ possède deux racines complexes conjuguées $\alpha+i\beta$ et $\alpha-i\beta$ alors l'ensemble des solutions à valeurs dans $\mathbb{R}$ est :

$$S_H = \left\{ t \mapsto \lambda e^{\alpha t} \cos \beta t + \mu e^{\alpha t} \sin \beta t ; (\lambda, \mu) \in \mathbb{R}^2 \right\}$$

Pour la démonstration, commençons par un lemme

Lemme - Solution réelle d'une équation réelle
Soit (H) l'équation différentielle homogène $ay''+by'+cy=0$ où $(a,b,c) \in \mathbb{R}^3, a \neq 0$.
Si f est solution de (H) à valeurs dans $\mathbb{C}$ alors $\mathrm{Re} f$ est solution de (H) à valeurs réelles.
Plus précisément l'ensemble des solutions réelles de (H) est exactement l'ensemble des parties réelles des solutions complexes de (H).

Pour aller plus loin - Paramètres
Considérons l'équation aux paramètres physiques : $y''+\frac{\omega}{Q}y'+\omega_0^2 y=0$.
On a $\Delta = \left(\frac{\omega_0}{Q}\right)^2 (1-4Q^2)$.

— *si $\Delta<0$ ($Q>\frac{1}{2}$) : régime périodique, les solutions sont de la forme $e^{-\frac{\omega_0}{2Q}t}(A\cos(\omega t+\varphi)$ avec $\omega = \frac{\omega_0}{Q}\sqrt{Q^2-\frac{1}{4}}$. Comme $\omega_0, Q > 0$, les solutions sont oscillantes, bornées et de fréquence $\omega = \frac{\omega_0}{Q}\sqrt{Q^2-\frac{1}{4}}$*

— *si $\Delta=0$ ($Q=\frac{1}{2}$) : régime critique, les solutions sont de la forme $e^{\omega_0 t}(At+B)$*

— *si $\Delta>0$ ($Q<\frac{1}{2}$) : régime a-périodique, les solutions sont de la forme $Ae^{-\frac{\omega_0}{Q}(1+\sqrt{\frac{1}{4}-Q^2})t} + Be^{-\frac{\omega_0}{Q}(1-\sqrt{\frac{1}{4}-Q^2})t}$*

Démonstration
Démonstration du lemme :
On rappelle que $a,b,c \in \mathbb{R}$.Supposons que $f=f_R+if_I$ est solution de (H).
Alors

$$0 = af''+bf'+cf = \left(af_R''+bf_R'+cf_R\right) + i\left(af_I''+bf_I'+cf_I\right)$$

Donc $af_R''+bf_R'+cf_R=0$ et f_R est à valeurs réelles.
On a donc : toute partie réelle de solutions complexes de (H) est une solution réelle de (H).
Réciproquement, si f est une solution réelle de (H),
alors $f+i0$ est une solution complexes de (H) dont f est la partie réelle.
Démonstration du théorème :
Les deux premiers cas se déduisent du cas complexe.
Supposons donc que le discriminant de l'équation caractéristique $\Delta<0$.
L'équation caractéristique est à coefficients réels donc les racines sont conjuguées : $\alpha+i\beta$ et $\alpha-i\beta$.
Les solutions sur $\mathbb{C}$ sont de la forme

$$f : t \mapsto \lambda e^{(\alpha+i\beta)t} + \mu e^{(\alpha-i\beta)t}$$

La partie réelle est alors une solution de (H) sur $\mathbb{R}$:

$$t \mapsto \frac{1}{2}(f+\overline{f})(t) = \frac{1}{2}(\lambda+\overline{\mu})e^{\alpha t+i\beta t} + \frac{1}{2}(\overline{\lambda}+\mu)e^{\alpha t-i\beta t}$$

$$t \mapsto e^{\alpha t}\Big(\mathrm{Re}(\lambda+\mu)\cos(\beta t) + \mathrm{Im}(\mu-\lambda)\sin(\beta t)\Big)$$

□

Remarque - Autre expression pour le cas $\Delta<0$
Dans le cas des racines complexes conjuguées, on peut poser $\lambda+i\mu = Ae^{-i\phi}$, où $A \geqslant 0$ et $\phi \in \mathbb{R}$. Les solutions s'écrivent alors $t \mapsto Ae^{\alpha t}\cos(\beta t+\phi)$
C'est un cas souvent préféré en physique.

Remarque - Base d'une espace vectoriel
Dans tous les cas (réel ou complexe) $S_H = \left\{\lambda f_1+\mu f_2 ; (\lambda,\mu) \in \mathbb{K}^2\right\}$ où f_1 et f_2 sont deux fonctions déterminées par l'équation caractéristique associée. On dira que la famille (f_1,f_2) est une *base* de S_H.
On en déduira que S_H est un espace vectoriel de dimension 2 (=nombre de vecteurs de la base).

Exercice
Résoudre les équations différentielles suivantes (on cherchera les solutions réelles).

1. $y''=\omega^2 y$
2. $y''=-\omega^2 y$
3. $y''-4y'+13y=0$
4. $y''-4y'+4y=0$

Correction

1. L'équation caractéristique est $x^2-\omega^2=0$, de racine $\pm\omega$.
   Les solutions de l'équation différentielle sont donc de la forme $t\mapsto Ae^{\omega t}+A'e^{-\omega t}$, avec $A,A'\in\mathbb{R}$
2. L'équation caractéristique est $x^2+\omega^2=0$, de racine $0\pm i\omega$.
   Les solutions de l'équation différentielle sont donc de la forme $t\mapsto A\cos(\omega t)+A'\sin(\omega t)$, avec $A,A'\in\mathbb{R}$
3. L'équation caractéristique est $x^2-4x+13=0$, de racine $2\pm 3i$.
   Les solutions de l'équation différentielle sont donc de la forme $t\mapsto e^{2t}(A\cos(3t)+A'\sin(3t))$, avec $A,A'\in\mathbb{R}$
4. L'équation caractéristique est $x^2-4x+4=0$, de racine double 2.
   Les solutions de l'équation différentielle sont donc de la forme $t\mapsto (A+A't)e^{2t}s$, avec $A,A'\in\mathbb{R}$

Résolution avec second membre

On considère l'équation complète $(E)\quad ay''+by'+cy=u(t)$ avec $(a,b,c)\in\mathbb{K}^3, a\neq 0$.

Théorème - Structure de l'ensemble S_E
La solution générale de l'équation (E) est la somme d'une solution particulière et de la solution générale de l'équation homogène associée (H), ce qui peut aussi s'écrire :
Si $\tilde{y}$ est une solution particulière de l'équation (E) et (f_1,f_2) une base de S_H alors

$$S_E=\left\{t\mapsto \lambda f_1(t)+\mu f_2(t)+\tilde{y}(t);(\lambda,\mu)\in\mathbb{K}^2\right\}.$$

Démonstration
On démontre que $y-\tilde{y}$ est solution de (H) et donc de la forme $\lambda f_1(t)+\mu f_2(t)$ □

Proposition - Principe de superposition des solutions
Si le second membre de l'équation (E) est de la forme $u(t)=u_1(t)+\cdots+u_n(t)$
et si l'on connaît des solutions particulières $\tilde{y}_1,\dots,\tilde{y}_n$ des équations avec les seconds membres $u_1(t),\dots,u_n(t)$,
alors une solution particulière de (E) est $\tilde{y}=\tilde{y}_1+\cdots+\tilde{y}_n$.

Remarque - Démonstration
Il s'agit exactement de la même démonstration que dans le cas $n=1$.
Ce qui compte ici c'est la linéarité. (Le fait que les coefficients soient constants ne changent rien concernant ce théorème de superposition)

Avec second membre : exponentielle-polynôme

On va s'intéresser au cas où $u(t)=e^{mt}P(t)$ avec $m\in\mathbb{C}$ et P une fonction polynomiale à valeurs complexes.

Savoir faire - Second membre $e^{mt}P(t)$
Soit $m\in\mathbb{C}$ et P une fonction polynomiale de degré n. Alors on peut trouver une solution particulière de l'équation

$$(E)\quad ay''+by'+cy=e^{mt}P(t)$$

de la forme $\tilde{y}(t)=e^{mt}Q(t)$ où Q est une fonction polynomiale
— de degré n si m n'est pas racine de $ar^2+br+c=0$

Pour aller plus loin - Méthode de Laplace
On note $\mathscr{L}:f\mapsto F$, telle que $F:p\mapsto\int_0^{+\infty}f(t)e^{-pt}\mathrm{d}t$ On montre (IPP) alors que $\mathscr{L}(f')(p)=pf(p)-f(0^+)$.
On transforme ainsi une dérivation en un produit et donc l'équation $ay''+by'+cy=g$ se transforme en

$$(ap^2+bp+c)F(p)=G(p)$$

et donc $F(p)=\dfrac{G(p)}{ap^2+bp+c}$, il reste donc à appliquer $\mathscr{L}^{-1}$ à F pour trouver la valeur de f...

— de degré $n+1$ si m est racine simple de $ar^2+br+c=0$
— de degré $n+2$ si m est racine double de $ar^2+br+c=0$

Démonstration
Le calcul donne :

$$a\tilde{y}''(t)+b\tilde{y}'(t)+c\tilde{y}(t)=[(am^2+bm+c)Q(t)+(2am+b)Q'(t)+aQ''(t)]e^{mt}$$

Si $d=\deg(Q)$, alors $\deg(Q')=d-1$ et $\deg(Q'')=d-2$.

$$\tilde{y} \text{ est solution de } (E) \iff \forall\ t\in I[(am^2+bm+c)Q(t)+(2am+b)Q'(t)+aQ''(t)]=P(t)$$

Nécessairement les polynômes doivent être de même degré. Donc :
— si m n'est pas racine de $ar^2+br+c=0$, $\deg(Q)=\deg(P)=n$
— si m est racine simple de $ar^2+br+c=0$, donc $2am+b\neq 0$, alors $\deg(Q')=\deg P$, donc $d-1=n$, i.e. $d=n+1$
— si m est racine double de $ar^2+br+c=0$ donc $2am+b=0$, alors $\deg(Q'')=\deg P$, donc $d-2=n$, i.e. $d=n+2$

La condition d'existence donne l'unicité (analyse). Il faut vérifier l'existence (synthèse). Il s'agit alors d'un système inversible de n équations à n inconnues à résoudre. □

Analyse - Polynôme × fonction trigonométrique

On peut utiliser ce qui précède pour le cas où $(a,b,c)\in\mathbb{R}^3$ et $u(t)=e^{\alpha t}(P(t)\cos\beta t+Q(t)\sin\beta t)$ avec P,Q des fonctions polynomiales à coefficients réels.
Dans ce cas le second membre est de la forme $T(t)e^{(\alpha+i\beta)t}$ où T est un polynôme complexe.
Or ce second membre est a priori réel. Le calcul suivant donne :

$$e^{\alpha t}(P(t)\cos\beta t+Q(t)\sin\beta t)=\mathrm{Re}\Big((P(t)-iQ(t))e^{(\alpha+i\beta)t}\Big)$$

on en déduit que $T=P(t)+iQ(t)$ et on peut chercher une solution particulière (sur $\mathbb{C}$) de la forme $R(t)e^{\alpha+i\beta t}$ avec :
— $\deg R=\deg T$ si $\alpha+i\beta$ n'est pas racine de $ar^2+br+c=0$
— $\deg R=\deg T$ si $\alpha+i\beta$ est racine simple $ar^2+br+c=0$.
(Elle ne peut être racine double, sinon l'équation serait à coefficients complexes)

Si $\tilde{y}$ est une solution à valeurs complexes de l'équation avec le second membre $(P(t)-iQ(t))e^{(\alpha+i\beta)t}$, alors $\mathrm{Re}(\tilde{y})$ est une solution particulière à valeurs réelles.

Savoir faire - Second membre de la forme $e^{\alpha t}(P(t)\cos\beta t+Q(t)\sin\beta t)$

On peut trouver (par identification) une solution particulière de l'équation

$$(E)\quad ay''+by'+cy=e^{\alpha t}(P(t)\cos(\beta t)+Q(t)\sin(\beta t))$$

de la forme $\tilde{y}(t)=e^{\alpha t}(T(t)\cos(\beta t)+R(t)\sin(\beta t))$ où T et R sont des fonctions polynomiales
— de degré $n=\max(\deg(P),\deg(Q))$ si $\alpha+i\beta$ n'est pas racine de $ar^2+br+c=0$
— de degré $n+1=\max(\deg(P),\deg(Q))+1$ si $\alpha+i\beta$ est racine simple de $ar^2+br+c=0$

Exercice
1. Résoudre $y''-y'-2y=3e^{-t}+1$.
2. Résoudre $y''+2y'+5y=\cos^2 t$.

Correction

1. L'équation homogène est (H) $y''-y'-2y=0$ d'équation caractéristique $r^2-r-2=(r-2)(r+1)$.
Les solutions de (H) sont de la forme $t\mapsto Ae^{2t}+A'e^{-t}$.
En exploitant le théorème précédent et le principe de superposition, on cherche une solution particulière sous la forme
$$\tilde{y}:t\mapsto C_1te^{-t}+C_2$$
Or
$$\tilde{y}''(t)-\tilde{y}'(t)-2\tilde{y}(t)=C_1[(-2-1)+(1+1-2)t]e^{-t}+C_2=-3C_1e^{-t}+C_2$$
Donc $C_1=-1$ et $C_2=1$ et les solutions de (E) sont :
$$t\mapsto Ae^{-2t}+(A'-t)e^{-t}+1\qquad \text{avec } A,A'\in\mathbb{R}$$

2. L'équation homogène est (H) $y''+2y'+5y=0$ d'équation caractéristique r^2+2r+5, de discriminant $\Delta=-16$ et donc de racine $-1\pm 2i$.
Les solutions de (H) sont de la forme $t\mapsto e^{-t}(A\cos 2t+A'\sin 2t)$.
En outre $\cos^2 t=\dfrac{1}{2}(\cos 2t+1)$, donc le second membre est (en partie) de la forme $e^{\alpha t}(P(t)\cos\beta t+Q(t)\sin\beta t)$, avec $\alpha=0$, $\beta=2$, $P(t)=\frac{1}{2}$ et $Q(t)=0$.
Pas besoin de monter en degré et on peut donc chercher une solution particulière de la forme

$$\tilde{y}: t\mapsto C_1\cos(2t)+C_2\sin(2t)+C_3$$

Or

$$\tilde{y}''(t)+2\tilde{y}'(t)+5\tilde{y}(t)=(C_1+4C_2)\cos(2t)+(C_2-4C_1)\sin(2t)+5C_3$$

Donc $C_1=\frac{1}{34}$, $C_2=\frac{4}{34}$ et $C_3=\frac{1}{10}$ et les solutions de (E) sont :

$$t\mapsto e^{-t}(A\cos 2t+A'\sin 2t)+\frac{1}{34}(\cos 2t+4\sin 2t)+\frac{1}{10}\qquad \text{avec } A,A'\in\mathbb{R}$$

Résolution du problème de Cauchy

Théorème - Conditions initiales
Soit (E) $\quad ay''+by'+cy=u(t)$ avec $(a,b,c)\in\mathbb{K}^3$, $a\neq 0$ et $u: I\to\mathbb{K}$ de l'une des formes précédentes (exponentielle-polynôme...).
Soit $(t_0,y_0,y_0')\in I\times\mathbb{K}\times\mathbb{K}$.
Alors il existe une unique solution $f: I\to\mathbb{K}$ telle que $f(t_0)=y_0$ et $f'(t_0)=y_0'$.

Le principe de la démonstration est simple : les deux conditions initiales fixent les deux valeurs des deux variables libres λ et μ.
Mais la démonstration est pénible selon la nature des racines de l'équation caractéristique.
Nous nous contenterons du cas $\Delta\neq 0$

Démonstration
Supposons que $\Delta\neq 0$. Soient α, β les deux racines de ax^2+bx+c.
Notons $F_{-\alpha}: y\mapsto y'-\alpha y$ et $F_{-\beta}: y\mapsto y'-\beta y$.
On a les équivalences :

$$y \text{ solution de } E \iff F_{-\alpha}\circ F_{-\beta}(y)=u \iff \exists\, C\in\mathbb{R} \text{ tq} F_{-\beta}(y): t\mapsto Ce^{\alpha(t-t_0)}+e^{\alpha(t-t_0)}\int_{t_0}^{t} v(u)e^{-\alpha u}\mathrm{d}u$$

$$\iff \exists\, C',C''\in\mathbb{R} \text{ tq} y: t\mapsto C''e^{\beta(t-t_0)}+C'e^{\alpha(t-t_0)}+V(t)$$

où V est une fonction explicite, qui dépend de v, α et β :

$$U: t\mapsto e^{\beta t}\int_{t_0}^{t} e^{(\alpha-\beta)s}\left(\int_{t_0}^{s} v(u)e^{-\alpha u}\mathrm{d}u\right)\mathrm{d}s$$

Les conditions initiales précisent ensuite les valeurs de C' et C''.
C'est un système de deux équations à deux inconnues à résoudre : $\begin{cases} C' & +C'' & =U(t_0) \\ \alpha C' & +\beta C'' & =U'(t_0)\end{cases}$
Il est de Cramer car $\alpha\neq\beta$.
□

Exercice
Résoudre le problème de Cauchy : $y''-2y'+y=te^t$ avec les conditions $y(0)=0$ et $y'(0)=1$.

Correction
L'équation homogène est (H) $y''-2y'+y=0$ d'équation caractéristique $r^2-r+1=(r-1)^2$.
Les solutions de (H) sont de la forme $t\mapsto(A+A't)e^t$.
En exploitant le théorème de la partie précédente, on cherche une solution particulière sous la forme

$$\tilde{y}: t\mapsto(C_1t^2+C_2t^3)e^t$$

Or

$$\tilde{y}''(t)-2\tilde{y}'(t)+\tilde{y}(t)=\left(C_1[(t^2-2t^2+t^2)+(4t-4t)+2)]+C_3[(t^3-3t^3+t^3)+(6t^2-6t^2)+6t]\right)e^t$$

Donc $C_1=0$ et $C_2=\frac{1}{6}$ et les solutions de (E) sont :

$$t\mapsto(A+A't+\frac{1}{6}t^3)e^t\qquad \text{avec } A,A'\in\mathbb{R}$$

Avec les condition de Cauchy, cela donne :

$$\begin{cases} A & & = & 0 \\ A & +A' & = & 1\end{cases}\Leftrightarrow\begin{cases} A=0 \\ A'=1\end{cases}$$

Donc la solution du problème de Cauchy est $y : t \mapsto (t + \frac{1}{6}t^3)e^t$

Exercice

Résoudre les équations différentielles de la physique (cadre)

Correction

5. Bilan

Synthèse

- ⇝ A toute fonction f, on associe (par un tableau) une fonction f' qui est sa dérivée. Réciproquement, on peut essayer de lui associer une fonction F dont elle (=f) serait la dérivée. Cette fonction F non unique s'appelle une primitive de f. On a un tableau de fonctions usuelles à APPRENDRE par coeur.
- ⇝ Il n'existe pas d'algorithme simple de primitivisation comme il en existe de dérivation. La seule chose que l'on sait (et que l'on démontrera plus tard) est que toute fonction continue sur un intervalle I admet une primitive sur cet intervalle I. On fait ce que l'on peut en reconnaissant localement des situations : pour une somme de fonctions, la linéarité suffit ; pour un produit, on exploite une intégration par parties ; pour une composition, on utilise un changement de variable.
- ⇝ Une équation différentielle est une équation (égalité) dont l'inconnue est une fonction et qui associe cette fonction à ses dérivées. L'équation peut être linéaire, d'ordre quelconque (un entier), écrite sous forme normale... On lui associe une courbe intégrale.
  L'ensemble des solutions apparait (pour les équations différentielles linéaires) sous forme d'un espace affine.
- ⇝ Le problème théorique consiste à trouver des hypothèses qui assure l'existence d'une solution (unique ?) à une équation différentielle, sur un intervalle le plu grand possible.
  Pour les équations différentielles linéaires, une théorie satisfaisante est donné par le théorème de Cauchy. A connaître par coeur avec toutes ses hypothèses et toutes ses conclusions (ordre 1 ou 2)...
- ⇝ Le problème pratique consiste à résoudre l'équation. Nous avons mis au point des stratégies : la méthode de la variation de la constante (pour les EDL1). Au passage, on rencontre le problème du recollement de solutions sur des intervalles joints ; ainsi que l'étude des inéquations différentielles.
  Pour les EDL2, nous nous limitons cette année aux équations à coefficients constants. La méthode est simple : il s'agit de mettre en parallèle cette équation différentielle et l'équation caractéristique (polynomiale de degré 2), puis de résoudre cette seconde équation. La forme des solutions et les solutions de cette seconde équation (polynomiale) donnent la forme des solutions et les solutions de la première (différentielle).

Savoir-faire et Truc & **Astuce du chapitre**

- — Savoir-faire - $e^{\alpha x}\cos\beta x$ ou $e^{\alpha x}\sin\beta x$
- — Savoir-faire - $\sin^n x\cos^m x$
- — Savoir-faire - Fractions rationnelles $\dfrac{1}{ax^2+bx+c}$
- — Truc & Astuce pour le calcul - Encadrer une intégrale
- — Savoir-faire - Inégalité des accroissements finis (avec f de classe $\mathscr{C}^1$)
- — Savoir-faire - Obtenir une primitive avec une IPP cachée
- — Savoir-faire - $f(x) = P(x)e^{\alpha x}$

— Savoir-faire - $f(x) = P(x)\ln(Q(x))$
— Savoir-faire - Changement de variable - dans la pratique
— Savoir-faire - Calculer une primitive par changement de variable
— Savoir-faire - Bijection par morceaux
— Savoir-faire - Calcul pour $f(t) = \sin^n t \cos^m t$ avec n ou m impair.
— Savoir-faire - Cas général ($\tan\frac{x}{2}$).
— Savoir-faire - Variable dans les bornes de l'intégrales.
— Savoir-faire - Découper I pour avoir des équations normalisées
— Savoir-faire - Méthode de « variation de la constante ».
— Savoir-faire - Etudier une inéquation différentielle.
— Savoir-faire - Cas non résolue (recollement).
— Savoir-faire - Second membre de la forme $e^{\alpha t}(P(t)\cos\beta t + Q(t)\sin\beta t)$.

Notations

Notations	*Définitions*	*Propriétés*	*Remarq*
$\left[\Phi(t)\right]_{\cdot}^{x} = \Phi(x)$	« Crochet » de Φ	Par extension : $[\Phi]_a^b = \Phi(b) - \Phi(a)$	On trou
$\int_a^b f(t)\mathrm{d}t$	Intégrale de f entre a et b. Il s'agit de l'aire « sous » la courbe	Pour tout $a \in \mathscr{D}_f$ et f continue, $x \mapsto \int_a^x f(t)\mathrm{d}t$ est une primitive de f	
$\int_{\cdot}^{x} f(t)\mathrm{d}t$	(Ensemble des) primitive(s) de f		Attentio primitiv

Retour sur les problèmes

25. Il est miraculeux qu'ils s'agissent des deux faces d'une même pièce. C'est Newton et Leibniz qui s'en rendirent compte les premiers, indépendamment et en prenant des chemins très différents. Certains pensent que Fermat l'avait compris… Un célèbre faux du XIX siècle fit croire à Michel Chasles que le premier avait été en réalité Blaise Pascal. Un mauvais pari de l'académicien français.
26. C'est l'application $x \mapsto 2^x u_0 = u_0 e^{x\ln 2}$
27. Rien de simple. D'où les deux méthodes : intégration par parties (pour un produit) et changement de variable (pour une composition).
28. Oui c'est possible, on l'aperçoit dans le cours. Mais le moins qu'on puisse dire c'est que la technique est très technique !
29. Toute fonction continue est intégrable (voir plus tard dans l'année). La réciproque est fausse : certaine fonction non continue sont néanmoins intégrable. En regardant la courbe, on trouve qu'il y a une compensation : $\int_0^\pi \tan(t)\mathrm{d}t = 0$.
    Mais concrètement, une primitive de tan = $\frac{\sin}{\cos}$ est $-\ln(|\cos|) + K$, mal définie en $\frac{\pi}{2}$. On ne peut inopinément écrire : $+\infty - \infty = 0$…
30. Cf. Cours de physique
31. Le théorème de Cauchy a pour but de répondre à cette question. Dans le cadre des équations linéaires, il allie nombre liberté (ordre de l'équation) et nombre de contraintes (conditions initiales).
32. Nous verrons une autre option pour étudier des équations différentielles : la force brute de calcul de l'ordinateur. La solution ne sera qu'approchée, mais l'approximation sera diablement efficace (méthode d'Euler, Heun…).

Troisième partie

Logique ensembliste

Chapitre 10

Structure logique

Résumé -

En mathématiques, on s'intéresse à des objets sur lesquelles on formule des assertions. Si, partant des axiomes ou des définitions on peut, en respectant des règles logiques, démontrer qu'une assertion est vraie, elle prend alors le nom de théoréme (avec un certain nombre de variantes sur ce nom : proposition (résultat considéré comme un peu moins important qu'un théorème), lemme (résultat qui est avant tout une étape intermédiaire pour arriver au résultat final), corollaire (conséquence plus ou moins immédiate d'un résultat précédemment démontré).
Le but de l'activité mathématique est de prouver de nouveaux résultats. Pour éviter les erreurs, il faut : du bon sens (i.e. de la logique), de la méthode, de la rigueur.
Quelques vidéos de youtubers :
— *Canal unisciel - Logique et raisonnement -*

Sommaire

1. Cours mathématiques

1.1. L'énigme mathématique

Pas de réponse, que des questions. Extrait de «l'efficacité des mathématiques est-elle déraisonnable» de D. Lambert.

? Problème 48 - Déraisonnable efficacité des mathématiques

Le développement des sciences physiques contemporaines a clairement manifesté l'efficacité surprenante des mathématiques. Dans un article souvent cité, E. P. Wigner parle à ce propos d'une « efficacité déraisonnable », d'une sorte de « miracle » qui est comme un « don magnifique que nous ne comprenons ni ne méritons ». Einstein lui-même manifeste son étonnement à cet égard ? : « Comment est-il possible que la mathématique, qui est un produit de la pensée humaine et indépendante de toute expérience, puisse s'adapter d'une si admirable manière aux objets de la réalité ? La raison humaine serait-elle capable, sans avoir recours à l'expérience, de découvrir par la pensée seule les propriétés des objets réels » ? Aujourd'hui cet étonnement est encore renforcé par les confirmations expérimentales très précises apportées à la mécanique quantique, à l'électrodynamique quantique, à la théorie unifiée des interactions électro-faibles (qui a permis la découverte effective des bosons vectoriels intermédiaires) ou encore à la théorie cosmologique standard (grâce aux mesures effectuées par le satellite COBE par exemple). De plus, cette efficacité se manifeste également, quoique de manière plus discrète, dans d'autres domaines des sciences. En biologie, par exemple, les mathématiques apportent des résultats surprenants au niveau de la compréhension des dynamiques de populations en écologie...

? Problème 49 - Un simple langage ? Ou plus

Les mathématiques sont-elles un langage ? Est-ce un jeu ? Un aperçu sur la/une vérité ? Autre chose...

? Problème 50 - Inspiration...

Les deux sources d'inspiration pour les mathématiques jusqu'au XXème siècle sont la physique dans sa globalité (mécanique, électricité, chimie...) et l'arithmétique (le calcul avec des nombres entiers). La géométrie a été aussi d'une certaine façon une source d'inspiration.
Il est donc nécessaire en filière mathématique d'étudier la physique et l'arithmétique...
La biologie et l'informatique (quoi que pour cette dernière, il est parfois compliqué de démêler informatique et mathématiques) sont de nouvelles sources d'inspiration...

1.2. Structure de cours

Analyse - Formalisation

Le cours se construit d'une façon systématique. La chronologie est importante.

1. Quelques axiomes : des résultats de base admis, sans démonstration, sur un *bon sens* commun.
2. A partir de ces axiomes, on développe des idées, des mots, des images mentales et des heuristiques.
   Ces images permettent de bien comprendre mais pas ne permettent pas les

démonstrations.
Une heuristique, c'est l'art de trouver, de découvrir. En pédagogie : l'adjectif heuristique signifie : « Qui consiste à faire découvrir par l'élève ce qu'on veut lui enseigner ».

3. Il faut formaliser les idées pour pouvoir agir dessus : on donne donc des définitions précises avec un langage très précis, mais finalement assez restreint. Une définition, c'est une *légalisation* d'une idée.
4. La manipulation des définitions pour créer des théorèmes ou propositions (moins importantes qu'un théorème, ou étape intermédiaire), des corollaires (résultats immédiats), des lemmes (propositions pour des démonstrations).
5. Les démonstrations sont les étapes intermédiaires entre définitions et propositions ou entre propositions et théorèmes. C'est l'essentiel de notre cours!
6. Pour bien comprendre et apprendre à manipuler les mathématiques, nous aurons beaucoup d'exercices d'applications, de moments d'analyse, des devoirs (maisons ou surveillés) et des colles évidemment!

Exercice
Pour ces six étapes, donner un pourcentage du temps passé en cours de mathématiques pendant le lycée ?
Reprendre l'exercice en fin d'année pour évaluer le cours de CPGE.

Correction

2. Quantificateurs et notations ensemblistes

2.1. Appartenance, éléments

Définition - Ensemble
Un ensemble est une "collection" d'objets appelés *éléments*. On introduit une relation particulière entre un élément x et un ensemble E, la *relation d'appartenance* :

$x \in E$, ce qui se lit "x appartient à E" ou "x est un élément de E.

La négation de la relation d'appartenance s'écrit $x \notin E$, ce qui signifie que $x \in E$ est faux.

Pour aller plus loin - Propriété essentielle, sinon, c'est la faillite...
Trés vite dans la théorie des ensembles sont apparus quelques paradoxes, ce qui a demandé de faire une théorie plus fine. Ici, on reste au stade naïf, car cela demanderait beaucoup de temps et il serait assez peu productif de prendre ce temps.
La contradiction est de cette forme (en notant catalogue ou lieu d'ensemble) : « Peut-on faire un catalogue des catalogues qui ne se citent pas? »

Proposition - Propriété essentielle
Un ensemble est défini dés que pour tout objet x, on **peut dire** si x est, ou n'est pas, un élément de cet ensemble.

Définition - Formalisation
On formalise les idées et objets pour signifier un certain type d'appartenance par :
$\forall x$, qui se lit « quel que soit x » ou pour « pour tout x »,
$\exists x$ se lit « il existe x »,
$\exists ! x$ se lit « il existe un unique x »,

Attention - Pas d'abus
On n'abuse pas de ce formalisme dans un texte en français.
Seul un « $x \in E$ » peut être toléré.

Exemple - Lecture
Par exemple,
$\forall x \in E$ se lit « quel que soit x appartenant à E », « quel que soit l'élément x de E » , ou encore « pour tout élément x de l'ensemble E »...
De plus si $P(x)$ désigne une propriété dépendant de x,
$\forall x \in E, P(x)$ se lit « pour tout x de E, la propriété P de $x\ldots$ »

et $\exists x \in E | P(x)$ (ou $\exists x \in E; P(x)$) se lit « il existe x dans E tel que la propriété P de $x\ldots$ »

Exemple - Ensembles classiques
Les plus connus : $\mathbb{N}$ (dont les éléments sont appelés "entiers naturels"), $\mathbb{Z}$ ("entiers relatifs"), $\mathbb{Q}$ ("nombres rationnels"), $\mathbb{R}$ ("nombres réels"), $\mathbb{C}$ ("nombres complexes"), mais également l'ensemble des élèves de la classe, l'ensemble des professeurs de la classe, l'ensemble des aliments contenus dans le frigo de la cuisine de vos parents...
Exercice
Que pensez-vous de l'affirmation suivante ?
On a donc

$$\forall x \in \mathbb{R}, x^2 \neq -1 \text{ mais } \exists x \in \mathbb{C} \mid x^2 = -1$$

Correction
Remarquez la négation du $\exists \ldots$

Remarque - Convention de notation
La lettre x peut être remplacée par n'importe quel autre symbole, bien que quelques conventions tacites existent en mathématiques (parfois différentes de la physique) : $n, m, p, i, j, k, \ell, \ldots$ pour des entiers, $x, y, z, s, t, \theta, \epsilon, \ldots$ pour des réels, z pour un complexe...

Définition - Règle de la théorie des ensembles
Quelques régles régissent les ensembles (dont certaines sont des axiomes de la théorie des ensembles) :
- régle n°1 : Deux ensembles qui ont les mêmes éléments sont égaux.
- régle n°2 : Il existe un ensemble qui n'admet aucun élément, soit

$$\exists E \mid \forall x, x \notin E$$

D'après la règle n°1, cet ensemble est unique, on l'appelle *ensemble vide* et on le note $\emptyset$.

2.2. Différentes manières d'écrire un ensemble

Descriptions : en extension, en compréhension, par image

Définition - Singleton, paire
Soit a un objet mathématique. L'ensemble dont a est l'unique élément s'appelle un singleton, on le note $\{a\}$.
Soient a et b deux objets distincts. L'ensemble dont ce sont les deux seuls éléments s'appelle la paire formée de a et b. On le note $\{a, b\}$

D'aprés la régle n°1, $\{a, b\} = \{b, a\}$, qu'il ne faut pas confondre avec le couple (a, b).
Si $a = b$, $\{a, b\} = \{a\}$.

Histoire - Asse
Depuis la fin du matique s'appui Le premier a les 1918).

On commence lant de cette th mathématicien nous exclure du

Définition - Définition en extension
On dit que l'on définit un ensemble en extension lorsque l'on énumère ses éléments :

$$\{a_1, a_2, \dots, a_n\}$$

Cette notation sous-entend que l'on sait interpréter les ... intermédiaires.

Exemple - Ensemble défini en extension
$\{1,2,\dots,n\}$ représente l'ensemble des entiers compris entre 1 et n, on le note parfois aussi $[\![1,n]\!]$.
$\{0,2,\dots,2n\}$ est l'ensemble des entiers de la forme $2k$ où k varie de 0 à n.
Par abus courant, la même notation s'emploie pour énumérer des ensembles infinis, comme
$\mathbb{N} = \{0,1,2,\dots\}$ ou $2\mathbb{N} = \{0,2,4,\dots\}$.
Exercice
$\{1,2,3,\dots\} =$, $\{0,1,-1,2,-2,\dots\} =$

Correction
$\{1,2,3,\dots\} = \mathbb{N}$, $\{0,1,-1,2,-2,\dots\} = \mathbb{Z}$

Définition - Définition en compréhension
On peut définir un ensemble en compréhension, c'est à dire par l'intermédiaire d'une propriété qui le caractérise : soit E un ensemble et $P(x)$ une propriété dépendant d'un objet x de E, alors

$$\{x \in E \mid P(x)\}$$

est l'ensemble des x éléments de E tels que $P(x)$ (sous-entendu "tels que $P(x)$ soit vraie").

Pour aller plus loin - Définition par image
On peut définir un ensemble par image, c'est à dire par l'intermédiaire d'une application qui le caractérise : soit E un ensemble et $f : E \to F$ et enfin $A \subset E$.
Alors

$$f(A) = \{f(x), x \in A\}$$

est le sous-ensemble de F dont les éléments sont des images d'éléments de A par f.

Exercice
$\{x \in \mathbb{R} | x^2 + 1 = 0\} =$
$\{x \in \mathbb{N} | x + 1 = 0\} =$
$\{x \in \mathbb{C} | x^2 + 1 = 0\} =$
$\{x \in \mathbb{Z} | x + 1 = 0\} =$
$\{a^2 + 2 | a \in [\![1,5]\!]\} =$

Correction
$\{x \in \mathbb{R} | x^2 + 1 = 0\} = \emptyset$
$\{x \in \mathbb{N} | x + 1 = 0\} = \emptyset$
$\{x \in \mathbb{C} | x^2 + 1 = 0\} = \{-i, i\}$
$\{x \in \mathbb{Z} | x + 1 = 0\} = \{-1\}$
$\{a^2 + 2 | a \in [\![1,5]\!]\} = \{3,6,11,18,27\}$

Le dernier ensemble de cet exercice est un exemple d'ensemble défini comme image.

Intervalle de $\mathbb{R}$

Définition - Intervalles de $\mathbb{R}$
Pour $a, b \in \mathbb{R}$, $a < b$, on définit les *intervalles* de $\mathbb{R}$, ce sont les ensembles suivants :

$$[a,b] = \{x \in \mathbb{R} \mid a \leqslant x \leqslant b\} \quad]a,b] = \{x \in \mathbb{R} \mid a < x \leqslant b\} \quad [a,b[,]a,b[\dots$$
$$]-\infty,a] = \{x \in \mathbb{R} \mid x \leqslant a\} \quad]-\infty,a[= \{x \in \mathbb{R} \mid x < a\} \quad [b,+\infty[,]b,+\infty[\dots$$

Les intervalles du type $[a,b],]-\infty,a], [b,+\infty[$ sont des *intervalles fermés*
Les intervalles du type $]a,b[,]-\infty,a[,]b,+\infty[$ sont des *intervalles ouverts*
Les intervalles du type $[a,b[$ ou $]a,b]$ sont dits *semi-ouverts (ou semi-fermés)*
$[a,b]$ s'appelle un *segment*.

Exemple - Autres exemples

$\emptyset =]0,0[$

$\mathbb{R}_- =]-\infty,0]$, $\mathbb{R}_-^* =]-\infty,0[$, $\mathbb{R}_+ = [0,+\infty[$, $\mathbb{R}_+^* =]0,+\infty[$ (on trouve aussi les notations $\mathbb{R}^+, \mathbb{R}^{+*}$...)

Savoir faire - Montrer que I est un intervalle de $\mathbb{R}$

Il suffit de montrer que pour tout $a,b \in I$, $[a,b] := \{t \in \mathbb{R} \mid a \leqslant t \leqslant b\} \subset I$.
C'est-à-dire :
« Soient $a,b \in I$ (quelconques puis fixés).
Soit $t \in [a,b]$ (i.e. $a \leqslant t \leqslant b$) alors..... et donc $t \in I$. »

Exercice
Montrer que $J = \{x \in \mathbb{R} \mid 1 \leqslant x + e^x \leqslant 10\}$ est un intervalle.

Correction
Soient $a,b \in J$. Soit $t \in [a,b]$.
Notons $\varphi : x \mapsto x + e^x$, croissante comme addition de deux fonctions croissantes.
Alors $a \leqslant t \leqslant b \Rightarrow \underbrace{1 \leqslant \varphi(a)}_{a \in J} \leqslant t + e^t \leqslant \underbrace{\varphi(b) \leqslant 10}_{b \in J}$

2.3. Utilisation de quantificateurs

Pour aller plus loin - Motivations
Il y a d'autres motivations.
Nous verrons plus loin l'articulation forte entre ces deux quantificateurs, en particulier pour la négation.
Nous verrons encore plus l'articulation très forte entre ces connecteurs et les opérations ensemblistes $\bigcap_{i\in I} A_i$ et $\bigcup_{i\in I} A_i$

Heuristique - Quantificateurs nécessaires

En mathématiques classiques, deux quantificateurs sont essentiels :
- $\forall$, lu « pour tout » ou « quel que soit », pour désigner qu'une propriété a un certain degré d'universalité :
  $\forall x, \mathscr{P}(x)$. La propriété $\mathscr{P}$ est donc toujours vraie, puisque vraie pour tout point.
  $\forall x \in E, x \in F$. Tous les éléments de E sont dans F. Dans sa totalité : $E \subset F$.
- $\exists$ lu « il existe » est la négation du précédent.

Remarque - Existence ET unicité

Il arrive qu'on ait besoin d'ajouter l'unicité à l'existence (cas des antécédents pour une fonction bijective...).
On écrit alors : $\exists !x$, pour dire « il existe un unique x qui... »

Analyse - Non commutativité des connecteurs

Les deux assertions suivantes ne sont pas identiques :

$$\forall x, \exists y \ldots \qquad \exists y, \forall x \ldots$$

L'une est plus forte que l'autre : la seconde. Pour la seconde assertion, il s'agit du même y pour tous les x. Dans la première assertion, chaque y dépend de chaque x.

Exemple - Suite majorée, ou rien

$\forall n \in \mathbb{N}, \exists M \in \mathbb{R}$ tel que $u_n \leqslant M$ est toujours vraie. Comme M peut dépendre de n, on peut prendre $M_n = x_n + 1$
$\exists M \in \mathbb{R}$ tel que $\forall n \in \mathbb{N}, u_n \leqslant M$ signifie qu'une suite est majorée. Ce n'est pas toujours vraie (ex : $(u_n) = (n)$).

Exercice
On considère la suite de FIBONACCI : $F_{n+2} = F_{n+1} + F_n$ et $F_0 = F_1 = 1$.
Que pensez-vous des deux assertions suivantes :
- $\forall n \in \mathbb{N}, \exists A,B \in \mathbb{R}$ tels que $F_n = A\left(\frac{1+\sqrt{5}}{2}\right)^n + B\left(\frac{1-\sqrt{5}}{2}\right)^n$
- $\exists A,B \in \mathbb{R}$ tels que $\forall n \in \mathbb{N}$, $F_n = A\left(\frac{1+\sqrt{5}}{2}\right)^n + B\left(\frac{1-\sqrt{5}}{2}\right)^n$

Correction
Le première est sans intérêt cela est toujours vrai. Il suffit de prendre $B_n = 0$ et $A_n = \dfrac{F_n}{\left(\frac{1+\sqrt{5}}{2}\right)^n}$.
La seconde est tout à fait intéressante, voire incroyable !

Savoir faire - Noter les dépendances

Si l'on écrit $\forall\ a, \exists\ b\ldots$, le nombre b en second dépend grandement de a. On devrait noter $b(a)$ ou b_a.
En revanche, si on écrit $\exists\ b$ tel que $\forall\ a\ldots$, le nombre a ne dépend pas (plus particulièrement que les autres) de b. La notation a_b ou $a(b)$ n'aurait pas d'intérêt. Rappelons que cette formulation est la plus forte.
De nombreux problèmes rencontrés en mathématiques en MPSI par les élèves reposent sur la non compréhension de cette (in)dépendance.

Une autre stratégie est de faire comme en Python, des indentations dans la démonstration.
A chaque paramètre introduit, on fait apparaître un petit retrait dans l'écriture de la démonstration qui permet de voir comme ces paramètres dépendent mutuellement les uns des autres...

2.4. Parties d'un ensemble

Définition - Partie d'un ensemble (sous ensemble)

On dit qu'un ensemble F est inclus dans un ensemble E, ce que l'on note $F \subset E$, si tous les éléments de F sont éléments de E.
On dit aussi que F est une partie ou un sous-ensemble de E.

Savoir faire - Montrer que $F \subset E$

Pour montrer que $F \subset E$, on démontre :
$$F \subset E \Leftrightarrow (\forall x,\ x \in F \Rightarrow x \in E).$$

Proposition - Relation d'ordre

Quelques propriétés :
- $\emptyset \subset E$ pour tout ensemble E
- $E \subset E$ (l'inclusion est une relation réflexive)
- Si on a $E \subset F$ et $F \subset G$ alors on a aussi $E \subset G$ (l'inclusion est une relation transitive)
- Si on a $E \subset F$ et $F \subset E$ alors $E = F$ (l'inclusion est une relation antisymétrique)

Savoir faire - Prouver l'égalité de deux ensembles

La dernière propriété sert souvent à prouver l'égalité de deux ensembles.

Exercice
A quelle condition a-t-on : $\{a\} \subset E$? $\{a,b\} \subset E$? $\{a\} \subset \{b\}$?

Correction
$\{a\} \subset E$ ssi $a \in E$ $\quad$ $\{a,b\} \subset E$ ssi $a,b \in E$ $\quad$ $\{a\} \subset \{b\}$ ssi $a = b$ ssi $\{a\} = \{b\}$

Définition - Ensemble des partes de E

$\mathscr{P}(E)$ est l'ensemble des parties de E : $X \in \mathscr{P}(E)$ signifie donc à $X \subset E$.

Exemple - Singleton...

$\mathscr{P}(\{a\}) = \{\emptyset, \{a\}\}$
$\mathscr{P}(\{a,b\}) = \{\emptyset, \{a\}, \{b\}, \{a,b\}\}$
$\mathscr{P}(\emptyset) = \{\emptyset\}$
$\mathscr{P}(\mathscr{P}(\emptyset)) = \{\emptyset, \{\emptyset\}\}$

2.5. Produit cartésien

Définition - Produit cartésien de deux ensembles
Soient E et F deux ensembles. On appelle produit cartésien de E et F, et on note $E \times F$, l'ensemble dont les éléments sont les couples formés d'un élément de E et d'un élément de F (dans cet ordre) :

$$E \times F = \{x | \exists a \in E, \exists b \in F;\ x = (a,b)\} = \{(a,b) | a \in E \text{ et } b \in F\}.$$

Remarque - Rôle de la ponctuation
Le " ;" dans la définition en compréhension peut être remplacé par " :", "tels que".
Plus généralement,

Définition - Produit cartésien de n ensembles
Soit $n \in \mathbb{N}$, si $E_1, \dots, E_n$ sont n ensembles, on définit le produit cartésien $E_1 \times \cdots \times E_n$ comme l'ensemble des n-uplets $(a_1, \dots, a_n)$ formés d'éléments $a_1 \in E_1, \dots, a_n \in E_n$.
Si les E_i désignent un même ensemble E, on note $E_1 \times \cdots \times E_n = E^n$ ($\mathbb{R}^2, \mathbb{R}^3$...)

Remarque - Associativité du produit cartésien
A priori $E \times (F \times G) \neq (E \times F) \times G$, mais on les identifie fréquemment à $E \times F \times G$.

2.6. Opérations sur les ensembles

Définition - Réunion, intersection, différence d'ensembles
Pour E et F deux ensembles on définit :
- la réunion (ou union) de E et F : $E \cup F = \{x | x \in E \text{ ou } x \in F\}$ (le "ou" est inclusif : on peut avoir les deux simultanément)
- l'intersection de E et F : $E \cap F = \{x | x \in E \text{ et } x \in F\}$
- la différence de E et de F : $E \setminus F = \{x | x \in E \text{ et } x \notin F\}$

Remarque - Interprétation avec une table de vérité
En d'autres termes on a :

$x \in E$	$x \in F$	$x \in E \cup F$	$x \in E \cap F$	$x \in E \setminus F$
V	V	V	V	F
V	F	V	F	V
F	V	V	F	F
F	F	F	F	F

Définition - Complémentaire d'un ensemble (dans un ensemble plus gros)
Lorsque $F \subset E$, l'ensemble $E \setminus F$ est appelé *complémentaire* de F dans E et noté $\complement_E F$.
On ne peut parler de complémentaire de l'ensemble F que relativement à un ensemble "contenant" ce dernier. Toutefois, en l'absence d'ambiguïté sur E, on peut noter $\complement F$ ou F^c ou $\overline{F}$ (cette dernière notation ayant cependant différents sens suivant le domaine des mathématiques...)

Attention - Le verbe contenir
Attention également à l'ambiguïté du verbe "contenir", parfois utilisé pour dire que x est élément de E, parfois pour dire que F est une partie de E.

Exercice
Soit $A = \{x \in \mathbb{N} \mid x/2 \in \mathbb{N} \text{ et } x \geqslant 10\}$. Que représente A ? Ecrire cet ensemble différemment. Ecrire en extension $\complement_{2\mathbb{N}} A$ puis déterminer $\complement_{\mathbb{N}} A$.
Déterminer les sous-ensembles X de $\mathbb{N}$ tels que $A \cup X = \mathbb{N}$.

Correction
A est l'ensemble des nombres paires plus grand que 10. $\complement_{2\mathbb{N}} A = \{0,2,4,6,8\}$.
$\complement_{\mathbb{N}} A = \{0,2,4,6,8\} \cup \{2k+1, k \in \mathbb{N}\}$. On doit prendre des ensembles qui contiennent au moins $\complement_{\mathbb{N}} A$

Exercice
Soit E un ensemble. Que peut-on dire de deux parties A et B de E vérifiant $A \cap B = A \cup B$?

Correction
$A = B$. En effet, nécessairement : $A \cup B \subset A \cap B \ldots$

Pour aller plus loin - Probabilité
Nous reprenons ces notions lors du cours de probabilité. Il y aura quelques modifications de vocabulaire

Exercice
On définit la différence symétrique de deux ensembles E et F par

$$E \Delta F = (E \setminus F) \cup (F \setminus E).$$

Ecrire $E \Delta F$ à l'aide de $E \cup F$ et $E \cap F$.

Correction
$E \Delta F = (E \cup F) \setminus (E \cap F)$

Proposition - Quelques règles de calcul
E, F et G désignent trois ensembles quelconques.

$$E \cup F = F \cup E \quad \text{(commutativité de la réunion)}$$
$$E \cup (F \cup G) = (E \cup F) \cup G \quad \text{(associativité de la réunion)}$$
$$\emptyset \cup E = E \cup \emptyset = E \quad \text{(l'ensemble vide est neutre pour la réunion)}$$
$$E \cap F = F \cap E \quad \text{(commutativité de l'intersection)}$$
$$E \cap (F \cap G) = (E \cap F) \cap G \quad \text{(associativité de l'intersection)}$$
$$\emptyset \cap E = E \cap \emptyset = \emptyset \quad \text{(l'ensemble vide est absorbant pour l'intersection)}$$
$$E \cap (F \cup G) = (E \cap F) \cup (E \cap G) \quad \text{(distributivité de l'intersection par rapport à la réunion)}$$
$$E \cup (F \cap G) = (E \cup F) \cap (E \cup G) \quad \text{(distributivité de la réunion par rapport à l'intersection)}$$

Pour A et B deux parties de E :

$$A \cap E = A$$
$$A \cup E = E$$
$$\complement_E(\complement_E A) = A$$
$$\complement_E(A \cup B) = (\complement_E A) \cap (\complement_E B)$$
$$\complement_E(A \cap B) = (\complement_E A) \cup (\complement_E B)$$

Les deux dernières formules sont connues sous le nom de *lois de Morgan.*

Pour la démonstration, on lie ces résultats aux affirmations correspondantes.
On peut aussi faire une table exhaustive de vérité

Démonstration
$x \in \complement(A \cap B) \iff x \notin A \cap B \iff \text{Non}(x \in A \cap B)$
$\iff \text{Non}(x \in A \text{ et } x \in B) \iff \text{Non}(x \in A) \text{ ou Non} x \in B$
$\iff x \notin A \text{ ou } x \notin B \iff x \in \complement(A) \text{ ou } x \in \complement(B) \iff x \in \complement(A) \cup \complement(B)$ □

3. Vocabulaire sur les assertions

3.1. Définitions

Définition - Assertion (proposition) et prédicat
Dans ce paragraphe une *proposition*, ou *assertion* est un énoncé qui peut prendre deux valeurs logiques : V (vrai) ou F (faux).
Si cette assertion dépend d'une variable x on parle alors de *prédicat*.

Exemple - Propositions
« tout entier est pair » est une assertion de valeur logique F ;
« tout réel est un complexe » est une assertion de valeur logique V ;
« 4 est pair » est une assertion de valeur logique V ;
« x est un multiple de 4' » est un prédicat dont la valeur logique dépend de la valeur de x.
Comme on peut le voir dans les parties suivantes, à partir de deux assertions A et B, on en définit d'autres dont la valeur logique est donnée par une *table de vérité*.

Exercice
Que pensez-vous de $\mathscr{P}_n$ dans l'énoncé formel suivant ?

$$\text{Notons, } \forall\, n \in \mathbb{N}, \mathscr{P}_n : \text{« } \exists\, k \in E \text{ tel que } a_n = k \text{ »}$$

Correction
Il s'agit d'une famille (suite) de prédicats, il y a donc $\mathscr{P}_0, \mathscr{P}_1 \dots \mathscr{P}_{1224} \dots$.
Peut-être que certains sont vraies et d'autres sont faux...

Pour aller plus loin - Logique floue (1)
Formalisée en 1965, la logique floue élargie ce point de vue : on associe un nombre entre 0 et 1 pour mesurer la véracité de chaque proposition. Une proposition de valeur 0 est fausse. Une proposition de valeur $\frac{1}{2}$ est à moitié fausse. L'arithmétique qui en découle est comparable à celle des probabilités d'événements.
Cela s'appuie sur les ensembles floues avec des degré d'appartenance entre le $\in$ et le $\notin$.
On l'utilise en automatique ou en médecine...

3.2. Négation

Définition - Négation d'une assertion
Considérons une assertion A.
On appelle négation de A l'assertion qui dit le contraire de A, c'est à dire qui est vraie exactement lorsque A est fausse, on la note "non$A$" (ou $\neg A$ en logique).
La *table de vérité* de non$A$:

A	non A
V	F
F	V

Exercice
La négation de « L'hiver dernier il a plu tous les jours à Toulouse » est :
La négation de « Chaque hiver, il neige au moins un jour en Aveyron » est :

Soit I un intervalle de $\mathbb{R}$.
La négation de « $0 \in I$ » est
La négation de « $\forall x \in I,\ x > 0$ » est
La négation de « $\exists x \in I \mid x \geqslant 0$ » est

Correction
La négation de « L'hiver dernier il a plu tous les jours à Toulouse » est « il n'a pas plu (au moins) un jour de l'hiver dernier à Toulouse » :
La négation de « Chaque hiver, il neige au moins un jour en Aveyron » est « il y a (eu) des hivers sans neige en Aveyron ».
Soit I un intervalle de $\mathbb{R}$.
La négation de « $0 \in I$ » est : « $0 \notin I$ »
La négation de « $\forall x \in I,\ x > 0$ » est « $\exists\, x \in I,\ x \leqslant 0$ »
La négation de « $\exists x \in I \mid x \geqslant 0$ » est « $\forall\ x \in I, x < 0$ »

D'une manière plus générale il faut savoir nier une proposition écrite avec des quantificateurs :
Exercice
P désignant une propriété dépendant de x ou de x, y suivant les cas, écrire la négation des assertions suivantes :

1. $\forall x \in E, P(x)$;
2. $\exists x \in E | P(x)$;
3. $\forall x \in E, \exists y \in E | P(x,y)$;
4. $\exists x \in E | \forall y \in E, P(x,y)$;
5. $\exists r \in \mathbb{R}, \exists s \in \mathbb{R} | \forall x \in \mathbb{R}, x \leqslant r$ et $s \leqslant r$.

Correction

1. $\exists\, x \in E, \neg P(x)$;
2. $\forall\, x \in E | \neg P(x)$;
3. $\exists\, x \in E, \forall\, y \in E | \neg P(x,y)$;
4. $\forall\, x \in E | \exists\, y \in E, \neg P(x,y)$;
5. $\forall\, r \in \mathbb{R}, \forall\, s \in \mathbb{R} | \exists\, x \in \mathbb{R}, x > r$ ou $s > r$.

L'exercice suivant permet de revoir également la table de vérité d'une conjonction (« et ») ou disjonction (« ou ») d'assertions.

Exercice

Compléter le tableau suivant :

A	B	A et B	A ou B	non (A et B)	non (A ou B)	non A	non B	
V	V	V	V					
V	F	F	V					
F	V	F	V					
F	F	F	F					

On a laissé une colonne pour des « tests ». Quelle relation remarquez-vous ?

Correction

A	B	A et B	A ou B	non (A et B)	non (A ou B)	non A	non B	non A et non B	non A ou non B
V	V	V	V	F	F	F	F	F	F
V	F	F	V	V	F	F	V	F	V
F	V	F	V	V	F	V	F	F	V
F	F	F	F	V	V	V	V	V	V

On démontre les lois de Morgan :

— non(A et B)$\Leftrightarrow$ (non A) ou (non B)

— non(A ou B)$\Leftrightarrow$ (non A) et (non B)

3.3. Implications et équivalence d'assertions

Heuristique - Comment exploiter une implication

On exploite une implication du type $A \Rightarrow B$, en règle générale lorsqu'on veut dire :

— A chaque fois que A est vraie, B est vraie. (c'est le vrai $\Longrightarrow$) ;

— et, dans ce cas A est vraie ;

alors, on peut conclure que B est nécessairement vraie.

1. Première conclusion : Il faut différencier $A \Longrightarrow B$ de [$A \Longrightarrow B$ et A]. Si vous souhaitez exprimer ce deuxième fait, vous aurez le droit (provisoirement) d'écrire : $A \overset{\&}{\Rightarrow} B$.

2. Deuxième conclusion : Si sA est faux alors on peut tout avoir pour B. Il est donc possible d'avoir A faux et B vrai lorsque $A \Rightarrow B$. En revanche, la seule impossibilité lorsque $A \Rightarrow B$ et d'avoir : A vrai et B faux.

Définition - Implication d'assertions

Considérons deux assertions A et B.

Si, lorsque l'assertion A est vraie, alors, nécessairement, l'assertion B l'est également, on dit que A implique B et l'on écrit $A \Rightarrow B$

(ce qui se lit donc "A implique B" ou "si A alors B").

Plus précisément, la table de vérité de $A \Rightarrow B$ est donnée par :

A	B	$A \Rightarrow B$
V	V	V
V	F	F
F	V	V
F	F	V

Pour aller plus loin - Logique floue (2)
A la place de V et F, on peut respectivement écrire 1 et 0.
Dans ce cas si p_i est une proposition et $v(p_i)$ sa valeur ($v(p_i) = 0$, si p_i est fausse...).
On a donc $v(p_1 \cap p_2) = v(p_1) \times v(p_2)$, $v(non(p_1)) = 1 - v(p_1)$ et $v(p_1 \cup p_2) = v(p_1) + v(p_2) - v(p_1)v(p_2)$.
Cet arithmétique se transpose aisément en logique floue.

Remarque - Et si A est fausse
On remarquera qu'en logique, dés que A est fausse, $A \Rightarrow B$ est évaluée vraie, en bref, vous pouvez construire sans problème une démonstration juste avec une hypothèse fausse, mais finalement c'est sans intérêt parce que vous n'avez aucun résultat à énoncer à la fin...c'est la raison pour laquelle usuellement "prouver $A \Rightarrow B$" sous entend "prouver A vraie $\Rightarrow B$ vraie ".
Et donc si $E = \emptyset$, tout énoncé "$\forall x \in E, P(x)$" (ou "$x \in E \Rightarrow P(x)$") a la valeur logique V ...

Exercice
Quelle est l'assertion qui a même table de vérité que « non$(A \Rightarrow B)$' » ?

Correction
Il s'agit de « A et non(B) » (il suffit de faire une table de vérité)

Avec deux implications, on a exactement une équivalence :

Définition - Equivalence d'assertions
Considérons deux assertions A et B.
On dit que A et B sont équivalentes si elles signifient la même chose, mais dite différemment, c'est à dire si, simultanément, A implique B et B implique A;
on note alors $A \Leftrightarrow B$.
Plus précisément, la table de vérité de $A \Leftrightarrow B$ est donnée par :

A	B	$A \Leftrightarrow B$
V	V	V
V	F	F
F	V	F
F	F	V

Exemple - Deux assertions équivalentes
On a par exemple pour x réel : $(x > 0) \Leftrightarrow (-x < 0)$.

Exercice
Comparer la table d'équivalence de $A \Leftrightarrow B$ avec celle de $[(A \Rightarrow B)$ et $(B \Rightarrow A)]$

Correction

A	B	$A \Leftrightarrow B$	$A \Rightarrow B$	$B \Rightarrow A$	$[A \Rightarrow B$ et $B \Rightarrow A]$
V	V	V	V	V	V
V	F	F	F	V	F
F	V	F	V	F	F
F	F	V	V	V	V

Attention - A démontrer?
Certaines équivalences correspondent en fait à la définition d'un objet mathématique, d'autres en revanche nécessitent une démonstration.

Attention - Ne pas abuser de $A \Longleftrightarrow B$
Les étudiants écrivent TROP souvent $A \Longleftrightarrow B$, en faisant un calcul dans leur tête (ou une démonstration) qui permet de passer de A à B, SANS vérifier si l'on passe aussi de B à A.
Il est important de ne pas faire cette erreur, surtout si l'on demande qu'une implication.... Il ne faut pas en faire trop, si c'est faux!

4. Principales méthodes de démonstration

4.1. Démonstration d'une implication

Démonstration directe

Considérons deux assertions A et B. On veut démontrer que $A \Rightarrow B$.

Savoir faire - $A \Rightarrow B$. Raisonnement direct

On suppose A vraie, par une succession d'implications connues (calculs, résultats de théorèmes...), on prouve qu'alors B est vraie.

Attention - Abus

Il existe un abus courant de la notation $\Rightarrow$ chez les étudiants (mais aussi les professeurs).
Comme réflexe de protection face à ces abus, il est conseillé de lire $A \Rightarrow B$ à voix haute en disant *« si A est vraie alors B est vraie »* plutôt que *« A implique B »*. Cela permet d'insister sur la prémisse.

Exercice

Montrer que si (x, y) est élément de $]0,2[\times]-2,0[$ alors $\dfrac{1}{x} - \dfrac{1}{y} > 1$

Correction

$$\left.\begin{array}{l} x \in]0,2[\Longrightarrow \dfrac{1}{x} > \dfrac{1}{2} \\ y \in]-2,0[\Longrightarrow \dfrac{1}{y} < \dfrac{-1}{2} \end{array}\right\} \Longrightarrow \frac{1}{x} - \frac{1}{y} > \frac{1}{2} - \frac{-1}{2} = 1$$

Exercice

Soit f la fonction définie sur $\mathbb{R}$ par $f(x) = |x + \frac{3}{2}| - \frac{1}{2}$. Montrer que

$$x \geqslant -1 \Rightarrow f(x) \geqslant 0$$

Correction

Si $x \geqslant -1$, alors $x + \frac{3}{2} \geqslant \frac{1}{2} \geqslant 0$, donc $f(x) = |x + \frac{3}{2}| - \frac{1}{2} = x + \frac{3}{2} f(x) - \frac{1}{2} \geqslant 0$

On peut-être amené à faire une disjonction de cas :

Exercice

Compléter l'énoncé suivant pour que la démonstration nécessite l'étude de deux cas séparés.
Soit f la fonction définie par ...
Montrer que $x \geqslant -1 \Rightarrow f(x) \geqslant 0$.

Correction

On peut par exemple prendre $f(x) = |x + \frac{3}{2}| + |x - 1| - \frac{1}{2}$.
Dans ce cas,

- si $x \in [-1,1]$, $x - 1 < 0$ et $x + \frac{3}{2} > 0$, donc $f(x) = (x + \frac{3}{2}) + (1 - x) - \frac{1}{2} = 2 \geqslant 0$
- si $x \geqslant 1$, $x - 1 > 0$ et $x + \frac{3}{2} > 0$, donc $f(x) = (x + \frac{3}{2}) + (x - 1) - \frac{1}{2} = 2x \geqslant 0$, car $x > 0$

Le chemin de la démonstration

Heuristique - Démontrer que ...

Lorsqu'on cherche à démontrer un résultat, il y a fondamentalement deux attentes :

1. Trouver la (une) démonstration,
   satisfaisante i.e. qui donne la certitude du fait
2. Ecrire la démonstration,
   de sorte que toute personne qui lise la démonstration soit également persuadé du fait

Avec le temps du passage de 1 à 2, il faut donc trois temps dans la recherche d'une démonstration.

Attention - Premier piège

Le temps 1 est le temps de l'analyse.
Le temps 2 est le temps de la synthèse.
Ce sont deux choses très différentes. Lorsque vous lisez une démonstration d'un théorème faite par un professeur, un corrigé dans un livre… vous ne voyez que le second temps, celui de la synthèse. Après la lecture, vous pouvez vous dire : « et oui, je vois que c'est vrai », mais vous n'avez pas appris *comment on trouve* la démonstration!!!
La seule solution est de chercher, chercher, chercher… et de ne pas se précipiter sur la (une) solution.
De même, si pour vous écrire une démonstration de cours lors d'une colle est uniquement un exercice de mémoire, alors c'est que vous n'avez pas compris le premier point, ni le point $1 \to 2$ de la démonstration du fait considéré. *Pouvez-vous donner un exemple d'une telle situation rencontrée ?*

Analyse - La métaphore du petit poucet

Faire une démonstration, c'est partir d'un point A (les hypothèses) pour arriver à un point B (la conclusion).
Il s'agit donc de **trouver un chemin**, sans se perdre en route…

- On peut avancer discrètement, les yeux fermés
  ou bien laisser des traces afin de pouvoir revenir (classique trace du petit poucet)
- On peut tourner autour du point de départ (on parle de mouvement brownien)
  ou bien fixer son cap et se diriger en direction du point B (même s'il peut y avoir quelques obstacles en chemin)
  (mettre un phare, prendre une boussole qui indique une direction)
  voire placer quelqu'un d'autre en B, le laisser venir vers nous et nous vers lui et chercher à faire la jonction (s'appeler, mettre un phare)
- On peut se précipiter sur son GPS
  ou bien chercher à reconnaître là où l'on se trouve, voir si on ne voit pas une route déjà connue (courage, familiarité avec le chemin) le monstre
- Prendre du recul, de la hauteur… prendre le temps de voir de plus haut (bottes de 7 lieux)

Exemple - Exercice « de base »

On suppose que $f : E \to E$ est surjective, montrer que $f^2 (= f \circ f)$ est surjective.
Si l'on n'écrit pas ce que l'on veut obtenir i.e. le point B : « f^2 est surjective », on ne peut s'en sortir uniquement en dérivant du point A : « f est surjective ».
Pour préparer la démonstration, il faut donc laisser des espaces et compléter au fur et à mesure.
Le temps joue un rôle important, or il n'y parait plus lorsque la démonstration est écrite. Pour le voir à l'oeuvre, nous allons présenter la démonstration comme succession de photos prises de son écriture.

1er 2eme 3eme 4eme 5eme 6eme 7eme
Soit $y \in E$
f est surjective donc $\forall\, b \in E, \exists\, a \in E$ tel que $b = f(a)$
ainsi $\exists\, x_1$ tel que $y = f(x_1)$
et $\exists\, x_2$ tel que $x_1 = f(x_2)$
donc $\exists\, x (= x_2)$ tel que $f^2(x) = f(x_1) = y$
$\forall\, y \in E, \exists\, x \in E$ tel que $y = f^2(x)$
donc f^2 est surjective

Cela s'écrit ensuite :
Soit $y \in E$
f est surjective donc $\forall\, b \in E, \exists\, a \in E$ tel que $b = f(a)$.
ainsi $\exists\, x_1$ tel que $y = f(x_1)$
et $\exists\, x_2$ tel que $x_1 = f(x_2)$
donc $\exists\, x (= x_2)$ tel que $f^2(x) = f(x_1) = y$
Donc : $\forall\, y \in E, \exists\, x \in E$ tel que $y = f^2(x)$ et finalement : f^2 est surjective

Exercice
Soit E un ensemble et $f : E \to E$, une application sur E.
On suppose que $f^3 = f$. Montrer que si f est injective alors f est surjective. Et réciproquement.

Pour aller plus loin - f surjective, injective

On verra plus loin les définitions. A ce stade, il suffit de savoir que par définition f est surjective de E sur F si $\forall\, y \in F, \exists\, x \in E$ tel que $y = f(x)$.
Et par définition f est injective de E (sur F) si $\forall\, x_1, x_2 \in E, f(x_1) = f(x_2) \Leftrightarrow x_1 = x_2$.

Correction

1er — 2eme — 3eme : Soit $y \in E$ — 4eme : $f^2(x) = f(y)$ et $f(x) = y$ car f injective — 5eme : Soit $x = f(y)$. Alors $f^2(x) = f^3(y)$ et puisque $f^3 = f$

2eme : $\forall\ y \in E, \exists\ x \in E$ tel que $y = f(x)$

1er : donc f^2 est surjective

Ce qui s'écrit :
Soit $y \in E$.
Prenons $x = f(y)$, alors $f^2(x) = f^3(y) = f(y)$ car $f^3 = f$.
Puis comme f est injective : $f(x) = y$.
Donc il existe $x \in E$ tel que $y = f(x)$.
Par conséquent : $\forall\ y \in E, \exists\ x \in E$ tel que $y = f(x)$, donc f surjective.
De même, réciproquement

1er — 2eme — 3eme : Soit $x_1, x_2 \in E$ tel que $f(x_1) = f(x_2)$ — 4eme : $\exists\ a_1, a_2 \in E$ tels que $x_1 = f(a_1)$ et $x_2 = f(a_2)$ car f surjective — 5eme : ainsi $f^2(a_1) = f(x_1) = f(x_2) = f^2(a_2)$, puis $f^3(a_1) = f^3(a_2)$ — 6eme : or $f^3 = f$, donc $x_1 = f(a_1) = f(a_2) = x_2$

3eme : donc $x_1 = x_2$

2eme : $\forall\ x_1, x_2 \in E,\ f(x_1) = f(x_2) \Rightarrow x_1 = x_2$

1er : donc f est injective

Ce qui s'écrit :
Soient $x_1, x_2 \in E$ tels que $f(x_1) = f(x_2)$.
Comme f est surjective, il existe $a_1, a_2 \in E$ tels que $x_1 = f(a_1)$ et $x_2 = f(a_2)$.
Et donc en composant par f : $f^2(a_1) = f(x_1) = f(x_2) = f^2(a_2)$, puis $f^3(a_1) = f^3(a_2)$
Mais comme $f^3 = f$, alors $f(a_1) = f(a_2)$, soit $x_1 = x_2$.
On a donc démontré : $\forall\ x_1, x_2 \in E$ tels que $f(x_1) = f(x_2)$, alors $x_1 = x_2$.
Donc f est injective.

Contraposée

Proposition - Contraposée
$(\text{non}\, B \Rightarrow \text{non}\, A)$ s'appelle la contraposée de $(A \Rightarrow B)$.
Il est équivalent de prouver l'une ou l'autre de ces deux implications

Démonstration
On complète la table de vérité :

A	B	non A	non B	$A \Rightarrow B$	non $B \Rightarrow$ non A
V	V	F	F	V	V
V	F	F	V	F	F
F	V	V	F	V	V
F	F	V	V	V	V

□

Savoir faire - $A \Rightarrow B$. Raisonnement par contraposée
On suppose donc B fausse et on prouve qu'alors A est fausse, comme précédemment

Le résultat de l'exercice suivant sera fréquemment exploité en analyse :
Exercice
On considére un nombre réel $x \geqslant 0$. Montrer que

$$(\forall \epsilon > 0, 0 \leqslant x \leqslant \epsilon) \Rightarrow x = 0.$$

Correction
On fait donc un raisonnement par contraposée (c'est presqu'un raisonnement par l'absurde ici).
Si $x>0$, alors avec $\epsilon=\frac{x}{2}$, on a la preuve de : $\exists\,\epsilon>0$ tel que $x>\epsilon$.
Donc l'assertion négative de notre hypothèse A est alors vérifiée.
Ainsi non$(B)\Rightarrow$ non(A) et par contraposée : $A\Rightarrow B$.

4.2. Démonstration d'une équivalence

Deux possibilités pour prouver l'équivalence $A\Leftrightarrow B$:

Savoir faire - $A\Leftrightarrow B$. On procède en deux temps

1. On montre $A\Rightarrow B$
2. On montre $B\Rightarrow A$

Exercice
On considère une fonction f définie sur $\mathbb{R}$ à valeurs dans $\mathbb{R}$. Montrer que

$$(f \text{ est une fonction paire et impaire }) \Leftrightarrow (f \text{ est la fonction nulle }).$$

Correction
En deux temps.

— Supposons que f est nulle.
Alors pour tout $x\in\mathbb{R}$, $f(-x)=0=f(x)$ donc f est paire et $f(-x)=0=-f(x)$ donc f est impaire. Bilan : f est nulle $\Longrightarrow$ f est une fonction paire et impaire.

— Réciproquement, si f est une fonction paire et impaire,
Alors pour tout $x\in\mathbb{R}$, $f(x)=f(-x)$ et $f(x)=-f(-x)$, donc $f(x)=-f(x)$ donc $2f(x)=0$ donc $f(x)=0$.
Ceci est vrai pour tout $x\in\mathbb{R}$, donc $f=0$ (f est identiquement la fonction nulle).
Bilan : f est une fonction paire et impaire $\Longrightarrow$ f est nulle.

Par double implication : f est nulle $\Longleftrightarrow$ f est une fonction paire et impaire.

Savoir faire - $A\Leftrightarrow B$. On procède par équivalences connues successives.

Cette méthode est surtout utilisée pour des résolutions calculatoires.
Attention de ne pas en abuser : *il faut à chaque étape être sûr que l'on peut « remonter » les équivalences.*

Exercice
Montrer que

$$\begin{cases} 2x+y=2 \\ 3x+4y=3 \end{cases} \Longleftrightarrow (x,y)=(1,0)$$

Correction
Les règles d'équivalence sur les systèmes sont strictes, il faut bien les respecter ! En particulier la substitution mal employée ne conserve pas les équivalences de système. Nous reverrons cela plus tard dans l'année.

$$\begin{cases} 2x+y=2 \\ 3x+4y=3 \end{cases} \Leftrightarrow \left\{\begin{array}{rcl|l} 2x & +y & = 2 & \\ -5x & & = -5 & L_2\leftarrow L_2-4L_1 \end{array}\right. \Leftrightarrow \left\{\begin{array}{rcl|l} y & = & 0 & L_1\leftarrow L_1+\frac{2}{5}L_2 \\ x & = & 1 & \end{array}\right.$$

Insistons :

Truc & Astuce pour le calcul - Ne pas abuser de $\Leftrightarrow$

Il faut éviter le plus possible d'écrire $\Leftrightarrow$ comme un tic de langage.

1. Si il n'est pas utile, on ne le note pas !
2. Si on choisit de le noter, on vérifie bien à chaque étape le sens $\Leftarrow$ tout particulièrement.

4.3. Raisonnement par l'absurde

Pour aller plus loin - Raisonnement par l'absurde, à accepter ?
Certaine axiomatique de mathématique refuse le raisonnement par l'absurde, en effet la notion d'existence qui en découle est quelque peut frustrante.
On sait que $\sqrt{2}$ n'est pas un nombre rationnel

Savoir faire - Raisonnement par l'absurde

Pour démontrer un certain énoncé, on fait l'hypothèse qu'il est faux et on aboutit à une contradiction.

Exercice

Montrer que le réel $\sqrt{2}$ est irrationnel (i.e. n'est pas rationnel)

Correction

Si $\sqrt{2} = \frac{p}{q}$, fraction irréductible, alors $2q^2 = p^2$, donc 2 divise p^2, puis p qui est donc pair.

Puis $p = 2p'$ et donc $2q^2 = 4p'^2$, donc $q^2 = 2p'^2$, puis 2 divise q^2, puis q qui est donc pair.

Et ainsi la fraction $\frac{p}{q}$ n'est pas irréductible...

Remarque - Différence avec la contraposée

Pour la contraposée, on suppose non$B$ et on aboutit à non$A$.

Pour l'absurde, on suppose non$B$ et A ensemble et on montre qu'il y a une contradiction. Dans ce second cas, on exploite plus d'hypothèses!

4.4. Conditions nécessaire, suffisante

Il s'agit simplement *d'un peu de bon sens* sur la signification des mots « nécessaire » et « suffisant ».

Exemple - A : « x est le carré d'un entier » et B : « $x \geqslant 0$ »

Considérons un réel x et posons

A : x est le carré d'un entier

B : $x \geqslant 0$

Qu'est-ce qui est nécessaire, qu'est-ce qui est suffisant (ou ne l'est pas)?

Définition - Condition nécessaire. Condition suffisante

Plus généralement $A \Rightarrow B$ peut se dire :

- B est une condition nécessaire (**CN**) pour avoir A (puisque si on A, nécessairement on a B)
- A est une condition suffisante (**CS**) pour avoir B (puisqu'il suffit d'avoir A pour avoir B

Rechercher une condition nécessaire et suffisante (**CNS**) pour avoir A revient donc à chercher B tel que $A \Leftrightarrow B$.

Savoir faire - Analyse-Synthèse

Certaines démonstrations, difficiles à gérer par équivalences, ou lorsque le résultat n'est pas donné, se font en deux phases :

1. phase d' "analyse" : on obtient une condition nécessaire (par implications successives par exemple) pour qu'une première hypothèse soit vérifiée
2. phase de "synthèse", ou phase de vérification : la condition précédemment obtenue est-elle suffisante?

Exercice

Montrer que toute fonction définie sur $\mathbb{R}$ à valeurs dans $\mathbb{R}$ s'écrit de manière unique comme somme d'une fonction paire et d'une fonction impaire.

On procédera de la manière suivante :

Première étape (analyse) : supposons qu'il existe deux fonctions f et g telles que ..., alors $f =$, $g =$

Deuxième étape (synthèse) : on vérifie que les solutions trouvées à la première étape conviennent

Correction

ANALYSE : Si $h = f + g$ avec f paire et g impaire,

alors $\forall\, x \in \mathbb{R}$, $h(-x) = f(-x) + g(-x) = f(x) - g(x)$ et $h(x) = f(x) + g(x)$.

En additionnant et retranchant : $f(x)=\frac{1}{2}\big(h(x)+h(-x)\big)$ et $g(x)=\frac{1}{2}\big(h(x)-h(-x)\big)$.

L'analyse a aboutit à une unique décomposition (sous reserve d'existence...).

SYNTHESE : Soit h une fonction de $\mathbb{R}$ et associons lui :

$f:x\mapsto\frac{1}{2}\big(h(x)+h(-x)\big)$ et $g:x\mapsto\frac{1}{2}\big(h(x)-h(-x)\big)$

alors par construction $h=f+g$,

mais aussi $f(-x)=\frac{1}{2}\big(h(-x)+h(x)\big)=\frac{1}{2}\big(h(x)+h(-x)\big)=f(x)$, donc f est paire

également $g(-x)=\frac{1}{2}\big(h(-x)-h(x)\big)=-\frac{1}{2}\big(h(x)-h(-x)\big)=g(x)$, donc g est impaire

Exercice

Soit A,B,C et D quatre points du plan tel que $AC=BD$ et (AB) non parallèle à (CD).
Alors il existe une unique rotation du plan r tel que $r(A)=B$ et $r(C)=D$.
Donner ses caractérisatiques

Correction

Nous allons raisonner par analyse-synthèse.

ANALYSE. Supposons que $r(A)=B$ et $r(C)=D$.

Notons Ω le centre de r et θ son angle de rotation.

Comme $r(A)=B$, on a donc $|\Omega A|=|\Omega B|$ et $(\overrightarrow{\Omega A},\overrightarrow{\Omega B})=\theta$.

Ainsi, Ω est sur la médiatrice de $[AB]$. De même Ω est sur la médiatrice de $[CD]$.

Ces deux médiatrices se coupent en un seul point, ssi elles ne sont pas parallèles.

Si ces médiatrices sont parallèles, alors (AB) et (CD) sont parallèles,

la contraposée donne donc :

si (AB) et (CD) ne sont pas parallèles, alors les deux médiatrices se coupent en un seul point.

Ainsi Ω est obtenu de manière unique et $\theta:=(\overrightarrow{\Omega A},\overrightarrow{\Omega B})$, bien défini.

SYNTHESE. Considérons la rotation, centrée Ω concours des médiatrices et d'angle $\theta:=(\overrightarrow{\Omega A},\overrightarrow{\Omega B})$.

On a par définition $r(A)=B$.

Comme $\Omega C=\Omega D$, $\Omega A=\Omega B$ et $AC=BD$, les triangles ΩAC et ΩBD sont semblables.

Donc $(\overrightarrow{\Omega A},\overrightarrow{\Omega C})=(\overrightarrow{\Omega B},\overrightarrow{\Omega D})$.

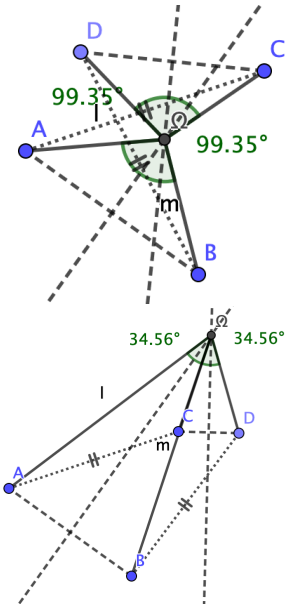
Ainsi, avec la relation de Chasles (des angles) :

$$\theta=(\overrightarrow{\Omega A},\overrightarrow{\Omega B})=(\overrightarrow{\Omega A},\overrightarrow{\Omega C})+(\overrightarrow{\Omega C},\overrightarrow{\Omega B})=(\overrightarrow{\Omega B},\overrightarrow{\Omega D})+(\overrightarrow{\Omega C},\overrightarrow{\Omega B})=(\overrightarrow{\Omega C},\overrightarrow{\Omega D})$$

Donc $r(C)=D$. BILAN : Il existe une unique rotation tel que $r(A)=B$ et $r(C)=D$, c'est la rotation dont le centre est Ω, le point de concours des médiatrices de $[AB]$ et de $[CD]$ et d'angle $\theta=(\overrightarrow{\Omega A},\overrightarrow{\Omega B})$.

