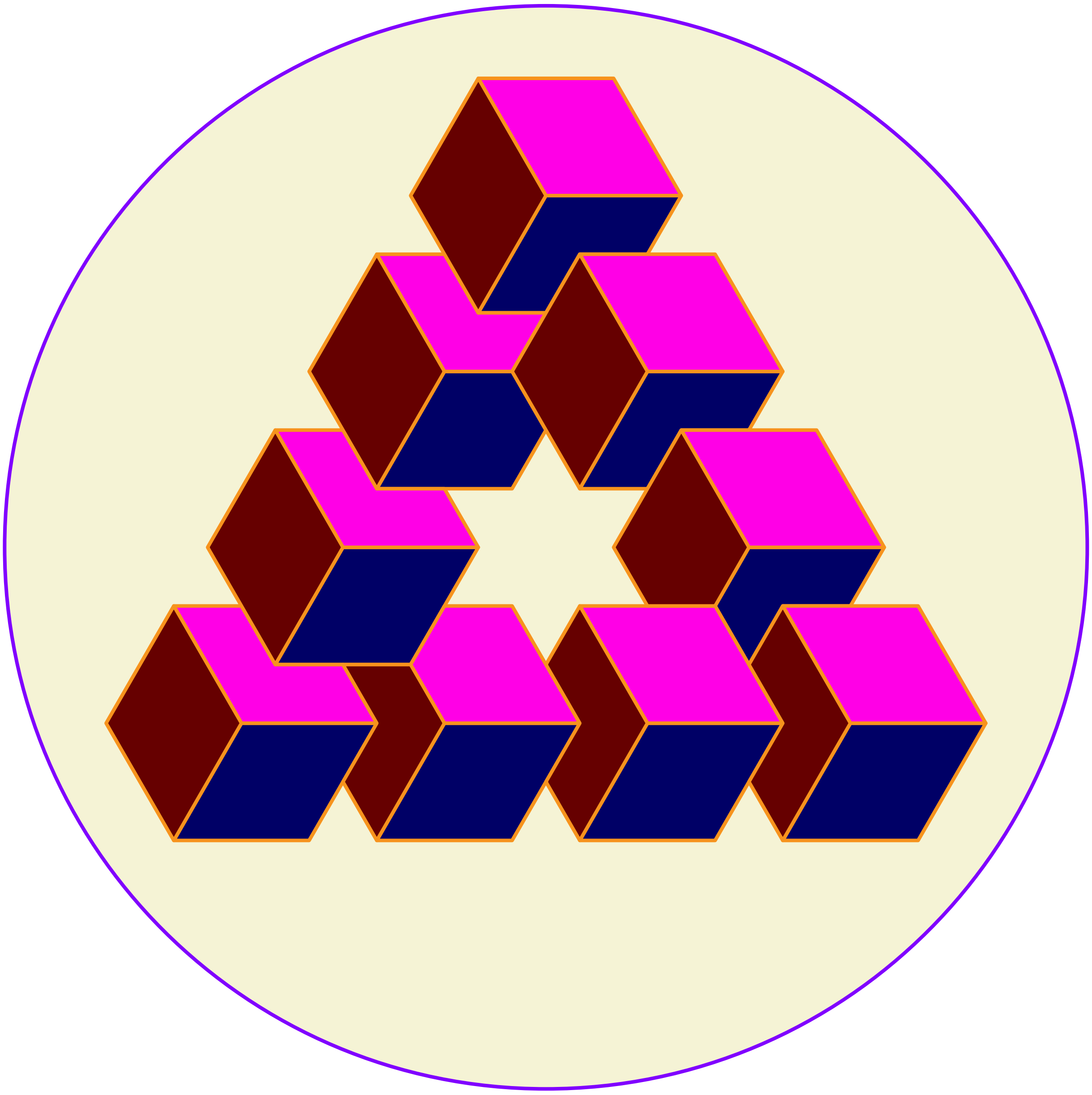
Cours de Mathématiques

Terminale

Spécialité Maths & Maths Expertes



Philippe Moutou
Lycée Henri IV

Mise à jour : 9 octobre 2024

Introduction

Ce cours de mathématiques a été écrit entre septembre 2023 et octobre 2024, lors de ma première année de retraite, à l'intention des élèves en classe de terminale. Il couvre le programme de la spécialité maths mais aussi celui des maths expertes, le tout regroupé en huit chapitres dont les quatre premiers et le dernier traitent le programme de la spécialité. Je l'ai écrit parce qu'il manquait, à mon avis, à mon actif que je partage depuis de nombreuses années sur mon site (mathadomicile.fr). Aussi et surtout, je l'avais promis à ma dernière classe d'élèves du Lycée Henri IV qui m'avait fait la gentillesse de me laisser croire que celui-ci leur serait indispensable... Merci à Alexis, Marguerite, Mathurine, Julie, Victoria, Jessica, Ilyana, Valentin, Valentine, Alice, Carmen, Malena, Soren, Raphaël, Léon, Jules, Martynas, Emma, les trois Thomas, Sami, Victor, Léopold, Mathilde, Louna, les deux Alexandre, Manel, Viviane, Mohamed, Jérémie, Tomi, Yuken, Chloé, Mathieu, Félix, Adèle, Zacharie, Rayan, Noah, Anir, je ne vous oublierai pas.

J'aime à penser qu'il sera éventuellement utile à d'autres élèves et m'excuse en guise d'introduction pour les erreurs qui s'y trouveraient. Personne n'est parfait et ce document ne l'est pas non plus. Vous pouvez aider à l'améliorer en me signalant ce qui est faux, incompréhensible ou ambigu. Il peut manquer aussi certaines parties que j'aurais oublié de traiter. Tout est possible, le document n'a été pour l'instant relu que par moi-même et je sais d'expérience que c'est insuffisant.

Les huit chapitres

Mes élèves me demandaient souvent quelles étaient mes sources. Est-ce que j'invente les exercices (pour l'instant il n'y en a pas pour ce cours de terminale) ? Est-ce que je m'inspire d'autres livres ou est-ce que tout cela sort directement de mon cerveau ? Bien sûr, il y a une part très personnelle dans mes documents, ne serait-ce que la mise en page, la plus ou moins grande concision, le choix et la façon de traiter les exemples, la rédaction des démonstrations... mais il y a aussi une grande part qui vient d'autres auteurs dont je cherche à faire la synthèse. Certains livres m'ont à cet égard été plus utiles que d'autres :

- *les livres de P. Sauser « Mathématiques Élémentaires », publiés chez Ellipses (édition de 1986)*
- *le cours de probabilités de Lebœuf, Roque et Guegand, publiés chez Ellipses (édition de 1987)*
- *les livres de terminale S de P.-H. Terracher, publiés chez Hachette (éditions de 1992 et 2002)*
- *les livres « Pour aller plus loin » de J. Wacksmann, publiés chez Ellipses (édition de 2022)*
- *les documents, publiés ou non, de certains de mes collègues (G. Dubost, P. Cote, L. Lemaire)*
- *les livres de terminale reçus en spécimen, je n'en cite qu'un (collection Hyperbole chez Nathan)*

Les mathématiques ne sont pas compliquées. Quand on les comprend elles deviennent même simples. J'espère ne pas avoir introduit plus de complexité que nécessaire dans cet enchaînement de définitions, de propriétés et d'exemples que constitue ce cours de mathématiques. Tout livre de maths peut paraitre indigeste ; il faut du temps pour se familiariser avec les objets abstraits qui y sont manipulés. Je vous encourage à faire cet effort nécessaire et espère que ce livre vous y aidera.

Philippe Moutou, 9 octobre 2024

Table des matières

Première partie

Fondements

1 Dénombrement & Suites

<u>**Plan du chapitre :**</u>

1.1 Le principe du raisonnement par récurrence

2.2 Dénombrement : cardinaux ensemblistes, produit cartésien, listes

3.3 Combinatoire : permutations, arrangements, combinaisons, coefficients binomiaux

4.4 Suites convergentes et suites divergentes, limite d'une suite et monotonie, suites adjacentes

<u>**Aperçu historique :**</u> Au début, les nombres entiers positifs étaient perçus implicitement – d'où leur appellation de « naturels » – mais au XIXe siècle, les mathématiciens Richard Dedekind (1831-1916) et Guiseppe Peano (1858-1932) établirent les fondements de la théorie des ensembles et l'axiomatique qui justifie le raisonnement par récurrence. Citons le 5e axiome de Peano : *Toute partie non vide de* $\mathbb{N}$ *admet un plus petit élément.* Le plus petit élément de $\mathbb{N}$ étant zéro, la construction complète de l'ensemble ne nécessite plus que le principe de récurrence : *le successeur de l'entier* n *est* $n+1$.

Un dénombrement est aisé si les nombres impliqués sont petits et si les structures sous-jacentes le facilitent. Ainsi, dénombrer le nombre de diagonales d'un polygone à n côtés est relativement simple quand on comprend que d'un sommet on peut en tracer $n-3$; de même, on peut assez facilement dénombrer le nombre de plaques d'immatriculation (dites minéralogiques car anciennement attribuées par le Service des mines) différentes dès lors que l'on en connait la structure (actuellement en France 2 lettres-3 chiffres-2 lettres, certaines combinaisons de lettres étant omises). Dessiner un arbre ou un tableau peut aider au dénombrement, mais certaines structures compliquées nécessitent des outils spécifiques et les mathématiciens s'employèrent à les mettre au point, d'abord dans le cadre de l'étude des jeux de hasard. En France, on retient notamment les noms de Blaise Pascal (1623-1663) et Pierre Fermat (1601-1665) qui échangèrent sur ce sujet. En fait, certains résultats combinatoires avaient déjà été établis durant l'antiquité en Inde, en Chine ou en Grèce. Les mathématiciens indiens Varāhamihira au VIe siècle et Bhāskara au XIIe s'intéressèrent aux nombres de combinaisons de k éléments choisis parmi n. À l'époque contemporaine, la combinatoire se trouve mêlée et s'enrichit de nombreux domaines : l'algèbre, l'analyse, les probabilités, l'informatique. Le mathématicien hongrois Paul Erdős (1913-1996), par exemple, développa la branche combinatoire de la théorie des nombres.

L'étude des suites est très anciennes. En Grèce, au IIIe siècle avant J.-C., Archimède approche le nombre π avec une belle précision grâce à un procédé impliquant deux suites adjacentes. Au Ier siècle après J.-C., Héron d'Alexandrie décrit une méthode très efficace pour extraire une racine carrée ; cette méthode utilise une suite récurrente. Léonard de Pise, dit Fibonacci, expose au XIIIe siècle sa célèbre suite récurrente d'ordre 2 ; au XIVe, Nicolas Oresme a clairement exposé les suites que l'on appelle désormais arithmétique et géométrique ; au XVIIe, Bernoulli, Newton, De Moivre, Stirling, Wallis, entre autres, s'intéressent aux suites pour approcher des valeurs numériques. C'est à ce moment notamment qu'est précisée la notion de limite. Au début du XIXe siècle, Joseph Louis Lagrange (1736-1813) utilise la notation indicielle u_n et Augustin Louis Cauchy (1789-1857) fonde la théorie des suites. À partir de la seconde moitié du XXe siècle, l'utilisation des ordinateurs va réactiver l'intérêt porté aux suites puisqu'ils permettent de pousser très loin et très vite les calculs, par nature répétitifs, qui leur sont liés.

1. Le raisonnement par récurrence

Ce chapitre traite principalement de nombres entiers.

- Dénombrer, c'est compter avec les entiers positifs, dits naturels
- L'analyse combinatoire ne manipule que des entiers naturels
- Les suites sont des applications de $\mathbb{N}$ (l'ensemble des naturels) ou d'une partie de $\mathbb{N}$ vers $\mathbb{R}$
- Le raisonnement par récurrence s'applique là où une propriété est valable pour tout entier à partir d'un certain rang, on va l'utiliser tout au long de ce chapitre

Contrairement à l'arithmétique qui étudie les propriétés intrinsèques des entiers naturels (diviseurs, congruences, etc.), on utilise ici les nombres entiers pour leurs fonctions, cardinale (pour compter) et ordinale (pour numéroter). Ces notions sont abordées dès les petites classes, mais elles deviennent rapidement techniques et complexes d'où l'étalement de leur étude en terminale et au-delà.

1.a. Récurrence simple

Propriété 1.1 (Démonstration par récurrence) Soit $\mathcal{P}(n)$ une propriété qui se décline pour toutes les valeurs d'un entier $n \in \mathbb{N}$.

- Initialisation : si $\mathcal{P}(0)$ est une proposition vraie
- Hérédité : et si $\mathcal{P}(n)$ étant supposée vraie, $\mathcal{P}(n+1)$ est aussi vraie

alors $\mathcal{P}(n)$ est vraie pour tout entier $n \in \mathbb{N}$.

Démonstration Le fondement de cette propriété est le 5^{e} axiome de Peano (toute partie non vide de $\mathbb{N}$ admet un plus petit élément). En effet, si la proposition $\mathcal{P}(n)$ n'était pas vraie pour certains entiers, il existerait un plus petit entier $p > 0$ pour lequel la proposition $\mathcal{P}(p)$ serait fausse.
Par conséquent, pour le rang $p-1$ la proposition $\mathcal{P}(p-1)$ serait encore vraie.
Or on sait que $\mathcal{P}(n)$ vraie $\Longrightarrow \mathcal{P}(n+1)$ vraie.
On doit en déduire que $\mathcal{P}((p-1)+1) = \mathcal{P}(p)$ est vraie, ce qui contredit l'hypothèse de l'existence d'un rang p pour lequel la proposition est fausse. Finalement, les deux conditions (initialisation et hérédité) suffisent à prouver que la propriété est vraie pour tout entier $n \in \mathbb{N}$.

Remarques :

- L'initialisation d'une démonstration par récurrence peut être élargie à un rang $n_0 \geqslant 0$ et formulée ainsi « si $\mathcal{P}(n_0)$ est une proposition vraie ». Dans ce cas, l'hérédité étant également assurée, la propriété est vraie pour tout entier $n \geqslant n_0$ (on note cet ensemble : $[\![n_0; +\infty[\![$).
Pour en donner la preuve, il suffit d'effectuer un glissement des indices : la proposition $\mathcal{P}(n)$ étant rebaptisée proposition $\mathcal{P}'(n-n_0)$, la proposition $\mathcal{P}'(0)$ est vraie car alors $n = n_0$; l'hérédité étant conservée, la propriété $\mathcal{P}'(k)$ est vraie pour tout $k \in \mathbb{N} \Longleftrightarrow n \geqslant n_0$ et donc aussi $\mathcal{P}(n)$ pour tout $n \geqslant n_0$.
- Méthode : commencer par montrer que la propriété est vraie à un rang n_0 (c'est l'initialisation) puis démontrer que si elle est vraie à un rang n, alors elle l'est aussi au rang $n+1$ (c'est l'hérédité : la véracité de la propriété se propage d'un rang où elle est vraie au rang suivant). Ceci permet de conclure qu'elles est vraie pour tous les rangs supérieurs ou égaux à n_0.

Exemple 1 – Prouvons par récurrence la propriété $\mathcal{P}(n) : \forall n \in \mathbb{N},\ 0+1+2+\cdots+n = \frac{n(n+1)}{2}$
Initialisation : $\mathcal{P}(0)$ est vraie car $0 = \frac{0(0+1)}{2}$.
Hérédité : supposons que $\mathcal{P}(n)$ soit vraie et établissons que $\mathcal{P}(n+1)$ est aussi vraie.
Pour cela calculons :
$0+1+2+\cdots+n+(n+1) = \frac{n(n+1)}{2} + (n+1) = \frac{n(n+1)}{2} + \frac{2(n+1)}{2} = \frac{n(n+1)+2(n+1)}{2} = \frac{(n+1)(n+2)}{2}$
Comme $\frac{(n+1)(n+2)}{2} = \frac{(n+1)((n+1)+1)}{2}$, l'expression obtenue est bien conforme à celle de $\mathcal{P}(n+1)$.
La propriété $\mathcal{P}$ est donc bien héréditaire.
Conclusion : la propriété $\mathcal{P}(n)$ est vraie pour tout entier $n \geqslant 0$.

On peut faire cette observation :

- $1^3 = 1^2$
- $1^3 + 2^3 = 1 + 8 = 9 = 3^2 = (1+2)^2$
- $1^3 + 2^3 + 3^3 = 1 + 8 + 27 = 36 = 6^2 = (1+2+3)^2$

Cette observation se généralise t-elle en une propriété $\mathcal{Q}(n)$ valable pour tous les entiers ?
A t-on : $0^3 + 1^3 + 2^3 + 3^3 + \cdots + n^3 = (0+1+2+3+\cdots+n)^2$?
Initialisation : $\mathcal{Q}(0)$ est évidemment vraie ($0^3 = 0^2$) et, comme on l'a vu, $\mathcal{Q}(1)$, $\mathcal{Q}(2)$ et $\mathcal{Q}(3)$ aussi.
Hérédité : supposons que $\mathcal{Q}(n)$ soit vraie.
On en déduit $1^3 + 2^3 + 3^3 + \cdots + n^3 + (n+1)^3 = (1+2+\cdots+n)^2 + (n+1)^3$ et, d'après la propriété $\mathcal{P}$ précédente, on obtient $1^3 + 2^3 + 3^3 + \cdots + n^3 + (n+1)^3 = \left(\frac{n(n+1)}{2}\right)^2 + (n+1)^3$.
Or $\left(\frac{n(n+1)}{2}\right)^2 + (n+1)^3 = \frac{n^2(n+1)^2}{4} + \frac{4(n+1)^3}{4} = \frac{n^2(n+1)^2+4(n+1)^3}{4} = \frac{(n+1)^2(n^2+4(n+1))}{4} = \frac{(n+1)^2(n+2)^2}{4}$.
Ainsi $1^3 + 2^3 + 3^3 + \cdots + n^3 + (n+1)^3 = \left(\frac{(n+1)((n+1)+1)}{2}\right)^2$ ce qui, d'après la propriété $\mathcal{P}$, est égal à $(1+2+\cdots+n+(n+1))^2$, soit l'expression attendue pour la proposition $\mathcal{Q}(n+1)$
La proposition $\mathcal{Q}$ est donc bien héréditaire.
Conclusion : la proposition $\mathcal{Q}(n)$ est vraie pour tout entier $n \geqslant 0$.
Finalement, on peut reformuler $\mathcal{Q}(n)$: $\forall n \in \mathbb{N},\ \ 0^3 + 1^3 + 2^3 + \cdots + n^3 = \left(\frac{n(n+1)}{2}\right)^2$

1.b. Récurrence double

La récurrence double est une généralisation assez naturelle du procédé de récurrence simple. Elle permet de démontrer des propriétés dont la véracité à un rang n dépend de sa véracité aux rangs $n-1$ et $n-2$.

Propriété 1.2 Soit $\mathcal{P}(n)$ une propriété qui se décline pour toutes les valeurs d'un entier $n \in \mathbb{N}$.

- Initialisation : si $\mathcal{P}(0)$ et $\mathcal{P}(1)$ est une proposition vraie
- Hérédité : et si $\mathcal{P}(n)$ et $\mathcal{P}(n+1)$ étant supposées vraies, $\mathcal{P}(n+2)$ est aussi vraie

alors $\mathcal{P}(n)$ est vraie pour tout entier $n \in \mathbb{N}$.

Démonstration Notons $\mathcal{Q}(n)$ la proposition qui est vraie lorsque $\mathcal{P}(n)$ et $\mathcal{P}(n+1)$ sont vraies.
D'après la 1re hypothèse, $\mathcal{P}(0)$ et $\mathcal{P}(1)$ sont vraies donc $\mathcal{Q}(0)$ est vraie.
La propriété $\mathcal{Q}$ est ainsi initialisée au rang 0.

D'après la 2e hypothèse, si $\mathcal{Q}(n)$ est vraie, c'est-à-dire si $\mathcal{P}(n)$ et $\mathcal{P}(n+1)$ sont vraies, alors $\mathcal{P}(n+2)$ est vraie.
On en déduit que $\mathcal{P}(n+1)$ et $\mathcal{P}(n+2)$ sont vraies donc $\mathcal{Q}(n+1)$ est vraie.
La propriété $\mathcal{Q}$ est donc héréditaire.

Conclusion : la propriété $\mathcal{Q}(n)$ est vraie pour tout $n \in \mathbb{N}$.
Cela signifie que $\mathcal{P}(n)$ (et $\mathcal{P}(n+1)$) est vraie pour tout $n \in \mathbb{N}$.

Remarques :

- L'initialisation d'une récurrence double peut être élargie à un rang $n_0 \geqslant 0$ (on formulerait cette initialisation : « si $\mathcal{P}(n_0)$ et $\mathcal{P}(n_0+1)$ sont des propositions vraies ». Dans ce cas, la propriété est alors vraie pour tout entier $n \geqslant n_0$.
- Méthode :
  - Énoncer clairement la propriété $\mathcal{P}(n)$ pour tout $n \in [\![n_0;+\infty[$
  - Faire une double initialisation : montrer que la propriété est vraie aux rangs n_0 et n_0+1
  - Démontrer l'hérédité avec une hypothèse de récurrence double. Soit $n \in [\![n_0;+\infty[$, en supposant que $\mathcal{P}(n)$ et $\mathcal{P}(n+1)$ sont vraies montrer que $\mathcal{P}(n+2)$ est vraie
  - Conclure d'une phrase : « par principe de récurrence double, la propriété $\mathcal{P}(n)$ est vraie pour tout $n \in [\![n_0;+\infty[$ »

Exemple 2 – On considère la suite (u_n) définie par :
$u_0 = 0$, $u_1 = 1$ et, pour tout $n \geqslant 2$, $u_{n+2} = u_{n+1} + u_n$
Cette suite (u_n) est la fameuse suite de Fibonacci dont les termes sont : 0, 1, 1, 2, 3, 5, 8, 13, 21, etc. et dont le rapport $\frac{u_{n+1}}{u_n}$ tend vers le « nombre d'or » $\varphi = \dfrac{1+\sqrt{5}}{2}$.
Pour un rang donné n, on pose $\mathcal{P}(n)$: $u_n = \dfrac{1}{\sqrt{5}} \times \left[\left(\dfrac{1+\sqrt{5}}{2}\right)^n - \left(\dfrac{1-\sqrt{5}}{2}\right)^n\right]$
Cette propriété est-elle vraie pour tout $n \in \mathbb{N}$?
Initialisation :
$\dfrac{1}{\sqrt{5}} \times \left[\left(\dfrac{1+\sqrt{5}}{2}\right)^0 - \left(\dfrac{1-\sqrt{5}}{2}\right)^0\right] = \dfrac{1}{\sqrt{5}} \times [1-1] = 0 = u_0$
$\dfrac{1}{\sqrt{5}} \times \left[\left(\dfrac{1+\sqrt{5}}{2}\right)^1 - \left(\dfrac{1-\sqrt{5}}{2}\right)^1\right] = \dfrac{1}{\sqrt{5}} \times [\dfrac{1+\sqrt{5}}{2} - \dfrac{1-\sqrt{5}}{2}] = \dfrac{1}{\sqrt{5}} \times \sqrt{5} = 1 = u_1$
Les propositions $\mathcal{P}(0)$ et $\mathcal{P}(1)$ sont donc vraies.

Hérédité :

$$\begin{aligned}
u_{n+2} &= u_{n+1} + u_n \\
&= \frac{1}{\sqrt{5}} \times \left[\left(\frac{1+\sqrt{5}}{2}\right)^{n+1} - \left(\frac{1-\sqrt{5}}{2}\right)^{n+1}\right] + \frac{1}{\sqrt{5}} \times \left[\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n\right] \\
&= \frac{1}{\sqrt{5}} \times \left[\left(\frac{1+\sqrt{5}}{2}+1\right)\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}+1\right)\left(\frac{1-\sqrt{5}}{2}\right)^n\right] \\
&= \frac{1}{\sqrt{5}} \times \left[\left(\frac{3+\sqrt{5}}{2}\right)\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{3-\sqrt{5}}{2}\right)\left(\frac{1-\sqrt{5}}{2}\right)^n\right]
\end{aligned}$$

Or $\left(\dfrac{1+\sqrt{5}}{2}\right)^2 = \dfrac{1+2\sqrt{5}+5}{4} = \dfrac{3+\sqrt{5}}{2}$ et $\left(\dfrac{1-\sqrt{5}}{2}\right)^2 = \dfrac{1-2\sqrt{5}+5}{4} = \dfrac{3-\sqrt{5}}{2}$
Finalement $u_{n+2} = \dfrac{1}{\sqrt{5}} \times \left[\left(\dfrac{1+\sqrt{5}}{2}\right)^{n+2} - \left(\dfrac{1-\sqrt{5}}{2}\right)^{n+2}\right]$
La proposition $\mathcal{P}(n+2)$ est vraie lorsque $\mathcal{P}(n+1)$ et $\mathcal{P}(n)$ sont toutes les deux vraies.

Conclusion : par principe de récurrence double la propriété $\mathcal{P}(n)$ est vraie pour tout $n \in \mathbb{N}$.

1.c. Récurrence forte

La récurrence simple (ou double) permet de démontrer des propriétés dont la véracité se propage d'un (ou deux) rang(s) au rang suivant. La récurrence forte permet de démontrer des propriétés dont la véracité à un rang donné dépend de la véracité à *tous* les rangs précédents.

Propriété 1.3 Soit $\mathcal{P}(n)$ une propriété qui se décline pour toutes les valeurs d'un entier $n \in \mathbb{N}$.
- Initialisation : si $\mathcal{P}(0)$ est une proposition vraie
- Hérédité : et si $\mathcal{P}(0)$, $\mathcal{P}(1)$, ..., $\mathcal{P}(n)$ étant supposées vraies, $\mathcal{P}(n+1)$ est aussi vraie

alors $\mathcal{P}(n)$ est vraie pour tout entier $n \in \mathbb{N}$.

Démonstration Notons $\mathcal{Q}(n)$ la proposition qui est vraie lorsque $\mathcal{P}(0)$, $\mathcal{P}(1)$, ..., $\mathcal{P}(n)$ sont vraies.
D'après la 1re hypothèse, $\mathcal{P}(0)$ est vraie donc $\mathcal{Q}(0)$ est vraie.
La propriété $\mathcal{Q}$ est ainsi initialisée au rang 0.

D'après la 2e hypothèse, si $\mathcal{Q}(n)$ est vraie, c'est-à-dire si $\mathcal{P}(0)$, $\mathcal{P}(1)$, ..., $\mathcal{P}(n)$ sont vraies alors $\mathcal{P}(n+1)$ est vraie. On en déduit que $\mathcal{P}(1)$, $\mathcal{P}(2)$, ..., $\mathcal{P}(n+1)$ sont vraies donc $\mathcal{Q}(n+1)$ est vraie.
La propriété $\mathcal{Q}$ est donc héréditaire.

Conclusion : la propriété $\mathcal{Q}(n)$ est vraie pour tout $n \in \mathbb{N}$.
Cela signifie que $\mathcal{P}(n)$ (et tous les $\mathcal{P}(k)$ pour $k \in [\![0; n-1]\!]$) est vraie pour tout $n \in \mathbb{N}$.