On appliquera ce résultat dans le cours sur les nombres complexes.

Représentation - Deux configurations de l'exercice



4.5. Exploiter un contre-exemple dans une démonstration

Savoir faire - Utilisation d'un contre-exemple

Lorsque l'on veut prouver qu'une implication est fausse, on cherche un exemple vérifiant l'hypothèse mais pas la conclusion, ce que l'on appelle un *contre-exemple*.

Exercice

Soit f la fonction définie sur $\mathbb{R}$ par $f(x)=|x+\frac{3}{2}|-\frac{1}{2}$. Montrer que $(x\geqslant -1)$ et $(f(x)\geqslant 0)$ ne sont pas équivalents.

Correction

$-10\leqslant -1$ et $f(-10)=\frac{7}{2}-\frac{1}{2}=3\geqslant 0$

4.6. Démonstration par récurrence

Proposition - Principe admis (axiome)

Soit $P(n)$ (parfois notée $\mathscr{P}_n$) une propriété portant sur l'entier n.

Si on a

$$\begin{cases} P(0) \text{ vraie} \\ \forall n\in\mathbb{N},(P(n) \text{ vraie } \Rightarrow P(n+1) \text{ vraie }) \end{cases}$$

alors $P(n)$ est vraie pour tout entier n.

Savoir faire - Rédaction d'un raisonnement par récurrence

• Pour $n = 0$, $P(0)$ est vérifiée, avec vérification effective! (souvent deux simples calculs)

• Supposons la propriété vérifiée pour **un certain** $n \geqslant 0$ (et surtout pas "pour tout", parce que là, ce n'est plus la peine de faire une démonstration!).

$\Rightarrow$ Montrons que $P(n+1)$ est vraie.

Conclusion : $\forall n \in \mathbb{N}$, $P(n)$ est vraie.

Remarque - Démarrer à un autre nombre

On peut aussi démarrer à une valeur de n autre que 0.

Exercice

Montrer que $\forall n \in \mathbb{N}$, $n < 2^n$.

Correction

Remarquons d'abord que pour tout entier n, $2^n \geqslant 1$ Notons pour tout entier $n \in \mathbb{N}$, $\mathscr{P}_n$: « $n < 2^n$ ».

— $0 < 1 = 2^0$, donc $\mathscr{P}_0$ est vraie.

— Soit $n \in \mathbb{N}$, supposons que $\mathscr{P}_n$ est vérifiée.
Donc $n < 2^n$ et donc $n+1 < 2^n + 1 \leqslant 2^n + 2^n = 2^{n+1}$
Donc $\mathscr{P}_{n+1}$ est alors vraie.

Le résultat est donc bien démontré par récurrence et on peut affirmer : $\forall\, n \in \mathbb{N}$, $n < 2^n$.

Exercice

Y a-t-il des erreurs dans les raisonnements suivants ? Où sont-elles ?

1. On a : $10^{n+1} + 1 = 10 \times 10^n + 1 = (9+1) \times 10^n + 1 = 9 \times 10^n + 10^n + 1$
Donc, si $10^n + 1$ est divisible par 9, il en est de même de $10^{n+1} + 1$, ce qui prouve que pour tout entier naturel n, $10^n + 1$ est divisible par 9.
2. Prouvons que tout ensemble fini a tous ses éléments égaux :
Si dans tout ensemble E_n à n éléments, tous les éléments sont égaux, alors, soit E_{n+1} un ensemble à $n+1$ éléments :
$$E_{n+1} = \{x_1, x_2, \dots, x_n, x_{n+1}\}.$$
Avec l'ensemble de n éléments $\{x_1, x_2, \dots, x_n\}$, on a, par hypothèse de récurrence : $x_1 = x_2 = \cdots = x_n$.
Avec l'ensemble de n éléments $\{x_2, x_3, \dots, x_{n+1}\}$, on a, par hypothèse de récurrence : $x_2 = x_3 = \cdots = x_{n+1}$.
Donc $x_1 = x_2 = \cdots = x_{n+1}$.
Comme la propriété est vraie pour $n = 1$ (cas d'un singleton), il en résulte que tout ensemble de n éléments a tous ses éléments égaux.

Correction

Dans le premier raisonnement, on montre bien le moteur de la récurrence : $\mathscr{P}(n) \Rightarrow \mathscr{P}(n+1)$.
Mais l'initialisation n'est pas démontrée ; et pour cause : elle est fausse.
Le résultat (pour tout $n \in \mathbb{N}$) n'est donc pas démontré (et d'ailleurs il est toujours faux).
Dans le second raisonnement, c'est plus subtile. Dans le passage $\mathscr{P}(n) \Rightarrow \mathscr{P}(n+1)$,
il faut nécessairement que E_{n+1} possède au moins 3 éléments, et donc $n \geqslant 2$.
Ainsi, l'initiation ne sert à rien ici, il faut commencer par démontrer $\mathscr{P}(2)$.

Histoire - Citation de Henri Poincaré, La science et l'hypothèse, 1902

« Le caractère du raisonnement par récurrence est qu'il contient, condensés, pour ainsi dire en une formule unique, une infinité de syllogismes.
Pour qu'on s'en puisse mieux rendre compte, je vais énoncer les uns après les autres ces syllogismes qui sont, si l'on veut me passer l'expression, disposés en cascade.
Ce sont bien entendu des syllogismes hypothétiques.
Le théorème est vrai du nombre 1.
Or si il est vrai de 1, il est vrai de 2.
Donc il est vrai de 2.
Or si il est vrai de 2, il est vrai de 3.
Donc il est vrai de 3, et ainsi de suite... »

Savoir faire - Récurrence à plusieurs pas (ou plusieurs termes)

Suivant la façon dont est énoncée la propriété de récurrence $P(n)$ il peut être nécessaire

• d'initialiser la récurrence avec plusieurs (k) valeurs de n

• de supposer $P(n), \dots, P(n+k-1)$ (il y en a aussi k) vraies pour un certain $n \geqslant 0$

• de prouver que ces k propriétés exactes entraînent $P(n+k)$ vraie

Savoir faire - Récurrence forte

• Pour $n = 0$, $P(0)$ est vérifiée (avec vérification effective!)

• Supposons la propriété vérifiée **jusqu'à** un certain $n \geqslant 0$ (i.e $P(0), P(1), \dots, P(n)$ vraies)

• Montrons que $P(n+1)$ est vraie.

On parle parfois alors de « récurrence à plusieurs pas » ou de « récurrence forte », par opposition à la récurrence du théorème dite « récurrence simple (ou faible) »).

Exercice

Soit (u_n) la suite définie par $u_0 = \frac{2}{5}, u_1 = 1$ et pour tout entier naturel n,

$$u_{n+2} = 5u_{n+1} - 6u_n.$$

Démontrer que pour tout $n \in \mathbb{N}$, $u_n = \frac{2^n + 3^n}{5}$.

Correction

Nous verrons par la suite d'autre méthode que la récurrence.
En attendant, nous devons employer ici une récurrence à deux pas.
Notons pour tout entier $n \in \mathbb{N}$, $\mathscr{P}_n$: « $u_n = \frac{2^n + 3^n}{5}$ ».

- $\frac{2^0 + 3^0}{5} = \frac{2}{5} = u_0$, donc $\mathscr{P}_0$ est vraie.
- $\frac{2^1 + 3^1}{5} = \frac{5}{5} = u_1$, donc $\mathscr{P}_1$ est vraie.
- Soit $n \in \mathbb{N}$, supposons que $\mathscr{P}_n$ et $\mathscr{P}_{n+1}$ sont vérifiées.
  Donc $u_{n+2} = 5\frac{2^{n+1} + 3^{n+1}}{5} - 6\frac{2^n + 3^n}{5} = \frac{2^n(10-6) + 3^n(15-6)}{5} = \frac{2^{n+2} + 3^{n+2}}{5}$ Donc $\mathscr{P}_{n+2}$ est alors vraie.

Le résultat est donc bien démontré par récurrence et on peut affirmer : $\forall\, n \in \mathbb{N}$, $u_n = \frac{2^n + 3^n}{5}$.

Exercice

Montrer que tout entier $n \geqslant 2$ se décompose en produit de nombres premiers.

Correction

On fait une récurrence forte.
Notons pour tout entier $n \geqslant 2$, $\mathscr{P}_n$: « n se décompose en produit de nombres premiers ».

- 2 est un nombre premier donc $\mathscr{P}_2$ est vraie.
- Soit $n \in \mathbb{N}$, $n \geqslant 2$ supposons que pour tout entier $k \leqslant n$, $\mathscr{P}_k$ est vérifiée.
  Si $n+1$ est un nombre premier, alors il est le produit de nombres premiers.
  Si $n+1$ n'est pas premiers, il existe $a, b \geqslant n$ tel que $n+1 = ab$.
  Or $\mathscr{P}_a$ et $\mathscr{P}_b$ sont vraies, donc on peut décomposer ces deux nombres, et par produit $n+1$ est également le produit de nombres premiers.
  Donc $\mathscr{P}_{n+1}$ est alors vraie.

Le résultat est donc bien démontré par récurrence (forte) et on peut affirmer que tout entier $n \geqslant 2$ se décompose en produit de nombres premiers.

Exercice

Montrer qu'un changement d'hypothèse de récurrence ramène une récurrence à plusieurs pas ou une récurrence forte à une récurrence faible.

Correction

Dans le cas d'une récurrence à h pas, on peut considérer $\mathscr{Q}_n$: « $\mathscr{P}_n$ et $\mathscr{P}_{n+1}$ et … $\mathscr{P}_{n+h}$ ».
Dans le cas d'une récurrence forte, on peut considérer $\mathscr{Q}_n$: « $\forall\, k \geqslant n$ $\mathscr{P}_k$ ».

Exercice

Montrer par récurrence forte que toute suite décroissante d'entiers est stationnaire (i.e. constante à partir d'un certain rang).

Correction

On note, pour tout entier $k \in \mathbb{N}$, $\mathscr{P}_k$: « si $(u_n) \in \mathbb{N}^{\mathbb{N}}$ avec $u_0 = k$ et (u_n) décroissante, alors (u_n) est stationnaire. »

- Si $u_0 = 0$, alors nécessairement, il s'agit de la suite stationnaire égal à 0, car $0 = u_0 \geqslant \underbrace{u_n}_{\in\mathbb{N}} \geqslant 0$.
  Donc $\mathscr{P}_0$ est vraie.
- Soit $k \in \mathbb{N}$. Supposons que pour tout $h \geqslant k$, $\mathscr{P}_h$ est vraie.
  Soit (u_n) suite d'entiers, décroissante telle que $u_0 = k+1$.
  • Ou bien pour tout $n \in \mathbb{N}$, $u_n = k+1$ et donc (u_n) est stationnaire.
  • Ou bien il existe $n_0 \in N$ tel que $u_{n_0} \neq k+1$, par décroissance, $u_{n_0} < k+1$.
  Considérons la suite (extraite) (v_n) telle que $v_n = u_{n+n_0}$.
  Alors (v_n) est une suite d'entiers, décroissante de premier terme $u_{n_0} \leqslant k$.
  On applique $\mathscr{P}_{u_{n_0}}$ à (v_n), qui est donc stationnaire à partir du rang N.
  Alors (u_n) est stationnaire à partir du rang $N + n_0$. Donc $\mathscr{P}_{k+1}$ est vraie.

4.7. Démonstration par algorithme

Algorithme

Heuristique - Exploitation d'un algorithme

Un algorithme peut permettre de démontrer, constructivement, l'existence d'un certain objet (ou d'une certaine fonction).
La difficulté est plutôt de démontrer que l'algorithme :

- se termine bien
- réalise bien ce que l'on désire

Définition - Algorithme

Un algorithme est une suite finie de règles à appliquer dans un ordre déterminé à un nombre fini de données pour arriver avec certitude (c'est-à-dire sans indétermination ou sans ambiguïté), en un nombre fini d'étapes, à un certain résultat et cela indépendamment des données.

Terminaison de l'algorithme

Pour démontrer que l'algorithme termine (que le nombre d'étapes est fini), on exploite un variant de boucle, en règle général, en suite d'entiers décroissante. Pour des boucles `for`, la terminaison de boucle est en règle générale immédiate.

Savoir faire - Démontrer qu'une boucle se termine

On identifie une expression (variable) qui :

- est entière
- décroît strictement à chaque étape de la boucle

Alors, nécessairement, la boucle se termine.

Exemple - Division euclidienne

On considère le bout de programme suivant :

```
q=0
r=n
while r>=d :
        q=q+1
        r=r-d
```

Ce programme calcul le quotient `q` et le reste `r` de la division euclidienne de `n` par `d`. (A démontrer)
Mais est-ce qu'il termine bien ? Oui
On constate que la suite des valeurs prises par la variable `r` est **strictement décroissante** de `n` à un nombre compris entier positif plus petit que `d`.
Si la boucle ne s'arrêtait pas, alors cette suite de valeur serait infini, ce qui est impossible.

Exercice

On considère le bout de programme suivant :

```
c=0
while p>0 :
        if c==0 :
                p=p-2
                c=1
        else :
                p=p+1
                c=0
```

1. Que fait ce programme ?
2. Les variables `p` et `c` sont-elles décroissantes, strictement ?
3. Montrer que la variable `t=2p+3c` est entière, strictement décroissante.
   En déduire la terminaison de l'algorithme.

Correction

1. On commence par faire le tableau pour comprendre, avec les premiers étapes ce qui peut se passer :

temps	p	c
0	p	0
1	p−2	1
2	p−2+1=p−1	0
3	p−3	1
4	p−2	0
5	p−4	1

   Le programme donne donc pour finir `p=0` (condition d'arrêt) et `c=0`, si `p` est pair initialement, et `c=1` sinon.
2. Les variables `p` et `c` ne sont pas décroissantes.
3. Il y a deux évolutions possibles pour `t=2p+3c`, selon la valeur de `c`.
   — si `c=0`, alors `t=2p+3c→2(p−2)+3(c+1)=2p+3c−1`
   — si `c=1`, alors `t=2p+3c→2(p+1)+3(c−1)=2p+3c−1`
   Donc la variable `t=2p+3c` est entière, strictement décroissante.
   L'algorithme se termine donc bien.

Correction de l'algorithme

Il faut aussi savoir **démontrer** que le programme (avec de nombreuses répétitions de la boucle) réalise se que l'on souhaite.
On utilise alors pour cela des *invariant de boucles*.
Comme son nom l'indique il s'agit d'identifier (créer) une expression-variable qui ne change pas de valeur tout au long du programme.
Puis lorsque le programme se termine, en exploitant cette expression, nous pourrons démontrer que l'on obtient bien le résultat attendu.

Savoir faire - Utilisation d'un invariant de boucle pour démontrer le résultat attendu

Pour démontrer qu'une boucle réalise bien le résultat attendu,

1. on cherche une expression qui reste constante tout au long des calculs de la boucle ;
2. on calcule sa valeur initiale (avant le début de la boucle) ;
3. on en déduit sa valeur finale ;
4. enfin connaissant les valeurs finales des variables du système (testées pour la sortie de boucle), on en déduit la valeur de la variable retournée par le programme.

Exemple - Retour sur la division euclidienne

Reprenons l'algorithme de la division euclidienne par soustraction successive :

```
q=0
r=n
while r>=d :
        q=q+1
        r=r-d
```

— La valeur qui n'évolue pas est `t=dq+r`.
   En effet, on passe à chaque boucle de `t=dq+r→d(q+1)+(r−d)=dq+d+r−d=dq+r`
— De plus initialement, `t=d×0+n=n`
— Et pour finir `r` appartient à l'intervalle $[0, d[$.

On a donc trouver le couple `(q,r)` tel que `n=dq+r` avec $r \in [0, d[$.
C'est la définition de la division euclidienne. Et le résultat obtenu est bien celui attendu.

Exercice

Montrer que le programme

```
1 def somme2(n) :
2     S=0
```

```
    for i in range(1,n+1):
        S=S+i**2
    return(S)
```

calcule bien $\sum_{i=1}^{100} i^2$

Correction

Notons $T=S-\sum_{k=1}^{i-1} k^2$.

Alors à chaque étape, on a $T=S-\sum_{k=1}^{i-1} k^2 \longrightarrow S+i^2-\sum_{k=1}^{i} k^2=S-\sum_{k=1}^{i-1} k^2$ Nous avons trouvé notre invariant : T.

Initialement, $T=0$, à la fin aussi et donc $S=\sum_{k=1}^{n} k^2$

Exercice

On cherche à écrire un programme qui calcul $n!$.

1. Ecrire un programme avec une boucle `while`.
2. **Démontrer** que le programme se termine bien.
3. **Démontrer** que le programme effectue bien ce que l'on souhaite.

Correction

1.
```
def factorielle2(n):
    """Calcul de la factorielle de n"""
    f=1
    k=1
    while k<n :
        f,k=f*(k+1),k+1
    return(f)
```

2. Il s'agit de trouver une variable entière qui décroit strictement.
   Ici, c'est clair, il faut prendre `k`.
   Elle commence à `n`, et décroit strictement à chaque étape. Donc à partir d'un certain moment (ici `n` étapes, puisqu'elle diminue de 1 à chaque étape), elle est nulle ; et la boucle s'arrête bien.
3. Notons $I=f*(k!)$.
   A l'instant initial, $k=n$, $f=1$ et il a pour valeur $I=1\times n!=n!$.
   A chaque instant, on a $I=f*(k!)\rightarrow(f*k)((k-1)!)=I$. Donc I est invariant.
   Enfin, en fin de course, pour la dernière boucle, $k=1$, et donc $I=f*1!=n!$. Ce qui implique que $f=n!$.
   Le programme renvoie la valeur de f, donc c'est bien la valeur de $n!$.

Proposition - Plus petit élément d'un ensemble fini
Considérons E un ensemble muni d'une relation d'ordre totale.
Un ensemble A de n éléments de E admet un plus petit élément.

Nous allons faire la démonstration par algorithme

Démonstration
Soit A un ensemble fini, on peut supposer que $A=\{x_1,x_2,\dots x_n\}$.
Considérons l'algorithme :
```
a=A[1]
   for k in range(n):
      if A[k]<a :
         a=A[k]
   return(a)
```
L'algorithme termine car il s'agit d'une boucle `for`
Il faut montrer la correction en trouvant un invariant de boucle. On va considérer pour le tour `k` la proposition $\mathscr{P}_k$:« $\forall\ i\leqslant k,\ a\leqslant x_i,\ a\in A$ ». Au premier tour, comme $A[1]=x_1=a$, $\mathscr{P}_1$ est vraie.
Si $\mathscr{P}_k$ est vraie, il en est de même de $\mathscr{P}_{k+1}$, selon que $x_{k+1}<a$ ou $x_{k+1}\geqslant a$.
En bout de course, on a donc $a\leqslant x_i$, pour tout $i\leqslant n$ et $a\in A$.
□

5. Bilan

Synthèse

- ⇝ En mathématiques, les raisonnements se fondent sur une vision ensembliste des objets ou des propositions. Nous faisons un premier passage, *de bon sens*, sur ce qu'est un ensemble et ce que signifie appartenir à un ensemble ou en être une partie ; ce qu'est un intervalle ou un produit cartésien d'ensembles.
- ⇝ Naturellement, apparait fréquemment dans les affirmations mathématiques ensemblistes deux notions : une notion d'universalité *pour tout* et une notion d'exception *il existe*. La formalisation qui est le langage écrit des mathématiques, réserve donc deux symboles pour ces notions : $\forall$ et $\exists$. On les retrouve tout le temps.
- ⇝ Les mathématiques donnent des relations (de vérité ?) entre les assertions. Nous voyons différentes méthodes exploitées dans *l'artisanat de la démonstration* : table de vérité (cas par cas), implication ou équivalence, analyse-synthèse, contraposée, raisonnement par l'absurde, contre-exemple ou récurrences…

Savoir-faire et Truc & **Astuce du chapitre**

- — Savoir-faire - Montrer que I est un intervalle de $\mathbb{R}$
- — Savoir-faire - Noter les dépendances
- — Savoir-faire - Montrer que $F \subset E$
- — Savoir-faire - Prouver l'égalité de deux ensembles
- — Savoir-faire - $A \Rightarrow B$. Raisonnement direct.
- — Savoir-faire - $A \Rightarrow B$. Raisonnement par contraposée.
- — Savoir-faire - $A \Leftrightarrow B$. On procède en deux temps.
- — Savoir-faire - $A \Leftrightarrow B$. On procède par équivalences connues successives.
- — Truc & Astuce pour le calcul - Ne pas abuser de $\Leftrightarrow$
- — Savoir-faire - Raisonnement par l'absurde
- — Savoir-faire - Analyse-Synthèse
- — Savoir-faire - Utilisation d'un contre-exemple
- — Savoir-faire - Rédaction d'un raisonnement par récurrence
- — Savoir-faire - Récurrence à plusieurs pas (ou plusieurs termes)
- — Savoir-faire - Récurrence forte
- — Savoir-faire - Démontrer qu'une boucle termine
- — Savoir-faire - Utilisation d'un invariant de boucle pour démontrer le résultat attendu

Notations

Notations	*Définitions*	*Propriétés*	*Remarques*
$\forall$ - $\exists$	Pour tout (ou quel que soit) - Il existe	Les affirmations mathématiques exploitent très souvent uniquement ces deux symboles	Attention, on $\forall a \exists b \Leftrightarrow \exists b$
$A \Longrightarrow B \equiv B \Longleftarrow A$	Implication de A vers B	A est suffisante pour B et B est nécessaire à A	Ex : $\exists b \forall a \Rightarrow$
$A \Longleftrightarrow B$	A et B sont équivalentes	Identique à $A \Rightarrow B \& B \Rightarrow A$	Ne pas en ab

Retour sur les problèmes

47. Quoi dire…
48. Ce n'est probablement pas qu'un langage, mais…
49. Ce n'est pas un problème

Chapitre 11

Applications (entre ensembles)

Résumé -

Nous complétons quelques notions essentielles du fondement des mathématiques (formalisé non sans mal à la fin du XIX-ième siècle). Ces fondements se basent sur les ensembles et sur les applications entre ces ensembles!
Quelles sont les applications qui ne transforment pas trop les ensembles?
Les ensembles images ou réciproques (retour) permet de décrire (par des ensembles) les qualités de l'application.
Une application classique est l'application qui compte les éléments. On précisera ici les notions intuitives de cardinaux des ensembles ici.
Nous terminerons pas étudier les familles Quelques vidéos :

- *Khan Academy - Intersection et union d'ensembles - https ://www.youtube.com/watch ?v=vcwMTpgNlQA*
- *Canal unisciel (P. Dehornoy) - La théorie des ensembles 50 ans après Cohen - https ://www.canal-u.tv/video/institut_fourier/patrick_dehornoy _la_theorie_des_ensembles_cinquante_ans_apres_cohen.41917*

Sommaire

1. Problèmes

? Problème 51 - Qualités des fonctions

Résoudre une équation, c'est trouver x (tous les x) tel que $f(x) = b$, où b et f sont connus.
Il est intéressant de savoir si :

- l'équation a (au moins) une solution.
- l'équation a exactement une solution.
- l'équation a (au plus) une solution.

Evidemment, la réponse dépend de b et de f, on peut la noter $f^{-1}(\{b\})$, c'est un ensemble de solution (qui peut être vide)!.
Si on reprend ces trois options, qu'on généralise à tout b, on trouve 3 qualités précises de f. Comment peut-on qualifier ces trois propriétés?

? Problème 52 - Description d'ensemble simple

Peut-on comparer deux ensembles facilement. Pour être un tant soit peu identique, ils doivent au moins avoir le même nombre d'éléments.
Comment fait-on pour savoir cela? Quelle est la nature du lien qui unit l'un à l'autre? Existe-t-il des ensembles de référence à k éléments?
Comment calculer formellement le nombre d'éléments d'un sous-ensemble?

? Problème 53 - Cardinal fini. Cardinal infini

Deux ensembles sont de taille identique s'il existe une application bijective de l'un sur l'autre.
Même l'existence d'une application bijective d'un ensemble à un autre peut très bien se produire même si les ensembles ne sont pas de taille fini.
Existe-t-il des ensembles infinis de même taille? Des ensembles infinis de tailles différentes? Existe-t-il une relation d'ordre (totale?) entre les ensembles de taille infini?

? Problème 54 - Nombres rationnels

Quelles sont les caractéristiques de l'application : $\varphi : \mathbb{Z} \times \mathbb{N}^* \to \mathbb{Q}$, $(a,b) \mapsto \frac{a}{b}$?

? Problème 55 - Famille $(O_i)_{i\in I}$

Lorsqu'une application dépend d'un nombre réel, on note $f : x \mapsto \ldots$ cette application.
Lorsqu'elle dépend d'un nombre entier naturel, on la note $(u_n)_{n\in\mathbb{N}}$.
Existe-t-il une notation officielle pour une application qui dépend d'un ensemble I de points. Et comment on appelle cette application : fonction, suite, autre chose?

2. Applications de E dans F

Il s'agit ici de donner une théorie plus précise sur les fonctions.

2.1. Vocabulaire lié aux applications

Heuristique - Application (définition non formelle)

Une application d'un ensemble E dans un ensemble F est un "procédé" qui associe à chaque élément $x \in E$ un élément $f(x) \in F$. Une telle application est notée

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

$f(x)$ est appelé image de x par f ;
l'ensemble E est appelé ensemble de départ de l'application f ;
l'ensemble F est appelé ensemble d'arrivée de f.
On a donc une application de E dans F dès qu'à tout élément $x \in E$ on peut associer sans ambiguïté un élément $f(x) \in F$ (c'est-à-dire s'il y en a un et un seul possible).

Une application est donc déterminée par la donnée des couples $(x, f(x))$ où x parcourt E, d'où la définition plus formelle :

Définition - Application

Soient E et F deux ensembles et $\mathcal{G} \subset E \times F$ vérifiant

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

La donnée d'un tel triplet $(E,F,\mathcal{G})$ s'appelle une application de E dans F.
On note

$$\begin{array}{rl} f: & E \to F \\ & x \mapsto y = f(x) \end{array}$$

où y est l'unique élément de F vérifiant $(x,y) \in \mathcal{G}$.
$\mathcal{G}$ s'appelle le graphe de l'application. On le note souvent Γ_f.
On a donc $\Gamma_f = \{(x,y) \in E \times F \,|\, y = f(x)\} = \{(x, f(x)) \,|\, x \in E\}$.

Remarque - Fonctions ou applications ?

D'après la définition, une fonction définie sur E, à valeurs dans F, est une application de E dans F.
De cette définition découle le résultat suivant :

Savoir faire - Montrer une égalité de deux fonctions

Soient $f : E \to F$ et $g : E' \to F'$ deux applications. f et g sont égales si et seulement si :
- $E = E'$ (même ensemble de départ),
- $F = F'$ (même ensemble d'arrivée),
- $\forall x \in E,\ f(x) = g(x)$.

Définition - Ensemble de fonctions

On notera $\mathcal{F}(E,F)$ (on trouve aussi les notations $\mathcal{A}(E,F)$ ou F^E) l'ensemble des applications (ou fonctions) de E dans F.

Exemple - Classiques

- Pour tout ensemble E, l'application $x \mapsto x$ de E dans lui-même est appelée application identité de E et notée Id_E, son graphe est la diagonale de E^2.
- Soient E et $F \neq \emptyset$ deux ensembles, et $a \in F$. L'application $x \mapsto a$ de E dans F s'appelle une application constante, son graphe est $E \times \{a\}$.

— Soient $E_1, \dots, E_n$ des ensembles. Pour chaque $i \in \{1, \dots, n\}$ l'application

$$\begin{array}{rccc} p_i : & E_1 \times \cdots \times E_n & \to & E_i \\ & (x_1, \cdots, x_n) & \mapsto & x_i \end{array}$$

s'appelle la i-ième projection ou la i-ième application coordonnée.

Définition - Restriction et prolongement
Soit $f : E \to F$ une application.
— Soit $A \subset E$. **La** restriction de f à A, notée $f_{|A}$, est l'application

$$\begin{array}{rccc} f_{|A} : & A & \to & F \\ & x & \mapsto & f(x) \end{array}$$

— Si $E \subset B$, une application $\tilde{f} : B \to F$ est **un** prolongement à B de l'application f si $\tilde{f}_{|E} = f$, c'est-à-dire si $\forall x \in E, \tilde{f}(x) = f(x)$.

Définition - Composée
Soient deux applications $f : E \to F, g : F \to G$. On définit l'application composée, notée $h = g \circ f$, de E dans G par

$$\forall x \in E, h(x) = g(f(x))$$

Exemple - Identité
Si on a une application $f : E \to F$, alors $Id_F \circ f = f$ et $f \circ Id_E = f$.

Remarque - Répresentation
Il est parfois utile de représenter les applications par un graphe.

Attention - Non commutativité
En général, même lorsque les deux applications $g \circ f$ et $f \circ g$ ont un sens, elles sont différentes.

Proposition - Associativité de $\circ$
Pour trois applications $E \xrightarrow{f} F \xrightarrow{g} G \xrightarrow{h} H$ on a

$$h \circ (g \circ f) = (h \circ g) \circ f.$$

On peut donc noter $h \circ g \circ f$.

Démonstration
$h \circ (g \circ f) = (h \circ g) \circ f$ sont toutes deux des applications de F dans H.
Soit $x \in F$,

$$[h \circ (g \circ f)](x) = h\big(g(f(x)\big) = (h \circ g)(f(x)) = [(h \circ g) \circ f](x)$$

□

2.2. Bijections (injections et surjections)

Applications injectives ou surjectives

Heuristique - Résoudre une équation

Une fonction est de la forme : $E \xrightarrow{f} F$. A tout x de E, f donne une valeur de F.
Résoudre une équation est toujours le problème inverse :
Sont données : $E \xrightarrow{f} F$ et $b \in F$. Il s'agit de trouver $x \in E$ tel que $f(x) = b$.
Les questions naturelles sont les suivantes :
— Cette équation admet-elle (au moins) une solution?
— Cette équation admet-elle au plus une solution?

— Cette équation admet-elle exactement une solution ? Ce qui évite les quiproquos…

Définition - Injection et surjection
Soit $f : E \to F$ une application. On dit que
— f est injective (est une injection) si $\forall (x,x') \in E^2, f(x) = f(x') \Rightarrow x = x'$;
— f est surjective (est une surjection) si $\forall y \in F, \exists x \in E \mid y = f(x)$.

Pour aller plus loin - Exemples
Donner E, F, f et b pour les équations :
— $y' + y = x$
— $x^2 + 3x - 4 = 2$
— $\begin{cases} 2x & +y & = & 1 \\ x & -y & = & 2 \end{cases}$

Remarque - Notation
On pourra noter $f : E \hookrightarrow F$ pour exprimer que f est une application injective de E sur F.
On pourra noter $f : E \twoheadrightarrow F$ pour exprimer que f est une application surjective de E sur F.
Comment écrire une flèche de bijection ? $E \hookrightarrow\!\!\!\!\twoheadrightarrow F$

Exercice
Montrer que f est injective ssi $\forall\ b \in F$, $f(x) = b$ admet au plus une solution.
Montrer que f est surjective ssi $\forall\ b \in F$, $f(x) = b$ admet toujours (au moins) une solution.

Correction
Supposons f injective.
Si $f(x) = b$ admet deux solutions x_1 et x_2, alors $f(x_1) = f(x_2)$, impossible si f injective.
Si $\forall\ b \in F$, $f(x) = b$ admet au plus une solution, alors $f(x) = f(x')(= b)$ donc $x = x^{p}rime$
La deuxième équivalence correspond exactement à la définition de la surjection.

Savoir faire - Autres formulations équivalentes (injectivité, surjectivité)
Il y a différentes façons équivalentes de formuler ces propriétés :
— Dire que f est injective revient à dire (par contraposée) que :

$$\forall (x,x') \in E^2,\ x \neq x' \Rightarrow f(x) \neq f(x')$$

c'est-à-dire que deux éléments distincts de l'ensemble de départ ont des images distinctes.
— Dire que f est surjective revient à dire que tout élément de l'ensemble d'arrivée possède au moins un antécédent.

Exemple - $x \mapsto x^2$, $x \mapsto x^3$, $x \mapsto \sin x$
Les applications de $\mathbb{R}$ dans $\mathbb{R}$ suivantes sont-elles injectives ? surjectives ?

$$f_1 : x \mapsto x^2, \quad f_2 : x \mapsto x^3, \quad f_3 : x \mapsto \sin x$$

f_1 n'est ni injective ni surjective de $\mathbb{R}$ sur $\mathbb{R}$, elle est surjective de $\mathbb{R}$ sur $\mathbb{R}_+$, puis même injective de $\mathbb{R}_+$ sur $\mathbb{R}_+$ ou de $\mathbb{R}_-$ sur $\mathbb{R}_+$.
f_2 est injective et surjective de $\mathbb{R}$ sur $\mathbb{R}$.
f_3 est ni injective ni surjective de $\mathbb{R}$ sur $\mathbb{R}$, elle est surjective de $\mathbb{R}$ sur $[-1,1]$, puis même injective de $[\frac{\pi}{2} + k\pi, \frac{\pi}{2} + (k+1)\pi]$ sur $[-1,1]$.

Exercice
L'application

$$\begin{array}{rccc} f : & \mathbb{R}^2 & \to & \mathbb{R}^2 \\ & (x,y) & \mapsto & (x-y, 2x+y) \end{array}$$

est-elle injective ? surjective ?
Mêmes questions avec

$$\begin{array}{rccc} f : & \mathbb{R}^2 & \to & \mathbb{R}^3 \\ & (x,y) & \mapsto & (x-y, 2x+y, x-3y) \end{array}$$

Correction
$(x-y, 2x+y) = (a,b) \iff (x,y) = (\frac{a+b}{3}, \frac{b-2a}{3})$.
Donc f est injective et surjective.
$(x-y, 2x+y, x-3y) = (a,b,c) \iff (x,y) = (\frac{a+b}{3}, \frac{b-2a}{3})$ avec $c = \frac{7a-2b}{3}$.
Donc f est injective mais pas surjective (si $c \neq \frac{7a-2b}{3}$, pas de solution…).

Nous ferons plus tard une étude plus complète de l'étude des systèmes linéaires.

Exercice

$$\begin{array}{rl} f: & \mathbb{C} \to \mathbb{C}^* \\ & z \mapsto \exp z \end{array}$$

est-elle injective ? surjective ?

Correction

Elle n'est pas injective : $1 = e^0 = e^{2i\pi}$.

En revanche, exp est bien surjective de $\mathbb{C}$ sur $\mathbb{C}^*$. En effet, si $z = \rho e^{i\theta} \neq 0$,

$\exp\left(\ln(\rho) + i\theta\right) = \exp^{\ln\rho} e^{i\theta} = z$.

Remarque - Les ensembles qui sont importants!

On remarquera que :

— c'est l'ensemble de départ joue un rôle important pour l'injectivité,

— c'est l'ensemble d'arrivée joue un rôle important dans la surjectivité.

Heuristique - Stratégies

Soit $f : E \to F$. Il est pratique que f soit bijective, mais cela ne nous appartient pas en règle générale.

Toutefois, il est possible de rendre :

— f surjective en restreignant « simplement » l'ensemble d'arrivée à $f(E)$.

— f injective en restreignant l'ensemble de départ, ou plus fréquemment en considérant non plus $f : E \text{ø} F$, mais $\hat{f} : \dfrac{E}{\mathscr{R}_f} \to F, \overline{x} \mapsto f(x)$.

Exercice

Montrer que pour cette dernière stratégie, la fonction $\hat{f}$ est bien définie et qu'elle est injective

---

Correction

Si $\overline{x} = \overline{y}$, alors $f(x) = f(y)$ et donc il n'y a pas de problème pour écrire $\hat{f}(\overline{x})$.

Elle est injective, car si $f(\overline{x}) = f(\overline{x'})$, alors $f(x) = f(x')$ et donc $x \mathscr{R}_f x'$ et donc $\overline{x} = \overline{x'}$.

Théorème - Propriétés des composées

Soient $f : E \to F$ et $g : F \to G$ deux applications.

Si f et g sont injectives (respectivement surjectives), alors $g \circ f$ est injective (resp. surjective).

Démonstration

Supposons que f et g sont injectives.

Soient $x, x' \in G$ tel que $g \circ f(x) = g \circ f(x')$.

Par injectivité de g : $f(x) = f(x')$

Par injectivité de f : $x = x'$

Donc $g \circ f$ est injective.

Supposons que f et g sont surjectives.

Soit $y \in G$,

par surjectivité de g, il existe $u \in F$ tel que $g(u) = y$.

par surjectivité de f, il existe $x \in E$ tel que $f(x) = u$.

Donc $g \circ f(x) = g(f(x)) = g(u) = y$

Donc $g \circ f$ est surjective.

□

Exercice

Soient $f : E \to F$ et $g : F \to G$ deux applications. Montrer que :

— si $g \circ f$ est injective, alors f est injective ;

— si $g \circ f$ est surjective, alors g est surjective.

Correction

Supposons que $g \circ f$ est injective.

Soient $x, x' \in E$ tel que $f(x) = f(x')$.

alors $g(f(x)) = g(f(x'))$, donc $x = x'$ car $g \circ f$ injective.

Par conséquent f est injective.

Supposons que $g \circ f$ est surjective.

Soit $y \in G$. Alors il existe $x \in E$ tel que $g \circ f(x) = y$.

donc avec $X = f(x)$, on a $g(X) = y$.

Par conséquent f est surjective.

Applications bijectives

Pour que l'equation $f(x) = b$ admette une unique solution, quel que soit $b \in F$, il faut (et il suffit) que f soit bijective :

Définition - Application bijective
Soit $f : E \to F$ une application. On dit que f est bijective (ou est une bijection) de E sur F si si f est injective et surjective.
Dire que f est bijective revient à dire que :

$$\forall\ y \in F, \exists\,! x \in E \quad \text{tel que} \quad y = f(x)$$

c'est-à-dire que tout élément de l'ensemble d'arrivée possède un et un seul antécédent.

Définition - Application réciproque
Soit f une bijection de E sur F, on définit une application g par

$$\begin{array}{rccl} g : & F \to & E & \\ & y \mapsto & x & | \, y = f(x) \text{ (unique antécédent de } y \text{ par } f) \end{array}$$

Cette application g est elle-même bijective et appelée bijection réciproque de f, et notée f^{-1}.

Démonstration
Il faut montrer que g ainsi définie est bien bijective.
Soit $x \in E$, alors prenons $y = f(x)$, on a donc $g(y) = x$. Donc g est surjective.
Si $g(y) = g(y') = x$, alors cela signifie que $y = f(x) = y'$. Donc g est injective. □

Savoir faire - Critère pour montrer la bijectivité
Soient $f : E \to F$ et $g : F \to E$ deux applications telles que $g \circ f = Id_E$ et $f \circ g = Id_F$.
Alors f et g sont bijectives et $g = f^{-1}$

Exercice
Soit

$$f : \begin{array}{c} \mathbb{N} \to \mathbb{N} \\ n \mapsto n+1 \end{array} \quad \text{et} \quad g : \begin{array}{rl} \mathbb{N} \to & \mathbb{N} \\ n \mapsto & \begin{cases} 0 \text{ si } n = 0 \\ n-1 \text{ si } n \neq 0 \end{cases} \end{array}$$

Etudier l'injectivité et la surjectivité des applications f et g.
Déterminer les applications $g \circ f$ et $f \circ g$. Conclusion ?

Correction
$g(0) = g(1) = 0$, donc g n'est pas injective (donc pas bijective).
0 n'a pas d'antécédent par f, donc f n'est pas surjective (ni bijective).
Pour tout $n \in \mathbb{N}$, $g \circ f(n) = n$ en revanche $f \circ g(n) = n$ pour $n \neq 0$, mis $f \circ g(0) = f(0) = 1$.
Conclusion : il faut bien les deux conditions $f \circ g = Id_F$ ET $g \circ f = Id_E$ pour affirmer la bijectivité.

Proposition - Relation entre f et f^{-1}
Si f est une bijection de E sur F, on a

$$\forall x \in E, (f^{-1} \circ f)(x) = x \qquad \text{soit } f^{-1} \circ f = Id_E;$$
$$\forall y \in F, (f \circ f^{-1})(y) = y \qquad \text{soit } f \circ f^{-1} = Id_F;$$
$$y = f(x) \iff x = f^{-1}(y);$$
$$(f^{-1})^{-1} = f.$$

Démonstration
Soit $x \in E$, avec $y = f(x)$, on $f^{-1}(f(x)) = f^{-1}(y) = x$ car $y = f(x)$. Donc $f^{-1} \circ f = Id_E$.
Soit $y \in F$, notons $x = f^{-1}(y)$, donc $y = f(x)$ et ainsi : $f(f^{-1}(y) = f(x) = y$. Donc $f \circ f^{-1} = Id_F$.
L'équivalence proposée découle de la définition.
$(f^{-1})^{-1} : E \to F, z \mapsto y$ tel que $f^{-1}(y) = z$, i.e. $y = f(z)$. Ainsi $\forall\, z \in E$, $(f^{-1})^{-1}(z) = f(z)$. □

Exercice
Soit

$$\begin{array}{rccc} f: & \mathbb{C}\setminus\{i\} & \to & \mathbb{C} \\ & z & \mapsto & \dfrac{z+2i}{z-i} \end{array}$$

Montrer que l'on peut trouver $F \subset \mathbb{C}$ tel que l'application $\hat{f}$ de $\mathbb{C}\setminus\{i\}$ dans F définie par $\hat{f}(z) = f(z)$ soit une bijection. Déterminer sa bijection réciproque.

Correction
$f(z) = Z \iff \dfrac{z+2i}{z-i} = Z \iff z+2i = Z(z-i) \iff z(1-Z) = -2i - Zi \iff z = \dfrac{(Z+2)i}{Z-1}$.
Donc si $Z \neq 1$, $f(z) = Z \implies h(Z) = z$, avec $h : Z \mapsto \frac{(Z+2)i}{Z-1}$.
Ainsi $\hat{f} : \mathbb{C}\setminus\{i\} \to \mathbb{C}\setminus\{1\}, z \mapsto f(z)$ est bijective et admet une application réciproque : $\hat{h}$.
Il est possible de donner une interprétation géométrique à ce calcul.

Théorème - Bijection réciproque d'une composée
Si $f : E \to F$ et $g : F \to G$ sont deux bijections, alors $g \circ f$ est une bijection et

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

Démonstration
$(f^{-1} \circ g^{-1}) \circ (g \circ f) = f^{-1} \circ g^{-1} \circ g \circ f = f^{-1} \circ f = Id_E$.
$(g \circ f) \circ (f^{-1} \circ g^{-1}) = g \circ f \circ f^{-1} \circ g^{-1} = g \circ g^{-1} = Id_F$.
On a ainsi et la bijectivité de $g \circ f$ et la valeur de $(g \circ f)^{-1}$. □

3. Image directe et image réciproque d'un ensemble

Heuristique - Si les applications ne sont pas bijectives
Si $f : E \to F$ n'est pas bijective, peut-on néanmoins trouver « comme » une application réciproque.
On a vu que tout n'est pas perdu, à condition de limiter l'ensemble de départ (pour tenter de gagner l'injectivité) et de réduire l'ensemble d'arrivée (pour gagner la surjectivité).
Dans le premier cas, on s'intéressera à l'ensemble réciproque de F par f, c'est un sous-ensemble de E.
Dans le second cas, on s'intéressera à l'ensemble image de E par f, c'est un sous-ensemble de F.

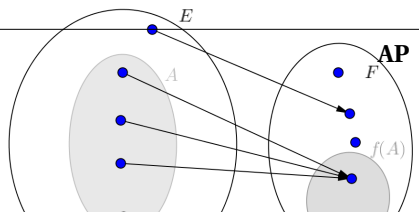
3.1. Image directe

Définition - Ensemble image
Soit $f : E \to F$ une application.
L'ensemble des éléments de F qui admettent un antécédent par f est une partie de F appelée ensemble image ou image de f et notée Im f :

$$\text{Im } f = \{y \in F \mid \exists x \in E; y = f(x)\}.$$

Remarque - Autre écriture
On peut écrire Im $f = \{f(x) \mid x \in E\}$

Représentation - Image directe
Avec un graphe :



Pour aller plu
Comme le mont
stabilité pour les
$\dfrac{az+b}{cz+d}$ *(réciproq*
Il s'agit d'une fa
très étudiée : en
rapport!
Notons que dans
sur $\overline{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$.

Savoir faire - Critère de surjectivité.

On a donc

$$f \text{ surjective } \Leftrightarrow \text{Im } f = F$$

Et toute application $f : E \to F$ définit une surjection en restreignant l'ensemble d'arrivée, c'est à dire en considérant l'application de E dans Im f qui a x associe $f(x)$ (que l'on continue usuellement à noter f, on dit alors que f est surjective de E sur Im f).

Définition - Image directe

Soient $f : E \to F$ une application et $A \subset E$. On appelle image (directe) de A par f la partie de F, notée $f(A)$, définie par

$$f(A) = \{y \in F \mid \exists x \in A; y = f(x)\} = \{f(x) \mid x \in A\} = \{f(x)\,;\, x \in A\}.$$

Remarque - Autres notations en exploitant Im f

On constate que l'on a également Im $f = f(E)$.

Donc f est surjective si et seulement si $f(E) = F$.

Et $f(A) = \text{Im } f_{|A}$.

Savoir faire - Montrer que $y \in f(A)$

Pour montrer que $y \in f(A)$ il faut trouver $x \in A$ tel que $f(x) = y$.

Proposition - Stabilité et image

Soit $f : E \to F$ une application et $A_1, A_2 \subset E$. Alors

$$A_1 \subset A_2 \Rightarrow f(A_1) \subset f(A_2);$$
$$f(A_1 \cup A_2) = f(A_1) \cup f(A_2);$$
$$f(A_1 \cap A_2) \subset f(A_1) \cap f(A_2).$$

Démonstration

Supposons que $A_1 \subset A_2$,

Soit $y \in f(A_1)$, alors il existe $x \in A_1$ tel que $y = f(x)$.

Mais $x \in A_2$ également, donc $y = f(x) \in f(A_2)$.

Donc $f(A_1) \subset f(A_2)$.

Soit $y \in f(A_1 \cup A_2)$ donc il existe $x \in A_1 \cup A_2$ tel que $y = f(x)$.

donc il existe $x \in A_1$ tel que $y = f(x)$ ou $x \in A_2$ tel que $y = f(x)$

donc $y \in f(A_1)$ ou $y \in f(A_2)$ i.e. $y \in f(A_1) \cup f(A_2)$.

Réciproquement, soit $y \in f(A_1) \cup f(A_2)$.

donc $y \in f(A_1)$ ou $y \in f(A_2)$

donc il existe $x_1 \in A_1$ tel que $y = f(x_1)$ ou $x_2 \in A_2$ tel que $y = f(x_2)$

donc il existe $x_1 \in A_1 \cup A_2$ tel que $y = f(x_1)$ ou $x_2 \in A_2 \cup A_1$ tel que $y = f(x_2)$

dans tous les cas $y \in f(A_1 \cup A_2)$.

Soit $y \in f(A_1 \cap A_2)$ donc il existe $x \in A_1 \cap A_2$ tel que $y = f(x)$.

donc il existe $x \in A_1$ tel que $y = f(x)$ et $x \in A_2$ tel que $y = f(x)$

donc $y \in f(A_1)$ et $y \in f(A_2)$ i.e. $y \in f(A_1) \cap f(A_2)$. □

Attention - Une seule inclusion pour l'intersection!

On fera bien attention qu'il n'y a pas l'égalité : $f(A_1 \cap A_2) = f(A_1) \cap f(A_2)$

Pour se convaincre on peut penser au cas où $A_1 \cap A_2 = \emptyset$.

Par exemple, avec $f : x \mapsto x^2$, $A_1 = [-2,-1]$ et $A_2 = [1,2]$,

alors $f(A_1 \cap A_2) = f(\emptyset) = \emptyset$, et $f(A_1) \cap f(A_2) = [1,4] \cap [1,4] = [1,4]$

Pour aller plus loin - Suite de la forme $u_{n+1} = f(u_n)$

Pour les suites définies par récurrence par une fonction f itérée, ces notions de partie stable (souvent intervalle) ou de points fixes sont très importants. Nous reviendrons sur ces notions

Définition - Partie stable et application induite
Soit $f : E \to E$ une application. On dit qu'une partie A de E est stable par f si $f(A) \subset A$.
On appelle application induite par f sur A l'application

$$\begin{aligned} A &\to A \\ x &\mapsto f(x) \end{aligned}$$

Exemple - Point fixe
Par exemple pour $x \in E$, le singleton $\{x\}$ est stable si et seulement si $f(x) = x$, c'est-à-dire si x est un *point fixe* de f.

3.2. Image réciproque d'un ensemble

Définition - Image réciproque
Soient $f : E \to F$ une application et $B \subset F$. On appelle image réciproque de B par f la partie de E, notée $f^{-1}(B)$ (ou $[f \in B]$ en probabilité), définie par

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

C'est donc l'ensemble formé des antécédents par f des éléments de B.

Exemple - Pour $x \mapsto x^2$ et $x \mapsto \exp x$
Par exemple pour

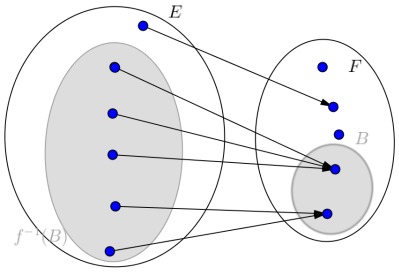
$$\begin{aligned} f : \quad \mathbb{R} &\to \mathbb{R} \\ x &\mapsto x^2 \end{aligned}$$

on a $f^{-1}(\{0,4\}) = \{-2,0,2\}$, $f^{-1}(\{4\}) = \{-2,2\}$, $f^{-1}([-1,4[) = [-2,2]$ et pour

$$\begin{aligned} g : \quad \mathbb{C} &\to \mathbb{C} \\ z &\mapsto \exp z \end{aligned}$$

on a $g^{-1}(\{0\}) = \emptyset$, $g^{-1}(\{1\}) = \{2ik\pi \mid k \in \mathbb{Z}\}$, $g^{-1}(\mathbb{R}) = \{L + 2ik\pi \mid L \in \mathbb{R}, k \in \mathbb{Z}\}$

Représentation - Image réciproque
Avec un graphe :



Attention - Ne pas confondre la fonction f^{-1} et l'ensemble $f^{-1}(B)$
Il s'agit d'une notation qui ne demande pas que f soit bijective. (voir l'exemple précédent)

Savoir faire - Montrer que $x \in f^{-1}(B)$
Pour montrer que $x \in f^{-1}(B)$ il faut montrer que $f(x) \in B$.

Exemple - Fonction sin
Soit

$$\begin{aligned} f : \quad \mathbb{R} &\to \mathbb{R} \\ x &\mapsto \sin x \end{aligned}$$

Déterminer, si cela a un sens : $f(0)$; $f(\{0\})$; $f([0,\pi[)$; $f(\mathbb{R})$; $f^{-1}(0)$; $f^{-1}(\{0\})$; $f^{-1}([0,+\infty[)$.
$f(0) = 0$; $f(\{0\}) = \{0\}$; $f([0,\pi[) = [0,1]$; $f(\mathbb{R}) = [-1,1]$; $f^{-1}(0)$ pas de sens ;
$f^{-1}(\{0\}) = \pi\mathbb{Z}$; $f^{-1}([0,+\infty[) = \bigcup_{k\in\mathbb{Z}} [2k\pi, (2k+1)\pi]$.

Remarque - Et si f est bijective
Soit $f : E \to F$ un bijection. Pour $B \subset F$ la notation $f^{-1}(B)$ n'est pas ambigüe.
En effet, ici f^{-1} existe, et $f^{-1}(B)$ est l'image de B par f^{-1}. Donc

$$\begin{aligned} f^{-1}(B) = \operatorname{Im} f^{-1}_{|B} &= \{f^{-1}(y), y \in B\} = \{x \in E \mid \exists\, y \in B, x = f^{-1}(y)\} \\ &= \{x \in E \mid f(x)(= y) \in B\} = f^{-1}(B) \end{aligned}$$

Proposition - Stabilité et image réciproque
Soit $f : E \to F$ une application et $B_1, B_2 \subset F$. Alors

$$B_1 \subset B_2 \Rightarrow f^{-1}(B_1) \subset f^{-1}(B_2);$$
$$f^{-1}(B_1 \cap B_2) = f^{-1}(B_1) \cap f^{-1}(B_2);$$
$$f^{-1}(B_1 \cup B_2) = f^{-1}(B_1) \cup f^{-1}(B_2).$$

Démonstration
Supposons que $B_1 \subset B_2$,
Soit $x \in f^{-1}(B_1)$, alors $f(x) \in B_1 \subset B_2$ Donc $x \in f^{-1}(B_2)$
Donc $f^{-1}(B_1) \subset f^{-1}(B_2)$.
On peut reconduire le même genre de démonstration que plus haut. Ou autre :

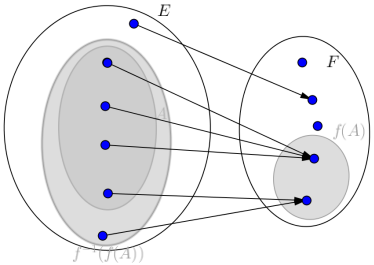
$$f^{-1}(B_1 \cap B_2) = \{x \in E \mid f(x) \in B_1 \cap B_2\} = \{x \in E \mid f(x) \in B_1 \text{ et } f(x) \in B_2\}$$
$$= \{x \in E \mid x \in f^{-1}(B_1) \text{ et } x \in f^{-1}(B_2)\} = f^{-1}(B_1) \cap f^{-1}(B_2)$$

(On a directement l'égalité sans faire la double inclusion).
De même :

$$f^{-1}(B_1 \cup B_2) = \{x \in E \mid f(x) \in B_1 \cup B_2\} = \{x \in E \mid f(x) \in B_1 \text{ ou } f(x) \in B_2\}$$
$$= \{x \in E \mid x \in f^{-1}(B_1) \text{ ou } x \in f^{-1}(B_2)\} = f^{-1}(B_1) \cup f^{-1}(B_2)$$

□

Représentation - Image directe et réciproque
Avec un graphe :



Attention - Attention $f^{-1}(f(A)) \neq A$ et $f(f^{-1})(B) \neq B$
Le schéma montre sur un exemple qu'on a pas l'égalité...
On a au mieux : $A \subset f^{-1}(f(A))$ et $f(f^{-1}(B)) \subset B$

Exercice
Démontrer ces inclusions. Donner des contre-exemple de l'inclusion réciproque

Correction
Soit $x \in A$, alors $f(x) \in f(A)$ et donc $x \in f^{-1}(f(A))$.
De même si $y \in f(f^{-1}(B))$, alors il existe $x \in f^{-1}(B)$ tel que $y = f(x)$.
Mais comme $x \in f^{-1}(B)$, cela signifie que $f(x) \in B$, donc $y = f(x) \in B$.
Prenons $f : x \mapsto x^2$, $A = [0,2]$, $f(A) = [0,4]$ et $f^{-1}(f(A)) = [-2,2] \sup A$.
C'est un problème d'injectivité De même avec $B = [-1,2]$, $f^{-1}(B) = [-\sqrt{2},\sqrt{2}]$, $f(f^{-1}(B)) = [0,2] \subset B$ *C'est un problème de surjectivité*

4. Fonction indicatrice

4.1. Définition

Définition - Fonction indicatrice
Soit E un ensemble. Pour $A \subset E$, on appelle fonction indicatrice de A l'application de E dans $\mathbb{R}$, notée $\mathbb{1}_A$ ou χ_A, définie par

$$\forall x \in E, \mathbb{1}_A(x) = \begin{cases} 1 \text{ si } x \in A \\ 0 \text{ si } x \notin A \end{cases}$$

Remarque - Pourquoi et comment exploiter une telle fonction ?
Cela permet de représenter certaines fonctions définies par morceaux par une seule expression (très utile en probabilités) sur $\mathbb{R}$, par exemple la loi uniforme sur $[a,b]$ a pour densité $\frac{1}{b-a}\mathbb{1}_{[a,b]}$ et la loi exponentielle de paramètre λ a pour densité la fonction définie sur $\mathbb{R}$ par $t \mapsto \lambda e^{-\lambda t}\mathbb{1}_{[0,+\infty[}(t)$.
Cela permet également de ramener certaines égalités d'ensemble à des calculs sur les fonctions.

4.2. Propriétés ensemblistes et calcul avec fonctions indicatrices

Proposition - Propriété essentielle de la fonction indicatrice
Soient A et B deux parties de E. Alors

$\mathbb{1}_A \leqslant \mathbb{1}_B \Leftrightarrow A \subset B$ $\quad$ $\mathbb{1}_A = \mathbb{1}_B \Leftrightarrow A = B$ (d'où le nom de fonction caractéristique);
$\mathbb{1}_{\complement_E A} = 1 - \mathbb{1}_A$;
$\mathbb{1}_{A\cap B} = \mathbb{1}_A \times \mathbb{1}_B$;
$\mathbb{1}_{A\cup B} = \mathbb{1}_A + \mathbb{1}_B - \mathbb{1}_A \times \mathbb{1}_B$.

Exercice
Soit E un ensemble. Pour deux parties A et B de E, on appelle différence symétrique de ces deux parties la partie de E définie par

$$A\triangle B = (A\cup B)\setminus(A\cap B).$$

Comment exprimer la fonction indicatrice de $A\triangle B$ à l'aide des fonctions indicatrices de A et B ?
En déduire que pour trois parties A, B, C de E, on a $(A\triangle B)\triangle C = A\triangle(B\triangle C)$.

Correction
$\mathbb{1}_{A\triangle B} = \mathbb{1}_{A\cup B} - \mathbb{1}_{A\cap B} = \mathbb{1}_A + \mathbb{1}_B - 2\mathbb{1}_A\mathbb{1}_B$ On a donc $A\triangle B = B\triangle A$ (ce qui est assez évident).
On a également :

$$\begin{aligned}\mathbb{1}_{(A\triangle B)\triangle C} &= \mathbb{1}_{A\triangle B} + \mathbb{1}_C - 2\mathbb{1}_{A\triangle B}\mathbb{1}_C\\ &= \mathbb{1}_A + \mathbb{1}_B + \mathbb{1}_C - 2[\mathbb{1}_A\mathbb{1}_B + \mathbb{1}_A\mathbb{1}_C\mathbb{1}_B\mathbb{1}_C]\\ &\quad +4\mathbb{1}_A\mathbb{1}_B\mathbb{1}_C\end{aligned}$$

Ce résultat est symétrique en A, b et C, donc : $(A\triangle B)\triangle C = (B\triangle C)\triangle A = A\triangle(B\triangle C)$.

Démonstration
Si $A \leqslant B$. Soit $x \in E$ Supposons que $\mathbb{1}_A(x) = 1$, alors $x \in A$ donc $x \in B$ donc $\mathbb{1}_B(x) = 1$.
Donc, comme les indicatrices sont à valeurs dans $\{0,1\}$, on a $\mathbb{1}_A \leqslant \mathbb{1}_B$.
Réciproquement, si $\mathbb{1}_A \leqslant \mathbb{1}_B$. Soit $x \in A$, alors $\mathbb{1}_B(x) \geqslant \mathbb{1}_A(x) = 1$, donc $x \in B$.
La double inclusion signifie la double inégalité. Donc $A = B$ ssi $\mathbb{1}_A = \mathbb{1}_B$.
$x \in \complement_E(A) \leftrightarrow x \notin A \leftrightarrow \mathbb{1}_A(x) = 0 \leftrightarrow 1 - \mathbb{1}_A(x) = 1$.
Quatrième proposition :

$$\mathbb{1}_{A\cap B}(x) = 1 \Longleftrightarrow x \in A\cap B \Longleftrightarrow x \in A \text{ et } x \in B$$

$$\Longleftrightarrow \mathbb{1}_A(x) = 1 \text{ et } \mathbb{1}_B(x) = 1 \Longleftrightarrow 1 = \mathbb{1}_A(x)\times\mathbb{1}_B(x) = \big(\mathbb{1}_A\times\mathbb{1}_B\big)(x)$$

car pour $a, b \in \{0,1\}$, $a \times b = 1$ si et seulement si $a = b = 1$ Comme il n'y a que deux valeurs, on peut affirmer $\mathbb{1}_{A\cap B} = \mathbb{1}_A \times \mathbb{1}_B$.
Cinquième proposition :
On peut aussi faire un autre type de démonstration, avec comme une table de vérité

$x \in ?$	$\mathbb{1}_{A\cup B}(x)$	$\mathbb{1}_A(x)$	$\mathbb{1}_B(x)$	$\boxed{\mathbb{1}_A + \mathbb{1}_B - \mathbb{1}_A\times\mathbb{1}_B}(x)$
$x\in A, x\in B$	1	1	1	$1+1-1=1$
$x\in A, x\notin B$	1	1	0	$1+0-0=1$
$x\notin A, x\in B$	1	0	1	$0+1-0=1$
$x\notin A, x\notin B$	0	0	0	$0+0-0=0$

□

On retrouve évidemment des résultats comparables à ceux vus en logique…

5. Cardinal d'ensemble fini

5.1. Principe des tiroirs

Nous considérons ici que l'ensemble $\mathbb{N}$ est bien connu. Nous expliquerons au chapitre suivant sa construction, en attendant, il nous faut savoir que c'est l'ensemble d'appui du raisonnement par récurrence….
On rappelle que l'on note $\mathbb{N}_k = \{1, 2, 3, \dots k-1, k\}$, l'ensemble des k premiers entiers naturels non nuls.
On commence par deux lemmes.

Lemme - Injection de $\mathbb{N}_n$ sur $\mathbb{N}_p$. Principe des tiroirs (DIRICHLET)
Soient $n, p \in \mathbb{N}^*$.
S'il existe une fonction $\varphi : \mathbb{N}_n \to \mathbb{N}_p$ injective, alors $n \leqslant p$.
Sous sa forme contraposée : si $\varphi : \mathbb{N}_n \to \mathbb{N}_p$ avec $n > p$,
Alors il existe $x \neq x\prime \in \mathbb{N}_n$ tel que $\varphi(x) = \varphi(x')$ (φ non injective).

Démonstration
On peut le démontrer par récurrence sur n.
Posons, pour tout entier $n \geqslant 1$, $\mathscr{P}_n$: « $\forall\ p \geqslant 1$, ($\exists\ \varphi : \mathbb{N}_n \to \mathbb{N}_p$ injective)$\Rightarrow n \leqslant p$ ».
— Le cas $n = 1$ est simple, car nécessairement $p \geqslant 1$.
Donc $\mathscr{P}_1$ est vraie.
— Soit $n \in \mathbb{N}^*$. On suppose que $\mathscr{P}_n$ est vraie.
Soit $p \in \mathbb{N}^*$. On suppose qu'il existe $\varphi : \mathbb{N}_{n+1} \to \mathbb{N}_p$ injective.
On note $r = \varphi(n+1) \in \mathbb{N}_p$.
On considère $\psi : \mathbb{N}_p \to \mathbb{N}_p$, $k \mapsto \begin{cases} p & \text{si } k = r \\ r & \text{si } k = p \\ k & \text{sinon} \end{cases}$.
En fait, ψ intervertit p et r. ψ est une bijection de $\mathbb{N}_p$ sur $\mathbb{N}_p$.
Donc, par composition, $\psi \circ \varphi : \mathbb{N}_{n+1} \to \mathbb{N}_p$, injective et $(\psi \circ \varphi)(n+1) = \psi(r) = p$.
Notons $\Psi = (\psi \circ \varphi)_{|\mathbb{N}_n}$, alors Ψ est également une injection.
Et $\Psi : \mathbb{N}_n \to \mathbb{N}_{p+1}$.
On peut appliquer $\mathscr{P}_n$ avec Ψ. Et donc $n \leqslant p-1$.
Donc $\mathscr{P}_{n+1}$ est vérifiée.
□

5.2. Classe des ensembles de même cardinal

Lemme - $\mathscr{R}$ comme relation d'équivalence
On note $\mathscr{R}$, la relation entre ensembles définies par :
$$E \mathscr{R} F \iff \exists\ \varphi : E \to F, \text{bijective}$$
$\mathscr{R}$ est une relation d'équivalence.

Démonstration
On vérifie les trois qualités d'une relation d'équivalence :
— Pour tout ensemble E, l'application $\varphi : E \to E, x \mapsto x$ est bijective de E sur E.
Donc pour tout E, $E \mathscr{R} E$ i.e. $\mathscr{R}$ est reflexive.
— Soient deux ensembles E et F tels que $E \mathscr{R} F$.
Alors il existe $\varphi : E \to F$ bijective de E sur F.
Elle admet une réciproque $\varphi^{-1} : F \to E$ bijective.
Donc pour tout E, F, $E \mathscr{R} F \Rightarrow F \mathscr{R} E$ i.e. $\mathscr{R}$ est symétrique.
— Soient trois ensembles E, F et G tels que $E \mathscr{R} F$ et $F \mathscr{R} G$.
Alors il existe $\varphi_1 : E \to F$ et $\varphi_2 : F \to G$ bijectives.
$\phi = \varphi_2 \circ \varphi_1$ est une bijection de E sur G.
Donc pour tout E, F, G, $E \mathscr{R} F$ et $F \mathscr{R} G \Rightarrow E \mathscr{R} G$ i.e. $\mathscr{R}$ est transitive.
□

Définition - Ensemble de cardinal n. Ensemble fini.
Soit $n \in \mathbb{N}$.
On dit qu'un ensemble E est fini de cardinal $n (\in \mathbb{N})$, si $E \mathscr{R} \mathbb{N}_n$. On note $\text{Card}(E) = n$
On dit qu'un ensemble E est fini, si il existe $n \in \mathbb{N}$ tel que E est fini de cardinal n.

Exemple - Cardinal de $E = \{1, 2, \dots k\}$
L'identité est bijective de E sur $\mathbb{N}_k$, donc $\text{card}(E) = k$.
Exercice
Montrer que si $n, p \in \mathbb{N}$ et $n < p$, alors on n'a pas $\mathbb{N}_n \mathscr{R} \mathbb{N}_p$.

Correction
On démontre le résultat contraposé.
Soient $n, p \in \mathbb{N}$. On suppose que $\mathbb{N}_n \mathscr{R} \mathbb{N}_p$.

Alors il existe $\varphi : \mathbb{N}_n \to \mathbb{N}_p$, bijective donc injective.
Et d'après le résultat admis : $n \leqslant p$.
Et par symétrie de la relation $\mathscr{R}$, on a de même : $p \leqslant n$.
Ainsi $n = p$. Donc par contraposée :

$$n < p \Longrightarrow n \neq p \Longrightarrow \text{NON } [\mathbb{N}_n \mathscr{R} \mathbb{N}_p]$$

Proposition - Classe d'équivalence pour $\mathscr{R}$
Sur l'ensemble des ensembles $E, F, \ldots$ de cardinaux finis, on a l'équivalence : $E \mathscr{R} F \Longleftrightarrow \text{Card}(E) = \text{Card}(F)$.
Les classes d'équivalence pour $\mathscr{R}$ sont formé des ensembles de même cardinaux.
On peut les *paramétrer* par leur cardinaux

Pour aller plus loin - Construction de $\mathbb{N}$
Un mode de construction de $\mathbb{N}$ *qui a fait grand bruit au début du XX*^{e} *siècle consistait à dire que l'ensemble des nombres entiers était en fait l'ensemble des cardinaux possibles pour les ensembles finis.*
Qu'est-ce que 3*, un représentant de tous les ensembles à 3 éléments. Si on veut...*

5.3. Cardinal, fonction indicatrice et somme (finie)

Proposition - Calcul avec l'indicatrice
Soit E, un ensemble fini et A un sous-ensemble de E.
Alors le calcul $\sum_{x \in E} \mathbb{1}_A(x)$ a un sens et il vaut $\text{Card}(A)$.

Démonstration
Comme E est un ensemble fini, il existe $k \in \mathbb{N}$ et $\varphi : E \to \mathbb{N}_k$ bijective.
On peut donc écrire que $E = \{\varphi^{-1}(1), \varphi^{-1}(2) \ldots \varphi^{-1}(k)\}$ que l'on préfère noter $\{x_1, x_2, \ldots x_k\}$.
On a alors la somme $\sum_{x \in E} \mathbb{1}_A(x)$ qui se calcule de la façon suivante : $\sum_{i=1}^{k} \mathbb{1}_A(x_i)$.
On note alors $\varphi : A \to \mathbb{N}$, $x_h (\in A \subset E) \mapsto \sum_{i=1}^{h} \mathbf{1}_A(x_i)$. Alors

— φ est injective :
si $x_h \neq x_\ell$, alors on peut supposer $h < \ell$ (SPDG) et donc $\varphi(x_\ell) = \varphi(x_h) + \sum_{i=h+1}^{\ell} \mathbf{1}_A(x_i) \geqslant \varphi(x_h) + \mathbf{1}_A(x_\ell) \geqslant \varphi(x_h) + 1$, donc $\varphi(x_h) \neq \varphi(x_\ell)$.

— Par ailleurs, par construction $\varphi(A)$ est de la forme $\mathbb{N}_s$ avec $s = \sum_{i=1}^{k} \mathbf{1}_A(k)$:
$\varphi(x_{m+1}) - \varphi(x_m) = \mathbf{1}_A(x_{m+1}) \in \{1, 0\}$. Il ne peut y avoir de trou.

La fonction φ établit donc une bijection de A sur $\mathbb{N}_s$. Nécessairement $s = \text{Card}(A) = \sum_{i=1}^{k} \mathbf{1}_A(k)$.
□

Corollaire - Inclusion et cardinaux
Si $A \subset B (\in E)$ avec E un ensemble fini, alors $\text{Card} A \leqslant \text{Card} B$.

Démonstration
On a vu $A \subset B$ implique $\mathbb{1}_A \subset \mathbb{1}_B$.
On somme en tout x de E. □

Exercice
Soient E, un ensemble fini de cardinal n et F, un ensemble fini de cardinal m.

1. On suppose que $f : E \to F$ est une application injective.
   (a) Montrer que $f(E)$ est un ensemble fini de cardinal n.
   (b) Montrer que $n \leqslant m$.
2. Que se passe-t-il si f est surjective ?
3. Démontrer le théorème de Cantor-Bernstein pour des ensembles E et F finis.

$$\exists\, i : E \to F, j : F \to E \text{ injectives} \Longrightarrow \exists\, b : E \to F \text{ bijective}$$

Correction

1. (a) $E \mathscr{R} \mathbb{N}_n$. Et on note $\hat{f} : E \to f(E), x \mapsto f(x)$.
   Par définition de $f(E)$: pour tout $y \in f(E)$, il existe $x \in E$ tel que $y = f(x) = \hat{f}(x)$.
   Donc $\hat{f}$ est surjective de E sur $f(E)$.
   Et si $\hat{f}(x) = \hat{f}(x')$, alors $f(x) = f(x')$ et donc $x = x'$, car f injective.
   Donc $\hat{f}$ est injective.
   Par conséquent $\hat{f}$ est bijective de E sur $f(E)$ et donc $E \mathscr{R} f(E)$.
   Par transitivité, $f(E) \mathscr{R} \mathbb{N}_n$, $f(E)$ est un ensemble fini de cardinal n.
   (b) On note $\varphi : E \to \mathbb{N}_n$, bijective (elle existe bien car $E \mathscr{R} \mathbb{N}_n$).
   Alors φ^{-1} est injective. Il en est de même de $f \circ \varphi^{-1}$ (f est injective).
   Puis on note $\psi : F \to \mathbb{N}_m$ bijective (elle existe bien car $F \mathscr{R} \mathbb{N}_m$).
   ψ est injective et par composition :
   $$\psi \circ f \circ \varphi^{-1} : \mathbb{N}_n \to \mathbb{N}_m$$
   est également injective. Donc $n \leqslant m$.
2. Si f est surjective : $f(E) = F$, et donc $\text{Card}F = \text{Card}f(E)$.
   Or $\text{Card}(f(E)) \leqslant \text{Card}E$, donc $\text{Card}F \leqslant \text{Card}E$.
3. Soient E et F deux ensembles finis.
   On suppose qu'il existe $i : E \to F$ injective. Donc d'après 2., $\text{Card}(E) \leqslant \text{Card}(F)$.
   On suppose qu'il existe $j : F \to E$ injective. Donc d'après 2., $\text{Card}(F) \leqslant \text{Card}(E)$.
   Donc, par double inégalité : $\text{Card}(E) = \text{Card}(F)$.
   D'après la question 1.(c), cela signifie que $E \mathscr{R} F$. Et par définition, cela signifie qu'il existe une bijection de E sur F.
   $\exists\, i : E \to F, j : F \to E$ injectives $\Longrightarrow \exists\, b : E \to F$ bijective

L'exercice donne le résultat suivant (dans le cas injectif)

Proposition - Cardinaux, injectivité et surjectivité
Si E et F sont des ensembles finis.
S'il existe une fonction $f : E \to F$ injective, alors $\text{card}E \leqslant \text{card}(F)$ (la réciproque est vraie).
S'il existe une fonction $f : E \to F$ surjective, alors $\text{card}F \leqslant \text{card}(E)$ (la réciproque est vraie).

Démonstration
On a toujours $\text{card}f(E) \leqslant \text{card}(E)$, avec égalité ssi f injective.
On a toujours $\text{card}f(E) \leqslant \text{card}(F)$, avec égalité ssi f surjective. □

6. Familles

6.1. Familles quelconques

On peut définir de manière formelle la notion de famille d'éléments ou de famille d'ensemble.

Définition - Familles
Soient I et E deux ensembles. On appelle famille d'éléments de E indexée par I toute "liste" (finie ou non, avec répétitions éventuelles), notée $(a_i)_{i \in I}$, telle qu'à tout élément de I (appelé indice) soit associé un unique élément a_i de E (appelé terme d'indice i de la famille).
Cette famille peut donc être considérée comme l'application

$$\begin{aligned} a : \quad & I \to E \\ & i \mapsto a_i = a(i) \end{aligned}$$

Pour aller plus loin - Cardinal d'un ensemble
Si $A \subset E$, sont des ensembles finis.
Alors $\text{Card}(A) = \sum_{x \in E} \mathbb{1}_A(x)$.
En déduire $\text{Card}(A \cup B) = \text{Card}(A) + \text{Card}(B) - \text{Card}(A \cap B)$.
Que vaut $\text{Card}(A \cup B \cup C) = ?$

Définition - Sous famille
Soit $(a_i)_{i\in I}$ une famille d'éléments d'un ensemble E.
Si $J \subset I$, on dit que $(a_i)_{i\in J}$ est une sous-famille de $(a_i)_{i\in I}$.
on dit également que $(a_i)_{i\in I}$ est une sur-famille de $(a_i)_{i\in J}$.

Exemple - $E = \mathbb{R}$ **et** $I = \mathbb{N}$
Par exemple, si $E = \mathbb{R}$ et $I = \mathbb{N}$, on définit ainsi une *suite* de réels.

Définition - Intersection et réunion d'une famille de parties
Soient un ensemble I (les indices) et un ensemble E.
On considère une famille de parties de E (c'est-à-dire une famille d'éléments de $\mathscr{P}(E)$) $(A_i)_{i\in I}$.
On note

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

Exemple - $E = \mathbb{R}$
Par exemple, si $E = \mathbb{R}$, on a

$$\bigcap_{k\in\mathbb{N}} [-k,k] = \quad , \bigcup_{k\in\mathbb{N}} [-k,k] = \quad , \bigcap_{k\in\mathbb{N}^*} \left]-\frac{1}{k},\frac{1}{k}\right[= \quad .$$

Exercice
On dit que la suite numérique (u_n) converge vers ℓ si :

$$\forall\, \epsilon > 0, \exists\, N \in \mathbb{N} \mid \forall\, n \geqslant N, \quad |u_n - \ell| < \epsilon$$

1. Montrer que $|u_n - \ell| < \epsilon \iff \ell \in]u_n - \epsilon, u_n + \epsilon[$.
2. En déduire que l'ensemble des limites possibles pour la suite (u_n) est l'intersection d'une réunion d'intersection d'ensembles

---

Correction

1. $|u_n - \ell| < \epsilon \iff -\epsilon < \ell - u_n < \epsilon \iff u_n - \epsilon < \ell < u_n + \epsilon \iff \ell \in]u_n - \epsilon, u_n + \epsilon[$
2. On a donc :

$$\forall\, \epsilon > 0, \forall\, N \in \mathbb{N}, \quad \{\ell \mid \forall\, n \geqslant N, \quad |u_n - \ell| < \epsilon\} = \bigcap_{n\geqslant N}]u_n - \epsilon, u_n + \epsilon[$$

Puis

$$\forall\, \epsilon > 0, \quad \{\ell \mid \exists\, N \in \mathbb{N} \mid \forall\, n \geqslant N, \quad |u_n - \ell| < \epsilon\} = \bigcup_{N\in\mathbb{N}} \left(\bigcap_{n\geqslant N}]u_n - \epsilon, u_n + \epsilon[\right)$$

Et

$$\{\ell \mid \forall\, \epsilon > 0, \exists\, N \in \mathbb{N} \mid \forall\, n \geqslant N, \quad |u_n - \ell| < \epsilon\} = \bigcap_{\epsilon>0} \left[\bigcup_{N\in\mathbb{N}} \left(\bigcap_{n\geqslant N}]u_n - \epsilon, u_n + \epsilon[\right) \right]$$

6.2. Famille indexée sur $\mathbb{N}$. Suites

Vocabulaire de base sur les suites (infinies)

L'ensemble E n'est pas précisé pour le moment. Cela peut être $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ ou autre chose ($\mathscr{M}_p(\mathbb{R})\ldots$).
Par la suite nous pourrons avoir besoin que l'ensemble E soit ordonné.

Définition - Suite et ensemble de suites
Soit E un ensemble.
Une suite est une application $u : \mathbb{N} \to E$ (on dira aussi qu'une application de $\{n \in \mathbb{N} | n \geqslant n_0\}$ dans E où $n_0 \in \mathbb{N}$ est une suite). On note cette application sous forme indicielle :

$$(u_n)_{n\in\mathbb{N}} \quad \text{(éventuellement } (u_n)_{n\geqslant n_0}\text{) ou } (u_n)$$

On note $E^{\mathbb{N}}$ l'ensemble des suites à valeurs dans E.

Attention - Avec ou sans parenthèses

(u_n) désigne une suite alors que u_n désigne un nombre de E (le n ou $n+1$-ième terme de la suite).

Exemple - Suite (ou famille) de fonctions

Il arrivera, de plus en plus souvent durant ces années de CPGE que vous rencontriez des suites de fonctions.
On les note en générale (f_n), avec pour tout $n \in \mathbb{N}$, $f_n : I \to \mathbb{R}$.
On peut, par exemple, avoir à étudier $M_n = \sup_{x\in I} f_n(x)$, puis le comportement de la suite (M_n).

$\mathbb{N}$ est ordonnée

Heuristique - Particularité des suites aux familles

L'ensemble d'indexation des suites $\mathbb{N}$ a une particularité très importante que n'a pas I.
Il est ordonné naturellement. Ainsi, une notion importante des suites qui n'existe pas pour les familles est la notion de suite croissante que l'on verra un peu plus bas ou encore les propriétés vraies à partir d'un certain rang.
Autre propriété : si $A \subset \mathbb{N}$, alors A est borné si et seulement si A est fini.

Définition - Propriété vraie à partir d'un certain rang...

On dit qu'une propriété $p(n)$ est vérifiée à partir d'un certain rang s'il existe $n_0 \in \mathbb{N}$ tel que la propriété $p(n)$ soit vraie pour $n \geqslant n_0$.

Exercice

Quelle est le contraire, *formalisée*, d'une propriété vraie à partir d'un certain rang ?
Que penser également de $A = \{n \in \mathbb{N} \mid p(n) \text{ vraie}\}$

Correction

La formalisation positive est : $\exists\, n_0 \in \mathbb{N}$ tel que $\forall\, n \geqslant n$, $p(n)$ est vraie. La négation donne : $\forall\, n \in \mathbb{N}$, $\exists\, n' \geqslant n$ tel que $p(n)$ fausse.
Dans le premier cas A est bornée (par n_0) et dans le second cas A est infini.