Remarques :

- L'initialisation d'une récurrence forte peut être élargie à un rang $n_0 \geqslant 0$ (on formulerait cette initialisation : « si $\mathcal{P}(n_0)$ est une proposition vraie ». Dans ce cas, la propriété est alors vraie pour tout entier $n \geqslant n_0$.
- Méthode :
  — Énoncer clairement, pour tout $n \in [\![n_0; +\infty[$ la propriété $\mathcal{P}(n)$
  — Faire une simple initialisation : montrer que la propriété est vraie au rang n_0
  — Démontrer l'hérédité avec une hypothèse de récurrence forte. Soit $n \in [\![n_0; +\infty[$, supposons que $\mathcal{P}(n_0)$, $\mathcal{P}(n_0+1)$, ..., $\mathcal{P}(n)$ sont vraies et montrons qu'alors $\mathcal{P}(n+1)$ est vraie
  — Conclure d'une phrase : « par principe de récurrence forte, la propriété $\mathcal{P}(n)$ est vraie pour tout $n \in [\![n_0; +\infty[$ »

Exemple 3 – Le théorème fondamental de l'arithmétique énonce que chaque entier naturel $n \geqslant 2$ peut se décomposer comme produit de facteurs premiers (un seul facteur pour les nombres premiers). L'existence de cette décomposition a été montrée dans le chapitre d'arithmétique par le procédé de descente infinie. On peut néanmoins prouver cette existence grâce à une récurrence forte.
Soit un entier $n \geqslant 2$ et $\mathcal{P}(n)$: n est, soit un nombre premier, soit un produit de nombres premiers.

Initialisation :
2 est un nombre premier. $\mathcal{P}(2)$ est donc vraie.

Hérédité :
Supposons que pour tout $k \in [\![2; n]\!]$, $\mathcal{P}(k)$ soit vraie.
Deux cas se présentent :

- Si $n+1$ est un nombre premier alors $\mathcal{P}(n+1)$ est bien vraie.
- Si $n+1$ n'est pas premier alors il existe au moins un diviseur m tel que $m \neq 1$, $m \neq (n+1)$ et $\exists p \in \mathbb{N},\ \ n+1 = p \times m$. Or tout diviseur m d'un entier $n+1$ est compris entre 1 et $n+1$.
  On a donc $1 \leqslant m \leqslant n+1$ mais comme $m \neq 1$ et $m \neq (n+1)$, on en déduit $m \in [\![2; n]\!]$.
  Ce qu'on vient de dire pour m est valable également pour p (car $m > 1 \Longrightarrow p \times m > p \Longrightarrow n+1 > p$ et $m < n+1 \Longrightarrow m < p \times m \Longrightarrow 1 < p$) donc $p \in [\![2; n]\!]$.
  Comme $m \in [\![2; n]\!]$ et $p \in [\![2; n]\!]$, d'après l'hypothèse de récurrence $\mathcal{P}(m)$ et $\mathcal{P}(p)$ sont vraies.
  Finalement, $n+1$ peut s'écrire comme produit de nombres premiers. $\mathcal{P}(n+1)$ est bien vraie.

Ainsi, dans tous les cas, $\mathcal{P}(n+1)$ est vraie.

Conclusion : d'après le principe de récurrence forte, pour tout $n \geqslant 2$ $\mathcal{P}(n)$ est vraie.
Tout entier $n \geqslant 2$ est, soit un nombre premier, soit un produit de nombres premiers.

2. Dénombrement

2.a. Cardinaux ensemblistes

Définition 1.1 (Cardinal) Soit E un ensemble fini d'éléments.
Le cardinal de E, noté $\mathrm{Card}(E)$ est le nombre d'éléments de E.

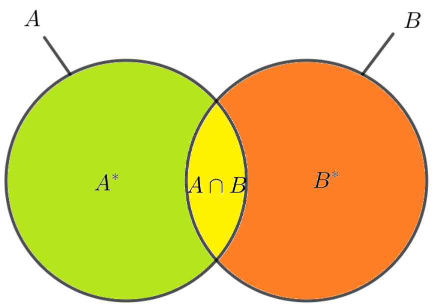
Propriété 1.4 L'*ensemble vide* $\emptyset$ ne contient aucun élément, d'où $\mathrm{Card}(\emptyset) = 0$

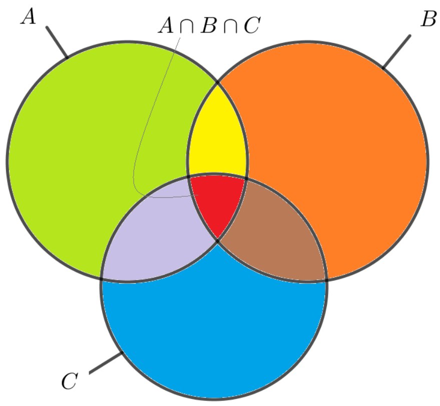
Cardinaux de l'union de deux ou trois ensembles :
Si les ensembles A et B sont *disjoints* alors $\mathrm{Card}(A \cup B) = \mathrm{Card}(A) + \mathrm{Card}(B)$
Pour deux ensembles A et B : $\mathrm{Card}(A \cup B) = \mathrm{Card}(A) + \mathrm{Card}(B) - \mathrm{Card}(A \cap B)$
Pour trois ensembles A, B et C :

$$\begin{aligned}\mathrm{Card}(A \cup B \cup C) = {} & \mathrm{Card}(A) + \mathrm{Card}(B) + \mathrm{Card}(C) \\ & - \mathrm{Card}(A \cap B) - \mathrm{Card}(A \cap C) - \mathrm{Card}(B \cap C) + \mathrm{Card}(A \cap B \cap C)\end{aligned}$$

Démonstration Deux ensembles disjoints :
Les éléments de $A \cup B$ sont tous ceux de A et tous ceux de B.
Il n'y a aucun doublon puisque les ensembles sont disjoints ($A \cap B = \emptyset$).
Par conséquent $\text{Card}(A \cup B) = \text{Card}(A) + \text{Card}(B)$.
Dans ce cas B est le *complémentaire* de A dans $E = A \cup B$, ce qu'on note $B = \bar{A}_E$.

Deux ensembles (cas général) :
Notons $E = A \cup B$ et $A^* = \bar{B}_E$ (voir l'illustration en bas à gauche).
On a $A = A^* \cup (A \cap B)$ avec $A^* \cap (A \cap B) = \emptyset$.
En appliquant la propriété précédente : $\text{Card}(A) = \text{Card}(A^* \cup (A \cap B)) = \text{Card}(A^*) + \text{Card}(A \cap B)$.
On en déduit $\text{Card}(A^*) = \text{Card}(A) - \text{Card}(A \cap B)$ or $A \cup B = A^* \cup B$ avec $A^* \cap B = \emptyset$
donc $\text{Card}(A \cup B) = \text{Card}(A^*) + \text{Card}(B) = \text{Card}(A) - \text{Card}(A \cap B) + \text{Card}(B)$.





Trois ensembles :
On utilise la propriété précédente deux fois, en remarquant que $A \cup B \cup C = A \cup (B \cup C)$ mais aussi que $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ et $(A \cap B) \cap (A \cap C) = A \cap B \cap C$.

$$\begin{aligned}
\text{Card}(A \cup (B \cup C)) &= \text{Card}(A) + \text{Card}(B \cup C) - \text{Card}(A \cap (B \cup C)) \\
&= \text{Card}(A) + [\text{Card}(B) + \text{Card}(C) - \text{Card}(B \cap C)] - \text{Card}(A \cap (B \cup C)) \\
&= \underbrace{\text{Card}(A) + \text{Card}(B) + \text{Card}(C)}_{\text{noté } ABC} - \text{Card}(B \cap C) - \text{Card}((A \cap B) \cup (A \cup C) \\
&= ABC - \text{Card}(B \cap C) - [\text{Card}(A \cap B) + \text{Card}(A \cap C) - \text{Card}((A \cap B) \cup (A \cap C))] \\
&= ABC - \text{Card}(B \cap C) - \text{Card}(A \cap B) - \text{Card}(A \cap C) + \text{Card}(A \cap B \cap C)
\end{aligned}$$

Exemple 4 – Dans le club de sport de ma ville, il y a 35 membres : 17 sont inscrits au foot, 14 au basket et 15 au tennis, 8 sont inscrits à la fois au foot et au basket, 7 à la fois au basket et au tennis, 6 à la fois au foot et au tennis et 5 sont inscrits à la fois aux 3 sports. Y a t-il des membres qui ne sont inscrits à aucun de ces trois sports ? Et si oui, combien ?

On va noter F, B et T les ensembles de membres inscrits au foot, basket et tennis. Ainsi on a :
- $\text{Card}(F) = 17$, $\text{Card}(B) = 14$, $\text{Card}(T) = 15$
- $\text{Card}(F \cap B) = 8$, $\text{Card}(B \cap T) = 7$, $\text{Card}(F \cap T) = 6$
- $\text{Card}(F \cap B \cap T) = 5$

On en déduit $\text{Card}(F \cup B \cup T) = 17 + 14 + 15 - 8 - 7 - 6 + 5 = 30$.
L'ensemble des membres du club, noté *Club*, est constitué des ensembles disjoints des membres inscrits à au moins un de ces trois sports, noté $F \cup B \cup T$, et de ceux qui ne sont inscrits à aucun des trois, noté $\overline{F \cup B \cup T}$.
On a donc $\text{Card}(Club) = \text{Card}(F \cup B \cup T) + \text{Card}(\overline{F \cup B \cup T})$ donc
$\text{Card}(\overline{F \cup B \cup T}) = \text{Card}(Club) - \text{Card}(F \cup B \cup T)$ soit $\text{Card}(\overline{F \cup B \cup T}) = 35 - 30 = 5$.
Il y a 5 membres qui ne sont inscrits à aucun des trois sports.

2.b. Produits cartésiens

DÉFINITION 1.2 Si A et B sont deux ensembles, le *produit cartésien* de A et B, noté $A \times B$, est l'ensemble dont les éléments sont les *couples* (a, b) où $a \in A$ et $b \in B$.

Soit n est un entier supérieur ou égal à 2 :
Si $A_1, A_2, \ldots, A_n$ sont n ensembles, le produit cartésien $A_1 \times A_2 \times \cdots \times A_n$ est l'ensemble des n-uplets $(a_1, a_2, \ldots, a_n)$ où $\forall i \in [\![1;n]\!],\ \ a_i \in A_i$.
$A^n = \underbrace{A \times A \times \cdots \times A}_{n fois}$, est l'ensemble des n-uplets $(a_1, a_2, \ldots, a_n)$ où $\forall i \in [\![1;n]\!],\ \ a_i \in A$.

PROPRIÉTÉ 1.5 Si A et B sont deux ensembles finis, $\mathrm{Card}(A \times B) = \mathrm{Card}(A) \times \mathrm{Card}(B)$

Soit n est un entier supérieur ou égal à 2 :
$\mathrm{Card}(A_1 \times A_2 \times \cdots \times A_n) = \mathrm{Card}(A_1) \times \mathrm{Card}(A_2) \times \mathrm{Card}(A_n)$
$\mathrm{Card}(A^n) = (\mathrm{Card}(A))^n$

DÉMONSTRATION Deux ensembles :
Pour constituer les *couples* (a, b) avec a décrivant l'ensemble A et b décrivant l'ensemble B, on peut commencer par fixer le 1^er^ élément a_1 de A et constituer les couples (a_1, b) avec b décrivant l'ensemble B, soit écrire les $\mathrm{Card}(B)$ couples (a_1, b_i) où $\forall i \in [\![1;\mathrm{Card}(B)]\!],\ \ b_i \in B$.
Ensuite on fixe le 2^e^ élément a_2 de A et constituer les couples (a_2, b_i) où $\forall i \in [\![1;\mathrm{Card}(B)]\!],\ \ b_i \in B$.
Ainsi de suite, jusqu'à constituer les couples $(a_{\mathrm{Card}(A)}, b_i)$ où $\forall i \in [\![1;\mathrm{Card}(B)]\!],\ \ b_i \in B$.
On a ainsi constitué $\mathrm{Card}(A)$ fois des ensembles contenant $\mathrm{Card}(B)$ éléments.
Cette méthode de construction peut être schématisée par un arbre contenant $\mathrm{Card}(A)$ branches sur chacune desquelles sont réparties $\mathrm{Card}(B)$ feuilles, ce qui conduit à obtenir $\mathrm{Card}(A) \times \mathrm{Card}(B)$ feuilles.

n ensembles :
On procède par récurrence, l'étape d'initialisation étant vérifiée d'après la propriété précédente.
Pour montrer l'hérédité de cette propriété, on suppose que
$\mathrm{Card}(A_1 \times \cdots \times A_n) = \mathrm{Card}(A_1) \times \cdots \times \mathrm{Card}(A_n)$ et on ajoute $\mathrm{Card}(A_{n+1})$ feuilles aux $\mathrm{Card}(A_1) \times \cdots \times \mathrm{Card}(A_n)$ extrémités de l'arbre précédent, en ajoutant à celles-ci les $\mathrm{Card}(A_{n+1})$ éléments du $(n+1)$^e^ ensemble, ce qui conduit à $\mathrm{Card}(A_1) \times \cdots \times \mathrm{Card}(A_{n+1})$ feuilles.

Lorsque $\forall i \in [\![1;n]\!],\ \ A_i = A$, la formule précédente s'écrit $\mathrm{Card}(A^n) = (\mathrm{Card}(A))^n$.

Remarques :

- Lorsqu'on tire deux dés à 6 faces numérotées, un résultat possible est un élément du produit cartésien A^2 avec $A = \{1, 2, 3, 4, 5, 6\}$. Explicitons cet ensemble selon le procédé de la démonstration : $A^2 = \{(1,1), (1,2), (1,3), (1,4), (1,5), (1,6), (2,1), (2,2), \ldots, (6,5), (6,6)\}$. Noter que les éléments $(1, 2)$ et $(2, 1)$ sont tous les deux présents car différents, même si dans la pratique ils peuvent paraître indiscernables : l'ordre des constituants importe.
- Lorsque les ensembles utilisés sont finis, le cardinal d'un produit cartésien est aussi fini. Mais si l'un des ensembles intervenant dans ce produit cartésien n'est pas fini, le produit l'est également. Par exemple $\{A, B\} \times \mathbb{N}$ contient une infinité d'éléments : les couples $(A, 0), (A, 1), (A, 2), \ldots$ et les couples $(B, 0), (B, 1), (B, 2), \ldots$. De même pour $\mathbb{R}^2$ ou $\mathbb{R}^3$ qui sont, de plus, indénombrables [1] comme $\mathbb{R}$.

1. Un ensemble est dénombrable si on peut numéroter ses éléments avec des entiers. Ainsi $\mathbb{Q}$ – l'ensemble des rationnels – est infini et dénombrable, mais pas $\mathbb{R}$ qui est indénombrable. Cette propriété a été démontrée par Georg Cantor (1845-1918) en utilisant l'argument de la *diagonale de Cantor*. Pour la dénombrabilité de $\mathbb{Q}^{+*}$, Cantor ordonna les rationnels positifs irréductibles selon la somme croissante du numérateur et du dénominateur ($\frac{1}{1}$, $\frac{1}{2}$, $\frac{2}{1}$, $\frac{1}{3}$, $\frac{3}{1}$, etc.) ce qui permet de les numéroter.

Exemple 5 (Immatriculations) – L'immatriculation d'un véhicule en France, depuis le 15 avril 2009, est une série de 7 caractères alphanumériques noirs sur fond blanc : 2 lettres, 1 tiret, 3 chiffres, 1 tiret et 2 lettres. Elle fait, de plus, apparaître, sur sa partie droite et sur fond bleu, un identifiant territorial.
Ce nouveau SIV (Système d'Immatriculation des Véhicules) s'est substitué à l'ancien dispositif de numérotation qui datait de 1950 et qui comprenait un numéro d'ordre d'un à quatre chiffres, une série d'une à trois lettres et un code départemental à deux chiffres, sauf exceptions (Corse et départements d'outre-mer).

Dénombrons les immatriculations possibles avec le nouveau SIV :
Il faut choisir 2 lettres, puis 3 chiffres et enfin 2 lettres.
L'identifiant territorial ne compte pas, il ne peut pas y avoir AB-123-AB avec deux identifiants territoriaux différents (un à Paris et l'autre dans le Loiret par exemple) car ce numéro est national.
On a $26^2 = 676$ façons de choisir 2 lettres ordonnées (AB différent de BA), $10^3 = 1000$ façons de choisir 3 chiffres ordonnées en supposant que l'on utilise des nombres comme 012, même écrit 12.
Finalement, on peut immatriculer $676 \times 1000 \times 676 = 456\,976\,000$ véhicules avec ce nouveau SIV.
En réalité il y a beaucoup moins d'immatriculations possibles du fait qu'on enlève les I, O et U (trop proches de 1, 0 et V) ainsi que les groupes SS et WW (utilisé pour les véhicules en attente d'immatriculation) du bloc de gauche et seulement SS du bloc de droite ; on enlève aussi le 000.
l y a donc $(23^2 - 2) \times 999 \times (23^2 - 1) = 277\,977\,744$ immatriculations possibles avec ce SIV. La durée de vie prévisible de ce système est estimée à 80 ans ; l'innovation portant également sur le fait qu'une immatriculation est inchangée si le véhicule change de propriétaire, contrairement à l'ancien système.

Dénombrons les immatriculations de l'ancien système :
Il y a $10^4 = 10000$ numéros d'ordre d'un à quatre chiffres. Pour les séries d'une à trois lettres, il y a 26 séries d'une lettre, $26^2 = 676$ séries de deux lettres et $26^3 = 17576$ séries de trois lettres, soit $26 + 676 + 17576 = 18\,278$ séries d'une à trois lettres. Il y a donc potentiellement $10000 \times 18278 = 182\,780\,000$ immatriculations possibles avec l'ancien système pour un département.
Dans chaque département, on peut trouver un même début : par exemple, 1975 GM 75 à Paris et 1975 GM 45 dans le Loiret. On constate que cet ancien système conduit à un nombre plus grand de possibilités car il y a 96 codes départementaux à deux chiffres (Corse $2A$ et $2B$ comprise) et une douzaine de codes à trois chiffres (départements d'outre-mer). Mais en réalité, ce système allait être dépassé à Paris (on s'est arrêté au numéro 523 RQD 75 en 2009) car cette immatriculation était changée lors d'un déménagement ou d'une vente du véhicule, dès lors que le département d'immatriculation changeait. Un même véhicule pouvait ainsi changer 1, 2 ou 3 fois d'immatriculation durant son utilisation, ou encore davantage.

Définition 1.3 (Liste avec répétitions) Si E est un ensemble et k un entier non-nul, une *liste avec répétitions* de k éléments de E – aussi appelée une k-liste avec répétitions – est une liste ordonnée d'éléments de E distincts ou non, c'est-à-dire un élément de E^n, un k-uplet $(e_1, e_2, \dots, e_k)$ où $\forall i \in [\![1;k]\!],\ e_i \in E$.

Remarques :

- On rencontre la notion de liste avec répétitions dans de nombreuses situations : on parle notamment d'*arrangement avec répétitions* dans un contexte combinatoire, et si on considère une *application*[2] définie sur $[\![1;k]\!]$ à valeurs dans E, celle-ci peut être considérée comme une k-liste avec répétitions. Le nombre d'applications de $[\![1;k]\!]$ vers E est donc donné par la propriété précédente, c'est $(\mathrm{Card}E)^k$.
- Une femme a eu trois enfants : l'aîné « A », le cadet « B » et le benjamin « C ». Elle considère le triplet (la 3-liste) de leurs dates d'anniversaire (A :12Jun, B :24Fev, C :2Nov) comme une application de $[\![1;3]\!]$ vers $E = \{1\mathrm{Jan}, 2\mathrm{Jan}, \dots, 31\mathrm{Dec}\}$, un ensemble de cardinal 366. Combien y a t-il d'applications de $[\![1;3]\!]$ vers E possibles : $366^3 = 49\,027\,896$.

2. Une application d'un ensemble E_1 vers un ensemble E_2 une fonction de E_1 vers E_2 telle que tous les éléments de E_1 aient une image dans E_2 (son domaine de définition et son ensemble de départ coïncident).

2.c. Listes sans répétition

DÉFINITION 1.4 (LISTE SANS RÉPÉTITION) Si E est un ensemble et k un entier non-nul, une *liste sans répétition* de k éléments de E – aussi appelée une k-liste sans répétition – est une liste ordonnée d'éléments distincts de E, c'est-à-dire un k-uplet $(e_1, e_2, \ldots, e_k)$ où $\forall i \in [\![1;k]\!],\ e_i \in E$ et $\forall (i,j) \in [\![1;k]\!]^2,\ i \neq j \Longrightarrow e_i \neq e_j$.

PROPRIÉTÉ 1.6 Soient E est un ensemble de cardinal n et $k \in \mathbb{N}^*$.

- Si $k > n$ aucune k-listes sans répétition d'éléments de E n'existe.
- Si $k \leqslant n$, le nombre de k-listes sans répétition d'éléments de E – noté A_n^k – est : $n \times (n-1) \times (n-2) \times \cdots \times (n-k+1) = \dfrac{n!}{(n-k)!}$

DÉMONSTRATION Pour constituer une k-liste sans répétition, on en choisit librement e_1, le 1^er^ élément, dans l'ensemble E : il y a n choix. On choisit alors e_2, le 2^e^ élément, dans $E \backslash \{e_1\}$ (E privé de l'élément e_1) dont le cardinal est $n-1$: il y a $n-1$ choix.
Ainsi de suite, jusqu'à choisir e_k, le k^{e} élément, dans $E \backslash \{e_1, e_2, \ldots, e_{k-1}\}$ (E privé des $k-1$ premiers éléments choisis) dont le cardinal est $n-(k-1) = n-k+1$: il y a $n-k+1$ choix.
Le produit de tous ces choix nous donne le nombre de possibilités.
Il y a donc $n \times (n-1) \times (n-2) \times \cdots \times (n-k+1)$ k-listes possibles.

La notation $n!$ – factorielle n – signifie $n \times (n-1) \times (n-2) \times \cdots \times 2 \times 1$. Ainsi :

$$\frac{n!}{(n-k)!} = \frac{n \times (n-1) \times (n-2) \times \cdots \times (n-k+1) \times (n-k) \times (n-k-1) \times \cdots \times 2 \times 1}{(n-k) \times (n-k-1) \times \cdots \times 2 \times 1}$$

D'où $\dfrac{n!}{(n-k)!} = n \times (n-1) \times (n-2) \times \cdots \times (n-k+1)$
On a simplifié par $(n-k)!$ qui n'est jamais nul quand $k \leqslant n$ (pour $k = n$, on a $0! = 1$).

Remarques :

- Dans un contexte combinatoire, pour une k-liste sans répétition, on parle d'*arrangement*. La notation A_n^k signifie d'ailleurs « nombre d'arrangements de k objets parmi n ».
  Dans un contexte probabiliste, une k-liste sans répétition est obtenue lors du tirage de k boules d'une urne en contenant n, *sans remise* de la boule tirée (si on effectue un tirage avec remise, on obtient une k-liste avec répétition).
  Dans le contexte des applications de $[\![1;k]\!]$ à valeurs dans E, une k-liste sans répétition est obtenue si l'application est *injective*[3].
- Lors d'une course de chevaux, s'il y a 15 chevaux au départ, donner le tiercé gagnant revient à énoncer un des $15 \times 14 \times 13 = 2\,730$ triplets sans répétition possibles.
  Dans une classe de 40 élèves, quelle est la probabilité de l'évènement E : deux élèves au moins ont la même date d'anniversaire ? L'évènement contraire est $\overline{E}$: chacun des élèves a une date d'anniversaire différente. En estimant qu'il n'y a que 365 dates, le nombre de cas possibles est $365^{40} \approx 3,1 \times 10^{102}$. Le nombre de cas favorables à $\overline{E}$ est le nombre de 40-listes sans répétition de dates : il y en a $365 \times 364 \times \cdots \times 326 = \frac{365!}{325!}$. Donc $P(\overline{E}) = \frac{365!}{325! \times 365^{12}} = \frac{365}{365} \times \frac{364}{365} \times \cdots \times \frac{326}{365}$.
  Ce simple programme Python nous fournit le résultat $P(\overline{E}) \approx 0,109$, d'où $P(E) = 1 - P(\overline{E}) \approx 0,891$.

```
N,P=12,1
for I in range(N) :
    P*=(365-I)/365
print("P(non E)=",P,"P(E)=",(1-P))
```

3. Une application est injective si chaque élément de l'ensemble de départ est associé à une image différente dans l'ensemble d'arrivée

2.d. Permutations

DÉFINITION 1.5 (PERMUTATION) Soit E en ensemble de cardinal n.
Une *permutation* des n éléments de E est une n-liste sans répétition des n éléments de E.

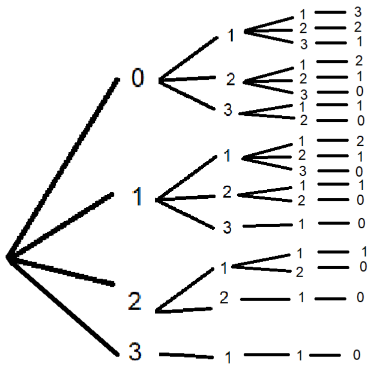
Remarques :

- Le nombre de permutations de n éléments est $A_n^n = n \times (n-1) \times \cdots \times 2 \times 1 = n!$
  En effet, une permutation des éléments de l'ensemble E est un arrangement de tous ces éléments, c'est-à-dire une liste ordonnée de ces éléments. Si $\text{Card}(E) = n$, une application de $[\![1;n]\!]$ vers E est *bijective*. Tous les éléments de E ont un antécédent (l'application est surjective) et un seul (l'application est injective).
- On a un ensemble de 8 points dans un plan et on souhaite connaître le nombre de polygones différents que l'on peut tracer avec ces points comme sommets : le nombre de permutations des 8 sommets est $8! = 40\,320$. Mais on a des doublons car $ABCDEFGH$ est le même polygone que $BCDEFGHA$; comme il y a 8 points de départ possible cela réduit le nombre à $7! = 5\,040$. Il reste encore des doublons car $ABCDEFGH$ est le même polygone que $HGFEDCBA$; comme il y a 2 sens possible cela réduit le nombre à $\frac{7!}{2} = 2\,520$.

EXEMPLE 6 – Le code du coffre-fort de ma grand-mère Mathilde est un anagramme de son prénom. Combien y a t-il de codes différents : $8! = 40\,320$ les huit lettres étant différentes.

Je me souviens qu'elle a dit que son code n'avait pas deux voyelles accolées.
Combien y a t-il de codes pour lesquels les 3 voyelles sont séparées ?
Cette question est un peu compliquée.

Je commence par dénombrer les emplacements possibles des voyelles : celles-ci doivent être séparées par au moins une consonne, elles peuvent être aux extrémités mais on peut disposer jusqu'à 3 consonnes à une extrémité. Les consonnes sont en 4 groupes : aux extrémités (le 1er et le 4e groupe) il peut y avoir de 0 à 3 consonnes, les groupes du centres (le 2e et le 3e) peuvent en contenir de 1 à 3 et le total des consonnes est 5. Le plus simple est de dresser un arbre avec ces contraintes.
Ainsi, je dénombre 18 emplacements possibles pour les voyelles.
Mathilde correspond à l'emplacement 1-2-2-0



Je dispose une permutation des 3 voyelles dans leurs emplacements : $3! = 6$ choix possibles.
Je dispose une permutation des 5 consonnes dans leurs emplacements : $5! = 120$ choix possibles.
Finalement, il y a $18 \times 6 \times 120 = 12\,960$ codes possibles. Si on pense que le code respecte la disposition des voyelles du prénom original, il n'y a plus que les $6 \times 120 = 720$ codes que voici :

```python
def permute(l,s,a,n):
    if a==n:
        ana="".join(s)
        l.append(ana)
    else:
        for b in range(a,n):
            s[a],s[b]=s[b],s[a]
            permute(l,s,a+1,n)
            s[a],s[b]=s[b],s[a]

consonnes = "mthld"
voyelles = "aie"
emp=[1,2,2,0]  #emplacement des consonnes
permutations_consonnes=[]
li,nb=list(consonnes),len(consonnes)
permute(permutations_consonnes,li,0,nb)
permutations_voyelles=[]
li,nb=list(voyelles),len(voyelles)
permute(permutations_voyelles,li,0,nb)
c,v,m=len(permutations_consonnes),len(permutations_voyelles),0
for i in range(c):
    con=list(permutations_consonnes[i])
    #print(con)
    for j in range(v):
        code=""
        voy=list(permutations_voyelles[j])
        #print(voy)
        m+=1
        for k in range(emp[0]):
            code+=con[k]
        code+=voy[0]
        for k in range(emp[0],emp[0]+emp[1]):
            code+=con[k]
        code+=voy[1]
        for k in range(emp[0]+emp[1],5-emp[3]):
            code+=con[k]
        code+=voy[2]
        for k in range(emp[0]+emp[1]+emp[2],5):
            code+=con[k]
        print(m,code)
```

1 mathilde	41 mehtilda	81 milhadte	121 madlihte	161 temlihda	201 tihladme	241 taldimhe	681 dilmahte
2 matheldi	42 mehtaldi	82 milhedta	122 madlehti	162 temlahdi	202 tihledma	242 taldemhi	682 dilmehta
3 mithalde	43 mahtidle	83 melhidta	123 midlahte	163 tamlidhe	203 tehlidma	243 tildamhe	683 delmihta
4 mithelda	44 mahtedli	84 melhadti	124 midlehta	164 tamledhi	204 tehladmi	244 tildemha	684 delmahti
5 methilda	45 mihtadle	85 maltihde	125 medlihta	165 timladhe	205 tahdilme	245 teldimha	685 damhilte
6 methaldi	46 mihtedla	86 maltehdi	126 medlahti	166 timledha	206 tahdelmi		686 damhelti
7 mathidle	47 mehtidla	87 miltahde	127 madlithe	167 temlidha	207 tihdalme	......	687 dimhalte
8 mathedli	48 mehtadli	88 miltehda	128 madlethi	168 temladhi	208 tihdelma		688 dimhelta
9 mithadle	49 mahlitde	89 meltihda	129 midlathe	169 tamdilhe	209 tehdilma		689 demhilta
10 mithedla	50 mahletdi	90 meltahdi	130 midletha	170 tamdelhi	210 tehdalmi	650 dalhetmi	690 demhalti
11 methidla	51 mihlatde	91 maltidhe	131 medlitha	171 timdalhe	211 tahdimle	651 dilhatme	691 damhitle
12 methadli	52 mihletda	92 maltedhi	132 medlathi	172 timdelha	212 tahdemli	652 dilhetma	692 damhetli
13 matlihde	53 mehlitda	93 miltadhe	133 madtilhe	173 temdilha	213 tihdamle	653 delhitma	693 dimhatle
14 matlehdi	54 mehlatdi	94 miltedha	134 madtelhi	174 temdalhi	214 tihdemla	654 delhatmi	694 dimhetla
15 mitlahde	55 mahlidte	95 meltidha	135 midtalhe	175 tamdihle	215 tehdimla	655 dalhimte	695 demhitla
16 mitlehda	56 mahledti	96 meltadhi	136 midtelha	176 tamdehli	216 tehdamli	656 dalhemti	696 demhatli
17 metlihda	57 mihladte	97 maldithe	137 medtilha	177 timdahle	217 talhimde	657 dilhamte	697 damlihte
18 metlahdi	58 mihledta	98 maldethi	138 medtalhi	178 timdehla	218 talhemdi	658 dilhemta	698 damlehti
19 matlidhe	59 mehlidta	99 mildathe	139 madtihle	179 temdihla	219 tilhamde	659 delhimta	699 dimlahte
20 matledhi	60 mehladti	100 mildetha	140 madtehli	180 temdahli	220 tilhemda	660 delhamti	700 dimlehta
21 mitladhe	61 mahdilte	101 melditha	141 midtahle	181 tahmilde	221 telhimda	661 daltihme	701 demlihta
22 mitledha	62 mahdelti	102 meldathi	142 midtehla	182 tahmeldi	222 telhamdi	662 daltehmi	702 demlahti
23 metlidha	63 mihdalte	103 maldihte	143 medtihla	183 tihmalde	223 talhidme	663 diltahme	703 damlithe
24 metladhi	64 mihdelta	104 maldehti	144 medtahli	184 tihmelda	224 talhedmi	664 diltehma	704 damlethi
25 matdilhe	65 mehdilta	105 mildahte	145 tamhilde	185 tehmilda	225 tilhadme	665 deltihma	705 dimlathe
26 matdelhi	66 mehdalti	106 mildehta	146 tamheldi	186 tehmaldi	226 tilhedma	666 deltahmi	706 dimletha
27 mitdalhe	67 mahditle	107 meldihta	147 timhalde	187 tahmidle	227 telhidma	667 daltimhe	707 demlitha
28 mitdelha	68 mahdetli	108 meldahti	148 timhelda	188 tahmedli	228 telhadmi	668 daltemhi	708 demlathi
29 metdilha	69 mihdatle	109 madhilte	149 temhilda	189 tihmadle	229 talmihde	669 diltamhe	709 damtilhe
30 metdalhi	70 mihdetla	110 madhelti	150 temhaldi	190 tihmedla	230 talmehdi	670 diltemha	710 damtelhi
31 matdihle	71 mehditla	111 midhalte	151 tamhidle	191 tehmidla	231 tilmahde	671 deltimha	711 dimtalhe
32 matdehli	72 mehdatli	112 midhelta	152 tamhedli	192 tehmadli	232 tilmehda	672 deltamhi	712 dimtelha
33 mitdahle	73 malhitde	113 medhilta	153 timhadle	193 tahlimde	233 telmihda	673 dalmithe	713 demtilha
34 mitdehla	74 malhetdi	114 medhalti	154 timhedla	194 tahlemdi	234 telmahdi	674 dalmethi	714 demtalhi
35 metdihla	75 milhatde	115 madhitle	155 temhidla	195 tihlamde	235 talmidhe	675 dilmathe	715 damtihle
36 metdahli	76 milhetda	116 madhetli	156 temhadli	196 tihlemda	236 talmedhi	676 dilmetha	716 damtehli
37 mahtilde	77 melhitda	117 midhatle	157 tamlihde	197 tehlimda	237 tilmadhe	677 delmitha	717 dimtahle
38 mahteldi	78 melhatdi	118 midhetla	158 tamlehdi	198 tehlamdi	238 tilmedha	678 delmathi	718 dimtehla
39 mihtalde	79 malhidte	119 medhitla	159 timlahde	199 tahlidme	239 telmidha	679 dalmihte	719 demtihla
40 mihtelda	80 malhedti	120 medhatli	160 timlehda	200 tahledmi	240 telmadhi	680 dalmehti	720 demtahli

2.e. Combinaisons

DÉFINITION 1.6 (COMBINAISON) Si E est un ensemble de cardinal n et un entier $k \in [\![0;n]\!]$. Une *combinaison* de k éléments de E est une liste non-ordonnée d'éléments distincts de E, c'est-à-dire un *sous-ensemble* de k éléments de E ou une *partie* à k éléments de E.

PROPRIÉTÉ 1.7 Soient E est un ensemble de cardinal n et $k \in [\![0;n]\!]$.
Le nombre de combinaisons de k éléments de E – noté C_n^k ou $\begin{pmatrix} n \\ k \end{pmatrix}$ (lire k parmi n) – est :

$$\frac{n \times (n-1) \times (n-2) \times \cdots \times (n-k+1)}{k \times (k-1) \times (k-2) \times \cdots \times 2 \times 1} = \frac{n!}{k!(n-k)!}$$

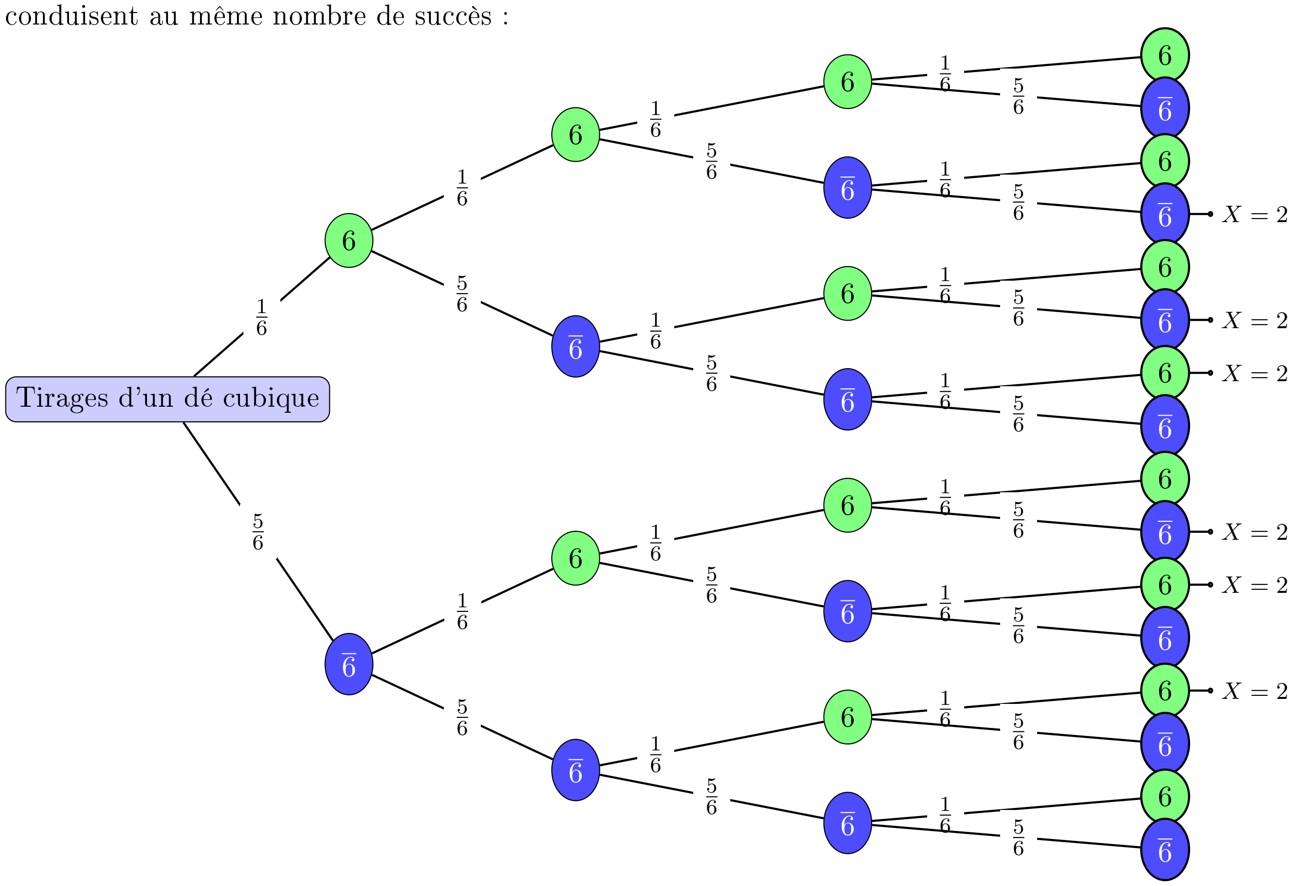
DÉMONSTRATION Les k éléments d'un sous-ensemble de E ne sont pas ordonnés :
le sous-ensemble $\{e_1, e_2, \ldots, e_k\}$ contient les mêmes éléments que les $k!$ permutations de ces éléments.
Par exemple, si $k = 3$, le sous-ensemble $\{e_1, e_2, e_3\}$ contient les mêmes trois éléments que ses $3! = 6$ permutations (e_1, e_2, e_3), (e_1, e_3, e_2), (e_2, e_1, e_3), (e_2, e_3, e_1), (e_3, e_1, e_2), (e_3, e_2, e_1).
Les C_n^k sous-ensembles de k éléments de E sont chacun en correspondance avec $k!$ permutations.
Il y a donc au total $C_n^k \times k!$ permutations, toutes distinctes, qui forment l'ensemble des k-listes sans répétition ou arrangements de k éléments de E.
Comme on sait qu'il en existe A_n^k, on en déduit :

$$A_n^k = C_n^k \times k! \Longleftrightarrow C_n^k = \frac{A_n^k}{k!} = \frac{n!}{k!(n-k)!}$$

Remarques :

- La notation est importante : des accolades pour les combinaisons, des parenthèses pour les arrangements. On retrouve d'ailleurs cette notation en Python ou les `list` sont des listes ordonnées (avec cette différence qu'un même élément peut y être répété) tandis que les `set` sont des ensembles non-ordonnés (un même élément ne peut y être répété).
  Exemple : $\{2, 4, 6\}$ est le sous-ensemble des éléments pairs des numéros d'un dé.
  $\{6, 4, 2\}$ est le même sous-ensemble, la même combinaison.
  Par contre, $(2, 4, 6)$ et $(6, 4, 2)$ ne sont pas les mêmes arrangements de ces trois éléments.
  En Python toujours, pour ajouter un élément à une liste, on utilise la méthode `append` qui l'ajoute à la fin ou `insert` qui l'insère à une position donnée ; pour ajouter un élément dans un ensemble on utilise la méthode `add` qui l'ajoute n'importe où.
- Dans toutes les situations où on rencontre des combinaisons, l'ordre n'a pas d'importance : les éléments d'une combinaison peuvent être réarrangés sans que celle-ci ne soit modifiée.
  Au poker, par exemple, un joueur reçoit une combinaison de 5 cartes choisies au hasard parmi 32 ; peu importe l'ordre dans lequel le joueur arrange ses cartes. Lors du tirage simultané de k boules d'une urne en contenant n, on obtient une combinaison de k boules.
  Trouver la combinaison gagnante au Loto consiste à trouver le sous-ensemble des numéros gagnants. Dans l'ancienne formule, on devait choisir 6 numéros d'une grille en contenant 49 ; il s'agit d'une partie à 6 éléments choisis parmi 49. Leur nombre est $\binom{49}{6} = \frac{49!}{6! \times 43!}$.
  Il ne faut pas calculer les nombres séparément car 49! s'écrit avec 63 chiffres (!)
  $\binom{49}{6} = \frac{49 \times 48 \times 47 \times 46 \times 45 \times 44}{6 \times 5 \times 4 \times 3 \times 2 \times 1} = 13\,983\,816$.
  Chaque grille étant équiprobable, il y avait une chance sur 13 983 816 de gagner le gros lot.
- Quelques propriétés immédiates de ces nombres C_n^k :
  — $C_n^0 = 1$, le seul sous-ensemble à 0 élément est $\emptyset$ (l'ensemble vide)
  — $C_n^1 = n$, il y a n façons de choisir un élément d'un ensemble à n éléments
  — $C_n^{n-1} = n$, il y a n façons de choisir l'élément qui est enlevé aux n éléments
  — $C_n^n = 1$, un seul sous-ensemble à n éléments d'un ensemble E à n éléments : E lui-même

Exemple 7 – On lance quatre fois de suite un dé équilibré à six faces. On appelle succès l'obtention d'un 6 et on considère la variable aléatoire X qui comptabilise le nombre de succès.
On verra, dans le chapitre « Probabilités », que cette variable suit une *loi binomiale* $\mathscr{B}(4, \frac{1}{6})$.
Mais on peut dors et déjà construire l'arbre aux $2^4 = 16$ chemins qui permet de dénombrer ceux qui conduisent au même nombre de succès :



- 1 seul chemin mène à 0 succès et 1 seul mène à 4 succès.
- 4 mènent à 1 succès et 4 mènent à 3 succès. Il suffit de choisir l'emplacement du dé qui conduit ou ne conduit pas au succès ; il y a 4 emplacements possibles.
- total provisoire : $1 + 1 + 4 + 4 = 10$ chemins ; il en reste $16 - 10 = 6$ qui réalisent 2 succès. Je les ai marqué sur l'illustration. Leur nombre correspond au nombre de combinaisons de 2 emplacements parmi 4 : il y en a 6, ce sont $\{1, 2\}$, $\{1, 3\}$, $\{1, 4\}$, $\{2, 3\}$, $\{2, 4\}$ et $\{3, 4\}$.

Les nombres de combinaisons $C_n^k = \begin{pmatrix} n \\ k \end{pmatrix}$ de $k \in [\![0, 4]\!]$ éléments parmi 4 sont :

$C_4^0 = \begin{pmatrix} 4 \\ 0 \end{pmatrix} = \mathbf{1}$, $C_4^1 = \begin{pmatrix} 4 \\ 1 \end{pmatrix} = \mathbf{4}$, $C_4^2 = \begin{pmatrix} 4 \\ 2 \end{pmatrix} = \mathbf{6}$, $C_4^3 = \begin{pmatrix} 4 \\ 3 \end{pmatrix} = \mathbf{4}$ et $C_4^4 = \begin{pmatrix} 4 \\ 4 \end{pmatrix} = \mathbf{1}$

Ces nombres sont les coefficients du développement du binôme $(a + b)^4$:
$(a + b)^4 = (a^2 + 2ab + b^2)^2 = \mathbf{1} \times a^4 + \mathbf{4} \times a^3b + \mathbf{6} \times a^2b^2 + \mathbf{4} \times ab^3 + \mathbf{1} \times b^4$

Pour cette raison, les nombres $C_n^k = \begin{pmatrix} n \\ k \end{pmatrix}$ sont appelés « coefficients binomiaux ». Ils s'appliquent au développement du binôme $(a + b)^n$ car, lorsqu'on doit calculer le coefficient du terme $a^k b^{n-k}$, on dénombre des parties à k éléments parmi n (les k facteurs où on prend le nombre a). Pour la loi binomiale, les parties à k éléments parmi n sont les places des k succès.

Propriété 1.8 Soient n et k deux entiers :

- si $k \leqslant n$ alors $\begin{pmatrix} n \\ k \end{pmatrix} = \begin{pmatrix} n \\ n - k \end{pmatrix}$
- si $k < n$ alors $\begin{pmatrix} n \\ k \end{pmatrix} + \begin{pmatrix} n \\ k + 1 \end{pmatrix} = \begin{pmatrix} n + 1 \\ k + 1 \end{pmatrix}$

DÉMONSTRATION On va raisonner ici dans le cadre de l'exemple précédent où on considère la répétition de n expériences aléatoires pour lesquelles deux issues sont possibles : le succès et l'échec (épreuves de Bernoulli).

- Il y a autant de chemins qui mènent à k succès que de chemins qui mènent à k échecs (c'est-à-dire à $n-k$ succès). Les qualificatifs de *succès* et *échec* sont subjectifs ; en permutant les mots, on ne change pas le nombre de leurs occurrences.
- Le coefficient $\binom{n+1}{k+1}$ est le nombre de chemins qui mènent à $k+1$ succès dans une répétition de $n+1$ épreuves de Bernoulli. À l'issue des n premières épreuves, il n'y a que trois possibilités :
  - on a déjà eu $k+1$ succès. Il y a $\binom{n}{k+1}$ possibilités, suivies d'un échec à la dernière étape
  - on a déjà eu k succès. Il y a $\binom{n}{k}$ possibilités, suivies d'un succès à la dernière étape
  - dans les autres cas (moins de k succès ou plus de $k+1$ succès), on ne peut aboutir à $k+1$ succès

  Finalement, le nombre de chemins cherché est la somme des dénombrements associés aux deux premières possibilités : $\binom{n}{k}+\binom{n}{k+1}$.

  Remarquons qu'il est très simple de prouver cette relation algébriquement.
  Il suffit d'écrire les deux expressions avec le même dénominateur, en remarquant que $(n-(k+1)) = (n-k-1)$ précède $(n-k)$:

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

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

Le triangle de Pascal[4] exploite ces propriétés pour le calcul des coefficients $\binom{n}{k}$ de proche en proche. Chaque nombre est obtenu en additionnant les deux nombres de la ligne du dessus (2[e] propriété 1.8) : celui de la même colonne et celui de la colonne précédente comme le montre sur fond jaune le $\binom{5}{2}=10$ qui est égal à la somme $\binom{4}{1}+\binom{4}{2}=4+6$.
Il suffit d'ajouter à cette propriété la 1[re] colonne de 1 qui vient de la propriété $C_n^0=1$ et la diagonale de 1 qui vient de la propriété $C_n^n=1$ pour construire le triangle en entier.

n \ k	0	1	2	3	4	5	...
0	1						
1	1	1					
2	1	2	1				
3	1	3	3	1			
4	1	4	6	4	1		
5	1	5	10	10	5	1	
$\vdots$	$\vdots$						$\ddots$

Utilisation de la calculatrice :
La calculatrice permet de déterminer les coefficients binomiaux $\binom{n}{k}$.

	TI	Casio	Numworks
$\binom{n}{k}$	touche math option PRB choisir 3 (combinaison)	menu RUN option PROB(F3) choisir nCr	menu Calculs, Toolbox (5[e] touche), Dénombrement, binomial(n,k), entrer n et k

Utilisation du tableur :
Avec le tableur *calc* d'OpenOffice, on a accès aux coefficients binomiaux $\binom{n}{k}$ en tapant la formule =COMBIN(n ;k)

4. Blaise Pascal (1623-1662) : philosophe et mathématicien français. Le triangle qui porte son nom était connu au Moyen-âge, par des mathématiciens persans et chinois et plus tard en Europe, au XVI[e] siècle, par Tartaglia et Stiefel. Pascal en étudia les propriétés dans son « Traité du triangle arithmétique (1654).

2.f. Développement du binôme

On l'a dit plus haut, les nombres $C_n^k = \binom{n}{k}$ sont appelés « coefficients binomiaux » car ce sont ceux qui interviennent dans le développement du binôme $(a+b)^n$: lorsqu'on doit calculer le coefficient du terme $a^k b^{n-k}$, on dénombre des parties à k éléments parmi n. Cet argument est de nature combinatoire mais on peut directement prouver ce résultat par un raisonnement algébrique.

DÉFINITION 1.7 (FORMULE DE NEWTON) Soient a et b deux réels et n un entier quelconque.
$(a+b)^n = a^n + na^{n-1}b + \cdots + \binom{n}{k} a^k b^{n-k} + \cdots + nab^{n-1} + b^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}$

DÉMONSTRATION Démontrons que cette propriété est vraie pour tout $n \in \mathbb{N}$ par récurrence.
Initialisation : Pour $n = 0$, on a $(a+b)^0 = 1$ or $\sum_{k=0}^{0} \binom{0}{k} a^k b^{-k} = \binom{0}{0} a^0 b^0 = 1$, la propriété est vraie.
Hérédité : Supposons que pour un entier $n \geqslant 0$ on a $(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}$
$(a+b)^{n+1} = (a+b) \times (a+b)^n = (a+b) \times \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}$
Développons et arrangeons cette somme pour utiliser la propriété de Pascal $\binom{n}{k} + \binom{n}{k+1} = \binom{n+1}{k+1}$ sous la forme $\binom{n}{k-1} + \binom{n}{k} = \binom{n+1}{k}$ (j'utilise aussi le changement de variable $k' = k+1$) :

$$\begin{aligned}
(a+b)^{n+1} &= a\sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k} + b\sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k} \\
&= \sum_{k=0}^{n} \binom{n}{k} a^{k+1} b^{n-k} + \sum_{k=0}^{n} \binom{n}{k} a^k b^{n+1-k} \\
&= \left(a^{n+1} + \sum_{k=0}^{n-1} \binom{n}{k} a^{k+1} b^{n-k}\right) + \left(\sum_{k=1}^{n} \binom{n}{k} a^k b^{n+1-k} + b^{n+1}\right) \\
&= a^{n+1} + \sum_{k'=1}^{n} \binom{n}{k'-1} a^{k'} b^{n+1-k'} + \sum_{k=1}^{n} \binom{n}{k} a^k b^{n+1-k} + b^{n+1} \\
&= a^{n+1} + \left(\sum_{k=1}^{n} \binom{n}{k-1} + \sum_{k=1}^{n} \binom{n}{k}\right) a^k b^{n+1-k} + b^{n+1} \\
&= a^{n+1} + \sum_{k=1}^{n} \left(\binom{n}{k-1} + \binom{n}{k}\right) a^k b^{n+1-k} + b^{n+1} \\
&= a^{n+1} + \sum_{k=1}^{n} \binom{n+1}{k} a^k b^{n+1-k} + b^{n+1} = \sum_{k=0}^{n+1} \binom{n+1}{k} a^k b^{n+1-k}
\end{aligned}$$

Ainsi la propriété est bien héréditaire.
Par principe de récurrence elle est donc vraie pour tout $n \in \mathbb{N}$.

Cas particuliers :

- En prenant $a = b = 1$ on obtient $2^n = \sum_{k=0}^{n} \binom{n}{k}$ ce qui montre que le nombre des combinaisons à $0, 1, \ldots, n$ éléments – soit le nombre de parties à k éléments avec $k \in [\![0, n]\!]$ – est 2^n.
- En prenant $b = 1$ on obtient $(1+a)^n = \sum_{k=0}^{n} \binom{n}{k} a^k = 1 + na + \sum_{k=2}^{n} \binom{n}{k} a^k$ et comme $\binom{n}{k} > 0$ pour $2 \leqslant k \leqslant n$, pour tout réel $a > 0$ on a $(1+a)^n > 1 + na$ (inégalité de Bernoulli), cette inégalité est vraie au sens large pour $n \in \mathbb{N}$ car pour $n = 0$ et $n = 1$ elle s'écrit $1 \geqslant 1$.

3. Suites

Définition 1.8 (suite numérique) Une suite est une fonction définie de $\mathbb{N}$ vers $\mathbb{R}$ à partir d'un certain rang, noté n_0. L'image d'un entier n par la suite u est notée u_n (lire « u indice n »). u_n est le « terme général » de la suite. Celle-ci est notée $(u_n)_{n \geqslant n_0}$.

Remarques :

- Généralement, une suite numérique commence à partir de $n = 0$: son 1er terme est alors u_0, le 2e terme est u_1, etc. Le ne terme est u_{n-1}. Elle peut être notée $(u_n)_{n \in \mathbb{N}}$. La suite de terme général $u_n = 2n$ est définie à partir de $n = 0$, c'est la suite des nombres pairs.
- Certaines suites ne sont définies qu'à partir d'un rang $n_0 \neq 0$, par exemple à partir de $n = 1$. Le ne terme de la suite est alors le terme de rang $n_0 + n - 1$. La suite de terme général $u_n = \frac{1}{n}$ est définie à partir de $n = 1$, c'est la suite des inverses d'entiers. La suite de terme général $v_n = \frac{1}{\sqrt{n-2}}$ n'est définie qu'à partir du rang $n = 3$, on la note $(v_n)_{n \geqslant 3}$.
- La notation indicielle est une spécificité des suites qui est commode et fixée par l'usage. On pourrait définir une suite (u) comme une fonction de $\mathbb{N}$ vers $\mathbb{R}$ (par exemple $u(n) = n^2$) mais on préfère définir une fonction f sur $\mathbb{R}$ ($f(x) = x^2$) et ensuite le terme général $u_n = f(n)$.