Analyse - Suites extraites (sous-suite)

Si on enlève des termes d'une suite, on obtient une suite extraite.
Soit E une partie infinie de $\mathbb{N}$. On dit $(u_n)_{n\in\mathbb{E}}$ est une suite extraite de $(u_n)_{n\in\mathbb{N}}$.
Le problème est que l'indexation sur E n'est pas très aisée à exploiter.
Par ailleurs, la particularité de la croissance doit être conservé. Il faut garder l'ordre.
On aimerait que si $\varphi : \mathbb{N} \to E(\subset \mathbb{N})$ est une bijection, elle conserve l'ordre :

$$k \leqslant h \iff \varphi(k) \leqslant \varphi(h) (\in E)$$

On a donc la définition suivante :

Définition - Suites extraites

On dit que $(v_n)_{n\in\mathbb{N}}$ est une suite extraite de $(u_n)_{n\in\mathbb{N}}$ si il existe $\varphi : \mathbb{N} \to \mathbb{N}$, strictement croissante telle que

$$\forall\, n \in \mathbb{N}, \qquad v_n = u_{\varphi(n)}$$

. On note parfois : pour tout $k \in \mathbb{N}$, $v_k = u_{n_k}$.

Exemple - Suites extraites paires et impaires
La suite extraite des termes d'indice pair de (u_n) est la suite $(u_{2n})_{n\in\mathbb{N}}$.
La fonction φ est ici $n \mapsto 2n\ldots$
Exercice
Montrer que si $p(n)$ est vraie à partir d'un certain rang alors elle est vraie pour tous les termes d'une suite extraite de $\mathbb{N}$ à préciser

Correction
Il s'agit simplement de la suite obtenue avec $\varphi : k \mapsto n_0 + k$

Proposition - Suite extraite et ensemble infini
Considérons une famille de propriété indexée par $\mathbb{N}$, notée $(P_n)_{n\in\mathbb{N}}$.
On note $A = \{n \in \mathbb{N} \mid P_n \text{ vraie.}\}$. Alors
A est infinie si et seulement si $\exists\, \varphi : \mathbb{N} \to \mathbb{N}$, strictement croissante telle que $\forall\, n \in \mathbb{N}$, $P_{\varphi(n)}$ est vraie.

Démonstration
Si il existe une suite extraite de (P_n) toujours vraie,
i.e. si il existe $\varphi : \mathbb{N} \to \mathbb{N}$, strictement croissante telle que $\forall\, n \in \mathbb{N}$, $P_{\varphi(n)}$ est vraie.
Notons $A' = \varphi(\mathbb{N}) = \{\varphi(0), \varphi(1), \ldots \varphi(n), \ldots\} = \{\varphi(k), k \in \mathbb{N}\}$, c'est un ensemble infini car φ est injectif.
Enfin $A' \subset A$. Donc A est infinie
Réciproquement, supposons que A est infini.
Il faut construire φ.
$A \subset \mathbb{N}$, donc A possède un plus petit élément $n_0 = \varphi(0)$.
Construisons φ par récurrence, i.e. pour tout $k \in \mathbb{N}$, on donne une valeur à $\varphi(k)$ bien déterminée, selon les valeurs de $\{\varphi(0), \ldots \varphi(k-1)\}$.
Ainsi fixons $k \in \mathbb{N}^*$ et supposons que $\varphi(0) < \cdots < \varphi(k-1)$ sont bien définis.
L'ensemble $A_k = \{\varphi(0), \ldots \varphi(k-1)\}$ est fini, donc $B_k = A \setminus [\![0, \varphi(k-1)]\!]$ est infini, inclus dans $\mathbb{N}$.
Il possède un plus petit élément noté $n_k = \varphi(k)$ et $\varphi(k) > \varphi(k-1)$ et $P_{\varphi(k)}$ est vraie. □

Suites bornées

On considère $E, \preceq$ un ensemble ordonnée.

Définition - Suite majorée, minorée, bornée
On dit qu'une suite (u_n) d'éléments de E est
- majorée s'il existe $M \in E$ tel que $\forall n \in \mathbb{N}$, $u_n \preceq M$.
- minorée s'il existe $m \in E$ tel que $\forall n \in \mathbb{N}$, $m \preceq u_n$.
- bornée si elle est majorée et minorée.

Remarque - Familles bornées, majorées...
Cette définition est également adaptable au cas des familles simples.
Exercice
Montrer qu'une suite majorée à partir d'un certain rang est une suite majorée.

Correction
Si (x_n) est majorée à partir du rang n_0,
Alors il existe M tel que pour tout $n \geqslant n_0$, $x_n \leqslant M$.
Notons $M' = \max(M, x_0, x_1, \ldots x_{n_0-1})$ (ensemble fini), alors : $\forall\, n \in \mathbb{N}$, $x_n \leqslant M'$

Suites monotones

Définition - Suite croissante, décroissante
On dit qu'une suite (u_n) d'éléments de E est
- croissante
  si $\forall n \in \mathbb{N}$, $u_n \preceq u_{n+1}$.
- décroissante
  si $\forall n \in \mathbb{N}$, $u_{n+1} \preceq u_n$.
- monotone si elle est croissante ou décroissante.
- stationnaire si elle est constante à partir d'un certain rang.

Exemple - Deux suites monotones :

— La suite (u_n) où $u_n = \binom{2n}{n}$ est monotone.
$u_{n+1} - u_n = \binom{2n}{n}\left(\frac{(2(2n+1)}{n+1} - 1\right) = \binom{2n}{n}\frac{3n+1}{n+1} > 0$

— Soit (u_n) une suite de nombres positifs et $m_p = \inf_{n \geqslant p} u_n$. Alors (m_p) est une suite croissante.
En effet : $m_p = \min(m_{p+1}, u_p)$, donc $m_p \leqslant m_{p+1}$

Nous élargirons ces notions, lorsque nous nous concentrerons sur les suites numériques, une fois que $\mathbb{R}$ sera construit…

7. Bilan

Synthèse

⇝ Depuis la fin du XIX, on travaille en mathématiques à partir d'ensembles. On peut aussi passer d'un ensemble à un autre par des applications. Les applications injectives ne mélangent pas les éléments de l'ensemble du départ, les applications surjectives sont complètes (vu de l'arrivée).

⇝ Très souvent en mathématiques (probabilité, construction de l'intégrale), on s'intéresse plutôt aux ensembles réciproques (ie des antécédents) $f^{-1}(B)$ qu'aux ensembles images directes $f(A)$.

⇝ La fonction indicatrice d'un ensemble A dans E est une projection naturelle de E sur A. Elle est d'une utilité essentielle en mathématique, par exemple pour calculer le cardinal (ou en probabilité).

⇝ On termine par décrire les propriétés pour des familles indexées sur un ensemble I fini ou dénombrable.

Savoir-faire et Truc & **Astuce du chapitre**

— Savoir-faire - Montrer une égalité entre deux fonctions.
— Savoir-faire - Autre formulations équivalentes (injectivité, surjectivité)
— Savoir-faire - Critère pour montrer la bijectivité
— Savoir-faire - Critère de surjectivité.
— Savoir-faire - Montrer que $y \in f(A)$
— Savoir-faire - Montrer que $x \in f^{-1}(B)$

Notations

	Propriétés	*Remarques*
définition (du terme de gauche)		autre notation : $\overset{\text{def.}}{s} =$
pplication de E sur F, injective		Cette notation donne deux informations
pplication de E sur F, surjective		Cette notation donne deux informations
de f à l'ensemble de départ A ment d'arrivée B)	$f_{\|A} : A \to F$, $x \mapsto f(x)$ et $f^{\|B} : E \to B$, $x \mapsto f(x)$	Pour la restriction d'arrivée, il faut vérifier que cela a du sens…
te de A par f	$\{f(x), x \in A\}$. $y \in f(A) \Leftrightarrow \exists\, x \in A \mid y = f(x)$	autre notation : $[\mathrm{Im}(f_{\|A}$
Image) réciproque de B par f	$\{x \in E \mid f(x) \in B\}$. $x \in f^{-1}(B) \Leftrightarrow f(x) \in B$	autre notation (probabilité) : $[f \in B]$
nion disjointe des ensembles A_i	$x \in C$ ssi $\exists\, !i \in \mathbb{N}_n$ tel que $x \in A_i$	Deux informations : C est la réunion des A_i & les A_i sont disjoints deux à deux.
$\Longrightarrow x \in A$	Codage numérique d'une propriété caractéristique	$\mathbb{1}_{A\cap B} = \mathbb{1}_A \times \mathbb{1}_B$, $\mathrm{Card} A = \sum_{x\in E} \mathbb{1}_A(x)$

Retour sur les problèmes

50. Voir cours

51. Une bijection montre que deux ensembles sont équivalents (si ce n'est égaux).

52. Là on sort du cours.
    $\mathbb{Z}$ et $\mathbb{N}$ ont la même puissance ($\approx$ cardinal) car on peut les mettre en bijection l'un par l'autre.
    Avec par exemple $\varphi : \mathbb{Z} \to \mathbb{N}$, $m \mapsto 2|m| + \mathbb{1}_{\mathbb{Z}_-}(m)$.
    De même (c'est plus subtile) $\mathbb{Z} \times \mathbb{N}$ est en bijection avec $\mathbb{N}$ (vous pouvez trouver une bijection?), donc $\mathbb{Q}$ est également en bijection avec $\mathbb{N}$.
    On dit que $\mathbb{N}$, $\mathbb{Z}$ et $\mathbb{Q}$ ont la puissance du dénombrable (le même cardinal infini).
    En revanche, il n'existe aucune bijection entre $\mathbb{R}$ et $\mathbb{N}$. $\mathbb{R}$ a un cardinal plus grand, on parle du cardinal du continue.

53. Cours.

Chapitre 12

Relations binaires sur un ensemble

Résumé -

Nous complétons quelques notions essentielles du fondement des mathématiques (formalisé non sans mal à la fin du XIX-ième siècle). Ces fondements se basent sur les ensembles !
Mais il faut agir sur ces ensemble. Nous commençons donc d'abord par voir deux notions dont l'emploi en mathématiques est fréquent (en particulier lorsqu'il s'agit de construire de nouvelles notions). Il s'agit des relations binaires : relation d'ordre et relation d équivalence.
Dans ce chapitre, il y a beaucoup de définitions. A apprendre !! Quelques vidéos :

- *Science4all - Les mathématiques modernes - https ://www.youtube.com/watch ?v=7fbn99V1f9U*
- *Maths Adultes - Relations binaires - https ://www.youtube.com/watch ?v=W7cH06qOImM*

Sommaire

1. Problèmes

? Problème 56 - Graphe

En option « maths expertes », les élèves étudient les graphes : sommets, arrêtes....
C'est une famille essentielle d'objets en mathématiques et en informatique (nous les y retrouverons en fin d'année). Comment formaliser proprement les graphes ? Comment passer de l'idée bien comprise (de sommets et d'arêtes/flèches) à une représentation mathématique/informatique acceptable ?

? Problème 57 - Forcer l'égalité. Qu'est-ce qu'une égalité ?

Pour résoudre un exercice, on exploite souvent des équivalences ($\Leftrightarrow$). La bijection de f permet d'écrire : $f(x) = y \Longleftrightarrow x = f^{-1}(y)$ en prenant x et y dans les bons ensembles.
Si f n'est pas surjective, il suffit de changer l'ensemble de définition de y et la résolution du problème se conserve.
Mais si f n'est pas injective, qu'il y a plusieurs solutions $x_1, x_2, \dots x_n$ à l'équation $f(x) = y$. Que faire ?
Une idée : forcer l'égalité et affirmer que $x_1 = x_2 = \dots x_n$, comme pour l'équation $\tan x = \sqrt{3}$ qui permet d'affirmer $x = \frac{\pi}{3}$ ou $x = \frac{4\pi}{3}$....
C'est choquant ! Comment rendre cela propre : en redonnant un sens nouveau à l'égalité $x_1 = x_2 = \dots x_n$. Qu'est-ce qu'une égalité ?

? Problème 58 - Relation d'ordre ?

Pour résoudre le problème précédent, nous créons la notion de relation d'équivalence.
Mais plus souvent, lorsque nous prenons deux objets nous ne pouvons pas affirmer qu'ils sont pareils. Souvent l'un est PLUS quelque chose que l'autre.
Comment formaliser cette idée ? Qu'est-ce qu'une relation d'ordre ? Et comment l'exploiter ?

? Problème 59 - Plus grand élément

Existe-t-il nécessairement un, un seul, plus grand élément à un ensemble ordonné ?
Par exemple : quel est le plus grand élément de $[0, 1[$?

? Problème 60 - Codage de graphe (ou relation) à partir d'applications. Et réciproquement...

Après avoir défini les ensembles, on peut ou bien définir les applications et à partir de là les relations (ou graphes), ou bien définir les relations et à partir de là les applications.
Comment faire naturellement ces deux implications ? (Evidemment, pas en même temps...)

2. Graphe

2.1. Formalisation

Heuristique - Images mentales des graphes!

Pour l'heuristique et les images mentales (à ne pas oublier et vraiment à garder en mémoire!!), il faut revoir le cours de mathématiques de terminale.

Définition - Graphe non orienté

On considère un ensemble S (de sommets), fini en règle générale. Puis un ensemble $A \subset \binom{S}{2}$ de paires d'arêtes.
On appelle graphe non orienté le couple (S, A).

Définition - Graphe orienté

On considère un ensemble S (de sommets), fini en règle générale. Puis un ensemble $A \subset S \times S$ de couples d'arêtes.
On appelle graphe orienté le couple (S, A).

Exercice

Donner la définition formalisée d'un graphe complet.

Correction

Un graphe complet est, heuristiquement, dont chaque sommet est relié à tous les autres.

Formellement : $A = S \times S$ (cas orienté) ou $A = \binom{S}{2}$ (cas non orienté).

2.2. Vocabulaire

Définition - Sommets reliés

Soit (S, A) un graphe (orienté ou non).
On dit que deux sommets $s_1, s_2 \in S$ sont reliés si $(s_1, s_2) \in A$ (cas orienté) ou $\{s_1, s_2\} \in A$ (cas non orienté)

Définition - Degré d'un sommet

Soit $s \in S$ un sommet d'un graphe non orienté (S, A) a pour degré $d(s) = \text{card}(A_s)$ où $A_s = \{a \in A \mid s \in a\}$.
Soit $s \in S$ un sommet d'un graphe orienté (S, A) a pour degré sortant $d_+(s) = \text{card}(A_s)$ où $A_s = A \cap (\{s\} \times S)$ et pour degré entrant $d_-(s) = \text{card}(A'_s)$ où $A'_s = A \cap (S \times \{s\})$.

Exercice

Comment définir chemin d'un sommet à un autre ?
Et graphe connexe ?

Correction

On dit que le graphe (S, A) admet un chemin de s à s' ($\in S$) s'il existe un entier n et une suite $(s_0, s_1, s_2, \dots s_n)$ d'éléments de S tels que $s_0 = s$, $s_n = s'$ et $\forall\, i \in \mathbb{N}_n$, $(s_{i-1}, s_i) \in A$ (cas orienté) ou $\{s_{i-1}, s_i\} \in A$ (cas non orienté).
Un graphe est connexe si pour tout sommets s, $s' \in S$, il existe un chemin de s à s'.

2.3. Applications

On retrouvera très vite les graphes dans le cours sur les relations binaires, plus loin en probabilité et algèbre linéaire (chaine de Markov), ou en informatique... A l'occasion, nous verrons en informatique, un façon supplémentaire et pratique de coder/définir un graphe à l'aide de matrice...

3. Relations binaires

3.1. Construction et représentation

Définition - Relation
Soit E un ensemble.
Une relation binaire sur E est un sous-ensemble G de $E \times E$. Si $(x,y) \in E^2$ on écrit $x\mathscr{R}y$ lorsque $(x,y) \in G$.

On peut représenter une relation par un graphe (diagramme sagittal) : une représentation de $E \times E$ et avec des flèches on indique que x (du premier E) est en relation à y (du second E).

Exemple - Stade Toulousain
Par exemple dans l'ensemble E=Boutique du Stade Toulousain où :

$$E = \{\text{beret rouge, chaussette blanche, maillot rouge, maillot noir, short rouge, cuissart noir}\} = \{B_R, CH_B, M_R, M_N, S_R, C_N\},$$

on définit la relation $\mathscr{R}_1$ par "est de la même couleur que" c'est-à-dire que l'on a

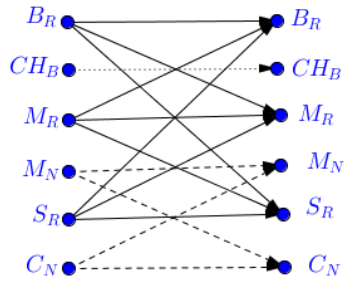
$$G_1 = \{(B_R,M_R),(M_R,B_R),(M_R,S_R),(S_R,M_R),(B_R,S_R),(S_R,B_R),(B_R,B_R),(M_R,M_R),(S_R,S_R),(CH_B,CH_B),(M_N,C_N),(C_N,M_N),(M_N,M_N),(C_N,C_N)\}$$

ou encore $B_R \mathscr{R}_1 M_R$, $M_R \mathscr{R}_1 B_R$, $M_R \mathscr{R}_1 S_R$...

Exercice
On peut définir dans l'ensemble $\{0,1,2,3,4,5,6\}$ les relations $\mathscr{R}_1$ "est un multiple de" ou $\mathscr{R}_2$ "est le double de".
Expliciter G_1, G_2 et les diagrammes sagittaux de ces deux relations.

Correction
$G_1 = \{(0,0),(0,1),(0,2),(0,3),(0,4),(0,5),(0,6),(1,1),(2,1),(2,2),(3,1),(3,3),(4,1),(4,2),(4,4),(5,1),(5,5),(6,1),(6,2),(6,3),(6,6)\}$ et $G_2 = \{(0,0),(2,1),(4,2),(6,3)\}$

Représentation - Graphe



C'est le graphe de l'exercice du Stade Toulousain.

3.2. Caractérisations

Définition - Propriétés des relations
Soit $\mathscr{R}$ une relation sur un ensemble E. On dit que $\mathscr{R}$ est :
- réflexive si $\forall x \in E$, $x\mathscr{R}x$;
- symétrique si $\forall (x,y) \in E^2$, $x\mathscr{R}y \Rightarrow y\mathscr{R}x$;
- antisymétrique si $\forall (x,y) \in E^2$, $x\mathscr{R}y$ et $y\mathscr{R}x \Rightarrow x = y$;
- transitive si $\forall (x,y,z) \in E^3$, $x\mathscr{R}y$ et $y\mathscr{R}z \Rightarrow x\mathscr{R}z$.

Exemple - Stade Toulousain
Dans l'exemple précédent, la relation est réflexive, symétrique et transitive.
Exercice
Comment se représentent pour un graphe les propriétés précédentes ?

Correction
En terme de graphe, cela signifie que :
- pour la réflexivité : chacun est relié à lui-même, donc le diagramme présente des lignes droites. (Dans le cas d'une représentation matricielle : que des 1 sur la diagonale)
- pour la symétrie : un flèche dans un sens, donne une flèche dans l'autre sens (on pourrait faire des doubles flèches).
  On parle alors de graphe non orienté.
  (Dans le cas d'une représentation matricielle : la matrice est symétrique)
- pour l'antisymétrie : il ne peut y avoir de double flèche, excepté sur eux-mêmes.
  (Dans le cas d'une représentation matricielle : la matrice est presque antisymétrique)
- pour la transitivité : il y a des blocs de points regroupés entre eux.

4. Relation d'ordre

4.1. Définitions

Définition - Relation d'ordre
Soit $\mathscr{R}$ une relation sur un ensemble E. On dit que c'est une relation d'ordre si elle est réflexive, antisymétrique et transitive.

Définition - Plus petit
Une relation d'ordre permet de *comparer* deux éléments. Lorsque $x\mathscr{R}y$ on dit que x est "plus petit" que y et on note usuellement $x \preceq y$.

Savoir faire - Montrer que $\mathscr{R}$ est une relation d'ordre.
Il s'agit de montrer, tour à tour, que la relation est réflexive, antisymétrique et transitive.

Pour aller plus loin - Treillis (de Galois)
Pour les relations d'ordre, au lieu du grpahe, on préfère une représentation graphique sous forme de treillis (de Galois).
Si $x \preceq y$, alors on représente x «sous» y, et l'on trace un lien entre les deux.
Si l'ordre est total, il n'y a qu'un élément à chaque hauteur.
Cette représentation n'a d'intérêt que pour E de cardinal fini (et petit), même si elle peut aussi donner de bonnes idées complémentaires.

4.2. Ensemble avec ordre total

Définition - Ordre total
Soit $\preceq$ une relation d'ordre sur un ensemble E. On dit que c'est une relation d'ordre total si
$$\forall (x,y) \in E^2,\ x \preceq y \text{ ou } y \preceq x$$
(c'est-à-dire si deux éléments quelconques de E sont comparables).

Remarque - Concernant le treillis (ou graphe)
Le fait que l'ordre soit total signifie que le treillis/graphe est connexe (en un seul morceau).

Exemple - Sur $\mathbb{R}$
Par exemple, dans $\mathbb{R}$ la relation $\leqslant$ est une relation d'ordre total, en revanche $<$ n'est pas une relation d'ordre.
En effet, $<$ n'est pas reflexive.

Exercice
Sur $E = \mathbb{R}^2$ on définit les deux relations suivantes :
— l'ordre produit :
$$(x,y) \preceq_1 (x',y') \Leftrightarrow x \leqslant x' \text{ et } y \leqslant y'$$
— l'ordre lexicographique :
$$(x,y) \preceq_2 (x',y') \Leftrightarrow (x < x') \text{ ou } (x = x' \text{ et } y \leqslant y')$$
Vérifier qu'il s'agit de relations d'ordre. S'agit-il d'ordre partiel ou d'ordre total ?

Correction
Première relation :
— Réflexive : Pour tout $x, y \in \mathbb{R}$, on a $x \leqslant x$ et $y \leqslant y$, donc $(x,y) \preceq_1 (x,y)$
— Antisymétrique :Soient $x, y, x'y' \in \mathbb{R}$ tels que $(x,y) \preceq_1 (x',y')$ et $(x',y') \preceq_1 (x,y)$
Donc $x \leqslant x'$, $x' \leqslant x$, $y \leqslant y'$ et $y' \leqslant y$.
Donc $x = x'$ et $y = y'$ (car $\leqslant$ est antisymétrique sur $\mathbb{R}$).
Finalement $(x,y) = (x',y')$
— Transitive : Soient $x, y, x', y', x'', y'' \in \mathbb{R}$, tels que $(x,y) \preceq_1 (x',y')$ et $(x',y') \preceq_1 (x'',y'')$.
On a donc $x \leqslant x' \leqslant x''$ et $y \leqslant y' \leqslant y''$, donc $(x,y) \preceq_1 (x'',y'')$
Cette relation n'est pas totale : il n'y a aucune relation entre $(1,0)$ et $(0,1)$ (et pas de relation d'ordre totale sur $\mathbb{C}$). Première relation :
— Réflexive : Pour tout $x, y \in \mathbb{R}$, on a $x = x$ et $y \leqslant y$, donc $(x,y) \preceq_2 (x,y)$
— Antisymétrique :Soient $x, y, x'y' \in \mathbb{R}$ tels que $(x,y) \preceq_2 (x',y')$ et $(x',y') \preceq_2 (x,y)$
On ne peut avoir $x < x'$, sinon, on aurait pas $(x',y') \preceq_2 (x,y)$, donc $x = x'$.
Puis $y \leqslant y'$ et $y' \leqslant y$ donc $y = y'$.
Finalement $(x,y) = (x',y')$
— Transitive : Soient $x, y, x', y', x'', y'' \in \mathbb{R}$, tels que $(x,y) \preceq_2 (x',y')$ et $(x',y') \preceq_2 (x'',y'')$.
Alors $x \leqslant x' \leqslant x''$.
- Ou bien $x < x''$ et donc $(x,y) \preceq_2 (x'',y'')$
- Ou bien $x = x''$ et donc $x = x' = x''$
et donc $y \leqslant y' \leqslant y''$ et ainsi $(x,y) \preceq_2 (x'',y'')$

Dans tous les cas $(x,y) \preceq_2 (x'',y'')$
Cette relation est totale : elle permet d'écrire le dictionnaire !
C'est aussi l'ordre total qui permet de classer les nombres écrits décimalement (ou une base quelconque, d'ailleurs).

4.3. Ensemble avec ordre partiel

Définition - Ordre partiel
Soit $\preceq$ une relation d'ordre sur un ensemble E. On dit que c'est une relation d'ordre partiel s'elle n'est pas total.
C'est-à-dire : il existe $(x,y) \in E^2$ tel que $x \not\preceq y$ et $y \not\preceq x$.

Exercice
Soit Ω un ensemble et $E = \mathscr{P}(\Omega)$. On définit sur E la relation $\mathscr{R}$ par

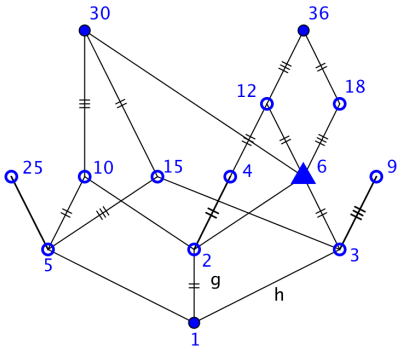
$$\forall (A,B) \in E^2, A \mathscr{R} B \Leftrightarrow A \subset B.$$

Vérifier que la relation $\mathscr{R}$ est une relation d'ordre. S'agit-il d'une relation d'ordre total ?

Correction
- Elle est réflexive : pour tout ensemble A, on a $A \subset A$.
- Elle est antisymétrique : soient A et B tels que $A \subset B$ et $B \subset A$, alors $A = B$.
- Elle est transitive : soient A, B et C tels que $A \subset B$ et $B \subset C$, alors $A \subset C$.

Mais ce n'est pas une relation d'ordre totale : si $\text{Card}(E) \geqslant 2$, il n'y a pas de relation entre $\{a\}$ et $\{b\}$ si $a \neq b$.

***Représentation - Treillis de diviseurs de** 30 **et** 36*



Exemple - Divisibilité sur $\mathbb{N}$
La relation « divise » : $n|m$, si il existe $k \in \mathbb{N}$ tel que $m = nk$ est une relation d'ordre partielle.
Exercice
Montrer ce résultat

Correction
C'est une relation reflexive, antisymétrique et transitive

Remarque - Treillis
On peut faire un treillis de divisibilité de certains nombres entiers.
Ce n'est pas une droite, comme avec $(\mathbb{R}, \leqslant)$ par exemple.

4.4. Eléments particuliers

Majorant/minorant

Définition - Majorants,minorants
Soit $\preceq$ une relation d'ordre sur un ensemble E. Pour $A \subset E$, on définit les éléments suivants :
- $M \in E$ est un majorant de A si $\forall x \in A, x \preceq M$;
- $m \in E$ est un minorant de A si $\forall x \in A, m \preceq x$;

Exemple - Majorant sur $(\leqslant, \mathbb{R})$
Pour la relation d'ordre $\leqslant$ sur $\mathbb{R}$,
3 est un majorant de l'ensemble $\{\frac{1}{n}, n \in \mathbb{N}\}$.
D'une certaine façon, ce n'est pas le meilleur.

Exemple - Majorant sur $(|, \mathbb{N})$
Pour la relation d'ordre $|$ sur $\mathbb{N}$,
- 120 est un majorant de $\{1,2,3,4,5\}$
  Ce n'est pas le meilleur. On dit que c'est un multiple.
  La majorant le plus intéressant serait le plus petit multiple. (PPCM)
- 3 est un minorant de $\{6,9,12\}$.
  On dit que c'est un diviseur.
  C'est d'une certaine façon le meilleur car c'est le plus grand des diviseurs (PGCD).

Exercice
Pour la relation d'ordre $\subset$ sur $\mathbb{R}$.
Donner un majorant et un minorant de $\big\{\{1,2,3\},\{2,3,5\}\big\}$

Correction
On peut prendre par exemple : $A = \{1,2,3,5\}$ et $B = \{2,3\}$ respectivement.
En fait tout majorant doit contenir A et tout minorant est inclus dans B

Plus grand/petit élément

Définition - Plus grand élément, plus petit élément
Soit $\preceq$ une relation d'ordre sur un ensemble E. Pour $A \subset E$, on définit les éléments suivants :
- $a \in E$ est un plus grand élément de A si $a \in A$ et $\forall x \in A, x \preceq a$;
- $a \in E$ est un plus petit élément de A si $a \in A$ et $\forall x \in A, a \preceq x$.

En fait a est respectivement un majorant de A et élément de A ou bien un minorant et élément de A

Théorème - Unicité
Un plus petit grand élément de $A \subset E$, lorsqu'il existe, est unique, noté $\max(A)$.
Un plus petit petit élément de $A \subset E$, lorsqu'il existe, est unique, noté $\min(A)$.

Démonstration
Supposons que A admettent (au moins) deux plus petits éléments : a_1 et a_2.
Alors pour tout $x \in A$, $a_1 \preceq x$, donc en particulier pour $x = a_2 \in A$: $a_1 \preceq a_2$.
Et pour tout $x \in A$, $a_2 \preceq x$, donc en particulier pour $x = a_1 \in A$: $a_2 \preceq a_1$.
Par antisymétrie : $a_1 = a_2$. A admet au plus un seul plus petit élément. □

Attention - Attention au mot
Ici il ya une source d'erreur classique. On fera bien attention aux mots définis ici : (un) majorant, (un) minorant, (le) plus grand élément, (le) plus grand élément.
S'ajoutent à ces mots : élément maximal, minimal…
Il y aura bientôt également l'expression borne supérieure, borne inférieure…

On rencontre très souvent le cas suivant :
Exercice
On suppose que $\preceq$ est une relation d'ordre total sur E.
Soit $A \subset E$. On suppose que A est fini.
Montrer que A admet nécessairement un plus grand élément

Correction
On fait une récurrence sur $\mathrm{Card}(A)$.
$\mathscr{P}_n$: Si $\mathrm{Card}(A) = n$, alors A admet un plus grand élément. Pour l'hérédité, on note que si $A = A_n \uplus \{a\}$, alors $\max(A) = \max(\max(A_n), a)$… Nous encourageons également à faire une démonstration par algorithme.

Exercice
On admet qu'un ensemble fini admet toujours un plus petit élément (il suffit de faire au plus $\frac{n(n-1)}{2}$ comparaisons - mais on peut se contenter de $n-1$ comparaisons…).
Montrer que tout sous-ensemble de $\mathbb{N}$ admet un plus petit élément.

Correction
Soit $A \subset \mathbb{N}$. Soit $a \in A$.
Alors $A' = A \cap [\![0, a]\!]$ est fini (car inclus dans $[\![0, a]\!]$), il admet un plus petit élément a'.
Pour tout $x \in A$,
- ou bien $x > a$ et donc $a' \leqslant a < x$,
- ou bien $x \leqslant a$ donc $x \in A'$ et donc $a' \leqslant x$.

Ainsi a' est le plus petit élément de A

Eléments maximaux/minimaux

Remarque - Généralisation : élément maximal ou minimal
On peut également définir la notion d'élément maximal ou minimal :
- $M \in A$ est un *élément maximal* de A si $\forall x \in A, M \preceq x \Rightarrow x = M$;
- $m \in A$ est un *élément minimal* de A si $\forall x \in A, x \preceq m \Rightarrow x = m$.

S'il s'agit d'une relation d'ordre total, ces éléments coïncident avec les plus grand et plus petit éléments (s'ils existent).

Exemple - Ensemble avec plusieurs éléments maximaux

Pour qu'il y en ait plusieurs, il ne faut pas qu'ils coïncident avec l'unique plus grand élément.

Donc d'après la remarque, l'ordre ne doit pas être total.

On peut prendre l'ordre produit sur $E = \mathbb{R}^2$ et $A = \{(x,y) \in E \mid x^2 + y^2 \leqslant 1\}$, le disque unité.

Alors A n'admet pas de plus grand élément et tous les éléments de $M = \{(x,y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1, x \geqslant 0, y \geqslant 0\}$ sont des éléments maximal de A.

En effet, si $(x_0, y_0) \in M$ et $(x,y) \in A$, avec $(x_0, y_0) \preceq_1 (x,y)$, alors $x_0 \leqslant x$ et $y_0 \leqslant y$.

Alors $1 = x_0 + y_0 \leqslant x^2 + y^2 \leqslant 1$, on a donc $x^2 + y^2 = 1 = x_0^2 + y_0^2$,

Ainsi $(x^2 - x_0^2) + (y^2 - y_0^2) = 0$, avec un somme de termes positifs, donc $x^2 - x_0^2 = 0$ et $y^2 - y_0^2 = 0$.

Enfin pat positivité de x_0 et y_0 : $x = x_0$ et $y = y_0$, donc $(x,y) = (x_0, y_0)$.

Borne supérieure/inférieure

Une dernière définition, pour des cas plus simples que celui de l'exemple précédent :

Définition - Borne inférieure, borne supérieure

Soit A un ensemble muni d'une relation d'ordre.

- Si l'ensemble des majorants de A est non vide et admet un plus petit élément a, a est appelé borne supérieure de A, on note $a = \sup A$.
- Si l'ensemble des minorants de A est non vide et admet un plus grand élément b, b est appelé borne inférieure de A, on note $b = \inf A$.

Cette définition sera au coeur de la définition de l'ensemble $\mathbb{R}$ (à partir de $\mathbb{Q}$ avec la relation d'ordre totale $\leqslant$). Mais elle sert aussi à d'autres moments du cours

Exemple - Borne inférieure et supérieure pour $(\mathbb{N}, |)$ (divisibilité)

Comme leur nom l'indique :

- La borne supérieure de l'ensemble fini d'entiers $\{u, v\}$ est le $PPCM(u,v)$.
- La borne inférieure de l'ensemble fini d'entiers $\{u, v\}$ est le $PGCD(u,v)$.

Savoir faire - Montrer que $a = \sup E$

On montre en deux temps :

1. $\forall\, x \in E,\ x \preceq a$
2. $\forall\, z$ tel que $\forall\, x \in E,\ x \preceq z$, alors $a \preceq z$
   (tout majorant z de E est plus grand que a)
   OU (de manière équivalente)
   $\forall\, u \preceq a, \exists\, x \in E$ tel que $u \preceq x$
   (tout élément plus petit que a ne peut pas être un majorant de E)

Exercice

Comment repérer sur le treillis des multiples et diviseurs de u et de v, leur PGCD et leur PPCM ?

Correction

C'est le grand père commun (PPCM) ou le petit fils commun (PGCD)

Exercice

Comment peut-on définir l'ensemble borne supérieure de deux ensembles A et B pour la relation $\subset$?

Même question avec la borne inférieure ?

Correction

Tout simplement : $A \cup B$ et $A \cap B$ respectivement

Pour aller plus loin - Espaces vectoriels

On indiquera que dans le cadre des espaces vectoriels, où l'on exige en outre une stabilité pour l'addition vectoriel, la borne supérieure des sous-espaces vectoriels F_1 et F_2 est donnée par l'ensemble $F_1 + F_2$

4.5. Ordre strict

Définition - Ordre strict
Soit $(E, \preceq)$ un ensemble ordonné. On définit la relation $\prec$ par :

$$x \prec y \Leftrightarrow (x \preceq y \text{ et } x \neq y)$$

ce n'est pas une relation d'ordre sur E car elle n'est pas réflexive.

Exemple - Sur $\mathbb{R}$
La relation $\leqslant$ est une relation d'ordre (totale).
Alors que $<$ est une relation d'ordre strict.

5. Relation d'équivalence

5.1. Propriétés caractéristiques

Définition - Relation d'équivalence
Soit $\mathscr{R}$ une relation sur un ensemble E. On dit que c'est une relation d'équivalence si elle est réflexive, symétrique et transitive.

Pour aller plus loin - Relation d'équivalence : volonté de définir =
Sur le site images des maths, un article intéressant : http ://images.math.cnrs.fr/Egalite

Exemple - Stade Toulousain
On a vu une première relation d'équivalence, avec l'exemple du stade Toulousain.

Exemple - Fractions rationnelles
Montrer que $\mathscr{R}$ définie sur $\mathbb{Z} \times \mathbb{N}$ par $(a,b)\mathscr{R}(c,d)$ ssi $a \times d = b \times c$ est une relation d'équivalence.
Montrons que $\mathscr{R}$ ainsi définie est :

- réflexive.
  $a \times b = b \times a (= ab)$ donc $(a,b)\mathscr{R}(a,b)$.
- symétrique.
  Supposons que $(a,b)\mathscr{R}(c,d)$, donc $a \times d = b \times c$
  donc $c \times b = d \times a$ donc $(c,d)\mathscr{R}(a,b)$.
- transitive.
  Supposons que $(a,b)\mathscr{R}(c,d)$ et $(c,d)\mathscr{R}(e,f)$ donc $a \times d = b \times c$ et $c \times f = d \times e$.
  donc $(a \times f) \times d = f \times (a \times d) = f \times (c \times b) = (f \times c) \times b = d \times e \times b$.
  et comme d est non nul, il est inversible (dans $\mathbb{R}$) et donc $a \times f = e \times b$ donc $(a,b)\mathscr{R}(e,f)$.

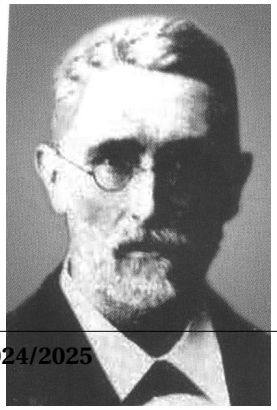
Savoir faire - Montrer que $\mathscr{R}$ est une relation d'équivalence
Il s'agit de montrer, tour à tour, que la relation est réflexive, symétrique et transitive.

Exercice
Montrer que $\mathscr{R}$ définie sur $\left(\mathbb{R}^{\mathbb{N}}\right)^2$ (ensemble des suites) par $(u_n)\mathscr{R}(v_n)$ ssi $\lim\limits_{n\to+\infty} \dfrac{u_n}{v_n} = 1$ est une relation d'équivalence.

Correction
Montrons que $\mathscr{R}$ ainsi définie est :

- réflexive.
  $\lim\limits_{n\to+\infty} \dfrac{u_n}{u_n} = 1$ donc $(u_n)\mathscr{R}(u_n)$.
- symétrique.
  Supposons que $(u_n)\mathscr{R}(v_n)$, donc $\lim\limits_{n\to+\infty} \dfrac{u_n}{v_n} = 1$
  donc $\lim\limits_{n\to+\infty} \dfrac{v_n}{u_n} = \dfrac{1}{1} = 1$ donc $(v_n)\mathscr{R}(u_n)$.
- transitive.
  Supposons que $(u_n)\mathscr{R}(v_n)$ et $(v_n)\mathscr{R}(w_n)$ donc $\lim\limits_{n\to+\infty} \dfrac{u_n}{v_n} = 1$ et $\lim\limits_{n\to+\infty} \dfrac{v_n}{w_n} = 1$.
  donc $\lim\limits_{n\to+\infty} \dfrac{u_n}{w_n} = \lim\limits_{n\to+\infty} \dfrac{u_n}{v_n} \times \lim\limits_{n\to+\infty} \dfrac{v_n}{w_n} = 1$ et donc $a \times f = e \times b$ donc $(u_n)\mathscr{R}(w_n)$.

Histoire - Construction de $\mathbb{Z}$
Etant donné $\mathbb{N}$, construit par un principe de type récurrence (par Péano) ; on peut suivre Dedekind est construire $\mathbb{Z}$.
On note sur $\mathbb{N}^2$, $\mathscr{R}$ telle que $(n,m)\mathscr{R}(n',m')$ ssi $n+m' = n'+m$
Les classes d'équivalence sont de la forme $(0,n) = \{(k, k+n), k \in \mathbb{N}\}$, représenté de l'entier relatif : $-n$.



Richard Dedekind (1831-1916) est un brillant mathématicien allemand, hanté par une question : « qu'est-ce que sont les nombres ? ».

Exercice
Soit $f : E \to F$.
On définit la relation $\mathscr{R}_f$ sur E par $x\mathscr{R}_f y$ ssi $f(x) = f(y)$.
Montrer que $\mathscr{R}_f$ est une relation d'équivalence sur E.

Correction
Trivial

5.2. Classes d'équivalence

Définition - Classe d'équivalence
Soit $\mathscr{R}$ une relation d'équivalence sur E.
Pour $a \in E$, on appelle classe d'équivalence de a l'ensemble $C(a) = \{x \in E \mid x\mathscr{R}a\}$. a est un représentant de $C(a)$.

Exemple - Fractions rationnelles
Les classes d'équivalence des fractions rationnelles sont exactement les couples dont les fractions sont égales.
On voit sur cette exemple le but de la notion de classe d'équivalence : préciser et généraliser la notion d'égalité !
Les fractions rationnelles simplifiés sont des représentants particulièrement simple de leur classe d'équivalence.

Exercice
Montrer que $\mathscr{R}$ définie sur $\mathbb{C}^{*2}$, par $z = a + \mathrm{i}b\mathscr{R}z' = a' + \mathrm{i}b'$ ssi $a \times b' = a' \times b$ est une relation d'équivalence.
Quelles sont les classes d'équivalence de $\mathscr{R}$?

Correction
Montrons que $\mathscr{R}$ ainsi définie est :
- réflexive.
  $a \times b = b \times a (= ab)$ donc $a + \mathrm{i}b\mathscr{R}a + \mathrm{i}b$.
- symétrique.
  Supposons que $a + \mathrm{i}b\mathscr{R}c + \mathrm{i}d$, donc $a \times d = b \times c$
  donc $c \times b = d \times a$ donc $c + \mathrm{i}d\mathscr{R}a + \mathrm{i}b$.
- transitive.
  Supposons que $a + \mathrm{i}b\mathscr{R}c + \mathrm{i}d$ et $c + \mathrm{i}d\mathscr{R}e + \mathrm{i}f$ donc $a \times d = b \times c$ et $c \times f = d \times e$.
  donc $(a \times f) \times d = f \times (a \times d) = f \times (c \times b) = (f \times c) \times b = d \times e \times b$.
  et comme d est non nul, il est inversible (dans $\mathbb{R}$) et donc $a \times f = e \times b$ donc $a + \mathrm{i}b\mathscr{R}e + \mathrm{i}f$.

Les classes d'équivalence peuvent représentées par les droites passant par 0.

Exercice
Montrer que $\mathscr{R}$ définie sur $\mathbb{R}$, par $\theta\mathscr{R}\theta'$ ssi $\exists\, k \in \mathbb{Z}$ tel que $\theta - \theta' = 2k\pi$ est une relation d'équivalence.
Quelles sont les classes d'équivalence de $\mathscr{R}$?

Correction
Montrons que $\mathscr{R}$ ainsi définie est :
- réflexive.
  $\theta - \theta = 0 = 0 \times 2\pi$ donc $\theta\mathscr{R}\theta$.
- symétrique.
  Supposons que $\theta\mathscr{R}\theta'$, donc $\theta - \theta' = 2k\pi$
  donc $\theta' - \theta = -2k\pi$ donc $\theta'\mathscr{R}\theta$.
- transitive.
  Supposons que $\theta\mathscr{R}\theta'$ et $\theta'\mathscr{R}\theta''$ donc $\theta - \theta' = 2k\pi$ et $\theta' - \theta'' = 2k'\pi$.
  donc $\theta - \theta'' = (\theta - \theta') + (\theta' - \theta'') = 2k\pi + 2k\imath\pi = 2(k + k')\pi$.
  et donc $\theta\mathscr{R}\theta''$.

Les classes d'équivalence peuvent représentées par les arguments principaux. En fait, on retrouve la même classe d'équivalence que précédemment.

Proposition - Caractéristique par les classes d'équivalence
Soient $\mathscr{R}$ une relation d'équivalence sur un ensemble E, a et b deux éléments de E. Alors
$$a\mathscr{R}b \Leftrightarrow C(a) = C(b)\,.$$

Démonstration
Supposons que $a\mathscr{R}b$.
Soit $x \in C(b)$, alors $x\mathscr{R}b$ et par transitivité (et réflexivité) : $x\mathscr{R}a$, donc $x \in C(a)$.

$\forall\, x \in C(b)$, $x \in C(a)$, donc $C(b) \subset C(a)$.
Et, si $y \in C(a)$ alors $y \mathscr{R} a$, donc $y \mathscr{R} b$, donc $y \in C(b)$. Et $C(a) \subset C(b)$.
Bilan : $C(a) = C(b)$.
Réciproquement, si $C(a) = C(b)$, alors $a \in C(a) \subset C(b)$, donc $a \in C(b)$ et $a \mathscr{R} b$. □

Exercice
On note $\mathscr{P}$ le plan usuel. On définit sur $\mathscr{P} \times \mathscr{P}$ la relation $\mathscr{R}$ par

$$(A,B)\mathscr{R}(C,D) \Leftrightarrow ABDC \text{ est un parallélogramme.}$$

Il s'agit d'une relation d'équivalence. Que représentent les classes d'équivalence de cette relation ?

Correction
Les classes d'équivalence sont données par l'ensemble de tous les vecteurs (pas d'origine).

Définition - Système de représentants
Soit $\mathscr{R}$ une relation d'équivalence définie sur un ensemble E.
Si il existe $S \subset E$ tel que pour tout $x \in E$, $\exists\, ! s \in S$ tel que $x \mathscr{R} s$,
on dit que S est un système de représentants de la classe d'équivalence $\frac{E}{\mathscr{R}}$.
(Il y a une bijection naturelle entre S et $\frac{E}{\mathscr{R}}$).
On note souvent $S_{E\setminus\mathscr{R}}$ un tel système.
Si E est fini, il existe toujours un système de représentants (sinon cela peut nécessiter l'axiome du choix).

Remarque - Notation floue
La notation $S_{E\setminus\mathscr{R}}$ est très imprécise (pas de différence entre un système et un autre).
Souvent ce qui compte pour les démonstrations est d'un considérer un système quelconque, sans précision

Exemple - Relation de congruence modulo $[0,2\pi]$
L'argument principal d'un nombre complexe noté Arg$z$ est choisi dans l'ensemble des représentants $[0,2\pi[$ (ou $]-\pi,\pi]$ selon les conventions).
Il représente sa classe d'équivalence des nombres réels associé à un même argument d'un nombre complexe.

5.3. Partition de E

Heuristique - Classe d'équivalence : partition de E
Avoir une relation d'équivalence, c'est faire l'assimilation entre différents objets à priori différents et finalement identique (ou plutôt équivalent) du point de vue de la relation. L'ensemble du départ est alors réduit en partie plus petite, ces éléments sont les classes d'équivalence. Elles forment une partition de l'ensemble initial.

Définition - Partition d'un ensemble
Une partition de E est un ensemble de sous-ensembles (non vides) de E tel que :
— leur réunion fait E
— leur intersection 2 à 2 est vide

Proposition - Partition de E
Si $\mathscr{R}$ est une relation d'équivalence sur E, alors ses classes d'équivalence forment une partition de E.

Démonstration
On note O_i, la famille des classes d'équivalence (la notion de famille est vue en fin de chapitre).
Pour tout $x \in E$, x appartient à sa propre classe d'équivalence, donc $x \in \bigcup_i O_i$,

donc $E \subset \bigcup_i O_i$.

Réciproquement, toutes les classes d'équivalences O_i sont des parties de E.

Donc $\bigcup_i O_i \subset E$.

Finalement $E = \bigcup_i O_i$.

Soit O_i et O_j deux classes d'équivalence.

Si $x \in O_i \cap O_j$, alors tous les éléments de O_i sont en relation avec x donc $O_i = C(x)$.
De même tous les éléments de O_j sont en relation avec x donc O_j est la classe de x.
Donc ou bien $O_i = O_j (= C(x))$, ou bien $O_i \cap O_j = \emptyset$. □

Remarque - La réciproque est vraie
Etant donnée une partition sur E : $E = \bigcup_{i \in I} O_i$.
Considérons alors $\mathscr{R}$: $(x \mathscr{R} y)$ ssi $\exists\, i \in I$ tel que $x \in O_i$ ET $y \in O_i$.
$\mathscr{R}$ est une relation d'équivalence sur E.

Remarque - Classes d'équivalence et dénombrement
Si E fini se décompose en classe d'équivalence $(O_i)_{i \in I}$, alors $\#E = \sum_{i \in I} \#O_i$.
Il arrive souvent que E se décompose en n classes d'équivalence toutes de même cardinal c. Dans ce cas : $\#E = c \times n$.

Application - Dénombrement et classe d'équivalence
La relation $\mathscr{R}$ définie sur $F \times G$ par :

$$(a,b)\mathscr{R}(c,d) \iff a = c$$

est une relation d'équivalence.
Les classes d'équivalence sont en nombre de $\#F$. Et chacune des classes possède exactement $\#G$ éléments.
Donc $\#(F \times G) = \sum_{i \in I} \#O_i = \sum_{a \in F} \#G = \#G \times \sum_{a \in F} 1 = \#G \times \#F$.
Il s'agit vraiment de la formalisation du premier résultat de dénombrement...

6. Bilan

Synthèse

⇝ Les ensembles, dont on a vu qu'ils étaient à la base des raisonnements mathématiques, peuvent être « travaillés ». Ils peuvent être coupés en morceaux, avec des classes d'équivalence ou bien structurés visuellement avec une relation d'ordre.

⇝ Selon chaque situation, on fait évoluer notre regard !

Savoir-faire et Truc & Astuce du chapitre

— Savoir-faire - Montrer que $\mathscr{R}$ est une relation d'ordre
— Savoir-faire - Montrer que $a = \sup E$
— Savoir-faire - Montrer que $\mathscr{R}$ est une relation d'équivalence

Notations

	Propriétés	*Remarques*
définition (du terme de gauche)		autre notation : $\overset{\text{def.}}{s} =$
équivalence variées	Réflexive, Symétrique, Transitive	
les classes d'équivalence sur E tion d'équivalence $\mathscr{R}$	Les classes d'équivalence forment une partition de E	
représentants de $\frac{E}{\mathscr{R}}$	$S_{E\setminus\mathscr{R}} \subset E$ et $\forall\, x \in E, \exists !\, s \in S_{\frac{E}{\mathscr{R}}}$ tel que $x\mathscr{R}s$.	
ordre variées	Réflexive, Antiymétrique, Transitive	
élément de E pour UNE relation e a priori	$\max E \in E$ et $\forall\, x \in E,\ x \leqslant \max E$ (ou $\subset$.....)	Si la relation est totale, $\max E$ est au plus unique.
lément de E pour UNE relation e a priori	$\min E \in E$ et $\forall\, x \in E,\ \min E \leqslant x$ (ou $\subset$.....)	Si la relation est totale, $\min E$ est au plus unique.
es majorants de E	$\forall\, x \in E,\ x \leqslant \sup E$ et $[\forall\, x \in E, x \leqslant y \Rightarrow \sup E \leqslant y$	Au plus unique.
des minorants de E	$\forall\, x \in E,\ \inf E \leqslant x$ et $[\forall\, x \in E, y \leqslant x \Rightarrow y \leqslant \inf E$	Au plus unique.

Retour sur les problèmes

54. Voir cours
55. Voir cours
56. Voir cours
57. Pas de plus grand élément à $[0,1[$. Mais un plus petit élément à tous ses majorants : $\sup[0,1[= 1$, on l'appelle la borne supérieure.
58. Si les applications sont bien définies.
    Alors une relation $\mathscr{R}$ est une application f de E^2 dans $\{0,1\}$, avec $f((a,b)) = 1$ ssi $a\mathscr{R}b$.
    Si les relations sont bien définies.
    Alors une application f de E sur F est définie par $f(a) = b$ ssi $a\mathscr{R}b$.

Quatrième partie

Arithmétique & Structures élémentaires

Chapitre 13

Groupes

Résumé -

Comme structure algébrique, nous devons étudier les groupes (une seule loi, interne), les anneaux et corps (deux lois internes) et les espaces vectoriels (deux lois internes et une externe). Nous parlerons sûrement à l'occasion d'algèbre et sûrement d'espace affine.
Le chapitre qui suit est assez court, et se concentre sur la structure de groupes.
La notion de groupe a été formalisée au début du XIX-ième siècle mais n'a trouvé toute sa force qu'à la fin du siècle. Nous préparerons quelques notions de seconde année. Et nous reprendrons cette notion en étudiant au second semestre le groupe des permutations, et ensuite nous verrons agir des groupes sur les matrices carrées.

- *Michaël Launay - Structure algébrique - https ://www.youtube.com/watch ?v=RaqlxOihGxw*
- *PoincareDuality - « Les maths ne sont qu'une histoire de Groupe » Poincaré par Etienne Ghys - https ://www.youtube.com/watch ?v=dLwi_opxLxs*
- *Interview Cirm - Claire voisin - https ://www.youtube.com/watch ?v=vcwMTpgNlQA*

Sommaire

1. Problèmes

? Problème 61 - Structure

En MPSI, la recherche de la démonstration, optimale la plupart du temps conduit à réfléchir précisément sur les hypothèses qui permet d'obtenir les résultats. Ainsi les objets sont épurés sur les hypothèses principales.
De quel ensemble et avec quelle loi (structure) minimale doit-on partir pour l'essentiel de nos théorèmes à l'origine ?

? Problème 62 - Résolution des équations polynomiales

Comment démontrer qu'en règle générale, un polynôme de degré 5 n'admet pas de formules explicites des racines du polynôme ?
C'est l'une des questions qui a conduit GALOIS à créer (inventer, découvrir) la notion de groupe... Quel chemin l'a conduit à cette construction ?

Histoire - Evariste Galois

La vie (très courte et très romantique) d'Evariste Galois (1811-1832) pourrait faire la base d'un excellent film...
On le présente souvent comme le premier, avec de nombreuses années d'avance, qui a compris le rôle fondamental de la structure de lois internes et de groupe. Il a ainsi démontré que pour $\geqslant 5$, il n'existe pas de formule explicite et avec des racines n^e exprimant les racines d'un polynôme quelconque de degré n.

? Problème 63 - Groupe de Poincaré

En physique relativiste, les éléments de l'espace sont des quadruplets (x,y,z,t).
On appelle groupe de Poincaré l'ensemble $(G,\circ)$ des transformations $\varphi : \mathbb{R}^4 \to \mathbb{R}^4 \in G$ tel que pour tout paire de quadruplets (x,y,z,t) et (x',y',z',t'), la distance est conservée :
$(x-x')^2+(y-y')^2+(z-z')^2-c^2(t-t')^2 = (X-X')^2+(Y-Y')^2+(Z-Z')^2-c^2(T-T')^2$
où $\varphi(x,y,z,t) = (X,Y,Z,T)$ et $\varphi(x',y',z',t') = (X',Y',Z',T')$.
Montrer qu'il s'agit bien d'un groupe.

? Problème 64 - Mariage chez les Murngin (citation)

A New York, [Lévi-Strauss] s'était lancé dans un vaste travail de sociologie théorique qui devint sa thèse de doctorat (aujourd'hui célèbre) sur les structures ?élémentaires de la parenté. Un jour, dans l'étude d'un certain type de mariage, il se heurta à des difficultés inattendues et pensa qu'un mathématicien pourrait lui venir en aide.
D'après ce qu'ont observé les sociologues travaillant ? sur le terrain ?, les lois de mariage des tribus indigènes d'Australie comportent un mélange de règles exogamiques et endogamiques dont la description et l'étude posent des problèmes combinatoires parfois compliqués. Le plus souvent le sociologue s'en tire par l'énumération de tous les cas possibles dans l'intérieur d'un système donné. Mais la tribu des Murngin, à la pointe Nord de l'Australie, s'était donné un système d'une telle ingéniosité que Lévi-Strauss n'arrivait plus à en dérouler les conséquences. En désespoir de cause il me soumit son problème.
Le plus difficile pour le mathématicien, lorsqu'il s'agit de mathématique appliquée, est souvent de comprendre de quoi il s'agit et de traduire dans son propre langage les données de la question. Non sans mal, je finis par voir que tout se ramenait à étudier deux permutations et le groupe qu'elles engendrent. Alors apparut une circonstance imprévue. Les lois de mariage de la tribu Murngin, et de beaucoup d'autres, comportent le principe suivant : ? Tout homme peut épouser la fille du frère de sa mère ? ou, bien entendu, l'équivalent de celle-ci dans la classification matrimoniale de la tribu. Miraculeusement, ce principe revient à dire que les deux permutations dont il s'agit sont échangeables, donc que le groupe qu'elles engendrent est abélien. Un système qui à première vue menaçait

> d'être d'une complication inextricable devient ainsi assez facile à décrire dès lors qu'on introduit une notation convenable. Je n'ose dire que ce principe a été adopté pour faire plaisir aux mathématiciens, mais j'avoue qu'il m'en est resteune certaine tendresse pour les Murngin.
> Oeuvres scientifiques, André Weil, 1979 p. 567-568

2. Lois de composition internes

2.1. Définitions

Définition - Loi de composition interne (Magma)
Une loi de composition interne sur un ensemble E est une application de $E \times E$ dans E : $\Phi : E \times E \to E, (x, y) \mapsto x \star y$.
On note, pour $(x, y, z) \in E^3$,
$$(x \star y) \star z = \Phi(\Phi(x, y), z) \text{ et } x \star (y \star z) = \Phi(x, \Phi(y, z)).$$
Un tel couple $(E, \star)$ est appelé un magma.
Quand la loi interne est clairement identifiée, par abus, on peut dire que E est un magma sans précision supplémentaire.

Exemple - $\mathbb{N}$
$(\mathbb{N}, +)$ ou $(\mathbb{N}, \times)$ sont des magmas.

Définition - Caractéristiques
On dit que le magma $(E, \star)$:
- est *commutatif* si $\forall (x, y) \in E \times E,\ x \star y = y \star x$
- est *associatif* si $\forall (x, y, z) \in E \times E \times E,\ x \star (y \star z) = (x \star y) \star z$
- est unifère ou possède un élément *neutre* s'il existe $e \in E$ tel que $\forall x \in E,\ e \star x = x \star e = x$ (e est alors l'élément neutre.)

Pour x élément de E, on dit qu'un élément y de E est un *symétrique* ou un *inverse* de x pour $\star$ si $x \star y = y \star x = e$, ($e$ neutre de E)

Définition - Monoïde
Un magma $(M, \star)$ associatif et unifère est appelé un monoïde.

Remarque - Notations
Plusieurs remarques

1. Les lois de composition interne sont usuellement notées $\star$, $\perp$, $\top$, $+$, $\times$, en notation multiplicative $x \star y = xy$.
2. La notation additive est usuellement réservée à une loi commutative et associative, dans ce cas le symétrique de x (s'il existe) est noté $-x$.
3. Lorsque la loi est commutative et associative, on peut écrire :
$$\sum_{i=1}^{n} x_i = x_1 + \cdots + x_n \text{ (notation additive) } ou \prod_{i=1}^{n} x_i = x_1 \dots x_n \text{ (notation multiplicative).}$$
4. Si la loi $\star$ est associative, on peut écrire $x^{\star n} = x \star \cdots \star x$ (n termes x).
En notation multiplicative on obtient ainsi x^n et en notation additive nx.

Définition - Distributivité
Supposons que l'ensemble E est muni de deux lois internes $\star$ et $\top$.
On dit que $\star$ est *distributive* par rapport à la loi interne $\top$ si :
$\forall (x, y, z) \in K^3,\ x \star (y \top z) = (x \star y) \top (x \star z)$ (distibutive à gauche)
$\forall (x, y, z) \in K^3,\ (x \top y) \star z = (x \star z) \top (y \star z)$ (distibutive à droite)

2.2. Propriétés directes

Proposition - Unicités (éléments neutres, symétrique)... si existence
Soit $(F,\top)$ un magma.
Si F est unifère, l'élément neutre pour $\top$ est unique.
Soit $(E,\star)$ un monoïde.
Si $x \in E$ admet un symétrique alors celui-ci est unique ;
Si $x, y \in E$ admettent des symétriques x^{-1} et y^{-1}
alors $x \star y$ admet un symétrique : $y^{-1} \star x^{-1}$.
Si x est symétrique alors x est régulier (à gauche et à droite) :
$\forall\, y,z \in E,\ x \star y = x \star z \Rightarrow y = z$ et $y \star x = z \star x \Rightarrow y = z$.

Démonstration
On va faire chacun des cas
- Supposons que e_1 et e_2 soit deux éléments neutres de $(F,\top)$.
  Alors comme e_2 est neutre : $e_1 = e_1 \top e_2$ et $e_1 \top e_2 = e_2$
  car e_1 est neutre.
  Donc $e_1 = e_2$. Il y a au plus un élément neutre.
- Si y_1 et y_2 deux symétriques de x : $y_1 = y_1 \star e = y_1 \star (x \star y_2) = (y_1 \star x) \star y_2 = e \star y_2 = y_2$
- $(y^{-1} \star x^{-1}) \star (x \star y) = y^{-1} \star (x^{-1} \star x) \star y = y^{-1} \star e \star y = y^{-1} \star y = e$
  car $\star$ est associative. De même : $(x \star y) \star (y^{-1} \star x^{-1}) = x \star (y \star y^{-1}) \star x^{-1} = x \star e \star x^{-1} = e$.
- Si $x \star y = x \star z$, alors en starisant par x^{-1} (x symétrique) à gauche : $y = x^{-1} \star x \star y = x^{-1} \star x \star z = z$.
  De même à droite.

□

2.3. Induction

Il s'agit de l'induction de la loi sur une partie ou une restriction d'un magma E.

Définition - Loi induite
Soit $A \subset E$, avec $(E,\top)$ magma.
A est stable par $\top$ (loi de composition interne sur E) si $\forall x, y \in A,\ x \top y \in A$.
$\top_A = \top_{|A\times A}$ s'appelle la loi induite (par $\top$ sur A).

Remarque - Transmission des propriétés
$\top_A$ est alors une loi interne sur A, i.e. $(A,\top_A)$ est un magma (induit).
Si $\top$ est commutative (resp. associative), $\top_A$ est nécessairement commutative (resp. associative).
En revanche l'élément neutre de $\top$ ou l'élément symétrique de $x \in A$ pour $\top$ (s'ils existent) ne sont pas nécessairement dans A.

3. Structure de groupe

3.1. Définition et propriétés

Un groupe est un monoïde dont tous les éléments sont inversibles :

Définition - Groupe
On appelle groupe un ensemble G muni d'une loi de composition interne $\top$ vérifiant :
- la loi $\top$ est associative ;
- G possède un élément neutre pour $\top$;
- tout élément x de G possède un symétrique pour $\top$ (ou tout élément de G est inversible, est symétrisable).

Si de plus la loi $\top$ est commutative, on dit que le groupe est abélien (ou

commutatif).

Exemple - Groupes des racines de l'unité

L'ensemble $(\mathbb{U},\times)$ est un groupe.

• En effet, $(\mathbb{C},\times)$ est associatif donc en prenant trois éléments z_1,z_2,z_3 de $\mathbb{U}$, donc de $\mathbb{C}$, on a $z_1\times(z_2\times z_3)=(z_1\times z_2)\times z_3$). Ainsi $(\mathbb{U},\times)$ est associative

• Si $z_1,z_2\in\mathbb{U}$, alors $|z_1z_2|=|z_1||z_2|=1$ donc $z_1z_2\in\mathbb{U}$.

• $1\in\mathbb{U}$ est un élément neutre de $\mathbb{U}$: $1\times z=z\times 1=z$.

• Soit $z\in\mathbb{U}$, alors $\frac{1}{z}=\overline{z}\in\mathbb{U}$. Pour les mêmes raisons, puisque si $z\in\mathbb{U}_n$, $\overline{z}\in\mathbb{U}_n$, on peut dire que $(\mathbb{U}_n,\times)$ est un groupe.

Remarque - Sous-groupe

Finalement, on a utilisé ici le fait que $(\mathbb{C},\times)$ était lui même un groupe, que $\mathbb{U}$ est stable par la loi induite, que $1_{\mathbb{C}}\in\mathbb{U}$ et pour tout $z\in U(\subset\mathbb{C})$, $\frac{1}{z}(\in\mathbb{C})\in\mathbb{U}$.

On reviendra sur ces propriétés plus loin.

Histoire - Plusieurs naissances

En fait, cela est comme toujours plus subtil. Il semble que l'idée de groupe est en germe à la fin du XVIII siècle (en particulier chez Lagrange) et donc le concept apparait clairement ensuite mais à plusieurs endroits en même temps : chez Galois en France, dans les écrits de Cayley en Grande-Bretagne ou dans les oeuvres de Dedekind en Allemagne.

Les définitions étant parfois légèrement différentes... (citation : http ://images.math.cnrs.fr/Un-concept-mathematique-trois)

Proposition - Régularité

Dans un groupe tous les éléments sont réguliers à gauche et à droite

Démonstration

Tous les éléments sont symétriques donc réguliers d'après un point précédent □

Comme des groupes sont des magmas unifère, où tous les éléments sont inversibles par définition :

Théorème - Existence et unicité

Soit $(G,\top)$ un groupe. Alors :

— L'élément neutre est unique.

— Tout élément possède un unique symétrique.

— En notant x^{-1} le symétrique (l'inverse) de x, on a $(x^{-1})^{-1}=x$.

— $(x\top y)^{-1}=y^{-1}\top x^{-1}$.

3.2. Groupes produits

Définition - Produit de groupes

Soient $(G,\perp)$ et $(H,\top)$ deux groupes.

On appelle groupe produit (de ces deux groupes) le groupe $(G\times H,\star)$ tel que

$$\forall\,(x_1,y_1),(x_2,y_2)\in G\times H,\qquad (x_1,y_1)\star(x_2,y_2)=(x_1\perp x_2,y_1\top y_2)$$

Pour aller plus loin - Produit direct. Produit indirect

En réalité, le produit comme il apparait ici est souvent qualifié de produit direct.

Le produit semi-direct est alors le produit de décomposition d'un groupe G sous la forme $G=HK$, où H est un sous-groupe normal ou distingué de G et K est isomorphe à $\frac{G}{H}$.

Il s'agit bien d'un groupe :

Démonstration

• Il est associatif. Pour tout $(x_1,y_1),(x_2,y_2),(x_3,y_3)\in G\times H$,

$$(x_1,y_1)\star\big((x_2,y_2)\star(x_3,y_3)\big)=(x_1,y_1)\star(x_2\perp x_3,y_2\top y_3)=(x_1\perp(x_2\perp x_3),y_1\top(y_2\top y_3))=(x_1\perp x_2\perp x_3,y_1\top y_2\top_3)$$

$$\big((x_1,y_1)\star(x_2,y_2)\big)\star(x_3,y_3)\big)=(x_1\perp x_2,y_1\top y_2)\star(x_3,y_3)=\cdots=(x_1\perp x_2\perp x_3,y_1\top y_2\top_3)$$

• Son élément neutre est (e_G,e_H), car pour tout $(x,y)\in G\times H$,

$$(e_G,e_H)\times(x,y)=(e_g\perp x,e_H\top y)=(x,y)=(x\perp e_g,y,\top e_h)=(x,y)\star(e_g,e_H)$$

• Tous ses éléments sont inversibles, pour tout $(x,y)\in G\times H$,

$$(x^{-1},y^{-1})\times(x,y)=(x^{-1}\perp x,y^{-1}\top y)=(e_g,e_H)=\cdots=(x,y)\star(x^{-1},y^{-1})$$

□

Remarque - Souvent $G=H$

On considère donc le groupe produit noté$(G^2,\perp)$ plutôt que $(G^2,\perp^2)$.

Et par récurrence, on peut étendre le produit de groupes à plus de deux groupes.

3.3. Exemples

Groupes triviaux

Exemple - Avec l'addition
Les ensembles de nombres munis de l'addition : $(\mathbb{Z},+),(\mathbb{Q},+),(\mathbb{R},+),(\mathbb{C},+)$ sont des groupes commutatifs.

Exemple - Avec la multiplication
Les ensembles de nombres munis de la multiplication : $(\mathbb{Q}^*,\times),(\mathbb{R}^*,\times),(\mathbb{C}^*,\times)$ sont des groupes commutatifs.
Moins trivialement : $(\mathbb{U}_n,\times)$ est un groupe.

Ensemble $\frac{\mathbb{Z}}{n\mathbb{Z}}$

Définition - Ensemble des classes d'équivalence modulo n
Soit $n\in\mathbb{N}$, fixé.
La relation $\equiv_n$ ou encore $\cdot\equiv\cdot[n]$ est une relation d'équivalence sur $\mathbb{Z}$.
L'ensemble des classes d'équivalence associées est noté $\dfrac{\mathbb{Z}}{n\mathbb{Z}}$.
Un système de représentant est $[\![0,n-1]\!]$, puisque $\dfrac{\mathbb{Z}}{n\mathbb{Z}}=\{\overline{0},\overline{1},\dots\overline{n-1}\}$,
où pour tout $k\in[\![0,n-1]\!]$, $\overline{k}=\{k,k+n,k+2n\dots k-n\dots\}=\{k+rn,r\in\mathbb{Z}\}$.
On peut alors définir sur $\dfrac{\mathbb{Z}}{n\mathbb{Z}}$ les lois $\overline{+}$ et $\overline{\times}$ par :

$$\overline{h}\overline{+}\overline{k}=\overline{h+k}\qquad \overline{h}\overline{\times}\overline{k}=\overline{h\times k}$$

Remarque - Le plus dur dans ce qui précède
est de démontrer que les lois $\overline{+}$ et $\overline{\times}$ sont bien définies.
C'est-à-dire qu'elles sont indépendantes des représentants de $\overline{k}$ et $\overline{h}$ choisis.
Nous ferons cette démonstration dans le cours d'arithmétique

Proposition - Groupe $\left(\dfrac{\mathbb{Z}}{n\mathbb{Z}},\overline{+}\right)$
Pour tout entier n, $\left(\dfrac{\mathbb{Z}}{n\mathbb{Z}},\overline{+}\right)$ est un groupe commutatif.
Son élément neutre est $\overline{0}$, et l'opposé de $\overline{k}$ est $\overline{n-k}$.

Démonstration
Pour tout $h,k\in\mathbb{Z}$, $\overline{h}\overline{+}\overline{k}=\overline{h+k}=\overline{k+h}=\overline{k}\overline{+}\overline{h}$.
Pour tout $k\in\mathbb{Z}$, $\overline{k}\overline{+}\overline{0}=\overline{k+0}=\overline{k}$ $\quad\overline{k}\overline{+}\overline{n-k}=\overline{k+n-k}=\overline{n}=\overline{0}$. □

Analyse - Et pour la multiplication $\overline{\times}$?
En revanche, il existe des diviseurs de 0, donc des éléments non inversibles, pour la multiplication modulo n : les diviseurs de n.
Mais si k est premier avec n, alors d'après le théorème de Bézout :

$$\exists\, u,v\in\mathbb{Z}\mid uk+vn=1\qquad \overline{u}\overline{\times}\overline{k}=\overline{uk}=\overline{uk+vn}=\overline{1}$$

Proposition - Groupe $\left(\dfrac{\mathbb{Z}}{p\mathbb{Z}},\overline{\times}\right)$ avec p premier
Pour tout nombre premier p, $\left(\dfrac{\mathbb{Z}}{p\mathbb{Z}},\overline{\times}\right)$ est un groupe commutatif.
Son élément neutre est $\overline{1}$, et l'opposé de $\overline{p}$ est obtenu en exploitant le théorème de Bézout.

On peut donc obtenir l'inverse de $\overline{k}$ en exploitant l'algorithme d'Euclide.

Exercice

A démontrer

Correction

Pour tout $h,k\in\mathbb{Z}$, $\overline{h}\overline{\times}\overline{k}=\overline{h\times k}=\overline{k\times h}=\overline{k}\overline{\times}\overline{h}$.

Pour tout $k\in\mathbb{Z}$, $\overline{k}\overline{+}\overline{1}=\overline{k\times 1}=\overline{k}$.

Pour l'existence de l'inverse, voir l'analyse précédente.

Remarque - Autre point de vue

Dans le cours sur les anneaux, nous définirons le groupe $\left(\frac{\mathbb{Z}}{n\mathbb{Z}}^{\star},\overline{\times}\right)$, valable pour tout entier n, en ne considérant que les classes inversibles, modulo n.

Dans le cas où n est premier, on retrouve le groupe précédent.

« Petits » groupes

Comme les groupes sont réguliers, il ne peut y avoir chaque lettre ne peut apparaite plus d'une fois par ligne et par colonne.

Remplir ces tableaux, c'est comme remplir un carré latin (nom mathématique des sudoku).

Analyse - Groupe à deux éléments

Essayons de construire un groupe à deux éléments : a et b.

L'un des deux éléments est le neutre, supposons qu'il s'agisse de a.

On a donc $a^2=a$ et $ab=ba=b$.

b est inversible, nécessairement $b^2=a$.

Ce groupe se résume dans le tableau :

$\nearrow_\top$	a	b
a	a	b
b	b	a

C'est le seul groupe à deux éléments. Notons qu'il est commutatif.

Mais surtout remarquons qu'il s'incarne ou a de nombreux avatars :

- a = fonction croissante, b = fonction décroissante et $\top=\circ$
- $a=1$ (ou classe des nombres positifs), $b=-1$ (ou classe des nombres négatifs) et $\top=\times$
- $a=0$ (ou classe des entiers pairs), $b=1$ (ou classe des entiers impairs) et $\top=+$ dans $\frac{\mathbb{Z}}{2\mathbb{Z}}$
- $a=\mathrm{id}_{\mathbb{C}}$ et b = symétrie (ou involution quelconque) et $\top=\circ$
- ...

Exercice

Construire un (le) groupe de 3 éléments

Correction

Par essais/erreurs, il s'agit de

$\nearrow_\top$	a	b	c
a	a	b	c
b	b	c	a
c	c	a	b

Dans la construction, on voit une notion importante : à cause de la régularité, il n'est pas possible de trouver sur une même ligne ou une même colonne deux valeurs identiques.

Par suite, dans chaque ligne et chaque colonne, on retrouve toutes les valeurs de G.

Une incarnation possible est $\frac{\mathbb{Z}}{3\mathbb{Z}}$.

Groupes matriciels

L'ensemble des matrices $(\mathcal{M}_{n,p}(\mathbb{K}),+)$ est un groupe. Il nous intéresse assez peu.

L'ensemble $(\mathcal{M}_n(\mathbb{K}),\times)$ n'est pas un groupe! (c'est un monoïde) Tous les éléments ne sont pas inversibles.

En revanche

- $(GL_n(\mathbb{K}),\times)$, ensemble des matrices inversibles est un groupe appelé, le groupe linéaire. Par définition : $GL_n(\mathbb{K})=\{M\in\mathcal{M}_n(\mathbb{K})\mid\exists N\in\mathcal{M}_n(\mathbb{K})\text{ tel que }M\times N=N\times M=I_n\}$

Pour aller plus loin - Au début d'année...

Nous avons vu apparaître à plusieurs reprises des calculs avec $\{1,j,j^2\}$ (expression des racines d'un polynôme de degré 3 ou bien calcul de la somme $S_k=\sum_{i\equiv k[3]}^{n}\binom{n}{i}$...). Cela était nécessaire, car il s'agit de la meilleure incarnation du groupe (unique) à trois éléments

Pour aller plus loin - Groupes à 4 éléments

Il existe deux groupes à 4 éléments : $D_4\simeq\{1,i,-1,-i\}\simeq\frac{\mathbb{Z}}{4\mathbb{Z}}$ et V_4 le groupe de Klein.

$V_4:=$

$\nearrow_\top$	a	b	c	d
a	a	b	c	d
b	b	a	d	c
c	c	d	a	b
d	d	c	b	a

Il s'incarne par exemple dans le groupe des symétries par rapports aux médiatrices d'un triangle équilateral et de la transformation identité.

$a:[BCD]\to[BCD]$, $b:[BCD]\to[BDC]$, $c:[BCD]\to[DCB]$ et $d:[BCD]\to[CBD]$

— $(\mathcal{O}_n(\mathbb{K}),\times)$, ensemble des matrices orthogonales est un groupe appelé, le groupe orthogonal.
Par définition : $\mathcal{O}_n(\mathbb{K}) = \{M \in \mathcal{M}_n(\mathbb{K}) \mid M^T \times M = I_n\}$

Groupes des permutations

Un groupe très important, on reviendra sur cette notion plus tard...

Proposition - Groupe des permutations d'un ensemble
Soit X un ensemble non vide. On note S_X l'ensemble des permutations de X (c'est-à-dire des bijections de X dans X). Alors $(S_X,\circ)$ est un groupe, généralement non commutatif, appelé groupe des permutations de X.

Remarque - Permutation
Qu'est-ce qu'une permutation de X ?
Il s'agit, par σ de changer « l'ordre » ou notre regard sur tous les éléments de X.
Donc cela signifie que

$$\{\sigma(x), x \in X\} = X$$

et donc σ est surjective. Il suffit de montrer que σ est injective : $\sigma(i) = \sigma(j) \Rightarrow i = j$..

Démonstration
Point par point :
— Pour tout $\sigma_1,\sigma_2 \in S_X$,

$$\forall\, x \in X, (\sigma_1 \circ \sigma_2)(x) = \sigma_1(\underbrace{\sigma_2(x)}_{\in X}) \in X$$

On notera que $\sigma_1 \circ \sigma_2$ est bien une bijection de X.
Donc $\circ$ est bien une loi de composition interne de S_X.
— Puis cette loi est associative.
Pour tout $\sigma_1,\sigma_2,\sigma_3 \in S_X$,

$$\forall\, x \in X, (\sigma_1 \circ (\sigma_2 \circ \sigma_3))(x) = \sigma_1(\sigma_2(\sigma_3(x))) = ((\sigma_1 \circ \sigma_2) \circ \sigma_3)(x)$$

— S_X possède un élément neutre : $id_X : x \mapsto x$:
$id \in S_X$ et pour tout $\sigma \in S_X$: $\forall\, x \in X$, $id(\sigma(x)) = \sigma(x) = \sigma(id(x))$.
— Last but not least : on note pour tout $\sigma \in S_X$,

$$\sigma' : x \mapsto y \text{ tel que } x = \sigma(y)$$

Clairement $\sigma^{-1} \circ (\sigma(y)) = \sigma^{-1}(x) = y$ et $\sigma \circ \sigma^{-1}(x) = \sigma(y) = x$,
donc σ est bien inversible et d'inverse σ^{-1}.
Mais il faudrait vérifier que σ^{-1} ainsi définie existe bien. Pour cela, il est nécessaire que tout $x \in X$ soit aussi l'image d'un y par σ, ou encore que $\sigma(X) = X$.
Or σ est une permutation de X, donc cette hypothèse est bien vérifiée (cf. remarque).
□

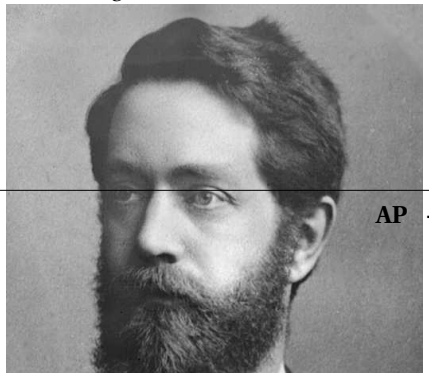
Ce groupe non commutatif sera longuement étudié plus tard dans l'année.

Groupes et géométrie

Proposition - Groupes des similitudes directes du plan $\mathbb{C}$
L'ensemble des similitudes directes est un groupe pour la loi $\circ$.
L'élément neutre est l'identité.
L'inverse de la similitude de centre Ω, d'angle θ et de rapport k est la similitude de centre Ω, d'angle $-\theta$ et de rapport $\frac{1}{k}$.
L'inverse de la translation de vecteur $\vec{u}$ est la translation de vecteur $-\vec{u}$.

Démonstration
Nous sommes bien obligé de considérer les translations car la composition de deux rotations de centre A et B respectivement, d'angle θ et $2\pi-\theta$ et de rapports 1 est la translation de vecteur $2\overrightarrow{AB}$.
Rappelons que la similitude de centre $\Omega(\omega)$, d'angle θ et de rapport k est $z \mapsto ke^{i\theta}(z-\omega)+\omega$.
A partir de cette notation, on peut tout démontrer. □

Pour aller plus loin - Programme d'Erlangen
Felix Klein (1849-1925) est un mathématicien allemand qui proposa dans le programme d'Erlangen de re« voir » toute les géométries en étudiant les groupes de symétrie qui agisse sur l'espace en question...
Il a eu une grosse influence sur Henri Poincaré.



4. Sous-groupe

4.1. Définition et caractérisations

Par la suite, on considérera $(G, \top)$ un groupe.

Définition - Sous-groupe
$H \subset G$ (non vide) est un sous-groupe de G si H est stable pour la loi interne et si la loi induite (restriction de la loi à H) munit H d'une structure de groupe. On note $H < G$

Proposition - Elément neutre et symétriques
Soit H un sous-groupe de G.
Alors l'élément neutre de H est l'élément neutre de G.
Si $x \in H$, le symétrique de x dans H est le symétrique de x dans G.