Méthode (Types de définition :)

- Explicite : soient $a \in \mathbb{N}$, et f une fonction définie sur l'intervalle $[a; +\infty[$.
  La suite $(u_n)_{n \geqslant a}$ de terme général $u_n = f(n)$ est définie *explicitement* par la fonction f.
- Implicite : une suite $(u_n)_{n \geq n_0}$ est définie *implicitement* si le terme général est donné par une description, valable à partir du rang n_0, qui permet de l'identifier, sans forcément fournir le moyen de le déterminer. Exemple : s_n est la somme des chiffres de n en base dix. Il n'y a pas d'expression algébrique qui corresponde à cette notion, pourtant on peut déterminer n'importe que terme de cette suite : $s_{101} = 2$, $s_{123\,456\,789} = 45$, etc.
- Récurrence : soient f une fonction définie sur un intervalle I telle que $f(I) \subset I$ et $a \in I$.
  La suite u définie par : $\begin{cases} u_0 = a \\ \forall n \in \mathbb{N}, u_{n+1} = f(u_n) \end{cases}$ est définie *par une récurrence simple.*
  On peut aussi définir une suite par une relation de récurrence portant sur plusieurs termes précédents, comme dans l'exemple 2 où la relation de récurrence $u_n = u_{n-1} + u_{n-2}$ implique de connaître les deux premiers termes pour calculer les autres.
- Mixte : une suite définie par une expression contenant à la fois une ou plusieurs valeur(s) précédente(s) et la valeur du rang. Par exemple, la suite (u_n) des factorielles $u_n = n!$ peut être définie par la relation mixte $u_n = u_{n-1} \times n$ et la condition initiale $u_0 = 1$.

Remarques :

- Une suite définie explicitement peut gagner à être définie par une récurrence ou une relation mixte, pour un calcul algorithmique plus efficace (ex : factorielles). Réciproquement, on cherche souvent à déterminer explicitement une suite définie par récurrence (exemple 2).
- Une même fonction f peut définir une infinité de suites récurrentes différentes car changer la valeur initiale change la suite. La suite des nombres pairs $u_n = 2n$ peut être définie par la relation de récurrence : $u_n = u_{n-1} + 2$ et la condition initiale $u_0 = 0$, mais la même fonction $(x \mapsto x + 2)$ avec la valeur initiale $u_0 = 1$ sert à définir la suite des nombres impairs.
- L'étude d'une suite peut bénéficier de suites intermédiaires. La suite (p_n) définie par $p_0 = 2$ et $p_n = \frac{1}{1+p_{n-1}}$ peut bénéficier de l'étude des suites (N_n) et (D_n) telles que $p_n = \frac{N_n}{D_n}$.
  On choisit $N_0 = 2$ et $D_0 = 1$ et comme $p_n = \frac{N_n}{D_n} = \frac{1}{1+\frac{N_{n-1}}{D_{n-1}}} = \frac{D_{n-1}}{D_{n-1}+N_{n-1}}$, on obtient les relations $N_n = D_{n-1}$ et $D_n = D_{n-1} + N_{n-1} = D_{n-1} + D_{n-2}$ qui permettent, à l'aide d'un programme, de déterminer la valeur exacte $p_{100} = \frac{792\,070\,839\,848\,372\,253\,127}{1\,281\,597\,540\,372\,340\,914\,251}$ (la définition initiale utilise des valeurs approchées).

3.a. Propriétés

DÉFINITION 1.9 (SENS DE VARIATION) Une suite est croissante (respect. strictement croissante) si, à partir d'un certain rang N, on a $u_{n+1} \geqslant u_n$ (respect. $u_{n+1} > u_n$).

Remarques :
Même définition pour une suite décroissante ou strictement décroissante, il suffit de mettre $\leqslant$ ou $<$.
Une suite est (strictement) *monotone* si elle est (strictement) croissante ou décroissante.
Une suite est *constante* si à partir d'un certain rang, on a $u_{n+1} = u_n$.

PROPRIÉTÉ 1.9 (CRITÈRES DE MONOTONIE) Une suite (u) est croissante si et seulement si :

$$\begin{array}{rl} \text{(i)} & \forall n \geqslant N, u_{n+1} - u_n \geqslant 0 \\ \text{(ii)} & \forall n \geqslant N, u_n > 0 \text{ et } \frac{u_{n+1}}{u_n} \geqslant 1 \\ \text{(iii)} & \forall n \geqslant N, u_n < 0 \text{ et } \frac{u_{n+1}}{u_n} \leqslant 1 \end{array}$$

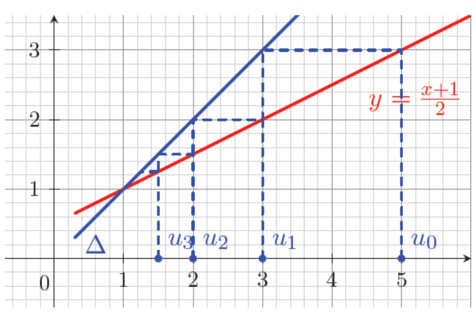
DÉMONSTRATION Évident, les critères ne faisant que traduire la définition 1.9.

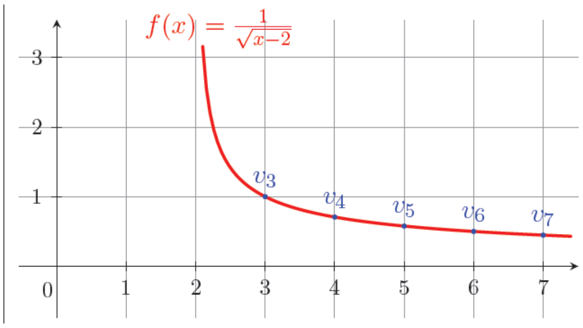
Remarque :
Changer le sens des inégalités pour montrer qu'une suite est décroissante (pour une monotonie stricte prendre des inégalités strictes) : $u_{n+1} - u_n \leqslant 0$; $u_n > 0$ et $\frac{u_{n+1}}{u_n} \leqslant 1$; $u_n < 0$ et $\frac{u_{n+1}}{u_n} \geqslant 1$

EXEMPLE 8 – ➲ Soit (u) la suite définie par $u_{n+1} = \frac{u_n+1}{2}$ et $u_0 = 5$.
Montrons que (u) est strictement décroissante à partir de $n = 0$ à l'aide du critère (i) :
$u_{n+1} - u_n = \frac{1-u_n}{2}$. Ainsi, $\forall n \geqslant 0$, $u_{n+1} - u_n < 0 \Longleftrightarrow \frac{1-u_n}{2} < 0 \Longleftrightarrow 1 - u_n < 0 \Longleftrightarrow u_n > 1$.
Initialisation : $u_0 = 5 > 1$.
Hérédité : Supposons $u_n > 1$ $(\star)$ pour $n > 0$. Pour $n+1$ on a $u_n + 1 > 2 \Longleftrightarrow \frac{u_n+1}{2} > \frac{2}{2}$, soit $u_{n+1} > 1$.
Par principe de récurrence la propriété est donc vraie pour tout $n \in \mathbb{N}$.

➲ Montrons que la suite (u) est strictement décroissante à l'aide du critère (ii) :
Ce critère nécessite qu'on s'assure de deux inégalités : $\forall n \in \mathbb{N}, u_n > 0$ et $\frac{u_{n+1}}{u_n} < 1$.
On a déjà montré $(\star)$ que $\forall n, u_n > 1$ donc, à fortiori, on a $\forall n, u_n > 0$.
Pour l'autre inégalité, $\frac{u_{n+1}}{u_n} = \frac{u_n+1}{2u_n} = \frac{1}{2} + \frac{1}{2u_n}$.
Or $u_n > 1 \Longleftrightarrow 2u_n > 2 \Longleftrightarrow \frac{1}{2u_n} < \frac{1}{2}$ car la fonction inverse est décroissante sur $]0; \infty[$.
On en déduit que $\frac{1}{2} + \frac{1}{2u_n} < \frac{1}{2} + \frac{1}{2} = 1$, soit $\frac{u_{n+1}}{u_n} < 1$.

➲ Montrons que la suite (v) définie par $v_n = \frac{1}{\sqrt{n-2}}$ est strictement décroissante à partir de $n = 3$.
On a $\frac{v_{n+1}}{v_n} = \frac{\frac{1}{\sqrt{n+1-2}}}{\frac{1}{\sqrt{n-2}}} = \frac{\sqrt{n-2}}{\sqrt{n-1}} = \sqrt{\frac{n-2}{n-1}}$.
Comme $n > 2$ on a $n - 2 > 0$, $n - 1 > n - 2$ et donc $0 < n - 2 < n - 1 \Longleftrightarrow \frac{n-2}{n-1} < 1$.
La fonction racine étant croissante sur $[0; \infty[$, on en déduit que $\sqrt{\frac{n-2}{n-1}} < \sqrt{1}$, soit $\frac{v_{n+1}}{v_n} < 1$.
Comme, de plus, $\forall n > 2, v_n > 0$, d'après (ii) la suite (v) est strictement décroissante.





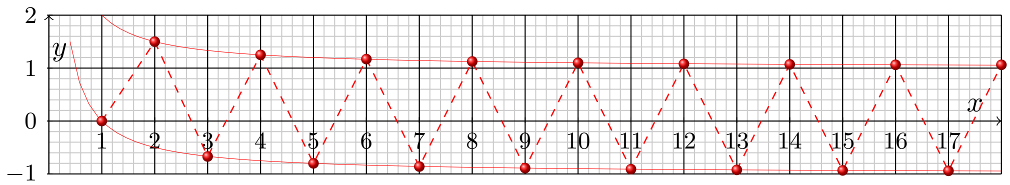
Ces suites sont représentées graphiquement ci-dessus de deux façons différentes.

Définition 1.10 (suite bornée) Une suite $(u_n)_{n\geqslant n_0}$ est *majorée* si $\exists M\in\mathbb{R},\forall n\geqslant n_0, u_n\leqslant M$.
Une suite $(u_n)_{n\geqslant n_0}$ est *minorée* si $\exists m\in\mathbb{R},\forall n\geqslant n_0, u_n\geqslant m$.
Une suite $(u_n)_{n\geqslant n_0}$ est *bornée* si elle est minorée *et* majorée.

Remarque :
Les nombres m et M qui interviennent dans cette définition sont appelés « minorant » et « majorant ». Il n'y a pas unicité de ces nombres : si une suite est majorée par M alors elle est majorée aussi par tous les nombres supérieurs à M. Il n'y a pas de plus grand majorant.
Par contre, une suite majorée admet toujours un plus petit majorant, appelé *borne supérieure* ; de même, une suite minorée admet toujours un plus grand minorant, appelé *borne inférieure*.
La suite des troncatures décimales de π admet pour borne supérieure le nombre π.

Exemples :

- La suite récurrente (u_n) définie par $u_{n+1}=\frac{u_n+1}{2}$ et $u_0=5$ est bornée car majorée par son premier terme $u_0=5$ et minorée par 1 (voir l'exemple 8). Ce minorant est également sa borne inférieure qui n'est jamais atteinte, sa limite quand n tend vers l'infini (voir plus loin).
- La suite $(t_n)_{n\geqslant 1}$, définie par $t_n=(-1)^n+\frac{1}{n}$ admet -1 comme borne inférieure et $\frac{3}{2}$ comme borne supérieure. Sur la représentation graphique ci-dessous, on remarque deux sous-suites :
  — la suite (p), associée à la fonction $f:x\mapsto -1+\frac{1}{x}$ contient tous les termes de rang pair
  — la suite (i) associée à la fonction $g:x\mapsto 1+\frac{1}{x}$ contient tous les termes de rang impair
  Les suites (p) et (i) sont décroissantes mais la suite (t), quant-à elle, n'est pas monotone.
  — Sa borne supérieure est le maximum des bornes supérieures des deux sous-suites.
  — Sa borne inférieure est le minimum des bornes supérieures des deux sous-suites.



Définition 1.11 (suite périodique) Une suite (u_n) est périodique à partir du rang N s'il existe un entier $p\geqslant 1$ tel que $\forall n\geqslant N, u_{n+p}=u_n$. L'entier p est alors la « période » de u.

Remarque :
L'entier p appelé période est « le plus petit » entier non nul vérifiant la définition 1.11. Il est évident qu'une suite de période $p>0$ vérifie aussi la propriété pour $2p$, $3p$, etc. La suite de terme général $(-1)^n$ est périodique de période 2 à partir de $n=0$, car $\forall n\geqslant 0, u_{n+2}=u_n$. Cette suite vérifie aussi $u_{n+4}=u_n$, mais la période est 2 et non 4.

Exemples :

- Les suites constantes, ou constantes à partir d'un certain rang, sont des suites périodiques de période 1. La suite q des décimales de $\frac{1}{15}=0,0666\ldots$ est une suite constante, donc périodique de période 1, à partir du rang 2 : $\forall n\geqslant 2,\ \ q_{n+1}=q_n$.
- La suite (p_n) donnant la n^{e} décimale d'un nombre rationnel non décimal est périodique à partir d'un certain rang. Celle qui est associée à $\frac{1}{7}=0,\overline{142857}$ (la séquence 142857 se répète) est périodique à partir du rang 1 ; sa période est 6. On a
  $\forall k\in\mathbb{N}, p_1=p_7=p_{6k+1}=1,\ \ p_2=p_8=p_{6k+2}=4,\ \ \ldots,\ \ p_6=p_{12}=p_{6k}=7.$
- Si k est un entier non nul, les suites c et s définies par $c_n=\cos(\frac{2n\pi}{k})$ et $s_n=\sin(\frac{2n\pi}{k})$ sont périodiques de période k car cos et sin sont des fonctions périodiques de période 2π :
  $c_{n+k}=\cos(\frac{2(n+k)\pi}{k})=\cos(\frac{2n\pi}{k}+2\pi)=\cos(\frac{2n\pi}{k})=c_n$ (même chose pour la suite s).
- La suite de Syracuse S d'un entier $S_0=p$ est un exemple intéressant de suite « finalement » périodique. Définie par $S_{n+1}=\frac{S_n}{2}$ si S_n pair, $3S_n+1$ si S_n impair, cette suite a un comportement périodique de période 3 qui tarde parfois à se manifester, le cycle final étant $[4,2,1]$ (cela reste une conjecture).

3.b. Limite d'une suite

3.b.1. Convergence

DÉFINITION 1.12 (SUITE CONVERGENTE) Une suite (u_n) converge vers une limite l si, à partir d'un certain rang N, tout intervalle ouvert contenant l contient toutes les valeurs u_n. Autrement dit : $\lim_{n \to +\infty} u_n = l \iff \forall \epsilon \in \mathbb{R}_+^*, \ \exists N \in \mathbb{N}, \ \forall n \in \mathbb{N}, \ n \geqslant N \Longrightarrow l - \epsilon < u_n < l + \epsilon$
Si la suite (u_n) converge ver l, on note $\lim_{n \to +\infty} u_n = l$.

Remarque :
À la fin de la dernière ligne, à la place de $l - \epsilon < u_n < l + \epsilon$, on aurait pu écrire $|u_n - l| < \epsilon$ ou encore $u_n \in]l - \epsilon; l + \epsilon[$ ou, en utilisant la notion de distance entre u_n et l, $d(u_n, l) < \epsilon$. Ces écritures traduisent l'expression *un intervalle ouvert contenant l contient le terme u_n*. Ici, l'intervalle est centré sur l, le réel positif ϵ est son rayon. Plus ϵ est petit, plus on s'approche de l. En exigeant qu'à partir d'un certain rang N tous les termes u_n soient dans cet intervalle, quelle que soit la petitesse de ϵ, on traduit analytiquement la notion intuitive de limite.

Suites de référence convergeant vers zéro :
La suite $(\frac{1}{n})_{n>0}$ des inverses d'entiers converge vers 0 : $\lim_{n \to +\infty} \frac{1}{n} = 0$.
Cette affirmation est fondée sur l'observation que les inverses d'entiers deviennent de plus en plus petites, tout en restant supérieures à zéro. Pour prouver cela, choisissons une valeur arbitraire de la précision, par exemple $\epsilon = 10^{-6}$, et montrons qu'il existe un rang N à partir duquel toutes les inverses sont plus petites que cette valeur : $\frac{1}{n} < 10^{-6} \iff n > \frac{1}{10^{-6}} \iff n > 10^6$; il suffit donc de choisir comme valeur de N le premier entier supérieur à 10^6, soit $10^6 + 1 = 1000001$.

Cela se généralise à toute valeur de ϵ car, comme $n > 0$ et $\epsilon > 0$, on a : $\frac{1}{n} < \epsilon \iff n > \frac{1}{\epsilon}$.
Le premier entier supérieur à $\frac{1}{\epsilon}$ convient et donc, pour tout entier $n \geqslant \frac{1}{\epsilon}$, on aura bien $0 < \frac{1}{n} < \epsilon$.

De la même façon, les suites $(\frac{1}{n^2})$, $(\frac{1}{n^3})$, etc., $(\frac{1}{\sqrt{n}})$ convergent toutes vers 0.

DÉFINITION 1.13 (SUITE DE LIMITE $+\infty$) Une suite (u_n) *diverge* et admet $+\infty$ comme limite si, à partir d'un certain rang N, tout intervalle de la forme $]A, +\infty[$ contient toutes les valeurs u_n. Autrement dit : $\lim_{n \to +\infty} u_n = +\infty \iff \forall A \in \mathbb{R}, \ \exists N \in \mathbb{N}, \ \forall n \in \mathbb{N}, \ n \geqslant N \Longrightarrow u_n > A$

Remarques :
- On a une définition analogue pour les suites de limite $-\infty$:
  $\lim_{n \to +\infty} u_n = -\infty \iff \forall A \in \mathbb{R}, \ \exists N \in \mathbb{N}, \ \forall n \in \mathbb{N}, \ n \geqslant N \Longrightarrow u_n < A$
- Une suite qui ne converge pas diverge. Il y a des suites qui divergent vers une limite infinie et des suites qui divergent sans avoir de limite comme la suite $((-1)^n)$ qui est périodique de période 2 ($u_{2k} = 1$ et $u_{2k+1} = -1$) ou la suite $((-1)^n + \frac{1}{n}))_{n>0}$ de l'illustration précédente.

Suites de référence divergeant vers l'infini :
La suite $(n^2)_{n \geqslant 0}$ des carrés diverge vers $+\infty$: $\lim_{n \to +\infty} n^2 = +\infty$.
Cette affirmation est fondée sur l'observation que les carrés des entiers deviennent de plus en plus grands et finissent par dépasser toute valeur fixée à l'avance. Pour prouver cela, choisissons un maximum arbitraire $A > 0$, par exemple $A = 10^6$, et montrons qu'il existe un rang N à partir duquel tous les carrés sont plus grands que cette valeur. Pour $A = 10^6$, il suffit de choisir $N = 1001$ car $\forall n \in \mathbb{N}, n \geqslant 1001 \Longrightarrow u_n \geqslant 1001^2 > 10^6$.
On généralise cela à toutes les valeurs possibles de A.
Comme n et A sont positifs tous les deux : $n^2 > A \iff n > \sqrt{A}$.
Il suffit de choisir comme valeur de N le premier entier supérieur à $\sqrt{A}$.
Ainsi, pour tout entier $n \geqslant N$, on aura bien $n^2 > A$.

De la même façon, les suites (n), (n^3), etc., $(\sqrt{n})$ sont des suites qui ont pour limite $+\infty$.

3.b.2. Théorèmes sur les limites

Théorème 1.1 (convergence monotone) Toute suite croissante et majorée est convergente vers une limite l qui est le plus petit majorant de la suite.
De même, pour toute suite décroissante et minorée : sa limite est le plus grand des majorants.

Démonstration La démonstration repose sur l'axiome de la borne supérieure : *toute partie E non vide de $\mathbb{R}$ et majorée admet une borne supérieure, notée* $\sup E$ *qui en est le plus petit des majorants.*
Soit $(u_n)_{n\geqslant n_0}$ une suite définie sur $\mathbb{N}$, l'ensemble $E = \{u_n, n \geqslant n_0\}$ des valeurs prises par la suite (u) n'est pas vide car $u_{n_0} \in E$ et il est majoré car (u_n) l'est.
E admet donc une borne supérieure $\mu = \sup E$.
Cela signifie que les valeurs u_n dépassent n'importe quelle valeur inférieure strictement à μ.
En d'autres termes $\forall \epsilon \in \mathbb{R}_+^*,\ \exists N \in \mathbb{N},\ \mu - \epsilon < u_N \leqslant \mu$.
Et, par conséquent, comme (u_n) est croissante, $\forall n \in \mathbb{N},\ n \geqslant N \Longrightarrow \mu - \epsilon < u_N \leqslant u_n < \mu$, soit $n \geqslant N \Longrightarrow \mu - \epsilon < u_n \leqslant \mu$.
A fortiori, comme $\mu + \epsilon > \mu$, on a même $n \geqslant N \Longrightarrow \mu - \epsilon < u_n < \mu + \epsilon$, ce qui, d'après la définition 1.12, prouve que la suite (u_n) converge vers μ, soit $\lim\limits_{n\to+\infty} u_n = \mu$.

Si $(u_n)_{n\geqslant n_0}$ est croissante et majorée alors elle converge vers une limite l.
La suite $(v_n)_{n\geqslant n_0} = (-u_n)_{n\geqslant n_0}$ est décroissante, minorée et converge vers la limite $-l$.

Exemple 9 – Montrons que la suite u définie par $u_n = \sum\limits_{k=1}^{n} \frac{1}{k^2}$ est majorée par 2.
Pour commencer, montrons que $\frac{1}{n^2} < \frac{1}{n(n-1)}$ pour tout $n > 1$. En effet, $n^2 > n(n-1) = n^2 - n$ pour tout entier n, leurs inverses sont donc rangées dans le sens inverse.
La somme $u_n = \sum\limits_{k=1}^{n} \frac{1}{k^2}$ est donc majorée par $1 + \sum\limits_{k=2}^{n} \frac{1}{k(k-1)}$.
Or $\frac{1}{k(k-1)} = \frac{1}{k-1} - \frac{1}{k}$ (mettre au même dénominateur), u_n est donc majoré par $1 + \sum\limits_{k=2}^{n} (\frac{1}{k-1} - \frac{1}{k})$.
En écrivant tous les termes de cette somme, on s'aperçoit qu'ils s'éliminent deux à deux :
$1 + (\frac{1}{1} - \frac{1}{2}) + (\frac{1}{2} - \frac{1}{3}) + \cdots + (\frac{1}{n-1} - \frac{1}{n}) = 2 - \frac{1}{n}$. Donc $u_n < 2$.
De plus, la suite u est croissante, car $u_{n+1} - u_n = \frac{1}{n^2} > 0$.
Par application du théorème 1.1, la suite converge vers une limite l.
Remarque : si le théorème annonce qu'il y a une limite, il ne donne pas sa valeur. Ici, on sait seulement que la limite est $l \leqslant 2$. puisque 2 est un majorant de cette suite. On peut s'approcher de la véritable limite de cette suite au moyen d'un programme. On en obtiendra ainsi une valeur approchée qui est bien inférieure à 2 puisqu'elle vaut environ 1,644 934 066 848 226 436 472 415 166 6.
Comme l'a montré Euler[5] en 1735, il s'agit de l'irrationnel $\frac{\pi^2}{6}$.

Théorème 1.2 (somme et produit par un réel) Soient u une suite qui diverge vers $+\infty$ et k un réel non nul.

- La suite $(u_n + k)$ diverge vers $+\infty$ et, si $k > 0$, la suite (ku_n) aussi
- Si $k < 0$, la suite (ku_n) diverge vers $-\infty$

Démonstration Examinons le cas $k > 0$: comme la suite u diverge vers $+\infty$, elle dépasse toute valeur arbitrairement fixée. Un réel $A > 0$ étant arbitrairement fixé.
La suite dépasse le réel $\frac{A}{k}$, à partir d'un certain rang N, et donc $\forall n > N,\ u_n > \frac{A}{k}$, soit $ku_n > A$. La suite (ku_n) diverge donc vers $+\infty$.
La suite dépasse le réel $A - k$, à partir d'un certain rang N', et donc $\forall n > N',\ u_n > A - k$, soit $u_n + k > A$. La suite $(u_n + k)$ diverge donc vers $+\infty$.
On raisonne de même dans le cas $k < 0$.

5. Leonhard Euler(1707 Bâle - 1783 St Pétersbourg), un des plus grands et des plus prolifiques mathématiciens

Ce théorème est un cas particulier du théorème suivant qui envisage les différents cas de figure pour la somme et le produit de deux suites.

THÉORÈME 1.3 (SOMME ET PRODUIT DE SUITES) Soient u et v deux suites ayant une limite, finie ou infinie. Les suites de terme général $u_n + v_n$ et $u_n \times v_n$, admettent ou non une limite selon les différents cas exposés par les deux tableaux ci-dessous :

Suite « somme » : $(u_n + v_n)$

$\lim u$ / $\lim v$	l	$+\infty$	$-\infty$
l'	$l + l'$	$+\infty$	$-\infty$
$+\infty$	$+\infty$	$+\infty$	
$-\infty$	$-\infty$		$-\infty$

Suite « produit » : $(u_n \times v_n)$

$\lim u$ / $\lim v$	$l \neq 0$	$l = 0$	$\pm\infty$
$l' \neq 0$	$l \times l'$	0	$\pm\infty$
$l' = 0$	0	0	
$\pm\infty$	$\pm\infty$		$\pm\infty$

DÉMONSTRATION ➲ Montrons que $\lim\limits_{n\to+\infty} u_n = l$ et $\lim\limits_{n\to+\infty} v_n = l' \Longrightarrow \lim\limits_{n\to+\infty} u_n + v_n = l + l'$.
D'après la définition des limites on sait que :
$\lim\limits_{n\to+\infty} u_n = l \Longrightarrow \forall \epsilon \in \mathbb{R}_+^*,\ \exists N \in \mathbb{N},\ \forall n \in \mathbb{N},\ n \geqslant N \Longrightarrow |u_n - l| < \frac{\epsilon}{2}$
$\lim\limits_{n\to+\infty} v_n = l' \Longrightarrow \forall \epsilon \in \mathbb{R}_+^*,\ \exists N' \in \mathbb{N},\ \forall n \in \mathbb{N},\ n \geqslant N' \Longrightarrow |v_n - l'| < \frac{\epsilon}{2}$
On en déduit que $n \geqslant \max\{N, N'\} \Longrightarrow |u_n - l| + |v_n - l'| < \epsilon$.
Inégalité triangulaire : $|u_n - l| + |v_n - l'| > |u_n - l + v_n - l'| = |(u_n + v_n) - (l - l')|$.
Finalement, par transitivité de la relation d'ordre, $n \geqslant \max\{N, N'\} \Longrightarrow |(u_n + v_n) - (l - l')| < \epsilon$.
Ceci étant vrai $\forall \epsilon \in \mathbb{R}_+^*$, on a $\lim\limits_{n\to+\infty} u_n + v_n = l + l'$.

➲ Appliquons le théorème d'encadrement (v. plus loin) pour montrer que $\lim\limits_{n\to+\infty} u_n \times v_n = l \times l'$.
On majore l'expression $|u_n \times v_n - l \times l'|$ par une expression de limite nulle :
$u_n \times v_n - l \times l' = u_n \times v_n - u_n \times l' + u_n \times l' - l \times l' = u_n(v_n - l') + l'(u_n - l)$
Inégalité triangulaire : $|u_n \times v_n - l \times l'| < |u_n(v_n - l')| + |l'(u_n - l)|$
Comme la suite (u_n) est convergente, elle est majorée.
En effet $\lim\limits_{n\to+\infty} u_n = l \Longrightarrow \forall \epsilon \in \mathbb{R}_+^*, \exists N \in \mathbb{N}, \forall n \in \mathbb{N}, n \geqslant N \Longrightarrow |u_n - l| < \epsilon \Longrightarrow u_n < l + \epsilon$
En notant M un de ses majorants, on obtient $|u_n(v_n - l')| + |l'(u_n - l)| < |M(v_n - l')| + |l'(u_n - l)|$.

Les expressions $v_n - l'$ et $u_n - l$ ont pour limite 0. D'après la propriété 1.2, les expressions $M(v_n - l')$ et $l'(u_n - l)$ ont donc aussi pour limite 0, de même pour leur somme, d'après la propriété ci-dessus. Ainsi $|u_n \times v_n - l \times l'|$ est encadrée entre 0 et une expression ayant pour limite 0 ; sa limite est donc 0.

Remarques :

- Les cases hachurées indiquent des cas indéfinis. Dans ces cas, on ne peut pas conclure directement à l'aide de ce théorème. Par exemple, si on a $u_n = n^2$ et $v_n = -2n$, on a $\lim u = +\infty$ et $\lim v = -\infty$. La somme $u + v$ a une limite indéterminée à l'égard de ce théorème. Il suffit, dans ce cas, de factoriser : $u_n + v_n = n(n - 2)$. La somme indéterminée est devenue déterminée puisque les suites r et s définies par $r_n = n$ et $s_n = n - 2$ forment une suite « produit » déterminée : chacune divergeant vers $+\infty$, la limite est trouvée à l'aide du tableau de droite, c'est $+\infty$. On verra dans l'exemple qui suit que les suites « polynomiale » admettent pour limite la limite de son « terme dominant ».
- Deux cas d'indétermination sont présents ici : $\infty - \infty$ et $\infty \times 0$.
  On peut lever la 1re par une factorisation : $n^2 - 3n = n(n - 3)$ qui n'est plus indéterminé.
  Pour la 2e, il peut suffire de simplifier : $n^2 \times \frac{2}{n} = 2n$ qui n'est plus indéterminé.
  Mais la plupart du temps, on aura recours aux théorèmes de comparaison pour lever les indéterminations.
- La notation $\pm\infty$ indique que la limite est $+\infty$ ou $-\infty$. La règle des signes étendue aux signes des limites, s'applique. Par exemple $\frac{n^3}{100}(-2 + \frac{1}{n})$ admet $-\infty$ comme limite, du fait que $\lim \frac{n^3}{100} = +\infty$ (positive) et $\lim(-2 + \frac{1}{n}) = -2$ (négative).

Exemple 10 – Quelle est la limite de la suite u définie par $u_n = -n^3 + 3n^2 - 2n + 1$?
Le théorème 1.3 ne permet pas de conclure directement, le terme général étant écrit sous cette forme développée. Il faut mettre la plus grande puissance de l'indice en facteur et utiliser le volet « produit » de ce théorème. Ici, on écrira $u_n = n^3(-1 + \frac{3}{n} + \frac{-2}{n^2} + \frac{3}{n^3})$.
Le 1er facteur tend vers $+\infty$ tandis que le 2d tend vers -1, du fait de la limite de chacun des termes de la somme et du théorème sur la somme : $\lim -1 = -1$, $\lim \frac{3}{n} = 0$, $\lim \frac{-2}{n^2} = 0$ et $\lim \frac{3}{n^3} = 0$.

Méthode (Suite polynomiale)
Lorsque le terme général a une forme polynomiale de degré p : $u_n = \sum_{k=0}^{p} a_k n^k$, le terme dominant au voisinage de $+\infty$ est toujours celui de plus haut degré, donc ici $a_p n^p$. Les autres termes sont vite négligeables devant lui lorsque n devient très grand. La limite de u sera celle du terme dominant :

$$a_p \neq 0 \Longrightarrow \lim \sum_{k=0}^{p} a_k n^k = \lim a_p n^p$$

L'expression « forme polynomiale » s'étend même ici aux puissances négatives de l'indice. La suite v définie par $v_n = -n^3 + 3n^2 - 2n + 1 - \frac{2}{n} + \frac{7}{n^2}$ aura le même comportement asymptotique que la suite u, au voisinage de $+\infty$, la différence $v_n - u_n$ ayant pour limite 0.

Théorème 1.4 (suite quotient) Soient u et v deux suites ayant une limite, finie ou infinie. La suite de terme général $\frac{u_n}{v_n}$ admet ou non une limite selon les différents cas exposés par le tableau ci-dessous :

$\lim v$ \ $\lim u$	$l \neq 0$	$l = 0$	$\pm\infty$
$l' \neq 0$	$\frac{l}{l'}$	0	$\pm\infty$
$l' = 0$	$\pm\infty$		$\pm\infty$
$\pm\infty$	0	0	

Démonstration Montrons simplement, comme précédemment, la 1re case de ce tableau.
Commençons pour cela à montrer $\lim\limits_{n \to +\infty} u_n = l \neq 0 \Longrightarrow \lim\limits_{n \to +\infty} \frac{1}{u_n} = \frac{1}{l}$
Il suffira ensuite d'appliquer le résultat sur la limite d'un produit à $\frac{u_n}{v_n} = u_n \times \frac{1}{v_n}$.
Comme la suite (u_n) est convergente, $(|u_n|)$ converge vers $|l|$ et est minorée par un réel $m > 0$:
On a donc $\forall \epsilon \in \mathbb{R}_+^*, \exists N \in \mathbb{N}, \forall n \in \mathbb{N}, n \geqslant N \Longrightarrow ||u_n| - |l|| < \epsilon \Longrightarrow |l| - \epsilon < |u_n| < |l| + \epsilon$
En prenant $\epsilon = \frac{|l|}{2}$, on obtient $|l| - \frac{|l|}{2} < |u_n| \Longleftrightarrow |u_n| > \frac{|l|}{2}$.
Ainsi, en notant $m = \frac{|l|}{2}$, à partir d'un rang N, $|u_n| > m$.

Comme $l \neq 0$, on en déduit que $m > 0$ et donc que $|u_n| \neq 0$ pour tout $n > N$.
$|u_n| > m > 0 \Longrightarrow \frac{1}{|u_n|} < \frac{1}{m}$, d'où $|\frac{1}{u_n} - \frac{1}{l}| = |\frac{l - u_n}{u_n \times l}| = \frac{|u_n - l|}{|u_n| \times |l|} = \frac{1}{|u_n| \times |l|} \times |u_n - l| < \frac{1}{m \times |l|} \times |u_n - l|$
Or $|u_n - l|$ tendant vers 0, il en est de même $\frac{1}{m \times |l|} \times |u_n - l|$.
D'après le théorème d'encadrement, la limite de $|\frac{1}{u_n} - \frac{1}{l}|$ est donc nulle ; $\frac{1}{u_n}$ a pour limite $\frac{1}{l}$.

Remarque :
Les cases hachurées indiquent ici deux autres cas d'indétermination : $\frac{\infty}{\infty}$ et $\frac{0}{0}$.
Dans le cas des suites rationnelles (quotients de deux expressions polynomiales), on ne conserve que les termes de plus haut degré du numérateur et du dénominateur, on simplifie l'expression obtenue et on lui applique le théorème sur le quotient.
Pour la suite u de terme général $u_n = \frac{-n^3 + 3n^2 - 2n + 1}{5n^4 + 12n - 7}$, on commence par dire que u a la même limite que la suite $(\frac{-n^3}{5n^4})$. L'expression du terme général se simplifiant en $\frac{-1}{5n}$, on applique le théorème sur le quotient : la limite est 0 (le numérateur est constant et le dénominateur tend vers $+\infty$).

Exemple 11 – Quelle est la limite de la suite u définie par $u_n = \frac{2^n-5}{3^n+1}$?
Le théorème 1.4 ne permet pas de conclure directement. En effet, les expressions au numérateur et au dénominateur tendent toutes les deux vers $+\infty$ (voir plus loin, les suites géométriques).
Mais si on divise le numérateur et le dénominateur par $3^n > 0$, on obtient $u_n = \frac{(\frac{2}{3})^n - \frac{5}{3^n}}{1+\frac{1}{3^n}}$.
Le numérateur tend maintenant vers 0 car les deux termes qui le composent tendent vers 0 ; le dénominateur, quant-à lui, tend vers 1.
Finalement, la suite u tend vers 0.

Théorème 1.5 (Majoration/Minoration) Soient u et v deux suites.
Si, à partir d'un certain rang N, on a $u_n \geqslant v_n$ et si $\lim\limits_{n\to+\infty} v_n = +\infty$, alors $\lim\limits_{n\to+\infty} u_n = +\infty$.
Si, à partir d'un certain rang N, on a $u_n \leqslant v_n$ et si $\lim\limits_{n\to+\infty} v_n = -\infty$, alors $\lim\limits_{n\to+\infty} u_n = -\infty$.

Démonstration Montrons seulement la 1re ligne de ce théorème.
Comme $\lim\limits_{n\to+\infty} v_n = +\infty$, on a $\forall A \in \mathbb{R},\ \ \exists N' \in \mathbb{N},\ \ \forall n \in \mathbb{N},\ \ n \geqslant N' \Longrightarrow v_n > A$,
et comme $\exists N \in \mathbb{N}, \forall n \in \mathbb{N}, n \geqslant N \Longrightarrow u_n \geqslant v_n$, on a $\forall A \in \mathbb{R}, \forall n \in \mathbb{N}, n \geqslant \max\{N, N'\} \Longrightarrow u_n > A$.
Ceci prouve bien que la suite u diverge vers $+\infty$ ($\lim\limits_{n\to+\infty} u_n = +\infty$).

Exemple 12 – Montrons que la suite u définie par $u_n = 2^n$ diverge vers $+\infty$.
Le terme général 2^n est supérieur à n pour toutes les valeurs de n car, en factorisant, on a :
$2^n - 1 = (2-1)(2^{n-1} + 2^{n-2} + \cdots + 2^1 + 2^0)$, et comme $\forall k \geqslant 0,\ \ 2^k \geqslant 1$, on a
$2^n = 1 + (2^{n-1} + 2^{n-2} + \cdots + 2^1 + 1) \geqslant n + 1$.
Comme la suite v définie par $v_n = n + 1$ diverge vers $+\infty$ et que $\forall n \geqslant 1,\ \ v_n < u_n$, d'après le théorème 1.5, la suite u diverge vers $+\infty$.

Méthode
➲ Limites des suites géométriques
Ce que nous venons de faire avec la suite (2^n) se généralise à toutes les suites (a^n) où $a \geqslant 1$.

L'inégalité préliminaire que l'on obtient s'appelle « inégalité de Bernoulli » :
$a^n - 1 = (a-1)(a^{n-1} + a^{n-2} + \cdots + a^1 + a^0)$, et comme $\forall k \geqslant 0,\ \ a \geqslant 1 \iff a^k \geqslant 1$, on a $\forall n \geqslant 0,\ \ a^n = 1 + (a-1)(a^{n-1} + a^{n-2} + \cdots + a^1 + a) \geqslant n(a-1) + 1$.

- Pour $a > 1$, d'après le théorème 1.5, la suite (a^n) diverge vers $+\infty$ car $a^n \geqslant n(a-1)+1$ et que $\forall a > 1,\ \ \lim n(a-1) + 1 = +\infty$.
- Pour $a = 1$, la suite (1^n) est constante. Elle converge vers 1.
- Pour $0 < a < 1$, la suite (a^n) converge vers 0. En effet, $a^n = \frac{1}{(\frac{1}{a})^n}$ et, comme $\frac{1}{a} > 1$, la suite de terme général $(\frac{1}{a})^n$ tend vers $+\infty$. Par application du théorème 1.4, la suite (a^n) étant obtenu par quotient d'une constante 1 par une expression tendant vers $+\infty$ tend vers 0.
- Pour $-1 < a < 0$, les termes de la suite sont alternativement positifs et négatifs, mais globalement la suite (a^n) tend vers 0.
- Pour $a < -1$, les termes de la suite sont alternativement positifs et négatifs, leur valeur absolue tend vers $+\infty$, mais la suite (a^n) diverge sans admettre de limite.

Comportement asymptotique d'une suite géométrique u de raison a et de premier terme u_0 :

a	$]-\infty;-1]$	$]-1;1[$	1	$]1;+\infty[$
$u_0 > 0$	pas de limite	0	u_0	$+\infty$
$u_0 < 0$	pas de limite	0	u_0	$-\infty$

Pour ce qui est de la somme S des termes d'une suite géométrique :
- Si $a > 1$, S diverge vers $\pm\infty$
- Si $|a| < 1$, la formule $S = u_0 \times \frac{1-a^{n+1}}{1-a}$ montre que, $1 - a^{n+1}$ tendant vers 1, $\lim S = \frac{1}{1-a}$.

Théorème 1.6 (gendarmes) Soient u, v et w trois suites. Si, à partir d'un certain rang N, on a $v_n \leqslant u_n \leqslant w_n$ et si $\lim\limits_{n\to+\infty} v_n = \lim\limits_{n\to+\infty} w_n = l$, alors $\lim\limits_{n\to+\infty} u_n = l$.

Démonstration Comme $\lim\limits_{n\to+\infty} v_n = \lim\limits_{n\to+\infty} w_n = l$, on a :

- $\forall \epsilon \in \mathbb{R}_+^*, \exists N' \in \mathbb{N}, \forall n \in \mathbb{N}, n \geqslant N' \Longrightarrow l - \epsilon < v_n < l + \epsilon$
- $\forall \epsilon \in \mathbb{R}_+^*, \exists N'' \in \mathbb{N}, \forall n \in \mathbb{N}, n \geqslant N'' \Longrightarrow l - \epsilon < w_n < l + \epsilon$

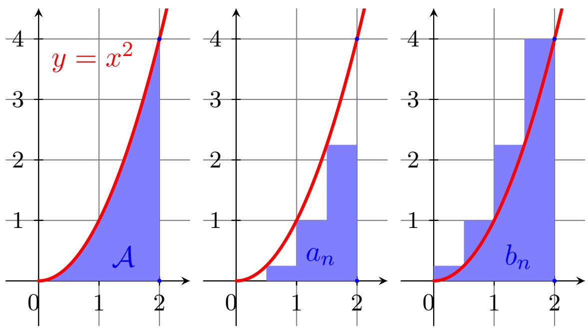
On en déduit que, pour tout entier $n \geqslant \max\{N, N', N''\}$ on a $l - \epsilon < v_n \leqslant u_n \leqslant w_n < l + \epsilon$, soit $l - \epsilon < u_n < l + \epsilon$, ce qui prouve bien que la suite u converge vers l ($\lim\limits_{n\to+\infty} u_n = l$).

Exemple 13 – ➲ Montrons que la suite u définie par $u_n = \frac{\sin n}{n}$ a pour limite 0.
Nous savons que $\forall n \geqslant 0$ on a $-1 \leqslant \sin n \leqslant 1$.
On en déduit l'encadrement, valable pour tout entier $n \geqslant 1$: $\frac{-1}{n} \leqslant \frac{\sin n}{n} \leqslant \frac{1}{n}$.
Or, $\lim\limits_{n\to+\infty} \frac{-1}{n} = \lim\limits_{n\to+\infty} \frac{-1}{n} = 0$. D'après le théorème 1.6, la suite u converge donc vers 0.

➲ Calcul d'aire
Calculons l'aire $\mathcal{A}$ du domaine limité par

- la parabole d'équation $y = x^2$
- les droites d'équation $x = 0$ et $x = 2$
- l'axe des abscisses $y = 0$

Pour cela, encadrons cette aire :
$\forall n \geqslant 0,\ \ a_n < \mathcal{A} < b_n$
les aires a_n et b_n étant faciles à calculer car constituées de rectangles
(sur notre illustration $n = 4$)



- L'aire a_n est obtenue en additionnant les aires de n rectangles de largeur $\frac{2}{n}$; les hauteurs de ces rectangles constituent la suite h : $h_0 = 0$, $h_1 = (\frac{1\times 2}{n})^2$, $h_2 = (\frac{2\times 2}{n})^2$, $\cdots$, $h_{n-1} = (\frac{(n-1)\times 2}{n})^2$
- L'aire b_n est obtenue en additionnant les aires de n rectangles de largeur $\frac{2}{n}$; les hauteurs de ces rectangles constituent la suite H : $H_0 = (\frac{1\times 2}{n})^2$, $H_1 = (\frac{2\times 2}{n})^2$, $H_2 = (\frac{3\times 2}{n})^2$,$\cdots$,$H_{n-1} = (\frac{n\times 2}{n})^2$

Ainsi on a $a_n = h_0 \times \frac{2}{n} + h_1 \times \frac{2}{n} + \cdots + h_{n-1} \times \frac{2}{n} = \frac{2}{n}[(0 + (\frac{1\times 2}{n})^2 + \cdots + (\frac{(n-1)\times 2}{n})^2]$
On obtient $a_n = (\frac{2}{n})^3[(1)^2 + (2)^2 + \cdots + (n-1)^2]$ et $b_n = (\frac{2}{n})^3[(1)^2 + (2)^2 + \cdots + (n-1)^2 + n^2]$.

L'identité $\sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}$ prouvée dans le cours de 1re (Ex.1.10 et prop.3.5) permet de simplifier $a_n = (\frac{2}{n})^3(\frac{(n+1)(n)[2(n-1)+1]}{6}) = (\frac{2}{n})^3(\frac{n(n+1)(2n-1)}{6})$ et $b_n = (\frac{2}{n})^3(\frac{n(n+1)(2n+1)}{6})$.
Déterminons alors les limites de a et b :

- $\lim a_n = \lim \frac{2^3\times n(n+1)(2n-1)}{6n^3} = \lim \frac{2^3\times 2n^3}{6n^3} = \frac{2^4}{6} = \frac{8}{3}$
- $\lim b_n = \lim \frac{2^3\times n(n+1)(2n+1)}{6n^3} = \lim \frac{2^3\times 2n^3}{6n^3} = \frac{2^4}{6} = \frac{8}{3}$

Ces deux suites encadrent $\mathcal{A}$ et convergent vers $\frac{8}{3}$; par application du théorème 1.6, $\mathcal{A} = \frac{8}{3}$.

Théorème 1.7 (Passage à la limite) Soient u et v deux suites qui convergent respectivement vers l et l'. Si, à partir d'un certain rang N, on a $u_n > v_n$, alors $l \geqslant l'$.

Démonstration Montrons d'abord par l'absurde, la propriété :
Soit u une suite qui converge vers l. Si, à partir d'un rang N, on a $u_n > 0$, alors $l \geqslant 0$.
Supposons $l < 0$. Comme $\lim\limits_{n\to+\infty} u_n = l \Longrightarrow \forall \epsilon \in \mathbb{R}_+^*, \exists N' \in \mathbb{N}, \forall n \in \mathbb{N}, n \geqslant N' \Longrightarrow l - \epsilon < u_n < l + \epsilon$, en prenant $\epsilon = -l > 0$, on obtient $2l < u_n < 0$, ce qui contredit l'hypothèse. Par conséquent $l \geqslant 0$.

En prenant $w_n = u_n - v_n$, à partir d'un certain rang N, on a $w_n > 0$ et comme, d'après la propriété 1.3, w est une suite qui converge vers $l - l'$, on lui applique le résultat précédent :
$l - l' \geqslant 0 \Longleftrightarrow l \geqslant l'$.

3.b.3. Suites adjacentes

Théorème 1.8 (Suites adjacentes) Soient (u_n) et (v_n) deux suites définies sur $\mathbb{N}$.
On dit que (u_n) et (v_n) sont adjacentes si (u_n) est croissante, (v_n) décroissante et $\lim\limits_{n\to+\infty} u_n - v_n = 0$

Remarques :

- Pour illustrer le théorème 1.6 des gendarmes, on a vu que le terme général de la suite $(\frac{\sin n}{n})$ était encadré par les termes des deux suites s $(\frac{-1}{n})$ et $(\frac{1}{n})$. La 1re est croissante tandis que la 2e est décroissante et la différence $\frac{1}{n} - \frac{-1}{n} = \frac{2}{n}$ tend vers zéro. Ces suites sont donc adjacentes. Ces deux suites tendent vers une même limite qui est, ici, nulle.
- Dans le même exemple 13, lorsqu'on a encadré l'aire du domaine situé entre la courbe de la fonction $x \longmapsto x^2$ et l'axe des abscisses pour $0 \leqslant x \leqslant 2$, on a construit les suites (a_n) et (b_n) qui convergeaient vers la même limite ; ces suites étaient adjacentes par construction car l'une minorait l'aire cherchée tandis que l'autre la majorait et, comme elles convergeait vers la même limite, leur différence tendait vers zéro.
- D'autres suites sont adjacentes de façon encore plus triviale : les suites p_n et q_n des approximations par défaut et par excès à 10^{-n} près d'un nombre réel quelconque.

Propriété 1.10 (Ordre) Soient (u_n) et (v_n) deux suites adjacentes définies sur $\mathbb{N}$.
Si (u) est croissante et (v) décroissante alors, pour tout $n \in \mathbb{N}$, on a $u_n \leqslant v_n$.

Démonstration Supposons qu'il existe un rang m pour lequel $u_m > v_m \Longleftrightarrow u_m - v_m > 0$.
Comme u est croissante, $\forall n \in \mathbb{N}, n > m \Longrightarrow u_n \geqslant u_m$
Comme v est décroissante, $\forall n \in \mathbb{N}, n > m \Longrightarrow v_n \leqslant v_m \Longleftrightarrow -v_n \geqslant -v_m$
Par addition des deux inégalités, il vient $u_n - v_n \geqslant u_m - v_m$ pour $n > m$.
Or, u et v étant adjacentes, on a $\lim\limits_{n\to+\infty} u_n - v_n = 0 \Longleftrightarrow \forall \epsilon \in \mathbb{R}_+^*, \exists N \in \mathbb{N}, -\epsilon < u_n - v_n < \epsilon$.
En prenant $\epsilon = u_m - v_m$, on a alors $u_n - v_n < u_m - v_m$.
Pour $n > \max\{m, N\}$, on a donc simultanément $u_n - v_n < u_m - v_m$ et $u_n - v_n \geqslant u_m - v_m$.
Cette contradiction prouve que l'hypothèse $\exists m \in \mathbb{N}, u_m > v_m$ est à rejeter.

Théorème 1.9 (Convergence) Soient (u_n) et (v_n) deux suites adjacentes définies sur $\mathbb{N}$.
Si (u) est croissante et (v) décroissante alors les deux suites convergent vers un même réel λ.
De plus, $\forall n \in \mathbb{N},\ \ u_n \leqslant \lambda \leqslant v_n$.

Démonstration Par définition, la suite (v_n) est décroissante, donc $\forall n \in \mathbb{N}, v_n \leqslant v_0$.
D'après la propriété 1.10, on a $\forall n \in \mathbb{N}, u_n \leqslant v_n$ donc $u_n \leqslant v_0$.
La suite (u_n) est donc croissante et majorée. D'après la propriété 1.1 elle converge vers un réel l.
De même, (v_n) étant décroissante et minorée (par u_0), converge vers un réel l'.
On sait que $\lim\limits_{n\to+\infty} u_n - v_n = 0$ or $\lim\limits_{n\to+\infty} u_n - v_n = \lim\limits_{n\to+\infty} u_n - \lim\limits_{n\to+\infty} v_n = l - l'$.
On en déduit que $l - l' = 0 \Longleftrightarrow l = l'$. Les deux suites convergent vers la même limite.

Exemple 14 – Soit $(u_n)_{n\geqslant 1}$ la suite de terme général $u_n = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots + \frac{(-1)^{n+1}}{n}$.
et soient (a_n) et (b_n) les suites définies par $a_n = u_{2n}$ et $b_n = u_{2n+1}$.

$$\begin{aligned} a_{n+1} - a_n =& u_{2(n+1)} - u_{2n} = \left[1 - \frac{1}{2} + \frac{1}{3} + \cdots + \frac{(-1)^{2n+3}}{2n+2}\right] - \left[1 - \frac{1}{2} + \frac{1}{3} + \cdots + \frac{(-1)^{2n+1}}{2n}\right] \\ =& \frac{(-1)^{2n+2}}{2n+1} + \frac{(-1)^{2n+3}}{2n+2} = \frac{1}{2n+1} - \frac{1}{2n+2} = \frac{2n+2-(2n+1)}{(2n+2)(2n+1)} = \frac{1}{(2n+2)(2n+1)} \end{aligned}$$

Comme $\forall n \in \mathbb{N}$, $\frac{1}{(2n+2)(2n+1)} > 0$, on en déduit que (a_n) est croissante.

Prouvons par récurrence que $a_n \leqslant 1 - \frac{1}{2n}$ pour tout $n \geqslant 1$.
$a_1 = u_2 = 1 - \frac{1}{2} = \frac{1}{2}$. La propriété est vraie pour $n = 1$.
Supposons que pour un rang $n \geqslant 1$ on ait $a_n \leqslant 1 - \frac{1}{2n}$.
Comme, d'après le calcul ci-dessus $a_{n+1} = a_n + \frac{1}{(2n+2)(2n+1)}$.
On a donc $a_{n+1} \leqslant 1 - \frac{1}{2n} + \frac{1}{(2n+2)(2n+1)} = 1 - \frac{(2n+2)(2n+1)-2n}{2n(2n+2)(2n+1)} = 1 - \frac{2n^2+2n+1}{2n(n+1)(2n+1)}$.
A-t-on $1 - \frac{2n^2+2n+1}{2n(n+1)(2n+1)} \leqslant 1 - \frac{1}{2(n+1)}$, soit $\frac{2n^2+2n+1}{2n(n+1)(2n+1)} \geqslant \frac{1}{2(n+1)}$ ou encore
$2(n+1)(2n^2+2n+1) \geqslant 2n(n+1)(2n+1) \iff 2n^2+2n+1 \geqslant n(2n+1) \iff n \geqslant -1$?
Oui, d'où la conclusion $a_n \leqslant 1 - \frac{1}{2n}$ pour tout $n \geqslant 1$.
La suite (a_n) est donc croissante et majorée par 1.
D'après le théorème 1.1 de convergence monotone et la propriété 1.7 de passage aux limites, la suite (a_n) converge vers un réel $\lambda \leqslant 1$.

$b_n = u_{2n+1} = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots + \frac{(-1)^{2n+1}}{2n} + \frac{(-1)^{2n+2}}{2n+1} = a_n + \frac{(-1)^{2n+2}}{2n+1} = a_n + \frac{1}{2n+1}$.
Comme (a_n) converge vers λ et $(\frac{1}{2n+1})$ converge vers 0 par addition des limites (b_n) converge vers λ.

$$\begin{aligned}
b_{n+1} - b_n &= \left[a_{n+1} + \frac{1}{2n+3}\right] - \left[a_n + \frac{1}{2n+1}\right] = a_{n+1} - a_n + \frac{(2n+1)-(2n+3)}{(2n+1)(2n+3)} \\
&= \frac{1}{(2n+2)(2n+1)} + \frac{-2}{(2n+1)(2n+3)} = \frac{(2n+3)-2(2n+2)}{(2n+3)(2n+2)(2n+1)} \\
&= \frac{-(2n+1)}{(2n+3)(2n+2)(2n+1)} = \frac{-1}{2(n+1)(2n+3)}
\end{aligned}$$

Comme $\forall n \in \mathbb{N}$, $\frac{-1}{2(n+1)(2n+3)} < 0$, on en déduit que (b_n) est décroissante.