Démonstration
Soit e_1 l'élément neutre de H et e celui de G. $\underbrace{e_1 \top e_1}_{\in H} = e_1 = \underbrace{e \top e_1}_{\in G}$
Et comme e_1 est régulier (dans G) : $e_1 = e$.
Puis si on note x' l'inverse de x dans H, $x \top x' = e = x \top x^{-1}$
Par régularité :
en « composant à gauche » par x^{-1}, inverse de x dans G (par la suite, tout se passe dans G)

$$x' = x^{-1} \top x \top x' = x^{-1} \top x \top x^{-1} = x^{-1}$$

□

Théorème - Caractérisation 1
Soit $H \subset G$. H est un sous-groupe de G si et seulement si il vérifie :
- $H \neq \emptyset$
- $\forall (x, y) \in H^2, x \top y \in H$
- $\forall x \in H, x^{-1} \in H$

Théorème - Caractérisation 2
Soit $H \subset G$. H est un sous-groupe de G si et seulement si il vérifie :
- $H \neq \emptyset$
- $\forall (x, y) \in H^2, x \top y^{-1} \in H$
  (en notation multiplicative : $\forall (x, y) \in H^2, xy^{-1} \in H$;
  en notation additive : $\forall (x, y) \in H^2, x - y \in H$)

Pour aller plus loin - Le « monstre »
Il s'agit du plus gros groupe fini « connu ». On l'appelle le Monstre M ou groupe de Fischer-Griess F_1. Son ordre (= son cardinal ici) est $246 \times 320 \times 59 \times 76 \times 112 \times 133 \times 17 \times 19 \times 23 \times 29 \times 31 \times 41 \times 47 \times 59 \times 71 = 808017424794512875886459904961710757005754\,368\,000\,000\,000 \approx 8 \times 10^{53}$.
C'est bien un nombre fini (mais c'est gros). Il a été découvert (est-ce le bon verbe?) ou construit en 1980

Savoir faire - Démontrer que H est un (sous-)groupe
Dans la pratique, lorsque H est une partie de E, groupe. On démontre qu'il s'agit d'un sous-groupe de E
- avec la caractérisation 1, lorsque H et E sont explicites
- avec la caractérisation 2, lorsque H et E sont théoriques

Démonstration
On a trois propositions équivalentes à démontrer, nous allons faire en trois temps

$$A_1 \Longrightarrow A_2 \Longrightarrow A_3 \Longrightarrow A_1$$

1. Supposons que H est un sous-groupe de G.
   - Nécessairement H n'est pas vide.
   - H est stable pour la loi interne donc $\forall (x, y) \in H^2, x \top y \in H$
   - Et le dernier point : $\forall x \in H, x^{-1} \in H$ à été démontré dans la proposition précédente.
2. Supposons que H vérifie les trois assertions de la caractérisation 1.

Pour aller plus loin - Action de groupe et représentation
Étant donné un groupe G, dont la loi est notée multiplicativement et dont l'élément neutre est noté e, on peut définir une action (ou opération) de G sur un ensemble E par une application : $G \times E \to E$, $(g, x) \mapsto g \cdot x$ *vérifiant les propriétés suivantes :* $\forall\, x \in E,\ e \cdot x = x$ *et* $\forall\, g, g' \in G, \forall\, x \in E,\ g' \cdot (g \cdot x) = (g \top g') \cdot e$.
De (très) nombreuses situations de présence de groupe sont de cette forme là, avec un ensemble E donné.
Réciproquement, si on connait très bien E sur lequel agit G, alors on apprend à connaître G.
La notion d'orbite, de permutations de E,

— $H \neq \emptyset$
— $\forall (x,y) \in H^2$, $z = y^{-1} \in H$, d'après la troisième assertion.
Puis $x \top z = x \top y^{-1} \in H$ d'après la seconde assertion.
Donc $\forall (x,y) \in H^2$, $x \top y^{-1} \in H$.

3. Supposons que H vérifie les deux assertions de la caractérisation 2.
— H n'est pas vide, on doit montrer qu'il est stable pour la loi interne.
La loi $\top$ est nécessairement associative sur $H \subset G$.
Soit $x \in H$, alors $x \in G$, et $e = x \top x^{-1} \in H$ d'après la seconde assertion.
En outre, tout élément $x \in H \subset G$ possède un symétrique x^{-1} dans G, qui est aussi dans H :

$$e, x \in H \Longrightarrow e \top x^{-1} = x^{-1} \in H$$

□

Exercice
Montrer que si $H_1 < (G_1, \perp)$ et $H_2 < (G_2, \top)$ alors $H_1 \times H_2$ est un sous-groupe du groupe produit $(G_1 \times G_2, \star)$.

Correction
On va exploiter la seconde caractérisation.
Puisque $e_1 \in H_1$, $e_2 \in H_2$, alors $(e_1, e_2) \in H_1 \times H_2$ et donc $H_1 \times H_2$ est non vide.
Puis si $(x_1, x_2), (y_1 y_2) \in H_1 \times H_2$, alors

$$(x_1, x_2) \star (y_1, y_2)^{-1} = (x_1, x_2) \star (y_1^{-1}, y_2^{-1}) = (\underbrace{x_1 \perp y_1^{-1}}_{\in H_1}, \underbrace{x_2 \perp y_2^{-1}}_{\in H_2}) \in H_1 \times H_2$$

Donc $H_1 \times H_2$ est un sous-groupe du groupe produit $(G_1 \times G_2, \star)$.

4.2. Intersection

Théorème - Intersection de deux sous-groupes
Soit H et K deux sous-groupes de $(G, \top)$.
Alors $H \cap K$ est un sous-groupe de G.

Démonstration
$e \in H$ et $e \in K$, donc $H \cap K$ n'est pas vide.
Soit $x, y \in H \cap K$, alors $xy^{-1} \in H$ et $xy^{-1} \in K$, donc $xy^{-1} \in H \cap K$. □

L'exercice suivant donne TOUS les sous-groupes de $(\mathbb{Z}, +)$:
Exercice

1. Soit $a \in \mathbb{Z}$. Montrer que $a\mathbb{Z}$ est un sous-groupe de $(\mathbb{Z}, +)$.
2. Soit G un sous-groupe de $(\mathbb{Z}, +)$, $G \neq \{0\}$.
Justifier que $G \cap \mathbb{N}^*$ a un plus petit élément $a > 0$.
Montrer que $G = a\mathbb{Z}$ (utiliser la division euclidienne).

Correction

1. $0 \in a\mathbb{Z}$ et si $x, y \in a\mathbb{Z}$, alors $x - y = ka - ha = (k-h)a \in a\mathbb{Z}$.
2. $G \cap \mathbb{N}^* \subset \mathbb{N}^*$, donc il a un plus petit élément, noté $a > 0$.
On a nécessairement $a\mathbb{Z} \subset G$ Soit $x \in G$, on applique la division euclidienne : $x = pa + q$.
Or $q = x - pa \in G$, par stabilité et donc $q \in G$, mais $q \in [\![0, a[\![$, donc $q = 0$, sinon, on a une contradiction avec la définition de a.
Donc $x = pa$ et $G \subset a\mathbb{Z}$.

La démonstration s'adpate à une infinité de sous-groupe.

Théorème - Intersection de sous-groupes
Soit $(H_i)_{i \in I}$ une famille de sous-groupes de $(G, \top)$. Alors $\bigcap_{i \in I} H_i$ est un sous-groupe de G.

Démonstration
$\forall\ i \in I$, $H_i < G$, donc $\bigcap_{i\in I} H_i \subset G$.
$\forall\ i \in I$, $e \in H_i$, car H_i est un groupe. Donc $e \in \bigcap_{i\in I} H_i$, qui n'est pas vide, donc.
Soit $x, y \in \bigcap_{i\in I} H_i$, alors pour tout $i \in I$, $x, y \in H_i$ et $xy^{-1} \in H_i$
Donc $xy^{-1} \in \bigcap_{i\in I} H_i$. □

4.3. Sous-groupe engendré

Première manipulation (trouver un candidat)

Analyse - Plus petit groupe contenant une partie A de G

Considérons une partie A (ensemble) d'un groupe G.
On note $\mathscr{A} = \{H < G \mid A \subset H\}$, l'ensemble de tous les sous-groupes de G contenant A.
Alors $\mathscr{A}$ est non vide car $G \in \mathscr{A}$.
On se souvient, qu'en théorie des ensembles, on a vu que $\inf(A, B) = A \cap B$
(inf est ici la borne inférieure pour la relation d'ordre de l'inclusion, i.e. le plus grand ensemble des plus petits ensembles que A et B).
Considérons alors $< A >:= \bigcap_{H\in\mathscr{A}} H$, l'intersection (non dénombrable, peut-être) de tous les sous-groupe contenant A.
Cette intersection est un sous-groupe et elle contient A.
En fait $< A >$ est le plus petit élément de $\mathscr{A}$.

Définition - Groupe engendré
Soit $(G, \top)$ un groupe. Soit A une partie de G.
On appelle groupe engendré par A, le plus petit sous-groupe de G, parmi les sous-groupes de G contenant A.
On le note $< A >$. On a donc $< A >= \bigcap_{H\in\mathscr{A}} H = \min \mathscr{A}$
(où $\mathscr{A} = \{H < G \mid A \subset H\}$).

Il faut démontrer que $\bigcap_{H\in\mathscr{A}} H$ est bien le plus petit sous-groupe de G contenant A.

Démonstration
Comme, pour tout $H \in \mathscr{A}$, $A \subset H$, on a bien $A \subset \bigcap_{H\in\mathscr{A}} H$.
Par ailleurs, $\bigcap_{H\in\mathscr{A}} H$ est une intersection de sous-groupes de G. Donc il forme bien d'un sous-groupe de G.
Il reste à démontrer que c'est le plus petit.
Soit K, un sous-groupe de G contenant A. Alors nécessairement $K \in \mathscr{A}$.
Donc K fait partie des sous-groupes donc l'intersection définie $< A >$,
on a donc $< A >= K \cap \left(\bigcap_{H\in\mathscr{A}\setminus K} H \right) \subset K$. □

Proposition - Croissance de l'engendrement
Si $A \subset B$ sont deux parties d'un groupe G.
Alors $< A > \subset < B >$

Démonstration
$< B >$ est le plus petit sous-groupe contenant B donc A car $A \subset B$.
Donc $< B >$ est un sous-groupe de G contenant A.
Mais $< A >$ est le plus petit des sous-groupe contenant A, donc il est plus petit que B : $< A > \subset < B >$. □

Application - Réflexes

Si A est un sous-groupe de G qui est contenu dans B, alors nécessairement $< A >= A \subset < B >$.
Si B est un sous-groupe de G qui est contient A, alors nécessairement $< A > \subset < B >= B$.

Savoir faire - Comment trouver le sous-groupe engendré par une partie A?

Il faut

1. *Pré-sentir* la bonne description (efficace) de ce sous-groupe. On donne alors un nom à cet ensemble K.
2. Montrer que K est bien un groupe, qui contient A
3. Montrer que K est nécessairement entièrement inclus dans $< A >$ ou dans tout sous-groupe de $\mathscr{A}$.

Comme $< A >$ est le plus petit sous-groupe contenant A, propriété vérifiée par K, alors $< A >= K$

Croissance par engendrement (deuxième point de vue)

Attention - Nom contradictoire?

Ce nom semble contradictoire. Par groupe engendré on entend plutôt quelque chose qui s'agrandit (pour l'inclusion) à partir de A, alors qu'il s'agit visiblement de quelque chose qui diminue à partir de G.
A-t-on la même chose?

Analyse - Engendrement

On aurait plutôt envie de considérer $\overline{A} := \{a_1 \star a_2 \star \cdots \star a_k \mid k \in \mathbb{N}, a_1, a_2, \cdots a_k \in \mathbb{N}\}$,
ensemble de tous les éléments obtenables à partir de répétition de produit (fini) de A.
C'est bien en ensemble engendré par A. En fat :
• cet ensemble contient bien A, avec $n = 1$ et a_1 décrivant tous les éléments de A. • cet ensemble est un groupe A, avec $n = 2$ et $a_2 = a_1^{-1}$, on a $e_G \in \overline{A}$.
si $a_1 \star a_2 \star \cdots \star a_k$, $b_1 \star b_2 \star \cdots \star b_m \in \overline{A}$, alors,
$(a_1 \star a_2 \star \cdots \star a_k) \star (b_1 \star b_2 \star \cdots \star b_m)^{-1} = a_1 \star a_2 \star \cdots \star a_k \star b_m^{-1} \star b_{m-1}^{-1} \star \cdots \star b_1^{-1} \in \overline{A}$.
Enfin, il contient tous les sous-groupe H contenant A,
car pour tout $i \in \mathbb{N}_k$, $a_i \in A \subset H$. H est un groupe donc $a_1 \star a_2 \star \cdots \star a_k \in H$.

Exemple - Groupes engendré par p dans $\mathbb{Z}$.

On considère le groupe $(\mathbb{Z}, +)$. Le groupe $< p >$ contient nécessairement tous les nombres de la forme : $\underbrace{p + p + \cdots + p}_{r \text{ fois}} = rp$, c'est à dire les multiples de $\mathbb{Z}$.
Donc $p\mathbb{Z} \subset < p >$.
Par ailleurs $p\mathbb{Z}$ est un sous-groupe de $\mathbb{Z}$. Donc $< p >= p\mathbb{Z}$.

Exercice
Quel est le sous-groupe engendré par $\{p, q\}$ dans $(\mathbb{Z}, +)$?

Correction
On pense qu'il s'agit du sous-groupe engendré par le PGCD de p et q, dont on a déjà un nom : $(p \wedge q)\mathbb{Z}$.
On sait déjà que c'est un groupe.
Comme $p \wedge q$ est un diviseur de p et de q, $p \in (p \wedge q)\mathbb{Z}$ et $q \in (p \wedge q)\mathbb{Z}$. Donc $\{p, q\} \subset (p \wedge q)\mathbb{Z}$.
Le groupe $< \{p, q\} >$ contient tous les nombres de la forme $up + vq$ où $u, v \in \mathbb{Z}$.
D'après le théorème de Bézout, il contient donc $p \wedge q$, le PGCD de p et q.
Comme il s'agit d'un groupe, il contient également tous les multiples de $p \wedge q$, donc tout $(p \wedge q)\mathbb{Z}$.
Donc $(p \wedge q)\mathbb{Z} \subset < \{p, q\} >$.
Et comme $< p \wedge q >$ est le plus petit sous-groupe contenant $\{p, q\}$: $< p, q > (= < \{p, q\} >) = (p \wedge q)\mathbb{Z}$

Groupe monogène

Définition - Groupe monogène
On dit que G est un groupe monogène, s'il existe $x \in G$ tel que $G =< x >$.
Dans ce cas $G = \{x^k, k \in \mathbb{Z}\}$.

Savoir faire - Etudier des groupes monogènes
Si G est un groupe que l'on sait monogène, alors il existe $x \in G$ (référence, que l'on fixe) tel que $G =< x >$.
L'application $\varphi : \mathbb{Z} \to G$, $k \mapsto x^k$ est bien définie, surjective. Elle peut être injective ou non (cas fini).
On transfère ensuite par φ (ou φ^{-1}) l'étude de G, à partir de propriétés de $\mathbb{Z}$.

Exercice
Montrer qu'un groupe monogène est nécessairement abélien

Correction
Soit $(G, \times)$ monogène, donc il existe $x \in G$ tel que $G =< x >$.
Soient $a, b \in G$, alors il existe $n, m \in \mathbb{Z}$ tel que $a = x^n$ et $b = x^m$.
Ainsi $a \times b = x^n \times x^m = x^{n+m} = x^m \times x^n = b \times a$

Exemples à partir de $\mathbb{U}$

L'exercice suivant nous aide à faire le point.

Exercice
On considère le groupe $(\mathbb{U}, \times)$.

1. Soit $z_k = e^{\frac{2i\pi}{k}}$. Que vaut le groupe $< z_k >$?
2. Avec les mêmes notations, que vaut le groupe $< z_r, z_s >$?
3. A-t-on pour tout $n \in \mathbb{N}$, pour tout $z \in \mathbb{U}_n$, $\mathbb{U}_n =< z >$?
   Sinon, à quelle condition sur z, a-t-on : $\mathbb{U}_n =< z >$?
4. Quel est le groupe $\mathbb{U}_n \cap \mathbb{U}_m$?
5. $\mathbb{U}$ est-il monogène ?

Pour aller plus loin - De qui parle Felix Klein ?
« Depuis longtemps déjà, il s'occupait à étudier des groupements de racines complexes de l'unité sur la base de sa théorie des racines primitives. Et voilà que ce matin-là (le 30 mars 1796) en se réveillant, il lui apparut clairement qu'à partir de sa théorie, on pouvait construire un polygone à 17 cotés[...]. Cet événement marqua un grand tournant dans sa vie : c'est précisément ce jour-là qu'il décida d'abandonner les langues pour se consacrer exclusivement aux mathématiques ». Les mathématiques y ont beaucoup gagné !

Correction

1. $< z_k >$ continent toutes les puissances de z_k, donc tous les nombres complexes z_k^r de $\mathbb{U}_k$.
   Donc $\mathbb{U}_k \subset < z_k >$ et $\mathbb{U}_k$ est un groupe, ainsi $< z_k >= \mathbb{U}_k$.
2. Tous les nombres complexes de la forme $z_r^u \times z_s^v = \exp 2i\pi \frac{us+vr}{sr}$ appartiennent à $< z_r, z_s >$, avec $u, v \in \mathbb{Z}$.
   D'après le théorème de BÉZOUT, il existe $u, v \in \mathbb{Z}$ tel que $us + vr = s \wedge r$, puis $sr = (s \wedge r)(s \vee r)$
   donc $z_{s \vee r} \in < z_r, z_s >$.
   Et donc $< z_{r \vee s} > \subset < z_r, z_s >$.
   Par ailleurs, $r | (r \vee s)$, donc il existe r' tel que $r \vee s = rr'$ et donc $z_{r \vee s}^{r'} = z_r$ donc $z_r \in < z_{r \vee s} >$.
   De même $z_s \in < z_{r \vee s} >$ et donc $\{z_r, z_s\} \subset < z_{r \vee s} >$, donc $< z_r, z_s > \subset < z_{r \vee s} >$.
   Par double inclusion $< z_r, z_s >=< z_{r \vee s} >$.
3. Non, clairement. En effet, pour tout $n \in \mathbb{N}$, $1 \in \mathbb{U}_n$ et $< 1 >= \{1\} \neq \mathbb{U}_n$.
   Lorsque $z \in \mathbb{U}_n$ vérifie $< z >= \mathbb{U}_n$, on dit que z est une racine *primitive* première de l'unité.
   On sait que $z \in \mathbb{U}_n$ signifie qu'il existe $k \in [\![0, n-1]\!]$ tel que $z = \exp \frac{2ik\pi}{n}$.
   Si $k \wedge n = 1$. Alors d'après BÉZOUT (toujours) : il existe $u, v \in \mathbb{Z}$ tel que $uk + vn = 1$
   et donc $z^u = \exp \frac{2iuk\pi}{n} \times \exp 2iv\pi = \exp \frac{2i(uk+vn)\pi}{n} = z_n$.
   Donc $z_n \subset < z >$ et donc $\mathbb{U}_n =< z_n > \subset < z >$.
   Comme $z \in \mathbb{U}_n$. On a donc $< z > \subset < \mathbb{U}_n >= \mathbb{U}_n$.
   Par double inclusion $\mathbb{U}_n =< z >$.
   Réciproquement, si $\mathbb{U}_n =< z >$, alors il existe $u \in \mathbb{Z}$ tel que $z^u = z_n (\in \mathbb{U})$
   et donc $uk \equiv 1[n]$ i.e. $k \wedge n = 1$ (réciproque de BÉZOUT).
4. Soit $z \in \mathbb{U}_{n \wedge m}$, alors $z^{n \wedge m} = 1$. Et comme $n \wedge m | n$,
   on a donc $z^n = (z^{n \wedge m})^{n/(n \wedge m)} = 1$ donc $z \in \mathbb{U}_n$. De même $z \in \mathbb{U}_m$.
   Par conséquent : $\mathbb{U}_{n \wedge m} \subset \mathbb{U}_n \cap \mathbb{U}_m$.
   Evidemment, si $z \in \mathbb{U}_n \cap \mathbb{U}_m$, alors pour tout $u, v \in \mathbb{Z}$, $z^{un+vp} = (z^n)^u (z^m)^v = 1$.
   Donc $z^{n \wedge m} = 1$ et $z \in \mathbb{U}_{n \wedge m}$.
   Par double inclusion $\mathbb{U}_n \cap \mathbb{U}_m = \mathbb{U}_{n \wedge m}$
5. Non, sinon $\mathbb{U}$ serait dénombrable.

4.4. Démontage d'un groupe

Théorème de Lagrange

Proposition - Relation d'équivalence modulo un sous-groupe
Soit $(G,*)$ un groupe et $H<G$, un sous-groupe de G.
On note $\mathscr{R}_H$, la relation définie sur G par :

$$a\mathscr{R}_H a' \iff a^{-1}*a' \in H$$

Alors $\mathscr{R}_H$ est une relation d'équivalence.

Démonstration
$\mathscr{R}_H$ est :
- réflexive : pour tout $a\in G$, $a^{-1}a=e\in H$. Donc $\forall\ a\in G$, $a\mathscr{R}_H a$.
- symétrique : si $a\mathscr{R}_H a'$, alors $a^{-1}*a'\in H$.
  Mais H est un groupe donc $(a^{-1}a')^{-1}=a'^{-1}a\in H$ donc $a'\mathscr{R}_H a$
- transitive : si $a\mathscr{R}_H a'$ et $a'\mathscr{R}_H a''$, alors $a^{-1}*a', a'^{-1}*a''\in H$.
  Mais H est un groupe donc $(a^{-1}*a')*(a'^{-1}*a''=a^{-1}*a''\in H$ donc $a\mathscr{R}_H a''$

□

Remarque - Lien relation d'équivalence et sous-groupe
On voit sur cette démonstration le lien très étroit qui unit les caractéristiques d'un sous-groupe et les propriétés d'une relation d'équivalence. Plus précisément : l'élément neutre à la réflexivité, l'inversibilité à la symétrie et la stabilité à la transitivité
Quand on a une relation d'équivalence, on a naturellement une décomposition en réunion disjointe

Proposition - Décomposition de G
Soit H un sous-groupe de G.
On note $S:=S_{\frac{G}{\mathscr{R}_H}}$ un système de représentant des classes d'équivalences de $\mathscr{R}_H$.
Alors $G=\uplus_{a\in S}\overline{a}$, la réunion disjointes des classes de a.
$\overline{a}$ n'est pas un groupe, mais il est en bijection avec H.
Par la suite, on notera aH, cet ensemble. On a donc $aH=a'H\Leftrightarrow a\mathscr{R}_H a'\Leftrightarrow a^{-1}a'\in H$
Le produit cartésien $H\times S$ et le groupe G sont en bijection (c'est la décomposition).

Démonstration
La première partie de la proposition a été donnée dans le cours sur les classes d'équivalence.
Soit $\varphi : H\mapsto\overline{a},\ x\mapsto ax$.
Montrons que φ est bien définie à valeurs dans $\overline{a}$.
On a, en effet, $a\mathscr{R}_H ax=\varphi(x)$, car $a^{-1}*ax=x\in H$.
φ est bijective car elle admet une application réciproque $\varphi^{-1}: b(\in\overline{a})\mapsto a^{-1}b(\in H)$.
On a bien $\varphi^{-1}(b)=a^{-1}b\in H$, car $b\in\overline{a}$, donc $a\mathscr{R}_H b$, i.e. $a^{-1}b\in H$.
Pour l'équivalence.
On considère $g\in G$, alors il existe un unique $a\in S$ tel que $\overline{g}=\overline{a}$,
i.e. il existe un unique $h\in H$ tel que $g=ah$.
On a alors une bijection naturelle $(a,h)\mapsto g=ah$
dont la réciproque est donnée sur la ligne précédente. □

En fait, les ensembles $\overline{a}$ sont comme des sous-groupes affines de G...

Proposition - Théorème de LAGRANGE
Si H un sous-groupe de G, groupe de cardinal fini, alors $\mathrm{card}H|\mathrm{card}G$.

Démonstration
$G = \uplus_{a\in S}\overline{a}$. La réunion est disjointe, donc en terme de cardinaux :

$$\text{card}(G) = \sum_{a\in S}\text{card}(aH) = \sum_{a\in S}\text{card}(H) = \text{card}(H)\times\sum_{a\in S}1 = \text{card}(H)\times\text{card}(S)$$

où on a exploité que H et aH sont en bijection, donc ont même cardinaux.
On aurait pu aussi exploiter le fait que $\text{card}(G) = \text{card}(H\times S) = \text{card}H\times\text{card}(S)$. □

Sous-groupe distingué

Analyse - S comme un groupe ?

On a dit et redit que sauf pour $\overline{a} = H$ (exemple : $a = e$, mais pas uniquement), $\overline{a} = aH$ n'est pas un groupe et donc φ n'est pas un morphisme de groupe (cf partie suivante).
Néanmoins, on peut se demander si S, l'ensemble des représentants ne pourrait pas être un groupe. Sous quelles condition ?
Pour que cela ait du sens, on devrait pouvoir créer l'opération $\overline{*}$ sur S de la façon suivante :

$$aH\overline{*}bH = (a*b)H$$

Dans ce cas, il faudrait que pour tout $a' = ax \in aH$ (avec $x\in H$) et $b' = by \in bH$ (avec $y\in H$), on ait : $a'*b' \in (a*b)H$, c'est-à-dire $axby \in (ab)H$.
Il suffit alors que pour tout $x\in H$, $xb\in bH$, ce que l'on note $Hb\subset bH \Leftrightarrow Hb = bH$.
Cette condition est également nécessaire : car elle doit être vérifiée pour tout $a\in G$ et donc $a = e$ (et $y = e$) aussi : $xb\in bH$.

Définition - Sous-groupe distingué
On dit que $H < G$ est un sous-groupe distingué (ou normal) de G si

$$\forall\ a\in G, \forall\ x\in H, \qquad a^{-1}xa\in H$$

On peut retenir que pour tout $a\in G$, $aH = Ha$.
On note alors $H\lhd G$

Démonstration
Montrons que : $\forall\ a\in G, x\in H, a^{-1}xa\in H$ correspond bien à ce que l'on souhaite.
Supposons donc que $\forall\ a\in G, x\in H, a^{-1}xa\in H$.
Soit $y\in aH$. Alors $\exists\ x\in H$ tel que $y = ax$.
Alors $ya^{-1} = axa^{-1}\in H$, donc $y = \underbrace{(axa^{-1})}_{z\in H}a\in Ha$. Ainsi $aH\subset Ha$.
Soit $y\in Ha$. Alors $\exists\ x\in H$ tel que $y = xa$.
Alors $a^{-1}y = a^{-1}xa\in H$, donc $y = a\underbrace{(a^{-1}xa)}_{z\in H}\in aH$. Ainsi $Ha\subset aH$.
Supposons que pour tout $a\in G$, $aH = Ha$.
Soit $a\in G$, $x\in H$. Alors $xa\in Ha = aH$, donc $\exists\ z\in H$ tel que $xa = az$ et donc $z = a^{-1}xa\in H$.
□

Proposition - Groupe quotient
Soit $(G,*)$ un groupe.
Si $H\lhd G$ est un sous-groupe distingué de G, alors $S = \dfrac{G}{\mathscr{R}_H}$, souvent noté $\dfrac{G}{H}$ est un groupe pour la loi $\overline{*}$ définie par $aH\overline{*}bH = (a*b)H$.

Exercice
A démontrer ! (Attention, ce n'est pas un sous-groupe. Il faut donc tout redémontrer à commencer par la bonne définition de la loi...)

Correction

Exemple - Groupe trivial

$H = \{e\}$ est un sous-groupe distingué. Mais cela n'a pas beaucoup d'intérêt.
Sauf si on l'associe avec la réciproque d'un morphisme (cf partie suivante).

Exemple - G abélien

Si G est abélien, alors tout sous-groupe H est distingué.
En effet, pour tout $a \in G$, $x \in H$: $a^{-1}xa = a^{-1}ax = x \in H$.
Ainsi, pour tout $n \in \mathbb{N}$, $n\mathbb{Z} \lhd \mathbb{Z}$ et donc $(\frac{\mathbb{Z}}{n\mathbb{Z}}, \overline{+})$ est un groupe.

Pour aller plus loin - Théorie de résolution par radicaux
Comme écrit wikipédia : « Pour n supérieur ou égal à 5, le groupe alterné sur n éléments $\mathcal{A}_n$ est simple. Ce résultat est à la base de la théorie de la résolution par radicaux. »

Définition - Groupe simple
On dit qu'un groupe est simple lorsqu'il ne possède pas de sous-groupe distingué autre que $\{e\}$ et lui-même

Cela ne vous rappelle pas une autre définition ?
Exercice
Soit G un groupe de cardinal p, premier.
Montrer que G est simple

Correction
Soit $H \lhd G$. Alors $\text{card}H | \text{card}G$, donc $\text{card}H \in \{1, p\}$, i.e. $H = \{e\}$ ou $H = G$ respectivement.
Donc G est nécessairement simple

5. Morphismes de groupes

5.1. Définition et propriété immédiate

Soient $(G, \star)$ et $(G', \top)$ deux groupes.

Définition - Morphisme de groupes
Une application f de G dans G' vérifiant :
$$\forall (x, y) \in G^2,\ f(x \star y) = f(x) \top f(y).$$
est appelé morphisme (de groupes) de $(G, \star)$ sur $(G', \top)$.

Proposition - Conservation du noyau
Soit $f : G \to G'$ un morphisme de groupes. Alors $f(e_G) = e_{G'}$.

Démonstration
Soient $x \in G$,
$$f(x) \top e_{G'} = f(x) = f(x \star e_G) = f(x) \top f(e_G)$$
Puis par régularité : $e_{G'} = e_G$ □

Proposition - Image de l'inverse
Soit $f : G \to G'$ un morphisme de groupes.
Alors pour tout $x \in G$, $f(x^{-1}) = \big(f(x)\big)^{-1}$

Démonstration
Soient $x \in G$,
$$f(x^{-1}) \top f(x) = f(x^{-1} \star x) = f(e_G) = e_{G'} = f(x)^{-1} \top f(x)$$
Donc par régularité : $f(x^{-1}) = \big(f(x)\big)^{-1}$ □

Exemple - exp
Le morphisme du groupe $(\mathbb{R}, +)$ vers $(\mathbb{R}^*, \times)$ est l'application :
$$\forall\ a, b \in \mathbb{R}, \qquad f(a+b) = f(a) \times f(b)$$
On a nécessairement $f(0) = 1$.
On montre :

1. par récurrence : $\forall\ n \in \mathbb{N}$, $f(n) = \big(f(1)\big)^n$
2. par passage au symétrique : $\forall\ n \in \mathbb{Z}$, $f(n) = \big(f(1)\big)^n$
3. pour $\mathbb{Q}$: pour tout $m \in \mathbb{Z}$, $n \in \mathbb{N}^*$,
   $(f(1))^m = f(m) = f(n \times \frac{m}{n}) = f(\frac{m}{n} + \cdots + \frac{m}{n}) = (f(\frac{m}{n}))^n$.
   donc pour tout $\forall\ q \in \mathbb{Q}$, $f(q) = \big(f(1)\big)^q$
4. si f est continue ou croissante, alors on peut passer à la limite $f : x \mapsto a^x$ avec $a = f(1)$

5.2. Image et noyau d'un morphisme

Proposition - Sous-groupe
Soit $f : G \to G'$ un morphisme de groupes.
Soit A un sous-groupe de G, alors $f(A)$ est un sous-groupe de G'.
Soit B un sous-groupe de G', alors $f^{-1}(B)$ est un sous-groupe de G.
En particulier :
- Im $f = f(G) = \{f(x), x \in G\}$ est un sous-groupe de G', appelé image de G
- $\operatorname{Ker} f = f^{-1}(e_{G'}) = \{x \in G \mid f(x) = e_{G'}\}$ est un sous-groupe de G, appelé noyau de f.

Exemple - $\mathbb{Z} \to \mathbb{U}_n$
$f : (\mathbb{Z}, +) \to (\mathbb{U}_n, \times)$, $k \mapsto e^{ki\pi/n}$ est un morphisme de groupe.
Son noyau est $n\mathbb{Z}$.

Démonstration
$f(A) \subset G'$, non vide ($e_G \in A$, donc $e_{G'} = f(e_G) \in f(A)$).
Soient $y_1, y_2 \in G'$, il existe $x_1, x_2 \in A$ tel que $f(x_1) = y_1$ et $f(x_2) = y_2$.
$y_1 \top y_2^{-1} = f(x_1) \top f(x_2)^{-1} = f(x_1) \top f(x_2^{-1}) = f(x_1 \star x_2^{-1}) \in f(A)$ Donc $f(A)$ est bien un sous-groupe de G'.
$f^{-1}(B) \subset G$, non vide ($e_{G'} \in B$, donc $e_G \in f^{-1}(B)$ car $f(e_G) = e_{G'}$).
Soient $x_1, x_2 \in G'$, alors $f(x_1) \in B$ et $f(x_2) \in B$. Ainsi : $f(x_1 \star x_2^{-1}) = f(x_1) \top f(x_2)^{-1} \in B$ Donc $f^{-1}(A)$ est bien un sous-groupe de G. □

Exercice
Montrer que $f : G \to G'$ morphisme de groupes est :
- surjective ssi Im $f = G'$
- injective ssi $\operatorname{Ker} f = \{e_G\}$

Correction
La première équivalence est évidente, simple écriture.
Pour la seconde,
on a si f injective,
pour tout $x \in \operatorname{Ker} f$, $f(x) = e_{G'} = f(e_g)$, donc $x = e_g$ et $\operatorname{Ker} f \subset \{e_G\}$.
L'inclusion réciproque est triviale, donc $\operatorname{Ker} f = \{e_G\}$
Réciproquement, si $\operatorname{Ker} f = \{e_G\}$.
Si $f(x) = f(x')$, alors $e_{G'} = f(x) \top f(x')^{-1} = f(x \star x'^{-1})$, donc $x \star x'^{-1} \in \operatorname{Ker} f = \{e_G\}$.
Ainsi $x \star x'^{-1} = e_G$, i.e. $x = x'$. Et f est injective

Exemple - $\operatorname{Ker} f$ **est un sous-groupe distingué**
On sait que $\operatorname{Ker} f < G$.
Soit $a \in G$ et $x \in \operatorname{Ker} f$,
$$f(a^{-1}xa) = f(a^{-1}) \underbrace{f(x)}_{=e_{G'}} f(a) = f(a)^{-1} f(a) = f(a^{-1}a) = f(e) = e_{G'}$$
Donc $a^{-1}xa \in \operatorname{Ker} f$.

5.3. Premier théorème d'isomorphisme

Théorème - Premier théorème
Soient G et G' deux groupes et $f : G \to G'$, un morphisme de groupes.
Alors f induit un isomorphisme de groupes : $\hat{f} : G/\operatorname{Ker} f \hookrightarrow\!\!\!\rightarrow \operatorname{Im} f$.

Démonstration
$\operatorname{Ker} f$ est un sous-groupe distingué de G.
$G/\operatorname{Ker} f$ est la notation du groupe quotient $\dfrac{G}{\operatorname{Ker} f}$ pour la relation d'équivalence $a \mathscr{R} b$ ssi $ab^{-1} \in \operatorname{Ker} f$ i.e $f(a) = f(b)$.
Soit donc $\overline{a}$, un élément de $G/\operatorname{Ker} f$ (c'est la classe de a) et $\hat{f} : \overline{a} \mapsto f(a)$.
Il faut donc commencer par vérifier que cette définition de $\hat{f}$ a bien un sens.
Compte-tenu de la définition de $\mathscr{R}$, c'est évident que $\hat{f}$ est bien définie.
$\hat{f}$ est clairement bijective, puisque $\hat{f}^{-1}$ est l'application qui $z \in \operatorname{Im} f \to \overline{z}$. □

Pour aller plus loin - Une deuxième, un troisième?
Le deuxième énonce :
si $N \lhd G$ et $H < G$, alors $N \cap H \lhd H$ et $H/(N \cap H) \hookrightarrow\!\!\!\rightarrow HN/N$.
Le troisième énonce :
si $N, M \lhd G$ tels que $M < N$, alors $N/M \lhd G/M$ et $(G/M)/(N/M) \hookrightarrow\!\!\!\rightarrow G/N$.

Remarque - Notation symbolique ou formelle
On retrouve dans la littérature (mathématique) le diagramme suivant que l'on qualifie de commutatif :

$$G \twoheadrightarrow \frac{G}{\operatorname{Ker} f} \hookrightarrow \operatorname{Im} f \hookrightarrow G'$$

Pour aller plus loin - Dans les espaces vectoriels...
Ce théorème est à comparer avec le lemme préparatoire qui conduit au si fondamental théorème du rang dans les espaces vectoriels de dimension finie.
Ainsi si f est une application linéaire de E sur F, avec E de dimension finie.
Alors $\dim E = \dim Im\, f + \dim \operatorname{Ker} f$ *car* $f_{|H} : H \to Im\, f$ *est une application bijective (où H tel que $E = H \oplus Ker\, f$.*

6. Bilan

Synthèse

- ⇝ La notion de groupe est la brique élémentaire des théories mathématique. C'est une notion primitive : un ensemble et une loi interne associative, unifère et dont tous les éléments sont symétriques. Cette structure est naturellement comparable à une relation d'équivalence : associativité↔ transitivité, élément neutre↔reflexivité et inversion↔symétrie.
- ⇝ On peut réduire des (sous-)groupes par intersection, ou bien générer des (sous-)groupes par engendrement de parties. Ces deux méthodes sont très classiques en algèbre ou en topologie.
- ⇝ On décompose ensuite les groupes en produit de sous-groupes à condition que le premier de ce produit soit un sous-groupe distingué. Finalement les sous-groupes distingués (ou normaux) sont comme des nombres premiers.
- ⇝ On peut aussi déplacer des structures avec des morphismes de groupes, voire comparer des groupes (est-ce que ces morphismes sont bijectives (isomorphisme) ?)

Savoir-faire et Truc & Astuce du chapitre

- — Savoir-faire - Démontrer que H est un (sous-)groupe
- — Savoir-faire - Comment trouver le sous-groupe engendré par une partie A ?
- — Savoir-faire - Etudier des groupes monogènes

Notations

Notations	*Définitions*	*Propriétés*	*Remarques*
$H < G$	H est un sous-groupe de G	$e_G \in H$ et $\forall\, x, y \in H,\ xy^{-1} \in H$	Parfois noté : $H \lhd G$
$H \lhd G$	H est sous-groupe distingué de G	$H \lhd G$ et $\forall\, x \in G,\ xHx^{-1} \subset H$	Parfois noté : $H \lhd \vert G$
$\frac{G}{H}$	Ensemble des classes d'équivalence sur G pour la relation d'équivalence $a\mathscr{R}_H b \iff ab^{-1} \in H$	Les classes d'équivalence sont chacune en bijection avec H.	On en déduit le théorème de LAGRANGE : $\operatorname{card} H \vert \operatorname{card} G$

Retour sur les problèmes

59. Groupes, anneaux, corps...
60. Pas facile. Il faut plonger dans un cours sur le corps de Galois. Lorsqu'on aura fait le cours sur le groupe symétrique (= groupe des permutations d'un ensemble à n éléments), cela sera peut-être plus simple... En gros une formule de résolution, c'est une décomposition du groupe des permutations (qui agit sur l'ensemble des racines) en sous-groupe distingué × groupe quotient. Si une telle décomposition n'est pas possible (le groupe est simple), nous sommes bloqués. Or pour les équations de degré 5, on regarde S_5 (cf second semestre), il se décompose en $A_5 t \times \{-1, 1\}$, mais A_5 ne se décompose pas plus...
61. La composition conserve nécessairement la distance.
La question fondamentale est dans l'existence de la transformation

inverse (bijectivité de φ).
A noter qu'il s'agit d'un groupe continue (ou de LIE) par opposition aux groupes discrets (et souvent finis comme S_n).

Chapitre 14

Construction d'ensembles numériques : des entiers à la droite réelle

Résumé -

Ce chapitre est comme une application du chapitre e sur les relations pour donner une assise mathématique satisfaisante aux ensembles bien connus des élèves car largement utilisés. Dans l'histoire, ces constructions se sont passés à la fin du XIX siècle. Cela faisait des siècles (voire des millénaires) que certains de ces ensembles étaient exploités...

Il n'y a au fond qu'un seul problème : comment donner du sens aux ensembles : $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{D}$, $\mathbb{Q}$, $\mathbb{R}$ *et* $\mathbb{C}$ *?*

Quelques vidéos sur internet :

- *Yvan Monka - Ils sont fous, ces nombres ! - Classification - https ://www.youtube.com/watch ?v=kL-eMNZiARM*
- *Exo7Maths - Nombres réels - https ://www.youtube.com/watch ?v=NCWWVven9Cs*
- *Science4all - La diagonale dévastatrice de Cantor - https ://www.youtube.com/watch ?v=xqSKawORrPo*

Sommaire

1. Problèmes

? Problème 65 - Construction des entiers naturels

Au milieu du XIXième siècle, les mathématiciens se sont rendus compte que leurs sens leur faisaient défaut lorsqu'ils ont découvert les géométrie non euclidienne. Il a fallu tout reconstruire sur des bases très solides.
Comment construire l'ensemble des entiers sans équivoque!

? Problème 66 - Construction des entiers relatifs, des rationnels

Une fois que les entiers $1,2,3\ldots$ sont définis, ainsi que le 0, comment définir proprement les nombres entiers négatifs.
A savoir que dans l'histoire, les fractions sont apparues bien avant les nombres négatifs!
Nous verrons en algèbre générale qu'il est souvent préférable d'avoir un corps plutôt qu'un anneau (tous les éléments sont inversibles).
Comment définir alors proprement l'ensemble des fractions $\frac{a}{b}$ et justifier l'égalité $\frac{a}{b}=\frac{c}{d}$, alors que $a\neq c$ et $b\neq d$?
Une fois que la construction est acquise (avec les lois + et $\times$), comment généraliser sur $\mathbb{Q}$ la relation d'ordre $\leqslant$ définie sur $\mathbb{Z}$?

? Problème 67 - Construction des réels

Pour le familier de la calculatrice (ou de Python), les nombres réels sont obtenus en écrivant les nombres décimalement quitte à ce que cette écriture soit infini. Cela marche bien; la preuve : le sur-développement des ordinateurs et autres objets numériques.
Comment faire cette construction et surtout comment gérer la problématique $1-0,999\ldots9\cdots=0$, donc la non unicité d'écriture de certains nombres réels.
Là aussi : comment définir la relation d'ordre?

? Problème 68 - Construction des réels

Une autre possibilité : exploiter le principe de la dichotomie. Cela rappelle la méthode des coupures de DEDEKIND, en séance de cours-TP.
Pouvons-nous dès maintenant anticiper cette méthode?

? Problème 69 - Densité de $\mathbb{Q}$ dans $\mathbb{R}$

La construction de $\mathbb{R}$ conduit à voir tous les éléments de $\mathbb{R}$ comme limite d'éléments de $\mathbb{Q}$.
On dit que $\mathbb{Q}$ est dense dans $\mathbb{R}$ (associé à la continuité, c'est une propriété très forte!).
Et pourtant, les cardinaux de $\mathbb{Q}$ et $\mathbb{R}$ sont-ils comparables? Existe-t-il une bijection de $\mathbb{N}$ sur $\mathbb{Q}$? de $\mathbb{Q}$ sur $\mathbb{R}$?

2. Nombres algébriques

2.1. Nombres entiers

Nombres entiers naturels

On commence par admettre la construction de l'ensemble des entiers naturels $\mathbb{N}$.

Heuristique - Nombres entiers naturels

La construction suivante est dûe à Péano. Soit E un ensemble non vide, possédant un élément de référence et dont tous les éléments ont un unique successeur (différent de l'élément de référence).
Cet élément de référence se note 0 (ou 1, selon). Puis on définit l'addition +1 comme le passage d'un nombre à son successeur.
On a ainsi les bases pour un raisonnement par récurrence et l'ensemble des entiers.
Ce qui suit est en fait assez naturel, même si cela peut paraître un peu compliqué la première fois qu'on le voit...

Théorème - Construction de PÉANO

Il existe un ensemble $\mathbb{N}$ non vide, munie d'une loi s (comme successeur) telle que :
- $\mathbb{N}$ étant non vide, il admet une élément premier noté 0.
- Pour tout élément $a \in \mathbb{N}$, il existe $b \in \mathbb{N}^*$ tel que $b = s(a)$.
- s est injective ($s(a) = s(a') \Rightarrow a = a'$)

Remarque - Addition +1

$s(a)$ correspond à la classique addition +1.
L'habitude consiste à associer à l'élément $s \circ s \circ \cdots s(0)$, le nombre n égal au nombre de fois que s est utilisé

Proposition - Opérations sur $\mathbb{N}$

On définit l'addition sur $\mathbb{N}$ par : $a + b = s^a(0) + s^b(0) = s^{a+b}(0)$.
On a $a + b = b + a$.
La multiplication est alors la répétition de l'addition : $a \times b = \underbrace{a + a + \cdots + a}_{b \text{ fois}}$
On a $a \times b = b \times a$.

Proposition - Relation d'ordre

$\mathbb{N}$ est naturellement ordonné (récursivement) :
$\forall\, a, b \in \mathbb{N}^*,\ a \leqslant b \iff s^{-1}(a) \leqslant s^{-1}(b)$.
L'ordre est total.

Et il existe un algorithme, qui termine, permettant de connaître le plus petit entre a et b :

Informatique - Ordre

```
def petit(a,b):
    c,d=a,b
    while c>0 and d>0 :
        c,d=c-1,d-1
    if c==0:
        return(a)
    else :
        return(b)
```

Nombres entiers naturels

Ensuite on construit l'ensemble des entiers relatifs. On propose ici un exercice.

Heuristique - Problématique

La problématique : l'addition à trou (ou recherche d'une opération réciproque) n'est qu'à moitié possible. En effet, elle dépend de la relation d'ordre entre les nombres soustraits.
Il faut donc créer un premier ensemble, afin que toute soustraction de nombres entiers soit possible.
Mais certaines soustractions peuvent conduire à un « même » résultat

Exercice
Sur $\mathbb{N}^2$,

1. Montrer que $\sim_1$ définie par :
$$(a,b) \sim_1 (c,d) \iff a+d = c+b$$
est une relation d'équivalence.
2. Montrer que tout couple (a,b) est dans la classe d'un couple $(0,d)$ ou $(d,0)$ selon que $a \leqslant b$ ou $a \geqslant b$
3. En déduire la construction de $\mathbb{Z}$ comme équivalent à l'ensemble $\dfrac{\mathbb{N}^2}{\sim_1}$

Correction

1. Elle est bien réflexive, symétrique et transitive (déjà démontré)
2. Si $a \leqslant b$, alors $b-a \geqslant 0$ et donc $(a,b) \sim_1 (0,b-a)$ car $a+(b-a) = b+0$.
Si $a \geqslant b$, alors $a-b \geqslant 0$ et donc $(a,b) \sim_1 (a-b,0)$ car $a = b+(a-b)$.
3. Aucun des couples $(0,c)$ ou $(d,0)$ n'est en relation avec un autre si les nombres c ou d sont différents.
On a donc trouvé un représentant de chacune des classes d'équivalences de $\mathbb{N}^2$ pour la relation $\sim_1$.
On peut noter $+c = (c,0)$ et $-c = (0,c)$. En on construit ainsi $\mathbb{Z}$.

Exemple - Le nombre -2

Selon cette construction, le nombre habituellement noté -2 correspond à la classe d'équivalence des nombres $(0,2)$, $(1,3)$,… $(n,n+2)$…
Et le nombre $+3$?

Proposition - Opération sur $\mathbb{Z}$

L'ensemble $\mathbb{Z}$ est l'ensemble $\dfrac{\mathbb{N}^2}{\sim_1}$ (des classes d'équivalence sur $\mathbb{N}^2$ de la loi $\sim_1$).
On définit alors la relation d'ordre $\leqslant_{\mathbb{Z}}$ par :
$$\overline{(a,b)} \leqslant_{\mathbb{Z}} \overline{(c,d)} \iff a+d \leqslant_{\mathbb{N}} c+b$$
L'addition est alors simplement : $\overline{(a,b)} +_{\mathbb{Z}} \overline{(c,d)} = \overline{(a +_{\mathbb{N}} c, b +_{\mathbb{N}} d)}$
La multiplication est plus compliquée :
$$\overline{(a,b)} \times_{\mathbb{Z}} \overline{(c,d)} = \overline{(a \times_{\mathbb{N}} c +_{\mathbb{N}} b \times_{\mathbb{N}} d, b \times_{\mathbb{N}} c +_{\mathbb{N}} a \times_{\mathbb{N}} d)}$$

Remarque - La difficulté : l'indépendance au représentant

Il faut bien vérifier que chacune de ces définitions est indépendantes du représentant de la classe d'équivalence.
Ainsi : si $\overline{(a,b)} = \overline{(a',b')}$ et $\overline{(c,d)} + \overline{(c',d')}$,
alors $(a+c)+(b'+d') = (a+b')+(c+d') = (a'+b)+(c'+d) = (a'+c')+(b+d)$.
donc on a bien : $(a+c,b+d) \sim_1 (a'+c',b'+d')$.
Ainsi la définition de $+_{\mathbb{Z}}$ est bien indépendante des représentants choisis.

Exemple - Multiplication

Vérifions sur un exemple que la multiplication donne le résultat attendu :

$$-3 \times 4 = \overline{(1,4)} \times \overline{(6,2)} = \overline{(1\times 6 + 4\times 2, 1\times 2 + 4\times 6)} = \overline{(14,26)} = -12$$

Démonstration

Il s'agit bien d'une relation d'ordre. Il faudrait montrer aussi qu'elle est indépendante des représentants choisis.

C'est l'opération d'ordre naturelle associée à une relation d'équivalence.

- Reflexive : $\overline{(a,b)} \leqslant_{\mathbb{Z}} \overline{(a,b)}$ car $a+b \leqslant_{\mathbb{N}} a+b$.
- Antisymétrique : $\overline{(a,b)} \leqslant_{\mathbb{Z}} \overline{(c,d)}$ et $\overline{(c,d)} \leqslant_{\mathbb{Z}} \overline{(a,b)}$ alors $a+d \leqslant_{\mathbb{N}} c+b$ et $c+b \leqslant_{\mathbb{N}} a+d$, donc $a+d=c+d$, donc $\overline{(a,b)} \sim_1 \overline{(c,d)}$.
- Transitive : $\overline{(a,b)} \leqslant_{\mathbb{Z}} \overline{(c,d)}$ et $\overline{(c,d)} \leqslant_{\mathbb{Z}} \overline{(e,f)}$ alors $a+d \leqslant_{\mathbb{N}} c+b$ et $c+f \leqslant_N e+d$.
  Donc, par transitivité de $\leqslant_{\mathbb{N}}$: $a+d+f \leqslant_{\mathbb{N}} c+b+f = c+f+b \leqslant e+d+b$,
  donc avec s^{-d} : $a+f \leqslant_N e+b$.
  Ainsi $\overline{(a,b)} \leqslant_{\mathbb{Z}} \overline{(e,f)}$

L'addition et la multiplication sont indépendants des représentants choisis.

□

Exercice

Montrer que l'ordre est total

Correction

L'ordre est total sur $\mathbb{N}$. Soient $\overline{(a,b)}, \overline{(c,d)} \in \mathbb{Z}$.

Alors on a ou bien $a+d \leqslant_{\mathbb{N}} b+c$ ou bien $b+c \leqslant_{\mathbb{N}} a+d$.

Donc l'ordre $\leqslant_{\mathbb{Z}}$ est bien total

2.2. Nombres rationnels

La construction de $\mathbb{Q}$ est en tout point équivalente à la construction de $\mathbb{Z}$, mais pour un problème lié à la multiplication (et donc division) au lieu de l'addition (et donc la multiplication).

Exercice

Sur $\mathbb{Z} \times \mathbb{N}^*$,

1. Montrer que $\sim_2$ définie par :
$$(a,b) \sim_2 (c,d) \iff a \times_{\mathbb{Z}} d = c \times_{\mathbb{Z}} b$$
est une relation d'équivalence.
2. Montrer la construction de $\mathbb{Q}$ comme équivalent à l'ensemble $\dfrac{\mathbb{Z}\times\mathbb{N}}{\sim_2}$
3. Montrer que $\leqslant_{\mathbb{Q}}$ définie par $(a,b) \leqslant_{\mathbb{Q}} (c,d)$ ssi $a \times_{\mathbb{Z}} d \leqslant_{\mathbb{Z}} b \times_{\mathbb{Z}} c$ définie bien une relation d'ordre sur $\mathbb{Q}$
4. Comment définir $+_{\mathbb{Q}}$ et $\times_{\mathbb{Q}}$

Correction

1. Elle est bien réflexive, symétrique et transitive (déjà démontré)
2. Supposons que $(a,b) \sim_2 (c,d)$ et $b \geqslant d$ (on ne perd pas de généralité).
   Si il existe $k \in \mathbb{N}$ tel que $d = bk$, alors $abk = cb$ et donc $c = ak$. On notera alors que k divise a et b.
   On appelle couple irréductible un couple $(a,b) \in \mathbb{Z}\times\mathbb{N}$ tel que $\forall\, k \in \mathbb{N},\ k \nmid a$ ou $k \nmid b$.
   Chaque classe d'équivalence a un unique représentant irréductible, que l'on note habituellement $\frac{a}{b}$.
   En on construit ainsi $\mathbb{Q}$.
3. C'est la relation d'ordre qui découle naturellement de la relation d'équivalence sur l'ensemble ordonné $\mathbb{Z}$
4. On peut imaginer : $\overline{(a,b)} +_{\mathbb{Q}} \overline{(c,d)} = \overline{(a\times_{\mathbb{Z}} d +_{\mathbb{Z}} b\times_{\mathbb{Z}} c, b\times_{\mathbb{Z}} d}$ et $\overline{(a,b)} \times_{\mathbb{Q}} \overline{(c,d)} = \overline{(a\times_{\mathbb{Z}} c, b\times_{\mathbb{Z}} d}$

Il faudrait vérifier que $\mathbb{Q}$ est bien un corps (addition, multiplication inversible) compatible avec $\leqslant_Q$. Cela se fait sans grande difficultés...

Pour aller plus loin - Comment définir π simplement ?

Nous savons que le périmètre d'une forme régulière est proportionnel à son agrandissement.

Donc pour un cercle, il existe une constante telle que son périmètre est égale au produit de cette constante par le diamètre : $p = C \times d$. Par définition, on peut appeler pi1 π_1 cette constante.

Ou bien, de même nous savons que la surface d'une forme régulière est proportionnel au carré de son agrandissement.

Donc pour un disque, il existe une constante telle que son aire est égale au produit de cette constante par le carré du rayon : $S = C' \times r^2$. Par définition, on peut appeler π_2 cette constante.

L'enjeu : montrer que $\pi_1 = \pi_2$...

Pour aller plus loin - Généralisation

Ce principe qui permet de passer de l'anneau $\mathbb{Z}$ au corps des fractions $\mathbb{Q}$ est un principe régulièrement repris en mathématiques.

On suivra exactement ce même principe pour décrire le corps des fractions de polynômes $\mathbb{K}(X)$ à partir de l'anneau des polynômes : $\mathbb{K}[X]$

Proposition - Opération sur $\mathbb{Q}$
L'ensemble $\mathbb{Q}$ est l'ensemble $\frac{\mathbb{Z}\times\mathbb{N}^*}{\sim_2}$ (des classes d'équivalence sur $\mathbb{Z}\times\mathbb{N}$ de la loi $\sim_2$).
On définit alors la relation d'ordre $\leqslant_{\mathbb{Q}}$ par :

$$\overline{(a,b)} \leqslant_{\mathbb{Q}} \overline{(c,d)} \iff a\times_{\mathbb{Z}} d \leqslant_{\mathbb{Z}} c\times_{\mathbb{Z}} b$$

La multiplication est alors simplement : $\overline{(a,b)}\times_{\mathbb{Q}}\overline{(c,d)} = \overline{(a\times_{\mathbb{Z}} c, b\times_{\mathbb{Q}} d)}$
L'addition est plus compliquée : $\overline{(a,b)}+_{\mathbb{Q}}\overline{(c,d)} = \overline{(a\times_{\mathbb{Z}} d +_{\mathbb{Z}} b\times_{\mathbb{Z}} c, b\times_{\mathbb{Z}} d)}$.
L'ordre est total (démonstration comme pour comme pour $\leqslant\mathbb{Z}$).

2.3. Nombres algébriques

La plupart du temps les nombres se définissent par une phrase, qui elle même se traduit en une équation (polynomiale).
On définit alors :

Définition - Nombre algébrique
Un nombre r est un nombre algébrique si il existe une fonction polynomiale P à coefficients entiers telle que $P(r)=0$.
Si n est le plus petit degré d'un polynôme vérifiant cette relation, on dit que r est algébrique d'ordre n

Exemple - Nombres rationnels
Les nombre rationnels sont des nombres algébriques d'ordre 1.
$\frac{p}{q}$ est racine du polynôme $x\mapsto qx-p$.

Exemple - Nombres quadratiques
Les nombre quadratiques sont les nombres algébriques d'ordre 2.
$\sqrt{2}$, racine de x^2-2 ou $\Phi=\frac{-1+\sqrt{5}}{2}$, racine de x^2+x-1 sont des nombres quadratiques.
D'une certaine façon, on peut également dire que i est quadratique.
D'autres nombres, toujours « naturels » ne s'expriment pas à partir d'une équation polynomiale à coefficients entiers. C'est le cas de π ou de e.

Pour aller plus loin - Nombres irrationnels
Ce fut un choc dans l'antiquité lorsqu'on compris que $\sqrt{2}$, qui existe bien (longueur de la diagonale du carré de côté 1), n'est pas une nombre rationnel. Quel type de nombre est-ce?
Si l'on considère des nombres irrationnels (i.e. non rationnels) dans leur singularité, on n'en trouve pas beaucoup, ils sont en effet difficile à définir.

Définition - Nombre transcendant
Si r n'est pas algébrique, on dit qu'il est transcendant.

Démontrer qu'un nombre donné est transcendant n'est pas une mission évidente.

Pour aller plus loin - Problème ouvert
Est-ce que γ est un nombre transcendant ou algébrique?
Par définition, $\gamma=\lim(\sum_{k=1}^{n}\frac{1}{k}-\ln n)$

3. Propriétés de $\mathbb{R}$

3.1. Principe de construction de $\mathbb{R}$

Heuristique - Un principe : les coupures de DEDEKIND. Rappel
$\mathbb{Q}$ est un ensemble, relativement naturel, muni d'une relation d'ordre totale $\leqslant$.
$\mathbb{R}\sim\mathscr{C}(\mathbb{Q})$ est l'ensemble des sections ouvertes commençantes sur $\mathbb{Q}$.

$$\mathscr{C}(\mathbb{Q})=\left\{T\subset\mathbb{Q}\mid \forall\, a\in T, b\leqslant a\Rightarrow b\in T \,\&\, \nexists\, m\in\mathbb{Q} \text{ tel que } T=\{x\in\mathbb{Q}\mid x\leqslant m\}\right\}$$

L'addition est assez naturellement prolongée, ainsi que la relation d'ordre totale.
La multiplication par des positifs est simple, ensuite c'est la règle des signes.
On obtient ensuite quelques résultats topologiques, nouveaux principes de bases ici. Cela

Pour aller plus loin - Autre complétion
L'ensemble $\mathbb{R}$ est l'ensemble que l'on obtient naturellement, à partir limites de suites de rationnels (éléments de $\mathbb{Q}$), limites définies à partir de la distance $d(a,b)=|a-b|$ où l'on retrouve la valeur absolue classique.
Mais il existe une (seule) autre façon de faire. On fixe un nombre premier p et on mesure la distance $d\left(\frac{r_1}{s_1},\frac{r_2}{s_2}\right) = p^{v_p(s_1)+v_p(s_2)-v_p(r_1s_2-r_2s_1)}$ où pour $a\in\mathbb{Z}$, $v_p(a)=\max\{\alpha\mid p^\alpha|a\}$.
A partir de cette distance, en complétant $\mathbb{Q}$

, est fondamentalement lié au fait qu'il existe des rationnels infiniment proches.

3.2. Fonctions classiques associées à $\mathbb{R}$

Valeur absolue

Définition - Valeur absolue

Pour $x \in \mathbb{R}$, on pose $|x| = \max(x,-x) = \begin{cases} x & \text{si } x \geqslant 0 \\ -x & \text{si } x < 0 \end{cases}$ (plus grand des deux réels x et $-x$).
$d(x,y) = |x-y|$ mesure la distance entre deux réels x et y de la droite réelle.

Définition - Partie positive, partie négative

Pour $x \in \mathbb{R}$, on pose $x^+ = \max(x,0)$ (partie positive du réel x)
et $x^- = \max(-x,0)$ (partie négative du réel x).
Ces deux réels sont POSITIFS.

Exercice
Ecrire $|x|$ en fonction de x^+ et x^-.
Ecrire x^+ en fonction de $|x|$ et de x.

Correction
$|x| = x^+ - x^-$, $x^+ = \frac{|x|+x}{2}$ et $x^- = \frac{|x|-x}{2}$

Proposition - Encadrements à connaître

$$|x| \leqslant M \iff -M \leqslant x \leqslant M$$
$$|x| \geqslant M \iff (x \geqslant M \text{ ou } x \leqslant -M)$$
$$\forall (x,y) \in \mathbb{R}^2, \quad |xy| = |x|\,|y|$$
$$\forall (x,y) \in \mathbb{R}^2, \quad \Big||x| - |y|\Big| \leqslant |x+y| \leqslant |x| + |y|$$
$$\forall (x,y) \in \mathbb{R}^2, d(x,y) \geqslant d(|x|,|y|) \text{ ou } |x-y| \geqslant \Big||x| - |y|\Big|$$

Démonstration
Première proposition :
Supposons que $|x| \leqslant M$.
- Si $x \geqslant 0$, $|x| = x$ et donc $|x| \leqslant M \Longrightarrow 0 \leqslant x \leqslant M \Longrightarrow -M \leqslant x \leqslant M$.
- Si $x \leqslant 0$, $|x| = -x$ et donc $|x| \leqslant M \Longrightarrow 0 \leqslant -x \leqslant M \Longrightarrow -M \leqslant x \leqslant 0 \Longrightarrow -M \leqslant x \leqslant M$.

Réciproquement si $-M \leqslant x \leqslant M$,
alors $M \geqslant -x \geqslant -M$, et donc x et $-x$ appartiennent à $[-M,M]$.
Comme $|x|$ est l'un des deux nombres x ou $-x$, alors $|x| \in [-M,M]$.
La seconde proposition est la contraposée de la première proposition.
Troisième proposition, par étude de cas :
- si $x, y \geqslant 0$, alors $|xy| = xy = |x|\,|y|$
- si $x \geqslant 0$, $y \leqslant 0$, alors $|xy| = -xy = x \times (-y) = |x|\,|y|$
- si $x \leqslant 0$, $y \geqslant 0$, alors $|xy| = -xy = (-x) \times y = |x|\,|y|$
- si $x, y \leqslant 0$, alors $|xy| = xy = (-x) \times (-y) = |x|\,|y|$

Quatrième proposition :
$x + y \leqslant |x| + |y|$ et $-(x+y) = -x + -y \leqslant |x| + |y|$,
donc $|x+y| = \max(x+y, -x-y) \leqslant |x| + |y|$
On en déduit que $|x| = |(x+y) + (-y)| \leqslant |x+y| + |y|$, donc $|x| - |y| \leqslant |x+y|$.
De même $|y| - |x| \leqslant |x+y|$.
Ainsi : $||x| - |y|| = \max(|x| - |y|, |y| - |x|) \leqslant |x+y|$. Cinquième proposition :
On a simplement : $||x| - |y|| = ||x| - |-y|| \leqslant |x + (-y)| = |x-y|$ □

Partie entière

Proposition - Corps archimédien
Comme $\mathbb{Q}$, $\mathbb{R}$ est archimédien :

$$\forall\ (a,A) \in\in \mathbb{R}_+^* \times \mathbb{R}_+, \qquad \exists\ n \in \mathbb{N},\ \text{tel que } na \geqslant A$$

Avec $a = 1$, cela conduit à la définition :

Définition - Partie entière
Soit $x \in \mathbb{R}$. Il existe un unique $n \in \mathbb{Z}$ vérifiant $n \leqslant x < n+1$.
n s'appelle la partie entière de x, on la note $\lfloor x \rfloor$. On a donc

$$\lfloor x \rfloor \leqslant x < \lfloor x \rfloor + 1 \quad \text{et} \quad x - 1 < \lfloor x \rfloor \leqslant x.$$

Démonstration

- Si $x \in \mathbb{Z}$, alors $n = x$ fonctionne.
- Si $x > 0$.
  L'ensemble $E_x = \{m \in \mathbb{N} \mid m \leqslant x\}$ est borné par k.
  Tout sous-ensemble non vide et majoré de $\mathbb{N}$ admet une borne supérieur n.
  On a alors $n \leqslant x$ et $n+1 \notin E_x$, donc $n+1 > x$.
- Et si $x < 0$ (avec $x \notin \mathbb{Z}$).
  Alors il existe $n' \in \mathbb{N}$ tel que $n' \leqslant -x < n'+1$.
  Donc $-n' \geqslant x > -n'-1$, et comme $x \notin \mathbb{Z}$, $x \neq -n'$.
  Ainsi avec $n = -n'-1$, on a $n < x < n+1$ ce qui implique $n \leqslant x < n+1$

□

Exercice
Pour tout entier $n \geqslant 1$, montrer :

$$\frac{1}{\sqrt{n+1}} < 2(\sqrt{n+1} - \sqrt{n}) < \frac{1}{\sqrt{n}}$$

En déduire la partie entière du réel $A = 1 + \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{3}} + \cdots + \frac{1}{\sqrt{10\,000}}$.

Correction
On peut faire une étude de deux fonctions.
On peut aussi penser à la quantité conjugué :

$$\sqrt{n+1} - \sqrt{n} = \frac{(\sqrt{n+1} - \sqrt{n})(\sqrt{n+1} + \sqrt{n})}{(\sqrt{n+1} + \sqrt{n})} = \frac{n+1-n}{(\sqrt{n+1} + \sqrt{n})} = \frac{1}{(\sqrt{n+1} + \sqrt{n})}$$

Et comme $2\sqrt{n} \leqslant \sqrt{n+1} + \sqrt{n} \leqslant 2\sqrt{n+1}$, on a

$$\frac{1}{\sqrt{n+1}} \leqslant 2\left(\sqrt{n+1} - \sqrt{n}\right) \leqslant \frac{1}{\sqrt{n}}$$

On peut alors sommer ces inégalités, par télescopage :

$$198 = 2(100-1) < 2\left(\sqrt{10\,001} - \sqrt{1}\right) = \sum_{n=1}^{10\,000} 2\left(\sqrt{n+1} - \sqrt{n}\right) < \sum_{n=1}^{10\,000} \frac{1}{\sqrt{n}} = A$$

Et de même :

$$A = 1 + \sum_{n=1}^{9\,999} \frac{1}{\sqrt{n+1}} < 1 + \sum_{n=1}^{9\,999} 2\left(\sqrt{n+1} - \sqrt{n}\right) = 1 + 2(100-1) = 199$$

Donc la partie entière de A vaut 198.

Savoir faire - Travailler avec la partie décimale
Fréquemment, on exploite également la fonction partie décimale θ : $\theta(x) = x - \lfloor x \rfloor$.
On voit que $\theta(x) \in [0,1[$, pour tout $x \in \mathbb{R}$

Exercice

Pour tout réel x, déterminer la limite $\lim_{n\to\infty} \frac{\lfloor x\rfloor + \lfloor 2x\rfloor + \cdots + \lfloor nx\rfloor}{n^2}$

Correction

On sait que $kx = \lfloor kx\rfloor + \theta(kx)$, avec $\theta(kx) \in [0,1[$.
Donc $\lfloor x\rfloor + \lfloor 2x\rfloor + \cdots + \lfloor nx\rfloor = x + 2x + \cdots + nx - \theta(x) - \theta(2x) - \ldots \theta(nx) = \frac{n(n+1)}{2}x - \Theta(x,n)$ avec $\Theta(x,n) \in [0,n]$.
Donc

$$\left|\frac{\lfloor x\rfloor + \lfloor 2x\rfloor + \cdots + \lfloor nx\rfloor}{n^2} - \frac{x}{2}\right| = \left|\left(\frac{n^2+n}{2n^2} - \frac{1}{2}\right)x - \frac{\Theta(x)}{n^2}\right| \leqslant |\frac{x}{2n}| + \frac{1}{n}$$

Et donc par encadrement, ceci tend vers 0 et donc la suite a pour limite $\frac{x}{2}$

4. Parties de $\mathbb{R}$ et topologie

4.1. Bornes supérieure et inférieure

On commence par quelques rappels de définitions, mais adaptés ici au cas réel :

Définition - Sous-ensemble majoré, minoré, borné
Soit A un sous-ensemble de $\mathbb{R}$. On dit que :
- A est *majoré* s'il existe un réel M tel que, pour tout x de A, on ait $x \leqslant M$.
  M est alors un majorant de A.
- A est *minoré* s'il existe un réel m tel que, pour tout x de A, on ait $m \leqslant x$.
  m est alors un minorant de A.
- Si A est majoré et minoré, on dit qu'il est borné.

Pour aller plus loin - Rappels
Un majorant M est le plus grand élément de A si et seulement si il appartient également à A.
Un minorant m est le plus petit élément de A si et seulement si il appartient à A

Remarque - Ensemble $\mathbb{N}$
Pour l'ensemble $\mathbb{N}$, on a quelques propriétés :
- Tout sous-ensemble non vide de $\mathbb{N}$ admet un plus petit élément.
- Tout sous-ensemble non vide et majoré de $\mathbb{N}$ admet un plus grand élément.

Ce résultat n'est pas vrai si l'on remplace $\mathbb{N}$ par $\mathbb{R}$.
Rappels :

Pour aller plus loin - Récurrence
C'est la première propriété ici (avec un raisonnement par l'absurde) qui permet de montrer l'exactitude du raisonnement par récurrence (et aussi de la descente infinie) et aussi la méthode du variant de boucle en informatique.

Définition - Borne inférieure, borne supérieure
Soit $A \subset \mathbb{R}$.
- Si l'ensemble des majorants de A est non vide
  et si il admet un plus petit élément a,
  alors a est appelé borne supérieure de A, on note $a = \sup A$.
  Formellement :

$$\sup A := \min\{M \in \mathbb{R} \mid \forall\ a \in A, a \leqslant M\} \text{ (si non vide)}$$

- Si l'ensemble des minorants de A est non vide
  et si il admet un plus grand élément b,
  alors b est appelé borne inférieure de A, on note $a = \inf A$.
  Formellement :

$$\inf A := \max\{m \in \mathbb{R} \mid \forall\ a \in A, a \geqslant m\} \text{ (si non vide)}$$

Attention - Borne supérieure
Comme son nom ne l'indique pas, la borne supérieure est par définition le plus **petit** élément d'un certain ensemble (celui des majorants).

L'exercice suivant donne des exemples à toujours bien garder dans un coin de sa tête…

Exercice

Déterminer, s'ils existent, le plus grand élément, le plus petit élément, la borne supérieure, la borne inférieure (sur $\mathbb{R}$) des parties suivantes :

$$A = [0,1], \quad B = [0,1[, \quad C = \left\{ \frac{1}{n} \mid n \in \mathbb{N}^* \right\}$$

Correction

- A admet 1 comme plus grand élément, 0 comme plus petit élément.
  Les majorants de A forment l'ensemble $[1,+\infty[$ et donc 1 est également la borne supérieure de A. Les minorants de A forment l'ensemble $]-\infty,0]$ et donc 0 est également la borne inférieure de A.
- B n'admet pas de plus grand élément, mais bien 0 comme plus petit élément.
  Les majorants de B forment l'ensemble $[1,+\infty[$ et donc 1 est également la borne supérieure de B. Les minorants de B forment l'ensemble $]-\infty,0]$ et donc 0 est également la borne inférieure de B.
- C n'admet pas de plus petit élément, mais bien $\frac{1}{1} = 1$ comme plus grand élément.
  Les majorants de C forment l'ensemble $[1,+\infty[$ et donc 1 est également la borne supérieure de C. Les minorants de C forment l'ensemble $]-\infty,0]$ et donc 0 est également la borne inférieure de C.

Proposition - Condition d'existence de la borne supérieure

Soit $A \subset \mathbb{R}$ non vide. On suppose que A possède un plus grand élément a (resp. plus petit élément b).
Alors A possède une borne supérieure (resp. inférieure) et $\sup A = a$ (resp $\inf A = b$).

Savoir faire - Etudier une borne supérieure

En règle générale, pour obtenir une égalité sur la borne supérieure, on exploite deux inégalités :

- $\forall\ a \in A,\ a \leqslant \sup A$ (minoration de $\sup A$)
- $\forall\ M \in \mathbb{R}$ tel que $\forall\ a \in A,\ a \leqslant M$, alors $M \geqslant \sup A$
  (majoration de $\sup A$, par tous les majorants de A)

On a évidemment des relations symétriques pour la borne inférieure…

Exercice

Soient A et B deux parties de $\mathbb{R}$ admettant des bornes supérieures. Montrer que

$$A \subset B \Rightarrow \sup A \leqslant \sup B.$$

Donner un résultat similaires avec les bornes inférieures.

Correction

Il s'agit de majorer $\sup A$.

On exploite donc la deuxième inégalité du savoir-faire (pour $\sup A$). On démontre donc que $\sup B$ est un majorant de A.

Et pour cela, on exploite la première inégalité du savoir-faire (pour $\sup B$) On « remonte » : $\forall\ a \in A$, $a \in B$ car $A \subset B$, donc $a \leqslant \sup B$.

Donc $\sup B$ est un majorant de A et $\sup A$ est le plus petit des majorants.

Ainsi $\sup A \leqslant \sup B$

Les deux propositions suivantes donnent des caractérisations *opératoires* (avec lesquelles travailler dans les démonstrations) et donc un nouveau savoir-faire :

Proposition - Caractérisation de la borne sup.

Soit $A \subset \mathbb{R}$ et $x \in \mathbb{R}$
Alors $x = \sup A$ si et seulement si

$$\begin{cases} \forall\ a \in A,\ a \leqslant x \\ \forall\ \epsilon > 0, \exists\ a_\epsilon \in A \mid , x - \epsilon < a_\epsilon \end{cases}$$

Proposition - Caractérisation de la borne inf.
Soit $A \subset \mathbb{R}$ et $x \in \mathbb{R}$
Alors $x = \inf A$ si et seulement si

$$\begin{cases} \forall a \in A,\ x \leqslant a \\ \forall \epsilon > 0,\ \exists a_\epsilon \in A \,|\, a_\epsilon < x + \epsilon \end{cases}$$

Démonstration
Si A admet une borne supérieure a.
Alors l'ensemble $M_A = \{M \mid \forall\, x \in A, x \leqslant M\}$, des majorants de A est non vide
et admet a comme plus petit élément.
Ainsi $a \in M_A$ et donc $\forall\, x \in A,\ x \leqslant a$.
Ensuite, considérons $\epsilon > 0$, alors $a - \epsilon \notin M_A$, et donc $\exists\, x_\epsilon \in A$ tel que $a - \epsilon < x_\epsilon$ (absurde).
Réciproquement, si il existe a tel que $\forall\, x \in A,\ x \leqslant a$ et $\forall\, \epsilon > 0, \exists\, x_\epsilon \in A \mid , a - \epsilon < x_\epsilon$.
Alors A est majorée et l'ensemble des majorants de A (M_A) est non vide : $a \in M_A$.
Soit $m \in M_A$, un majorant de A. Supposons que $m < a$.
Soit $\epsilon = \frac{a-m}{2} > 0$.
On a alors $a - \epsilon < x_\epsilon \leqslant m$ et donc $a \leqslant m + \epsilon = m + \frac{a-m}{2} = \frac{a+m}{2} < \frac{a+a}{2} = a$.
Ceci est contradictoire. Donc si m est un majorant, alors $m \geqslant a$. □

Le théorème suivant est parfois pris comme caractérisation de $\mathbb{R}$. Nous en avons vu la démonstration dans le cours-TD (il faut une défintion pour $\mathbb{R}$)

Théorème - Existence de la borne supérieure
Toute partie non vide majorée de $\mathbb{R}$ admet une borne supérieure.
Toute partie non vide minorée de $\mathbb{R}$ admet une borne inférieure.

Attention - Propriété non vérifiée par $\mathbb{Q}$
Cette propriété différencie $\mathbb{R}$ et $\mathbb{Q}$:
— $\{x \in \mathbb{R} \,|\, x^2 < 2\}$ admet une borne supérieure (dans $\mathbb{R}$), que l'on notera : $\sqrt{2}$
— mais $\{x \in \mathbb{Q} \,|\, x^2 < 2\}$ n'admet pas de borne supérieure dans $\mathbb{Q}$ (non existence d'un plus petit élément dans $\mathbb{Q}$ de l'ensemble des majorants rationnels).

Heuristique - Manipuler l'ensemble des majorants et non l'ensemble E lui-même
L'ensemble E peut être très compliqué, un ensemble à trous par exemple $\bigcup_{n\in\mathbb{N}} \left[(1+\frac{1}{n})^n - \frac{1}{n^2}; (1+\frac{1}{n})^n + \frac{1}{n^2}\right]$.
Il vaut mieux raisonner sur l'ensemble des majorants $\mathcal{M}$: celui-ci est nécessairement un intervalle :
Mieux (mais on ne le sait pas encore) il s'agit de l'intervalle fermé $[\sup E, +\infty[$.

Exercice
Soit A une partie de $\mathbb{R}$. On note $\mathcal{M}(A)$, l'ensemble des majorants de A.
A quoi ressemble $\mathcal{M}(A)$?

Correction
Si A n'est pas majoré alors $\mathcal{M}(A) = \emptyset$.
Sinon, alors A admet une borne supérieure, notée $\sup A$. Nécessairement $\sup(A) \in \mathcal{M}(A)$.
Si $m \geqslant \sup A$, alors par transitivité, $\forall\, a \in A,\ a \leqslant \sup A \leqslant m$, donc $m \in \mathcal{M}(A)$. Ainsi $[\sup A, +\infty[\subset \mathcal{M}(A)$.
Réciproquement si $m \in \mathcal{M}(A)$, alors par minimalité de $\sup A$, $m \geqslant \sup A$. Donc $\mathcal{M}(A) \subset [\sup A, +\infty[$.
Dans ce cas $\mathcal{M}(A) = [\sup A, +\infty[$.

Pour aller plus loin - Démonstration de ce résultat
$E = \{r \in \mathbb{Q}^+ \mid r^2 < 2\}$ n'admet pas de borne supérieure dans $\mathbb{Q}$ en considèrant $s = \frac{p}{q} \mapsto \frac{3p+4q}{2p+3q}$.
Si $s = \frac{p}{q} \in E$, alors $\frac{p^2}{q^2} < 2$ et $s' = \frac{3p+4q}{2p+3q} \in \mathbb{Q}$.
Par ailleurs, $s^2 < s'^2 < 2$.
En effet : $s' > 0$ et $s < s' \Leftrightarrow p(2p+3q) < q(3p+4q) \Leftrightarrow 2p^2 < 4q^2 \Leftrightarrow \frac{p^2}{q^2} < 2$
et $s'^2 < 2 \Leftrightarrow (3p+4q)^2 < 2(2p+3q)^2 \Leftrightarrow 9p^2 + 16q^2 + 24pq < 8p^2 + 18q^2 + 24pq \Leftrightarrow \frac{p^2}{q^2} < 2$
Par conséquent, si $s = \frac{p}{q}$ est le plus petit des majorants de E, alors nécessairement $s \notin E$.
Mais de même si $s = \frac{p}{q} \notin E$, alors $\frac{p^2}{q^2} > 2$ puis

Corollaire - Critère de nullité d'un nombre
Un réel a vérifiant $\forall \epsilon > 0, |a| \leqslant \epsilon$ est nul.

Démonstration
Si a vérifie $\forall \epsilon > 0, |a| \leqslant \epsilon$, alors $|a|$ est la borne supérieur de $\{0\}$.
Or cette borne supérieur est 0, donc $a = 0$. □

On avait déjà fait une démonstration ici par contraposée.

4.2. Densité de $\mathbb{D}$ ou $\mathbb{Q}$ dans $\mathbb{R}$

Ensemble $\mathbb{D}$

Définition - Ensemble des décimaux
Soit $x \in \mathbb{R}$.
On dit que x est un nombre décimal s'il existe $p \in \mathbb{Z}, n \in \mathbb{N}$ tels que $x = \dfrac{p}{10^n}$.
On note $\mathbb{D}$ l'ensemble des nombres décimaux. On a $\mathbb{Z} \subset \mathbb{D} \subset \mathbb{Q}$.

Remarque - Un nombre décimal...
... c'est tout simplement un nombre qui s'écrit avec une virgule et une fin dans le développement.
Par exemple : $25,456394 = \dfrac{25\,456\,394}{10^6}$

Définition - Valeur décimale approchée
Si $p \in \mathbb{Z}$ est tel que $\dfrac{p}{10^n} \leqslant x \leqslant \dfrac{p+1}{10^n}$,
on dit que $\dfrac{p}{10^n}$ (resp. $\dfrac{p+1}{10^n}$) est une valeur décimale approchée par défaut (resp. par excès) de x à la précision 10^{-n}.

Proposition - Obtenir la valeur décimale approchée
Soit $x \in \mathbb{R}$. Pour tout $n \in \mathbb{N}$, $\dfrac{\lfloor 10^n x \rfloor}{10^n}$ (resp. $\dfrac{\lfloor 10^n x \rfloor + 1}{10^n}$) est une valeur approchée de x par défaut (resp. par excès) à la précision 10^{-n}.

Démonstration
On fait le calcul $p = \lfloor 10^n x \rfloor \in \mathbb{Z}$, on a donc $p \leqslant 10^n x < p+1$,
puis en divisant par 10^n : $\dfrac{p}{10^n} \leqslant x < \dfrac{p+1}{10^n}$ □

$\mathbb{D}$ (et $\mathbb{Q}$) denses dans $\mathbb{R}$

Heuristique - Une partie dense ?
Une partie X est dense dans $\mathbb{R}$ si elle peut toucher (à $\epsilon > 0$ près - choisi par avance, aussi petit qu'on veut) tous les éléments de $\mathbb{R}$ avec les propres de X.

$$\forall\, x \in \mathbb{R}, \qquad \forall \epsilon > 0, \exists r \in X, |x - r| < \epsilon$$

Analyse - Vers une définition équivalente
Si X est dense dans $\mathbb{R}$, alors,
pour tout $x \in \mathbb{R}$ et tout $\epsilon > 0$, $\exists\, r \in X$ tel que $|x - r| < \epsilon$, donc $x - \epsilon < r < x + \epsilon$.
pour tout $x \in \mathbb{R}$ et tout $\epsilon > 0$, $]x - \epsilon, x + \epsilon[\cap X \neq \emptyset$.
Ainsi pour tout $a < b \in \mathbb{R}$, en prenant $x = \frac{a+b}{2}$ et $\epsilon = \frac{b-a}{2}$, on a $]a, b[\cap X \neq \emptyset$.
Réciproquement, si pour tout $a < b \in \mathbb{R}$, on a $]a, b[\cap X \neq \emptyset$,
alors pour tout $x \in \mathbb{R}$, $\epsilon > 0$, en prenant $a = x - \epsilon$ et $b = x + \epsilon$, on a $x - \epsilon < r < x + \epsilon$.

Définition - Partie dense dans $\mathbb{R}$
Une partie non vide X de $\mathbb{R}$ est dite dense dans $\mathbb{R}$
si elle rencontre tout intervalle ouvert non vide,
c'est-à-dire si pour deux réels a et b, $a<b$, il existe $x\in X\cap]a,b[$.

Théorème - Parties denses dans $\mathbb{R}$
$\mathbb{D}$, $\mathbb{Q}$ et $\mathbb{R}\setminus\mathbb{Q}$ sont denses dans $\mathbb{R}$.

Démontrons la densité de $\mathbb{D}$ et celle de $\mathbb{R}\setminus\mathbb{Q}$. Comme $\mathbb{D}\subset\mathbb{Q}$, on en déduira la densité de $\mathbb{Q}$

Démonstration
Soient $x\in\mathbb{R}$. Soit $\epsilon>0$.
Il existe $n\in\mathbb{N}$ tel que $\epsilon>10^{-n}$. Puis il existe $p\in\mathbb{Z}$ tel que $\frac{p}{10^n}\leqslant x\leqslant\frac{p+1}{10^n}$.
Puis on a $\frac{p}{10^n}\in\mathbb{D}$ et
$$\left|x-\frac{p}{10^n}\right|=x-\frac{p}{10^n}\leqslant\frac{p+1}{10^n}-\frac{p}{10^n}=10^{-n}<\epsilon$$
Donc $\mathbb{D}$ est dense dans $\mathbb{R}$.
Puis comme $\mathbb{D}\subset\mathbb{Q}$, on a donc aussi $\mathbb{Q}$ dense dans $\mathbb{R}$. Enfin, on sait que $\sqrt{2}\notin\mathbb{Q}$ (raisonnement par l'absurde). Donc si $r\in\mathbb{Q}$, $\sqrt{2}+r\notin\mathbb{Q}$
Soit $x\in\mathbb{R}$ et $\epsilon>0$. Il existe $r\in\mathbb{Q}$ tel que $|(x-\sqrt{2})-r|<\epsilon$.
Donc $|x-(r+\sqrt{2})|<\epsilon$.
Ainsi $\sqrt{2}+\mathbb{Q}\subset(\mathbb{R}\setminus\mathbb{Q})$ est dense dans $\mathbb{R}$, donc $\mathbb{R}\setminus\mathbb{Q}$ aussi. □

Par l'absurde :

Corollaire -
Tout intervalle de $\mathbb{R}$ contient donc au moins un rationnel et un irrationnel.
On en déduit qu'il y a un rationnel (ainsi qu'un irrationnel) « aussi proche que l'on veut » d'un réel x donné :

$$\text{Soit } x\in\mathbb{R}\quad:\quad\forall\epsilon>0,\exists r\in\mathbb{Q},|x-r|<\epsilon,\qquad\exists\xi\in\mathbb{R}\setminus\mathbb{Q},|x-\xi|<\epsilon$$

5. Bilan

Synthèse

- ⇝ Les ensembles numériques classiques (de $\mathbb{N}$ à $\mathbb{R}$), se déduisent les uns des autres à partir de relation d'équivalence, qui permet d'étendre la relation d'ordre $\leqslant$ toujours totale sur chaque ensemble. Au commencement, l'ensemble $\mathbb{N}$ est la répétition (récursive) de l'addition +1.
- ⇝ Nous proposons ici une construction originale et complète de $\mathbb{R}$, à partir de suite de bissecantes de rationnels. Le processus n'est pas nécessairement à retenir, mais il permet de TOUT démontrer, là où le programme demande d'admettre le théorème de la borne supérieure sur $\mathbb{R}$.
- ⇝ On termine par définir la fonction valeur absolue, la partie entière sur $\mathbb{R}$. On étend aussi la notion d'intervalle numérique en sur-ensemble et sous-ensemble de $\mathbb{R}$.

Savoir-faire et Truc & Astuce du chapitre

- — Savoir-faire - Manipuler des nombres réels
- — Savoir-faire - Travailler avec la partie décimale
- — Savoir-faire - Etudier une borne supérieure

Notations

Notations	*Définitions*	*Propriétés*	*Remarq*
$\lvert x\rvert$	Valeur absolue de x, $\lvert x\rvert = \max(x,-x)$,	$\lvert x\rvert \geqslant 0$	On not $\max(-x$ $x_+ - x_-$
$\lfloor x\rfloor$	Partie entière de x. $\lfloor x\rfloor = \sup\{n \in \mathbb{N} \mid n \leqslant x\}$	$\lfloor x\rfloor \in \mathbb{N}$ et $\lfloor x\rfloor \leqslant x < \lfloor x\rfloor + 1$	On note fraction
$\sup A$	Borne supérieure de A.	Le plus PETIT des éléments plus grand que TOUS les éléments de A	m $\begin{cases} \forall\, a \in \\ M \geqslant \end{cases}$

Retour sur les problèmes

62. Par récurrence...

63. Vu en cours (avec deux relations d'équivalence)

64. Vu en cours (avec une relation d'équivalence)

65. Vu en TD-cours

66. Bien que $\overline{Q} = \mathbb{R}$, on a une bijection de $\mathbb{N}$ sur $\mathbb{Q}$, mais pas de $\mathbb{N}$ sur $\mathbb{R}$.
D'après un théorème de Bernstein, il suffit de montrer qu'il existe une injection de $\mathbb{Q}$ sur $\mathbb{N}$.

On peut prendre $\mathbb{Q} \to \mathbb{N}$, $(\frac{a}{b}) \mapsto 2^{[a\geqslant 0]}3^{|a|}5^b$, injective (non surjective),

on voit qu'en fait pour tout $q \in \mathbb{N}$, $\mathbb{N}^q$ s'injecte dans $\mathbb{N}$ (infinité de nombres premiers).

Par ailleurs, si il existe $\varphi : \mathbb{N} \to [0,1]$ bijective. On note $\varphi(i) = x_i = 0, x_1^i x_2^i \dots x_n^i \cdots \in [0,1]$ (écriture décimale)

Considérons alors le nombre $X = 0, X_1 X_2, \dots X_n \dots$ tel que $X_i \equiv \varphi(i)_i + 5[10]$.

Nécessairement, pour tout $i \in \mathbb{N}$, $\varphi(i)_i \neq X_i$, donc $\varphi(i) \neq X$.

Ainsi φ n'est pas surjective. Contradiction.

Chapitre 15

Divisibilité et congruence sur $\mathbb{Z}$. PGCD & PPCM

Résumé -

Nous plongeons ici dans une des parties mathématiques les plus ancestrales : la théorie des nombres (entiers) ou arithmétique. Beaucoup de résultats présentés ici datent (au moins) d'Euclide (3-ième siècle avant notre ère) : la notion de divisibilité, associée à la division euclidienne.
Etonnament, le meilleur point de vue sur la question est assez récent : il date des Disquisitiones arithmeticae de GAUSS, *publié à l'aube du XIX siècle. Ce point de vue insiste sur les congruences (modulo n) comme des nouvelles égalités.*
Nous nous concentrons ensuite, dans ce chapitre sur la notion de PGCD de deux (ou plusieurs nombres) et l'algorithme d'Euclide pour l'obtenir. Un théorème clé, ignoré des mathématiciens grecs, est la décomposition de (Bachet-)Bézout. C'est le théorème clé de ce chapitre.
On termine par l'étude du PPCM (sorte de complément symétrique du PGCD)
Youtube (parodies ?) :

- *Canal Universaitaire - Arithmétique dans Z - https ://www.canal-u.tv/chaines/canal-unisciel/arithmetique-dans-z/chapitre-arithmetique-partie-1-division-euclidienne-et*
- *Optimal sup-spé - Arithmétique modulaire - https ://www.youtube.com/watch ?v=jZoOGB9WUms*

Sommaire

1. Problèmes

? Problème 70 - Pairs, impairs et . . .

L'addition de deux nombres pairs ou de deux nombres impairs donnent **toujours** un nombre pair.
L'addition d'un nombre pair et d'un nombre impair deux toujours un nombre impair.
Une multiplication donne un nombre impair si et seulement si les deux nombres multipliés sont impairs.
Comment démontrer ces résultats ? Existe-t-il un résultat équivalent pour des multiples de 3, 4, 5, n ?
A propos de multiples, on sait que les nombres divisibles par 9 (et 3) sont exactement ceux dont la somme des chiffres est divisible par 9 (respectivement par 3). Est-ce vrai ? Pourquoi ? Existe-t-il d'autres règles équivalentes ?

? Problème 71 - Table de multiplication modulo n

Lorsqu'on trace les tables des multiplications modulo 5 ou modulo 6, on trouve les deux tableaux suivants :

$\times_{[5]}$	1	2	3	4
1	1	2	3	4
2	2	4	1	3
3	3	1	4	2
4	4	3	2	1

$\times_{[6]}$	1	2	3	4	5
1	1	2	3	4	5
2	2	4	0	2	4
3	3	0	3	0	3
4	4	2	0	4	2
5	5	4	3	2	1

on voit que sur certaines lignes, on retrouve tous les nombres possibles et pas sur d'autres. Pourquoi ?
Existe-t-il des tableaux ou sur toutes les lignes on retrouve tous les nombres (comme celui de la multiplication modulo 5) ?

? Problème 72 - Représentation sous un treilli

La divisibilité est l'exemple typique d'une relation d'ordre non totale.
Le treillis est le bon outil pour visualiser les relations d'ordre non totales.
Est-il possible de convertir tous les théorèmes qui suivent en un schéma visuel basé sur des treillis ?

? Problème 73 - Division euclidienne et PGCD

Si la division de a par b donne combien de fois on peut placer b dans a (quotient) et la place qui reste (reste), on peut continuer l'algorithme en plaçant ensuite r dans b...
On crée ainsi l'algorithme d'Euclide. Est-ce que cela (se) termine ?

. Qu'est-ce qu'on obtient au bout du compte

? Problème 74 - Monde fictif

Dans un monde où il n'y aurait que des pièces de 3 euros et 5 euros, quel montant pourrions-nous payer? (Sachant que le vendeur pourrait nous rendre la monnaie).
Et s'il y a des pièces de 6 et 9 euros?
Et des pièces de 6, 10 et 15 euros?

2. Divisibilité dans $\mathbb{Z}$

2.1. Intégrité de $\mathbb{Z}$ et régularité

Proposition - Intégrité de l'anneau $\mathbb{Z}$
$\mathbb{Z}$ est un anneau intègre.
Formellement :

$$\forall\ a,b\in\mathbb{Z},\qquad a\times b=0\Longrightarrow a=0 \text{ ou } b=0$$

Démonstration
On raisonne par contraposée.
Si a et b sont non nuls.
On suppose que $a>0$ et $b>0$. Donc $a\in\mathbb{N}$.
Par récurrence : $a\times b\geqslant b>0$. Donc $a\times b\neq 0$.
Si $b<0$, de même $a\times b\leqslant b<0$. Donc $a\times b\neq 0$.
Si $a<0$, alors $-a>0$ et avec le même raisonnement $a\times b\neq 0$. □

Comme pour tout anneau intègre :

Proposition - Régularité
Les éléments non nuls de $\mathbb{Z}$ sont réguliers. Formellement :

$$\forall\ a\in\mathbb{Z},\ a\neq 0,\qquad \forall\ b,c\in\mathbb{Z},\ a\times b=a\times c\Longrightarrow b=c$$

Démonstration
Si $ab=ac$, alors par distributivité : $a\times(b-c)=0$.
Puis $a\neq 0$ donc $b-c=0$ i.e. $b=c$. □

Pour aller plus loin - Anneaux, Corps...
Nous reviendrons sur ces propriétés formalisées lorsque nous étudierons les anneaux (euclidiens).
On a vu dans le cours sur les groupes, que l'inversibilité entraine la régularité. On voit ici que la réciproque est fausse.
Il existe donc des anneaux intègres qui ne sont pas des corps

Remarque - A quoi sert cette propriété?
Ce résultat assure l'unicité de la décomposition, si a divise d, il n'y a qu'une décomposition possible de d sous la forme $a\times b$.
Par ailleurs, on remarque aussi, qu'à ce stade, n'avons pas besoin de plonger dans $\mathbb{Q}$ l'inégalité $a\times b=a\times c$ en faisant une division par a. Cela est rassurant.

2.2. Diviseurs, multiples

Définition - Diviseur, multiple
Soit $(a,b)\in\mathbb{Z}^2$. On dit que b *divise* a s'il existe $k\in\mathbb{Z}$ tel que $a=kb$ et on note $b|a$.
On dit aussi que b est un *diviseur* de a, ou que a est un *multiple* de b.
On note $b\mathbb{Z}=\{b\times k;\ k\in\mathbb{Z}\}$ l'ensemble des multiples de b.

Savoir faire - Montrer que b divise a

(Sous-entendu en arithmétique entière) Le plus important n'est pas de montrer l'existence de k tel que b divise a, mais bien montrer qu'il s'agit d'un nombre entier.

Définition - Ensemble des diviseurs

On notera par la suite $\mathscr{D}(a)$ l'ensemble des diviseurs de a. Pour $a \neq 0$, cet ensemble ne contient qu'un nombre fini d'éléments puisque

$$d|a \Rightarrow |d| \leqslant |a|.$$

Application - Majoration du cardinal

On a donc $\text{Card}(\mathscr{D}(a)) \leqslant 2a$. Le pire étant $\mathscr{D}(2) = \{-2;-1;1;2\}$.

Remarque - Le cas de 0 et 1

— 1 et -1 divisent tous les entiers mais ne sont divisibles que par 1 et -1.
— 0 est un multiple de tous les entiers mais n'est diviseur que de lui même.

Définition - Nombres associés

Soit $(a,b) \in \mathbb{Z}^2$. on dit que a et b sont associés si $a|b$ et $b|a$. On a la caractérisation suivante :

$$(a|b \text{ et } b|a) \Leftrightarrow |a| = |b| \Leftrightarrow a = \epsilon b \text{ avec } \epsilon \in \{-1,1\} = \mathbb{Z}^x$$

Proposition - Division de combinaison linéaire

Soit $(a,b) \in \mathbb{Z}^2$. Pour $(u,v) \in \mathbb{Z}^2$, $n \in \mathbb{Z}$, $n \neq 0$ on a :

$$(d|a \text{ et } d|b) \Longrightarrow d|au+bv$$

$$an|bn \Longleftrightarrow a|b$$

Démonstration

Si $d|a$ et $d|b$, alors il existe $a_1, b_1 \in \mathbb{Z}$ tels que $a = da_1$ et $b = db_1$.
Donc $au+bv = d \times (a_1 u + b_1 v)$, et donc d divise $au+bv$.
Si $a|b$, alors il existe $k \in \mathbb{Z}$ tel que $b = ka$, donc $nb = nka$, donc $na|nb$.
et si $na|nb$, alors il existe $k' \in \mathbb{Z}$ tel que $nb = k'na$, donc $b = k'a$, donc $a|b$ □

Exercice

Montrer que la relation « divise » est une relation d'ordre

Correction

Elle est clairement réflexive : $a \times 1 = a$.
Elle est antisymétrique ce qu'on a vu sur les nombres associés.
Elle est transitive : si $a|b$ et $b|c$, alors $b = ka$, $c = k'b$, donc $c = k'ka$, donc $a|c$.

2.3. Division euclidienne de a par b

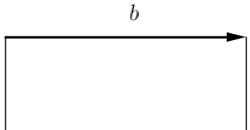
Théorie

Heuristique - La division euclidienne : des soustractions!

A cause de l'algorithme de la division présentée au XIV-ième par Fibonacci, très efficace, on a tendance à considérer la division euclidienne comme une vraie division de a par b, qui s'arrête avant d'écrire les chiffres derrière la virgule.
Mais il est souvent beaucoup plus efficace de considérer l'algorithme d'Euclide comme une succession de soustraction de a par b (ou d'addition de b à a si $a < 0$).
L'algorithme suivant le précise.

Représentation - Première représentation de la division euclidienne

sur l'exemple de 15 divisé par 4 :



Théorème - Division euclidienne
Soient $a \in \mathbb{Z}$, $b \in \mathbb{N}^*$. Il existe un unique couple $(q,r) \in \mathbb{Z} \times \mathbb{N}$ tel que

$$a = bq + r \text{ et } 0 \leqslant r < b.$$

q et r sont appelés respectivement *quotient* et *reste* de la division euclidienne de a par b.
Dans ce cours, nous noterons comme en Python : $a//b$ pour désigner q et $a\%b$ pour désigner r.

Démonstration
Démontrons *l'existence*. On note $R = \{a - bk, k \in \mathbb{Z}\}$.
Alors $R \subset \mathbb{Z}$, et donc $R \cap \mathbb{N}$ est un sous-ensemble de $\mathbb{N}$.
Ce dernier ensemble admet une valeur minimale : $r \geqslant 0$.
Il existe donc $q \in \mathbb{Z}$ tel que $r = a - bq$.
Si $r \geqslant b$, alors $\overline{r} = r - b \geqslant 0$, et $\overline{r} = a - b(q+1) \in R \cap \mathbb{N}$.
on a donc un nouvel élément de $R \cap \mathbb{N}$ plus petit que r.
Cela est impossible, donc $r < b$.
Démontrons *l'unicité*. Si $a = bq + r = bq' + r'$.
On a donc $b(q - q') = r' - r$. Ainsi b divise $r' - r$.
Or $r, r' \in \{0, 1, \dots b-1\}$, on a donc $r' - r \in \{-b+1; -b+2; \dots; -1; 0; 1; \dots b-2; b-1\}$.
Et le seul nombre de cet ensemble divisible par b est 0.
Donc $r - r' = 0$ i.e. $r = r'$, puis $bq = bq' = a - r$ □

Proposition - Critère de divisibilité et division euclidienne
Soient $a \in \mathbb{Z}$, $b \in \mathbb{N}^*$.
$b|a$ si et seulement si le reste de la division euclidienne de a par b vaut 0.

Démonstration
Si le reste vaut 0, alors avec le quotient q : $a = bq$ donc b divise a.
Réciproquement, si b divise a,
il existe $q \in \mathbb{Z}$ tel que $a = bq = bq + 0$.
Par unicité de la division euclidienne, on peut identifier : $r = 0$. □

Algorithme

Cette démonstration n'est pas explicite. On préfèrera le programme suivant :

Informatique - Division euclidienne

```
def div_eucl(a,b):
    """division euclidienne de a par b"""
    #Principe : on soustrait b autant que necessaire a a
    d,k=a,0
    if a<0 :
        c,eps=-b,-1 # si a<0, il faudra additionner b
    else :
        c,eps=b,1
    while d>=b or d<0:
        d=d-c
        k=k+eps #k = nbre de soustractions = quotient
    return(k,d)
```

Application - Déroulement de `div_eucl(12,5)` et `div_eucl(-12,5)`

Pour suivre un algorithme, on fait un tableau :

d	k	c	eps	test
12	0	5	1	$12 > 5$: vrai
7	1	5	1	$7 > 5$: vrai
2	2	5	1	$2 < 5$ et $2 > 0$: faux

d	k	c	eps	test
-12	0	-5	-1	$-12 < 0$: vrai
-7	-1	-5	-1	$-7 < 0$: vrai
-2	-2	-5	-1	$-2 < 0$: vrai
3	-3	-5	-1	$3 > 0$ et $3 < 5$: faux

La première application du programme renvoie $(2,2)$. C'est correct : $12 = 2\times 5+2$
La seconde application du programme renvoie $(-3,3)$. C'est correct : $-12 = -3\times 5+3$

Informatique - Récursivité

Une autre possibilité est d'exploiter la récursivité :

```
def div_eucl_rec(a,b):
    """calcul de la division euclidienne de a par b, par recursivite"""
    if a<b and a>-1:
        return(0,a)
    elif a>b :
        m,n=div_eucl_rec(a-b,b)
        return(m+1,n)
    else :
        m,n=div_eucl_rec(a+b,b)
        return(m-1,n)
```

Pour aller plus loin - Démontrer qu'un programme fait bien ce qu'il faut...

Pour démontrer qu'un programme fait bien ce qu'il faut, on a besoin :

- *d'un variant de boucle.*
  Il nous assure la terminaison du programme.
  Ici on prend `d`
- *d'un invariant de boucle.*
  Connaissant sa valeur initiale et finale, on obtient le résultat final donné par l'algorithme. On compare avec le résultat attendu.
  Ici, on considère `kb+d=a`

2.4. Arithmétique modulaire

Relation d'équivalence

Définition - a congru à b modulo n
Soient $n\in\mathbb{N}$, $a,b\in\mathbb{Z}$.
On dit que *a est congru à b modulo n* si n divise $a-b$ ($\iff a-b\in n\mathbb{Z}$), c'est-à-dire s'il existe $k\in\mathbb{Z}$ tel que $a=b+kn$.
On note $a\equiv b[n]$.

Pour tout $n\in\mathbb{N}^*$, la relation de congruence modulo n est une relation d'équivalence (nous l'avons déjà vu).

Remarque - $\mathbb{U}_n$

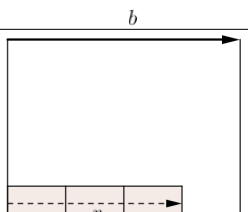
Le groupe des racines n-ieme de l'unité (avec la multiplication) est bien un lieu où les congruences modulo n sont naturelles.
En effet, si on note $\xi_r = \exp(\frac{2ri\pi}{n}$, on a $\xi_r=\xi_{r'}$ ssi $r\equiv r'[n]$.
Puis, si on considère $\xi_r\times\xi_s$, on trouve $\xi_{r+s}=\xi_{(r+s)\%n}$.
La proposition suivante donne un système de représentant naturel (et donc le nombre de classe d'équivalence)

Proposition - Reste
Soit $n\in\mathbb{N}^*$. Pour tout $a,b\in\mathbb{Z}$
$$a\equiv b[n]\iff a\%n = b\%n.$$
Ainsi, $[\![0,n-1]\!]$ est un système de représentant de $\dfrac{\mathbb{Z}}{\cdot\equiv\cdot[n]}$, ensemble possédant donc n éléments

Démonstration
Pour tout $a,b\in\mathbb{Z}$, $a=(a//n)\times n+(a\%n)$, et $b=(b//n)\times n+(b\%n)$,
donc $a\equiv(a\%n)[n]$ et $b\equiv(b\%n)[n]$.
Par transitivité :
$$a\equiv b[n]\iff (a\%n)\equiv(b\%n)[n]$$
Or $(a\%n)-(b\%n)\in[\![-(n-1),n-1]\!]$ et donc $n|((a\%n)-(b\%n))\implies(a\%n)-(b\%n)=0$.
Et donc $a\equiv b[n]\iff(a\%n)=(b\%n)$ □

Représentation - Voir les congruences modulo n

Si on prend l'habitude de regarder les congruences modulo n dans un damier (comme lors de la première représentation).
Alors $a\equiv b[n]$ ssi le denier carré de a et de b est dans la même colonne. Ici $15\equiv 7[4]$.



Arithmétique modulaire

Proposition - Opérations : arithmétique modulaire
Soit $n\in\mathbb{N}$. La congruence modulo n est compatible avec l'addition et la multiplication :

Pour a, a', b, b' entiers relatifs on a

$$\begin{cases} a \equiv a'[n] \\ b \equiv b'[n] \end{cases} \Longrightarrow \begin{cases} a+b \equiv a'+b'[n] \\ a \times b \equiv a' \times b'[n] \end{cases}$$

Démonstration
Si $a = a' + kn$, $b = b' + hn$.
Alors $a + b = a' + b' + (k+h)n$ et $a \times b = a'b' + n(a'h + b'k + nhk)$.
□

Truc & Astuce pour le calcul - Réduction modulo n
La réduction modulo n réduit les calculs : les nombres ne dépassent pas la valeur n.
En revanche, on perd des informations ou bien parfois, on supprime les informations inutiles (à quoi sert de connaître exactement un nombre, alors que seul son dernier chiffre est demandé ?)

Remarque - Vérifier un calcul
Dans l'art de calculer, nous avons vu l'importance d'avoir des clés de vérification de calcul.
Ainsi d'après la formule précédente, si $a \equiv a'[n]$, alors $a^k \equiv a'^k[n]$,
et par linéarité, pour tout polynôme P : $P(a) \equiv P(a')[n]$.
Par exemple, montrer qu'une racine entière x d'un polynôme $P = a_0 + a_1 X + \cdots + a_d X^d \in \mathbb{Z}[X]$ vérifie : $x | a_0$.
Il suffit d'écrire : $P(x) = 0 = a_0 + a_1 x + \cdots + a_d x^d \equiv a_0[x]$.

Exercice
Avons-nous l'équivalence $a \equiv b[n] \iff ca \equiv cb[n]$?
Quelle est l'implication. Donner un contre-exemple de l'implication réciproque. Une condition pour l'équivalence ?

Correction
On a d'après la proposition précédente : pour tout $c \in \mathbb{Z}$, $a \equiv b[n] \Longrightarrow ca \equiv cb[n]$.
La réciproque est fausse : $0 \times 1 \equiv 0 \times 2[3]$, mais $1 \not\equiv 2[3]$.
Si $c \wedge n = 1$, alors $ca \equiv cb[n] \Longrightarrow n | c(a-b) \Longrightarrow n | (a-b)$ (Gauss) $\Longrightarrow a \equiv b[n]$.

Remarque - Réduction modulo p, $p \in \mathscr{P}$
Si p est premier, donc premier avec tous les nombres, pour tout $a \in \mathbb{Z}$, $a \wedge p = 1$, donc d'après le théorème de Bézout il existe $u, v \in \mathbb{Z}$ tels que $ua + vp = 1$, donc modulo p : $au \equiv 1[p]$.
Ainsi, a est inversible. Donc $\frac{\mathbb{Z}}{p\mathbb{Z}}$ est un corps (tous les éléments sont inversibles, donc régulier.

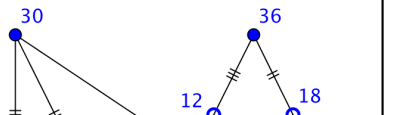
3. Plus Grand Commun Diviseur de deux nombres

3.1. $PGCD$ de deux nombres. Définition « naturelle »

Il s'agit du plus grand pour la relation d'ordre classique : $\leqslant$.

Analyse - Construction du $PGCD$
Soient a et b deux entiers relatifs non nuls.
L'intersection des deux ensembles des diviseurs de a et de b : $\mathscr{D}(a) \cap \mathscr{D}(b)$ est non vide.
En effet, 1 appartient à ces deux ensembles.
Puis, l'ensemble des diviseurs entiers de a et de b : $\mathscr{D}(a) \cap \mathscr{D}(b) \cap \mathbb{N}$ est un sous-ensemble de $\mathbb{N}$, non vide majorée par $|a|$ (et/ou $|b|$).
Cet ensemble admet donc un plus grand élément, c'est le plus grand diviseur commun à a et b.

Pour aller plus loin - Treillis des diviseurs
Il est possible de représenter sous forme de treillis les diviseurs de 30 et 36 et chercher le PGCD.
C'est le plus grand pour $a \leqslant b$, mais aussi $a|b$



Définition - $PGCD$ de a et b
Soient a et b deux entiers relatifs non nuls.
On appelle $PGCD$ (Plus Grand Commun Diviseur) de a et b, le nombre
$$PGCD(a,b) = \max\left(\mathscr{D}(a) \cap \mathscr{D}(b) \cap \mathbb{N}\right).$$
On le note également $a \wedge b$.
On a clairement $b \wedge a = a \wedge b = |a| \wedge |b|$.
Par convention on pose pour $a \in \mathbb{Z}$, $a \wedge 0 = |a|$ (y compris si $a = 0$).

Exemple - $a = 36$, $b = 30$
$\mathscr{D}(a) = \{-36,-18,-12,-9,-6,-3,-2,-1,1,2,3,4,6,9,12,18,36\}$.
$\mathscr{D}(b) = \{-30,-15,-10,-6,-5,-3,-2,-1,1,2,3,5,6,10,15,30\}$.
$\mathscr{D}(a) \cap \mathscr{D}(b) \cap \mathbb{N} = \{1,2,3,6\}$,
et donc $PGCD(36,30) = \max\left(\mathscr{D}(a) \cap \mathscr{D}(b) \cap \mathbb{N}\right) = \max(\{1,2,3,6\}) = 6$

Remarque - Divisibilité des nombres de $\mathscr{D}(30) \cap \mathscr{D}(36)$
6, le $PGCD(36,30)$ est divisible par tous les éléments de $\mathscr{D}(30) \cap \mathscr{D}(36)$.
Et c'est le seul (avec son associé : -6) !

Remarque - Algorithme des facteurs premiers
Plus jeunes, les étudiants exploitaient l'algorithme de la décomposition en facteurs premiers.
Il s'agit de faire deux listes : celles des facteurs premiers de chacun des deux nombres.
Puis on multiplie tous ceux qui sont en commun (avec leur multiplicité) pour obtenir le PGCD.
Tant que nous ne savons pas ce qu'est un nombre premier, cet algorithme attendra...

3.2. Algorithme d'Euclide

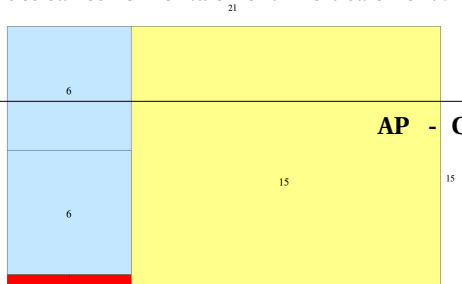
Algorithme des divisions successives pour obtenir le $PGCD$

Lemme - Stabilité par division euclidienne
Soit $(a,b) \in \mathbb{Z} \times \mathbb{N}^*$.
Si $a = bq + r$, alors $\mathscr{D}(a) \cap \mathscr{D}(b) = \mathscr{D}(r) \cap \mathscr{D}(b)$ et donc $a \wedge b = b \wedge r$.

Démonstration
Supposons que $a = bq + r$.
Si $d \in \mathscr{D}(a) \cap \mathscr{D}(b)$,
$d|a$ et $d|b$, alors $d|1a - qb = r$, donc $d \in \mathscr{D}(r) \cap \mathscr{D}(b)$.
Réciproquement, si $d \in \mathscr{D}(r) \cap \mathscr{D}(b)$,
$d|r$ et $d|b$, alors $d|qb + 1r = a$, donc $d \in \mathscr{D}(b) \cap \mathscr{D}(a)$.
Donc $\mathscr{D}(a) \cap \mathscr{D}(b) \cap \mathbb{N} = \mathscr{D}(b) \cap \mathscr{D}(r) \cap \mathbb{N}$
et nécessairement, par égalité des maximum : $a \wedge b = b \wedge r$ □

Heuristique - Algorithme pour obtenir le $PGCD$ de deux nombres
La proposition qui donne un algorithme qui exploite une suite de division euclidienne pour trouver le $PGCD(a,b)$.
Il permet également d'obtenir les coefficients de Bézout des nombres a et b.
On notera qu'il s'applique à deux nombres a et b vérifiant : $0 < b < a$. Toute recherche de $PGCD(a,b)$ se ramène à ce cas là, en effet :
- Si $b = 0$ alors $PGCD(a,b) = |a|$
- Si $b \neq 0$, notons alors que $PGCD(a,b) = PGCD(|a|,|b|) = PGCD(|b|,|a|)$,
on peut donc supposer que les deux nombres sont positifs et que le premier est le plus grand.

Représentation - Illustration de l'algorithme d'Euclide (Wikipedia)
pour $a = 21$ et $b = 15$. On complète les trous par des carrés horizontalement↔verticalement :



Proposition - Algorithme d'Euclide
Soient $a, b \in \mathbb{N}^*$. Supposons que $0 < b < a$. L'algorithme d'Euclide consiste en un succession de division euclidienne :
- On commence par poser $r_0 = a$ et $r_1 = b$;
- ensuite, k désignant un entier naturel non nul (étape de l'algo-

rithme),
tant que $r_{k+1} \neq 0$,
on note r_{k+2} le reste de la division euclidienne de r_k par r_{k+1}
(on a donc $r_{k+2} < r_{k+1}$).
Il existe $N \in \mathbb{N}^*$ tel que $r_N = 0$,
r_{N-1} est alors le dernier reste non nul de la suite (r_k), et : $a \wedge b = r_{N-1}$

Démonstration
Il faut démontrer qu'il existe bien N tel que $r_N = 0$.
En fait, à l'issue de l'algorithme, on crée une suite strictement décroissante d'entiers : $r_0 \geqslant r_1 > r_2 > \cdots \geqslant 0$.
Une telle suite est nécessairement nulle à partir d'un certain rang : il existe $N \in \mathbb{N}^*$ tel que $r_N = 0$.
(Mieux : ce rang maximal $N \leqslant r_0$).
Il reste ensuite à appliquer le lemme stabilité par division euclidienne :

$$PGCD(a,b) = PGCD(r_0,r_1) = PGCD(r_1,r_2) = \cdots = PGCD(r_{N-1},r_N)$$
$$= PGCD(r_{N-1},0) = r_{N-1}$$

□

Application - $PGCD(1542,58) = 2$
On calcule une succession de division euclidienne, le diviseur et le reste de la division euclidienne k deviennent respectivement le dividende et le diviseur de la division euclidienne $k+1$:

$1542 = 26\times58+34$, $58 = 1\times34+24$, $34 = 1\times24+10$, $24 = 2\times10+4$, $10 = 2\times4+2$, $4 = 2\times2+0$

d'où

$PGCD(1542,58) = PGCD(58,34) = PGCD(34,24) = PGCD(24,10) = PGCD(10,4) = PGCD(4,2) = PGCD(2,0) = 2.$

La division euclidienne s'exploite en pratique, c'est-à-dire avec un calcul réel, à savoir-faire (on fait plutôt des divisions que des soustractions). Ou bien

Savoir faire - Exploiter la division euclidienne (théorie)
Si $f : \mathbb{Z}^2 \to E$ (E quelconque) tel que $f(a,b) = f(b, a\%b)$, alors $f(a,b) = f(a \wedge b, 0)$.

Truc & Astuce pour le calcul - Trouver son erreur dans un algorithme d'Euclide
Une fois l'algorithme terminé, il est important de vérifier que le dernier reste non nul est bien un diviseur de a et de b.
Si ce n'est pas le cas, il faut revenir à la ligne de calcul où le reste obtenu n'est pas divisible par le PGCD-candidat.

Informatique - Division euclidienne
On peut écrire l'algorithme d'Euclide sous Python. Remarquons que dans le programme ainsi écrit, on tient compte des cas $b = 0$ et a ou b négatif.

```
1  def alg_eucl(a,b):
2      """pgcd(a,b) par algorithme d'Euclide"""
3      if b==0:
4          return(a)
5      a1,a2=max(abs(a),abs(b)),min(abs(a),abs(b))
6      ra,rb=a1,a2
7      while rb!=0:
8          rc=div_eucl(ra,rb)[1]
9          ra,rb=rb,rc
10     return(ra)
```

Pour aller plus loin - Fractions continues
L'algorithme d'Euclide a une autre application importante : la décomposition en fractions continues. C'est clairement le cas concernant les fractions comme le montre l'exemple :

$$\frac{1542}{58} = 26 + \cfrac{1}{1+\cfrac{1}{1+\cfrac{1}{2+\cfrac{1}{2+\frac{1}{2}}}}}$$

Mais c'est aussi le cas concernant les nombres irrationnels

3.3. Couple de Bézout

Exploitation plus fine de l'algorithme d'Euclide

Analyse - Exploitation plus approfondie encore de l'algorithme d'Euclide

Lorsqu'on applique l'algorithme d'Euclide, on se trouve avec une suite de relations :

$$r_k = q_k r_{k+1} + r_{k+2} \qquad 0 \leqslant r_{k+2} < r_{k+1}$$

Avec $r_0 = a > b$, $r_1 = b > 0$, $r_N = 0$ et $r_{N-1} = a \wedge b$.
On peut alors affirmer (constructivement) qu'il existe u_k et $v_k \in \mathbb{Z}$ tels que

$$\forall\, k \in \mathbb{N} \qquad r_k = u_k a + v_k b$$

On a, en effet :

- $r_0 = a$, donc $u_0 = 1$ et $v_0 = 0$,
- $r_1 = b$, donc $u_1 = 0$ et $v_1 = 1$,
- si $r_k = u_k a + v_k b$ et $r_{k+1} = u_{k+1} a + v_{k+1} b$, alors

$$r_{k+2} = r_k - q_{k+1} r_{k+1} = (u_k - q_{k+1} u_{k+1}) a + (v_k - q_{k+1} v_{k+1}) b$$

Et en particulier pour $k = N-1$:

$$\exists\, u, v \in \mathbb{Z} \mid a \wedge b = ua + vb$$

Si $a < b$, ou $a < 0 \ldots$, on applique l'algorithme à $|a|$ et $|b|$ et si nécessaire on multiplie u et v par -1.

Théorème - Coefficients de Bézout
Soit $(a, b) \in \mathbb{Z}^2$. Il existe des entiers u et v tels que $au + bv = a \wedge b$.
Un tel couple (u, v) est appelé un couple de coefficients de Bézout de a et b.

Histoire - Bézout
Etienne Bézout (1730-1783) est un mathématicien français. Il généralisa l'identité de Bachet, c'est pourquoi elle lui est couramment attachée.

Histoire - Bachet de Méziriac ou Bézout
Le théorème de Bézout est en fait obtenu pour la première fois par Gaspard Bachet de Méziriac (1581-1638). C'est un mathématicien, poète et traducteur français. Il a en particulier traduit l'arithmetica de Diophante (là où Fermat a écrit l'énoncé de son grand théorème), et est l'auteur de Problèmes plaisants et délectables qui se font par les nombres

Savoir faire - Pas unicité du couple de Bézout. Les obtenir tous

Il n'y a pas unicité du couple (u, v) :
si (u_0, v_0) est un couple de Bezout, en divisant par $\delta = a \wedge b$:

$$ua + bv = u_0 a + v_0 b \Rightarrow \frac{a}{\delta}(u - u_0) = \frac{b}{\delta}(v_0 - v) \underbrace{\Rightarrow}_{\text{lemme de Gauss}} \frac{a}{\delta} | (v - v_0) \ldots$$

Alors pour tout $n \in \mathbb{Z}$, $(u_0 + n\frac{b}{\delta}, v_0 - n\frac{a}{\delta})$ en est aussi un .

Informatique - Division euclidienne

En étendant la division euclidienne (elle garde en mémoire les calculs des quotients), on peut obtenir les coefficients de Bézout

```
def Bezout(a,b):
    """pgcd(a,b)+coefficient de Bezout"""
    if b==0:
        return(a)
    a1,a2=max(abs(a),abs(b)),min(abs(a),abs(b))
    u,uu=1,0
    v,vv=0,1
    ra,rb=a1,a2
    while rb!=0:
        q,rc=div_eucl(ra,rb)
        ra,rb=rb,rc
        u,uu=uu,u-q*uu
        v,vv=vv,v-q*vv
    return(ra,u,v)
```

Truc & Astuce pour le calcul - Obtenir un couple de Bézout

En pratique, il y a deux stratégies.

• Celle plutôt vue en terminale, assez naturelle et peut-être bien ancrée en vous. On trouve (u,v) en EN REMONTANT l'algorithme d'Euclide.

(à la main, il vaut mieux partir des valeurs numériques, elles n'arrivent qu'en fin d'algorithme).

Par exemple pour 1542 et 58 on a (en remontant les division euclidienne) :

$$\begin{aligned}2 &= 10-2\times 4 = 10-2(24-2\times 10) = -2\times 24+5\times 10 = -2\times 24+5\times(34-24)\\ &= 5\times 34-7\times 24 = 5\times 34-7\times(58-34) = -7\times 58+12\times 34\\ &= -7\times 58+12\times(1542-26\times 58) = 12\times 1542-319\times 58\end{aligned}$$

Donc $1542u+58v=2$, avec (par exemple) : $u=12$ et $v=-319$.

• Celle qui découle ici, avec un tableau à remplir directement :

On rappelle qu'à chaque étape $k\in\mathbb{N}^*$: $r_{k+1}=q_kr_k+q_{k-1}$ et (u_k) et (v_k) vérifient la même relation de récurrence $x_{k+1}=-q_kx_k+x_{k-1}$.

k	r_k	q_k	u_k	v_k	$1542\times u_k+58\times v_k=r_k$
0	1542		1	0	1542
1	58	26	0	1	58
2	34	1	1	−26	34
3	24	1	−1	27	24
4	10	2	2	−53	10
5	4	2	−5	133	4
6	2	0	12	−319	2

Algorithme d'Euclide, arithmétique modulaire et carrelage

Analyse - Relation de Bézout modulaire

Si on a $ua+bv=a\wedge b$, on pourrait passer cette relation modulo a :

$$bv\equiv a\wedge b[a]$$

Si la relation modulo a permet de trouver $a\wedge b$ et v (et u), cela peut valoir le coup de représenter l'algorithme d'Euclide.

Exercice

Donner une représentation théorique de la recherche du PGCD de a et de b, par algorithme d'Euclide dans un carrelage.

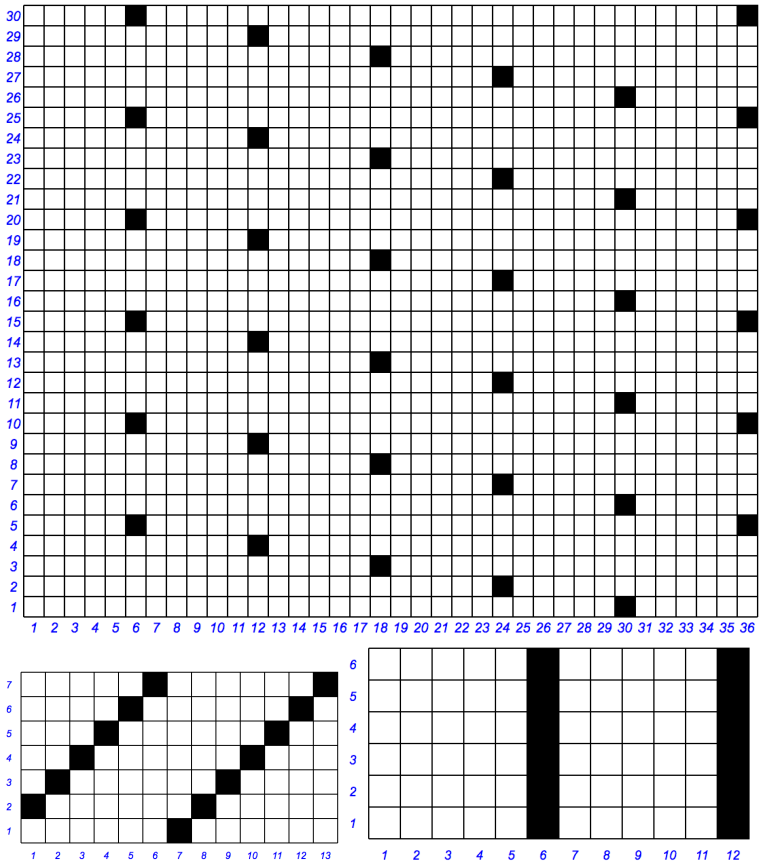
On pourra appliquer la méthode pour donner le $PGCD$ et les coefficient de Bézout pour les couples $(13,7)$ et $(12,6)$

Correction

On cherche le PGCD de a et de b.

On se place sur un carrelage de taille $a\times b$ et on repère les carreaux tous les b déplacements. Ainsi on trouve les restes $k\times b$ divisé par a, pour k de 1 à a.

Le figure obtenu nous donne des informations. Voici trois exemples avec $a=36$ et $b=30$ - $a=13$ et $b=7$ et $a=12$ et $b=6$.



Soient a et $b \in \mathbb{N}$.
Pour tout $u, v \in \mathbb{Z}$, $ua + vb = c \Rightarrow vb \equiv c[a]$
et donc la v-ième case noircie sur le carrelage se trouve en colonne c (reste de la division).
Or on sait que $a \wedge b = \min\{au + bv, u, v \in \mathbb{Z}\}$.
Donc le numéro de la colonne ayant la case noircie la plus à gauche indique le $PGCD$ de a et b.
(Il faut montrer que $\{au + bv, u, v \in \mathbb{Z}\} = \{au + bv, u \in \mathbb{Z}, v \in [\![1, a]\!]\}$, mais c'est possible...).
Puis v, le facteur devant b, donne le numéro de la case noircie (c'est la v-ième case noircie).
Enfin, si $ua + bv = c$, on a $bv = -ua + c$
et donc $-u = q$ est le quotient de la division euclidienne de vb par a.
Or ce quotient est exactement égale à ℓ, le numéro de la ligne -1. Donc $u = -q = -(\ell - 1) = 1 - \ell$
On remarque ainsi que le $PGCD$ des deux nombres a et b est toujours égal au numéro de la colonne noircie la plus à gauche. On trouve alors u et v en faisant :

- u (coeffcient de a) : est l'opposé du numéro de la colonne de la case noiricie indiquant le $PGCD$ +1.
- v (coefficient de b) : est le numéro de la case noircie, rappelons que l'on va d'abord de gauche à droite, puis de bas en haut.

Sur le premier exemple, la case noircie la plus à gauche, puis la plus basse se trouve en $(6,5)$. Donc $36 \wedge 30 = 6$.
Puis $u = -5 + 1 = -4$ et $v = 5$, car c'est case noircie est la 5-ième.
Et en effet : $-4 \times 36 + 5 \times 30 = -144 + 150 = 6$ On le lit sur le schéma :

- $13 \wedge 7 = 1$ et on a $-1 \times 13 + 2 \times 7 = 1$.
- $12 \wedge 6 = 6$ et on a $0 \times 12 + 1 \times 6 = 6$.

3.4. Deux caractérisations essentielles du PGCD

Caractérisation par divisibilité

En fait la définition de plus grand diviseur commun peut s'entendre d'une seconde façon : plus grand pour la relation d'ordre de divisibilité.

Remarque - Relation binaire

La relation binaire "est un diviseur de", définie par $a \mathcal{R} b \Leftrightarrow a|b$, est une relation d'ordre partiel sur $\mathbb{N}$.
Pour cette relation d'ordre $\mathcal{D}(a) \cap \mathcal{D}(b)$ admet un plus grand élément qui est $a \wedge b$.

Proposition - Caractérisation essentielle du $PGCD$
Soit $(a,b)\in\mathbb{Z}^2$.
Alors $a\wedge b$ est le seul entier naturel dont les diviseurs sont exactement les diviseurs communs à a et b :

$$\mathscr{D}(a\wedge b)=\mathscr{D}(a)\cap\mathscr{D}(b)$$

autre façon de l'écrire :

$$\forall d\in\mathbb{Z},\ (d|a \text{ et } d|b)\iff d|a\wedge b$$

Remarque - Force de cette propriété
Dans l'équivalence trouvée (ou la double inclusion d'ensemble),
une équivalence est triviale : $d|a\wedge b\implies d|a$ et $d|b$.
L'usage fréquent que l'on fera donc de cette proposition est : $d|a$ et $d|b\implies d|a\wedge b$.

Démonstration
D'après le lemme de stabilité :

$$\mathscr{D}(a)\cap\mathscr{D}(b)=\mathscr{D}(b)\cap\mathscr{D}(r)=\cdots=\mathscr{D}(r_{N-1})\cap\mathscr{D}(r_N)$$

(avec les notations de l'algorithme d'Euclide.
Or $r_{N+1}=0$, donc r_N divise r_{N-1}, et $\mathscr{D}(r_{N-1})\cap\mathscr{D}(r_N)=\mathscr{D}(r_N)$.
Et comme $r_N=a\wedge b$, on en conclue :

$$\mathscr{D}(a\wedge b)=\mathscr{D}(a)\cap\mathscr{D}(b)$$

En ce qui concerne l'équivalence :

$$d|a \text{ et } d|b)\iff d\in\mathscr{D}(a)\cap\mathscr{D}(b)=\mathscr{D}(a\wedge b)\iff d|a\wedge b.$$

Autre démonstration possible :
Si $d|a\wedge b$, alors comme $a=a'(a\wedge b)$, alors $d|a$, de même $d|b$.
Si $d|a$ et $d|b$, alors pour tout u,v, $d|au+bv$, en particulier avec le couple de Bézout. □

Savoir faire - Trouver un PGCD. Exploiter un PGCD.
On note $\delta=a\wedge b$.
Pour trouver le PGCD de a et b :
on démontre que si $m|a$ et $m|b$, alors nécessairement $m|\delta$.
Puis on cherche m le plus grand possible.
(si $m=1$, a et b sont premiers entre eux).
Pour exploiter le PGCD de a et b :
on exploite le fait que si $m|\delta$, alors $m|a$ et $m|b$.
Et on travaille sur cette co-divisibilité.

Caractérisation par combinaison linéaire

Proposition - Combinaison linéaire
Soient $a,b\in\mathbb{Z}$.
On note $a\mathbb{Z}+b\mathbb{Z}$ l'ensemble $\{au+bv,(u,v)\in\mathbb{Z}\}$ des combinaisons linéaires de a et b.
Alors

$$a\mathbb{Z}+b\mathbb{Z}=(a\wedge b)\mathbb{Z}$$

Démonstration
Notons $\delta=a\wedge b$. D'après le proposition de Bézout, il existe $u_0,v_0\in\mathbb{Z}$ tels que $\delta=au_0+bv_0$.
Comme $\delta|a$ et $\delta|b$, alors pour tout $u,v\in\mathbb{Z}$, $\delta|au+bv$,
donc pour tout $u,v\in\mathbb{Z}$ $au+bv=\delta w\in\delta\mathbb{Z}$.
Ainsi $a\mathbb{Z}+b\mathbb{Z}\subset\delta\mathbb{Z}$.
Réciproquement, pour tout $w\in\mathbb{Z}$, $\delta w=a(u_0w)+b(v_0w)\in a\mathbb{Z}+b\mathbb{Z}$.
Donc $\delta\mathbb{Z}\subset a\mathbb{Z}+b\mathbb{Z}$ □

Pour aller plus loin - Construction classique
La tradition mathématique utilise donc une autre stratégie pour ***définir*** *le PGCD de a et b :*

1. *on s'intéresse à l'ensemble $a\mathbb{Z}+b\mathbb{Z}=\{ua+vb,u,v\in\mathbb{Z}\}$*
2. *on* ***démontre*** *qu'il existe $\delta\in\mathbb{N}^*$ tel que $a\mathbb{Z}+b\mathbb{Z}=\delta\mathbb{Z}$*

On appelle alors δ le PGCD de a,b

Savoir faire - Combinaisons linéaires (entières)
Dès que l'on a des combinaisons linéaires d'entiers (on appelle cela un réseau d'entiers), il faut penser que la maille élémentaire est donnée par le PGCD des nombres. C'est par exemple, le cas du problème « monde fictif »

4. Entiers premiers entre eux. Factorisation

4.1. Définition et critère de Bézout

Définition - Couple d'entiers premiers entre eux
a et b sont dits *premiers entre eux* si $a \wedge b = 1$.

On note que pour le théorème suivant, nous avons bien une équivalence :

Théorème - Théorème (ou identité) de Bézout
Soient a et b deux entiers relatifs non nuls. Alors

$$a \wedge b = 1 \iff \exists (u,v) \in \mathbb{Z}^2 \text{ tel que } au + bv = 1$$

Démonstration
On applique la décomposition de Bézout : si $a \wedge b = 1$ alors $\exists (u,v) \in \mathbb{Z}^2$ tel que $au + bv = 1$.
Réciproquement si $au + bv = 1$ alors $d|a$ et $d|b \Rightarrow d|au + bv = 1$ donc $d = 1$. □

Comme il est plus simple de travailler avec des nombres premiers entre eux, on exploite souvent le savoir-faire suivant :

Savoir faire - Réduction à des entiers premiers entre eux (1)
Soient $a, b \in \mathbb{Z}$. On note $\delta = a \wedge b$.
Alors, $a = \delta a'$ et $b = \delta b'$, avec $a' \wedge b' = 1$.
Puis on travaille avec les a' et b'

4.2. Lemme de Gauss et décomposition en facteurs relativement premiers

Enoncé

Théorème - Lemme de Gauss
Soient $a, b, c \in \mathbb{Z}$. Alors

$$(a \wedge b = 1 \text{ et } a|bc) \Longrightarrow a|c.$$

Démonstration
Si $a \wedge b = 1$ alors $\exists (u,v) \in \mathbb{Z}^2$ tel que $au + bv = 1$
d'où $acu + bcv = c$ et $a|bc \Longrightarrow a|bcv \Longrightarrow a|c - acu$.
Or $a|acu$ d'où $a|c = (c - acu) + acu$. □

Facteur relativement premier

Proposition - Facteur relativement premier
Soient $a, b, c \in \mathbb{Z}$. Alors

$$(a \wedge b = 1 \text{ et } a \wedge c = 1) \Longrightarrow a \wedge bc = 1 \text{ (réciproque vraie)}$$

Histoire - Disquisitiones Arithmeticae
En 1801, Gauss âgé de 24 ans publie les Disquisitiones Arithmeticae (recherches arithmétiques) et révolutionne totalement le genre. Pour la première fois, on pense les nombres à l'aide de l'arithmétique modulaire (congruence - voir fin de chapitre) !

$$(a \wedge b = 1,\ a|c,\ b|c) \Longrightarrow ab|c$$

Démonstration
D'après l'identité de Bézout il existe u, v, u', v' tels que $au + bv = 1$ et $au' + cv' = 1$,
d'où par produit $a(u(au' + cv')) + a(bu'v) + bcvv' = 1$,
ce qui peut s'écrire $aU + bcV = 1$ avec $(U, V) \in \mathbb{Z}^2$,
et par réciproque de l'identité de Bézout $a \wedge bc = 1$.
Pour la seconde implication : $a|c$ donc $c = aq$; $b|c = aq$ et $b \wedge a = 1$
donc $b|q$; $q = bp$ et $c = abp$ d'où $ab|c$. □

Exercice
On note $\mathscr{P}_a$, l'ensemble des nombres premiers avec a et $a\mathbb{Z}$, les multiples de a.
Ré-écrire les théorèmes de GAUSS avec ces ensembles.

Correction
Le lemme de Gauss : $b \in \mathscr{P}_a$. $bc \in a\mathbb{Z} \Rightarrow c \in a\mathbb{Z}$.
Facteur (1) : $b \in \mathscr{P}_a, c \in \mathscr{P}_a \Rightarrow bc \in \mathscr{P}_a$.
Facteur (2) : $b \in \mathscr{P}_a (\Rightarrow a \in \mathscr{P}_b)$,$c \in a\mathbb{Z} \cap b\mathbb{Z} \Rightarrow c \in ab\mathbb{Z}$ ou $a, b \in \mathscr{D}(c) \Rightarrow ab \in \mathscr{D}(c)$.

Plusieurs facteurs (relativement premiers)

Corollaire - Facteurs premiers
Soient $a, c, b_1, \ldots, b_n$ des entiers relatifs.

$$(\forall i \in [\![1, n]\!], a \wedge b_i = 1) \Longrightarrow a \wedge \prod_{i=1}^{n} b_i = 1$$

$$(\forall (i, j) \in [\![1, n]\!]^2, i \neq j \Rightarrow b_i \wedge b_j = 1 \text{ et } \forall i \in [\![1, n]\!], b_i|c) \Longrightarrow \prod_{i=1}^{n} b_i \Big| c$$

Démonstration
Il suffit de faire une récurrence sur $n = \text{Card}(\{b_i\})$ □

Corollaire - Forme irréductible d'un rationnel
Soit $r \in \mathbb{Q}$. Il existe un unique couple $(p, q) \in \mathbb{Z} \times \mathbb{N}^*$ tel que $r = \dfrac{p}{q}$ et tel que p et q soient premiers entre eux.
Cette écriture est appelée la forme irréductible de r.
Les autres écritures fractionnaires sont de la forme $r = \dfrac{\lambda p}{\lambda q}$ avec $\lambda \in \mathbb{Z}^*$.

Démonstration

— existence : $r = \dfrac{a}{b} = \dfrac{dp}{dq}$ avec $d = PGCD(a, b)$ et donc $p \wedge q = 1$.

— unicité : $\dfrac{p}{q} = \dfrac{p'}{q'}$ avec $p \wedge q = p' \wedge q' = 1$ implique $pq' = p'q$ et donc $p|p'$ (car $p \wedge q = 1$) et $p'|p$ (car $p' \wedge q' = 1$), d'où $p = p'$ et $q = q'$.

□

5. Généralisation à plusieurs entiers

5.1. $PGCD$ d'un nombre fini d'entiers relatifs

Heuristique - Définition

Au lieu de construire le $PGCD$ de k nombres en prenant :

$$\delta = \max\left(\mathscr{D}(a_1) \cap \mathscr{D}(a_2) \cdots \mathscr{D}(a_k) \cap \mathbb{N})\right)$$

nous allons prendre pour définition, une méthode récursive.
On notera ensuite l'extension de la caractéristique vue par la suite :

$$d|a \text{ et } d|b \Longrightarrow d|\delta$$

Dans ce cas le terme « plus grand » ne doit pas être pris pour la relation d'ordre $n \leqslant n$, mais pour la relation d'ordre $n|m$.

Pour aller plus loin - Lien PGCD et inf

inf E *est le plus grand des minorants de* E.

- $\forall\, x \in E$, $\inf E \leqslant x$
- $\forall\, m$ *tel que* $\forall\, x \in E$, $m \leqslant x$, *alors* $m \leqslant \inf E$.

$PGCD(E)$ *est le plus grand des diviseurs de* E.

- $\forall\, x \in E$, $PGCD(E) \mid x$
- $\forall\, m$ *tel que* $\forall\, x \in E$, $m|x$, *alors* $m|PGCD(E)$.

Ce n'est pas la méthode la plus naturelle (compte-tenu du nom PGCD), mais la plus pratique pour les démonstrations...

Définition - Définition par récurrence

Soient $k \in \mathbb{N}^*$, $k \geqslant 2$, et $(a_1, a_2, \ldots, a_k) \in \mathbb{Z}^k$. Alors, on définit récursivement :

$$\bigwedge_{i=1}^{k} a_i := \left(\bigwedge_{i=1}^{k-1} a_i\right) \wedge a_k$$

L'identité de Bézout se généralise également :

Proposition - Décomposition de Bézout

Soient $k \in \mathbb{N}^*$, $k \geqslant 2$, et $(a_1, a_2, \ldots, a_k) \in \mathbb{Z}^k$.

$$\exists (u_1, \ldots, u_k) \in \mathbb{Z}^k \quad \Big| \quad \bigwedge_{i=1}^{k} a_i = \sum_{i=1}^{k} u_i a_i.$$

Démonstration

Il faut faire une récurrence sur k.
Notons, pour tout $k \in \mathbb{N}^*$, $k \geqslant 2$:
$\mathscr{P}_k$: « Pour $(a_1, a_2, \ldots, a_k) \in \mathbb{Z}^k$, $\exists\, (u_1, \ldots, u_k) \in \mathbb{Z}^k \mid \bigwedge_{i=1}^{k} a_i = \sum_{i=1}^{k} u_i a_i$. »

- $\mathscr{P}_2$ est vraie. C'est le théorème de Bézout vu plus haut.
- Soit $k \in \mathbb{N}$, $k \geqslant 2$. Supposons que $\mathscr{P}_k$ est vraie.
  Soient $(a_1, a_2, \ldots, a_k, a_{k+1}) \in \mathbb{Z}^k$.
  Notons $\delta_1 = \bigwedge_{i=1}^{k} a_i$ et $\delta = \bigwedge_{i=1}^{k+1} a_i$.
  Par définition de δ, $\delta = \delta_1 \wedge a_{k+1}$.
  D'après l'identité de Bézout, il existe $u, v \in \mathbb{Z}$ tel que $\delta = u\delta_1 + v a_{k+1}$.
  Puis on applique $\mathscr{P}_k$, à $(a_1, a_2, \ldots, a_k)$: $\delta_1 = u_1 a_1 + \cdots + u_k a_k$.
  Finalement $\delta = \sum_{i=1}^{k+1} u'_i a_i$, avec $u'_{k+1} = v$ et $u'_j = u_j \times u$ pour $j \leqslant k$.
  Donc $\mathscr{P}_{n+1}$ est vraie.

Notons qu'on aurait pu commencer à $k = 1$, mais le résultat obtenu n'a pas d'intérêt □

Proposition - Avec l'ensemble des diviseurs

Pour tout $j \in \mathbb{N}_k$, $\bigwedge_{i=1}^{k} a_i$ est un diviseur de a_j.
Mieux : $\bigwedge_{i=1}^{k} a_i = \max\left(\mathscr{D}(a_1) \cap \mathscr{D}(a_2) \cdots \mathscr{D}(a_k) \cap \mathbb{N}\right)$,
au choix pour max : au sens de la relation d'ordre | ou $\leqslant$.
On a $\mathscr{D}(\bigwedge_{i=1}^{k} a_i) = \bigcap_{j=1}^{k} \mathscr{D}(a_j)$.

Démonstration

On fait la démonstration par récurrence sur k.
Le résultat est vrai par définition pour $k = 2$.
Supposons que le résultat soit vrai pour un nombre $k \in \mathbb{N}$ et $k \geqslant 2$.
Soit $a_1, a_2, \ldots a_{k+1} \in \mathbb{N}$.
Alors par définition du PGCD à deux éléments : $\bigwedge_{i=1}^{k+1} a_i$ divise a_{k+1},
et $\bigwedge_{i=1}^{k+1} a_i$ divise $\bigwedge_{i=1}^{k} a_i$.
Mais par récurrence, $\bigwedge_{i=1}^{k} a_i$ divise a_j, pour tout $j \leqslant k$
et par transitivité de la divisibilité : $\bigwedge_{i=1}^{k+1} a_i$ divise donc a_j, pour tout $j \leqslant k$.
Ce qui permet de montrer l'hérédité.
Cela montre également que $\bigwedge_{i=1}^{k} a_i \in \bigcap_{j=1}^{k} \mathscr{D}(a_j)$, donc $\mathscr{D}(\bigwedge_{i=1}^{k} a_i) \subset \bigcap_{j=1}^{k} \mathscr{D}(a_j)$.

Il reste à montrer la maximalité du PGCD.
Soit $n \in \mathscr{D}(a_1) \cap \mathscr{D}(a_2) \cdots \mathscr{D}(a_k) \cap \mathbb{N}$, alors $n|a_1$, $n|a_2 \ldots n|a_k$ et n est positif.
Donc n divise toute combinaison linéaire de ces nombres,
et en particulier $\bigwedge_{i=1}^{k} a_i$, d'après la décomposition de Bézout généralisé.
Mais on vient de voir que $\bigwedge_{i=1}^{k} a_i \in \mathscr{D}(a_1) \cap \mathscr{D}(a_2) \cdots \mathscr{D}(a_k) \cap \mathbb{N}$, c'est donc bien le plus grand élément pour la divisibilité et pour la relation d'ordre classique.
Et on a également $\bigcap_{j=1}^{k} \mathscr{D}(a_j) \subset \mathscr{D}(\bigwedge_{i=1}^{k} a_i)$ □

5.2. Deux caractérisation du $PGCD(a_1, a_2, \ldots a_k)$

Réécriture de la propriété précédente :

Proposition - Critère caractéristique du $PGCD(a_1, a_2, \ldots a_k)$
Soient $k \in \mathbb{N}^*$, $k \geqslant 2$, et $(a_1, a_2, \ldots, a_k) \in \mathbb{Z}^k$.
$\left(\bigwedge_{i=1}^{k} a_i\right)$ est l'unique entier naturel d dont les diviseurs sont exactement les diviseurs communs à tous les a_i, c'est-à-dire tel que

$$\forall n \in \mathbb{Z}, \qquad (\forall i \in [\![1, k]\!],\ n|a_i) \iff n|d.$$

Démonstration
Notons $\delta = \bigwedge_{i=1}^{k} a_i$,
Pour tout $i \in \mathbb{N}_k$, $\delta \in \mathscr{D}(a_i)$, i.e. $\delta|a_i$.
Si $m|\delta$, alors, par transitivité pour tout $i \in \mathbb{N}_k$, $m|a_i$,
Réciproquement, si pour tout $i \in \mathbb{N}_k$, $m|a_i$. (On suppose aussi que $m \geqslant 1$)
alors m divise toute combinaison linéaire entière des a_i.
δ est une de ces combinaisons linéaires (Bézout), donc $m|\delta$. □

Cela donne bien une caractérisation

Savoir faire - Démonter/utiliser le PGCD de $(a_1, a_2, \ldots a_k)$
Tout diviseur du PGCD, divise chaque a_i.
Toute diviseur de tous les a_i est un diviseur de leur PGCD.
Et le PGCD est le plus grand de tous les diviseurs au sens de la relation d'ordre de la division (c'est une borne inférieure).

Proposition - Linéarité absolue
Le PGCD de $(xa_1, xa_2, \ldots xa_k)$ est $x| \times \bigwedge_{i=1}^{k} a_i$

Démonstration
Notons d'abord qu'il existe $u_1, \ldots u_k \in \mathbb{Z}$ tel que $\bigwedge_{i=1}^{k} a_i = \sum_{i=1}^{k} u_i a_i$.
1. Notons que $|x| \times \bigwedge_{i=1}^{k} a_i$ est un entier naturel.
2. Puis pour tout $j \in \mathbb{N}_k$, $\bigwedge_{i=1}^{k} a_i | a_j$, donc $|x| \times \bigwedge_{i=1}^{k} a_i | xa_j$.
3. Enfin, si pour tout $j \in \mathbb{N}_k$, $z|xa_j$, alors $z|\sum_{i=1}^{k} u_i xa_i = x\bigwedge_{i=1}^{k} a_i$, donc $z||a|\bigwedge_{i=1}^{k} a_i$.
(Attention, ce qui compte au point 3., ce sont bien les implications!!).
Donc $x| \times \bigwedge_{i=1}^{k} a_i$ est bien le plus grand des diviseurs communs des $(xa_i)_{1 \leqslant i \leqslant k}$ □

Proposition - Sous-groupe $a_1\mathbb{Z} + a_2\mathbb{Z} + \cdots + a_k\mathbb{Z}$
Soient $a_1, a_2, \ldots a_k \in \mathbb{N}$.
Alors $a_1\mathbb{Z} + a_2\mathbb{Z} + \cdots + a_k\mathbb{Z}$ est exactement le sous-groupe de $(\mathbb{Z}, +)$: $\left(\bigwedge_{i=1}^{k} a_i\right)\mathbb{Z}$.
Autrement écrit, le groupe engendré par $\{a_1, a_2, \ldots a_k\}$ est le groupe $\left(\bigwedge_{i=1}^{k} a_i\right)\mathbb{Z}$.

$$< a_1, a_2, \ldots a_k >_{\mathbb{Z}} = \left(\bigwedge_{i=1}^{k} a_i\right)\mathbb{Z}$$

Démonstration
L'addition de sous-groupes est un sous-groupe de $\mathbb{Z}$.
Comme $\delta := \bigwedge_{i=1}^{k} a_i$ est une combinaison linéaire entière de $a_1, a_2, \dots a_k$ (Bézout),
on a donc $\delta \in a_1\mathbb{Z} + a_2\mathbb{Z} + \dots + a_k\mathbb{Z}$,
Par stabilité de groupe : $\delta\mathbb{Z} \subset a_1\mathbb{Z} + a_2\mathbb{Z} + \dots + a_k\mathbb{Z}$.
Réciproquement, puisque $\delta | a_i$, alors pour tout $i \in \mathbb{N}_k$, $a_i \in \delta\mathbb{Z}$.
Par stabilité de groupe : $a_1\mathbb{Z} + a_2\mathbb{Z} + \dots + a_k\mathbb{Z} \subset \delta\mathbb{Z}$ □

5.3. Entiers premiers entre eux dans leur ensemble

Définition - Entiers premiers entre eux dans leur ensemble
Les entiers $a_1, \dots, a_k$ sont dits *premiers entre eux dans leur ensemble* si leur PGCD vaut 1.

Attention - Ne pas confondre « premiers entre eux dans leur ensemble » et « premiers deux à deux »
Il ne faut pas confondre « premiers entre eux dans leur ensemble » et « premiers deux à deux ».
Par exemple $a =$, $b =$, $c =$
sont
ne sont pas

Tous les savoir-faire donnés précédemment sur les PGCD se généralisent à plusieurs nombres. Nous ne les ré-écrirons pas mais il faudra savoir y penser. Voici un exemple

Savoir faire - Réduction à des entiers premiers entre eux (2)
Soient $a_1, a_2, \dots a_k \in \mathbb{Z}$. On note $\delta = \bigwedge_{i=1}^{k} a_i$.
Alors pour tout $i \in [\![1, k]\!]$, $a_i = \delta a_i'$, avec $(a_1', a_2', \dots, a_k')$ premiers entre eux dans leur ensemble.
Puis on travaille avec les a_i'

Proposition - Théorème (identité) de Bézout
Soient $a_1, \dots, a_k$ des entiers relatifs. Alors

$$\bigwedge_{i=1}^{k} a_i = 1 \iff \exists (u_1, \dots, u_k) \in \mathbb{Z}^k \mid \sum_{i=1}^{k} u_i a_i = 1$$

Démonstration
Si $\bigwedge_{i=1}^{k} a_i = 1$, alors d'après la décomposition de Bézout, $\exists (u_1, \dots, u_k) \in \mathbb{Z}^k \mid \sum_{i=1}^{k} u_i a_i = \bigwedge_{i=1}^{k} a_i = 1$.
Réciproquement, si $\exists (u_1, \dots, u_k) \in \mathbb{Z}^k \mid \sum_{i=1}^{k} u_i a_i = 1$,
alors si pour tout $i \in [\![1, k]\!]$, $n | a_i$, on a $n | \sum_{i=1}^{k} u_i a_i = 1$, donc $n|1$, i.e. $n = 1$.
Par définition du PGCD : $\bigwedge_{i=1}^{k} a_i = 1$ □

Selon que l'on cherche à montrer que des nombres sont premiers entre eux, ou utiliser cette connaissance, on exploite ce l'identité de Bézout de l'une ou l'autre de ces façons :

Savoir faire - Exploiter l'identité de Bézout
Pour montrer que $(a_1, \dots a_n)$ sont premiers entre eux, ils arrivent, qu'on montre :

$$\exists\, u_1, u_2, \dots u_n \in \mathbb{Z} \text{ tels que } \sum_{i=1}^{n} u_i a_i = 1$$

Quand, on sait que $(a_1, \dots a_n)$ sont premiers entre eux, on génère alors

$u_1, u_2, \dots u_n \in \mathbb{Z}$ tels que $\sum_{i=1}^{n} u_i a_i = 1$ et on exploite ces (u_i).

6. Plus Petit Commun Multiple

6.1. Construction

On considère une borne supérieure...

Analyse - Construction du $PPCM$

Soient a et b deux entiers relatifs non nuls.
$a\mathbb{Z} \cap b\mathbb{Z} \cap \mathbb{N}^*$ est une partie non vide de $\mathbb{N}^*$.
En effet, elle contient au moins$|ab|$, donc elle admet un plus petit élément non nul.
Ce plus petit élément est le $PPCM$ de a et b

Définition - $PPCM$ de a et de b
Soient a et b deux entiers relatifs non nuls.
On appelle $PPCM$ (Plus Petit Commun Multiple) de a et b, le nombre

$$PPCM(a,b) = \min\left(a\mathbb{Z} \cap b\mathbb{Z} \cap \mathbb{N}^*\right)$$

On le note également $a \vee b$.
On a clairement $b \vee a = a \vee b = |a| \vee |b|$.

Les deux caractéristiques du PGCD se fusionne en une seule, concernant le PPCM :

Proposition - Caractérisation essentielle du $PPCM$
Soit $(a,b) \in \mathbb{Z}^2$. Alors $a \vee b$ est le seul entier naturel dont les multiples sont exactement les multiples communs à a et b

$$a\mathbb{Z} \cap b\mathbb{Z} = (a \vee b)\mathbb{Z}$$

c'est-à-dire tel que

$$\forall m \in \mathbb{Z},\ (a|m \text{ et } b|m) \iff a \vee b|m$$

Démonstration
Si $a|m$ et $b|m$, alors m est un multiple commun de a et de b.
Il est a priori plus grand que $a \vee b$, le plus petit de tous.
La division euclidienne donne : $m = q(a \vee b) + r$ avec $r < a \vee b$.
Donc $r = m - q(a \vee b)$ et divisible par a et par b. C'est un commun multiple de a et de b.
Or $r < a \vee b$, la seule possibilité est $r = 0$ et donc $a \vee b$ divise m.
Réciproquement, si $a \vee b$ divise m.
Alors $m = a \vee b \times k$. Mais $a \vee b = aa' = bb'$
Donc $m = a(a'k) = b(b'k)$. Ainsi m est un multiple de a et de b. □

Savoir faire - Trouver un PPCM. Exploiter un PPCM.
On note $\mu = a \vee b$.
Pour trouver le PPCM ,
on démontre que si $a|m$ et $b|m$, alors nécessairement $\mu|m$.
Puis on cherche le m le plus petit possible.
Pour exploiter le PPCM,
on exploite le fait que si $\mu|m$, alors $a|m$ et $b|m$.
Et on travaille sur cette co-divisibilité.

Proposition - Linéarité
Soient $a, b, \lambda \in \mathbb{Z}$, alors $\lambda a \vee \lambda b = |\lambda|(a \vee b)$.

Démonstration
Si $\lambda a | m$, alors $\lambda | m$, donc $\frac{m}{\lambda} \in \mathbb{Z}$.
On a donc les équivalentes :

$$\lambda a | m, \lambda b | m \iff a | \frac{m}{\lambda}, b | \frac{m}{\lambda} \iff a \vee b | \frac{m}{\lambda} \iff \lambda(a \vee b) | m$$

On peut identifier, en faisant attention à la positivité : $\lambda a \vee \lambda b = \lambda(a \vee b)$. □

6.2. Relation PPCM et PGCD

Proposition - Relation $PPCM$ et $PGCD$
Soient $a, b \in \mathbb{Z}$.
- si $a \wedge b = 1$ alors $|ab| = (a \vee b)$.
- dans le cas général, $|ab| = (a \wedge b) \times (a \vee b)$.

Démonstration

- ab est un multiple commun à a et b.
  Soit n un autre multiple commun. Alors $n = qa$ et $b | qa$, donc $b | q$ et $ab | n$.
  Tous les autres multiples de a et b, sont des multiples de ab.
- Soit $d = a \wedge b$, $a = da'$, $b = db'$ avec $a' \wedge b' = 1$,
  d'où $a' \vee b' = |a'b'|$ et $a \vee b = da' \vee db' = d(a' \vee b') = |da'b'|$ d'où le résultat.

□

Remarque - Généralisation à un nombre fini d'entiers
On peut également généraliser la notion de $PPCM$ à un nombre fini d'entiers relatifs. Il existe un unique entier naturel m dont les multiples sont exactement les multiples communs à tous les a_i, c'est-à-dire tel que

$$\forall n \in \mathbb{Z}, (\forall i \in [\![1, k]\!],\ a_i | n) \iff m | n.$$

On l'appelle PPPCM de $a_1, a_2, \ldots, a_k$ et on le note $a_1 \vee a_2 \vee \ldots \vee a_k$ ou $\bigvee_{i=1}^{k} a_i$.

7. Bilan

Synthèse

- ⇝ Les nombres entiers relatifs sont les premiers objets reconnus comme mathématiques rencontrés. Ils sont donc comme à la base du sentiment mathématique de tout apprenti mathématicien.
- ⇝ Comme un jeu, leur manipulation est ludique. On travaille uniquement avec addition et soustraction, puis multiplication ou division euclidienne (soustractions répétées). Pour deux (voire n) nombres quelconques, le commun est le PGCD. C'est comme la plus grosse molécule constitutive de chacun de ces nombres ; avec ces deux nombres, il n'est pas possible de dégager des nombres plus fins que cette molécule. De là, de nombreux lemmes, théorèmes découlent dont les essentielles : théorème de Bézout et lemme de Gauss.
  On peut également réfléchir en terme de PPCM.
- ⇝ Reprenant une idée fondamentale simple et simplifiante de GAUSS, nous pouvons étudier les nombres entiers réduits modulo n. La plupart des propriétés arithmétiques se prolongent bien lors de cette réduction : addition et multiplication (pas très bien pour la division, car tous les nombres ne sont pas nécessairement inversibles...).

Savoir-faire et Truc & Astuce du chapitre

- Savoir-faire - Montrer que b divise a
- Truc & Astuce pour le calcul - Réduction modulo n
- Savoir-faire - Exploiter la division euclidienne (théorie)
- Truc & Astuce pour le calcul - Trouver une erreur dans un algorithme d'Euclide
- Savoir-faire - Pas d'unicité du couple de Bézout. Les obtenir tous
- Truc & Astuce pour le calcul - Obtenir un couple de Bézout
- Savoir-faire - Trouver un PGCD. Exploiter un PGCD.
- Savoir-faire - Combinaisons linéaires (entières)
- Savoir-faire - Réduction à des entiers premiers entre eux (1)
- Savoir-faire - Démontrer/utiliser le PGCD de $(a_1, a_2, \dots a_k)$
- Savoir-faire - Réduction à des entiers premiers entre eux (2)
- Savoir-faire - Exploiter l'identité de Bézout
- Savoir-faire - Trouver un PPCM. Exploiter un PPCM.