Les suites (a_n) et (b_n) sont adjacentes et convergent vers une même limite $\lambda \leqslant 1$ que cette étude ne permet pas de déterminer algébriquement. On peut, tout au plus en donner des valeurs approchées puisque, d'après le théorème 1.9, on a l'encadrement $a_n \leqslant \lambda \leqslant b_n$.
Voici les premières valeurs obtenues pour cet encadrement :

n	1/n	u_n	a_n	b_n
1	1	1	0,5	0,8333333333
2	0,5	0,5	0,5833333333	0,7833333333
3	0,3333333333	0,8333333333	0,6166666667	0,7595238095
4	0,25	0,5833333333	0,6345238095	0,7456349206
5	0,2	0,7833333333	0,6456349206	0,7365440115
6	0,1666666667	0,6166666667	0,6532106782	0,7301337551
7	0,1428571429	0,7595238095	0,6587051837	0,7253718504
8	0,125	0,6345238095	0,6628718504	0,7216953798
9	0,1111111111	0,7456349206	0,6661398242	0,7187714032
10	0,1	0,6456349206	0,6687714032	0,7163904508
11	0,0909090909	0,7365440115	0,6709359053	0,7144141662
12	0,0833333333	0,6532106782	0,6727474995	0,7127474995

Ainsi, pour $n = 12$ par exemple on obtient $a_{12} \leqslant \lambda \leqslant b_{12}$, soit environ $0,6727 \leqslant \lambda \leqslant 0,7127$.
On constate que l'amplitude de l'encadrement ne diminue pas rapidement : au rang n il vaut $\frac{1}{2n+1}$.
Pour $n = 12$, sur notre illustration, il vaut $\frac{1}{25} = 0,4$ et il faut atteindre $n = 50$ pour que cette amplitude soit inférieure à $0,01$.
La valeur exacte de la limite λ est le nombre transcendant $\ln(2) \approx 0,6931471806$.

2 Fonctions (1)

Plan du chapitre :

Aperçu historique : La notion de limite s'est précisée lentement, au cours des XVIIe et XVIIIe siècles, pour trouver une définition précise au début du XIXe avec les mathématiciens Augustin-Louis Cauchy (1789-1857), Bernard Bolzano (1781-1848) et surtout Karl Weierstrass (1815-1897), « le père de l'analyse moderne », qui définit également la notion de continuité.

Jusqu'à Bolzano et Weierstrass, la notion de continuité était intuitive : on considérait des courbes tracées sans lever le crayon, et le théorème des valeurs intermédiaires était perçu comme une évidence géométrique. Pourtant ces notions sont fondamentales et il convient de les aborder avec rigueur. L'étude des limites en un point est l'outil qui permit cette conceptualisation. Si une fonction est continue sur un intervalle, sa monotonie en fait une bijection, ce qui a pour conséquence qu'on peut en définir la réciproque ; cette propriété a notamment permis de définir la fonction ln (les initiales de *logarithme népérien*) comme réciproque de la fonction continue et strictement croissante exp.

La notion de dérivée a aussi émergé de façon intuitive avec les travaux sur l'*infiniment petit* de Newton et Leibniz au XVIIe siècle. Elle s'est ensuite précisée avec Jean Le Rond d'Alembert (1717-1783) qui fait le lien avec le taux d'accroissement, et c'est encore Weierstrass qui finalise les définitions dans des termes encore utilisés aujourd'hui.

L'autre notion fondamentale de l'analyse, l'intégration – notion que nous étudierons au chapitre suivant – a des racines anciennes puisque, notamment, les travaux d'Archimède sur le calcul de l'aire d'un disque la préfigure. Le calcul infinitésimal de Newton et Leibniz va relier le calcul différentiel et le calcul intégral mais c'est Bernhard Riemann (1826-1856) qui va, au XIXe siècle, exposer la théorie de l'intégrale utilisée aujourd'hui.
Les fonctions logarithmes furent inventées au début du XVIIe siècle par le mathématicien astronome écossais John Napier (1550-1617). Logarithmes vient du grec *logos*(relation) et *arithmeticos*(nombre) Napier publie avec Henry Gibbs (1561-1631) des tables de correspondances entre deux séries de valeurs possédant la propriété de faire correspondre un produit à une somme, dans le but de simplifier les calculs astronomiques. Les tables trigonométriques existaient depuis plusieurs siècles, mais la relation fondamentale $\log(a \times b) = \log(a) + \log(b)$ permet de remplacer une multiplication par une addition. Ces tables furent utilisées jusqu'au début de l'ère informatique.
Les fonctions exponentielles et logarithmes deviennent rapidement des incontournables des problèmes liés au « nouveau calcul »

1. Limite d'une fonction

1.a. Au voisinage de l'infini

Définition 2.1 ($+\infty$) Une fonction f, définie sur un intervalle de la forme $]a;+\infty[$ a pour limite $+\infty$ en $+\infty$, ce que l'on note $\lim_{x\to+\infty} f(x) = +\infty$ si et seulement si, quelque soit le nombre A, à partir d'une certaine valeur $x_0 > a$ on a $x > x_0 \Longrightarrow f(x) > A$.
Autrement dit : $\lim_{x\to+\infty} f(x) = +\infty \Longleftrightarrow \forall A \in \mathbb{R},\ \exists x_0 \in \mathbb{R},\ \forall x > x_0, f(x) > A$

Remarques : On a une définition analogue pour les fonctions de limite $-\infty$ en $+\infty$:
$\lim_{x\to+\infty} f(x) = -\infty \Longleftrightarrow \forall A \in \mathbb{R},\ \exists x_0 \in \mathbb{R},\ \forall x > x_0, f(x) < A$
De même, on définit ce que signifie, pour une fonction numérique, avoir une limite $\pm\infty$ en $-\infty$:
$\lim_{x\to-\infty} f(x) = +\infty \Longleftrightarrow \forall A \in \mathbb{R},\ \exists x_0 \in \mathbb{R},\ \forall x < x_0, f(x) > A$
$\lim_{x\to-\infty} f(x) = -\infty \Longleftrightarrow \forall A \in \mathbb{R},\ \exists x_0 \in \mathbb{R},\ \forall x < x_0, f(x) < A$
On dit « f a pour limite $+\infty$ en $+\infty$ » ou bien « f tend vers $+\infty$ quand x tend vers $+\infty$ ».

Propriété 2.1 Chacune des fonctions suivantes a pour limite $+\infty$ en $+\infty$:
$x \mapsto x$, $x \mapsto x^2$, $x \mapsto x^3$, etc. (soit $x \mapsto x^n$ avec $n \in \mathbb{N}^*$) et $x \mapsto \sqrt{x}$

Démonstration ➟ Montrons cela pour la fonction $x \mapsto x^2$.
Pour tout nombre $A > 0$ on peut trouver x_0 à partir duquel tous les carrés sont plus grands que A : il s'agit de $x_0 = \sqrt{A}$ car, pour $x > 0$, $x^2 > A \Longleftrightarrow x > \sqrt{A}$.

➟ Pour les autres valeurs de l'exposant $n \geqslant 1$: comme $x^n > x^{n-1}$ pour tout $x \geqslant 1$ (en multipliant par $x^{n-1} > 0$), $x > \sqrt{A} \Longrightarrow x^n > x^2 > A$ et donc, x^n dépasse n'importe quelle valeur A arbitrairement choisie pourvu que x soit assez grand.

➟ Pour la fonction racine, la valeur x_0 à partir de laquelle $\sqrt{x}$ dépasse n'importe quelle valeur A arbitrairement choisie, est $x_0 = A^2$.
Les fonctions puissances et la fonction racine tendent donc toutes vers $+\infty$ quand x tend vers $+\infty$.

Définition 2.2 (Fonction de limite finie) Une fonction f à valeurs positives sur $[a;+\infty[$, admet 0 comme limite en $+\infty$ si et seulement si, quel que soit le nombre $\epsilon > 0$, à partir d'une certaine valeur x_0 de x, on a $f(x) < \epsilon$.
Autrement dit : $\lim_{x\to+\infty} f(x) = 0 \Longleftrightarrow \forall\epsilon > 0,\ \exists x_0 > a,\ \forall x > x_0, f(x) < \epsilon$
l étant un réel, une fonction f admet l comme limite en $+\infty$ si la fonction $f - l$ admet pour limite 0 en $+\infty$. Autrement dit, s'il existe une fonction φ de limite 0 en $+\infty$, telle que $f(x) = l + \varphi(x)$.

Remarques :

- De la première partie, on tire une définition de la limite 0 en $+\infty$ pour une fonction de signe quelconque : $\lim_{x\to+\infty} f(x) = 0 \Longleftrightarrow \lim_{x\to+\infty} |f(x)| = 0$. On peut exprimer cela explicitement :
  $\lim_{x\to+\infty} f(x) = 0 \Longleftrightarrow \forall\epsilon > 0,\ \exists x_0 > a,\ \forall x > x_0, f(x) \in]-\epsilon, +\epsilon[$.
  Pour une limite finie quelconque :
  $\lim_{x\to+\infty} f(x) = l \Longleftrightarrow \forall\epsilon > 0,\ \exists x_0 > a,\ \forall x > x_0, f(x) \in]l - \epsilon, l + \epsilon[$
  ce qu'on peut noter aussi $|f(x) - l| < \epsilon$.
- Notons que toutes les fonctions n'ont pas forcément une limite en $+\infty$:
  $x \mapsto \sin(x)$ par exemple n'a pas de limite.

PROPRIÉTÉ 2.2 Chacune des fonctions suivantes a pour limite 0 en $+\infty$:
$x \mapsto \frac{1}{x}$, $x \mapsto \frac{1}{x^2}$, $x \mapsto \frac{1}{x^3}$, etc. (soit $x \mapsto \frac{1}{x^n}$ avec $n \in \mathbb{N}^*$) et $x \mapsto \frac{1}{\sqrt{x}}$

DÉMONSTRATION ➠ Montrons cela pour la fonction $x \mapsto \frac{1}{x}$.
Pour tout nombre $\epsilon > 0$ on peut trouver x_0 à partir duquel toutes les inverses sont plus petites que ϵ : il s'agit de $x_0 = \frac{1}{\epsilon}$ car, pour $x > 0$, $\frac{1}{x} < \epsilon \iff x > \frac{1}{\epsilon}$.

➠ Pour les autres valeurs de l'exposant $n > 0$, il suffit de remarquer encore que, comme $x^n > x^{n-1}$ pour tout $x \geqslant 1$, on déduit $\frac{1}{x^n} < \frac{1}{x^{n-1}} < \cdots < \frac{1}{x^2} < \frac{1}{x}$ et donc $\forall n > 0$, $\lim\limits_{x \to +\infty} \frac{1}{x^n} = 0$.

➠ Pour la fonction racine, la valeur x_0 à partir de laquelle $\frac{1}{\sqrt{x}}$ devient inférieur à n'importe quelle valeur ϵ arbitrairement choisie, est $x_0 = \frac{1}{\epsilon^2}$.

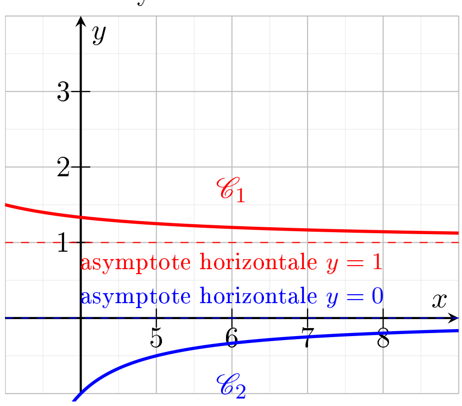
Limite en $+\infty$:

- D'une façon générale tout ce qui a été énoncé concernant la limite en $+\infty$ est transposable au voisinage de $-\infty$. Des limites énoncées dans la propriété 2.1, on déduit notamment :
  - comme la fonction carré est paire, $\lim\limits_{x \to -\infty} x^2 = +\infty$
  - comme la fonction cube est impaire, $\lim\limits_{x \to -\infty} x^3 = -\infty$
- On peut étendre la propriété $\lim\limits_{x \to +\infty} \frac{1}{x^n} = 0$ aux fonctions $x \mapsto \frac{1}{(x+\alpha)^n}$ avec $n \in \mathbb{N}^*$ et α un réel quelconque. En effet $\lim\limits_{x \to +\infty} \frac{1}{(x+\alpha)^n} = 0$ car $\forall \epsilon > 0$, pour avoir $|\frac{1}{(x+\alpha)^n}| < \epsilon$ il suffit de prendre $|(x+\alpha)^n| > \frac{1}{\epsilon}$ soit $x > \sqrt[n]{\frac{1}{\epsilon}} - \alpha$.

Exemples et interprétations graphiques : Si une fonction f a une limite finie l en $+\infty$, la courbe de f se rapproche de plus en plus de la droite $\mathscr{D}$ d'équation $y = l$. Cette droite est appelée « asymptote horizontale » pour la courbe $\mathscr{C}_f$. On dit que la courbe tend asymptotiquement vers la droite : elle s'en rapproche sans jamais être confondue avec elle.

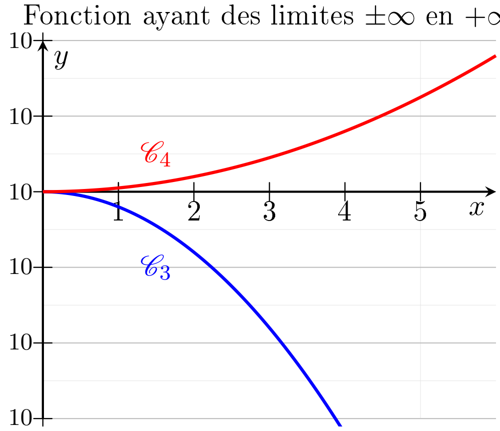
- La fonction g_1 définie par $g_1(x) = \frac{x}{x-1}$ a pour limite $l = 1$ en $+\infty$, car $\frac{x}{x-1} = \frac{x-1+1}{x-1} = 1 + \frac{1}{x-1}$. Cette fonction s'écrit $g_1(x) = 1 + \varphi(x)$ où $\varphi(x) = \frac{1}{x-1}$ est une fonction de limite nulle en $+\infty$ (notre remarque précédente justifie cela). La courbe de g_1 tend asymptotiquement vers la droite d'équation $y = 1$ (notée $\mathscr{C}_1$ sur le graphique de gauche).
- La fonction g_2 définie par $x \mapsto \frac{1}{3-x}$ tend, quant-à elle, vers 0 en $+\infty$ (notée $\mathscr{C}_2$).
- Les fonction g_3 et g_4 définies par $g_3(x) = -2x^2$ et $g_4(x) = \frac{x^2}{2}$, ont une limite infinie. On a $\lim\limits_{x \to +\infty} g_3(x) = -\infty$ et $\lim\limits_{x \to +\infty} g_4(x) = +\infty$ (la détermination de ces limites est très simple à l'aide des théorèmes qui suivent). Les courbes sont notées $\mathscr{C}_3$ et $\mathscr{C}_4$ sur le graphique de droite.

Fonction ayant des limites finies en $+\infty$:



La courbe s'approche de son asymptote horizontale (limite finie).

Fonction ayant des limites $\pm\infty$ en $+\infty$:



La courbe dépasse toute valeur arbitraire A vers le haut (limite $+\infty$) ou vers le bas (limite $-\infty$).

Propriété 2.3 (Unicité de la limite) Si une fonction f admet une limite finie l en $+\infty$ (respectivement en $-\infty$) alors cette limite est unique.

Démonstration Supposons que f admette deux limites l et l' en $+\infty$.
On a alors $\forall \epsilon > 0,\ \exists x_0 > a,\ \forall x > x_0, |f(x) - l| < \epsilon$
Et de même $\forall \epsilon > 0,\ \exists x'_0 > a,\ \forall x > x'_0, |f(x) - l'| < \epsilon$
D'après l'inégalité triangulaire $(\forall (a,b) \in \mathbb{R}^2,\ \ |a+b| < |a| + |b|)$, on a alors
$|l - l'| = |l - f(x) + f(x) - l'| < |l - f(x)| + |f(x) - l'| < 2\epsilon$.
Par conséquent $\forall \epsilon > 0,\ \forall x > \max(x_0, x'_0),\ \ |l - l'| < 2\epsilon$.
$|l - l'|$ peut donc être rendu aussi petit qu'on veut, donc plus petit que lui-même...
Cela étant impossible, on doit donc avoir $l = l'$.

1.b. Limite en un point

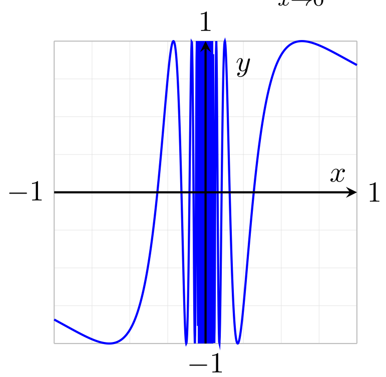
Définition 2.3 (Limite en un point) Une fonction f admet 0 comme limite en a si, quel que soit le nombre A, dès que x est assez proche de a, la valeur absolue de $f(x)$ est plus proche de 0 que A. Autrement dit : $\lim\limits_{x \to a} f(x) = 0 \iff \forall A \in \mathbb{R},\ \ \exists \epsilon > 0, |x - a| < \epsilon \Longrightarrow |f(x)| < A$

l étant un réel, une fonction f admet l comme limite en a si la fonction $f - l$ admet pour limite 0 en a. Autrement dit, s'il existe une fonction φ de limite 0 en a, telle que $f(x) = l + \varphi(x)$.

Remarques :

- La question de la limite en un point a est un problème à résoudre lorsqu'on est au bord de l'ensemble de définition de la fonction car $\forall x \in \mathscr{D}_f, \lim\limits_{x \to a} f(x) = f(a)$. La fonction $f : x \mapsto \frac{(1+x)^3 - 1}{x}$ par exemple n'est pas définie pour $x = 0$. On va simplifier par $x \neq 0$ et faire « tendre » x vers 0 : pour $x \neq 0$, $f(x) = \frac{(1+x)^3-1}{x} = \frac{1+3x^2+3x+x^3-1}{x} = \frac{x(3x+3+x^2)}{x} = 3x + 3 + x^2$.
La simplification finale n'est possible que si
$x \neq 0$ et, quand x s'approche infiniment près de 0, $3x+3+x^2$ s'approche de 3, soit $\lim\limits_{x \to 0} f(x) = 3$.

- Toutes les fonctions n'ont pas nécessairement une limite en un point.
Par exemple la fonction $x \mapsto \sin(\frac{1}{x})$ n'a pas de limite en 0 : quand x s'approche de 0, $\frac{1}{x}$ tend vers l'infini, et la fonction sin continue bravement à osciller entre -1 et $+1$ pour chaque intervalle d'amplitude 2π. La courbe obtenue contient une infinité d'oscillations dans la zone bleutée. On peut dilater cette zone autant qu'il nous plait, il y aura toujours cette zone au voisinage de 0.



- En considérant que $\lim\limits_{x \to a} f(x) = \lim\limits_{h \to 0} f(a + h)$, on peut ramener la recherche d'une limite en un point a, à la recherche de la limite en 0. Il suffit pour cela d'effectuer le changement de variable $x = a + h$. Lorsque h tend vers 0, x tend bien vers a.
La fonction $f : x \mapsto \frac{x^2-x-2}{x-2}$, par exemple, n'est pas définie pour $x = 2$.
En écrivant $x = 2 + h$, on a $f(2 + h) = \frac{(2+h)^2-(2+h)-2}{(2+h-2)^2} = \frac{h^2+3h}{h} = h + 3$.
$\lim\limits_{x \to 2} f(x) = \lim\limits_{h \to 0} f(2 + h) = \lim\limits_{h \to 0} h + 3 = 3$.
Ici $f(x) = \frac{(x-2)(x+1)}{x-2}$ se simplifie directement : si $x \neq 2$ on a $f(x) = x + 1$, et $\lim\limits_{x \to 2} x + 1 = 3$.
La fonction f peut être « prolongée par continuité » en la fonction affine $x \mapsto x + 1$.

Définition 2.4 (Limite infinie en un point) Une fonction f admet $+\infty$ comme limite en a si, quel que soit le nombre A, dès que x est assez proche de a, on a $f(x) > A$.
Autrement dit : $\lim\limits_{x\to a} f(x) = +\infty \iff \forall A \in \mathbb{R},\ \exists \epsilon > 0,\ |x-a| < \epsilon \Longrightarrow f(x) > A$

Remarques :

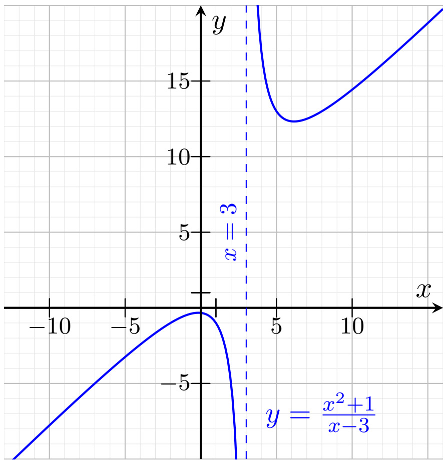
- On définit de même ce que signifie $\lim\limits_{x\to a} f(x) = -\infty$.
- Lorsque la limite d'une fonction f en a est $\pm\infty$, on dit que la droite d'équation $x = a$ est une « asymptote verticale » à la courbe représentative de f (la courbe s'approche de son asymptote sans jamais l'atteindre).

Exemple 15 – La fonction $f : x \mapsto \frac{x^2+1}{x-3}$ n'est pas définie pour $x = 3$.

En écrivant $x = 3 + h$, on a
$f(3+h) = \frac{(3+h)^2+1}{3+h-3} = \frac{h^2+6h+10}{h} = h + 6 + \frac{10}{h}$
Quand on fait tendre x vers 3, soit h vers 0, $h+6$ tend vers 6, mais $\frac{10}{h}$ tend vers l'infini. Il parait alors évident que $f(3+h)$ tend vers l'infini : $\lim\limits_{x\to 3} f(x) = \pm\infty$.
La droite d'équation $x = 3$ est une asymptote verticale à la courbe représentative de f. Comme f est définie des deux côtés de la valeur interdite $x = 3$, il y a deux limites :

- la limite à gauche notée $\lim\limits_{x\to 3^-} f(x)$ ou $\lim\limits_{\substack{x\to 3\\ x<3}} f(x)$
- la limite à droite notée $\lim\limits_{x\to 3^+} f(x)$ ou $\lim\limits_{\substack{x\to 3\\ x>3}} f(x)$

Ici, on a $\lim\limits_{x\to 3^-} f(x) = -\infty$ et $\lim\limits_{x\to 3^+} f(x) = +\infty$.



Propriété 2.4 (Limites de référence en 0) Chacune des fonctions suivantes a pour limite 0 en 0 : $x \mapsto x$, $x \mapsto x^2$, $x \mapsto x^3$, etc. (soit $x \mapsto x^n$ avec $n \in \mathbb{N}^*$) et $x \mapsto \sqrt{x}$.
n étant un entier non nul, les fonctions $x \mapsto \frac{1}{x^n}$ ont pour limite $+\infty$ en 0, excepté lorsque n est impair, la limite à gauche de 0 est alors $-\infty$.

Démonstration Pour ne pas trop nous répéter, nous laissons cette démonstration au lecteur.

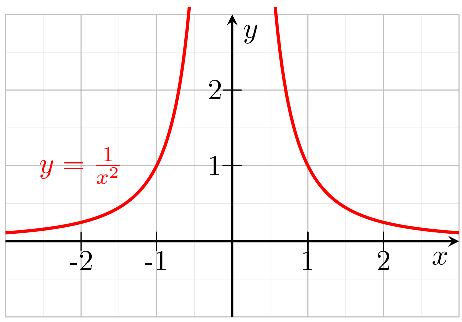
Remarque :

La fonction inverse correspond à $x \mapsto \frac{1}{x^n}$ pour $n = 1$.
n étant impair, on a $\lim\limits_{x\to 0^-} \frac{1}{x} = -\infty$, ce que l'examen de la courbe laisse déjà supposer.
On peut alors compléter le tableau de variation de cette fonction, en y faisant figurer les limites en 0 et aussi en $\pm\infty$.

x	$-\infty$		0		$+\infty$
$\frac{1}{x}$	0 ↘	$-\infty$	‖	$+\infty$ ↘	0

Pour $x \mapsto \frac{1}{x^2}$, $n = 2$ étant pair, la propriété affirme que les deux limites (à gauche et à droite de 0) sont $+\infty$.

Cela se voit sur la courbe et se note dans le tableau des variations de cette fonction.



x	$-\infty$		0		$+\infty$
$\frac{1}{x^2}$	0 ↗	$+\infty$	‖	$+\infty$ ↘	0

1.c. Théorèmes sur les Limites

Dans cette section, nous reprenons sans les démontrer les différents théorèmes qui sont exposés dans le chapitre sur les suites. Pour une suite, il ne s'agissait que de la limite éventuelle en $+\infty$. Il faut ici les transposer pour des limites éventuelles en $-\infty$ et aussi au voisinage d'un point a (à gauche ou à droite, là où la fonction n'est pas définie).

Théorème 2.1 (croissance majorée) Toute fonction croissante et majorée sur un intervalle $[a;+\infty[$ admet une limite finie l en $+\infty$ qui est le plus petit majorant de la fonction.
De même, pour toute fonction décroissante et minorée : sa limite est le plus grand des minorants.

Remarques :

- Exemples de transposition :
  - Toute fonction décroissante et minorée, définie sur un intervalle de la forme $[a;b[$ admet une limite à gauche finie l en b qui est le plus grand des minorants.
  - Toute fonction croissante et minorée, définie sur un intervalle de la forme $]-\infty;a]$ admet une limite finie l en $-\infty$ qui est le plus grand des minorants.
- Une fonction f a une limite finie l en $\pm\infty$ si $\lim\limits_{x\to\pm\infty} f(x) - l = 0$. Dans ce cas la courbe $\mathscr{C}_f$ admet une asymptote horizontale δ d'équation $y = l$ dans ce voisinage et on détermine la position de $\mathscr{C}_f$ par rapport à δ en examinant le signe de cette limite.

Théorème 2.2 (Comparaison) Soient f et g deux fonctions.
Si, à partir d'un certain réel x_0, on a $f \geqslant g$ et si $\lim\limits_{x\to+\infty} g(x) = +\infty$, alors $\lim\limits_{x\to+\infty} f(x) = +\infty$.
De même, si à partir d'un réel x_0, on a $f \leqslant g$ et si $\lim\limits_{x\to+\infty} g(x) = -\infty$, alors $\lim\limits_{x\to+\infty} f(x) = -\infty$.

Théorème 2.3 (Encadrement) Soient f, g et h trois fonctions. Si, à partir d'un certain réel x_0, on a $g \leqslant f \leqslant h$ et si $\lim\limits_{x\to+\infty} g(x) = \lim\limits_{x\to+\infty} h(x) = l$, alors $\lim\limits_{x\to+\infty} f(x) = l$.

Exemple 16 – La fonction $f : x \mapsto \frac{\sqrt{1+x^2}}{x}$ est définie sur $\mathbb{R}^*$.
Déterminons les limites de f aux bornes de son ensemble de définition, soit en 0 et en $\pm\infty$.

Lorsque $x > 0$, on a $0 < 1 < 2x + 1$ et, en ajoutant $x^2 > 0$, on obtient l'encadrement $x^2 < 1 + x^2 < (1 + x)^2$.
Comme la fonction racine carrée est croissante, cela implique que $x < \sqrt{1 + x^2} < 1 + x$.
En divisant par $x > 0$, on obtient l'encadrement de $f(x)$

$$1 < f(x) < 1 + \frac{1}{x}$$

Comme $\lim\limits_{x\to+\infty} 1 + \frac{1}{x} = 1$, d'après le théorème 2.3, on peut donc conclure que $\lim\limits_{x\to+\infty} f(x) = 1$.