Notations

	Propriétés	*Remarques*
les diviseurs de a		
les multiples de a		
reste (resp.) de la division eucli- a par b	$a = (a//b) \times b + (a\%b)$ avec $a\%b \in [\![0, b-1]\!]$	Notations Python. Non officielle.
$\Rightarrow \exists\, k \in \mathbb{Z}$ tel que $a = b + nk$	Relation de congruence modulo n (relation d'équivalence)	$\forall\, a \in \mathbb{Z}$, $n \in \mathbb{N}$, $a \equiv (a\%n)\ [n]$
et b (généralisable)	$a \wedge b = \max(\mathscr{D}(a) \cap \mathscr{D}(b))$ $d\|a$ et $d\|b$ ssi $d\|a \wedge b$	$a\mathbb{Z} + b\mathbb{Z} = (a \wedge b)\mathbb{Z}$
et b (généralisable)	$a \vee b = \min(a\mathbb{Z} \cap b\mathbb{Z})$ $a\|m$ et $b\|m$ ssi $a \vee b\|m$	$a\mathbb{Z} \cap b\mathbb{Z} = (a \vee b)\mathbb{Z}$
les nombres premiers avec a	$b \in \mathscr{P}_a \iff a \wedge b = 1$	Equivalent à $a \in \mathscr{P}_b$ (symétrie)

Retour sur les problèmes

67. Oui, on peut tout faire. Sinon, on rajoute le nombre dans les premiers.
68. C'est les congruences.
69. Par exemple, les nombres premiers sont les nombres au premier étage du treillis des diviseurs.
    Le PGCD de deux nombres a et b est celui situé le plus haut dans les racines communes de a et de b.
    Le PPCM de deux nombres a et b est celui situé le plus bas dans les décédants commus de a et de b...
70. Oui cela termine (suite d'entiers naturels, strictement décroissantes). On obtient en dernier reste non nul, le PGCD.
71. Avec 3 et 5 euros, on peut obtenir tous les nombres entiers d'euros.
    Avec 6 et 9 euros, on peut obtenir tous les multiples de 3 euros.
    Avec 6, 10 et 15 euros, on peut obtenir tous les nombres entiers d'euros $(6 + 10 - 15)$.

Chapitre 16

Nombres premiers

Résumé -

Nous continuons notre plongée dans la reine des mathématiques : l'arithmétique. Euclide s'est particulièrement concentré sur l'ensemble des nombres premiers. C'est le but de ce chapitre.
Nous verrons que cet ensemble reste toujours le graal des mathématiciens.
Ces dernières années, les nombres premiers sont revenus au centre des mathématiques comme le coeur des applications numériques de codes secrets (internet...).
Cela nous permettra d'étudier l'énoncé et des applications du petit théorème de Fermat. Nous prendrons le temps de comprendre les idées du plus grand mathématicien toulousain !
Youtube (parodies ?) :

- *Panpan1963 - Petit théorème de Fermat - https ://www.youtube.com/watch ?v=Ei4PMxddm9Q*
- *ScienceEtonnante - Les nombres premiers - https ://www.youtube.com/watch ?v=R37JHiA-HOg*

Sommaire

1. Problèmes

? Problème 75 - Produit de premiers

Est-ce qu'on peut vraiment tout faire (multiplicativement) qu'avec des nombres premiers ?

? Problème 76 - Construction du cours

La notion de divisibilité est toujours première dans un cours d'arithmétique d'entiers.
Puis, ici nous avons choisi (selon le programme officielle) un cours sous la forme : PGCD (et PPCM) $\Longrightarrow$ Nombres premiers $\Longrightarrow$ congruence.
Dans son cours, GAUSS avait choisi : congruence $\Longrightarrow$ PGCD (et PPCM) $\Longrightarrow$ Nombres premiers. Quant à EUCLIDE, il commence par les nombres premiers.
Quel est le choix qui vous semble plus naturel ? Comment les démonstrations en sont-elles changées ?

? Problème 77 - Fonctions arithmétiques

Si une fonction $f : \mathbb{N} \to \mathbb{Z}$ est un morphisme : $f(ab) = f(a) \times f(b)$ ou $f(ab) = f(a) + f(b)$, que peut-on dire de f, en particulier concernant son image sur $\mathscr{P}$, l'ensemble des nombres premiers qui génèrent $\mathbb{N}$ par multiplication ?

2. Théorèmes d'Euclide

2.1. Définition

Il ne s'agit plus d'une notion relative comme précédemment (a et b sont premiers entre eux). Ici, il s'agit d'une notion absolue.

Définition - Nombre premier
Soit $p \in \mathbb{N}^*$.
On dit que p est (un nombre) premier
si $p \neq 1$ et si les seuls diviseurs de p dans $\mathbb{N}$ sont 1 et p.
On note souvent $\mathscr{P}$, l'ensemble des nombres premiers.

Le premier théorème d'Euclide est la version « nombre premier » du lemme de Gauss.

Pour aller plus loin - Nombres premiers
Dans un anneau euclidien (muni d'une division euclidienne), un nombre premier est un nombre qui n'est divisible que par des inversibles de l'anneau : si $p = a \times b$, alors a ou $b \in A^*$.
Ici $\mathbb{Z}^* = \{-1, 1\}$. Et pour les polynômes, un polynôme premier n'est divisible que par des constantes : $(\mathbb{K}[X])^* = \mathbb{K}$, c'est un polynôme irréductible.
Un autre anneau euclidien bien connu est $\mathbb{Z}[i] = \{a + ib \mid a, b \in \mathbb{Z}\}$, l'anneau des entiers de GAUSS.

2.2. Lemmes d'Euclide

On reconnait le lemme de Gauss et ses corollaires, mais associés à des nombres premiers.

Proposition - Premier théorème d'Euclide
Un nombre premier est premier avec tous les entiers qu'il ne divise pas.
Un nombre premier divise un produit si et seulement si il divise l'un des facteurs.

Démonstration
Soit p un nombre premier et $a \in \mathbb{Z}$.
Notons $\delta = p \wedge a$. Alors $\delta | p$, donc $\delta = 1$ ou p.
Si $\delta = p$, alors $p|a$.
Et si $\delta = 1$, alors a et p sont premiers entre eux.
Supposons $p|a \times b$.
Ou bien $p|a$,
ou bien p ne divise pas a, alors $p \wedge a = 1$ et d'après le lemme de Gauss $p|b$. □

L'intérêt des nombres premiers est d'être des briques élémentaires multiplicatives de $\mathbb{Z}$, on ne peut la couper en deux.
La remarque suivante est un savoir-faire, puisqu'elle donne un truc de manipulation de nombres premiers. Mais c'est aussi un ATTENTION, car il y est associée une erreur fréquente, dans le cas de nombres non premiers.

Savoir faire - Trouver un facteur premier avec un nombre premier
Soit $p \in \mathscr{P}$ tel que $p|ab$ alors p divise a ou p divise b.
Ce n'est pas le cas de $p = 12$ qui divise 6×4 avec $a = 6$ et $b = 4$, par exemple

Théorème -
Tout entier naturel $n \geqslant 2$ possède au moins un diviseur premier.

Démonstration
Soit $n \geqslant 2$.
L'ensemble $\mathscr{D}(n) \cap \mathbb{N} = \{d \in \mathbb{N}; d|n, d \geqslant 2\}$ est une partie non vide de $\mathbb{N}$,
donc admet un plus petit élément $p \geqslant 2$.
Or si $k \geqslant 2$ est tel que $k|p$ alors $k|n$ d'où $k \geqslant p$ et finalement $k = p$ donc p est premier. □

Corollaire - Critère de primalité entre deux entiers
Soient $a, b \in \mathbb{Z}$.
Alors a et b sont premiers entre eux si et seulement s'ils n'ont aucun diviseur premier en commun.

Pour aller plus loin - Idéaux premiers
On retrouve aussi la même notion dans la définition des idéaux premiers d'un anneau intègre (vu en seconde année).
De même un groupe simple est un groupe qui ne peut se décomposer en produit de sous-groupes (distingués). La décomposition d'un groupe en produit de sous-groupe simple suit à la même philosophie…

Démonstration
On raisonne en contraposée. Le théorème est équivalent à :
$a \wedge b \neq 1$ si et seulement si $\exists\, p$, premier tel que $p|a$ et $p|b$.
Cette équivalence est vraie en prenant p, un facteur premier de $a \wedge b$.
(c'est possible d'après le théorème) □

2.3. Théorème fondamental

Puis le théorème fondamental de l'arithmétique :

Théorème - Décomposition en produit de facteurs premiers
Soit $n \in \mathbb{N}$, $n \geqslant 2$.
Alors il existe $r \in \mathbb{N}$, $p_1, p_2, \dots p_r$, r nombres premiers et $\alpha_1, \dots \alpha_r \in \mathbb{N}^*$ tel que

$$n = \prod_{i=1}^{r} p_i^{\alpha_i}$$

Cette écriture est unique.

Démonstration
On démontre d'abord l'existence (récurrence) de la décomposition puis son unicité.
- Posons, pour tout $n \in \mathbb{N}$, $n \geqslant 2$, $\mathscr{H}_n$:« tout entier compris entre 2 et n admet une telle décomposition. »
  - $\mathscr{H}_2$ est vraie car 2 est premier.
  - supposons $\mathscr{H}_n$ vraie pour un certain $n \geqslant 2$.
    au rang $n+1$: si $n+1$ est premier, c'est fini,
    sinon $n+1 = pq$ avec p premier et $2 \leqslant q \leqslant n$ et on applique $\mathscr{H}_n$ à q.
- Si
$$n = \prod_{i=1}^{r} p_i^{\alpha_i} = \prod_{j=1}^{s} p_j^{\beta_j}$$
alors pour tout i $p_i | n$ d'où il existe j tel que $p_i | q_j$ (p_i premier) soit $p_i = q_j$ (q_j premier),
donc en fait on a les mêmes polynômes irréductibles dans les deux décompositions.
Reste à prouver que les puissances sont les mêmes.
Supposons pour un i que l'on ait alors $\alpha_i > \beta_i$, en simplifiant par $p_i^{\beta_i}$ on a
$$p_i^{\alpha_i - \beta_i} \prod_{k \neq i} p_k^{\alpha_k} = \prod_{k \neq i} p_k^{\beta_k}$$
d'où $p_i | \prod_{k \neq i} p_k^{\beta_k}$ ce qui est absurde car p_i premier et $p_i \wedge p_k = 1$ pour $k \neq i$. D'où $\alpha_i = \beta_i$.

□

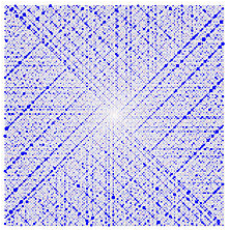
3. L'ensemble des nombres premiers

3.1. Ensemble infini

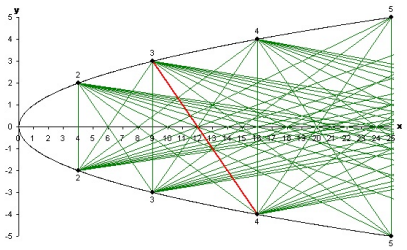
Proposition - Théorème d'Euclide
L'ensemble des nombres premiers, noté $\mathscr{P}$ (parfois $\mathbb{P}$) est infini.

Démonstration
Par l'absurde :
Supposons qu'il n'y ait qu'un nombre fini de nombres premiers $p_1 < p_2 < \cdots < p_n$ (ordonné).
Posons $N = p_1 p_2 \dots p_n + 1$.
On a $N \geqslant 2$ donc N possède un diviseur premier p, et il existe j tel que $p = p_j$
alors $p | p_1 p_2 \dots p_n$ et $p | N$ donc $p | = N - \prod_{i=1}^{n} p_i = 1$ et $p = 1$, ce qui est impossible.
Donc l'ensemble des nombres premiers est infini. □

Représentation - Deux visions de $\mathscr{P}$
Sans commentaires :
Spirale d'Ulam :



Crible de Matiyasevich



Cet ensemble mystérieux représente une sorte de graal du mathématicien.

Remarque - Ensemble dénombrable
Comme tout ensemble inclus dans $\mathbb{N}$ et infini, $\mathscr{P}$ est dénombrable, c'est-à-dire il existe une bijection de $\mathbb{N}$ dans $\mathscr{P}$. Et cela signifie qu'on peut donc écrire $\mathscr{P} = \{p_1, p_2, \dots p_r, \dots\}$, où $p_1 = 2$ est le premier des nombres premiers, $p_2 = 3$, $p_3 = 5 \dots$, p_r est le r-ième nombre premier.

Proposition - Enumération des nombres premiers
Il existe une bijection $\rho : \mathbb{N}^* \to \mathscr{P}$, $i \mapsto p_i$, le i-ième nombre premier

Exercice
Que vaut p_{10} ?

Correction
On n'a pas d'autre option que de faire la liste des 20 premiers nombres premiers.
$p_1 = 2$, $p_2 = 3$, $p_3 = 5$, $p_4 = 7$, $p_5 = 11$, $p_6 = 13$, $p_7 = 17$, $p_8 = 19$, $p_9 = 23$ et $p_{10} = 29$

Démonstration
On construit par récurrence une suite d'ensembles $\mathbb{P}_i$ et une suite d'éléments (p_i) où $p_i \in \mathbb{P}_i$:
- Au rang 1 : $\mathbb{P}_1 = \mathscr{P}$ et $p_1 = \rho(1) = \min \mathbb{P}_1$, on a bien $p_1 \in \mathbb{P}_1$
- On suppose qu'au rang n, $\mathbb{P}_n$ et p_n sont bien définies (avec $p_n \in \mathbb{P}_n$).
  On définit alors $\mathbb{P}_{n+1} = \mathbb{P}_n \setminus \{p_n\}$ et $p_{n+1} = \rho(n+1) = \min \mathbb{P}_{n+1}$.
  On a bien $p_{n+1} \in \mathbb{P}_{n+1}$

La construction ne s'épuise pas car $\mathscr{P}$ est infini, et qu'à chaque étape, $\mathscr{P}_n$ reste infini. □

3.2. Crible d'Eratosthène

Informatique - Crible d'Eratosthène

Pour obtenir la liste des nombres premiers, nous n'avons pas trouvé beaucoup mieux que le crible d'Eratosthène (3-ième siècle avant J-C).
Il s'agit d'écrire la liste des entiers de 1 à n. Puis d'enlever les multiples des nombres qui restent. Une fois terminée (deux boucles), il ne reste que la liste des nombres premiers plus petit que n.

```
def Eratosthene(n):
    """Crible d'Erathostene adapte"""
    L=[1]*n
    for k in range(2,n):
        if L[k]==1 :
            h=2*k
            while h<n :
                L[h]=0
                h=h+k
    P=[]
    for k in range(1,n):
        if L[k]==1 :
            P=P+[k]
    return(P)
```

Remarque - Conjectures

L'ensemble $\mathscr{P}$ est infini, mais une question résiste : contient-il une infinité de nombres premiers jumeaux?
On dit que $(p, p+2)$ est un couple de nombres premiers jumeaux si ils sont tous les deux premiers. Par exemple $(3,5)$ ou $(11,13)$ sont des nombres premiers jumeaux.
Ben Green(1977-) et Terence Tao (1975-) ont démontré ce qui est désormais connu sous le nom de théorème de Green-Tao (pré-publié en 2004 et publié en 2008). Ce théorème établit qu'il existe des progressions arithmétiques de nombres premiers arbitrairement longues.
Un résultat culturel (mais pas nécessairement à retenir)

Théorème - Hadamard - De la Vallée Poussin (1896)

Notons $\pi(n)$, le nombre de nombres premiers plus petit que n. Alors

$$\pi(n) \underset{n\to+\infty}{\sim} \frac{n}{\ln n}$$

Le livre Merveilleux nombres premiers de Jean-Paul Delahaye donne une foule d'informations et d'anecdotes concernant les nombres premiers (par exemple : l'histoire des jumeaux John et Michael qui voient les nombres premiers...).

Exercice

En exploitant le théorème de Hadamard-De la vallée Poussin, donner une valeur approchée du 1000 nombre premiers (i.e. de p_{1000})

Correction

On a $\pi(p_{1000})$ qui est le nombre de nombres premiers plus petit que p_{1000}, par construction, il y en a exactement 1000.
Donc si $n = p_{1000}$, on a $\dfrac{n}{\ln n} \equiv 1000$.
Il faut inverser la fonction $n \mapsto \frac{n}{\ln n}$. Ce qui est difficile.
Nécessairement, $n \geqslant 1000$, on peut penser que si $n = 1000a$, on a $\ln n = \ln 1000 + \ln a$ et donc

$$\frac{1000a}{\ln 1000 + \ln a} = 1000 \Rightarrow \frac{a}{\ln 1000 + \ln a} \approx 1 \Rightarrow a \approx \ln 1000 = 6,90775$$

Donc $n \approx 6900$

Histoire - Conjecture de Riemann

Parmi les problèmes qui résistent aux mathématiciens, ce trouve la fameuse conjecture de Riemann. Elle concerne directement les nombres premiers même si elle s'énonce avec des nombres complexes. L'esthétisme de cette conjecture réside en partie dans le lien miraculeux entre l'arithmétique à l'ensemble des fonctions de la variable complexe. L'énoncé est :
Les zéros non triviaux de la fonction

$$\zeta : \begin{array}{ccc} \mathbb{C} & \to & \mathbb{C} \\ s & \to & \displaystyle\sum_{n=1}^{+\infty} \frac{1}{n^s} \end{array}$$

ont une partie réelle exactement égale à $\frac{1}{2}$.
A lire : la symphonie des nombres premiers de Marcus du Sautoy

4. Valuation (p-adique)

4.1. Fonction valuation (en base p)

Définition - Valuation p-adique
Soit p un nombre premier. Pour $n \in \mathbb{N}$, on appelle *valuation p-adique* de n le plus grand entier $k \in \mathbb{N}$ tel que p^k divise n. On le note $v_p(n)$.
Donc pour p premier, $p|n \Leftrightarrow v_p(n) \geqslant 1$.

On a les caractérisations suivantes, qu'il faut savoir exploiter comme savoir-faire :

Savoir faire - Caractérisations de $v_p(a)$.

$$\begin{array}{ll} p^h|a & \iff v_p(a) \geqslant h. \\ p^h q = a \text{ et } p \wedge q = 1 & \iff v_p(a) = h. \end{array}$$

Démonstration
• $v_p(a) = \max\{h \mid p^h|a\}$ et ainsi $\{h \mid p^h|a\} = [\![0, v_p(a)]\!]$.
Donc $p^h|a \iff h \in [\![0, v_p(a)]\!] \iff h \leqslant v_p(a)$.
• Si $p^h \times q = a$, alors $p^h|a$, et donc $h \leqslant v_p(a)$.
Notons $r = v_p(a) - h \in \mathbb{N}$, on a donc $p^h p^r = p^{v_p(a)}|a = p^h q$ donc $p^r|q$.
Or $p \wedge q = 1$ donc $p^r \wedge q = 1$ et donc $p^r = 1$, donc $r = 0$ et $h = v_p(a)$.
Réciproquement, si $v_p(a) = h$, alors $p^h|a$. On note $q \in \mathbb{N}$ tel que $p^h q = a$.
Soit $\delta = p \wedge q$, alors $\delta|p$, et p premier, donc $\delta = 1$ ou $\delta = p$.
Si $\delta = p$, alors $p|q$ et donc $p^{h+1}|a$, impossible. Donc $\delta = 1$. □

Exercice
Montrer que $v_p(pb) = 1 + v_p(b)$

Correction
$p^{v_p(pb)}|pb \Rightarrow p^{v_p(b)-1}|b \Rightarrow v_p(b) \geqslant v_p(pb) - 1$
$p^{v_p(b)}|b \Rightarrow p^{v_p(b)+1}|pb \Rightarrow v_p(pb) \geqslant v_p(b) + 1$.
Par double inégalité : $v_p(b) = v_p(pb) - 1$.

Remarque - Réécriture du théorème fondamentale de l'arithmétique
Pour tout $n \in \mathbb{N}$,

$$n = \prod_{i=1}^{+\infty} p_i^{v_{p_i}}(n) = \prod_{p \in \mathscr{P}} p^{v_p}(n)$$

le produit en réalité, nécessairement (à support) fini puisqu'à partir d'un certain rang, $v_{p_k}(n) = 0$.

4.2. Morphisme (de monoïde)

Proposition - Valuation, fonction logarithmique
Pour tout $a, b \in \mathbb{Z}$ et $p \in \mathscr{P}$, $v_p(a \times b) = v_p(a) + v_p(b)$

Démonstration
Si on décompose a et b : $a = p^{v_p(a)} q$ et $b = p^{v_p(b)} r$ avec $q \wedge p = 1$ et $r \wedge p = 1$.
Donc $ab = p^{v_p(a)+v_p(b)} qr$.
Et comme d'après le lemme de Gauss : $p \wedge qr = 1$, on a donc $v_p(ab) = v_p(a) + v_p(b)$. □

4.3. Factorisation en produit de premiers

Proposition - Liste des diviseurs
Soient $a, b \in \mathbb{N}$ non nuls. Si

$$a = p_1^{\alpha_1} p_2^{\alpha_2} \dots p_k^{\alpha_k} \text{ et } b = p_1^{\beta_1} p_2^{\beta_2} \dots p_k^{\beta_k}$$

où les p_i sont des nombres premiers distincts deux à deux, $\alpha_i, \beta_i \in \mathbb{N}$ (éventuellement nuls), alors

$$a|b \iff \forall i \in [\![1, k]\!],\ \alpha_i \leqslant \beta_i$$

$$a \wedge b = PGCD(a,b) = \prod_{i=1}^{k} p_i^{\min(\alpha_i,\beta_i)}$$

$$a \vee b = PPCM(a,b) = \prod_{i=1}^{k} p_i^{\max(\alpha_i,\beta_i)}$$

Ce qui peut aussi s'écrire :

Corollaire - Liste des diviseurs, avec la valuation

$$a|b \Longleftrightarrow \forall p \text{ premier }, v_p(a) \leqslant v_p(b)$$

$\forall p$ premier , $v_p(a \wedge b) = \min(v_p(a), v_p(b))$ et $v_p(a \vee b) = \max(v_p(a), v_p(b))$

$$a \wedge b = \prod_{p \in \mathscr{P}} p^{\min(v_p(a),v_p(b)} \qquad a \vee b = \prod_{p \in \mathscr{P}} p^{\max(v_p(a),v_p(b)}$$

Cette dernière formule se généralisant sans difficulté au cas de k entiers distincts.

Application - Algorithme de recherche du PGCD
On retrouve un algorithme bien connu pour réduire les fractions (simplifier par le PGCD) :
il suffit de chercher la liste des facteurs premiers du numérateur et dénominateur…

Démonstration
Si a divise b, alors $p_i^{\alpha_i}$ divise b et donc $\alpha_i \leqslant \beta_i$.

La réciproque est trivial, avec $c = p_1^{\beta_1-\alpha_1} p_2^{\beta_2-\alpha_2} \dots p_k^{\beta_k-\alpha_k}$, on a $a \times c = b$ Supposons que $d|a$, et $d|b$,

alors $d = p_1^{\delta_1} p_2^{\delta_2} \dots p_k^{\delta_k}$, avec $\delta_k \leqslant \alpha_k$.
de même $\delta_k \leqslant \beta_k$. Donc $\delta_k \leqslant \min(\alpha_k, \beta_k)$.
C'est le cas pour $d = a \wedge b$.

Par ailleurs : $\prod_{i=1}^{k} p_i^{\min(\alpha_i,\beta_i)}$ divise a et b, donc divise $a \wedge b$.

On peut donc assimiler : $a \wedge b = \prod_{i=1}^{k} p_i^{\min(\alpha_i,\beta_i)}$ La démonstration pour le $PPCM$ est laissé en exercice. □

Exercice
Soit $n \in \mathbb{N}^*$. Donner une expression de $\text{card}\mathscr{D}(n)$, utilisant les valuations $v_p(n)$

Correction
D'après la proposition : $m|n \Longleftrightarrow \forall\ p$, premier $v_p(m) \leqslant v_p(n)$.
Il y a donc pour chacun de ces facteurs p tel que $p|n$, $v_p(n)+1$ choix possibles (de 0 à $v_p(n)$).
Puis d'après le principe de décomposition, il faut tenir compte du produit de ces nombres. Ainsi
$\text{card}\mathscr{D}(n) = \prod_{p|n} (v_p(n)+1)$

4.4. Formule de Legendre

On termine par un exercice
Exercice
Soit p un nombre premier.
Montrer que $v_p(n!) = \sum_{k=1}^{+\infty} \lfloor \frac{n}{p^k} \rfloor$ (la somme étant en réalité finie).
On pourra s'intéresser aux ensembles $N_a = \{k \in \mathbb{N}_n \mid p^a | k\}$.

Correction
Par morphisme : $v_p(n!) = \sum_{k=1}^{n} v_p(k)$. Considérons $N_a = \{k \in \mathbb{N}_n \mid p^a | k\}$. Cet ensemble se décrit facilement :

$$N_a = \{p^a, 2p^a, 3p^a, \dots r_a p^a\}$$

où $r_a p^a \leqslant n$ et $(r_a+1)p^a > n$, donc $r_a = \lfloor \frac{n}{p^a} \rfloor = \text{card} N_a$.
Notons maintenant $M_a = \{k \in \mathbb{N}_n \mid v_p(k) = a\}$.
On a $k \in M_a \Longleftrightarrow k \in N_a$ et $k \notin N_{a+1}$.
Donc comme $N_{a+1} \subset N_a$, on a exactement : $N_a = M_a \uplus N_{a+1}$.
Ainsi, en terme de cardinal : $\text{card}(M_a) = \text{card}(N_a) - \text{card}(N_{a+1})$.
Enfin, tous les nombres plus petit que n ont une et une seule valuation p-adique,

$$\mathbb{N}_n = \biguplus_{a \in \mathbb{N}} M_a \qquad \text{(réunion finie)}$$

On note s, la plus grande puissance de p qui divise un nombre plus petit que n : donc $p^s \leqslant n < p^{s+1}$, i.e. $s = \lfloor \frac{\ln n}{\ln p} \rfloor$ Ainsi (par sommation par paquets) :

$$\begin{aligned} v_p(n!) &= \sum_{k\in\mathbb{N}_n} v_p(k) = \sum_{a=0}^{s}\left(\sum_{k\in M_a} v_p(k)\right) = \sum_{a=0}^{s}\left(a\sum_{k\in M_a} 1\right) = \sum_{a=0}^{s}(a\text{card}(M_a)) \\ &= \sum_{a=1}^{s} a(\text{card}(N_a) - \text{card}(N_{a+1})) = \sum_{a=1}^{s} a\text{card}(N_a) - \sum_{a=0}^{s} a\text{card}(N_{a+1}) \\ &= \sum_{a=1}^{s} a\text{card}(N_a) - \sum_{a=1}^{s+1}(a-1)\text{card}(N_a) = \sum_{a=1}^{s}(a-(a-1))\text{card}(N_a) + s\text{card}(N_{s+1}) \\ &= \sum_{a=1}^{s}\text{card}(N_a) + 0 = \sum_{a=1}^{s}\left\lfloor \frac{n}{p^a}\right\rfloor \end{aligned}$$

5. Garantir que des nombres sont premiers

5.1. Motivations

Heuristique - Décomposable vs. décomposition

L'analyse de la primalité éventuelle d'un nombre n entier peut déboucher sur quatre questions de complexités différentes :

1. Prouver que n n'est pas premier
2. Si n n'est pas premier, le décomposer
3. Garantir, avec un risque d'erreur faible que n est premier
4. Certifier que n est premier

Cela semble comparable, mais le petit théorème de Fermat permet de répondre « aisément » aux questions 1 et 3. Les deux autres questions qui semblent équivalentes sont bien plus compliquées…

Remarque - Obtenir des nombres premiers

L'arithmétique et avec la géométrie la plus vieille branche des mathématiques. Pendant deux millénaires, on s'y intéresse « pour le plaisir ». Mais avec le renouveau de l'algorithmique et la développement exponentielle des ordinateurs, se sont développés deux branches :

— la cryptanalyse
— les codes correcteurs.

Dans ces disciplines, on a très largement besoin de grands nombres premiers. C'est le cas du codage RSA, grand consommateur de nombres premiers qui permet de coder et décoder les messages numérisés (internet, banque…). Il faut trouver des nombres premiers, savoir les reconnaitre. Peut-on faire mieux qu'avec le simple crible d'Eratosthène ?

Analyse - Et pour Fermat ?

Pierre de Fermat était fasciné par la proposition de Diophante :
« si $s(n) = 1 + 2 + \cdots + 2^{n-1}$ est un nombre premier, alors $2^{n-1}s(n)$ est un nombre parfait »
Or $s(n) = 2^n - 1$. Il fallait savoir si ce nombre est premier. Les nombres premiers de cette forme sont appelés les nombres de Mersenne.
En 1640, Frénicle de Bessy défie Fermat et lui demande de trouver un nombre parfait de 20 chiffres environ.
Pour Fermat, il s'agit de trouver un nombre parfait compris entre 10^{20} et 10^{22}, et donc un nombre premier $s(n) = 2^n - 1$ avec $10^{20} < 2^{n-1}(2^n - 1) < 10^{22}$.
Donc $10^{20} < 2^{2n-1} - 2^{n-1} < 10^{22}$.
Or on sait qu'approximativement $2^{10} = 10^3$.

Cela donne donc $10^{20} = 10^{3^6} \times 100 \approx 2^{60} \times 100 > 2^{66}$ et $10^{22} = 10^{3^7} \times 10 < 2^{74}$.
Comme le premier ordre est donné par 2^{2n-1}, on trouve $66 < 2n - 1 < 74$, $33,5 < n < 37,5$.
Et donc : $34 \leqslant n \leqslant 37$. Il y a 4 nombres $2^n - 1$ avec $n \in \{34, 35, 36, 37\}$ et il faut savoir s'ils sont premiers.
Remarquons d'abord que si $n = ab$ n'est pas premier, par télescopage :

$$2^{ab} - 1 = 2^{ab} - 1^b = (2^a - 1)\sum_{k=0}^{b-1} 2^{ak}$$

Ainsi, le seul possible serait $2^{37} - 1$.

Pour aller plus loin - Plus grand nombre premier connu (janvier 2016)

C'est un nombre de Mersenne : $2^{74\,207\,281} - 1$, il contient plus de 22 millions de chiffres.

Fermat affirme qu'il n'est pas premier! Comment s'y est-il pris?

5.2. Enoncé et applications

Théorème - Petit théorème de Fermat (1640)
Pour p premier et $n \in \mathbb{Z}$, on a $n^p \equiv n[p]$.
Si $n \wedge p = 1$ (i.e. p ne divise pas n) , alors $n^{p-1} \equiv 1[p]$

Histoire - Nombres parfaits
Depuis l'antiquité, de nombreux mathématiciens se sont intéressés aux nombres parfaits. Ce sont les nombres égaux à la somme de leurs diviseurs stricts (sans eux-mêmes).
Ainsi $6 = 1+2+3$ *est parfait. C'est aussi le cas de* 28 *ou* 496*...*

Savoir faire - Exploiter le petit théorème de Fermat
Il y a deux façons d'exploiter ce théorème pour un nombre N :
- Trouver une factorisation (plus exactement un facteur premier) du nombre N de la forme $a^n - 1$ (voir l'application qui suit)
- Montrer que le nombre N est probablement premier; dans ce cas N joue le rôle de p (voir la remarque : nombre de Carmichaël)

Application - $2^{37}-1$ est-il premier?
Ici, on applique le théorème de Fermat avec $n = 2$ et p, un diviseur premier (s'il existe) de $N = 2^{37} - 1$.
On a donc
- $p|N = 2^{37} - 1$,
- mais aussi $p|2^{p-1} - 1$ ($p \neq 2$, donc $2 \wedge p = 1$).

Alors Fermat affirme que $37|p-1$.
En effet, on considère le PGCD de 37 et $p-1$: $r = 37 \wedge (p-1)$.
Comme 37 est un nombre premier, alors $r = 37$ ou $r = 1$.
On applique l'identité de Bézout (modulo p) :

$$2^r \equiv 2^{37u+v(p-1)} \equiv (2^{37})^u \times (2^{p-1})^v \equiv 1[p]$$

L'hypothèse $r = 1$ est impossible : $2^1 = 2 \not\equiv 1[p]$.
finalement $r = 37$ et nécessairement $37|p-1$.
Puis comme $p-1$ est pair, divisible par 2, donc $p - 1 = 2 \times 37 \times q = 74q$.
Et finalement $p = 74q + 1$. Il faut cherche $q \in \mathbb{N}$.
- 75 n'est pas premier.
- Il essaye donc avec $149 = 74 \times 2 + 1$ (nombre premier), cela ne marche.
- Il essaye ensuite le calcul (division) avec $223 = 74 \times 3 + 1$ (nombre premier) et trouve

$$2^{37} - 1 = 223 \times 616\,318\,177$$

Histoire - De la main de Fermat
« Tout nombre premier mesure infailliblement une des puissances −1 de quelques progression que ce soit; et l'exposant de la dite puissance est sous-multiplie du nombre premier donné −1; et, après qu'on a trouvé la première puissance qui satisfait à la question, toutes celles dont les exposants sont multiples de l'exposant de la première satisfont tout de même à la question »

Remarque - Réciproque? Nombre de Carmichaël
Malheureusement, la réciproque du théorème de Fermat est fausse.
Il existe des nombres non premiers P tels que $P|2^P - 2$:

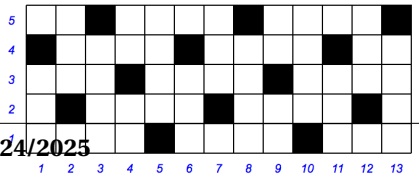
$$341|2^{341} - 2$$

C'est le plus petit contre-exemple. On se restreint ici au cas $n = 2$.
Mais il y a pire, des nombres p, non premiers qui vérifient : $\forall\, a \in \mathbb{N}$, $p|a^p - a$.
Ces nombres sont appelés *les nombres de Carmichaël.*
Le plus petit connu est $p = 561 = 3 \times 11 \times 17$.
On sait depuis peu (1994) qu'il en existe une infinité.

Exercice
Montrer que $341|2^{341} - 2$ sans que 341 soit premier.

Correction
On note que $341 = 31 \times 11$.
Puis $2^{341} = 2^{10\times 34+1} = 2^{10^{34}} \times 2 \equiv 1^{34} \times 2 = 2$ [11] d'après le petit théorème de Fermat.
Puis $2^{341} = 2^{30\times 11+11} = 2^{33^{11}} \times 2^5 \times 2^5 \times 2 \equiv 1^{11} \times 1 \times 1 \times 2 = 2$ [31] d'après le petit théorème de Fermat et $2^5 = 32 \equiv 1[31]$.

Représentation - Illustration de la démonstration
Dans le cas de $p = 13$ *et* $n = 31$.
On a alors $n \equiv 5[13]$ *et donc* $kn \equiv r_k$ *se voit sur le carrelage* ($a = p = 13, b = 5$).
On remarque alors que tous les restes sont obtenus, et une et une seule fois



5.3. Démonstrations

Nous ferons plusieurs démonstrations. Chacune apporte un résultat mathématique différent.

Démonstration
Si n est divisible par p, alors $n^p \equiv 0[p]$ et $n \equiv 0[p]$, donc $n^p \equiv n[p]$.
Supposons n non divisible par p.
Notons $N = n \times 2n \times 3n \times \cdots \times (p-1)n$.
Considérons r_k, le reste de la division euclidienne de kn par p. $kn \equiv r_k[p]$.
— Le calcul direct donne $N = (p-1)! \times n^{p-1}$.
— Mais aussi d'après l'arithmétique modulaire : $N \equiv r_1 \times r_2 \times \cdots \times r_{k-1}[p]$.
— $r_i \neq 0$, sinon : $p | i \times n$, impossible : p premier, n non divisible par p et $i < p$.
— Or si $r_i = r_j$, on a $(i-j)n \equiv 0[p]$, donc $p|(i-j)$, car p est premier et ne divise pas n.
Or $i, j \in [\![1, p-1]\!]$, donc $i - j \in [\![-p+1, p-1]\!]$ et ainsi $i - j = 0$, i.e. $i = j$.
— Comme tous les i sont distincts, il en est de même des r_i dans $[\![1, p-1]\!]$.
$N \equiv (p-1)![p]$
— Conclusion : $n^{p-1}(p-1)! \equiv (p-1)![p]$, i.e. $p|(n^{p-1}-1) \times (p-1)!$.
Comme p premier, pour tout $k < p$, $p \wedge k = 1$ et donc $p|n^{p-1} - 1$.
En multipliant par n : $n^p \equiv n[p]$
□

Pour aller plus loin - Fécondité d'un théorème
On peut choisir de mesurer l'intérêt d'un résultat mathématique aux domaines touchés.
Le petit théorème de FERMAT *dont la plupart des démonstrations datent d'*EULER *est assurément un théorème fécond. Chacune des démonstrations présentés ici exploite une partie différente des mathématiques : la structure du groupe $(\frac{\mathbb{Z}}{p\mathbb{Z}}, \times)$ pour la première, le morphisme de* FROEBENIUS *: $a \mapsto a^p$ dans le groupe $(\frac{\mathbb{Z}}{p\mathbb{Z}}, +)$ pour la seconde. La troisième démonstration assez proche de la première exploite un raisonnement de type combinatoire. La quatrième fait le lien entre la première et la troisième. Elle montre mieux : le petit théorème d'*EULER*, $n^{\varphi(p)} = 1[p]$.*

Voici la première démonstration historique d'Euler et probablement de Leibniz :
Exercice
Considérons p un nombre premier.

1. Montrer que $\forall k \in [\![1, p-1]\!], p|\binom{p}{k}$
2. $\forall (a,b) \in \mathbb{Z}^2, (a+b)^p \equiv a^p + b^p[p]$
3. En déduire le petit théorème de Fermat

Correction

1. Pour $k \neq 0$, $k \neq p$:
$$k\binom{p}{k} = k\frac{p!}{k!}(p-k)! = p\frac{(p-1)!}{(k-1)!(p-1-k+1)!} = p\binom{p-1}{k-1}$$
Or le coefficient $\binom{p-1}{k-1}$ est un nombre entier, donc p divise $k \times \binom{p}{k}$. Mais p premier, $k < p$, donc d'après le premier théorème d'Euclide : p divise $\binom{p}{k}$.
2. Soient $(a,b) \in \mathbb{Z}^2$, d'après le binôme de Newton,
$$(a+b)^p = \sum_{k=0}^{n} \binom{p}{k} a^k b^{p-k} \equiv a^p + 0 + b^p[p]$$
d'après l'arithmétique modulaire.
3. Démontrons maintenant par récurrence (sur n), le petit théorème de Fermat.
Posons, pour tout $n \in \mathbb{N}^*$, $\mathscr{P}_n$: « $n^p \equiv n[p]$ ».
— $1^p = 1$, donc $\mathscr{P}_1$ vraie.
— Soit $n \in \mathbb{N}^*$, supposons $\mathscr{P}_n$ vraie. On a d'après le résultat plus haut :
$$(n+1)^p \equiv n^p + 1^p \equiv n + 1[p]$$
Donc $\mathscr{P}_{n+1}$ est vraie.

Remarque - Morphisme de Fröbenius
Si p est premier, alors $(a+b)^p = a^p + b^p[p]$ et $(ab)^p = a^p b^p$.
Donc $x \mapsto x^p$ est un morphisme de corps $\frac{\mathbb{Z}}{p\mathbb{Z}}$. C'est un morphisme de corps, a priori non trivial (différent de l'identité). Il s'appelle le morphisme de Fröbenius.
Autre méthode, par double dénombrement :
Exercice
Soit p un nombre premier et n, un entier quelconque.
Considérons un alphabet de n lettres.

1. Combien y a-t-il de mots (possibles) de p lettres avec au moins deux lettres distinctes ? On notera N ce nombre.
2. On note $\mathscr{R}$, la relation d'équivalence sur l'ensemble des mots de p lettres $A = a_1 a_2 \dots a_p$:
$$A \mathscr{R} B \iff \exists\, h \mid B = a_{h+1} a_{h+2} \dots a_p a_1 \dots a_h$$
(Il existe une permutation circulaire des lettres pour passer des mots A à B.
On peut aussi noter que $A = A_1 A_2$ et $B = A_2 A_1$ avec A_1 premier sous-mot de A de longueur h et A_2 de longueur $p-h$.) Montrer qu'il s'agit bien d'une relation d'équivalence.

3. Combien existe-t-il de mots différents dans chaque classe d'équivalence ?
4. On note H le nombre de classe d'équivalence avec au moins deux mots.
   Quelle relation existe-t-il entre H, p et N ?
5. En déduire le petit théorème de Fermat.

Correction

1. Il y a n^p mots distincts, mais il y en n ayant exactement les mêmes lettres.
   Donc $N = n^p - n$.
2. Elle est réflexive et symétrique clairement.
   Si $B = a_{h+1}a_{h+2}\dots a_p a_1 \dots a_h = b_1 \dots b_p$ et $C = b_{k+1}\dots b_p b_1 \dots b_k = a_{h+k+1}\dots a_p a_1 \dots a_{k+h}$
   (si $h+k > p$, il faut prendre plutôt $h+k-p$.)
   donc la relation est transitive.
3. Au plus, il y a p mots différents dans chaque classe d'équivalence : celui qui commence par a_1, celui qui commence par $a_2 \dots$ par a_p.
   Supposons que $A = B$, avec $A = a_1 a_2 \dots a_p$ et $B = a_{h+1}a_{h+2}\dots a_p a_1 \dots a_h$.
   Alors $a_i = a_{h+i}$ pour tout $i \in [\![1, p-h]\!]$ et $a_j = a_{h+j-p}$ pour tout $j \in [\![p-h+1, p]\!]$
   Puis, se crée alors un cycle de même lettre (notons K sa taille) :

   $$a_1 = a_{h+1} = a_{2h+1} = \dots = a_{kh+1[p]} = \dots = a_{(K-1)h+1[p]} = a_{(Kh+1[p]} := a_1$$

   Et finalement $Kh + 1 \equiv 1[p]$ i.e. p divise Kh.
   D'après le premier théorème d'Euclide : $p|K$ et donc $K = p$:
   toutes les lettres sont les mêmes. Au final : ou bien les classes d'équivalence ne contiennent qu'un mot : ceux avec une seule lettre, ou bien ces classes ne contiennent que des mots différents, en nombre p.
4. Il y a n classes avec un seul mot,
   et il y a H classes avec p différents.
   On a donc $H \times p + n = n^p$
5. Donc $p|n^p - n$ i.e. $n^p \equiv n[p]$

La seconde démonstration d'Euler exploite (sans le savoir) un théorème dû par la suite à Lagrange.
Cet exercice est un bilan algébrique de l'exercice précédent et de la première démonstration.

Exercice

Soit p premier. Soit $n \in \mathbb{Z}$, supposons que p ne divise pas n.
Notons r le reste de la division euclidienne de n par p.

1. On note $G = ([\![1, p-1]\!], \times)$. Montrer que $(G, \times)$ est un groupe
2. Montrer qu'il existe $k \in \mathbb{N}$ tel que $r^k \equiv 1[p]$. On note $k_0 = \min\{k \in \mathbb{N} \mid r^k \equiv 1[p]\}$
3. Montrer que $\mathscr{R}$: $a\mathscr{R}b$ ssi $\exists\, k \in \mathbb{Z}$ tel que $a = r^k b$
   est une relation d'équivalence.
4. Quelle est la taille de chacune des classe d'équivalence ?
5. On note H le nombre de classe d'équivalence. Montrer que $p - 1 = k_0 \times h$.
   En déduire r^{p-1}, puis le théorème de Fermat

> ***Pour aller plus loin - Théorème d'Euler***
> *En affinant ce dernier exercice, on montre que : pour $t > 0$, n tel que $n \wedge t = 1$, alors $n^{\varphi(t)} \equiv n[t]$, avec $\varphi(a)$ est le nombre de diviseur de a.*
> *Pour cela on considère le groupe des $\varphi(t)$ éléments de $[\![1, t-1]\!]$ inversible (i.e. premier avec t).*
> *On notera que si p est premier : $\varphi(p) = p - 1$.*

Correction

1. D'après Bézout, il existe $u, v \in \mathbb{Z}$ tel que $ur + vp = 1$.
   Donc $ur \equiv 1[p]$, quitte à prendre $u_0 \equiv u[p]$, il existe $u_0 \in G$ tel que $u_0 \times r = 1$.
   La stabilité par produit est simple.
2. $\{r^k[p], k \in \mathbb{N}\} \subset G$ est un ensemble fini, il a au plus $p-1$ éléments.
   Donc il existe $k_1 < k_2 \in \mathbb{N}$ tel que $r^{k_1} \equiv r^{k_2}[p]$, alors $r^{k_2-k_1} \equiv 1[p]$.
3. Ok
4. Chacune des classes d'équivalence à k_0 éléments.
5. On a donc $\mathrm{Card}(G) = p - 1 = H \times k_0$, donc $r^{p-1} \equiv r^{H \times k_0} \equiv (r^{k_0})^H \equiv 1^H = 1[p]$.
   Enfin $n^{p-1} \equiv r^{p-1} \equiv 1[p]$ et l'on retrouve le théorème de Fermat.

En fait, si p est premier, on démontre que $r_0 = p - 1$ et $H = 1$.

6. Bilan

Synthèse

- ⇝ Les nombres entiers relatifs sont les premiers objets reconnus comme mathématiques rencontrés. Ils sont donc comme à la base du sentiment mathématique de tout apprenti mathématicien.
- ⇝ Plus généralement, au lieu d'étudier des nombres premiers relativement à a, on peut étudier les nombres premiers relativement à tous les nombres. Ce sont les nombres premiers, atomes minimaux des multiplications/divisions d'entiers.
- ⇝ Gauss a eu la merveilleuse idée de plonger $\mathbb{Z}$ dans des sous-ensemble modulo le nombre a qui nous intéresse. Cela marche bien car la structure d'anneau est conservée, mais cela peut être encore plus fort si a est premier, car on crée une structure de corps! Avec cette façon de penser, beaucoup de théorèmes devient faciles à démontrer voire à comprendre : c'est le cas du petit théorème de Fermat.
  Nous re-investirons ce point de vue lors de l'étude des polynômes
- ⇝ Nous terminons par l'étude hors-programme de quelques fonctions arithmétiques. La motivation est la même que celle pour les séries génératrices. Il est souvent mathématiquement plus simple d'étudier directement toute la suite (u_n) plutôt qu'un seul de ces éléments u_n...

Savoir-faire et Truc & **Astuce du chapitre**

- — Savoir-faire - Trouver un facteur premier, avec un nombre premier.
- — Savoir-faire - Caractérisations de $v_p(a)$.
- — Savoir-faire - Exploiter le petit théorème de Fermat

Notations

Notations	*Définitions*	*Propriétés*	*Remarques*
$\mathscr{P}_a$	Ensemble des nombres premiers avec a	$b \in \mathscr{P}_a \iff a \wedge b = 1$	Equivalent à $a \in \mathscr{P}_b$ (symétrie)
$\mathscr{P}$	Ensemble des nombres premiers	$\mathscr{D}(p) = \{1,-1,p,-p\}$ et $p \in \mathbb{N}$, $p \geqslant 2$	$p \in \mathscr{P} \Leftrightarrow \forall\ k \in [\![2, p-1]\!]$, $k \in \mathscr{P}_p$
$v_p(a)$	Valuation p-adique de a	$v_p(a) = r \Leftrightarrow a = p^r q$ avec $p \wedge q = 1$	On raisonne plutôt par double inégalité

Retour sur les problèmes

72. En effet, en simplifiant par $(n-1)!$, premier avec n (si n est premier), on trouve $k^{n-1} \equiv 1[n]$
73. Par exemple si l'on commence par les nombres premiers, le lemme de Gauss s'applique avec le théorème 1 d'Euclide...
    Nous ne faisons pas ici tous les 6(=3!) cas possibles
74. $f(p_1 p_2) = f(p_1) \times f(p_2)$ et il suffit de connaître $f(p)$, pour tout $p \in \mathbb{N}$: $f(\prod_{i=1}^{r} p_i^{r_i}) = \prod_{i=1}^{r} f(p_i)^{r_i}$.

Chapitre 17

Anneaux et corps

Résumé -

Après les groupes, deux nouvelles structures jouent un rôle important en mathématiques : les anneaux et les corps. Ils possèdent chacun deux lois internes et quelques régularités.
L'exemple type d'anneau est l'ensemble $\mathbb{Z}$. Nous baserons notre étude des anneaux sur ce que l'on a pu faire en arithmétique. En retour, nous verrons que le second exemple de l'année $\mathbb{K}[X]$ (anneau intégre des polynômes) possède beaucoup de points commun (notion de PGCD, primamlité...).
Les corps sont des anneaux où tous les éléments sont inversibles pour la seconde loi. C'est un ensemble fréquent : $\mathbb{Q}$, $\mathbb{R}$ ou $\mathbb{C}$, et il sera important pour l'étude des espaces vectoriels...

Sommaire

1. Problèmes

? Problème 78 - Structures fondamentales associées à $\mathbb{Z}$

Les deux chapitres précédents nous ont prouvé qu'il est possible de faire beaucoup de chose avec une structure aussi limitée que $\mathbb{Z}$.
Pouvons-nous généraliser? Quelles sont les propriétés fondamentales vérifiées par $\mathbb{Z}$?
Rappelons que $\mathbb{Z}$ est créé par addition $+1$. Mais que très vite, c'est la multiplication qui nous intéresse et en particulier la décomposition (unique) en facteurs premiers.

? Problème 79 - Théorème de Bézout et le lemme de Gauss dans un anneau

Notons (a) et (b), l'ensemble des multiples de a et b respectivement.
Nous avons vu que le théorème de Bézout était d'une importance capitale. Il exprime que $\mathbb{Z} = (a) + (b)$ lorsque $a \wedge b = 1$.
Plus généralement, comment s'exprime-t-il pour des anneaux? Et le lemme de Gauss?

? Problème 80 - Quotienter un anneau

Considérons A, un anneau et munissons le d'une relation d'équivalence $\mathscr{R}$.
A quelle condition suffisante (nécessaire) sur $\mathscr{R}$, peut-on transformer l'ensemble des classes d'équivalence $\frac{A}{\mathscr{R}}$ (avec les lois induites) en anneau voire en corps?

2. Structures d'anneau

2.1. Définitions et propriétés premières

Définition d'un anneau

Définition - Anneaux
Soit A un ensemble muni de deux lois de composition internes notées $+$ et $\star$. On dit que $(A,+,\star)$ est un *anneau* si :
- $(A,+)$ est un groupe commutatif;
- la loi $\star$ est associative;
- la loi $\star$ est distributive par rapport à la loi $+$:

$$\forall (x,y,z) \in A^3,\ x \star (y+z) = x \star y + x \star z \quad \text{(distibutive à gauche)}$$

$$\forall (x,y,z) \in A^3,\ (x+y) \star z = x \star z + y \star z \quad \text{(distibutive à droite)}$$

- A possède un élément neutre pour $\star$, noté 1.

Si de plus la loi $\star$ est commutative, on dit que $(A,+,\star)$ est un anneau commutatif.

Exemple - Anneaux classiques

Les deux anneaux essentiels vus cette année : $\mathbb{Z}$ et $\mathbb{K}[X]$.
Un autre anneau important (mais qui est surtout une algèbre) : $\mathcal{M}_n(\mathbb{K})$.
Un dernier exemple : $\frac{\mathbb{Z}}{n\mathbb{Z}}$ (on en reparlera plus loin)
Soit $(A,+,\star)$ un anneau. On note 0 l'élément neutre de + et 1 celui de $\star$.

Règles de calcul immédiates

Proposition - Lien + et $\star$
On a les relations suivantes :
- $\forall x \in A,\ x\star 0 = 0\star x = 0$ (on dit que 0 est absorbant).
- $\forall (x,y)\in A^2,\ (-x)\star y = x\star(-y) = -(x\star y)$
- $\forall (x,y)\in A^2,\ (-x)\star(-y) = x\star y$

Démonstration
Pour tout $x\in A$, $0+0=0$, car $(A,+)$ est un groupe. Donc

$$0\star x = (0+0)\star x = 0\star x + 0\star x \Longrightarrow 0\star x = 0$$

par régularité dans le groupe $(A,+)$.
De même, on montre que $x\times 0 = 0$.
$(x\star y)+((-x)\star y) = (x+(-x))\star y = 0\star y = 0$, donc $(-x)\star y = -(x\star y)$.
de même $(x\star y)+(x\star(-y)) = (x)\star(y+(-y)) = x\star 0 = 0$, donc $x\star(-y) = -(x\star y)$ Puis $(-x)\star(-y) = -(x\star(-y)) = -(-(x\star y)) = x\star y$ □

Intégrité

Définition - Diviseur de 0 et anneau intègre
Soit $(A,+,\star)$ un anneau.
- $a\in A\setminus\{0\}$ est un diviseur de 0 si il existe $b\in A\setminus\{0\}$ tel que $a\star b = 0$ ou $b\star a = 0$.
- A est dit intègre s'il est commutatif et sans diviseurs de 0.

Pour aller plus loin - Commutativité
Il arrive que certains mathématiciens n'exigent pas la commutativité comme condition à l'intégrité (ex : cours d'Algèbre de Roger Godement), mais c'est une exception à la tradition largement adoptée.

Proposition - Simplification (division)
Si A est un anneau intègre, tout élément non nul a de A est régulier pour $\star$, c'est-à-dire que l'on peut simplifier par a :

$$a\star b = a\star c \Rightarrow b = c.$$

Démonstration
$a\star b = a\star c \Rightarrow a\star b + a\star(-c) = a\star c + a\star(-c) \Rightarrow a\star(b-c) = 0 \Rightarrow b = c$
car $a\neq 0$ et A intègre. □

Exemple - $\mathbb{Z}$ et $\mathcal{M}_n(\mathbb{K})$

$\mathbb{Z}$ est un anneau intègre. $\mathcal{M}_n(\mathbb{K})$ n'est pas intègre ($n\geqslant 2$)

Savoir faire - Exploiter l'intégrité

On exploite l'intégrité dans son sens contraposée : $a\neq 0$ et $b\neq 0 \Longrightarrow ab\neq 0$.
En particulier, si on sait qu'un ensemble est un corps (comme $\frac{\mathbb{Z}}{p\mathbb{Z}}$, alors il est intègre.

Règles fréquentes

Proposition - Quelques règles de calcul
Les règles de calculs fréquentes :
— $(a_i)_{1\leqslant i\leqslant n}$ et $(b_j)_{1\leqslant j\leqslant p}$ deux familles d'éléments de A. Alors on peut écrire :
$$\left(\sum_{i=1}^{n} a_i\right)\left(\sum_{j=1}^{p} b_j\right) = \sum_{(i,j)\in[\![1,n]\!]\times[\![1,p]\!]} a_i b_j$$
— formule du binôme : si a et b **commutent pour** $\star$ alors
$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}$$
— factorisation : si a et b **commutent pour** $\star$ alors
$$a^n - b^n = (a-b)\sum_{k=0}^{n-1} a^{n-k-1} b^k$$
en notant $xy = x\star y$ et les puissances étant au sens de la loi $\star$.

Remarque - Démonstration ?
Il s'agit exactement des mêmes démonstrations que celles vues pour les nombres réels en début d'année.
En effet, les seules hypothèses qui ont été mobilisées étaient celles de commutativité des nombres (objets).

Un groupe pour $\star$: le groupe des inversibles

Proposition - Groupe des inversibles
L'ensemble des éléments inversibles de $(A,\star)$ est un groupe pour la loi $\star$.
Classiquement, ce groupe est noté $A^\times$.

Démonstration
A est non vide et possède au moins 1.
Soient $x, y \in A$, inversibles,
$$(yx^{-1})\star(xy^{-1}) = yx^{-1}xy^{-1} = yy^{-1} = 1 \qquad (xy^{-1})\star(yx^{-1}) = \cdots = 1$$
Donc xy^{-1} est inversible et donc $xy^{-1}\in A^*$.
Dans A^*, $\times$ est associatif.
L'ensemble des inversibles de A est un groupe. □

Exemple - $\mathbb{Z}^*$
$\mathbb{Z}^* = \{-1, 1\}$

Exemple - $(\mathcal{M}_n(\mathbb{K}))^*$
L'ensemble des matrices inversibles se note : $\mathcal{M}_n(\mathbb{K})^* = GL_n(\mathbb{K})$.
On parle du groupe linéaire des matrices inversibles (à coefficients dans le corps $\mathbb{K}$.

2.2. Construction d'anneaux

Sous-anneaux

Définition - Sous-anneau
Soit $(A,+,\star)$ un anneau. $B\subset A$ est un *sous-anneau* de A si B est stable pour les lois internes $+$ et $\star$ et si ces lois induites munissent B d'une structure d'anneau, donc, avec $1\in B$.

$(B,+)$ est nécessairement un groupe, stable également pour $\star$: la réciproque est suffisante :

Savoir faire - Caractérisation des sous-anneaux

Soit B une partie de A. B est un sous-anneau de A si et seulement si il vérifie :

— $1 \in B$
— $\forall (x,y) \in B^2,\ x-y \in B$
— $\forall (x,y) \in B^2,\ x \star y \in B$

Démontrons que ce savoir-faire est juste.

Démonstration

Soit B une partie de A. Supposons que $(B,+,\star)$ (lois induites donc stabilités...) soit un sous-anneau

Par hypothèse (définition de sous-anneau) : $1 \in B$.

Alors $(B,+)$ est un sous-groupe de $(A,+)$ et donc on la caracrétisation 2 du sous-groupe :
$\forall (x,y) \in B^2,\ x-y \in B$. Puisque $\star_B = \star$ est une loi interne sur B : $\forall (x,y) \in B^2,\ x \star y \in B$.

Réciproquement supposons que les trois points caractéristiques du savoir-faire sont vérifiés.

Alors B est stable par $\star$ d'après le point 3 respectivement,

il est également stable par $+$ car $0_A = 0_B$, puis si $y \in B$, $-y = 0_B - y \in B$ et $x+y = x-(-y) \in B$ pour $x, y \in B$.

Puis $(B,+_B,\star_B) = (B,+,\star)$ est bien un anneau car :

$(B,+)$ est un groupe avec la caractérisation du point 2.

$\star_B = \star$ reste associatif et distributif par rapport à $+_B = +$.

$(B,\star)$ possède l'élément neutre $1_B = 1$. □

Exercice

Montrer que l'intersection de deux sous-anneaux de A est un sous-anneau de A.
(On pourrait étendre à une intersection d'une famille de sous-anneaux

Correction

Notons les B_1 et B_2. Alors • $1 \in B_1$ et $1 \in B_2$, donc $1 \in B_1 \cap B_2$.
• Si $x, y \in B_1 \cap B_2$, alors $x, y \in B_1$ donc $x-y \in B_1$ et $x, y \in B_2$ donc $x-y \in B_2$.
Donc $x-y \in B_1 \cap B_2$.
• Si $x, y \in B_1 \cap B_2$, alors $x, y \in B_1$ donc $x \star y \in B_1$ et $x, y \in B_2$ donc $x \star y \in B_2$.
Donc $x \star y \in B_1 \cap B_2$.

Exemple - $2\cdot\mathbb{Z}$ est-il un sous-anneau de $(\mathbb{Z},+,\times)$

Si $a, b \in 2\mathbb{Z}$, alors $2|a-b$ et $2|a \times b$ mais $1 \notin 2\cdot\mathbb{Z}$, ce n'est pas un sous-anneau.
En fait un s'agit d'un idéal, on en reparlera avec les anneaux euclidiens $\mathbb{Z}$ et $\mathbb{K}[X]$.

Pour aller plus loin - Idéal

On dit que I $(\subset A)$ est un idéal de A, si $(I,+) < (A,+)$ et $\forall\ x \in I, y \in A,\ x \star y \in I$ et $y \star x \in I$ (propriété d'absorbance. Voir plus loin...

Morphisme (et image) d'anneaux

Définition - Morphisme d'anneaux

Soient $(A,+_A,\star_A)$ et $(A',+_{A'},\star_{A'})$ deux anneaux. Un morphisme d'anneaux de A dans A' est une application f de A dans A' vérifiant :

— $\forall (x,y) \in A^2,\ f(x+_A y) = f(x) +_{A'} f(y)$
— $\forall (x,y) \in A^2,\ f(x\star_A y) = f(x) \star_{A'} f(y)$
— $f(1_A) = 1_{A'}$

Exemple - Projection canonique

Soit $F : \mathbb{Z} \to \frac{\mathbb{Z}}{n\mathbb{Z}},\ m \mapsto \overline{m}$.
Alors $F(m_1+m_2) = \overline{m_1+m_2} = \overline{m_1} + \overline{m_2}$ (on fera une vraie démonstration plus loin).
Alors $F(m_1 \times m_2) = \overline{m_1 \times m_2} = \overline{m_1} \times \overline{m_2}$.

Exemple - Sur $\mathbb{C}$

L'application $z \mapsto \overline{z}$ est également un morphisme d'anneaux (de corps ici).
Avec l'identité, ce sont les seuls morphismes d'anneaux surjectifs de $\mathbb{C}$ sur $\mathbb{C}$.

Exemple - Morphisme de Fröbenius

Nous verrons, dans l'une des démonstrations du petit théorème de Fermat que si p est premier, alors p divise $\binom{k}{p}$, pour tout $k \notin \{0,p\}$.

Alors $\frac{\mathbb{Z}}{p\mathbb{Z}} \to \frac{\mathbb{Z}}{p\mathbb{Z}},\ m \mapsto m^p$ (p premier) est un morphisme d'anneaux.

On le démontre à l'aide du binôme de Newton.

Proposition - Transfert par morphisme d'anneaux
Soit $f : A \to A'$ un morphisme d'anneaux. Alors
— $f(0_A) = 0_{A'}$
— $\forall\, x \in A,\ f(-x) = -f(x)$ et $\forall\, x \in A^*\ f(x^{-1}) = f(x)^{-1}$.

Démonstration
C'est un morphisme de groupe, donc $f(0) = 0$ ($f(x) = f(x+0) = f(x) + f(0)$. Donc $f(0) = 0$).
De même $0 = f(0) = f(x + (-x)) = f(x) + f(-x)$. Donc $f(-x) = -f(x)$.
$(A^*, \times)$ est un groupe. On applique exactement la même démonstration. □

Proposition - Morphisme du groupe $(A^x, \star)$
Soit $f : A \to A'$ un morphisme d'anneaux. Alors $f^\times : A^\times \to A'^\times$, $x \mapsto f(x)$ est un morphisme de groupes.

Démonstration
$f(1_A) = 1_{A'}$.
Soient $x, y \in A^x$.
$1_{A'} = f(1_A) = f(y \star_A y^{-1}) = f(y) \star_{A'} f(y^-1)$, donc $f(y^{-1}) = \big(f(y)\big)^{-1}$. Puis

$$f(x \star_A y^{-1}) = f(x) \star_{A'} f(y^{-1}) = f(x) \star_{A'} \big(f(y)\big)^{-1}$$

A noter que f n'est ni nécessairement surjectif, ni injectif… □

Exercice
On dit que $I \subset A$ est un idéal de l'anneau A, si
— $(I, +)$ est un sous-groupe de $(A, +)$.
— $\forall\, x \in I, y \in A,\ x \star y \in I$ et $y \star x \in I$
Soit $f : A \to A'$ un morphisme d'anneaux. Montrer que $\mathrm{Ker}\, f$ est un idéal de A.

Correction
f est également un morphisme de groupes. Donc $(\mathrm{Ker}\, f, +)$ est un sous-groupe de $(A, +)$.
Soient $x \in \mathrm{Ker}\, f$ et $y \in A$.
Alors $f(x \star y) = f(x) \star f(y) = 0 \star f(y) = 0$, donc $x \star y \in \mathrm{Ker}\, f$. De même pour $y \star x$.
Donc $\mathrm{Ker}\, f$ est bien un idéal de A.

2.3. Idéaux

Dans la suite les anneaux sont considérés commutatifs.

Extension de la congruence

on étend à tous les anneaux, la notion de congruence vu dans $\mathbb{Z}$.

Pour aller plus loin - Cas A ***non commutatif***
Si A n'est pas commutatif, il faut étudier les multiples à droite et les multiples à gauche…

Définition - Multiple dans un anneau
Soit a un élément d'un anneau A. On appelle multiple de a, les éléments de l'ensemble $(a) = \{a \times d, d \in A\}$, (parfois noté aA).
On dit que a divise b (noté $a|b$ si b est un multiple de a.

Définition - Congruence dans un anneau
Soit m un éléments d'un anneau A. Soient a, b deux éléments de A.
On dit que a est congru à b modulo m, noté $a \equiv b[m]$ ssi $b - a \in (m)$ (ou $m|b-a$).

Remarque - Notation multiple
On peut écrire au choix : $m|n$ ou $n \in (m)$ ou encore $(n) \subset (m)$.

Proposition - Relation d'équivalence
Dans un anneau, la relation de congruence modulo m est une relation d'équivalence

Exercice
Faire la démonstration

Correction
- $0 \in (m)$, donc $\equiv [m]$ est réflexif.
- si $k \in (m)$ alors $-k \in (m)$ donc $\equiv [m]$ est symétrique.
- si $k_1, k_2 \in (m)$ alors $k_1 + k_2 \in (m)$ donc $\equiv [m]$ est transitif.

Compatibilité

Théorème - Compatibilité
Si A est un anneau (commutatif) et $m \in A$.
Alors l'addition et la multiplication sont compatibles pour la relation d'équivalence $\equiv [m]$.
Autrement écrit, l'addition et la multiplication sont indépendants du choix du représentant de la classe d'équivalence ; on peut donc définir une addition et une multiplication sur les classes d'équivalence :

$$\overline{x}+\overline{x'} = \overline{x+x'} \qquad \overline{x}\times\overline{x'} = \overline{x\times x'}$$

Démonstration
Supposons que $x \equiv y[m]$ et $x' \equiv y'[m]$.
On peut note $x - y = mk$ et $x' - y' = mk'$, alors

$$x + x' = y + y' + m(k + k') \Longrightarrow x + x' \equiv y + y'$$

Donc la relation de congruence est compatible avec l'addition.

$$x \times x' = (y + mk) \times (y' + mk') = y \times y' + m(ky' + k'y + mkk')$$

Donc la relation de congruence est compatible avec la multiplication.
□

Idéaux

Analyse - Ce qui a marché
Si l'on relit la démonstration, la réussite de compatibilité es lié aux faits que :
- Si $a, a' \in (m)$ alors $a + a' \in (m)$.
  Ici $a = x - y$ et $a' = x' - y'$
- Si $a \in (m)$ et $b \in A$ alors $a + b \in (m)$
  Ici $a = mk$ et $b = y'$, puis $a = mk'$ et $b = y$

Définition - Idéal de A
Soit A un anneau.
On appelle idéal de A, toute partie I de A tel que :
- $0 \in I$
- $(I, +)$ est un sous-groupe de $(A, +)$ (noté $I < A$)
- $\forall\ a \in I, \forall\ b \in A, ab \in I$

Pour aller plus loin - Cas A non commutatif
Si A n'est pas commutatif, il faut étudier les idéaux à droite ($ba \in I$ si $a \in I$) et les idéaux à gauche ($ab \in I$ si $a \in I$)...

On se souvient que l'on a déjà rencontré $2\cdot\mathbb{Z}$ qui est un idéal mais pas un sous-anneau de $\mathbb{Z}$. Exercice
Quels sont les idéaux de $\mathbb{Z}$?

Correction
Exactement les ensembles $a\mathbb{Z}$, pour $a\in\mathbb{Z}$

Remarque - Idéal engendré
Comme pour les sous-groupes engendrés (et plus tard les sous-espaces vectoriels engendrés), on définit les idéaux engendrés par une partie B de l'anneau $(A,+,\times)$ comme le plus petit des idéaux contenant B.
On l'obtient comme intersection (décroissante) des tous les idéaux contenant B.
Mais c'est aussi l'ensemble obtenu en faisant agir les éléments de B le plus largement possible... Exercice
Montrer que si I et J sont deux idéaux de $(A,+\times)$, alors $I\cap J$ et $I+J(:=\{a+b\mid a\in I, b\in J\})$ sont des idéaux de A.

Correction
• $I\cap J$.
$0\in I$, $0\in J$, donc à $\in I\cap J$.
$(I,+)<(A,+)$, $(J,+)<(A,+)$ donc $(I\cap J,+)<(A,+)$ par propriété des sous-groupes.
Soient $a\in I\cap J$ et $b\in A$, alors $a\in I$, donc $ab\in I$ et $a\in J$ donc $ab\in J$ donc $ab\in I\cap J$.
• $I+J$.
$0=0+0\in I+J$, car $0\in I$ et $0\in J$
Soient $x,y\in I+J$, il existe $a,c\in I$, $b,d\in J$ tel que $x=a+b$ et $y=c+d$.
Puis $x-y=(a-b)+(c-d)\in I+J$. Donc $I+J<A$.
Soient $x=a+b\in I+J$ et $y\in A$, alors $x\times y=ay+by\in I+J$.

Sous-anneaux quotients

Heuristique - Quotientage d'un anneau par un idéal
Parmi les sous-groupes, les sous-groupes distingués permettaient de prolonger la loi interne (par compatibilité) à la structure quotiente qui devenait ainsi un groupe (quotient).
Formellement : si $(H,+)\lhd(G,+)$, alors $\left(\frac{G}{H},\overline{+}\right)$ est un groupe.
Il en est de même pour le quotient d'un anneau par un idéal

Proposition - Anneau quotient
Soit $(A,+,\star)$ un anneau (commutatif) et I un idéal.
Alors $\left(\frac{A}{I},\overline{+},\overline{\star}\right)$ est un anneau (quotient).

Rappelons que $\frac{A}{I}$ désigne l'ensemble des classes d'équivalence de A pour la relation $a\equiv b\iff a-b\in I$.

Démonstration
Le plus dur est de montrer la comptabilité des règles opératoires. Mais I étant idéal, tout va bien.
En effet, par construction :
$$\overline{a}\overline{+}\overline{b}=\overline{a+b}\qquad \overline{a}\overline{\star}\overline{b}=\overline{a\star b}$$
Puis l'élément neutre pour $\overline{+}$ est $\overline{0}$, celui pour $\overline{\star}$ est $\overline{1}$.
Plus précisément encore, $\frac{A}{I}$ est un anneau en tant qu'image de A par la projection (surjection) canonique $a\mapsto\overline{a}$. □

Exercice
Montrer que si f est un morphisme d'anneaux A sur B.
Alors $\operatorname{Ker}f=\{x\in A\mid f(x)=0_B\}$ est un idéal de A.
Puis en déduire que $\frac{A}{\operatorname{Ker}f}$ est un anneau

Correction
C'est comme pour les groupes distingués.
Si $x,y\in\operatorname{Ker}f$, alors $f(x-y)=f(x)-f(y)=0-0=0$ car f est un morphisme.
Donc $x-y\in\operatorname{Ker}f$.
Si $x\in\operatorname{Ker}f$ et $a\in A$, alors $f(xa)=f(x)f(a)=0\times f(a)=0$.
Donc $xa\in\operatorname{Ker}f$.
Nous avons enfin une définition propre d'un ensemble dont on a beaucoup parlé.
Puis, comme $a\mathbb{Z}$ est un idéal :

Corollaire - Anneau quotient de $\mathbb{Z}$
Soit $a \in \mathbb{Z}$, l'ensemble $\left(\frac{\mathbb{Z}}{a\mathbb{Z}}, \overline{+}, \overline{\times}\right)$ des classes d'équivalence de $\mathbb{Z}$ est un anneau

2.4. Anneau euclidien. Anneau principal

Proposition - Anneau principal
Soit A un anneau.
On dit qu'un idéal I de A est principal si il existe $a \in I$ tel que $I = (a) (= aA)$.
On dit qu'un anneau est principal s'il est intègre et que tous ses idéaux sont principaux.

Pour aller plus loin - Anneau non principal
L'anneau des polynômes à coefficients entiers $\mathbb{Z}[X]$ n'est pas prinicipal.
En effet, l'idéal engendré dans $\mathbb{Z}[X]$ par $< 2, X >$ n'est pas principal.

Exemple - $\mathbb{Z}$ est principal
Les idéaux de $\mathbb{Z}$ sont les sous-groupes de $(\mathbb{Z}, +)$ donc de la forme $a\mathbb{Z}$.

Savoir faire - Montrer qu'un anneau est principal
Une méthode qui ne marche pas toujours est de montrer qu'un tel anneau est d'abord euclidien.

Définition - Anneau euclidien
Soit A un anneau intègre. On dit que A est euclidien s'il existe une application $\varphi : A \setminus \{0\} \to \mathbb{N}$ (appelé **stathme** euclidien) telle que :

$\forall\ (a,b) \in A \times A \setminus \{0\}, \exists\ (q,r) \in A^2$ tel que $a = bq + r$ avec $r = 0$ ou $\varphi(r) < \varphi(b)$

Notons que l'unicité n'est pas demandé.

Proposition - Anneau euclidien $\Rightarrow$ Anneau principal
Si A est euclidien, alors A est principal

Démonstration
Soit I un idéal de A euclidien.

$$\{\varphi(r), r \in I \setminus \{0\}\} \subset \mathbb{N}$$

Cet ensemble est non vide, inclus dans $\mathbb{N}$ donc admet un plus petit élément $s = \varphi(m)$.
Soit $a \in I$. Faisons la division euclidienne de a par m.
Il existe donc $q, r \in A$ tel que $a = mq + r$, avec $\varphi(r) < \varphi(m)$ ou $\varphi(r) = 0$.
Or I est un idéal, donc $mq \in I$ puis $a - mq \in I$, donc $r \in I$.
Comme $s = \varphi(m)$ est minimal, nécessairement, $r = 0$, et donc $m|a$.
Réciproquement, comme $(m) \subset I$.
On a donc $I = (m)$. I est principal.
Ainsi A est un anneau principal □

Exemple - Nombreux
$\mathbb{Z}$ et $\mathbb{K}[X]$ sont euclidiens donc principaux.
Exercice
On note $\mathbb{Z}[i] = \{a + ib; a, b \in \mathbb{Z}\}$, l'ensemble des entiers de Gauss.

1. Trouver une division euclidienne sur $\mathbb{Z}[i]$
   On prendra, le carré de la fonction module comme stathme
2. En déduire que $\mathbb{Z}[i]$ est principal

Correction

1. Supposons que $a=u+iv$ et $b=x+iy$, alors $\dfrac{a}{b}=\dfrac{(ux-vy)+i(uy+vx)}{x^2+y^2}$.
Notons $m=\begin{cases}\lfloor\frac{ux-vy}{x^2+y^2}\rfloor & \text{si } \theta(\frac{ux-vy}{x^2+y^2})\leqslant\frac{1}{2}\\ \lfloor\frac{ux-vy}{x^2+y^2}\rfloor+1 & \text{si } \theta(\frac{ux-vy}{x^2+y^2})>\frac{1}{2}\end{cases}$ et $n=\begin{cases}\lfloor\frac{uy+vx}{x^2+y^2}\rfloor & \text{si } \theta(\frac{uy+vx}{x^2+y^2})\leqslant\frac{1}{2}\\ \lfloor\frac{uy+vx}{x^2+y^2}\rfloor+1 & \text{si } \theta(\frac{uy+vx}{x^2+y^2})>\frac{1}{2}\end{cases}$,
puis $s=\dfrac{a}{b}-(m+in)$.
Alors $a=bq+r$ (avec $q=m+in\in\mathbb{Z}[i]$) et $r=sb$.
Puis $|r|^2=(sb)\overline{(sb)}=|s|^2|b|^2$.
Or $|s|^2=\mathrm{Re}^2(s)+\mathrm{Im}^2(s)=\theta_1^2+\theta_2^2\leqslant 2\times\dfrac{1}{2^2}=\dfrac{1}{2}<1$ donc $|r|^2<|b|^2$.

2. $\mathbb{Z}[i]$ est eucliden donc principal.

3. Structures de corps

3.1. Corps

Définition - Corps
Un corps est un anneau commutatif $(K,+,\times)$ dans lequel tous les éléments autres que 0 sont inversibles pour $\times$ c'est-à-dire que :
$(K,+,\times)$ est un corps si :
- $(K,+)$ est un groupe commutatif;
- $(K^*,\times)$ est un groupe commutatif, où 0 désigne l'élément neutre de K pour + et $K^*=K\setminus\{0\}$.
- la loi $\times$ est distributive par rapport à la loi +;

Exemple - Nombreux
Le corps $\mathbb{Q}$ est créé à partir de $\mathbb{Z}$, qui n'admet pas d'inverse, dans le but de rendre tous les éléments (sauf 0) inversibles.
De même, le corps $\mathbb{K}(X)$ des fractions rationnelles a le même rôle pour l'ensemble des polynômes $\mathbb{K}[X]$ que Q pour $\mathbb{Z}$.

Proposition - Tout élément est régulier
Un corps n'a pas de diviseurs de 0. Tout élément autre que 0 est donc régulier (on peut simplifier).

Démonstration
Si $ab=0$, alors si $a\neq 0$, donc inversible : $b=a^{-1}ab=a^{-1}0=0$.
Et donc $b=0$.
Ainsi aucun corps n'a de diviseurs de 0. □

3.2. Idéaux maximaux

Analyse - A quel condition un anneau quotient est-il un corps?
Il faut que tout élément $\overline{x}$ soit inversible. Donc qu'il existe $y\in A$ tel que $xy-1\in I$.
On a donc $1\in I+(x)$ et ainsi $A=(1)\subset I+(x)$.
Nécessairement, il faut donc, pour que tout $\overline{x}$ soit inversible, que (x) ne contienne par I (sinon $I+(x)=(x)\neq A$).

Définition - Idéal maximal
Soit I un idéal de A.
On dit que I est maximal s'il $I\neq A$ et A est le seul idéal distinct de I, contenant I

Pour aller plus loin - Idéal engendré
Si I_1 et I_2 sont deux idéaux, alors $I_1+I_2=\{a_1+a_2;a_1\in I_1,a_2\in I_2\}$ est un idéal; c'est le plus petit des idéaux qui contient à la fois I_1 et I_2.

Proposition - Corps
Soit I un idéal maximal de A. Alors $\left(\frac{A}{I}, \overline{+}, \overline{\times}\right)$ est un corps.

Remarque - Réciproque
Comme le montre l'analyse la réciproque est vrai.

Exemple - $6\mathbb{Z}$ n'est pas maximal
Les multiples de 6 sont des multiples de 3 : $6\mathbb{Z} \subset 3\mathbb{Z} \neq \mathbb{Z}$. Donc $6\mathbb{Z}$ n'est pas maximal.
D'après la réciproque : $\frac{\mathbb{Z}}{6\mathbb{Z}}$ n'est pas un corps.

Démonstration
Supposons que I est maximal.
Soit $\overline{x} \in \frac{A}{I}$.
Soit $J = I + (x)$ (plus petit idéal contenant I et (x)). Alors J est un idéal de A contenant I et différent de I (si $x \notin I$).
Or I est maximal, $J = A$ et donc $1 \in I + (x)$. Et donc il existe $a \in I$, $u \in A$ tel que $1 = a + ux$ et donc $1 = \overline{u \times x}$.
Donc $\overline{x}$ est inversible.
□

Exemple - $\frac{\mathbb{Z}}{p\mathbb{Z}}$ avec p premier
Dernier exemple et non des moindres : $\frac{\mathbb{Z}}{p\mathbb{Z}}$, avec p premier.

$$\frac{\mathbb{Z}}{p\mathbb{Z}} = \{0, 1, 2, \dots p-1\} \qquad \text{avec } u + v \equiv u + v\,[p] \text{ et } u \times v \equiv u \times v\,[p].$$

On a bien : Tout élément est inversible :
Soit $u \in \{0, 1, 2, \dots p-1\}$, et p premier, donc $u \wedge p = 1$.
D'après Bézout : il existe $a, b \in \mathbb{Z}$ tels que $au + bp = 1$, donc $au \equiv 1\,[p]$
Et donc u est inversible d'inverse a (ou son congru dans $\{0, 1, 2, \dots p-1\}$

Pour aller plus loin - Idéal premier
On dit que I est premier ssi $x \notin I$ et $y \notin I \Rightarrow xy \notin I$.
On a, par exemple, $n\mathbb{Z}$ est premier ssi n est premier.
Et plus généralement, $\frac{A}{I}$ est un anneau intègre ssi I idéal premier.
Dans le cadre de $A = \mathbb{Z}$, $\frac{A}{n\mathbb{Z}}$ est intègre ssi $\frac{A}{n\mathbb{Z}}$ est un corps…

3.3. Sous-corps. Morphisme (de corps)

Définition - Sous-corps, morphisme
On peut généraliser les définitions précédentes.
- Un sous-corps est un sous-anneau muni d'une structure de corps.
- Un morphisme de corps est un morphisme d'anneaux.
- L'image d'un corps par un morphisme de corps est un corps.

Le dernier point est un exercice à démontrer.

4. Bilan

Synthèse

⇝ Les anneaux sont les structures naturelles pour deux lois internes (addition, et multiplication, ou composition). Nous avons de nombreux exemples : $\mathbb{Z}$, $\mathbb{K}[X]$, $\mathcal{M}_n(\mathbb{K})$…
Avec l'inversibilité de tous les éléments non nuls, la structure est encore plus riche, elle s'appelle un corps. Les exemples classiques sont $\mathbb{R}$, $\mathbb{Q}$, $\mathbb{C}$, voire $\frac{\mathbb{Z}}{p\mathbb{Z}}$ (p premier)

⇝ Mais pour ce dernier exemple, il faut commencer par s'assurer de la bonne définition des lois $\overline{+}$ et $\overline{\times}$, lorsqu'on passe de $\mathbb{Z}$ à $\frac{\mathbb{Z}}{n\mathbb{Z}}$.
Nous avons vu qu'il fallait que l'ensemble $n\mathbb{Z}$, soit un d'abord un idéal pour que le calcul ait un sens.
Mieux pour que $\frac{\mathbb{Z}}{n\mathbb{Z}}$ soit un corps, (il faut et)il suffit que $n\mathbb{Z}$ soit un idéal maximal (équivalent à idéal premier, dans ce contexte)

⇝ Les structures d'anneaux ou de corps, se transfère par morphisme (d'anneaux) ou par restriction.