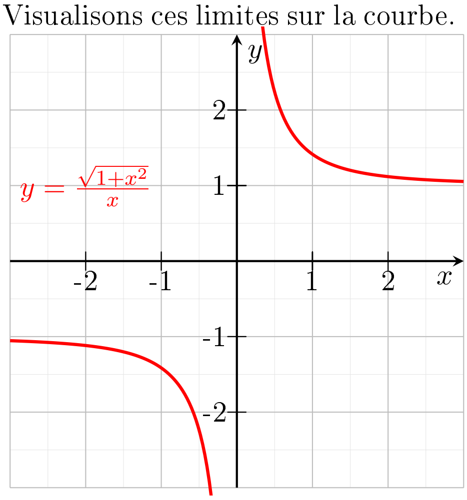
Au voisinage de $0+$ (x tend vers 0 et $x > 0$), la majoration $f(x) < 1 + \frac{1}{x}$ ne permet pas de conclure.
Par contre, on peut remarquer que $1 < \sqrt{1 + x^2}$ et donc

$$\frac{1}{x} < f(x)$$

Comme $\lim\limits_{x\to0^+} \frac{1}{x} = +\infty$, d'après le théorème 2.2, on peut donc conclure que $\lim\limits_{x\to0^+} f(x) = +\infty$.

Voilà pour les deux bornes ouvertes positives de $\mathscr{D}_f$. Pour les deux autres, il suffit de remarquer que la fonction est impaire. En effet, $f(-x) = \frac{\sqrt{1+(-x)^2}}{-x} = -\frac{\sqrt{1+x^2}}{x} = -f(x)$.
Les limites seront donc « symétriques » par rapport à O : $\lim\limits_{x\to0^-} f(x) = -\infty$ et $\lim\limits_{x\to-\infty} f(x) = -1$.

Visualisons ces limites sur la courbe.



Complétons le tableau de variation.

x	$-\infty$		0		$+\infty$
$\frac{\sqrt{1+x^2}}{x}$	-1	$\searrow$	$-\infty$ ‖ $+\infty$	$\searrow$	1

Théorème 2.4 (somme et produit par un réel) Soient f une fonction qui tend vers $+\infty$ et k un réel non nul.

- La fonction $(f+k)$ tend vers $+\infty$ et, si $k>0$, la fonction (kf) aussi
- Si $k<0$, la fonction (kf) tend vers $-\infty$

Théorème 2.5 (Somme et Produit de fonctions) Soient f et g deux fonctions ayant une limite, finie ou infinie. Les fonctions $f+g$ et $f\times g$, admettent ou non une limite selon les différents cas exposés par les deux tableaux ci-dessous :

Fonction « somme » : $(f+g)$

$\lim f$ / $\lim g$	l	$+\infty$	$-\infty$
l'	$l+l'$	$+\infty$	$-\infty$
$+\infty$	$+\infty$	$+\infty$	
$-\infty$	$-\infty$		$-\infty$

Fonction « produit » : $(f\times g)$

$\lim f$ / $\lim g$	$l\neq 0$	$l=0$	$\pm\infty$
$l'\neq 0$	$l\times l'$	0	$\pm\infty$
$l'=0$	0	0	
$\pm\infty$	$\pm\infty$		$\pm\infty$

Théorème 2.6 (Quotient de fonctions) Soient f et g deux fonctions ayant une limite, finie ou infinie. La fonction quotient $\frac{f}{g}$ admet ou non une limite selon les différents cas exposés par le tableau ci-dessous :

$\lim f$ / $\lim g$	$l\neq 0$	$l=0$	$\pm\infty$
$l'\neq 0$	$\frac{l}{l'}$	0	$\pm\infty$
$l'=0$	$\pm\infty$		$\pm\infty$
$\pm\infty$	0	0	

Remarques :

- Pour un produit ou un quotient, c'est la règle des signes qui détermine le signe des limites.
- Les *formes indéterminées* (les cas hachurés ci-dessus) sont de quatre types :
  - $\pm(\infty-\infty)$, par exemple $\lim\limits_{x\to+\infty} x-\sqrt{x^2+1}$ ou $\lim\limits_{x\to-\infty} x^2+x+1$
  - $\pm(0\times\infty)$, par exemple $\lim\limits_{x\to0^+}(e^x-1)(\ln x+1)$ ou $\lim\limits_{x\to-\infty} e^x(x^2+1)$
  - $\pm\frac{0}{0}$, par exemple $\lim\limits_{x\to1^+}\frac{x^2+x-2}{\sqrt{x-1}}$ ou $\lim\limits_{x\to0}\frac{\sin x}{x}$
  - $\pm\frac{\infty}{\infty}$, par exemple $\lim\limits_{x\to+\infty}\frac{e^x}{x}$ ou $\lim\limits_{x\to-\infty}\frac{x^2+1}{x^3-1}$

  Une factorisation suivie d'une simplification suffit parfois à lever l'indétermination.

Méthode (Fonction polynôme)

Pour une fonction polynôme de degré p : $f : x\mapsto\sum_{k=0}^{p}a_kx^k$, la limite au voisinage de $\pm\infty$ est celle du terme dominant (celui de plus haut degré) :

$$\lim_{x\to\pm\infty}\sum_{k=0}^{p}a_kx^k=\lim a_px^p$$

- Au voisinage de $+\infty$: si $a_p>0$, $\lim\limits_{x\to+\infty}f(x)=+\infty$ mais si $a_p<0$, $\lim\limits_{x\to+\infty}f(x)=-\infty$.
- Au voisinage de $-\infty$, on doit tenir compte de la parité de p :
  - Si $a_p>0$ et p pair, $\lim\limits_{x\to-\infty}f(x)=+\infty$, si $a_p>0$ et p impair, $\lim\limits_{x\to-\infty}f(x)=-\infty$.
  - Si $a_p<0$ c'est l'inverse : si p pair, $\lim\limits_{x\to-\infty}f(x)=-\infty$; si p impair, $\lim\limits_{x\to-\infty}f(x)=+\infty$).

La généralité énoncée s'étend aux fonctions polynomiales ayant des termes de degré négatif.
La fonction f définie par $f(x)=-x^3+3x^2-2x+1-\frac{2}{x}+\frac{7}{x^2}$, par exemple, a un comportement au voisinage de $\pm\infty$ qui est celui du terme de plus haut degré (de $-x^3$ ici).

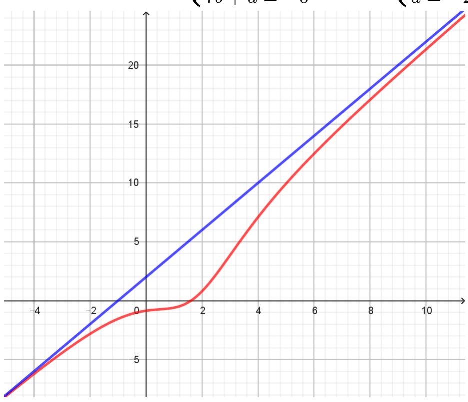
Exemple 17 – ➲ Justifions la propriété sur les limites de polynômes par un exemple où l'indétermination est du type $\infty - \infty$: $\lim\limits_{x \to +\infty} 2x^3 - 4x^2 + 5x - 6 = \lim\limits_{x \to +\infty} x^3\left(2 - \frac{4}{x} + \frac{5}{x^2} - \frac{6}{x^3}\right)$.
Or, $x \longmapsto \frac{1}{x}$ étant une fonction de référence de limite 0 en $+\infty$: $\lim\limits_{x \to +\infty} -\frac{4}{x} = 0$.
De même $\lim\limits_{x \to +\infty} \frac{5}{x^2} = 0$ et $\lim\limits_{x \to +\infty} \frac{6}{x^3} = 0$. On en déduit $\lim\limits_{x \to +\infty} 2 - \frac{4}{x} + \frac{5}{x^2} - \frac{6}{x^3} = 2$.
Finalement, d'après les théorèmes sur les opérations avec les limites :
$\lim\limits_{x \to +\infty} x^3\left(2 - \frac{4}{x} + \frac{5}{x^2} - \frac{6}{x^3}\right) = \lim\limits_{x \to +\infty} 2x^3 = +\infty$. Et de même
$\lim\limits_{x \to -\infty} x^3\left(2 - \frac{4}{x} + \frac{5}{x^2} - \frac{6}{x^3}\right) = \lim\limits_{x \to -\infty} 2x^3 = -\infty$.
Le même procédé de factorisation du monôme de plus haut degré conduit toujours au même résultat : c'est lui qui détermine la limite du polynôme en $\pm\infty$.

➲ On peut aussi généraliser la détermination de la limite d'un quotient de deux polynômes en $\pm\infty$ (à priori, elle est du type $\pm\frac{\infty}{\infty}$) : elle est égale à la limite du quotient des monômes de plus haut degré.
$\lim\limits_{x \to +\infty} \frac{2x^3-4x^2+5x-6}{x^2-3x+7} = \lim\limits_{x \to +\infty} \frac{x^3\left(2-\frac{4}{x}+\frac{5}{x^2}-\frac{6}{x^3}\right)}{x^2\left(1-\frac{3}{x}+\frac{7}{x^2}\right)}$.
Comme précédemment $\lim\limits_{x \to +\infty} 2 - \frac{4}{x} + \frac{5}{x^2} - \frac{6}{x^3} = 2$ et $\lim\limits_{x \to +\infty} 1 - \frac{3}{x} + \frac{7}{x^2} = 1$.
On en déduit que $\lim\limits_{x \to +\infty} \frac{x^3\left(2-\frac{4}{x}+\frac{5}{x^2}-\frac{6}{x^3}\right)}{x^2\left(1-\frac{3}{x}+\frac{7}{x^2}\right)} = \lim\limits_{x \to +\infty} \frac{2x^3}{x^2} = \lim\limits_{x \to +\infty} 2x = +\infty$.

Á propos de cette fonction, montrons que sa courbe admet une *asymptote oblique*.
Montrons qu'on peut écrire l'expression $\frac{2x^3-4x^2+5x-6}{x^2-3x+7}$ sous la forme $ax + b + \frac{cx+d}{x^2-3x+7}$.
En mettant la 2ᵉ expression au même dénominateur, on doit avoir : $\frac{2x^3-4x^2+5x-6}{x^2-3x+7} = \frac{ax^3+(b-3a)x^2+(c+7a-3b)x+7b+d}{x^2-3x+7}$

Par identification des coefficients des monômes de même degré, on doit avoir $\begin{cases} a = 2 \\ b - 3a = -4 \\ c + 7a - 3b = 5 \\ 7b + d = -6 \end{cases} \iff \begin{cases} a = 2 \\ b = 2 \\ c = -3 \\ d = -20 \end{cases}$

Conclusion : $\frac{2x^3-4x^2+5x-6}{x^2-3x+7} = 2x + 2 - \frac{3x+20}{x^2-3x+7}$
Comme $\lim\limits_{x \to +\infty} \frac{3x+20}{x^2-3x+7} = \lim\limits_{x \to +\infty} \frac{3x}{x^2} = \lim\limits_{x \to +\infty} \frac{3}{x} = 0^+$,
on en déduit $\lim\limits_{x \to +\infty} \frac{2x^3-4x^2+5x-6}{x^2-3x+7} - (2x+2) = \lim\limits_{x \to +\infty} \frac{-3}{x} = 0^-$.
De même $\lim\limits_{x \to -\infty} \frac{2x^3-4x^2+5x-6}{x^2-3x+7} - (2x+2) = \lim\limits_{x \to -\infty} \frac{-3}{x} = 0^+$.
La droite d'équation $y = 2x + 2$ est donc asymptote à la courbe de la fonction $x \longmapsto \frac{2x^3-4x^2+5x-6}{x^2-3x+7}$ au voisinage de l'infini ; la courbe est au dessous de son asymptote au voisinage de $+\infty$ et, même si ça ne se voit pas sur l'illustration, au dessus au voisinage de $-\infty$.



➲ Montrons un autre procédé pour lever une indétermination sur les limites avec des racines carrées. Comme souvent avec les racines carrées, on utilise la *quantité conjuguée*.
Cherchons la limite de $x \longmapsto x - \sqrt{x^2+1}$ en $+\infty$ (elle est du type $\infty - \infty$).
$\lim\limits_{x \to +\infty} x - \sqrt{x^2+1} = \lim\limits_{x \to +\infty} \frac{(x-\sqrt{x^2+1})(x+\sqrt{x^2+1})}{x+\sqrt{x^2+1}} = \lim\limits_{x \to +\infty} \frac{x^2-(x^2+1)}{x+\sqrt{x^2+1}} = \lim\limits_{x \to +\infty} \frac{-1}{x+\sqrt{x^2+1}}$.
Ainsi transformée, l'expression a perdu son indétermination : comme $\lim\limits_{x \to +\infty} x + \sqrt{x^2+1} = +\infty$, on en déduit que $\lim\limits_{x \to +\infty} x - \sqrt{x^2+1} = 0^-$ (le signe de la limite peut être utile).

➲ Un dernier exemple de levée d'indétermination sur les limites avec une fonction exponentielle.
Cherchons la limite de $x \longmapsto \frac{e^x}{x}$ en $+\infty$ (elle est du type $+\frac{\infty}{\infty}$).
Étudions les variations de la fonction f définie par $f(x) = e^x - \frac{x^2}{2}$. On a $f'(x) = e^x - x$.
On sait que $\forall x \in \mathbb{R}, e^x \geqslant x + 1 > x$ (par l'étude de la fonction $x \longmapsto e^x - x - 1$).
On en déduit que $\forall x \in \mathbb{R}, e^x - x > 0 \iff f'(x) > 0$; f est strictement croissante sur $\mathbb{R}$.
Ainsi, $x > 0 \implies f(x) > f(0) \iff e^x - \frac{x^2}{2} > e^0 - \frac{0^2}{2} \iff e^x - \frac{x^2}{2} > 1 \iff e^x > \frac{x^2}{2} + 1 > \frac{x^2}{2}$.
En divisant par $x > 0$ on obtient $\forall x > 0, \frac{e^x}{x} > \frac{x}{2}$.
Comme $\lim\limits_{x \to +\infty} \frac{x}{2} = +\infty$, d'après le théorème de comparaison, on en déduit $\lim\limits_{x \to +\infty} \frac{e^x}{x} = +\infty$.

Propriété 2.5 (Fonction composée) Soient a, l et l' des nombres finis ou infinis, soit f une fonction définie sur un intervalle semi-ouvert $I =]a, b]$ ou $[b, a[$ et soit g une fonction définie sur un intervalle $J = f(I)$.
Si $\lim\limits_{x\to a} f(x) = l$ et si $\lim\limits_{x\to l} g(x) = l'$ alors $\lim\limits_{x\to a} g \circ f(x) = l'$.

Démonstration Nous admettrons cette propriété qui admet de nombreuses situations (a, l et l' pouvant être finis ou infinis)

Remarques :

- J'ai déjà implicitement employé cette propriété :
  Pour la 3e limite du précédent exemple (la limite de $x \longmapsto x - \sqrt{x^2+1}$) j'ai affirmé sans preuve que $\lim\limits_{x\to+\infty} \sqrt{x^2+1} = +\infty$ mais il s'agit de la limite d'une fonction composée $g \circ f$ avec la fonction f définie sur $[0, +\infty[$ par $f(x) = x^2 + 1$ telle que $\lim\limits_{x\to+\infty} f(x) = +\infty$ et la fonction g définie sur $[0, +\infty[$ par $g(x) = \sqrt{x}$ telle que $\lim\limits_{x\to+\infty} g(x) = +\infty$.
  La propriété énoncée nous permet alors de conclure : $\lim\limits_{x\to+\infty} g \circ f(x) = +\infty$
  (dans ce cas, on a $a = l = l' = +\infty$)
- Cette propriété peut servir à déterminer la limite d'une suite (v_n) définie par la relation $v_n = f(u_n)$ où (u_n) est une suite de limite $l \in \mathbb{R} \cup \{-\infty, +\infty\}$ et f une fonction telle que $\lim\limits_{x\to l} f(x) = l' \in \mathbb{R} \cup \{-\infty, +\infty\}$. Dans ces conditions, la suite (v_n) admet pour limite l'.

Exemple 18 – Étudions le comportement asymptotique de la suite (v_n) définie par $v_n = n \sin\left(\frac{1}{n}\right)$.
On écrit $v_n = \dfrac{\sin\left(\frac{1}{n}\right)}{\frac{1}{n}}$ et on reconnait que $v_n = \dfrac{\sin u_n}{u_n} = f(u_n)$ avec $u_n = \frac{1}{n}$ défini pour $n \in \mathbb{N}^*$ et f la fonction définie par $f(x) = \dfrac{\sin x}{x}$.
Les termes u_n appartiennent tous à l'intervalle $]0, 1]$ et la fonction f est définie sur cet intervalle.

Nous l'avons déterminée en classe de 1re et nous le rappelons : $\lim\limits_{x\to 0} \dfrac{\sin x}{x} = 1$.
On peut comprendre cela si on admet que $\sin'(x) = \cos(x)$ car cette limite équivaut au nombre dérivé en 0 de la fonction sinus : $\lim\limits_{h\to 0} \dfrac{\sin h - \sin 0}{h - 0} = \sin'(0) = \cos 0 = 1$.
Comme $\lim\limits_{n\to\infty} u_n = \lim\limits_{n\to\infty} \frac{1}{n} = 0$ et $\lim\limits_{x\to 0} f(x) = 1$, on en déduit que $\lim\limits_{n\to\infty} v_n = \lim\limits_{n\to\infty} f(u_n) = 1$.

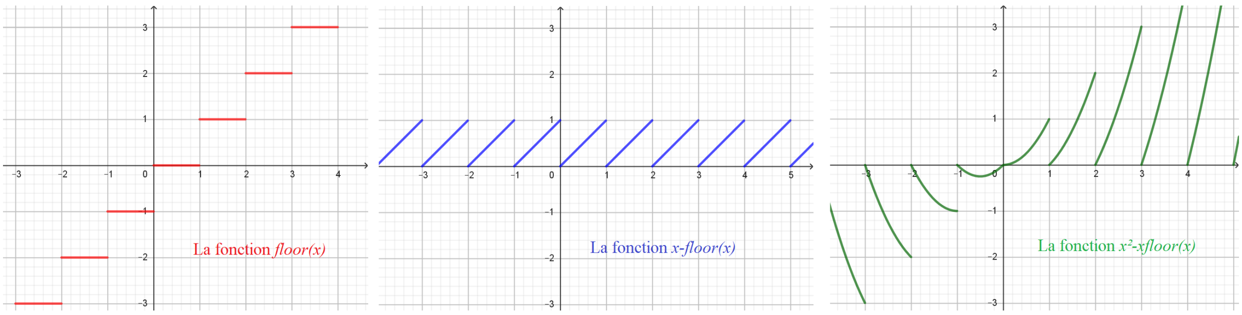
2. Fonctions continues

2.a. Continuité

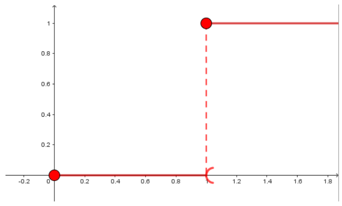
Définition 2.5 (Continuité en un point) Une fonction f, définie sur un intervalle I est continue en un réel $x_0 \in I$ si et seulement si : $\lim\limits_{x\to x_0} f(x) = f(x_0)$.

Remarques :
On définit la continuité *à droite* en $x_0 \in I$ si la fonction f vérifie :
$\lim\limits_{\substack{x\to a\\ x>a}} f(x) = f(a)$ (on note aussi cela $\lim\limits_{x\to a^+} f(x) = f(a)$).
De même, on définit la continuité *à gauche* en $x_0 \in I$ si la fonction f vérifie :
$\lim\limits_{\substack{x\to b\\ x<b}} f(x) = f(b)$ (on note aussi cela $\lim\limits_{x\to a^-} f(x) = f(a)$).
Ainsi, la continuité en $x_0 \in I$ est la conjonction d'une continuité à droite et à gauche en x_0.

Inversement, si une fonction est définie en un point $x_0 \in I$ mais si elle n'est pas continue d'un côté (à droite ou à gauche), alors elle n'est pas continue en x_0. Dans ce cas, lorsqu'on trace la courbe de cette fonction, on doit lever le stylo en passant par la valeur x_0. Ci-dessous, j'ai tracé trois fonctions discontinues en tous points entiers de $\mathbb{R}$. Elles sont obtenues à partir de la fonction *floor* – la fonction « plancher« : $x \longmapsto \lfloor x \rfloor$) – discontinue à gauche car $\forall n \in \mathbb{N}$, $\lim\limits_{x \to n^-} \lfloor x \rfloor = n - 1$ alors que $\lfloor n \rfloor = n$.



La discontinuité est parfois représentée par un procédé graphique tel que celui ci-contre mais les traceurs de courbe (logiciels, calculatrices, etc.) n'entrent généralement pas dans ces considérations (ci-dessus tracés avec Geogebra).



Propriété 2.6 Soit I un intervalle où la fonction f est définie et soit $x_0 \in I$.
f est continue en x_0 si et seulement si l'une des propositions suivantes est vraie :
- $\lim\limits_{h \to 0} f(x_0 + h) = f(x_0)$
- $\forall A > 0, \exists \epsilon > 0, \forall x \in I, |x - x_0| < \epsilon \Rightarrow |f(x) - f(x_0)| < A$

Démonstration La première proposition traduit simplement la définition d'une fonction continue en x_0 en posant $x = x_0 + h$ et en faisant tendre h vers 0 on fait tendre x vers x_0.

La seconde traduit la définition 2.3 : une fonction f admet une limite l en x_0 si et seulement si
$\forall A > 0, \ \exists \epsilon > 0, |x - x_0| < \epsilon \Longrightarrow |f(x) - l| < A$
On voit qu'il suffit ici de remplacer l par $f(x_0)$ qui existe puisque la fonction est définie en x_0.

Définition 2.6 (Continuité sur un intervalle) Soient I un intervalle et f une fonction définie sur I. f est continue sur I si et seulement si elle est continue en tous points $x_0 \in I$.

La fonction $x \longmapsto \lfloor x \rfloor$ est continue sur $[n, n+1[$ pour tout entier $n \in \mathbb{N}$ et discontinue sur $\mathbb{R}$.

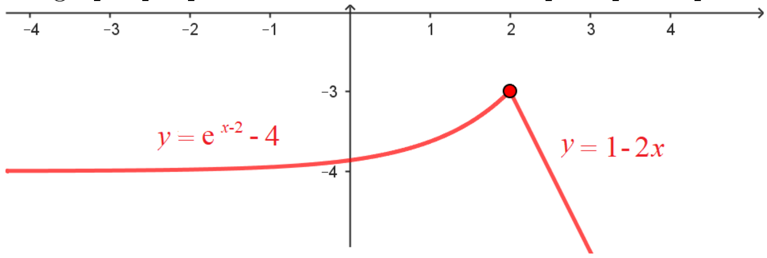
Propriété 2.7 (Fonctions de référence) Les fonctions de référence suivantes sont continues sur leur ensemble de définition :
- Continues sur $\mathbb{R}$: les fonctions polynômes, la fonction exponentielle, les fonctions sinus et cosinus, la fonction valeur absolue
- Continue sur $\mathbb{R}^+$: la fonction racine carrée

Toute fonction obtenue par opérations ou composition avec les fonctions de référence est continue sur tout intervalle I où elle est définie.

Démonstration Cette propriété est admise.

Remarques :

- La fonction $x \longmapsto \frac{e^x+1}{x^2-1}$ est définie sur $\mathbb{R}\backslash\{-1,1\}$. Sur chacun des intervalles où elle est définie ($]-\infty,-1[$, $]-1,1[$ et $]1,+\infty[$) cette fonction est continue car c'est le quotient de deux fonctions continues $x \longmapsto e^x+1$ et $x \longmapsto x^2-1$. D'une façon générale une fonction rationnelle (quotient de deux polynômes) est continue sur tout intervalle inclus dans son ensemble de définition.
- La question de la continuité sur un intervalle est ramenée à la continuité en un point pour des fonctions définies par morceaux, chacun étant continus d'après la propriété 2.7.
  Exemple : la fonction f est définie par $\begin{cases} 1-2x & \text{si } x \geqslant 2 \\ e^{x-2}-4 & \text{si } x < 2 \end{cases}$
  Cette fonction est continue sur $]-\infty,2[$ car elle y est définie par composition et somme de fonctions continues sur $\mathbb{R}$; f est aussi continue sur $[2,+\infty[$ car c'est alors une fonction affine, continue sur $\mathbb{R}$.
  Pour montrer que f est continue sur $\mathbb{R}$, il faut montrer la continuité à gauche en 2 :
  $\lim\limits_{x\to 2^-} f(x) = \lim\limits_{x\to 2^-} e^{x-2}-4 = e^0-4 = -3$. Or $f(2) = 1-2\times 2 = -3$.
  Donc $\lim\limits_{x\to 2^-} f(x) = f(2)$: la fonction f est continue à gauche et par conséquent continue sur $\mathbb{R}$.
  Un graphique peut le confirmer mais il ne peut pas le prouver.

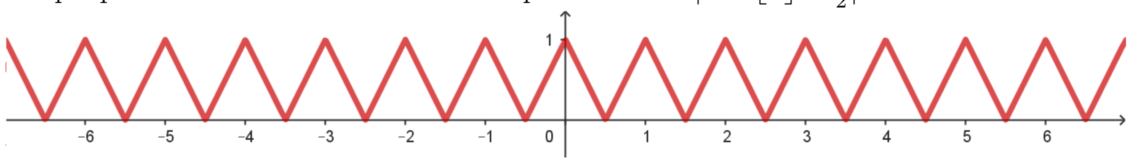


Propriété 2.8 (Fonction dérivable) Soient I un intervalle ouvert et f une fonction.
Si la fonction f est dérivable sur I alors elle est continue sur I.

Démonstration Si la fonction est dérivable en un point $x_0 \in I$, cela signifie que :
$f(x_0+h) = f(x_0) + hf'(x_0) + h\varphi(h)$ où φ est une fonction de limite 0 en 0 (voir le cours de 1^re).
On en déduit $\lim\limits_{h\to 0} f(x_0+h) = \lim\limits_{h\to 0} f(x_0) + \lim\limits_{h\to 0} hf'(x_0) + \lim\limits_{h\to 0} h\varphi(h) = f(x_0)$ ce qui est, d'après la propriété 2.6, une caractéristique de la continuité en x_0.

Remarques :

- Cette déduction se fait sur un intervalle ouvert I car il peut se trouver qu'une fonction ne soit pas dérivable aux bornes d'un intervalle fermé où elle est définie. La fonction $x \longmapsto \sqrt{x}$ par exemple est définie et continue sur $[0,+\infty[$ alors qu'elle n'est dérivable que sur $]0,+\infty[$. En effet, $\lim\limits_{h\to 0^+} \frac{\sqrt{h}-\sqrt{0}}{h-0} = \lim\limits_{h\to 0^+} \frac{1}{\sqrt{h}} = +\infty$, sa dérivée en 0 n'est pas définie.
- Une fonction continue sur un intervalle ouvert I n'est pas forcément dérivable en tous points : il suffit de considérer la fonction $x \longmapsto |x|$ qui est continue sur $\mathbb{R}$ mais pas dérivable en 0 car $\lim\limits_{x\to 0^-} |x|' = \lim\limits_{x\to 0^-} (-x)' = -1$ alors que $\lim\limits_{x\to 0^+} |x|' = \lim\limits_{x\to 0^-} (x)' = 1$.
  Et que penser de cette fonction dont l'expression est $2|x - \lfloor x \rfloor - \frac{1}{2}|$?