Savoir-faire et Truc & Astuce du chapitre

— Savoir-faire - Caractérisation des sous-anneaux
— Savoir-faire - Montrer qu'un anneau est principal

Notations

Notations	*Définitions*	*Propriétés*	*Remarques*
$(A,+,\star)$ parfois A	Anneau	$(A,+)$ groupe; $\star$ l.c.i. associative et unifère; distributivité	Exemples courants : $\mathbb{Z}$, $\mathbb{K}[X]$, $\mathcal{M}_n(\mathbb{K})$, $\frac{\mathbb{Z}}{n\mathbb{Z}}$
I	Idéal de A	$(I,+) < (A,+)$ & $\forall\ a \in A, x \in I,\ ax \in I$	
$(\mathbb{K},+,\star)$ parfois $\mathbb{K}$	Corps	$(\mathbb{K},+,\star)$ anneau, tout élément non nul inversible	Exemples courants : $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$, $F_p = \frac{\mathbb{Z}}{p\mathbb{Z}}$

Retour sur les problèmes

75. Cours
76. Bézout : $(a)+(b)=A$ et lemme de Gauss : $bc \in (a)$ et $(b)+(a)=A$ (a et b étrangers) alors $c \in (a)$.
77. Cours

Cinquième partie

Analyse réelle

Sixième partie

Algèbre linéaire & bilinéaire

Septième partie

Combinatoire et groupe fini

Huitième partie

Polynômes et fractions rationnelles

Neuvième partie

Probabilités

Dixième partie

Analyse (2)

Annexes

Annexe A

Suite de variable aléatoire. Convergence (HP)

Résumé -

Totalement hors-programme (sauf la loi faible des grands nombres, au programme de seconde année, mais sans le vocabulaire bien adapté), on peut considérer cette partie comme une application des deux chapitres précédents.
Ce qui unit les différents résultats ici, est la notion de suite de variables aléatoires. Et en particulier, la question de la convergence associée.
Plusieurs modes de convergence sont assez naturels, nous en proposons trois ici. En seconde année, les différents modes de convergence de suites de fonctions sont également un morceau important. On peut faire un parallèle entre ces deux parties.
C'est aussi l'occasion de revenir sur un résultat parachuté en terminale : le théorème de Moivre-Laplace, conséquence d'un sommet des mathématiques : le théorème limite central, mais qui nécessite, à notre avis de passer par un certain nombre d'étapes intermédiaires (définition, vocabulaire, démonstration…), pour vraiment comprendre ce que l'on manipule.
Quelques vidéos :

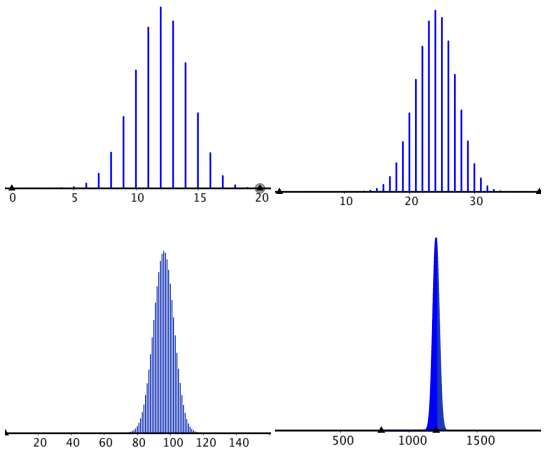
- *Maths en tête : Alexandre Morgan - La dynastie des Bernoulli / Maths C qui ? #10 - https ://www.youtube.com/watch ?v=jWut-6jBl3U*
- *Sur le Chemin des Maths - Histoire des Mathématiques 03 : Abraham de Moivre de la loi binomiale à la loi normale - https ://www.youtube.com/watch ?v=uLbPKe3LT28*
- *Statoscope - Convergence en loi vs convergence en probabilité - https ://www.youtube.com/watch ?v=9O1ves_L2eM*

Sommaire

1. Problèmes

? Problème A1 - $\frac{S_n}{n} \to$?

On a fait la représentation pour $p = 0,6$ et $n = 20$, $n = 40$, $n = 160$ et $n = 2000$:



L'expérience montre que M_n devrait tendre vers p, ce qu'illlustre le resserrement du pic.
Qu'en penser ?

? Problème A2 - Convergence (en loi)

En point d'orgue du programme de mathématique en terminale, figure le théorème de Moivre-Laplace.
Quel est sa sginification ? Et comment se démontre-t-il ?
Il s'agit d'une limite pour des variables aléatoires. Que signifie qu'une suite de variables aléatoires convergent ?

? Problème A3 - Estimateur

Quel rapport entre l'espérance mathématique et l'estimateur d'une expérience ?
Lorsqu'on répète un grand nombre de fois une même expérience, quels sont les résultats qu'on obtient ? Comment faire le lien entre

- la probabilité d'une variable aléatoire, souvent choisis comme modèle selon les symétries de l'expérience,
- son espérance mathématique
- la moyenne expérimentale des résultats obtenus, lors de la répétition un grand nombre de fois de l'expérience associée ?

2. Suite de variable aléatoire

2.1. Suite de variables aléatoires

La définition suivante est un rappel :

Définition - Indépendance pour une suite de v.a.
Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires définies sur $(\Omega, \mathbf{P})$.
$(X_n)_{n\in\mathbb{N}}$ est une suite de variables aléatoires mutuellement indépendantes si, pour tout $n \in \mathbb{N}^*$, les variables aléatoires $X_1, \dots, X_n$ sont mutuellement indépendantes.

Heuristique - Cas classique
Très souvent, n représente le temps.
On répète un certain nombre de fois une même expérience élémentaire. On note X_k, le résultat d'une certaine mesure sur l'expérience k.
On obtient ainsi une suite de variable aléatoire.
Très souvent, les expériences répétées sont indépendantes. On a alors une suite (X_k) de variable aléatoire indépendante, identiquement distribué (notée « *v.a.i.i.d.* » dans la littérature probabiliste).
C'est le cas par exemple lorsqu'on répète une même expérience en science physique. Le résultat est d'une certaine façon une variable aléatoire (même si son espérance est très précise…)

Remarque - Cas classique (2). Ω infini (dénombrable)
Il n'est pas rare que Ω soit en fait infini, même si pour tout n, X_n est définie sur Ω_n, un ensemble fini.
On aurait alors $\Omega_n = \pi_n(\Omega)$, où pour $\omega = (\omega_i)_{i\in\mathbb{N}} \in \Omega$, $\pi_n\big((\omega_i)_{i\in\mathbb{N}}\big) = (\omega_1, \omega_2, \dots \omega_n)$.
C'est hors programme de première année, mais pas de seconde année (donc au programme des concours).

2.2. Exemples de convergence de variable aléatoire

Remarque - Différents modes de convergence
Une des difficultés est qu'il existe des modes de convergence variés pour les variables aléatoires. Ils sont plus ou moins « forts » … Pour aborder ces différentes notions, nous allons considérer une famille d'exemple.

Exemple - Variable de Bernoulli répétée
Considérons une première suite (X_i) de variable aléatoire indépendante et de même loi $\mathscr{B}(p)$.
Notons alors $S_n = \sum_{k=1}^{n} X_k$. On sait que $S_n \hookrightarrow \mathscr{B}(n, p)$.
Représentons différentes valeurs de S_n (pour différentes valeurs de n). Le bon mode de représentation est l'histogramme.
Mais « trichons un peu », de manière à ce que chaque représentation soit comparable. Donc comme $S_n(\Omega) = [\![1, n]\!]$, on représente plutôt (et finalement assez naturellement) $M_n = \frac{1}{n} S_n$.

Heuristique - Convergence de variable aléatoire
Comment interpréter ce « devrait tendre » :
- Interprétation déterministe : pour toute réalisation ω du hasard, on a $M_n(\omega) \to p$.
  Ceci est clairement faux. Dans Ω, on a par exemple ω tel que $X_i(\omega) = \frac{1}{2}(1 + (-1)^i)$ (une fois sur 2 X vaut 0 ou 1).
- Interprétation forte : avec une probabilité égale à 1, la suite $M_n(\omega) \to p$.
  $$\mathbf{P}(\{\omega \mid M_n(\omega) \to p\}) = 1$$
  Il s'agit de la convergence presque sûree que nous verrons plus loin.
  Elle conduit à la loi forte des grands nombres.
- Interprétation faible : l'ensemble des moyennes qui reste écartées de p de plus de

ϵ (quelconque) tend vers 0.

$$\forall\, \epsilon > 0, \qquad \mathbf{P}(\{\omega \mid |M_n(\omega) - p| > \epsilon\}) \longrightarrow 0$$

Il s'agit de la convergence en probabilité que nous verrons plus loin.
Elle conduit à la loi faible des grands nombres.

Remarque - On notera l'interversion des ϵ (et plus) entre les convergences p.s. et en probabilité

Pour une convergence presque sûre, on a :

$$\mathbf{P}\left(\{\omega \mid \forall\, \epsilon, \exists\, N_\epsilon(\omega) \mid \forall\ n \geqslant N_\epsilon(\omega), |M_n(\omega) - p| < \epsilon\}\right) = 1$$

ou en passant au complémentaire :

$$\begin{aligned} 0 \quad &= \mathbf{P}\left(\overline{\{\omega \mid \forall\, \epsilon, \exists\, N_\epsilon(\omega) \text{ tq } \forall\ n \geqslant N_\epsilon(\omega), |M_n(\omega) - p| < \epsilon\}}\right) \\ &= \mathbf{P}\left(\{\omega \mid \exists\, \epsilon > 0 \text{ tq} \forall\ n \in \mathbb{N}, \exists\, N > n \text{ tq } |M_N(\omega) - p| > \epsilon\}\right) \end{aligned}$$

Ce qui peut s'écrire avec des suites extraites :

$$\mathbf{P}(\{\omega \mid \exists\, \epsilon > 0, \exists\, \varphi : \mathbb{N} \to \mathbb{N} \nearrow\nearrow \text{ tq} \forall\ n \in \mathbb{N}, |M_{\varphi(n)}(\omega) - p| > \epsilon\} = 0$$

Pour une convergence en probabilité, on a :

$$\forall\, \epsilon > 0, \qquad u_n(\epsilon) := \mathbf{P}(\{\omega \mid |M_n(\omega) - p| > \epsilon\} \longrightarrow 0$$

$$\forall\, \epsilon > 0, \forall\, \alpha > 0, \exists\, N_\alpha \text{ tq } \forall\ n \geqslant N_\alpha, |u_n(\epsilon)| \leqslant \alpha$$

Heuristique - Convergence de loi (de var) (» convergence en loi »)

On peut s'intéresser à plus que la limite de M_n. En effet, on peut s'intéresser aux fluctuations autour de p. L'échelle de fluctuations de M_n autour de sa valeur est de l'ordre de l'écart-type : $\sigma(M_n) = \dfrac{\sigma(X_i)}{\sqrt{n}}$.

Pour comparer ces M_n, on s'intéresse donc à sa version centrée réduite : $M_n^* = \dfrac{S_n - np}{\sqrt{n}\sqrt{p(1-p)}}$.

On ne peut pas espérer une convergence de variable, car les fluctuations restent importantes : de l'ordre de l'unité.

Mais on peut espérer une estimation de la répartition, ou encore une loi (limite) de probabilité de la variable aléatoire de répartition.

On a donc besoin d'une nouvelle notion de convergence : la convergence en loi. Le théorème clé est alors le théorème limite centrale.

3. Différents modes de convergence

On considère dans cette section des variables aléatoires réelles $X_1, X_2, \dots X_n \dots$ et X, toutes définies sur un même espace de probabilité.

3.1. convergence presque sûre

La première convergence dérive d'une convergence d'événements

Pour aller plus loin - Convergence naturelle simple

Evidement la convergence la plus naturelle, mais également trop rare pour être rencontrée est donnée par :

$$\forall\, \omega, X_n(\omega) \to X(\omega)$$

Définition - convergence presque sûre

La suite (X_n) **converge presque sûrement** vers X si

$$\mathbf{P}(\lim_{n\to\infty} X_n = X) = \mathbf{P}\left(\{\omega \mid \lim_{n\to\infty} X_n(\omega) = X(\omega)\}\right) = 1$$

On note alors $X_n \xrightarrow{ps} X$.

Dans une remarque précédente, on a vu que l'étude de la convergence par suite extraite pourrait être intéressante.

Définition - Infiniment souvent
Soit $\omega \in \Omega$, $\epsilon > 0$ et une suite Y_n de v.a.
On dit que $|Y_n(\omega)| > \epsilon$ infiniment souvent,
si il existe $\varphi : \mathbb{N} \to \mathbb{N} \nearrow\nearrow$ telle que pour tout $n \in \mathbb{N}$ $|Y_{\varphi(n)}(\omega)| > \epsilon$.
si $\{n \in \mathbb{N} \mid |Y_n(\omega)| > \epsilon\}$ n'est pas borné.
L'événement associé se note : $[|Y_n| > \epsilon \text{ i.s.}]$

Proposition - Événement équivalent
Pour tout $\omega \in \Omega$, et tout $\epsilon \in \mathbb{R}_+^*$

$$|Y_n(\omega)| > \epsilon \text{ i.s.} \iff \forall\ n \in \mathbb{N}, \exists\ p \geqslant n \text{ tel que } |Y_p(\omega)| > \epsilon$$

Ainsi :

$$\{\omega \in \Omega \mid |Y_n(\omega)| > \epsilon \text{ i.s.}\} = \bigcap_{n\in\mathbb{N}} \left(\bigcup_{p\geqslant n} \{\omega \mid |Y_p(\omega)| > \epsilon\} \right)$$

Démonstration
Il s'agit simplement de mettre en forme ensembliste les quantificateurs : $\exists \leftrightarrow \cup$ et $\forall \leftrightarrow \cap$ □

Heuristique - Réunion décroissante. Passage à la limite
Si on note A_n, l'événement $\bigcup_{p\geqslant n} \{\omega \mid |Y_p(\omega)| > \epsilon\}$.
Alors $\omega \in A_{n+1} \Rightarrow \exists\ p \geqslant n+1 (\geqslant n)$ tel que $|Y_p(\omega)| > \epsilon \Rightarrow \omega \in A_n$.
Donc $A_{n+1} \subset A_n$, i.e. la suite (A_n) est décroissante.
Donc pour tout $N \in \mathbb{N}$, $\bigcap_{n=1}^{N} A_n = A_N$. Pour passer à la limite, on utilise le théorème de convergence décroissante :

$$\mathbf{P}\left(\bigcap_{n\in\mathbb{N}} A_n\right) = \lim_{N\to+\infty} \mathbf{P}(A_N)$$

Proposition - Critère équivalent à la convergence presque sûre
La suite (X_n) **converge presque sûrement** vers X
si et seulement si $\forall\ \epsilon > 0$, $\mathbf{P}(|X_n - X| > \epsilon, \text{infiniment souvent}\} = 0$
si et seulement si $\forall\ \epsilon > 0$, $\lim_{N\to\infty} \mathbf{P}(\exists\ n \geqslant N \mid |X_n - X| > \epsilon) = 0$.

Démonstration
Avec des suites extraites, dire que $X_n(\omega)$ ne tend pas vers $X(\omega)$ c'est dire :

$$\exists\ \epsilon > 0, \exists\ \varphi : \mathbb{N} \to \mathbb{N} \nearrow\nearrow \text{ tq } |X_{\varphi(n)}(\omega) - X(\omega)| > \epsilon$$

Donc en terme d'événement (ou d'ensemble)

$$\overline{[X_n \to X]} = \bigcup_{\epsilon>0} [|X_n - X| > \epsilon, \text{infiniment souvent}]$$

Donc on a équivalence entre :
— $X_n \xrightarrow{ps} X$
— $\mathbf{P}([X_n \to X]) = 1$
— $\mathbf{P}(\overline{[X_n \to X]}) = 0$
— $\mathbf{P}\left(\bigcup_{\epsilon>0} [|X_n - X| > \epsilon, \text{infiniment souvent}]\right) = 0$

On a donc les équivalences :

$$X_n \xrightarrow{ps} X \quad \iff \mathbf{P}\left(\bigcup_{\epsilon>0} [|X_n - X| > \epsilon, \text{infiniment souvent}]\right) = 0$$

Sens direct.
On fixe $\epsilon_0 > 0$.
, $[|X_n - X| > \epsilon_0, \text{infiniment souvent}] \subset \bigcup_{\epsilon>0} [|X_n - X| > \epsilon, \text{infiniment souvent}]$,
on a donc $\forall\ \epsilon_0 > 0$, $\mathbf{P}(|X_n - X| > \epsilon_0, \text{infiniment souvent}) = 0$.
Réciproquement : si $\forall\ \epsilon > 0$, $\mathbf{P}(|X_n - X| > \epsilon, \text{infiniment souvent}) = 0$,

alors $\forall\, \epsilon_0 > 0, \mathbf{P}\left(\bigcup_{\epsilon>\epsilon_0} [|X_n - X| > \epsilon, \text{infiniment souvent}]\right) = \mathbf{P}(|X_n - X| > \epsilon_0, \text{infiniment souvent}) = 0$.

Il s'agit d'une suite décroissante d'événements : $\epsilon_1 < \epsilon 2$, $\bigcup_{\epsilon>\epsilon_2} [|X_n - X| > \epsilon, \text{infiniment souvent}] \subset \bigcup_{\epsilon>\epsilon_1} [|X_n - X| > \epsilon, \text{infiniment souvent}]$.

On peut passer à la limite dans la probabilité : $\lim_\epsilon \mathbf{P}(A_\epsilon) = \mathbf{P}(\lim_\epsilon A_\epsilon)$.

Donc $\mathbf{P}(\bigcup_{\epsilon>0} [|X_n - X| > \epsilon, \text{infiniment souvent}]) = 0$.

Enfin, si l'on souhaite préciser cet événement : Dire qu'infiniment souvent, $|X_n - X| > \epsilon$, cela signifie que pour tout N que l'on se fixe, il existe toujours $n > N$ tel que $|X_n - X| > \epsilon$.
En faisant tendre N vers l'infini, on retrouve le critère précédent. □

Savoir faire - Démontrer une convergence presque sûre

On exploite souvent le lemme de Borel-Cantelli.
En effet, il donne un argument de convergence infiniment souvent d'une probabilité

Pour aller plus loin - Réciproque ?

Est-ce que la réciproque est vraie : si $\mathbf{P}(A_n, \text{i.s.}) = 0$, *alors* $\sum_{n\geqslant n_0} \mathbf{P}(A_n)$ *converge ?*

On a une condition suffisante : la suite (A_p) *est une suite d'événements indépendants. On étudie les événements contraires*

$$\begin{aligned}\mathbf{P}\left(\bigcup_{p\geqslant N} A_p\right) &= 1 - \mathbf{P}\left(\overline{\bigcup_{p\geqslant N} A_p}\right)\\ &= 1 - \mathbf{P}\left(\bigcap_{p\geqslant N} \overline{A_p}\right)\\ &= 1 - \prod_{p\geqslant N} \mathbf{P}(\overline{A_p})\end{aligned}$$

Puis on compose par le ln, *comme on compare série et produit infini*

Théorème - Borel-Cantelli

Soit (A_n) une suite d'événements.
Si $\sum_{n\geqslant n_0} \mathbf{P}(A_n)$ converge, alors $\mathbf{P}(A_n, \text{ i.s.}) := \mathbf{P}(\{\omega \mid \omega \in A_n \text{ i.s.}\}) = 0$

Démonstration

On a vu que $[A_n \text{ i.s.}] = \bigcap_{n\in\mathbb{N}} \left(\bigcup_{p\geqslant n} A_p\right)$.
On note $B_n = \bigcup_{p\geqslant n} A_p$. Il s'agit d'une suite d'événements décroissants.
Donc $\mathbf{P}(\bigcap_{n\in\mathbb{N}} B_n) = \lim_{N\to+\infty} \mathbf{P}(B_N)$.
Or $\mathbf{P}(B_N) = \mathbf{P}\left(\bigcup_{p\geqslant N} A_p\right) \leqslant \sum_{p\geqslant N} \mathbf{P}(A_p) = R_N$.
Mais la série $\sum_{n\geqslant n_0} \mathbf{P}(A_n)$ converge, donc les restes tendent vers 0.
On trouve donc bien : $\lim_{N\to+\infty} \mathbf{P}(B_N) = 0$ et donc $\mathbf{P}(\bigcap_{n\in\mathbb{N}} B_n) = 0$ i.e. $\mathbf{P}(A_n, \text{ i.s.}) = 0$.
□

L'exercice suivant donne une application directe.

Exercice

Soit (X_i) une suite de variables aléatoires indépendantes identiquement distribuée admettant un moment d'ordre 4.
On note $S_n = \frac{1}{n}\sum_{k=1}^{n} X_k$.
En exploitant l'inégalité de Markov pour $\mathbf{P}((S_n - p)^4 > \epsilon^4)$, puis le lemme de Borel-Cantelli, montrer que S_n converge presque sûrement vers $p = \mathbf{E}(X_i)$

Correction

On applique l'inégalité de Markov à $(S_n - p)^4$ qui admet une espérance :

$$\forall\, \epsilon > 0, \qquad \mathbf{P}\big((S_n - p)^4 > \epsilon^4\big) \leqslant \frac{\mathbf{E}((S_n - p)^4)}{\epsilon^4}$$

Or $[S_n - p]^4 > \epsilon^4 \iff [S_n - p] > \epsilon$.
Puis $S_n - p = \frac{1}{n}\sum_{i=1}^{n} (X_i - p)$.
Par linéarité de l'espérance et indépendance des X_i :

$$\begin{aligned}\mathbf{E}\big((S_n - p)^4\big) &= \mathbf{E}\left(\frac{1}{n^4}[\sum_{i=1}^{n} (X_i - p)]^4\right)\\ &= \frac{1}{n^4}\mathbf{E}\left(\sum_{i=1}^{n} X_i^{*4} + \sum_{i\neq j} 4X_iX_j^3 + 6X_i^2X_j^2 + \sum_{i\neq j\neq k} 12X_iX_jX_k^2 + \sum_{i\neq j\neq k\neq h} 24X_iX_jX_hX_k\right)\\ &= \frac{1}{n^3}\mu_4(X) + \frac{n-1}{2n^3}4\mu_1(X)\mu_3(X) + \frac{n-1}{2n^3}6\mu_2^2(X)\\ &\quad + \frac{(n-1)(n-2)}{6n^2}12\mu_1^2(X)\mu_2(X) + \frac{(n-1)(n-2)(n-3)}{24n^3}24\mu_1^4(X)\\ &= \frac{1}{n^3}\mu_4(X) + \frac{3(n-1)}{n^3}\mu_2^2(X)\end{aligned}$$

où $X_i^* = X_i - p$, donc $\mathbf{E}(X_i^*) = \mu_1(X)0$.

Donc, pour tout $\epsilon > 0$, $\mathbf{P}(|S_n - p| > \epsilon) \leqslant \frac{1}{\epsilon^4}\left(\frac{1}{n^3}\mu_4(X) + \frac{3(n-1)}{n^3}\mu_2^2(X)\right)$.

Donc la série $\sum_{n\geqslant 1} \mathbf{P}(|S_n - p| > \epsilon)$ converge (comparaison à une série de Riemann).

On applique le Lemme de Borel-Cantelli à $A_n = [|S_n - p| > \epsilon]$.

Ainsi, $\mathbf{P}([S_n - p > \epsilon \text{ i.s.})] = 0$. Ceci est vrai pour tout $\epsilon > 0$

3.2. Convergence en probabilité

Définition - Convergence en probabilité

La suite (X_n) **converge en probabilité** vers X si

$$\forall\, \epsilon > 0, \qquad \lim_{n\to+\infty} \mathbf{P}(|X_n - X| \geqslant \epsilon) = 0$$

On note alors $X_n \xrightarrow{\mathbf{P}} X$.

Savoir faire - Démontrer la convergence en probabilité

Très fréquemment, on emploie les inégalités de Markov ou de Bienaymé-Tchebychev pour démontrer qu'une suite de variable aléatoire converge en probabilité vers une constante.

Proposition - Stabilité

Si $X_n \xrightarrow{\mathbf{P}} X$ et $Y_n \xrightarrow{\mathbf{P}} Y$, respectivement $X_n \xrightarrow{ps} X$ et $Y_n \xrightarrow{ps} Y$.

Alors pour tout $\alpha, \beta \in \mathbb{R}$, $\alpha X_n + \beta Y_n \xrightarrow{\mathbf{P}} \alpha X + \beta Y$ resp. : $\xrightarrow{ps} \alpha X + \beta Y$.

Si $g : \mathbb{R} \to \mathbb{R}$ est continue alors $g(X_n) \xrightarrow{\mathbf{P}} g(X)$ resp. : $g(X_n) \xrightarrow{ps} g(X)$

Pour aller plus loin - Lien entre convergence

Si (X_n) converge presque sûrement vers X, alors (X_n) converge en probabilité vers X.

Si (X_n) converge en probabilité vers X, alors il existe $\varphi \nearrow\nearrow$, $(X_{\varphi(n)})$ converge presque sûrement vers X.

La linéarité est assez simple à démontrer.

La composition par une fonction continue est plus compliquée… On l'admet. (On pourrait utiliser l'image réciproque de tout ouvert est un ouvert…).

3.3. Lien entre ces deux convergences de variables aléatoires

Commençons par analyser un exemple.

Analyse - Une suite de variable aléatoire qui converge en probabilité mais pas presque sûrement.

On note $T_n = \frac{n(n+1)}{2} = \sum_{k=1}^{n} k$.

Pour tout $k \in \mathbb{N}$, $\exists\, n(k) \in \mathbb{N}$ tel que $T_{n(k)-1} < k \leqslant T_n(k)$.

Considérons alors $X_k = \mathbb{1}_{\left[\frac{k-T_{n(k)-1}-1}{n(k)}, \frac{k-T_{n(k)-1}}{n(k)}\right]}$, indicatrice Pour bien comprendre ce que signifie cette suite, on peut résumer sa construction de la façon suivante :

1. On coupe l'intervalle $[0,1]$ en un morceau et on considère $X_1 = \mathbb{1}_{[0,1]}$
2. On coupe l'intervalle $[0,1]$ en deux morceaux et on considère $X_2 = \mathbb{1}_{\left[0,\frac{1}{2}\right]}$ et $X_3 = \mathbb{1}_{\left[\frac{1}{2},1\right]}$
3. On coupe l'intervalle $[0,1]$ en trois morceaux et on considère $X_4 = \mathbb{1}_{\left[0,\frac{1}{3}\right]}$, $X_5 = \mathbb{1}_{\left[\frac{1}{3},\frac{2}{3}\right]}$ et $X_6 = \mathbb{1}_{\left[\frac{2}{3},1\right]}$
4. …
n On coupe l'intervalle $[0,1]$ en n morceaux et on considère $X_{T_{n-1}+1} = \mathbb{1}_{\left[0,\frac{1}{n}\right]}$, $X_{T_{n-1}+2} = \mathbb{1}_{\left[\frac{1}{n},\frac{2}{n}\right]}$,… et $X_{T_n} = \mathbb{1}_{\left[\frac{n-1}{n},1\right]}$

Alors pour la plupart des ω, $X_h(\omega)=0$, mais régulièrement, $X_h(\omega)=1$ (une et une seule fois entre chaque T_n et T_{n+1}).
Ainsi on peut considérer $X=0$. Soit $\epsilon>0$ et $\epsilon<1$.

$$[|X_k-X|>\epsilon]=\{\omega\in\Omega\mid X_k(\omega)=1\} \qquad \mathbf{P}(|X_k-X|>\epsilon)=\mathbf{E}\left(\mathbb{1}_{\left[\frac{k-T_{n(k)-1}-1}{n(k)},\frac{k-T_{n(k)-1}}{n(k)}\right]}\right)=\frac{1}{n(k)}\underset{n\to\infty}{\longrightarrow}0$$

Donc X_n converge en probabilité vers $X=0$.
En revanche, pour tout $\omega\in\Omega$, il existe une suite (v_n) strictement croissante (infinie) tel que $X_{v_n}(\omega)=1$.
Donc pour tout $\epsilon>0$ (et inférieur à 1), on n'a pas $\lim\limits_{N\to\infty}\mathbf{P}([\exists\, n\geqslant N\mid |X_n-X|>\epsilon)=0$.
Ainsi X_n ne tend pas presque sûrement vers $X=0$

Proposition - Implication des convergences
Si $(X_n)\xrightarrow{ps}X$, alors $(X_n)\xrightarrow{\mathbf{P}}X$.

Attention - Réciproque fausse
La réciproque est fausse comme le montre le contre-exemple vu en analyse

Démonstration
Notons $Y_n=|X_n-X|$. Soit $\epsilon>0$ et pour tout $N\in\mathbb{N}$, $B_N(\epsilon)$ l'événement « il existe $n\geqslant N$ tel que $Y_n\geqslant\epsilon$ » La convergence presque sûre de (Y_n) vers 0 montre que $\lim\limits_{N\to+\infty}\mathbf{P}(B_N(\epsilon))=0$.
Or pour tout entier n, $[Y_n>\epsilon]\subset B_n(\epsilon)$ ce qui prouve que par majoration que $\mathbf{P}([Y_n>\epsilon])\to 0$ □

Toutefois, il y a presqu'équivalence ;

Proposition - Condition suffisante pour l'implication réciproque
Si $(X_n)\xrightarrow{\mathbf{P}}X$,
alors on peut extraire de X_n une sous-suite convergent presque sûrement vers X. Autrement écrit :

$$\exists\,\varphi\nearrow\nearrow\mid (X_{\varphi(n)})\xrightarrow{ps}X$$

On ne fait pas la démonstration.

3.4. Convergence en loi

Comme son nom l'indique, la convergence en loi indique que la suite des lois de X_n converge vers la loi de X.
L'objet n'est plus directement la variable aléatoire…

Définition - Convergence en loi (variable discrète)
On note $E=\bigcup_{n\in\mathbb{N}^*}X_n(\Omega)\cup X(\Omega)$.
La suite (X_n) **converge en loi** vers X si

$$\forall\, x\in E,\qquad \lim_{n\to+\infty}\mathbf{P}(X_n=x)=\mathbf{P}(X=x)$$

Si $x\notin X_n(\Omega)$, on note $\mathbf{P}(X_n=x)=0$ et de même si $x\notin X(\Omega)$, on note $\mathbf{P}(X=x)=0$
On note alors $X_n\xrightarrow{\mathscr{L}}X$.

convergence
as tout à fait
convergences

Remarque - Cas des variables continues

Lorsque $X(\Omega)$ n'est pas dénombrable, mais un ensemble compact dans $\mathbb{R}$ (comme en terminale), on ne calcule pas $\mathbf{P}(X=x)$, celle-ci est toujours nul.
Le calcul qui joue le rôle équivalent est donnée par la fonction de répartition : $\mathbf{P}(X \leqslant x) = \int_{-\infty}^{x} f(t)\mathrm{d}t$.
Ainsi le calcul d'espérance de X est $\mathbf{E}(X) = \int_{-\infty}^{+\infty} t f(t)\mathrm{d}t$ au lieu de $\sum_{k\in\mathbb{N}} k\mathbf{P}(X=k)$.
Et de même pour la définition de la convergence en loi :

Définition - Convergence en loi (variable continue)

On note $E = \bigcup_{n\in\mathbb{N}^*} X_n(\Omega) \cup X(\Omega)$.
La suite (X_n) **converge en loi** vers X si

$$\forall\, x \in E \text{ (non discontinuité)}, \quad \lim_{n\to+\infty} \mathbf{P}(X_n \leqslant x) = \mathbf{P}(X \leqslant x)$$

Il s'agit de la convergence simple des fonctions de répartition. On note alors $X_n \overset{\mathscr{L}}{\Longrightarrow} X$.

Attention - Pas de structure algébrique

Considérons X et Y indépendantes qui suivent la même loi binomiale $\mathscr{B}(n,p)$.
Considérons pour tout $n \in \mathbb{N}$, $X_n = X$ et $Y_n = n - X$.
Alors $X_n \hookrightarrow \mathscr{B}(n,p)$ et $Y_n \hookrightarrow \mathscr{B}(n,p)$.
Donc la convergence en loi de $\lim_{\mathscr{L}}(X_n) + \lim_{\mathscr{L}}(Y_n) \hookrightarrow \mathscr{B}(2n,p)$.
Alors que $X_n + Y_n = n$ alors donc $\lim_{\mathscr{L}}(X_n + Y_n) \hookrightarrow n$.

Proposition - Stabilité

Si $g : \mathbb{R} \to \mathbb{R}$ est continue alors $g(X_n) \overset{\mathscr{L}}{\Longrightarrow} g(X)$

Proposition - Convergence en probabilité implique la convergence en loi

Supposons que la suite (X_n) converge en probabilité vers X.
Alors (X_n) converge en loi vers X

On fait la démonstration dans le cas discret.

Démonstration

Soit (X_n) qui converge en probabilité vers X.
Soit $\epsilon > 0$. Soit $x \in X(\Omega)$,

$$[X_n = x] = [X_n = x, X < x-\epsilon] \cup \big[X_n = x, X \in [x-\epsilon, x+\epsilon]\big] \cup [X_n = x, X > x+\epsilon]$$

$$\mathbf{P}(X_n = x) = \mathbf{P}(X_n = x, X_n - X > \epsilon) + \mathbf{P}(X_n = x, X \in [x-\epsilon, x+\epsilon]) + \mathbf{P}(X_n = x, X - X_n > \epsilon)$$

$$\mathbf{P}(X_n = x) \leqslant \mathbf{P}(X \in [x-\epsilon, x+\epsilon]) + \mathbf{P}(|X - X_n| > \epsilon)$$

Comme $X(\Omega)$ est fini, l'ensemble $\{|x_i - x_j|\}$ est fini, il existe donc η tel que $\forall\, i,j, |x_i - x_j| > \eta$.
Donc avec $\epsilon < \eta$:

$$\mathbf{P}(X_n = x) \leqslant \mathbf{P}(X = x) + \mathbf{P}(|X - X_n| > \epsilon)$$

De même, on trouve :

$$\mathbf{P}(X = x) \leqslant \mathbf{P}(X_n = x) + \mathbf{P}(|X_n - X| > \epsilon)$$

Donc

$$|\mathbf{P}(X_n = x) - \mathbf{P}(X = x)| \leqslant \mathbf{P}(|X - X_n| > \epsilon)$$

On achève la démonstration en exploitant la convergence en probabilité. □

Un exemple très classique :

Proposition - Approximation poissonnienne
Notons X_n une variable aléatoire qui suit une loi binomiale de paramètre $(n, \frac{\lambda}{n})$.
Alors $X_n \overset{\mathscr{L}}{\Longrightarrow} X$, où $X \hookrightarrow \mathscr{P}(\lambda)$ la loi de Poisson de paramètre λ, c'est-à-dire :

$$X(\Omega) = \mathbb{N} \qquad \mathbf{P}(X = k) = e^{-\lambda}\frac{\lambda^k}{k!}$$

Démonstration
Soit $k \in \mathbb{N}$. Pour $n \geqslant k$, $k \in X_n(\Omega)$.

$$\begin{aligned}\mathbf{P}(X_n = k) &= \frac{n!}{k!(n-k)!}\frac{\lambda^k}{n^k}\left(1-\frac{\lambda}{n}\right)^{n-k} = \frac{\lambda^k}{k!}\frac{n\times(n-1)\dots(n-k+1)}{n\times n\times\dots n}\left(1-\frac{\lambda}{n}\right)^{n-k}\\ &= \frac{\lambda^k}{k!}\prod_{h=0}^{k-1}\left(1-\frac{h}{n}\right)\left(1-\frac{\lambda}{n}\right)^{n-k} = \frac{\lambda^k}{k!}\left(1-\frac{\lambda}{n}\right)^{n}\prod_{h=0}^{k-1}\left((1-\frac{h}{n})(1-\frac{\lambda}{n})^{-1}\right)\end{aligned}$$

Or k est fixé, le produit est fini et pour $n \to +\infty$: $\prod_{h=0}^{k-1}\left((1-\frac{h}{n})(1-\frac{\lambda}{n})\right)^{-1} \to 1$.
Donc $\mathbf{P}(X_n = k) \underset{n\to+\infty}{\longrightarrow} \frac{\lambda^k}{k!}e^{-\lambda}$. □

4. Loi (faible) des grands nombres et estimateurs

4.1. Lois des grands nombres

Théorème - Loi faible des grands nombres
Soit (X_n) une suite de variables aléatoires indépendantes de même espérance m et variance σ^2.
Alors la suite $M_n = \frac{1}{n}\sum_{k=1}^{n} X_k$ converge en probabilité vers la variable constante égale à m

Exercice
Ecrire ce résultat avec des ϵ

Correction
$\forall\, \epsilon > 0, \lim_{n\to+\infty} \mathbf{P}(|M_n - m| > \epsilon) = 0$

Démonstration
On exploite l'inégalité de Bienaymé-Tchebychev : car par linéarité de l'espérance $\mathbf{E}(M_n) = \frac{1}{n}\sum_{k=1}^{n}\mathbf{E}(X_k) = m$

$$\mathbf{P}(|M_n - m| > \epsilon) = \mathbf{P}(|M_n - \mathbf{E}(M_n)| > \epsilon) \leqslant \frac{\mathbf{V}(M_n)}{\epsilon^2}$$

Or $\mathbf{V}(M_n) = \frac{1}{n^2}n\sigma^2$, par indépendance des X_i.

$$\mathbf{P}(|M_n - m| > \epsilon) \leqslant \frac{\sigma^2}{n\epsilon^2} \to 0$$

□

Histoire - La loi des grands nombres et l'histoire des probabilités
On pourrait faire un cours de probabilité en suivant le fil de l'histoire et comment les mathématiciens ont, depuis Jacques Bernoulli, eu comme unique mission d'alléger les hypothèses de ce théorème ou de le rendre plus fort

La loi forte des grands nombres existe (Kolmogorov) avec quasiment aucune hypothèse supplémentaire (mais avec une démonstration plus costaud !).

4.2. Estimateurs

Heuristique - Faire la moyenne
Ce théorème est essentiel. Il affirme que si on fait une moyenne des résultats obtenus en répétant une même expérience aléatoire (des mesures après une répétition d'une même expérience), alors celle-ci va converger vers l'espérance de la loi.
Une autre application est la suivante. On réalise un très grand nombre de simulations informatiques (avec Python), on fait la moyenne des résultats, et on obtient un résultat

, mathématique précis. C'est la philosophie des méthodes dites de Monte-Carlo

Voilà enfin démontrer la raison pour laquelle nous calculons les moyennes des élèves… Mais il peut y avoir d'autres estimateurs.

Heuristique - Contexte

On considère un phénomène aléatoire et on s'intéresse à une variable aléatoire réelle X qui lui est liée, dont on suppose que la loi de probabilité n'est pas complètement spécifiée et appartient à une famille de lois dépendant d'un paramètre θ décrivant un sous-ensemble $\Theta \in \mathbb{R}$. Le paramètre θ est une quantité inconnue, fixée dans toute l'étude, que l'on cherche à déterminer ou pour laquelle on cherche une information partielle.
Le problème de l'estimation consiste alors à estimer la vraie valeur du paramètre θ (ou de $g(\theta)$ -fonction à valeurs réelles du paramètre θ), à partir d'un échantillon de données $x_1,\dots,x_n$ obtenues en observant n fois le phénomène. Cette fonction du paramètre représentera en général une valeur caractéristique de la loi inconnue comme son espérance, sa variance, son étendue...

Définition - Estimateur de $g(\theta)$

Un estimateur de $g(\theta)$ est une variable aléatoire de la forme $T_n = \Phi(X_1,\dots,X_n)$.
La réalisation $\Phi(x_1,\dots,x_n)$ de l'estimateur T_n est l'estimation de $g(\theta)$. Cette estimation ne dépend que de l'échantillon $(x_1,x_2,\dots,x_n)$ observé.

Exemple - Moyenne

Si $X_1,\dots X_n$ sont des variables aléatoires indépendantes de même espérance m.
Alors la moyenne de $X_1,\dots X_n$ est un estimateur (empirique) de m.

Définition - Biais et estimateur sans biais

Si pour tout $\theta \in \Theta$, l'estimateur T_n admet une espérance, on appelle **biais de T_n en** $g(\theta)$ le réel

$$b_\theta(T_n) = \mathbf{E}_\theta(T_n) - g(\theta)$$

L'estimateur T_n de $g(\theta)$ est dit estimateur sans biais si pour tout $\theta \in \Theta$, $\mathbf{E}_\theta(T_n) = g(\theta)$.

Définition - Estimateur convergent

Une suite d'estimateurs $(T_n)_{n\geqslant 1}$ de $g(\theta)$ est convergente si pour tout θ, la suite $(T_n)_{n\geqslant 1}$ converge en probabilité vers $g(\theta)$.
Par abus de langage on dit rapidement que l'estimateur est convergent.

Exercice

On considère une suite (X_i) de variables aléatoires indépendantes qui suivent toute la même loi de Benoulli de paramètre p.
Et l'on cherche un estimateur de la variance V de cette loi.

1. On considère $M_n = \dfrac{1}{n}\sum_{k=1}^{n} X_k$.
   Montrer que M_n est un estimateur sans biais de p.
   Est-il convergent ?
2. On considère $V_n = \dfrac{1}{n}\sum_{k=1}^{n} (X_k - M_n)^2$.
   V_n est-il un estimateur sans biais de V ? Sinon calculer biais.
   Est-il convergent ?
3. On considère $W_n = \dfrac{1}{n-1}\sum_{k=1}^{n} (X_k - M_n)^2$.
   W_n est-il un estimateur sans biais de V ? Sinon calculer le biais.

Correction

1. $\mathbf{E}(M_n) = \frac{1}{n}\sum_{k=1}^{n}\mathbf{E}(X_k) = m$.
   Donc M_n est un estimateur sans biais de p. D'après la loi faible des grands nombres, il est convergent (en probabilité) vers p.
2. On considère $V_n = \frac{1}{n}\sum_{k=1}^{n}(X_k - M_n)^2$.
   $$\mathbf{E}(V_n) = \frac{1}{n}\sum_{k=1}^{n}\mathbf{E}((X_k - M_n)^2)$$
   Or
   $$(X_k - M_n)^2 = X_k^2 + M_n^2 - 2X_k M_n = \left(1 - \frac{2}{n}\right)X_k^2 + M_n^2 - \frac{2}{n}\sum_{i\neq k} X_k X_i$$
   Donc
   $$\mathbf{E}(V_n) = \frac{1}{n}\sum_{k=1}^{n}\left[\left(1-\frac{2}{n}\right)\left(\mathbf{V}(X_k) + [\mathbf{E}(X_k)]^2\right) + \left(\mathbf{V}(M_n) + [\mathbf{E}(M_n)]^2\right) - \frac{2}{n}\sum_{i\neq k}\mathbf{E}(X_k)\mathbf{E}(X_i)\right]$$
   $$= \left[\frac{n-2}{n}\left(pq+p^2\right) + \left(\frac{1}{n}pq + p^2\right) - \frac{2(n-1)}{n}p^2\right] = \frac{1}{n}\left((n-2)p + pq + np^2 - 2(n-1)p^2\right)$$
   $$= \frac{1}{n}\left((n-1)p - (n-1)p^2\right) = \frac{n-1}{n}pq = \frac{n-1}{n}V$$
   Donc V_n est un estimateur avec biais égal à $\mathbf{E}(V_n) - V = \frac{-1}{n}V$.
   $\forall\, \epsilon > 0$
   $$|V_n - V| > \epsilon = |V_n - \frac{n-1}{n}V + \frac{1}{n}V| > \epsilon \Longrightarrow |V_n - \mathbf{E}(V_n)| + \frac{1}{n}V > \epsilon$$
   $$\mathbf{P}(|V_n - V) > \epsilon) \leqslant \mathbf{P}(|V_n - \mathbf{E}(V_n)| > \epsilon - \frac{1}{n}V) \leqslant \frac{\mathbf{V}(V_n)}{(\epsilon - \frac{1}{n}V)^2}$$
   d'après l'inégalité de Bienaymé-Tchebychev : Il reste à calculer la variance de V_n, ce n'est pas aisé...
3. L'espérance étant linéaire, comme $W_n = \frac{n}{n-1}V_n$, on a donc $\mathbf{E}(W_n) = \frac{n}{n-1}\mathbf{E}(V_n) = V$.
   Donc W_n est-il un estimateur sans biais de V.

5. Théorème limite central

5.1. Rappels sur la loi normale

On a besoin de quelques rappels de terminale sur la loi normale : Lorsque $X(\Omega)$ est un intervalle de $\mathbb{R}$, nous devons penser autrement les lois de probabilité. On exploite les fonctions de répartition et non les lois ; la notion de densité devient alors importante.

Définition - Variable aléatoire à densité
Une variable aléatoire continue X est définie à partir d'une densité f vérifiant :
- $f : \mathbb{R} \to \mathbb{R}^+$
- f est continue par morceaux sur $\mathbb{R}$
- $\lim\limits_{x\to-\infty, y\to+\infty}\int_x^y f(t)\mathrm{d}t = 1$. Ce nombre est notée $\int_{-\infty}^{+\infty} f(t)\mathrm{d}t$.

On a alors

$$\mathbf{P}(X \leqslant x) = \int_{-\infty}^{x} f(t)\mathrm{d}t \qquad \mathbf{P}(X \in [a,b]) = \int_a^b f(t)\mathrm{d}t$$

Par ailleurs, si « les intégrales suivantes sont convergentes » (i.e. les limites existent) :

$$\mathbf{E}(X) = \int_{-\infty}^{+\infty} tf(t)\mathrm{d}t \qquad \mathbf{E}(\varphi(X)) = \int_{-\infty}^{+\infty}\varphi(t)f(t)\mathrm{d}t \qquad \mathbf{V}(X) = \mathbf{E}(X^2) - \mathbf{E}(X)^2$$

Exemple - Loi exponentielle
cf. Le cours de terminale
Un exemple classique de loi à densité :

Définition - Loi normale
Pour tout $\mu \in \mathbb{R}$, $\sigma^2 > 0$, la fonction

$$f_{\mu,\sigma^2} : t \mapsto \frac{1}{\sqrt{2\pi}} \exp(-\frac{(t-\mu)^2}{2\sigma^2}$$

est une fonction de densité.
On dit qu'une variable aléatoire X qui admet pour densité f_{μ,σ^2} suit une loi normale (ou gaussienne ou Laplace-Gauss).
On note $X \hookrightarrow \mathcal{N}(\mu, \sigma^2)$.
On a alors $\mathbf{E}(X) = \mu$ et $\mathbf{V}(X) = \sigma^2$

Remarque - Convergence en loi pour les variables à densité
Nous venons de l'écrire : pour des variables à densité, on exploite plus les lois mais les fonctions de répartitions.
La convergence en loi s'écrit donc autrement dans ce cas là

Définition - Convergence en loi
On suppose $X(\Omega$ et $X_n(\Omega) \subset \mathbb{R}$.
La suite (X_n) **converge en loi** vers X (à densité) si

$$\forall\, x \in \mathbb{R}, \qquad \lim_{n\to+\infty} \mathbf{P}(X_n \leqslant x) = \mathbf{P}(X \leqslant x)$$

5.2. Enoncé

Il est très général, il est plus large que la loi des grands nombres mais ne donne une convergence « qu'en loi »

Proposition - Théorème central limite
Soit $(X_n)_{n\in\mathbb{N}^*}$ est une suite de variables aléatoires indépendantes et de même loi, admettant une espérance m et une variance σ^2 non nulle.
Notons $\overline{X_n} = \dfrac{X_1 + \cdots + X_n}{n}$ et $\overline{X_n}^* = \sqrt{n}\left(\dfrac{\overline{X_n} - m}{\sigma}\right)$, variable aléatoire centrée et réduite.
Alors $(\overline{X_n}^*)$ converge en loi vers une variable aléatoire suivant la loi normale centrée réduite.

$$X_n^* \overset{\mathscr{L}}{\Longrightarrow} \mathcal{N}(0;1)$$

Nous n'en ferons pas de démonstration. Il nous manque quelques outils…
Insistons : ce résultat est indépendant de la loi suivie par X_n !
On peut par contre l'appliquer afin d'obtenir :

Théorème - Théorème de De Moivre-Laplace
Soit $p \in]0,1[$ et S_n une suite de variables aléatoires telles que $S_n \hookrightarrow \mathscr{B}(n,p)$.
Alors pour tout réel x,

$$\mathbf{P}\left(\frac{S_n - np}{\sqrt{np(1-p)}} \leqslant x\right) \underset{n\to\infty}{\longrightarrow} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}}\, \mathrm{d}t$$

Démonstration
S_n est une somme de Bernoulli indépendante. On applique le théorème limite central. Notons $Z_n = \dfrac{S_n - np}{\sqrt{np(1-p)}}$.
Par linéarité $\mathbf{E}(Z_n) = \dfrac{1}{\sqrt{np(1-p)}}(\mathbf{E}(S_n) - np) = 0$ et $\mathbf{V}(Z_n) = \dfrac{1}{np(1-p)}\mathbf{V}(S_n) = 1$ Donc Z_n est une suite de va centrée réduite. Donc

$$Z_n \xrightarrow{\mathscr{L}} \mathcal{N}(0;1)$$

Pour des variables à densité, la convergence en loi ne s'exprime plus avec des lois de probabilités mais avec des densités.
□

5.3. Intervalle de confiance

Heuristique - Fluctuations autour d'une valeur estimée
L'estimateur théorique, basée ensuite sur des mesures, donne une estimation d'un paramètre $g(\theta)$. Mais parfois, on souhaite plutôt une connaissance des fluctuations autour de la valeur limite. La convergence en loi joue un peu ce rôle comparée à la convergence en probabilité.
On peut donc espérer élargir la notion d'estimateur à l'aide du théorème limite central ou le théorème de De Moivre-Laplace si l'on se concentre sur des variables de Bernoulli.
C'est la notion d'intervalle de confiance qui joue se rôle

On notera que l'intervalle de confiance est **une variable aléatoire**, c'est ce qui le différencie de l'intervalle de fluctuation (aux bornes fixées).
Le principe est le suivant :

Définition - Intervalle de confiance
Soit (X_n) une suite de variables aléatoires indépendants de même loi admettant un moment d'ordre 2. Notons $m = \mathbf{E}(X_1)$ et $\sigma^2 = \mathbf{V}(X_1)$.
$M_n = \dfrac{1}{n}\sum_{k=1}^{n} X_k$ est un estimateur de m.
Soit $\alpha > 0$, un risque. L'intervalle de confiance au risque α est défini comme

$$I_n(\omega) = \left[M_n - \frac{\epsilon\sigma}{\sqrt{n}}, M_n + \frac{\epsilon\sigma}{\sqrt{n}}\right]$$

où ϵ est définie par $\mathbf{P}(|N| > \epsilon) = \alpha$, avec $N \hookrightarrow \mathcal{N}(0;1)$

Remarque - ϵ ?
$\mathbf{P}(|N| > \epsilon = 1 - \displaystyle\int_{-\epsilon}^{\epsilon} n(t)\mathrm{d}t = 1 - 2\int_0^{\epsilon} n(t)\mathrm{d}t$, car n est paire.
Donc $\mathbf{P}(|N| > \epsilon) = \alpha \iff N(\epsilon) = \dfrac{1-\alpha}{2} \iff \epsilon = N^{-1}\left(\dfrac{1-\alpha}{2}\right)$
où N est la primitive de n, strictement croissante (car $n > 0$) et continue donc bijective

Proposition - Encadrement de m
Avec les hypothèses de la définition, on a

$$\mathbf{P}(m \in I_n) \underset{n\to\infty}{\longrightarrow} \int_{-\epsilon}^{+\epsilon} n(t)\mathrm{d}t = 1 - \alpha$$

Démonstration
On a applique le TLC (ou Moivre-Laplace dans le cas de Bernoulli).
On a donc, par convergence en loi :

$$\mathbf{P}(m \in I_n) = \mathbf{P}\left(m \in \left[M_n - \frac{\epsilon\sigma}{\sqrt{n}}, M_n + \frac{\epsilon\sigma}{\sqrt{n}}\right]\right) \underset{n\to\infty}{\longrightarrow} \int_{-\epsilon}^{+\epsilon} n(t)\mathrm{d}t = 1 - \alpha$$

□

Pour des exercices, voir le cours de terminale.

6. Bilan

Synthèse

⇝ La convergence simple des suites de fonctions (X_n) n'est pas satisfaisante (trop exigeante) pour l'étude des suites de variables aléatoires. Il suffit de se contenter de convergence sur des parties de Ω de mesure égale à 1, donc des convergence presque sûree. Voire une convergence en probabilité. Ces convergences ne sont pas sans lien.
On associe à ces convergences les convergentes fortes (p.s.) ou faibles (en **P**). Par exemple les lois des grands nombres ou la convergence (faible) d'estimateurs..

⇝ On peut aussi s'intéresser aux lois des variables aléatoire et voir si ces fonctions convergent. Cela donne une convergence en loi, en réalité indépendante des variables d'une certaine façon. Une application classique est le théorème central limite dont une des conséquences est le théorème de Moivre-Laplace vu (trop rapidement) en terminale.

Savoir-faire et Truc & Astuce du chapitre

— Savoir-faire - Démontrer une convergence presque sûre
— Savoir-faire - Démontrer une convergence en probabilité

Notations

Notations	*Définitions*	*Propriétés*	*Remarques*
$i.s.$	ω fixé. On a la propriété $\mathscr{P}(X_n(\omega))$ infiniment souvent si $\{n \in \mathbb{N} \mid \mathscr{P}(X_n(\omega))\}$ n'est pas borné	Equivalent à $\exists\, \varphi : \mathbb{N} \to \mathbb{N}$ ↗↗ tel $\forall\, n \in \mathbb{N}$, $\mathscr{P}(X_{\varphi(n)}(\omega))$	$\{\omega \mid \mathscr{P}(X_n(\omega)) \ i.s.\} = \bigcap_{n\in\mathbb{N}} \bigcup_{p\geqslant n} \{\omega \mid \mathscr{P}(X_n(\omega))\}$.
$(X_n) \xrightarrow{p.s.} X$	La suite de v.a. (X_n) converge presque sûrement vers X	$\mathbf{P}(\lim_{n\to+\infty} X_n = X) = 1$	
$(X_n) \xrightarrow{\mathbf{P}} X$	La suite de v.a. (X_n) converge en probabilité vers X	$\forall\, \epsilon > 0$, $\lim_{n\to+\infty} \mathbf{P}(\lvert X_n - X\rvert \geqslant \epsilon) = 0$	$(X_n) \xrightarrow{p.s.} X \Rightarrow (X_n) \xrightarrow{\mathbf{P}} X$. On exploite souvent l'inégalité de B.T.
$(X_n) \xRightarrow{\mathscr{L}.} X$	La suite de v.a. (X_n) converge en loi vers X	(Cas au plus dénombrable) : $\forall\, x \in E$, $\lim_{n\to+\infty} \mathbf{P}(X_n = x) = \mathbf{P}(X = x)$	Utilisation pour l'approximation poissonnienne ou le TCL.

Retour sur les problèmes

A.18 Ici on a une convergence faible, comme un estimateur. C'est (au moins) la loi faible des grands nombres. Mais la loi forte peut également s'appliquer.

A.19 Gros morceau du cours hors-programme. Si les démonstrations ont échappées au lecteur, on espère au moins que les notions ont été comprises en profondeur!

A.20 Même réponse qu'à la question précédente.

Annexe B

Fonctions holomorphes

Résumé -

Dans ce complément hors-programme de CPGE, nous étudions les fonctions dérivables de la variable complexe. On qualifie ces fonctions d'holomorphes (« de forme entière »). Nous verrons d'abord que ces fonctions sont nombreuses (toutes nos fonctions usuelles sont holomorphes sur un ouvert plus ou moins grand) et qu'elles possèdent une certaine rigidité (conformes), cela leur donne des propriétés nouvelles : elles sont nécessairement analytiques et donc de classe $\mathscr{C}^\infty$, elles vérifient le principe harmonique...

Pour obtenir ces relations, nous exploiterons l'intégration sur un chemin complexe. Nous devons faire un détour par les intégrales curvilignes (et en bonus, le théorème de Green-Riemann). Nous serons armés avec les formules intégrales de Cauchy pour faire les démonstrations, dont le prolongement analytique (en gros : deux fonctions holomorphes sur un (petit) ouvert sont égales sur toute la composante connexe) et une méthode superpuissante pour calculer des intégrales immondes : le théorème de résidus.

Un chaîne youtube d'analyse complexe - https ://www.youtube.com/playlist ?list=PLErC88eFpes2gkbeh-RFAJqoQ5vYrwUBz

Sommaire

1. Problèmes

? Problème A4 - Division par un nombre complexe et passage à la limite

Les fonctions de plusieurs ne sont pas dérivables car la division par un vecteur n'a pas de sens : $\dfrac{f(\vec{x}+\vec{h})-f(\vec{x})}{\vec{h}}$ ne signifie rien...
En revanche, la division par un nombre complexe a bien une signification. Que dire des fonctions $\mathbb{C}$-dérivable ?
Et si on note $F(x,y)=f(x+iy)$, que signifie pour $\dfrac{\partial F}{\partial x}$ et $\dfrac{\partial F}{\partial y}$ le fait que $\dfrac{\partial f}{\partial z}$ existe ?

? Problème A5 - Intégration complexe

Si la dérivation complexe semble avoir un sens facilement accessible : $\lim\limits_{|a|\to 0}\dfrac{f(z+a)-f(z)}{a}$, qu'en est-il de son opération réciproque.
Que signifie $\displaystyle\int_\Gamma f(z)\mathrm{d}z$ pour $f:\mathbb{U}\subset C\to\mathbb{C}$ et Γ, un chemin dans le plan complexe ou encore $\displaystyle\iint_K f(x+iy)\mathrm{d}x\mathrm{d}y$ pour K une partie de $\mathbb{C}$?

? Problème A6 - Théorème de D'Alembert-Gauss

Le corps $\mathbb{C}$ a été crée comme l'extension de $\mathbb{R}$ afin que tous les polynômes à coefficients entiers ou réels de degré n aient exactement n racines (certaines pouvant être multiples).
Comment démontrer ce résultat ? Les fonctions holomorphes aident-elles pour faire cette démonstration ?

? Problème A7 - Extension de ln

Pourquoi la fonction exp s'étend sans difficulté sur $\mathbb{C}$: $\exp(x+iy)=e^x(\cos y+i\sin y)$?
Et que se passe-t-il pour sa fonction réciproque $\ln z$? Est-elle bien définie, ou pourquoi ne l'est-elle pas ?
Et à quoi « ressemble » la représentation graphique associée à ln ?

? Problème A8 - Fonction Γ

La fonction $\Gamma : x\mapsto\displaystyle\int_0^{+\infty}t^{x-1}e^{-t}\mathrm{d}t$ est central en mathématiques. En particulier, elle interpole la factorielle (précisément : $\Gamma(n+1)=n!$), elle semble définie sur $\mathbb{R}_+^*$ entier où elle est de classe $\mathscr{C}^\infty$.
Peut-elle être prolongée sur $\mathbb{C}$, de manière unique ?

Même question pour $\zeta : s \in \mathbb{R}$ ou $\mathbb{C} \mapsto \sum_{n=1}^{+\infty} \frac{1}{n^s}$?

2. Holomorphie = Dérivation complexe

2.1. Fonctions holomorphes

Remarque - Division par un nombre complexe : possible

Pour les fonctions de $\mathbb{R}^n \to \mathbb{R}^p$, il n'est pas possible d'étudier la limite $\frac{f(u+h)-f(u)}{h}$, pour h vecteur de $\mathbb{R}^n$ tendant vers 0. En effet, cela nécessiterait de donner un sens à la division par un vecteur.

En revanche, $\mathbb{C}$ étant un corps, il est possible d'effectuer $\frac{f(z_0+h)-f(z_0)}{h}$ et de regarder la limite pour $h \to 0 (\in \mathbb{C})$.

Définition - Fonctions holomorphes en un point (nombre complexe)

Soit $f : \mathbb{C} \to \mathbb{C}$, une fonction de la variable complexe (et à valeurs dans $\mathbb{C}$).
Soit U un ouvert de $\mathbb{C}$ (i.e. $\forall\, z \in U, \exists\, \rho > 0$ tel que $|z - z_0| < \rho \Rightarrow z \in U$.)
On dit que f est $\mathbb{C}$-dérivable ou (plutôt) holomorphe en $z_0 \in U$ si $\frac{f(z)-f(z_0)}{z-z_0}$ admet une limite (dans $\mathbb{C}$) pour $z \underset{z\in U}{\longrightarrow} z_0$.
Autrement écrit, on a les équivalences :

- f est holomorphe en z_0 de dérivée égale à $f'(z)$
- $\forall\, \epsilon > 0, \exists\, \rho > 0$ tel que $|z - z_0| < \rho \Longrightarrow \left| \frac{f(z)-f(z_0)}{z-z_0} - f'(z) \right| < \epsilon$.
- $\exists\, \epsilon : U \to \mathbb{C}$ tel que $\forall\, z \in U$, $f(z) = f(z_0) + f'(z) \times (z - z_0) + (z - z_0)\epsilon(z)$ avec $\epsilon(z) \underset{z\to z_0}{\longrightarrow} 0$.

Définition - Fonction holomorphe sur un ouvert

Soit $f : \mathbb{C} \to \mathbb{C}$, une fonction de la variable complexe (et à valeurs dans $\mathbb{C}$).
Soit U un ouvert de $\mathbb{C}$.
On dit que f est holomorphe sur U, si f est holomorphe en tout point z_0 de U.
On pourra noter $\mathscr{H}(U)$, l'ensemble des fonctions holomorphes définies sur l'ouvert U.

Pour aller plus loin - Connexe

Dans la pratique, il sera souvent nécessaire de nous restreindre à des ouverts qui sont connexes (parfois on se contentera de convexe).

Exemple - Fonction puissance entière

Soit $n \in \mathbb{N}$. On peut supposer $n \geqslant 2$ (sinon, c'est sans intérêt).
Pour $z \neq z_0 \in \mathbb{C}$:

$$\begin{aligned}\frac{z^n - z_0^n}{z - z_0} - n z_0^{n-1} &= \sum_{k=0}^{n-1} z^k z_0^{n-1-k} - \sum_{k=0}^{n-1} z_0^{n-1} = \sum_{k=1}^{n-1} (z^k - z_0^k) z^{n-k-1} \\ &= (z - z_0) \sum_{k=1}^{n-1} \sum_{h=0}^{k-1} z^h z_0^{k-h-1} z_0^{n-k-1} = (z - z_0) \sum_{h=0}^{n-2} (n-h-1) z^h z_0^{n-h-2}\end{aligned}$$

car on peut commencer la somme à $k = 1$. Pour z suffisamment proche de z_0 (en module) i.e. $|z - z_0| < \eta$, donc $|z| < |z_0| + \eta$:

$$\left| \sum_{h=0}^{n-2} (n-h-1) z^h z_0^{n-h-2} \right| \leqslant (|z_0|+\eta)^{n-2} \sum_{h=0}^{n-2} (n-h-1) = (|z_0|+\eta)^{n-2} \sum_{j=1}^{n-1} j = \frac{n(n-1)}{2} |z_0+\eta|^{n-2}$$

Donc pour $|z - z_0| < \eta$

$$\left| \frac{z^n - z_0^n}{z - z_0} - n z_0^{n-1} \right| \leqslant \eta \frac{n(n-1)}{2} |z_0 + \eta|^{n-2} \underset{\eta\to 0}{\longrightarrow} 0$$

Ainsi $z \mapsto z^n$ est holomorphe sur $\mathbb{C}$ (entier) et sa dérivée est $z \mapsto n z^{n-1}$.

Exercice
Si $n \in \mathbb{Z}$ et $n < 0$, montrer que $f_n : z \mapsto z^n$ est également holomorphe sur $\mathbb{C} \setminus \{0\}$

Correction
On note $m = -n \in \mathbb{N}^*$. $f_n(z) - f_n(z_0) = \dfrac{1}{z^m} - \dfrac{1}{z_0^m} = \dfrac{-(f_m(z) - f_m(z_0))}{(zz0)^m}$. Donc

$$\frac{f_n(z) - f_n(z_0)}{z - z_0} = \frac{-1}{z^m z_0^m} \frac{f_m(z) - f_m(z_0)}{z - z_0} \underset{z \to z_0}{\longrightarrow} \frac{-1}{z_0^{2m}} f'_m(z_0) = \frac{-m z_0^{m-1}}{z_0^{2m}} = \frac{-m}{z_0^{m+1}} = n z_0^{n-1}$$

2.2. Stabilité et premiers exemples

Comme l'application de passage à la limite est linéaire, on retrouve les mêmes résultats que pour les fonctions dérivables (sur $\mathbb{R}$)

Proposition - Stabilité
Soit f et g holomorphe sur un ouvert U de $\mathbb{C}$. Alors :
- Pour tout $a, b \in \mathbb{C}$, $af + bg$ est holomorphe sur U et
  $\forall\ z \in U,\ (\lambda f + \mu g)'(z) = \lambda f'(z) + \mu g'(z)$.
- $f \times g$ est holomorphe sur U et
  $\forall\ z \in U,\ (f \times g)'(z) = f'(z)g(z) + f(z)g'(z)$.
- si g ne s'annule pas sur U, $\dfrac{f}{g}$ est holomorphe sur U et
  $\forall\ z \in U,\ \left(\dfrac{f}{g}\right)'(z) = \dfrac{f'(z)g(z) - f(z)g'(z)}{g^2(z)}$.

On démontre les deux premiers résultats

Démonstration
On suppose que f et g sont holomorphes sur U.
Soit $z_0 \in U$. Il existe $\epsilon_1, \epsilon_2 : U \to \mathbb{C}$ tel que $\forall\ z \in U$, $f(z) = f(z_0) + f'(z) \times (z - z_0) + (z - z_0)\epsilon_1(z)$ et $g(z) = g(z_0) + g'(z_0)(z - z_0) + (z - z_0)\epsilon_2(z)$ avec $\epsilon_1(z) \underset{z \to z_0}{\longrightarrow} 0$ et $\epsilon_2(z) \underset{z \to z_0}{\longrightarrow} 0$. Donc, pour $\lambda, \mu \in \mathbb{C}$:

$$(\lambda f + \mu g)(z) = \lambda f(z) + \mu g(z) = (\lambda f + \mu g)(z_0) + [\lambda f'(z_0) + \mu g'(z_0)](z - z_0) + (z - z_0) \times \epsilon_3(z)$$

où $\epsilon_3(z) = \lambda\epsilon_1(z) + \mu\epsilon_2(z) \underset{z \to z_0}{\longrightarrow} 0$ par addition.
Et selon le même principe :

$$(f \times g)(z) = f(z) \times g(z) = f(z_0)g(z_0) + (z - z_0)[f'(z_0)g(z_0) + f(z_0)g'(z_0)] + (z - z_0)\epsilon_4(z)$$

où $\epsilon_4(z) = \epsilon_1(z) \times g(z) + \epsilon_2(z) \times f(z) + (z - z_0)f'(z_0)g'(z_0) \underset{z \to z_0}{\longrightarrow} 0$ par addition. □

Exercice
Démontrer le dernier résultat

Correction
On rappelle qu'il est malin d'écrire : $f(z)g(z) - f(z_0)g(z_0) = f(z)(g(z) - g(z_0)) + (f(z) - f(z_0))g(z_0)\ldots$
Comme les puissances entières sont holomorphes, par combinaison linéaire :

Corollaire - Polynômes (et fractions rationnelles)
Les fonctions polynomiales : $p : z \mapsto \sum_{k=0}^{n} a_k z^k$ sont holomorphes sur $\mathbb{C}$ et

pour tout $z \in \mathbb{C}$, $p'(z) = \sum_{k=0}^{n} k a_k z^{k-1} = \sum_{k=1}^{n} k a_k z^{k-1} = \sum_{h=0}^{n-1} (h+1) a_{h+1} z^h$.

Les fractions rationnelles sont holomorphes sur tout ouvert de $\mathbb{C}$ ne contenant aucun pôle de la fraction.
La fonction dérivée associée est « comme on l'imagine ».

Pour aller plus loin - Fonction analytique
On sait décrire parfaitement les fonctions holomorphes. Il s'agit des fonctions analytiques, i.e. des séries localement entière (ou polynôme de degré infini). Comme nous le verrons plus loin

Une fonction holomorphe sur $\mathbb{C}$, en entier, est qualifiée de fonction entière.
Une fonction polynomiale est donc une fonction entière.

2.3. Condition de Cauchy-Riemann

Remarque - Bijection $\mathbb{C} \to \mathbb{R}^2$

Il existe un isomorphisme (canonique/ naturelle) entre $\mathbb{R}^2$ et $\mathbb{C}$: $(x,y) \mapsto x+iy$.
On peut donc considérer $F : \mathbb{R}^2 \to \mathbb{R}^2$, $(x,y) \mapsto \big(\underbrace{\text{Re}(f(x+iy))}_{P(x,y)}; \underbrace{\text{Im}(f(x+iy))}_{Q(x,y)}\big)$.
Est-ce équivalent de dire que f est holomorphe en $z_0 = x_0 + iy_0$ ou F est différentiable en (x_0, y_0) ?
On conserve, pour la suite, les notations de la remarque

Proposition - Condition de Cauchy-Riemann

f est holomorphe en z_0 ssi F est différentiable en (x_0, y_0) avec

$$\begin{cases} \dfrac{\partial P}{\partial x}(x_0,y_0) = \dfrac{\partial Q}{\partial y}(x_0,y_0) \\ \dfrac{\partial P}{\partial y}(x_0,y_0) = -\dfrac{\partial Q}{\partial x}(x_0,y_0) \end{cases}$$

Savoir faire - Montrer l'holomorphie (ou non) avec $H : \mathbb{R}^2 \to \mathbb{C}$, $(x,y) \mapsto f(x+iy)$

On notera l'équivalence entre le système de la proposition et :
$$\left[\frac{\partial P}{\partial x}(x,y) + i\frac{\partial Q}{\partial x}(x,y)\right] + i\left[\frac{\partial P}{\partial y}(x,y) + i\frac{\partial Q}{\partial y}(x,y)\right] = 0.$$
Ce qui est encore équivalent à $\dfrac{\partial H}{\partial x} + i\dfrac{\partial H}{\partial y} = 0_{\mathbb{C}}$, pour $H : (x,y) \mapsto f(x+iy)$
(i.e. $H : \mathbb{R}^2 \to \mathbb{C}$)

Démonstration

Si f est holomorphe, il existe $\epsilon : U \to \mathbb{C}$ tel que $\epsilon(z) \underset{z \to z_0}{\longrightarrow} 0$, telle que $f(z) = f(z_0) + (z-z_0)f'(z) + (z-z_0)\epsilon(z)$.

$$\begin{aligned} P(x,y) - P(x_0,y_0) &= \text{Re}\big(f(x+iy) - f(x_0+iy_0) + ((x-x_0)+i(y-y_0))\epsilon(x+iy)\big) \\ &= \text{Re}\Big(\big((x-x_0)+i(y-y_0)\big) \times f'(x_0+iy_0) + ((x-x_0)+i(y-y_0))\epsilon(x+iy)\Big) \\ &= (x-x_0)\text{Re}\big(f'(x_0+iy_0)\big) - (y-y_0)\text{Im}\big(f'(x_0+iy_0)\big) + o(\|(x,y)\|) \\ Q(x,y) - Q(x_0,y_0) &= \text{Im}\big(f(x+iy) - f(x_0+iy_0) + ((x-x_0)+i(y-y_0))\epsilon(x+iy)\big) \\ &= \text{Im}\Big(\big((x-x_0)+i(y-y_0)\big) \times f'(x_0+iy_0) + ((x-x_0)+i(y-y_0))\epsilon(x+iy)\Big) \\ &= (x-x_0)\text{Im}\big(f'(x_0+iy_0)\big) + (y-y_0)\text{Re}\big(f'(x_0+iy_0)\big) + o(\|(x,y)\|) \end{aligned}$$

Donc F est différentiable en (x_0,y_0), avec $\dfrac{\partial P}{\partial x}(x_0,y_0) = \text{Re}\big(f'(x_0+iy_0)\big) = \dfrac{\partial Q}{\partial y}(x_0,y_0)$.

mais aussi $\dfrac{\partial P}{\partial y}(x_0,y_0) = -\text{Im}\big(f'(x_0+iy_0)\big) = -\dfrac{\partial Q}{\partial x}(x_0,y_0)$.

Réciproquement, on suppose que F est différentiable en (x_0,y_0) avec les relations de Cauchy-Riemann.

On note $A = \dfrac{\partial P}{\partial x}(x_0,y_0) + i\dfrac{\partial Q}{\partial x}(x_0,y_0) = \dfrac{\partial Q}{\partial y}(x_0,y_0) - i\dfrac{\partial P}{\partial y}(x_0,y_0)$.

Pour $z = x+iy$ proche de $z_0 = x_0+iy_0$,

$$\begin{aligned} f(z) - f(z_0) &= \big[F(x,y) - F(x_0,y_0)\big]_1 + i\big[F(x,y) - F(x_0,y_0)\big]_2 \\ &= \left[(x-x_0)\frac{\partial P}{\partial x}(x_0,y_0) + (y-y_0)\frac{\partial P}{\partial y}(x_0,y_0)\right] + i\left[(x-x_0)\frac{\partial Q}{\partial x}(x_0,y_0) + (y-y_0)\frac{\partial Q}{\partial y}(x_0,y_0)\right] \\ &\quad + ((x-x_0)+i(y-y_0))\epsilon(x+iy) \\ &= (x-x_0)\times A + i(y-y_0)\times A + ((x-x_0)+i(y-y_0))\epsilon(x+iy) = (z-z_0)\times A + (z-z_0)\epsilon(z) \end{aligned}$$

Donc f est holomorphe en z_0, avec $f'(z_0) = A$. □

Application - $z \mapsto e^z$ est holomorphe

Ici $H : (x,y) \mapsto e^x\cos y + ie^x \sin y$, différentiable sur $\mathbb{R}^2$.
Et $\dfrac{\partial H}{\partial x} + i\dfrac{\partial H}{\partial y} = (e^x\cos y + ie^x\sin y) + i(-e^x\sin y + ie^x\cos y) = 0$.
Donc $f : z \mapsto e^z$ est holomorphe sur $\mathbb{C}$.

Application - $z \mapsto \overline{z}$ n'est pas holomorphe

Même stratégie : ici $H : (x,y) \mapsto x - iy$, différentiable sur $\mathbb{R}^2$.
Et $\dfrac{\partial H}{\partial x} + i\dfrac{\partial H}{\partial y} = 1 + i(-i) = 2 \neq 0$.
Donc $f : z \mapsto \overline{z}$ n'est pas holomorphe sur $\mathbb{C}$

Heuristique - Du plan tangent à la similitude (locale)

Le fait que F soit différentiable nous dit, géométriquement, que P et Q sont comparables chacun à un plan au voisinage de tout (x_0, y_0), ou bien que P et Q admetten un plan tangent en (x_0, y_0).
Qu'ajoute l'information sur les relations de Cauchy-Riemann pour les fonctions holomorphes?

Les vecteurs normaux au plan (∇P et ∇Q) sont orthogonaux : $\dfrac{\partial P}{\partial x}\dfrac{\partial Q}{\partial x} + \dfrac{\partial P}{\partial y}\dfrac{\partial Q}{\partial y} = 0$.

Ainsi la matrice jacobienne est une matrice orthogonale (à condition de la normalisée par la racine carrée de son déterminant) du plan, il s'agit donc d'une rotation.
Une fonction holomorphe est localement la composée d'une rotation et d'une homothétie (normalisation) : c'est une similitude.

Remarque - Application conforme

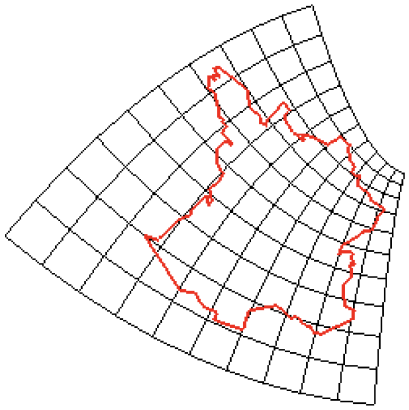
On a vu en début d'année, qu'une similitude conserve les angles; i.e. si s est une similitude :

$$\left(\overrightarrow{s(A)s(B)}, \overrightarrow{s(C), s(D)}\right) = \arg\frac{s(d)-s(c)}{s(b)-s(a)} = \arg\frac{d-c}{b-a} = \left(\overrightarrow{AB}, \overrightarrow{C, D}\right)$$

on rappelle que $s : z \mapsto z_0 + re^{i\theta}(z - z_0)$. Donc les similitudes conservent les angles (souvent on prend $A = C$, ici). Une telle application est dite application conforme.
Une fonction holomorphe est donc (localement) une application conforme : elle conserve les formes des petites figures (mais pas les longueurs)

Représentation - Application conforme

Une application conforme conserve les angles localement. L'image de la carte de la France par une application conforme donne quelque chose de reconnaissable.



2.4. Fonction analytique e(s)t fonction holomorphe

Un mot sur les séries de fonctions (de référence)

Définition - Séries entières. Rayon de convergence

On appelle série entière les fonctions de la forme $S : z \in \mathbb{C} \mapsto \sum_{k=0}^{+\infty} a_n z^n$ où $(a_n) \in \mathbb{C}^{\mathbb{N}}$ est une suite de complexe.
On pose $R_S = \sup\{r \in \mathbb{R}_+ \mid (|a_n| r^n)_{n\in\mathbb{N}} \text{ bornée}\}$, appelé rayon de convergence de la série (entière).

On énonce une série de propriétés pour les séries entières, elles seront démontrées en seconde année.

Proposition - Régularité d'une série entière

La fonction $S_{\mathbb{R}}$ est définie, continue et de classe $\mathscr{C}^\infty$ sur le « disque » ouvert $]-R, R[$. On a alors

$$\forall\, n \in \mathbb{N},\ \forall\, x \in]-R, R[,\ S^{(n)}(x) = \sum_{k=n}^{+\infty} \frac{k!}{(k-n)!} a_k x^{k-n} = \sum_{h=0}^{+\infty} \frac{(h+n)!}{h!} a_{h+n} x^h.$$

Heuristique - Idées de démonstration

1. D'abord, notons que pour $|x| > R$ la série diverge grossièrement. $\mathscr{D}_{S_{\mathbb{R}}} \subset [-R, R]$.
Si $|x| < R$, alors avec $\rho \in]|x|, R[$, on a $|a_n x^n| \leqslant a_n \rho^n \times \left|\frac{|x|}{\rho}\right|^n \leqslant M\left|\frac{|x|}{\rho}\right|^n$. Et la série entière est **normalement** convergente. Tout se passera bien : continuité, dérivabilité… Ce qui se passe en R et $-R$ dépend de chaque série considérée. On peut tout avoir.
2. On calcule la dérivée, comme limite de la série des dérivées. C'est une nouvelle série entière.
3. On trouver pour $S'_{\mathbb{R}}$, le même rayon de convergence (petite manipulation).
Et on continue ainsi pour toutes les dérivées!

De la même façon, on a le théorème suivant (qui ne sera pas démontré…) :

Proposition - Série entière et fonctions holomorphes
La fonction S est définie, continue, holomorphe et de classe $\mathscr{C}^\infty$ (dérivation complexe) sur le disque ouvert $\mathscr{D}(0,R_S)=\{z\in\mathbb{C}\mid |z|<R_S\}$. Alors

$$\forall\, n\in\mathbb{N},\ \forall\, z\in\mathscr{D}(0,R_S),\ S^{(n)}(z)=\sum_{k=n}^{+\infty}\frac{k!}{(k-n)!}a_k z^{k-n}=\sum_{h=0}^{+\infty}\frac{(h+n)!}{h!}a_{h+n}z^h.$$

Définition - Fonction analytique
On dit que $f:U\subset\mathbb{C}\to\mathbb{C}$ est analytique sur l'ouvert U si
pour tout $z_0\in U$, il existe $R>0$ et $(a_n)\in\mathbb{C}^{\mathbb{N}}$ tel que :

$$\forall\, z\in\mathscr{D}(z_0,R),\qquad f(z)=\sum_{n=0}^{+\infty}a_n(z-z_0)^n$$

(Il s'agit du disque ouvert de $\mathbb{C}$ centré en z_0 et de rayon R.
Si f est analytique sur U, alors f est holomorphe sur U

On reviendra en fin de chapitre sur les propriétés analytiques (prolongement…) des fonctions holomorphes. C'est l'un des pierres précieuses des fonctions holomorphes. En attendant, mais sans pouvoir faire démonstration, on énonce le théorème suivant, dont la connaissance est importante à ce stade du chapitre :

Théorème - Holomorphie et analycité
Si f est holomorphe sur un ouvert U, alors f est analytique sur U.
Ainsi, pour tout $z_0\in U$, il existe $R>0$ et (a_n) tel que $\forall\, z\in\mathscr{D}(z_0,R)$, $f(z)=\sum_{n=0}^{+\infty}a_n(z-z_0)^n$.
Et on a, pour tout $r\in]0,R[$:

$$a_n=\frac{1}{2r^{n+1}i\pi}\int_0^{2\pi}f(z_0+re^{i\theta})e^{-i(n+1)\theta}\mathrm{d}\theta=\frac{1}{2i\pi}\oint_{\mathscr{C}(z_0,r)}\frac{f(z)}{(z-z_0)^{n+1}}\mathrm{d}z$$

Attention - Dépendance des (a_n) par rapport à z_0
On rappelle que chaque z_0 définit un rayon distinct et une suite (a_n) propre.
On aurait du noter $(a_n(z_0))_n$.

3. Chemins dans le plan complexe et intégrale curviligne

3.1. Arc du plan

On commence par beaucoup de vocabulaire dont le sens géométrique ne devrait pas échapper au lecteur.

Définition - Arc du plan complexe ou chemin
Soit $[a,b]$ un compact de $\mathbb{R}$ et $\gamma:[a,b]\to\mathbb{C}$, **continue**.
L'image du compact $[a,b]$ de $\mathbb{R}$ par γ est appelé arc de $\mathbb{C}$, noté Γ, γ est alors appelé (une) paramétrisation de Γ.
On dit que l'arc est fermé si $\gamma(a)=\gamma(b)$.
On dit que l'arc est simple si γ est injectif (i.e. $\gamma(t)\neq\gamma(s)$ dès que $s\neq t\in]a,b[$).

Si γ est de classe $\mathscr{C}^1$, on dit que γ est un chemin. Un chemin fermé est parfois qualifié de lacet.

Pour aller plus loin - Chemins continument différentiable
Pour les chemins, le cours devrait se limiter aux applications γ dérivables par morceaux.
Nous simplifions ici le cours en ne prenant que des chemins de classe $\mathscr{C}^1$. En revanche, nous couperons nos intégrales curvilignes en morceaux si l'on rencontre des coupures naturelles de chemin

Attention - Tolérance
Parfois on appelle chemin, à la fois la paramétrisation de la courbe et sa représentation géométrique.
Le contexte est en règle générale suffisant pour faire la distinction.
Mais cette tolérance peut créer une confusion pour le débutant; on essayera de bien s'en garder.

Remarque - Interprétation des contions γ de classe $\mathscr{C}^1$ et $\gamma'(t) \neq 0$
La paramétrisation γ s'appelle parfois le mobile (comme un point mobile qui se déplace). A chaque instant t il laisse sa trace sur la courbe. Le fait que γ soit dérivable signifie que le mobile a une vitesse. $\gamma'(t) \neq 0$ signifie qu'elle ne s'annule pas et comme le déplacement du mobile est continue, cela implique qu'il ne revient pas en arrière.

Exemple - Paramétrisation du cercle unité
L'application : $[0, 2\pi] \to \mathbb{C}$, $t \mapsto e^{it}$ est une paramétrisation de l'arc : cercle unité.
Une autre, associée au compact $[0,1]$ est $t \mapsto e^{2i\pi t}$.

Proposition - Paramétrisation à partir de $[0,1]$
Soit Γ un chemin de $\mathbb{C}$.
Alors il existe une paramétrisation γ de Γ, à partir du compact $[0,1]$.

Démonstration
Considérons une paramétrisation $\gamma_1 : [a,b] \to \mathbb{C}$ de Γ.
Considérons $\varphi : [0,1] \to [a,b]$, $t \mapsto a + (b-a)t$, alors $\gamma : t \mapsto \gamma_1 \circ \varphi$ est une paramétrisation de Γ à partir du compact $[0,1]$ □

Définition - Chemins équivalents
Soient $\gamma_1 : [a,b] \to \mathbb{C}$ et $\gamma_2 : [c,d] \to \mathbb{C}$ deux chemins.
Si il existe une bijection $\varphi : [a,b] \to [c,d]$, de classe $\mathscr{C}^1$ et croissante telle que $\gamma_1 = \gamma_2 \circ \varphi$, alors on dit que les chemins sont équivalents.
Il s'agit d'une relation d'équivalence.

Exercice
intégrale curviligne

Correction
Fiche de TD

3.2. Intégration le long d'un chemin

Analyse - Intégration d'une fonction holomorphe le long d'un chemin dans $\mathbb{C}$
Comment définir $\displaystyle\int_\Gamma f(z)\mathrm{d}z$?
Si on exploite la même idée que pour les sommes de Riemann, on se trouve en présence de $\displaystyle\sum_{k=0}^{n} f(u_k)(y_{k-1} - y_k)$ où $\tau_p = \big(([y_{k-1}, y_k], u_k), k \in \mathbb{N}_n\big)$ est une subdivision pointée de Γ.
Or la seule chose qu'on peut dire des points x_k et t_k de Γ, c'est qu'ils sont des images par le paramétrage γ du compact $[a,b]$.
Ainsi, en simplifiant à la condition que le paramétrage soit croissant, il existe une subdivision pointée $\sigma_p = \big(([x_{k-1}, x_k], t_k), k \in \mathbb{N}_n\big)$ de $[a,b]$ telle que $\gamma(\sigma_p) = \tau_p$, i.e. $y_k = \gamma(x_k)$ et $u_k = \gamma(t_k)$.
On a donc (en exploitant l'égalité des accroissements finis) :

$$\int_\Gamma f(z)\mathrm{d}z = \lim_{\sigma_p \to 0} \sum_{k=1}^{n} f(\gamma(t_k))(\gamma(x_k) - \gamma(x_{k-1}) \lim_{\sigma_p \to 0} \sum_{k=1}^{n} f(\gamma(t_k))\gamma'(c_k)(x_k - x_{k-1}) = \int_a^b f(\gamma) \times \gamma'$$

La définition suivante, reste vraie pour des fonctions f non nécessairement holomorphe.

Définition - Intégration curviligne le long d'un chemin
Soit U un ouvert de $\mathbb{C}$. Soient $f : U \to \mathbb{C}$ et $\gamma : [a,b] \to U$, un chemin.
On définit l'intégrale de g sur le chemin Γ :

$$\int_\Gamma g(z)\mathrm{d}z = \int_a^b g(\gamma(t))\gamma'(t)\mathrm{d}t$$

Le résultat ne dépend pas du choix du paramétrage γ du chemin Γ.

Le résultat qui suit sera très souvent exploité !

Application - $\oint_{\mathbb{U}} z^n \mathrm{d}z$, pour $n \in \mathbb{Z}$
On note $\mathbb{U}$ le cercle trigonométrique dont une paramétrisation est $\gamma : [0,1] \to \mathbb{C}$, $t \mapsto e^{2i\pi t}$.
On a alors, pour tout entier $n \in \mathbb{Z}$:

$$\oint_{\mathbb{U}} z^n \mathrm{d}z = \int_0^1 \left(e^{2i\pi t}\right)^n \times (2i\pi) e^{2i\pi t}\mathrm{d}t = 2i\pi \int_0^1 e^{2i\pi(n+1)t}\mathrm{d}t$$

Si $n+1 \neq 0$ i.e. $n \neq -1$:

$$\oint_{\mathbb{U}} z^n \mathrm{d}z = \left[\frac{1}{n+1} e^{2i\pi(n+1)t}\right]_0^1 = \frac{1}{n+1}[1-1] = 0$$

Si $n+1 = 0$ i.e. $n = 0$:

$$\oint_{\mathbb{U}} z^n \mathrm{d}z = 2i\pi \int_0^1 \mathrm{d}t = 2i\pi$$

Reste à démontrer l'indépendance du choix du paramétrage.

Démonstration
Soient γ_1 et γ_2 deux paramétrages de Γ.
Alors il existe une bijection $\varphi : [c,d] \to [a,b]$ de classe $\mathscr{C}^1$ tel que $\gamma_2 = \gamma_1 \circ \varphi$.
On a alors $\gamma_2' = \varphi' \times \gamma_1' \circ \varphi$ et donc

$$\int_c^d f(\gamma_2(t))\gamma_2'(t)\mathrm{d}t = \int_c^d f(\gamma_1(\varphi(t))\gamma_1'(\varphi(t))\varphi'(t)\mathrm{d}t = \int_a^b f(\gamma_1(s))\gamma_1'(s)\mathrm{d}s$$

en faisant le changement de variable $s = \varphi(t)$ dans le calcul intégral. □

Proposition - Intégrale curviligne d'une dérivée
Soit f une fonction holomorphe sur U et $\gamma : [a,b] \to U$, un chemin.
Alors $\displaystyle\int_\Gamma f'(z)\mathrm{d}z = f(\gamma(b)) - f(\gamma(a))$.

Démonstration
On fait le calcul

$$\int_\Gamma f'(z)\mathrm{d}z = \int_a^b f'(\gamma(t)) \times \gamma'(t)\mathrm{d}t = \left[f(\gamma(t))\right]_a^b = f(\gamma(b)) - f(\gamma(a))$$

□

Corollaire - Intégrale d'une dérivée sur un chemin fermé
Si Γ est un chemin fermé et que f est la dérivée d'une fonction holomorphe, alors $\oint_\Gamma f = 0$.

Savoir faire - Démontrer qu'une fonction n'est pas la dérivée d'une fonction holomorphe

Si l'intégrale d'une fonction f sur un chemin fermé n'est pas nulle, alors, elle ne peut être la dérivée d'une fonction holomorphe.

Exemple - z^{-1}

Pour tout $n \in \mathbb{Z} \setminus \{-1\}$, $z \mapsto z^n$ est la dérivée d'une fonction holomorphe : $z \mapsto \dfrac{1}{n+1} z^{n+1}$.
En revanche, ce n'est pas le cas de $z \mapsto \frac{1}{z} = z^{-1}$.

Exercice

Après l'intégration, plusieurs ont eu en colle : calculer $\displaystyle\int_0^\pi \ln(a^2 - 2a\cos t + 1)\mathrm{d}t$.
Voici une nouvelle question : interpréter en terme de intégrale curviligne le calcul

Correction

Pour aller plus loin - Invariance et homotopie (marge)

Une homotopie est une déformation continue entre deux applications, notamment entre les chemins à extrémités fixées et en particulier les lacets. Cette notion topologique permet de définir des invariants algébriques utilisés pour classifier les applications continues entre espaces topologiques dans le cadre de la topologie algébrique.

3.3. Longueur d'une courbe et majoration

Analyse - Longueur d'un chemin dans $\mathbb{C}$

Considérons le chemin Γ paramétré par γ, définie sur $[a,b]$.
On note ℓ la longueur de Γ que l'on aimerait calculer.
Soit $\sigma = (x_k)_{k\in[\![0,n]\!]}$, une subdivision de $[a,b]$.
La longueur de Γ entre les points $M_{k-1} = \gamma(x_{k-1})$ et $M_k = \gamma(x_k)$ peut être approchée par $\|M_{k-1}M_k\| = |z_k - z_{k-1}| = |\gamma(x_k) - \gamma(x_{k-1})|$.
D'après l'inégalité des accroissements finis :

$$\sum_{k=1}^{n} \inf_{[x_{k-1},x_k]} |\gamma|(x_k - x_{k-1}) \leqslant \sum_{k=1}^{n} |\gamma(x_k) - \gamma(x_{k-1})| \leqslant \sum_{k=1}^{n} \sup_{[x_{k-1},x_k]} |\gamma|(x_k - x_{k-1})$$

On reconnait des sommes de Darboux, elles convergent (quand tout va bien) vers $\displaystyle\int_a^b |\gamma'(x)|\mathrm{d}x$.

Proposition - Longueur d'un chemin de $\mathbb{C}$

La longueur d'un chemin Γ de $\mathbb{C}$ paramétré par γ (de classe $\mathscr{C}^1$) est donné par

$$\ell(\Gamma) = \int_a^b |\gamma'(t)|\mathrm{d}t$$

indépendant du choix de γ.

Application - Périmètre d'une ellipse

Soient $a, b > 0$. La chemin $z : t \mapsto a\cos t + i\sin t$ décrit une ellipse, pour $t \in [0, 2\pi]$.
Sa longueur (périmétre) est

$$\ell = \int_0^{2\pi} |-a\sin t + ib\cos t|\mathrm{d}t = \int_0^{2\pi} \sqrt{a^2\sin^2 t + b^2\cos^2 t}\,\mathrm{d}t = b\int_0^{2\pi} \sqrt{1 + e^2\sin^2 t}\,\mathrm{d}t$$

où $e = \dfrac{\sqrt{a^2 - b^2}}{b}$. Et on ne sait pas faire plus...

Proposition - Inégalité des accroissements finis

Soit Γ un chemin paramétré par γ de classe $\mathscr{C}^1$.
Alors, pour toute fonction f holomorphe sur U contenant Γ,

$$\left|\int_\Gamma f(z)\mathrm{d}z\right| \leqslant \sup_\Gamma |f| \times \ell(\Gamma)$$

Démonstration
Le calcul donne :

$$\left|\int_\Gamma f(z)\mathrm{d}z\right| = \left|\int_a^b f(\gamma(t))\gamma'(t)\mathrm{d}t\right| \leqslant \sup_\Gamma |f| \int_a^b |\gamma'(t)|\mathrm{d}t = \sup_\Gamma |f| \times \ell(\Gamma)$$

□

4. Théorème(s) de Cauchy

4.1. Indice de Cauchy

Définition - Indice (de Cauchy) d'un chemin par rapport à un point z_0
Soient Γ un chemin fermé de $\mathbb{C}$, de paramétrisation γ.
On note U, le complémentaire de Γ dans $\mathbb{C}$. Soit $z_0 \in \mathbb{C}$.
On appelle indice de Γ (ou Γ) par rapport à z_0, le nombre.

$$\mathrm{Ind}_\Gamma(z_0) = \frac{1}{2i\pi}\oint_\Gamma \frac{\mathrm{d}z}{z-z_0}$$

Remarque - Autre nom
On parle parfois de l'indice de z_0 par rapport à Γ (ou γ).

Exemple - $\Gamma = \mathbb{U}$, le cercle unité et $z_0 \in \mathbb{C}$
Si $z_0 = 0$, $\mathrm{Ind}_\Gamma(z_0) = \dfrac{1}{2i\pi}\displaystyle\int_0^{2\pi} \frac{ie^{i\theta}}{e^{i\theta}-0}\mathrm{d}\theta = 1$.
Si $z_0 = \frac{1}{2}$,

$$\begin{aligned}
\mathrm{Ind}_\Gamma(z_0) &= \frac{1}{2i\pi}\int_0^{2\pi} \frac{ie^{i\theta}}{e^{i\theta}-\frac{1}{2}}\mathrm{d}\theta \\
&= \frac{1}{\pi}\int_0^{2\pi} \frac{e^{i\theta}}{2e^{i\theta}-1}\mathrm{d}\theta = \frac{1}{\pi}\int_0^{2\pi} \frac{e^{i\theta}(2e^{-i\theta}-1)}{5-4\cos\theta}\mathrm{d}\theta = \frac{1}{\pi}\int_0^{2\pi} \frac{2-e^{i\theta}}{5-4\cos\theta}\mathrm{d}\theta \\
&= \frac{1}{2\pi}\int_0^{\pi} \frac{2-\cos\theta}{5-4\cos\theta}\mathrm{d}\theta - \frac{i}{\pi}\int_0^{2\pi} \frac{\sin\theta}{5-4\cos\theta}\mathrm{d}\theta \qquad t=\tan\frac{\theta}{2} \\
&= \frac{1}{2\pi}\int_0^{+\infty} \frac{1+3t^2}{(1+9t^2)(1+t^2)}\mathrm{d}t - \frac{i}{4}\left[\ln(5-4\cos(\theta))\right]_0^{2\pi} \\
&= \frac{1}{2\pi}\int_0^{+\infty} \frac{1}{1+t^2}+\frac{3}{1+9t^2}\mathrm{d}t = \frac{1}{2\pi}\left[\arctan t + \arctan 3t\right]_0^{+\infty} = 1
\end{aligned}$$

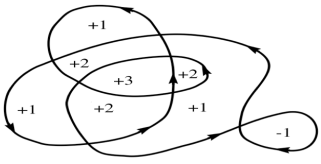
Si $z_0 = 2$,

$$\begin{aligned}
\mathrm{Ind}_\Gamma(z_0) &= \frac{1}{2i\pi}\int_0^{2\pi} \frac{ie^{i\theta}}{e^{i\theta}-2}\mathrm{d}\theta = \frac{1}{2\pi}\int_0^{2\pi} \frac{e^{i\theta}(e^{-i\theta}-2)}{5-4\cos\theta}\mathrm{d}\theta = \frac{1}{2\pi}\int_0^{2\pi} \frac{1-2e^{i\theta}}{5-4\cos\theta}\mathrm{d}\theta \\
&= \frac{1}{2\pi}\int_0^{\pi} \frac{1-2\cos\theta}{5-4\cos\theta}\mathrm{d}\theta - \frac{i}{\pi}\int_0^{2\pi} \frac{\sin\theta}{5-4\cos\theta}\mathrm{d}\theta \qquad t=\tan\frac{\theta}{2} \\
&= \frac{1}{2\pi}\int_0^{+\infty} \frac{-1+3t^2}{(1+9t^2)(1+t^2)}\mathrm{d}t - \frac{i}{4}\left[\ln(5-4\cos(\theta))\right]_0^{2\pi} \\
&= \frac{1}{4\pi}\int_0^{+\infty} \frac{1}{1+t^2}-\frac{3}{1+9t^2}\mathrm{d}t = \frac{1}{4\pi}\left[\arctan t - \arctan 3t\right]_0^{+\infty} = 0
\end{aligned}$$

Avec les mêmes notations

Proposition - Valeur de l'indice
$\mathrm{Ind}_\Gamma(z_0)$ est à valeurs entières et constant sur chaque composante connexe de U et il est nul sur la composante connexe non bornée de U.

Représentation - Indice d'un point
Il s'agit du nombre de tour que la boucle effectue autour d'un point.
Cela est géométriquement facilement visible, même sans calcul :



Démonstration
• $\mathrm{Ind}_\Gamma(z_0)$ est à valeurs entières.
Notons $\gamma : [a,b] \to \mathbb{C}$ une paramétrisation de Γ de classe $\mathscr{C}^1$, donc $\mathrm{Ind}_\Gamma(z) = \dfrac{1}{2i\pi}\displaystyle\int_a^b \frac{\gamma'(t)}{\gamma(t)-z}\mathrm{d}t$.
Considérons, pour $z \in U$ ($z \notin \Gamma$), $\varphi_z : u \mapsto \exp\left(\displaystyle\int_a^u \frac{\gamma'(t)}{\gamma(t)-z}\mathrm{d}t\right)$.
Par composition φ_z est dérivable sur $[a,b]$ et

$$\varphi_z'(u) = \frac{\gamma'(u)}{\gamma(u)-z}\varphi_z(u)$$

La fonction $\Psi_z : t \mapsto \dfrac{\varphi_z(t)}{\gamma(t)-z}$ est dérivable également et

$$\Psi'_z(t) = \frac{\varphi'_z(t)(\gamma(t)-z) - \varphi_z(t)\gamma'(t)}{(\gamma(t)-z)^2} = 0$$

Donc Ψ est constante sur $[a,b]$. Or $\Psi(a) = \dfrac{\varphi_z(0)}{\gamma(0)-z} = \exp(0) = 1$, donc pour tout $t \in [a,b]$:

$$\Psi_z(t) = 1 = \frac{\varphi_z(t)}{\gamma(t)-z}$$

Ainsi $\varphi_z(t) = \gamma(t) - z$. Et donc $\varphi_z(b) = \gamma(b) - z = \gamma(a) - z = \varphi_z(a) = 1$.
Donc $\exp(2i\pi \mathrm{Ind}_\Gamma(z)) = 1$ donc $\mathrm{Ind}_\Gamma(z) \in \mathbb{Z}$.
• $\mathrm{Ind}_\Gamma(z_0)$ est constante sur les composantes connexes.

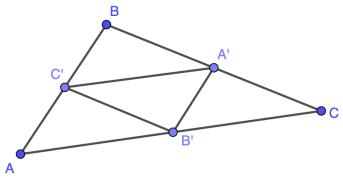
$$z \mapsto \int_a^b \frac{\gamma'(t)}{\gamma(t)-z}\mathrm{d}t = \sum_{k=0}^{+\infty}\left(\int_a^b \frac{\gamma'(t)}{\gamma^{n+1}(t)}\right)z^k$$

Donc Ind_Γ est fonction analytique, donc holomorphe donc continue.
Une fonction continue sur une partie connexe a une image connexe (T.V.I.). Mais comme Ind_Γ est à valeurs dans un ensemble à valeurs séparées, nécessairement Ind_Γ est constante sur chaque composante connexe.
• Sur la composante non bornée.
Si z est suffisamment grand, $\mathrm{Ind}_\Gamma(z)$ est un entier de valeurs absolues strictement plus petite que 1.
Donc $\mathrm{Ind}_\Gamma(z) = 0$. □

4.2. Lemme de Goursat et holomorphie

Proposition - Lemme de Goursat
Soit Δ un triangle (fermé) de U ouvert de C. Soit f une fonction holomorphe sur U.
Alors $\oint_\Delta f(z)\mathrm{d}z = 0$

❄ ***Représentation - Décomposition du triangle***



Démonstration
On note A, B, C les trois sommets du triangle (plein!!) $\Delta = [A \to B \to C \to A]$.
On note respectivement A', B' et C' les milieux de $[BC]$, $[AC]$ et $[AB]$ (cf figure dans la marge).
On note $\Delta_1 = [A \to C' \to B' \to A]$, $\Delta_2 = [C' \to B \to A' \to C']$, $\Delta_3 = [A' \to C \to B' \to A']$ et enfin $\Delta_4 = [C' \to A' \to B' \to C']$.
Par relation de Chasles (dans $\mathbb{R}$, avec le paramétrage) :

$$\begin{aligned}
&\oint_{\Delta_1} f(z)\mathrm{d}z + \oint_{\Delta_2} f(z)\mathrm{d}z + \oint_{\Delta_3} f(z)\mathrm{d}z = \left(\int_{A\to C'} f(z)\mathrm{d}z + \int_{C'\to B'} f(z)\mathrm{d}z + \int_{B'\to A} f(z)\mathrm{d}z\right) + \cdots \\
&= \left(\int_{A\to C'} f(z)\mathrm{d}z + \int_{C'\to B} f(z)\mathrm{d}z + \int_{B\to A'} f(z)\mathrm{d}z + \int_{A'\to C} f(z)\mathrm{d}z + \int_{C\to B'} f(z)\mathrm{d}z + \int_{B'\to A} f(z)\mathrm{d}z\right) \\
&\qquad - \left(\int_{C'\to A'} f(z)\mathrm{d}z + \int_{A'\to B'} f(z)\mathrm{d}z + \int_{B'\to C'} f(z)\mathrm{d}z\right) \\
&= \oint_\Delta f(z)\mathrm{d}z - \oint_{\Delta_4} f(z)\mathrm{d}z
\end{aligned}$$

Supposons, par l'absurde que $\oint_\Delta f(z)\mathrm{d}z \neq 0$, donc $C := \left|\oint_\Delta f(z)\mathrm{d}z\right| > 0$ On a alors

$$C := \left|\oint_{\Delta_1} f(z)\mathrm{d}z + \oint_{\Delta_2} f(z)\mathrm{d}z + \oint_{\Delta_3} f(z)\mathrm{d}z + \oint_{-\Delta_4} f(z)\mathrm{d}z\right| > 0$$

Par inégalité triangulaire :

$$0 < C \leqslant \left|\oint_{\Delta_1} f(z)\mathrm{d}z\right| + \left|\oint_{\Delta_2} f(z)\mathrm{d}z\right| + \left|\oint_{\Delta_3} f(z)\mathrm{d}z\right| + \left|\oint_{-\Delta_4} f(z)\mathrm{d}z\right|$$

Notons $C' = \max\left|\oint_{\Delta_i} f(z)\mathrm{d}z\right|$, on a donc $C \leqslant 4C'$.
Donc il existe i tel que $\left|\oint_{\Delta_i} f(z)\mathrm{d}z\right| > \dfrac{C}{4}$.
On continue ainsi, en coupant Δ_i en 4 morceaux comme précédemment, et ainsi, par récurrence, on crée une suite $(T_i)_{i\in\mathbb{N}}$, de triangles emboités tels que :

$$\text{pour tout } n \in \mathbb{N},\ \left|\oint_{T_n} f(z)\mathrm{d}z\right| > \frac{C}{4^n} \text{ et également } \ell(T_n) = \tfrac{1}{2}\ell(T_{n-1}) = \frac{\ell(\Delta)}{2^n}.$$

En réalité, chaque T_n, triangle plein est un compact de $\mathbb{C}$.
On a une suite emboitée de compact de $\mathbb{C}$, donc son intersection est un fermé non vide.
La longueur de ces compacts tend vers 0 donc (T_n) converge vers un unique point a.
f est holomorphe en a. Donc, pour tout $\epsilon > 0$,

il existe V voisinage de a, tel que : $\forall\ z\in V, |f(z)-f(a)-f'(a)\times(z-a)|\leqslant\epsilon|z-a|$.
Or, il existe n tel que $T_n\subset V$, et donc

$$\left|\oint_{T_n} f(z)\mathrm{d}z - f(a)\oint_{T_n}\mathrm{d}z + f'(a)\oint_{T_n}(z-a)\mathrm{d}z\right|\leqslant\epsilon\oint_{T_n}|z-a|\mathrm{d}z\leqslant\epsilon\sup_{T_n}|z-a|\times\ell(T_n)$$

Et comme $f(a)\oint_{T_n}\mathrm{d}z = f'(a)\oint_{T_n}(z-a)\mathrm{d}z = 0$, on a :

$$\left|\oint_{T_n} f(z)\mathrm{d}z\right|\leqslant\epsilon\ell(T_n)^2=\frac{\epsilon\ell^2(\Delta)}{2^{2n}}$$

Ainsi, pour tout $\epsilon>0$

$$0\leqslant C<4^n\times\epsilon\frac{\ell^2(D_n)}{4^n}=\epsilon\ell^2(\Delta)$$

Donc $C=0$. □

Remarque - Allégement des hypothèses

En fait, on peut considérer que f est holomorphe sur $U\setminus\{p\}$ où p est un point quelconque de Δ.
Mais la démonstration est plus compliquée.

Proposition - Dérivabilité

Soit U un ouvert, étoilé (i.e. : $\exists\ A\in U$ tel que $\forall\ M\in U$, $[AM]\subset U$.
On suppose que f est holomorphe sur U, alors f admet une primitive sur U.

Démonstration

Considérons le point $A(a)$ à l'origine de l'étoile incluse dans U.
Notons, pour tout $z\in U$ (z affixe du point M),

$$F:z\mapsto\int_{[AM(z)]}f(u)\mathrm{d}u=\int_0^1(z-a)\times f\big(a+t(z-a)\big)\mathrm{d}t$$

car $\gamma:[0,1]\to U$, $t\mapsto(1-t)a+tz=z_0+(z-a)t$ est une paramétrisation de $[AM]$.
Fixons $z_0\in U$ et z, au voisinage de z_0, inclus dans U (ouvert).

$$\begin{aligned}F(z)-F(z_0)=&\int_{[AM(z)]}f(u)\mathrm{d}u-\int_{[AM_0(z_0)]}f(u)\mathrm{d}u=\int_{[AM(z)]}f(u)\mathrm{d}u+\int_{[M_0(z_0)A]}f(u)\mathrm{d}u\\=&-\int_{[M(z)M_0(z_0)]}f(u)\mathrm{d}u=\int_{[M_0M]}f(u)\mathrm{d}u=\int_0^1(z-z_0)\times f\big(z_0+t(z-z_0)\big)\mathrm{d}t\end{aligned}$$

puisque, comme f est holomorphe : $\oint_{[AMM_0A]}f(\tau)\mathrm{d}\tau=0$.
Donc :

$$\frac{F(z)-F(z_0)}{z-z_0}-f(z_0)=\int_0^1\big(f(z_0+t(z-z_0))-f(z_0)\big)\mathrm{d}t$$

Puis f continue au voisinage de z_0.
Soit $\epsilon>0$. Il existe V, voisinage de z_0, tel que $\forall\ u\in V$, $|f(u)-f(z_0)|\leqslant\epsilon$.
Puis, pour tout $z\in V$, $\forall\ t\in[0,1]$, $|z_0+t(z-z_0)-z_0|\leqslant|z-z_0|$, donc $u:=z_0+t(z-z_0)-z_0\in V$.
Ainsi :

$$\left|\frac{F(z)-F(z_0)}{z-z_0}-f(z_0)\right|\leqslant\int_0^1\epsilon\mathrm{d}t=\epsilon$$

Ainsi F est une primitive de f. □

On a le corollaire suivant, puisque $f=F'$:

Théorème - Première formule de Cauchy

Si U est un ouvert étoile, ou donc un ouvert convexe voire un ouvert simplement connexe. Si f est holomorphe sur U et Γ est un chemin fermé ou lacet de U :

$$\oint_\Gamma f(z)\mathrm{d}z=0$$

Remarque - Hypothèses du théorème

Sur les points de non-holomorphie de f (dont on parlera plus loin), le théorème indique qu'il n'y a aucun point singulier de f à l'intérieur de Γ.

Pour aller plus loin - Fonction harmonique

Les applications qui vérifie la propriété de la moyenne sur la boule sont des fonctions harmoniques (vérifiant également $\Delta f=0$ (lire Laplacien de f)). Elles sont très importantes en physique où elles sont associées à tout type de phénomènes ondulatoires.

Théorème - Formule intégrale de Cauchy
Soit f une fonction holomorphe sur un ouvert U simplement connexe.
Soit $z_0 \in U$ et Γ un chemin fermé ne contenant pas z_0. Alors

$$f(z_0) \times \mathrm{Ind}_\Gamma(z_0) = \frac{1}{2i\pi} \oint_\Gamma \frac{f(z)}{z-z_0} \mathrm{d}z$$

Démonstration
On considère $g : z \mapsto \begin{cases} \frac{f(z)-f(z_0)}{z-z_0} & \text{si } z \neq z_0 \\ f'(z_0) & \text{si } z = z_0 \end{cases}$ définie sur U.
Alors g est continue et holomorphe sur $U \setminus \{z_0\}$ (au moins).
On a donc

$$0 = \oint_\Gamma g(z)\mathrm{d}z = \int_\Gamma \frac{f(z)}{z-z_0}\mathrm{d}z - f(z_0) \oint_\Gamma \frac{\mathrm{d}z}{z-z_0}$$

En multipliant par $2i\pi$, on reconnait l'indice de z_0 par rapport à Γ. □

Application - Interprétation nouvelle de $\int_0^{2\pi} f(z_0 + re^{i\theta})\mathrm{d}\theta$

Notons $\Gamma = \{z = z_0 + re^{i\theta}, \theta \in [0, 2\pi]\}$, $\gamma : \theta \mapsto z_0 + re^{i\theta}$.
Γ est un chemin fermé, γ en est un paramétrage.
Donc, pour tout f holomorphe sur U, ouvert contenant Γ,

$$f(z_0)\mathrm{Ind}_\Gamma(z_0) = \frac{1}{2i\pi} \oint_\Gamma \frac{f(z)}{z-z_0}\mathrm{d}z = \frac{1}{2i\pi} \int_0^{2\pi} \frac{f(\gamma(\theta)) \times ire^{i\theta}\mathrm{d}\theta}{re^{i\theta}} = \frac{i}{2i\pi} \int_0^{2\pi} f(z_0 + re^{i\theta})\mathrm{d}\theta$$

Donc $\int_0^{2\pi} f(z_0 + re^{i\theta})\mathrm{d}\theta = 2\pi \mathrm{Ind}_\Gamma(z_0) \times f(z_0)$ Mais, compte-tenu de la définition de Γ (cercle de centre z_0) : l'indice $\mathrm{Ind}_\Gamma(z_0) = \frac{1}{2i\pi} \int_0^{2\pi} \frac{rie^{i\theta}\mathrm{d}\theta}{re^{i\theta}} = \frac{1}{2\pi}$.
Donc $\int_0^{2\pi} f(z_0 + re^{i\theta})\mathrm{d}\theta = f(z_0)$

Analyse - Accès à $f^{(n)}(z_0)$

Si on veut avoir accès à la valeur de $f^{(n)}(z_0)$, on peut avoir intérêt à calculer

$$\oint_\Gamma \frac{f^{(n)}(z)}{z-z_0}\mathrm{d}z$$

On aurait envie de faire une intégration par parties : en intégrant $f^{(n)}$ et dérivant $\frac{1}{(z-z_0)}$ et en effet :

$$\oint_\Gamma \frac{f^{(n)}(z)}{z-z_0}\mathrm{d}z = \oint_\Gamma \frac{f^{(n-1)}(z)}{(z-z_0)^2}\mathrm{d}z = \cdots = \oint_\Gamma \frac{n! \times f(z)}{(z-z_0)^{n+1}}\mathrm{d}z$$

Pour les crochet, comme l'intégrale est sur un contour fermé, ils sont tous nuls. On retrouve une partie d'un résultat énoncé plus haut. Il faudrait ajouter que cela permet de voir comme une fonction analytique.

Théorème - Formule intégrale de Cauchy d'ordre n
Soient f est holomorphe sur U, et $z_0 \in U$.
Soit $r > 0$, tel que $\Gamma = \{z_0 + re^{i\theta}, \theta \in [0, 2\pi[\} \subset U$, alors f est infiniment dérivable en z_0 et

$$\forall\, n \in \mathbb{N}, \qquad f^{(n)}(z_0) = \frac{n!}{2\pi} \oint_\Gamma \frac{f(z)}{(z-z_0)^{n+1}}\mathrm{d}z = \frac{n!}{2r^n\pi} \int_0^{2\pi} f(z_0 + re^{i\theta})e^{-in\theta}\mathrm{d}\theta$$

Démonstration

Pour faire cette démonstration, considérons $a_n = \dfrac{n!}{2\pi}\oint_\Gamma \dfrac{f(u)}{(u-z_0)^{n+1}}\mathrm{d}u$.

Puis notons $g : z \mapsto \sum_{n=0}^{+\infty} a_n (z-z_0)^n$.

Soit z, à l'intérieur de Γ, i.e. $|z-z_0| < r$. On a alors pour tout $u \in \Gamma$ et $\rho := \dfrac{z-z_0}{u-z_0}$, $|\rho| < \dfrac{|z-z_0|}{r} < 1$.

Ainsi la série géométrique $\sum_{n\geqslant 0} \dfrac{1}{u-z_0}\rho^n$ converge vers $\dfrac{1}{u-z_0}\times\dfrac{1}{1-\rho} = \dfrac{1}{u-z}$.

Cette convergence est uniforme en u, donc on peut (théorème de seconde année) intervertir série et intégrale :

$$f(z_0) = \frac{1}{2i\pi}\oint_\Gamma \frac{f(u)}{u-z_0}\mathrm{d}u = \frac{1}{2i\pi}\oint_\Gamma \sum_{n=0}^{+\infty}\frac{(z-z_0)^n f(u)}{(u-z_0)^{n+1}}\mathrm{d}u = \frac{1}{2i\pi}\sum_{n=0}^{+\infty}(z-z_0)^n\oint_\Gamma\frac{f(u)}{(u-z_0)^{n+1}}\mathrm{d}u = \sum_{n=0}^{+\infty}\frac{a_n}{n!}(z-z_0)^n = g(z_0)$$

La fonction g est une série entière donc de classe $\mathscr{C}^\infty$ sur son disque de convergence.
g, donc f est de classe $\mathscr{C}^\infty$, puis on sait que le DSE de g est unique et donne $a_n = g^{(n)}(z_0) = f^{(n)}(z_0)$.
□

4.3. Petit détour : Formule de Green-Riemann

Voici un théorème qui sort du cadre des fonctions holomorphes, mais il permet d'avoir une nouvelle vision sur le théorème de Cauchy. Il fait aussi le lien avec le calcul vectoriel dont on abuse en physique en seconde année…
La force du théorème suivant : il permet de passer à des intégrales sur une surfaces (flux) à une intégrale sur le bord (circulation)…

Théorème - Théorème de Green-Riemann

Soient $P, Q : \mathbb{R}^2 \to \mathbb{C}$ des fonction $\mathbb{R}^2$-différentiables et K un compact suffisamment régulier, de frontière Γ. Alors

$$\oint_\Gamma (P\mathrm{d}x + Q\mathrm{d}y) = \iint_K \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right)\mathrm{d}x\mathrm{d}y$$

.

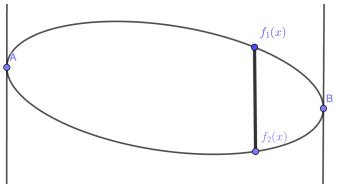
Pour aller plus loin - Théorème de Stokes

La formule de Stokes est le nom générique donné à toutes les formules du type $\int_\Omega \mathrm{d}\omega = \int_{\partial\Omega}\omega$ *qui dit que l'intégrale sur une surface est égale à la différence de la variation d'une certaine primitive.*
C'est le cas de la formule de Newton, ou des plus fameuses formule de Stokes ou formule d'Ostrograski vu en cours de physique de spé.

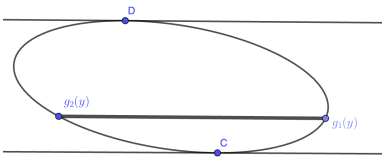
Remarque - Explication des termes

L'intégrable double, de droite, est une intégrale d'une variable du plan. On peut l'interpréter simplement en deux intégrations selon les deux variables x et y qui la composent (c'est ce qui est fait pour la pseudo-démonstration). Mais cela est trop restrictif (cf le calcul de flux en physique).
L'intégrable de droite et une double intégrale curviligne, on regarde en fonction de x et de y, séparément.
Il s'agit plus d'un éclairage que d'une démonstration. Par exemple le compact considéré sera considéré sans trou et avec une frontière biunivoque (à chaque y, correspond deux et seulement deux x sur Γ ; de même pour chaque x - excepté les quatre points de rebroussement).

Représentation - Représentation de K



Intégration $\mathrm{d}x$



Intégration $\mathrm{d}y$

Démonstration

On se place donc sur la figure de la marge.

On peut couper les deux intégrales en deux : $\oint_\Gamma P\mathrm{d}x = \iint_K -\dfrac{\partial P}{\partial y}\mathrm{d}x\mathrm{d}y$ et $\oint_\Gamma Q\mathrm{d}y = \iint_K \dfrac{\partial Q}{\partial x}\mathrm{d}x\mathrm{d}y$.

Pour la première, on prend les notations de la figure.

$$\begin{aligned}\iint_K -\frac{\partial P}{\partial y}\mathrm{d}x\mathrm{d}y &= -\int_a^b\left(\int_{f_2(x)}^{f_1(x)}\frac{\partial P}{\partial y}\mathrm{d}y\right)\mathrm{d}x = -\int_a^b P(x,f_1(x)) - P(x,f_2(x))\mathrm{d}x\\ &= \int_a^b P(x,f_2(x))\mathrm{d}x - \int_a^b P(x,f_2(x))\mathrm{d}x = \int_{[AB]_2} P\mathrm{d}x + \int_{[BA]_1} P\mathrm{d}x = \oint_\Gamma P\mathrm{d}x\end{aligned}$$

$$\begin{aligned}\iint_K \frac{\partial Q}{\partial x}\mathrm{d}x\mathrm{d}y &= -\int_c^d\left(\int_{g_2(y)}^{g_1(y)}\frac{\partial Q}{\partial x}\mathrm{d}x\right)\mathrm{d}y = \int_c^d Q(g_1(y),y) - Q(g_2(y),y)\mathrm{d}y\\ &= \int_c^d Q(y,g_1(y))\mathrm{d}y - \int_c^d Q(g_2(y),y)\mathrm{d}y = \int_{[CD]_1} Q\mathrm{d}y + \int_{[DC]_2} Q\mathrm{d}y = \oint_\Gamma Q\mathrm{d}y\end{aligned}$$

□

Application - (Première) formule de Cauchy

Soit f une fonction holomorphe sur U.

Soient P et Q tel que $P(x,y)=\text{Re}(f(x+iy))$ et $Q(x,y)=\text{Im}(f(x+iy))$.
Les formules de Cauchy-Riemann donne $\dfrac{\partial P}{\partial x}=\dfrac{\partial Q}{\partial y}$ et $\dfrac{\partial Q}{\partial x}=-\dfrac{\partial P}{\partial y}$.
Soit Γ, la frontière d'un compact K de U (donc Γ est un chemin fermé) et γ un paramétrage de Γ définit sur $[a,b]$. Supposons que $\gamma(t)=x(t)+iy(t)$, donc $\gamma'(t)=x(t)+iy(t)$

$$\begin{aligned}\oint_\Gamma f(z)\mathrm{d}z &= \int_a^b f(\gamma(t))\gamma'(t)\mathrm{d}t=\int_a^b (P(x(t),y(t))+iQ(x(t),y(t)))(x'(t)+iy'(t))\mathrm{d}t\\ &= \int_a^b P(x(t),y(t))x'(t)\mathrm{d}t-Q(x(t),y(t))y'(t)\mathrm{d}t\\ &\qquad +i\int_a^b P(x(t),y(t))y'(t)\mathrm{d}t+Q(x(t),y(t))y'(t)\mathrm{d}t\\ &= \oint_\Gamma P\mathrm{d}x-Q\mathrm{d}y+i\oint_\Gamma Q\mathrm{d}x+P\mathrm{d}y\\ &= \iint_K -\frac{\partial Q}{\partial x}-\frac{\partial P}{\partial y}\mathrm{d}x\mathrm{d}y+i\iint_K \frac{\partial P}{\partial x}-\frac{\partial Q}{\partial y}\mathrm{d}x\mathrm{d}y=\iint_K 0\mathrm{d}x\mathrm{d}y\end{aligned}$$

d'après les relations de Cauchy-Riemann, puisque f est holomorphe sur K.

Remarque - Interprétation et formule d'Ampère

Peut-être avez-vous rencontré le théorème d'Ampère : $\oint_{\mathscr{C}}\vec{H}\mathrm{d}\vec{\ell}=\iint_S\vec{j}\mathrm{d}\vec{S}$ ou sa version différentielle $\vec{\text{rot}}\vec{H}=\vec{\nabla}\wedge\vec{H}$.
Le premier terme s'appelle la circulation du champ vectoriel $\vec{H}$ sur le contour $\mathscr{C}$. Il s'agit d'une l'intégrale d'un produit scalaire, donc du type $(P,Q)\cdot(\mathrm{d}x,\mathrm{d}y)$, ce qui donne le premier terme de la formule de Green-Riemman.
On a alors pour le second terme un produit vectoriel (dimension 3) ou un déterminant (dimension 2) : $\begin{vmatrix}\frac{\partial}{\partial x} & P\\ \frac{\partial}{\partial y} & Q\end{vmatrix}=\dfrac{\partial Q}{\partial x}-\dfrac{\partial P}{\partial y}$ pour l'intégration sur une surface (appelé un flux, par les physiciens.

5. Théorème des résidus

5.1. Principe du théorème des résidus

Heuristique - Une idée pour calculer $\int_\Gamma f(z)\mathrm{d}z$, directement

Si on reprend la formule précédente, on voit que pour $n\in\mathbb{N}$, la valeur de $\dfrac{n!}{2\pi}\oint_\Gamma\dfrac{f(z)}{(z-z_0)^{n+1}}\mathrm{d}z$, donne accès à $f^{(n)}(z_0)$.
Et par ailleurs, ce nombre $f^{(n)}(z_0)$, ce trouve dans le développement de Taylor (série entière) $f(z)=\sum_{k=0}^{+\infty}\dfrac{f^{(k)}(z_0)}{k!}(z-z_0)^k$.
Mais, ce que l'on rencontre souvent c'est plutôt : $\int_\Gamma f(z)\mathrm{d}z$, comme si $n=-1$, dans le calcul précédent.
N'aurait-on pas : $\int_\Gamma f(z)\mathrm{d}z=\text{Ind}_\Gamma(z_0)f^{(-1)}(z_0)$?

Pour aller plus loin - Développement de Laurent

Le développement considéré ici s'appelle le développement de Laurent, il s'étend normalement sur un ouvert u de U contenant z_0, f étant holomorphe sur $u\setminus\{z\}$.

Définition - Résidu de f en z_0

On appelle résidu de f au point z_0 de $\mathbb{C}$, noté $\text{Res}(f,z_0)$, le coefficient devant $(z-z_0)^{-1}$ dans le développement asymptotique de f au voisinage de z_0.

Exemple - Fraction $F(z)=\dfrac{z^2+z-1}{z^4-z^3-z+1}$

Comme $z^4-z^3-z+1=(z^2+z+1)(z^2-2z+1)$, cette fraction admet trois pôles, deux sont simples : j et j^2 et un est double : 1.
On a la décomposition en éléments simples

$$F(z)=\frac{j^2+j-1}{j^3-3j^2-j}=\frac{c}{z-j}+\frac{\overline{c}}{z-j^2}+\frac{a}{z-1}+\frac{b}{(z-1)^2}=\frac{j-1}{3(z-j)}+\frac{j^2-1}{3(z-j^2)}-\frac{1}{z-1}+\frac{1}{3(z-1)^2}$$

car $c = \dfrac{j^2+j-1}{j^3-3j^2-j} = \dfrac{-2}{2-2j^2} = \dfrac{1}{j^2-1} = \dfrac{j-1}{3}$, $b = \dfrac{1+1-1}{1+1+1} = \dfrac{1}{3}$ et $a+c+\overline{c} = \lim_{z\to+\infty} zf(z) = 0$, donc $a = -\frac{1}{3}$.

On a donc pour tout $z \notin \{1, j, j^2\}$, $\mathrm{Res}(f,z) = 0$, $\mathrm{Res}(f,j) = \dfrac{j-1}{3}$, $\mathrm{Res}(f,j^2) = \dfrac{j^2-1}{3}$, et $\mathrm{Res}(f,1) = \dfrac{-1}{3}$. En effet :

$$\begin{aligned} F(z) = & \frac{1}{3(z-1)^2} - \frac{1}{3(z-1)} + \frac{j-1}{3j}\sum_{k=0}^{+\infty}\left(\frac{z}{j}\right)^k + \frac{j^2-1}{3j^2}\sum_{k=0}^{+\infty}\left(\frac{z}{j^2}\right)^k \\ = & \frac{1}{3(z-1)^2} - \frac{1}{3(z-1)} + \sum_{h=0}^{+\infty} b_h (z-1)^h \end{aligned}$$

Heuristique - Extension : des polynômes aux fractions rationnelles. Des séries entières, au séries de Laurent.

En première partie, nous avons vu que les fonctions holomorphes sont la bonne façon de voir les séries entières, extension degré infini des polynômes.
Or les polynômes connaissent une autre extension intéressante (anneau→corps) : les fractions rationnelles, avec la notion de pôles et de degré négatif. Les séries de Laurent sont localement (autour des pôles ou des points normaux) les extensions naturelles de fractions rationnelles. Il est important que les pôles soient isolés.
Quant aux fonctions holomorphes avec des pôles, elles sont appelées : fonctions méromorphes. Les pôles sont isolés.

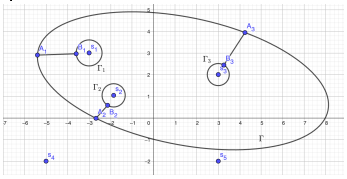
Théorème - Résidus
Soit f une fonction holomorphe sur U, un ouvert étoilé, présentant des singularités en des points isolés de $S = \{s_1, \dots s_m\} \subset U$.
Soit Γ un chemin fermé tracé dans U, ne rencontrant pas S. Alors

$$\int_\Gamma f(z)\mathrm{d}z = 2i\pi \sum_{i=1}^{m} \mathrm{Ind}_\Gamma(s_i) \times \mathrm{Res}(f, s_i)$$

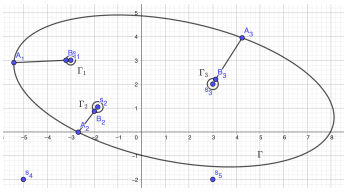
Pour aller plus loin - Zéros isolés
Nous verrons plus loin (classification des zéros) que si f n'est pas nulle, alors ses zéros sont isolés et au plus dénombrables.
On reviendra sur ce résultat profond d'analycité avec le théorème de prolongement analytique.

Ce théorème a de multiplies applications que nous verrons plus loin.
Plutôt qu'une démonstration complète, nous nous contenterons d'un dessin animé commenté…

Représentation - Théorème des résidus



Décomposition de Γ, sans singularité.



Avec $n \to +\infty$.

Démonstration
En notant, pour tout $n \in \mathbb{N}^*$ et $i \in \mathbb{N}_m$, $\Gamma_i^n = \mathscr{C}(s_i, \frac{1}{n})$, le cercle centré en s_i de rayon $\frac{1}{n}$, tournant dans le sens inverse de Γ, ainsi que $[A_i B_i^n]$, le segment joignant Γ à Γ_i^n.
Enfin, on considère le chemin

$$\Gamma^n = [A_1 \to B_1^n \to \Gamma_i^n \to B_1^n \to A_1] \underset{\text{sur } \Gamma}{\to} [A_2 \to B_2^n \to \Gamma_i^n \to B_2^n \to A_2]$$

$$\underset{\text{sur } \Gamma}{\to} \cdots \underset{\text{sur } \Gamma}{\to} [A_m \to B_m^n \to \Gamma_i^n \to B_m^n \to A_m] \underset{\text{sur } \Gamma}{\to} [A_1]$$

A l'intérieur de ce chemin, f n'a pas de singularité, donc son intégrale est nulle : $\oint_{\Gamma^n} f(z)\mathrm{d}z = 0$.
Or on peut recoller l'intégrale et comme sur les chemins $\int_{A_iB_i^n} f(z)\mathrm{d}z + \int_{B_i^nA_i} f(z)\mathrm{d}z = 0$, on trouve

$$0 = \int_{\Gamma^n} f(z)\mathrm{d}z = \int_\Gamma f(z)\mathrm{d}z + \sum_{i=1}^{m} \oint_{\Gamma_i^n} f(z)\mathrm{d}z$$

$$\oint_\Gamma f(z)\mathrm{d}z = \sum_{i=1}^{m} \oint_{-\Gamma_i^n} f(z)\mathrm{d}z$$

Pour tout $\epsilon > 0$, pour n suffisamment grand (i.e. Γ_i^n proche de s_i), en remplaçant f par son développement analytique local :

$$\left| \oint_\Gamma f(z)\mathrm{d}z - \sum_{i=1}^{m} \oint_{-\Gamma_i^n} \sum_{h=-\infty}^{+\infty} a_h(s_i)(z-s_i)^h \mathrm{d}z \right| = \left| \oint_\Gamma f(z)\mathrm{d}z - \sum_{i=1}^{m} (2i\pi)\mathrm{Ind}_{-\Gamma_i^n}(s_i) \times a_{-1}(s_i) \right| \leqslant \epsilon$$

Ainsi, en passant à la limite sur n : $\oint_\Gamma f(z)\mathrm{d}z = \sum_{i=1}^{m} 2i\pi \mathrm{Ind}_{-\Gamma_i^n}(s_i) \times a_{-1}(s_i)$ Il faut donc tourner dans le même sens entre Γ et Γ_i.
Notons ici que la démonstration sous-entend un seul tour Γ, mais elle s'adapte à plusieurs tours… □

5.2. Calculer les résidus

On admet :

Proposition - Linéarité du résidu
Si f et g sont méromorphes et $\lambda, \mu \in \mathbb{C}$,
$\operatorname{Res}(\lambda f + \mu g, a) = \lambda \operatorname{Res}(f, a) + \mu \operatorname{Res}(g, a)$.

La méthode du calcul connue pour les pôles de fractions s'adaptent aux fonctions holomorphes plus compliquées :

Savoir faire - Calculer les résidus de f en a
Plusieurs méthode :

1. On exploite un développement asymptotique de f au voisinage de a.
2. Si f a en a un pôle d'ordre 1 : $\operatorname{Res}(f, a) = \lim_{z \to a} (z-a) f(z)$.
3. Si f a en a un pôle d'ordre n : $\operatorname{Res}(f, a) = \dfrac{1}{(n-1)!} \lim_{z \to a} \dfrac{\partial^{n-1}}{\partial z^{n-1}} \big((z-a)^n f(z)\big)$.
4. Si h possède en a une racine d'ordre 1 : $\operatorname{Res}\left(\frac{g}{h}, a\right) = \dfrac{g(a)}{h'(a)}$.

Remarque - Explication du second résultat
Supposons que $f(z) = \sum_{k=1}^{n} \dfrac{\alpha_k}{(z-a)^k} + g(z)$ où a n'est pas pôle de g. On cherche α_1.
En multipliant par $(z-a)^n$, on trouve :
$(z-a)^n f(z) = \sum_{k=1}^{n} \alpha_k (z-a)^{n-k} + (z-a)^n g(z)$ puis
$\dfrac{\partial^{n-1}}{\partial z^{n-1}}\big((z-a)^n f(z)\big) = (n-1)!\alpha_1 + 0 + \cdots + 0 + \sum_{k=0}^{n-1} \binom{n-1}{k} \dfrac{n!}{(n-k)!} (z-a)^{n-k} g^{(n-k)}(z)$.

Exercice
Que vaut $\operatorname{Res}\left(\frac{\cos t}{t}, 0\right)$?

Correction
$\dfrac{\cos t}{t} = \dfrac{1}{t} - \dfrac{t}{2} + \dfrac{t^3}{24} + \ldots$ Donc $\operatorname{Res}\left(\frac{\cos t}{t}, 0\right) = 1$.

5.3. Applications (aux calculs d'intégrales immondes et sommes effrayantes)

Le principe est (presque) toujours le même : on prolonge l'intégrande en une fonction holomorphe sur une partie $\mathbb{C}$. Puis on calcul ses résidus.

Pour aller plus loin - Calculer des sommes par les résidus
On peut aussi exploiter le théorème des résidus pour évaluer quelques sommes du type $\sum_{n \in \mathbb{Z} \setminus S} f(n)$ ou $\sum_{n \in \mathbb{Z} \setminus S} (-1)^n f(n)$. Voir la page Théorème des résidus de Wikipedia.

Savoir faire - Intégrale de premier type
Il s'agit de calculer $I = \displaystyle\int_0^{2\pi} R(\cos t, \sin t)\mathrm{d}t$ où R est une fraction rationnelle ayant un nombre fini de pôle : s_j et donc aucun n'appartient à $\mathbb{U} = \mathscr{C}(0,1)$, le cercle unité.
Alors $I = 2i\pi \sum_{|s_j|<1} \operatorname{Res}(f, s_z)$ où $f(z) = \dfrac{1}{iz} R\left(\frac{z+z^{-1}}{2}, \frac{z-z^{-1}}{2i}\right)$.

Exemple - $\displaystyle\int_0^{2\pi} \frac{\mathrm{d}x}{a + \sin x}$ avec $a > 1$.
La fonction $R : (x, y) \mapsto \dfrac{1}{a+y}$, on a bien l'intégrande qui est $t \mapsto R(\cos t, \sin t)$.
Posons z de tel sorte que t vérifie $z = e^{it} = \gamma(t)$, donc $\cos t = \frac{z+z^{-1}}{2}$ et $\sin t = \frac{z-z^{-1}}{2i}$ et $\gamma'(t) = i\gamma(t) = iz$. $f(z) = \dfrac{1}{iz} R\left(\frac{z+z^{-1}}{2}, \frac{z-z^{-1}}{2i}\right) = \dfrac{1}{iz} \dfrac{1}{a + \frac{z-z^{-1}}{2i}} = \dfrac{2}{z^2 + 2iaz - 1}$ et

$$\int_0^{2\pi} \frac{\mathrm{d}x}{a + \sin x} = \oint_\Gamma f(z)\mathrm{d}z = 2i\pi \sum_{s \in S} \operatorname{Ind}_\Gamma(s) \operatorname{Res}(f, s)$$

Or $z^2+2iaz-1=(z+ia)^2+(a^2-1)=(z+ia-i\sqrt{a^2-1})(z+ia+i\sqrt{a^2-1})$.
Les pôles sont donc $i(-a-\sqrt{a^2-1})$ de module $a+\sqrt{a^2-1}>1$,
$i(-a+\sqrt{a^2-1})$ de module $a-\sqrt{a^2-1}\in]0,1]$ car $a>1$. La fonction f ne présente qu'une singularité dans $\mathbb{U}$ en $i(-a+\sqrt{a^2-1})$ (l'autre point à un indice nul).
Le cercle $\Gamma=\mathbb{U}$ fait un tour autour de ce point donc $\text{Ind}_\Gamma(i(-a+\sqrt{a^2-1}))=1$.
Puis $\text{Res}(f,i(-a+\sqrt{a^2-1}))=\left[(z+ia-i\sqrt{a^2-1})f(z)\right]_{\cdot}^{i(-a+\sqrt{a^2-1})}=\dfrac{1}{i\sqrt{a^2-1}}$.
Donc

$$\int_0^{2\pi}\frac{\mathrm{d}x}{a+\sin x}=\oint_\Gamma f(z)\mathrm{d}z=2i\pi\times 1\times\frac{1}{i\sqrt{a^2-1}}=-\frac{2\pi}{\sqrt{a^2-1}}$$

Savoir faire - Intégrale de deuxième type

Il s'agit de calculer $I=\displaystyle\int_{-\infty}^{+\infty}f(x)\mathrm{d}x$.
On suppose que cette intégrale est convergente, ce qui impose nécessairement que $zf(z)\underset{|z|\to\infty}{\longrightarrow}0$.
En considérant le « demi-cercle » de diamètre l'axe des réels et la partie positive (resp. négative), on trouve :

$$I=2i\pi\sum_{\text{Im}(s_j)>0}\text{Res}(f,s_z)=-2i\pi\sum_{\text{Im}(s_j)<0}\text{Res}(f,s_z).$$

Exemple - $\displaystyle\int_{-\infty}^{+\infty}\frac{\mathrm{d}x}{x^2+a^2}$ avec $a>0$.

L'idée est de ne pas exploiter arctan.
On considère Γ_r, le demi-cercle de centre 0, de rayon r de partie imaginaire positive.
Un paramètrage de Γ_r est par exemple $\gamma_r:[0,\pi+2]\to\mathbb{C}$ tel que $\gamma_r(t)=re^{it}$ si $t\in[0,\pi]$, puis $\gamma_r(t)=-r+(t-\pi)r$ sur $[\pi,\pi+2]$.
$f:z\mapsto\dfrac{1}{z^2+a^2}=\dfrac{1}{2ia(z-ia)}-\dfrac{1}{2ia(z+ia)}$. Si $r>a$, f admet un seul pôle dans Γ_r : le point ia.
On a alors $\text{Ind}_{\Gamma_r}(ia)=1$ et $\text{Res}(f,ia)=\dfrac{1}{2ia}$. Alors

$$\oint_{\Gamma_r}f(z)\mathrm{d}z=2i\pi\times 1\times\frac{1}{2ia}=\frac{\pi}{a}$$

Par ailleurs, pour tout $\epsilon>0$, il existe r tel que $|f(z)|z||<\epsilon$ pour $|z|>r$, donc :

$$\oint_{\Gamma_r}f(z)\mathrm{d}z=\int_0^\pi f(re^{it})rie^{it}\mathrm{d}t+\int_\pi^{\pi+2}f(-r+(t-\pi)r)r\mathrm{d}t$$

$$\left|\oint_{\Gamma_r}f(z)\mathrm{d}z-\underbrace{\int_{-r}^{r}f(u)\mathrm{d}u}_{u=-r+(t-\pi)r}\right|<\epsilon\times\pi$$

Ainsi pour r suffisamment grand : $\displaystyle\int_{-r}^{r}f(u)\mathrm{d}u\underset{r\to+\infty}{\longrightarrow}\frac{\pi}{a}$ et $\displaystyle\int_{-\infty}^{+\infty}\frac{\mathrm{d}x}{x^2+a^2}=\frac{\pi}{a}$
Pour le dernier savoir-faire, on fera attention au moment du calcul du résidu que la fonction $z\mapsto f(z)e^{iaz}$ n'est pas, en règle générale, une fraction rationnelle.

Savoir faire - Intégrale de troisième type

Il s'agit de calculer $I=\displaystyle\int_{-\infty}^{+\infty}f(x)e^{iax}\mathrm{d}x$.
On suppose que cette intégrale est convergente, ce qui impose nécessairement que $zf(z)\underset{|z|\to\infty}{\longrightarrow}0$. On suppose également que $S\subset\mathbb{C}\setminus\mathbb{R}$ (pas de points singuliers (pôles) réels).
En considérant le « demi-cercle » de diamètre l'axe des réels et la partie positive (resp. négative), on trouve :

$$I=2i\pi\sum_{\text{Im}(s_j)>0}\text{Res}(fe^{ia\cdot},s_z)\text{ si }a>0.$$

$$I = -2i\pi \sum_{\mathrm{Im}(s_j)<0} \mathrm{Res}(fe^{ia\cdot}, s_z) \text{ si } a<0 \text{ (resp.).}$$

Souvent on exploite ce savoir-faire à une recherche de cos ou sin au numérateur (en prenant la partie réelle ou imaginaire).

Exemple - $\displaystyle\int_{-\infty}^{+\infty} \frac{\cos(bx)\mathrm{d}x}{x^2+a^2}$ **avec** $a, b>0$**.**

L'idée est considérer $x \mapsto \dfrac{e^{bx}}{x^2+b^2}$, puis de prendre la partie réelle.

On considère de nouveau Γ_r, le demi-cercle de centre 0, de rayon r de partie imaginaire positive. Un paramètrage de Γ_r est par exemple $\gamma_r : [0,\pi+2] \to \mathbb{C}$ tel que $\gamma_r(t) = re^{it}$ si $t \in [0,\pi]$, puis $\gamma_r(t) = -r+(t-\pi)r$ sur $[\pi,\pi+2]$.

$f : z \mapsto \dfrac{1}{z^2+a^2}$ admet deux pôles ia et donc si $r>a$, f admet un seul pôle dans Γ_r : le point ia.

On a alors $\mathrm{Ind}_{\Gamma_r}(ia) = 1$ et $\mathrm{Res}(g, ia) = \lim\limits_{z\to ia}(z-ia)\dfrac{e^{ibz}}{z^2+a^2} = \dfrac{e^{-ab}}{2ia}$

$$\oint_{\Gamma_r} f(z)\mathrm{d}z = 2i\pi \times 1 \times \frac{e^{-ab}}{2ia} = \frac{\pi e^{-ab}}{a}$$

Par ailleurs, pour tout $\epsilon>0$, il existe r tel que $f(z)|z|<\epsilon$ pour $|z|>r$, donc :

$$\oint_{\Gamma_r} f(z)\mathrm{d}z = \int_0^{\pi} f(re^{it})e^{ibre^{it}} r\mathrm{d}t + \int_{-r}^{r} f(u)e^{ibu}\mathrm{d}u$$

Ainsi pour r suffisamment grand : $\displaystyle\int_{-r}^{r} f(u)\mathrm{d}u \underset{r\to+\infty}{\longrightarrow} \frac{\pi e^{-ab}}{a}$ et $\displaystyle\int_{-\infty}^{+\infty} \frac{\mathrm{d}x}{x^2+a^2} = \frac{\pi e^{-ab}}{a}$. Rest à prendre la partie réelle : $\displaystyle\int_{-\infty}^{+\infty} \frac{\cos(bx)\mathrm{d}x}{x^2+a^2} = \frac{\pi e^{-ab}}{a}$

6. Prolongement analytique

6.1. Zéros d'une fonction holomorphe

Avant d'étudier le théorème de résidus, nous avons rencontré les fonctions méromorphes. Il s'agit de fonctions holomorphes sauf en un nombre finis de point de $\mathbb{C}$. L'étude des pôles et l'étude des racines d'une fonction holomorphe sont certainement intéressants.

Remarque - Rappels topologique 1 : points d'accumulation et points isolés

On rappelle qu'un points adhérent est la limite d'une suite de points d'un ensemble.

Les points d'accumulation a sont des points adhérents par une suite de points de A, tous différent du point limite.

$\exists\, (a_n) \in (A\setminus\{a\})^{\mathbb{N}}$ telle que $(a_n) \to a$.

Ou encore : $\forall\, \epsilon>0, \exists\, a_\epsilon \in A$ tel que $0<|a-a_\epsilon|<\epsilon$.

Les points isolés sont des points qui ne sont pas des points d'accumulation

$\exists\, \epsilon>0$ tel que $B(x,\epsilon)\cap A = \{x\}$.

Remarque - Rappels topologique 2 : parties connexes

On a vu que E est connexe (dans $\mathbb{R}$), si on ne peut pas l'écrire comme réunion de deux sous-ensembles séparés non vide.

On n'a pas $E = A\cap B$ avec $\overline{A}\cap B = \emptyset$ et $A\cap\overline{B} = \emptyset$.

Une méthode pour démontrer que E est connexe (par arcs) consiste à montrer que pour tout $a, b \in E$, il existe $\varphi : [0,1] \to [a,b]$ continue avec $\varphi(0) = a$ et $\varphi(1) = b$.

Une autre méthode consiste souvent à faire un raisonnement par l'absurde et à travailler à partir du nombre x_0 qui est obtenu comme borne supérieure d'un ensemble A (à inventer) et élément de B ou bien élément de A et borne inférieure de B. A partir de ce x_0, trouver une contradiction.

Analyse - Factorisation par $(z-a)^m$

On sait que f est holomorphe donc analytique sur U.

Soit $a \in Z_f$, f admet un développement autour de a,

$\exists\, r>0$ tel que $\forall\, z\in B(a,r)$ (U est un ouvert) : $f(z)=\sum_{n=0}^{+\infty} a_n(z-a)^n$.

$a\in Z_f$, donc $f(a)=0=a_0$. Soit $K=\{n\in\mathbb{N}^*\mid a_n\neq 0\}\subset\mathbb{N}$.

• Ou bien K est majoré, et donc K admet un plus grand élément >0 : $m(a)-1$.

On a au voisinage de a : $f(z)=\sum_{n=m(a)}^{+\infty} a_n(z-a)^n=(z-a)^{m(a)}\sum_{n=0}^{+\infty} a_{n+m}(z-a)^n$.

La fonction $g:z\mapsto\dfrac{f(z)}{(z-a)^{m(a)}}$ est bien définie au voisinage de a et sur $U\setminus\{a\}$.

Et f ne s'annule qu'en a sur le voisinage de a.

Le nombre de zéros de f est au plus dénombrable.

• Ou bien K n'est pas majoré et donc $f=0$ est identiquement nulle au voisinage de a.

Ce point a est un point d'accumulation.

Le théorème suivant reprend cette idée, mais il annonce plus : s'il existe un point d'accumulation, alors f est nécessairement nulle sur tout le connexe.

Notons $Z_f=\{a\in U\mid f(a)=0\}$, l'ensemble des zéros de f.
Puisque f est holomorphe en a, elle est développable en série entière autour de a

Définition - Ordre des zéros de f
Soit U un ouvert connexe de $\mathbb{C}$ et f une fonction holomorphe définie sur U.
Soit $a\in Z_f$.

— Ou bien : $\forall\, n\in\mathbb{N}$, $f^{(n)}(a)=0$ Dans ce cas il existe V voisinage de a tel que $\forall\, x\in V\setminus\{a\}$, $f(x)=0$.

— Ou bien : il existe $m(a)\in\mathbb{N}^*$ et $g\in\mathscr{H}$ tel que $g(a)\neq 0$ et $\forall\, z\in U$, $f(z)=g(z)(z-a)^{m(a)}$.
Le nombre $m(a)$ s'appelle l'ordre du zéro de f au point a.
Dans ce cas il existe V voisinage de a tel que $\forall\, x\in V\setminus\{a\}$, $f(x)\neq 0$.

Proposition - Zéros isolés
Soit U un ouvert connexe de $\mathbb{C}$. Soit f holomorphe sur U.
S'il existe $a\in Z_f$, d'ordre infini, alors f est identiquement nulle sur U.

Démonstration
Supposons que a est d'ordre infini. Donc pour tout $n\in\mathbb{N}$, $f^{(n)}(a)=0$.
Posons $A=\{x\in U\mid\forall\, n\in\mathbb{N}, f^{(n)}(x)=0\}$ A est non vide, puisque $a\in A$. Soit $b\in U$, quelconque et $\gamma:[0,1]\to[a,b]$ un chemin (non nécessairement droit de a à b dans le connexe U).
$\gamma(0)\in A$. Notons $T=\sup\{t\in[0,1]\mid\gamma(t)\in A\}$.
Il existe une suite (t_n) d'éléments de $[0,1]$ tel que $(t_n)\to T$ et $\gamma(t_n)\in A$.
Comme f est continue, ainsi que toutes ses dérivées : $\forall\, k\in\mathbb{N}$, $f^{(k)}(t_n)=0\to f^{(k)}(T)$. Donc $\gamma(T)\in A$.
Ainsi $\gamma(T)$ est un zéro d'ordre infini de f. Il existe une voisinage de $\gamma(T)$ sur lequel f est nul.
Comme T est une borne supérieure, cela signifie que $T=1$, nécessairement et donc $f(b)=0$. □

6.2. Prolongement analytique

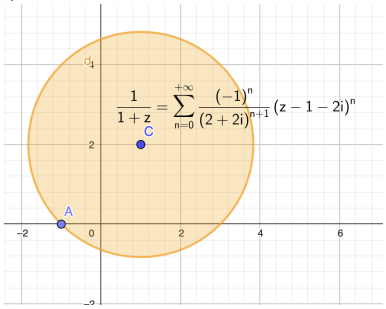
Analyse - L'application $z\mapsto\dfrac{1}{1+z}$

C'est bien connu, $f(z)=\sum_{k=0}^{+\infty}(-1)^k z^k$, pour $z\in\mathscr{D}(0,1)$ (ouvert).
Peut-on étendre plus? Ici, on est centré en 0 et le rayon est $R=1$, à cause du pôle en -1.
Et centré en a quelconque (dans $\mathbb{C}$? On pose $z=a+h$, on a

$$\frac{1}{1+z}=\frac{1}{1+a+h}=\frac{1}{1+a}\times\frac{1}{1+\frac{h}{1+a}}=\frac{1}{1+a}\sum_{n=0}^{+\infty}\left(\frac{-1}{1+a}\right)^n h^n=\sum_{n=0}^{+\infty}\frac{(-1)^n}{(1+a)^{n+1}}(z-a)^n$$

On a trouvé le développement en série analytique de $\frac{1}{1+z}$ au point a.
La série obtenue est géométrique, elle converge pour tout z tel que $\left|\frac{z-a}{1+a}\right|<1\iff z\in$

Représentation - DES en $1+2i$



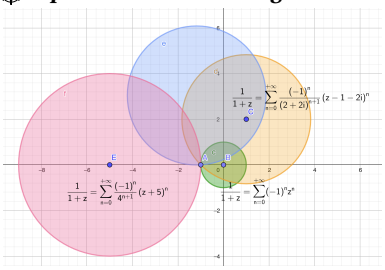
$\mathscr{B}(a,1+a)$ En fait, le rayon $R=a+1$ peut s'interpréter comme la valeur maximale de manière à ne pas avoir $z=-1$. Il est certain que $\frac{1}{1+z}$ n'est pas défini en -1.

Proposition - Prolongement analytique
Soient f et g deux fonctions holomorphes ou analytiques.
Soit U un ouvert connexe contenant $\mathscr{D}_f$ et $\mathscr{D}_g$.
Si $\{z\in U \mid f(z)=g(z)\}$ admet une point d'accumulation, alors $f=g$ sur U entier.

Remarque - Domaine d'égalité
On exploite ce théorème souvent sur $U=\mathscr{D}_f\cap\mathscr{D}_g$ directement (à condition que cet ensemble soit (ouvert et) connexe.

Démonstration
Considérons $h=f-g$, alors Z_h admet un point d'accumulation, on a donc un zéro non isolé.
Par conséquent $h=0$ sur tout U, i.e : $f_{|U}=g_{|U}$. □

Représentation - Prolongement analytique



Application - $z\mapsto\dfrac{1}{1+z}$
On peut passer d'une zone à une autre par connexité.
Cela permet aussi d'élargir la définition d'une fonction. En effet, il n'existe en fait qu'une seule fonction holomorphe ayant un développement particulier en un point.
Dans le cadre des fonctions holomorphes, le rêve de Cauchy se réalise!

Exemple - $z\mapsto e^{-1/z^2}$
Cette fonction f vérifie pour tout $z\in\mathbb{C}^*$, $f^{(n)}(z)\mapsto_0 0$ et pourtant $f=\neq 0$.
0 est la fonction holomorphe qui s'annule ainsi que toutes ses dérivées en 0.
f n'est nécessairement pas holomorphe.

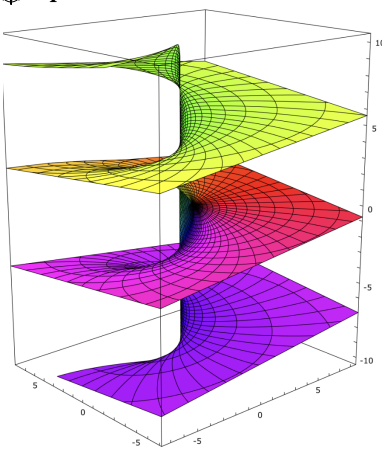
6.3. Fonctions à singularités. Vers les surfaces de Riemann

On termine par un preview de la saison 2.

Heuristique - Tourner autour d'un pôle

> Lorsqu'on fait le tour autour de 0 de la fonction $z\mapsto\dfrac{1}{1+z}$, on retrouve la même série entière.
> La fonction $z\mapsto\dfrac{1}{1+z}$ est univariée.
> Il n'en est pas de même de son intégrale : $\displaystyle\int_\Gamma\frac{dz}{1+z}$

Représentation - $z\mapsto\ln(z)$



Analyse - $z\mapsto\ln z$
La fonction $z\mapsto\ln z$ qui est « la » primitive de $\frac{1}{z}$ ne peut être définie En tournant autour du pôle, un décallage de $2i\pi$ se mesure.
La fonction ln est en fait multivariée : $\forall\, k\in\mathbb{Z}, \exp(\ln a)=\exp(\ln(a)+2ik\pi)$ Riemann propose un concept pour dépasser cette limite, a priori.

Définition - Surface de Riemann
Une surface de Riemann est un espace topologique séparé X, admettant un atlas modelé sur le plan complexe $\mathbb{C}$ dont les applications de changement de cartes sont des applications biholomorphes.
Autrement dit X admet un recouvrement par des ouverts U_i homéomorphes à des ouverts de $\mathbb{C}$; ces cartes dites holomorphes $f_i:U_i\to V_i$ sont telles que les fonctions de changement de cartes $f_i\circ f_j^{-1}$ soient des fonctions holomorphes entre ouverts de $\mathbb{C}$.

En fait, en géométrie différentielle et géométrie analytique complexe, une surface de Riemann est une variété complexe de dimension 1. Vivement la prochaine saison!

7. Bilan

Synthèse

- ⇝ Une fonction holomorphe sur un ouvert U est une fonction dérivable en tout $z_0 \in U$, de la variable complexe. Cette forme de dérivation est plus exigeante que la dérivation à deux variables : localement la transformation n'est pas simplement affine, elle est conforme (similitude). Il existe de nombreuses fonctions holomorphes (toutes nos fonctions usuelles, entre autres).
- ⇝ Cette rigidité donne aux fonctions de a-la variable complexe une propriété impensable pour les fonctions réelles : si elles sont une fois dérivable (holomorphe) alors elles sont infiniment dérivable, et même analytique. Cela signifie qu'en tout point de u, la fonction holomorphe est comparable à une série entière (centrée en ce point) de rayon > 0.
- ⇝ Pour montrer ce résultat important, il a d'abord fallut faire un détour par les intégrales curvilignes, i.e. les intégrales sur des chemins (lignes) de $\mathbb{C}$. Là, une foule de théorème, tous plus ou moins dûs à Cauchy s'observent : ils lient les intégrales de chemin d'une fonction aux valeurs aux extrémités de la primitive (comme toute intgrale) mais également aux nombres de points singuliers (pôles) de la fonction intégrée à l'intérieur du chemin d'intégration (indice d'un point par rapport à une courbe fermée). En découle des tas de théorèmes comme le principe du maximum que nous verrons en TD.
- ⇝ Une première application a été le théorème de résidus qui permet, lorsqu'on le maitrise bien de calculer les valeurs exactes d'intégrales difficiles voire de sommes infinies immondes.
- ⇝ Une seconde application est celui du recollement par prolongement analytique sur un convexe. En fait, les éeros d'une fonction holomorphe sont soient isolés soit dense. Ou encore : il n'existe qu'une fonction holomorphe sur $\mathbb{C} \setminus E$ (où E est au plus dénombrable) connaissant la valeur de toutes les dérivées en un point.

Savoir-faire et Truc & Astuce du chapitre

- — Savoir-faire - Montrer l'holomorphie (ou non) avec $H : \mathbb{R}^2 \to \mathbb{C}$, $(x, y) \mapsto f(x + iy)$
- — Savoir-faire - Démontrer qu'une fonction n'est pas la dérivée d'une fonction holomorphe
- — Savoir-faire - Calculer les résidus de f en a
- — Savoir-faire - Intégrale de premier type (Résidus)
- — Savoir-faire - Intégrale de deuxième type (Résidus)
- — Savoir-faire - Intégrale de troisième type (Résidus)

Notations

Notations	*Définitions*	*Propriétés*	*Remarques*
$\mathscr{H}(U)$	Ensemble des fonctions holomorphe définies sur un ouvert $U \subset \mathbb{C}$	$\forall\, z_0 \in U\ f(z) = f(z_0) + A(z - z_0) + (z - z_0)\epsilon(z)$ où $A \in \mathbb{C}$ (similitude) et $\epsilon(z) \underset{z \to z_0}{\longrightarrow} 0$	$(\mathscr{H}(U), \cdot, +, \times)$ est une algèbre
$\int_\Gamma g(z)\mathrm{d}z$	Intégrale curvilgne de g le long de Γ	Si $\Gamma = \mathrm{Im}\,(\gamma)$ avec $\gamma : [a, b] \to \mathbb{C}$, de classe $\mathscr{C}^1$ sans point stationnaire : $\int_\Gamma g(z)\mathrm{d}z = \int_a^b g(\gamma(t))\gamma'(t)\mathrm{d}t$	Ce nombre est indépendant du choix de γ.
$\mathrm{Ind}_\Gamma(z_0)$	Indice de Γ par rapport à z_0	$\mathrm{Ind}_\Gamma(z_0) = \frac{1}{2i\pi}\oint_\Gamma \frac{\mathrm{d}z}{z - z_0} \in \mathbb{Z}$	Il indique le nombre de tour (sens direct) de Γ autour de z_0
$\mathrm{Res}(f, z_0)$	Résidu de f en z_0	Le coefficient devant $(z - z_0)^{-1}$ dans le DAS de f au voisinage de z_0	Si z_0 est pôle d'ordre n, $\mathrm{Res}(f, z_0) = \frac{1}{(n-1)!}\lim_{z \to z_0} \frac{\partial^{n-1}}{\partial z^{n-1}}((z - z_0)^n f(z))$

Retour sur les problèmes

A.4 C'est l'enjeu de ce cours

A.5 Intégration curviligne (sur une ligne ou un chemin), ou bien intégration sur une surface comme pour le théorème de Green-Riemann

A.6 Si P ne s'annule pas alors $\frac{1}{P}$ est holomorphe et verifie donc le principe de maximum, . Voir TD

A.7 Question avancée dans la partie sur le prolongement analytique.

A.8 Oui, la relation $\Gamma(z+1) = z\Gamma(z)$ permet de définir pour $z \in \mathbb{R}^-$ et $z \notin \mathbb{Z}$, $\Gamma(z) = \dfrac{1}{(z+1)(z+2)\cdots(z+\lfloor z \rfloor)}\Gamma(z+\lceil z \rceil)$, il n'y a qu'une fonction holomorphe qui égale avec Γ sur le connexe $\mathbb{R}_+^*$.

Table de matières