PROPRIÉTÉ 2.9 (SUITES) Soient I un intervalle ouvert, f une fonction définie sur I et (u_n) une suite telle que $\forall n,\ u_n \in I$. Si f est continue sur I et si $\lim\limits_{n\to+\infty} u_n = l$ avec $l \in I$, alors :

- La suite (v_n) définie par la relation $v_n = f(u_n)$ converge vers $f(l)$.
- La suite (u_n) définie par une valeur initiale et la relation $u_{n+1} = f(u_n)$ converge vers une des solutions de l'équation $l = f(l)$.

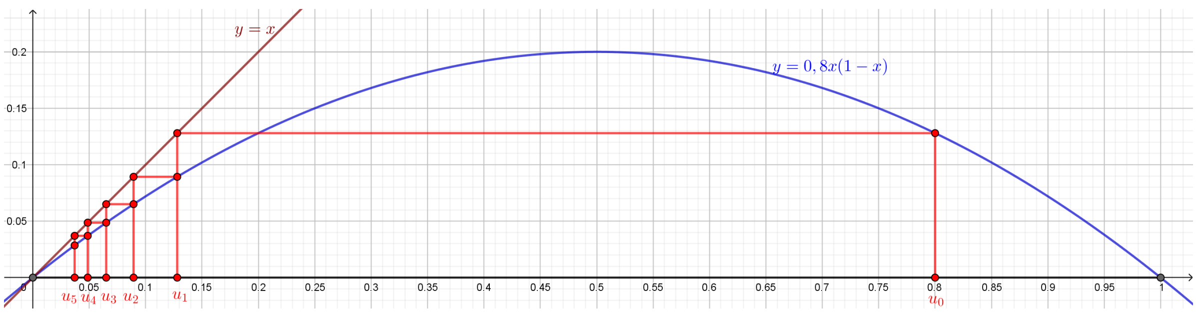
DÉMONSTRATION ➠ Pour la suite (v_n) définie par $v_n = f(u_n)$:
Comme $\lim\limits_{n\to+\infty} u_n = l$ et comme f est continue en $l \in I$, d'après la propriété 2.5 on a $\lim\limits_{n\to+\infty} v_n = \lim\limits_{n\to+\infty} f(u_n) = f(l)$.

➠ Pour la suite (u_n) définie par $u_{n+1} = f(u_n)$ et une valeur initiale, par exemple u_0 :
D'après la propriété précédente, comme $\lim\limits_{n\to+\infty} u_n = l$, on a $\lim\limits_{n\to+\infty} f(u_n) = f(l)$.
La suite (u_{n+1}) est une suite dont tous les termes sont identiques à ceux de la suite (u_n), excepté le terme u_0 qui manque à la suite (u_{n+1}), on peut donc affirmer que $\lim\limits_{n+1\to+\infty} u_{n+1} = \lim\limits_{N\to+\infty} u_N = l$.
La propriété 2.3 d'unicité de la limite permet de conclure : $f(l) = l$.

Remarques :

- Donnons un exemple d'application de la 1^re propriété : la suite (u_n) définie par $u_n = e^{1+\frac{1}{2}+\frac{1}{4}+\cdots+\frac{1}{2^n}}$ est une suite du type $u_n = e^{v_n}$ où v_n est la somme des $n+1$ premiers termes d'une suite géométrique de raison $\frac{1}{2}$ et de premier terme 1 ; on en déduit que $v_n = \frac{1-\frac{1}{2^{n+1}}}{1-\frac{1}{2}}$; la suite (v_n) converge vers $l = \lim\limits_{n\to+\infty} v_n = \frac{1-0}{\frac{1}{2}} = 2$; comme la fonction $x \longmapsto e^x$ est continue sur $\mathbb{R}$, on en déduit que (u_n) converge vers $e^l = e^2$.
- Dans les conditions de la 2^e propriété, la suite (u_n) définie par $u_{n+1} = f(u_n)$ converge vers une des solutions de l'équation $f(x) = x$ (il peut y avoir plusieurs solutions mais il ne peut y avoir qu'une seule limite). La réciproque est fausse : le fait qu'il existe une solution à l'équation $f(x) = x$ ne signifie pas que la suite (u_n) converge.

EXEMPLE 19 (SUITE LOGISTIQUE) – On considère une suite (x_n) définie par la relation de récurrence $x_{n+1} = \mu x_n(1-x_n)$, $x_0 \in [0;1]$ où μ est un paramètre qui appartient à l'intervalle $[0;4]$.
Je choisis $\mu = 0,8$ et $x_0 = 0,8$ et trace la courbe d'équation $y = 0,8x(1-x)$.
On constate que la suite (x_n) est décroissante et semble converger vers 0.

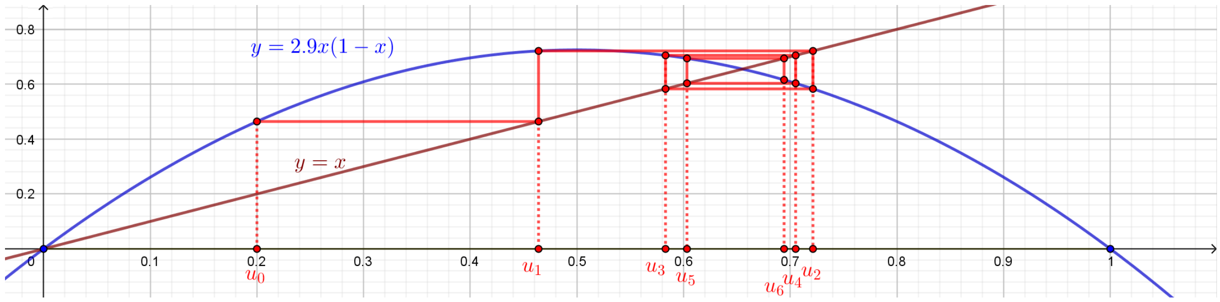


Si la suite converge vers une limite l, celle-ci est solution de l'équation $x = \mu x(1-x)$ qui a pour solution évidente $x = 0$, l'autre étant la solution de $1 = \mu(1-x) \iff x = 1 - \frac{1}{\mu}$.
Ainsi, les deux valeurs potentielles pour la limite de cette suite sont 0 et $1 - \frac{1}{\mu}$ (ici $1 - \frac{1}{0{,}8} = -0,25$).

Montrons que, lorsque $0 < \mu < 1$, la suite converge.
Montrons par récurrence que les termes de la suite (x_n) sont tous compris entre 0 et 1 :
En effet, on a $x_0 \in [0;1]$. Supposons que $0 \leqslant x_n \leqslant 1$ et montrons que $0 \leqslant x_{n+1} \leqslant 1$
$-1 \leqslant -x_n \leqslant 0 \Longrightarrow 0 \leqslant 1 - x_n \leqslant 1 \Longrightarrow 0 \leqslant \mu x_n(1-x_n) \leqslant \mu x_n \leqslant \mu < 1 \Longrightarrow 0 \leqslant x_{n+1} < 1$.
Si $x_n \neq 0$ on a $\frac{x_{n+1}}{x_n} = \mu(1-x_n)$ et comme $0 \leqslant \mu(1-x_n) \leqslant \mu < 1$ on en déduit que $0 \leqslant \frac{x_{n+1}}{x_n} \leqslant 1$.
La suite (x_n) est donc décroissante et minorée par 0. Sa limite est donc supérieure ou égale à 0.

L'autre limite potentielle étant négative ($0 < \mu < 1 \Longrightarrow \frac{1}{\mu} > 1 \Longrightarrow 1 - \frac{1}{\mu} < 0$) sa limite est 0.

Ci-dessous j'ai représenté le cas $\mu = 2,9$ avec $x_0 = 0,2$ donc lorsque $x_{n+1} = 2,9x_n(1-x_n)$.
La suite semble s'approcher alternativement de la 2e limite potentielle qui est ici $1 - \frac{1}{2,9} = \frac{19}{29} \approx 0,655$.



2.b. Théorème des Valeurs Intermédiaires (TVI)

Théorème 2.7 (Valeurs Intermédiaires) Soit f une fonction continue sur un intervalle $[a,b]$. Pour tout réel λ entre $f(a)$ et $f(b)$ il existe au moins une valeur $\alpha \in [a,b]$ telle que $f(\alpha) = \lambda$.

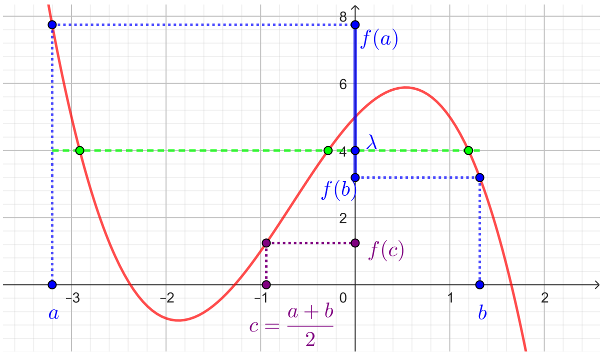
Démonstration Dans le cas $f(a) = \lambda$ (respect. $f(b) = \lambda$), on en déduit que $\alpha = a$ (resp. $\alpha = b$).
Dans les autres cas, supposons que l'on ait $f(a) < \lambda < f(b)$ (même raisonnement si $f(b) < \lambda < f(a)$).
Écrivons cela $f(a) - \lambda < 0 < f(b) - \lambda$; en définissant $g : x \longmapsto f(x) - \lambda$, cela s'écrit $g(a) < 0 < g(b)$.
$g(a) < 0$ et $g(b) > 0$, autrement dit $g(a) \times g(b) < 0$.
Notons $a_0 = a$ et $b_0 = b$ ces premières valeurs de deux suites (a_n) et (b_n) que l'on va construire.
Prenons $c = \frac{a_0+b_0}{2}$ et calculons $g(c)$:

- Si $g(c) = 0$ on a trouvé la valeur $\alpha = c$ cherchée et on s'arrête (sinon on continue)
- Si $g(c) > 0$ on donne à b_1 la valeur de c (a_1 conserve la valeur a_0)
- Si $g(c) < 0$ on donne à a_1 la valeur de c (b_1 conserve la valeur b_0)

Dans les deux premiers cas, le nouvel intervalle $[a_1, b_1]$ a une amplitude moitié de l'intervalle $[a_0, b_0]$: $b_1 - a_1 = \frac{b_0 - a_0}{2}$. De plus, on remarque que $a_0 \leqslant a_1$ et $b_0 \geqslant b_1$. Ce choix implique $g(a_1) < 0$ et $g(b_1) > 0$, on a donc $g(a_1) \times g(b_1) < 0$ et on se retrouve dans la situation initiale.
Selon le signe de $g(\frac{a_1+b_1}{2})$ on détermine l'intervalle $[a_2, b_2]$ tel que $b_2 - a_2 = \frac{b_1-a_1}{2} = \frac{b_0-a_0}{4}$ avec $a_0 \leqslant a_1 \leqslant a_2$, $b_0 \geqslant b_1 \geqslant b_2$, $g(a_2) < 0$ et $g(b_2) > 0$, puis on détermine $[a_3, b_3]$, $[a_4, b_4]$, etc. $[a_n, b_n]$ tels que $b_n - a_n = \frac{b_0-a_0}{2^n}$ avec $a_0 \leqslant a_1 \leqslant \cdots \leqslant a_n$ et $b_0 \geqslant b_1 \geqslant \cdots \geqslant b_n$, autrement dit (a_n) et (b_n) sont deux suites adjacentes qui convergent vers une même limite car (a_n) est croissante, (b_n) décroissante et, comme $\lim\limits_{n\to+\infty} b_n - a_n = \lim\limits_{n\to+\infty} \frac{b_0-a_0}{2^n} = 0$ on en déduit que $\lim\limits_{n\to+\infty} a_n = \lim\limits_{n\to+\infty} b_n$.
Pour tout n, $g(a_n) < 0$ et $g(b_n) > 0$, cette limite commune notée α vérifie $g(\alpha) = 0 \Longleftrightarrow f(\alpha) = \lambda$.

Remarques :

- La démonstration de ce théorème est constructive : elle fournit un procédé de calcul – appelé « algorithme de dichotomie » – qui détermine une valeur approchée d'un antécédent de l'équation $f(x) = \lambda$ à la précision souhaitée. Le programme Python ci-dessous traduit cet algorithme et l'applique à la fonction $f : x \longmapsto -x^3 - 2x^2 + 3x + 5$ représentée sur l'illustration pour $\lambda = 4$ et $[a,b] = [-3,2;1,32]$ à la précision 10^{-9}.
- Dans les conditions de cette proposition, la solution de l'équation $f(x) = \lambda$ n'est pas forcément unique. La figure ci-dessous montre un cas où il y a 3 solutions dans l'intervalle $[a,b]$. On constate, en observant ce graphique que l'algorithme de dichotomie aboutira à la solution la plus petite (environ $-2,9$) car dans cette situation, comme $f(a) > \lambda$ et $f(\frac{a+b}{2}) < \lambda$, on aura $a < \alpha < \frac{a+b}{2}$ (les deux autres solutions sont supérieures à $\frac{a+b}{2}$).



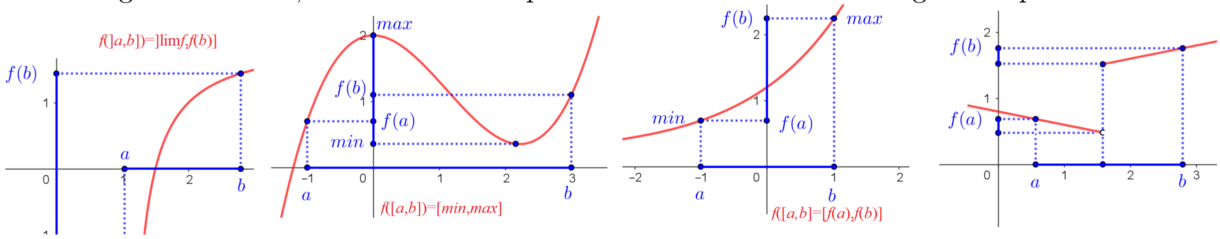

```
def f(x):
    return -x**3-2*x**2+3*x+5

def dichomie(a,b,k,p):
    while b-a>p :
        c=(a+b)/2
        if (f(a)-k)*(f(c)-k)<=0:
            b=c
        else:
            a=c
    return (a,b)

print(dichomie(-3.2,1.32,4,10**-9))
```

```
(-2.9122291789203896, -2.9122291783941923)
```

PROPRIÉTÉ 2.10 (INTERVALLE IMAGE) Soit f une fonction continue sur un intervalle I.

- $f(I)$ – l'ensemble image de I – est un intervalle.
- Si I est un intervalle fermé alors $f(I)$ est l'intervalle fermé $[\min, \max]$ où min et max sont le minimum et le maximum de f sur I.
- Si $I = [a, b]$ est un intervalle fermé et si f est monotone alors $f(I)$ est fermé. ($f(I) = [f(a), f(b)]$ si f croissante ou $[f(b), f(a)]$ si f décroissante)

DÉMONSTRATION Cette propriété est la conséquence directe du TVI ; en voici des illustrations. Dans la figure de droite, la fonction n'est pas continue et l'ensemble image n'est pas un intervalle.



THÉORÈME 2.8 (BIJECTION) Soit f une fonction continue et strictement monotone sur un intervalle $[a, b]$. Pour tout réel λ entre $f(a)$ et $f(b)$ il existe un unique réel $\alpha \in [a, b]$ tel que $f(\alpha) = \lambda$.

DÉMONSTRATION D'après le TVI, comme f est continue, il existe au moins une valeur $\alpha \in [a, b]$ telle que $f(\alpha) = \lambda$. Pour montrer que cette valeur est unique, supposons que f est strictement croissante :

- $\forall x \in [a, b],\ \ x < \alpha \Longrightarrow f(x) < f(\alpha) = \lambda$; l'équation $f(x) = \lambda$ n'a pas de solution sur $[a, \alpha[$.
- $\forall x \in [a, b],\ \ x > \alpha \Longrightarrow f(x) > f(\alpha) = \lambda$; l'équation $f(x) = \lambda$ n'a pas de solution sur $]\alpha, b]$.

Le même raisonnement conduit à l'unicité dans le cas où f est strictement décroissante.

Remarques :

- Si f est strictement croissante (respect. décroissante) alors $\forall \lambda \in [f(a), f(b)]$ (resp. $[f(b), f(a)]$), l'équation $f(x) = \lambda$ admet une solution unique dans $[a, b]$.
  Pour prouver l'existence et l'unicité d'une solution à une équation $f(x) = \lambda$ dans un intervalle $[a, b]$, il faut donc :
  — Prouver la continuité de f sur $[a, b]$
  — Prouver la stricte monotonie de f sur $[a, b]$ (notamment par l'étude du signe de f')
  — Calculer $f(a)$ et $f(b)$ puis montrer que $f(a) \leqslant \lambda \leqslant f(b)$ ou $f(b) \leqslant \lambda \leqslant f(a)$
  — Conclure en rappelant ou en citant le théorème de la bijection.
- On appelle cette propriété, le *théorème de la bijection* car f réalise alors une bijection de I (l'ensemble de départ) sur $J = f(I)$ (l'ensemble d'arrivée). Pour qu'une fonction soit une bijection de I sur J il faut et il suffit que tout élément de J ait une unique antécédent dans I, ce qu'assure ce théorème. Noter qu'on peut décliner cette propriété sur un intervalle ouvert

$]a,b[$ ou semi-ouvert $]a,b]$ ou $[a,b[$. Exemples de bijections (toutes les fonctions strictement monotones sur un intervalle) : $x \longmapsto \frac{1}{x}$ sur $]0,+\infty[$, $x \longmapsto \sqrt{x}$ sur $[0,+\infty[$, $x \longmapsto e^x$ sur $]-\infty,+\infty[$, $x \longmapsto x^2$ sur $[-\infty,0]$, $x \longmapsto \sin x$ sur $[\frac{-\pi}{2},\frac{\pi}{2}]$, etc.

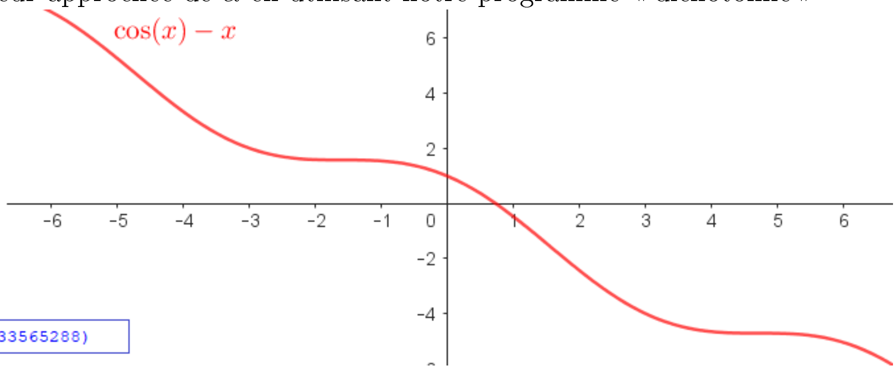
Exemple 20 – Montrons que l'équation $\cos x = x$ a une solution unique dans $[0,\pi]$.
Pour cela définissons la fonction $f : x \longmapsto \cos x - x$.
La question revient à prouver que l'équation $f(x) = 0$ a une unique solution dans $[0,\pi]$.

- f est continue sur $[0,\pi]$ comme somme de fonctions continues sur $[0,\pi]$
- $f'(x) = \sin x - 1$ or $\forall x \in [0,\pi],\ \ \sin x \leqslant 1 \Longleftrightarrow \sin x - 1 \leqslant 0$ avec égalité seulement pour $x = \frac{\pi}{2}$ donc f est strictement décroissante sur $[0,\pi]$.
- $f(0) = 1 - 0 = 1$ et $f(\pi) = -1 - \pi \approx -4,14 < 0$; on en déduit $f(\pi) \leqslant 0 \leqslant f(0)$
- Comme f est une fonction continue et strictement décroissante sur l'intervalle $[0,\pi]$, elle définit une bijection entre $[0,\pi]$ et $[f(\pi), f(0)]$ et comme, de plus, $0 \in [f(\pi), f(0)]$, il existe un unique réel $\alpha \in [0,\pi]$ tel que $f(\alpha) = 0$.

On peut déterminer une valeur approchée de α en utilisant notre programme « dichotomie »

```
from math import cos,pi
def f(x):
    return cos(x)-x

def dichomie(a,b,k,p):
    while b-a>p :
        c=(a+b)/2
        if (f(a)-k)*(f(c)-k)<=0:
            b=c
        else:
            a=c
    return (a,b)

print(dichomie(0,pi,0,10**-9))
```

```
(0.739085132833829, 0.739085133565288)
```

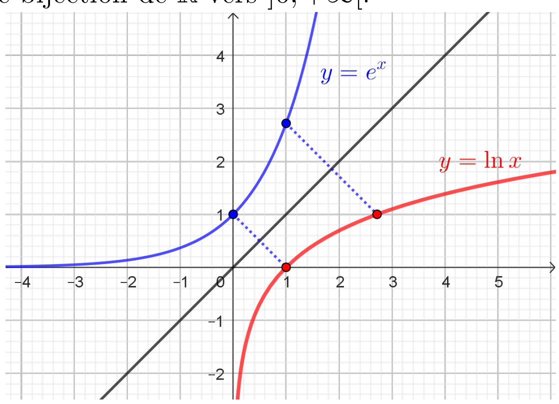


Propriété 2.11 (Réciproque) Soient $\mathscr{D}$ et $\mathscr{D}'$ deux parties de $\mathbb{R}$ et f une fonction définie sur $\mathscr{D}$ à valeurs dans $\mathscr{D}'$. Les deux propositions suivantes sont équivalentes :

- f est une bijection de $\mathscr{D}$ vers $\mathscr{D}'$
- f admet une bijection réciproque, notée f^{-1}, définie sur $\mathscr{D}'$ à valeurs dans $\mathscr{D}$, c'est-à-dire que $\forall y \in \mathscr{D}',\ \ \exists! x \in \mathscr{D},\ \ y = f(x)$; cet antécédent unique de y est $x = f^{-1}(y)$.

Remarques :

- Il est parfois simple de déterminer l'expression de la fonction réciproque.
  — $x \longmapsto \frac{1}{x}$ de $]0,+\infty[$ vers $]0,+\infty[$ est sa propre réciproque car $y = \frac{1}{x} \Longleftrightarrow x = \frac{1}{y}$
  — $x \longmapsto \sqrt{x}$ de $[0,+\infty[$ vers $[0,+\infty[$ a pour réciproque $x \longmapsto x^2$ car $y = \sqrt{x} \Longrightarrow x = y^2$
  — $x \longmapsto x^2$ définie sur $\mathbb{R}^+$ a pour réciproque $x \longmapsto \sqrt{x}$ mais $x \longmapsto -\sqrt{x}$ définie sur $\mathbb{R}^-$
  — $x \longmapsto \sin x$ de $[\frac{-\pi}{2},\frac{\pi}{2}]$ vers $[-1,1]$ a pour réciproque $x \longmapsto \sin^{-1} x$, aussi noté asin$x$
  — $x \longmapsto ax + b$ de $\mathbb{R}$ vers $\mathbb{R}$ a pour réciproque $x \longmapsto \frac{x-b}{a}$ car $y = ax + b \Longleftrightarrow x = \frac{y-b}{a}$
- La fonction exponentielle $x \longmapsto e^x$ définie une bijection de $\mathbb{R}$ vers $]0,+\infty[$.
  Elle admet donc une bijection réciproque de $]0,+\infty[$ vers $\mathbb{R}$: cette nouvelle fonction est appelée « logarithme népérien » et notée ln et on a $y = e^x \Longleftrightarrow x = \ln y$.
  La courbe représentative de la fonction ln est symétrique de celle de la fonction exp. Le point de coordonnées $(0,1)$ étant sur $\mathscr{C}_{\exp}$, $(1,0)$ est sur $\mathscr{C}_{\ln}$ (ce qui signifie $\ln 1 = 0$). De même, le point $(1,e)$ étant sur $\mathscr{C}_{\exp}$, $(e,1)$ est sur $\mathscr{C}_{\ln}$ (ce qui signifie $\ln e = 1$).



3. Dérivation

3.a. Dérivabilité

3.a.1. Nombre dérivé (rappel)

DÉFINITION 2.7 (NOMBRE DÉRIVÉ) Soit f une fonction définie sur un intervalle I contenant le réel x_0. On dit que f est dérivable en x_0 si le taux d'accroissement de la fonction au voisinage de x_0 admet une limite finie. Cette limite est appelée *nombre dérivé* de f en x_0 et notée $f'(x_0)$.

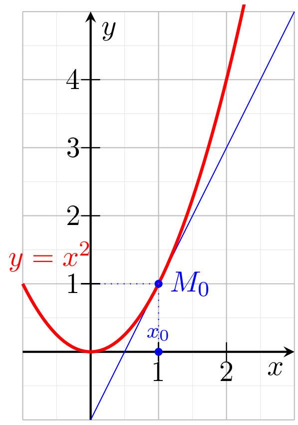
$$f'(x_0) = \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} = \lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h}$$

Exemple :

La fonction $f : x \mapsto x^2$, est dérivable en x_0 et on a $f'(x_0) = 2x_0$ car $\lim\limits_{x \to x_0} \frac{f(x)-f(x_0)}{x-x_0} = \lim\limits_{h \to 0} \frac{(x_0+h)^2-(x_0)^2}{h} = \lim\limits_{h \to 0} \frac{2x_0h+h^2}{h} = \lim\limits_{h \to 0} 2x_0 + h = 2x_0$

$2x_0$ est la pente de la tangente à la courbe $\mathscr{C}_f$ au point $M_0(x_0, f(x_0))$.
Cette tangente est une droite d'équation :
$y = f'(x_0)(x - x_0) + f(x_0) = 2x_0(x - x_0) + x_0^2 = 2x_0x - x_0^2$
Cette équation fournit l'approximation affine $x^2 \approx 2x_0x - x_0^2$

Pour $x_0 = 1$, cette tangente a pour équation $y = 2x - 1$
et pour x proche de 1, x^2 est proche de $2x - 1$
($1,1^2 = 1,21$ et $2 \times 1,1 - 1 = 1,2$).



PROPRIÉTÉ 2.12 (INTERPRÉTATIONS) Soit f une fonction dérivable en $x_0 \in I$, un intervalle.
$\boxed{\text{Graphique}}$ La tangente en x_0 à la courbe $\mathscr{C}_f$ a pour pente $f'(x_0)$
l'équation de cette tangente est $y = f'(x_0)(x - x_0) + f(x_0)$
$\boxed{\text{Cinématique}}$ $f(x_0)$ étant la distance parcourue en x_0 secondes par un objet en mouvement, la *vitesse instantanée* de cet objet à cet instant est $v(x_0) = f'(x_0)$
$\boxed{\text{Numérique}}$ $\varphi(h) = \frac{f(x_0+h)-f(x_0)}{h} - f'(x_0)$ est une expression de limite nulle en 0.
Ainsi $f(x_0 + h) = f(x_0) + hf'(x_0) + h\varphi(h)$ est le *développement limité à l'ordre 1* de f en x_0, $f(x_0) + hf'(x_0)$ est une valeur approchée de $f(x_0 + h)$ dont $|h\varphi(h)|$ en est l'erreur commise.

Remarques :

- Si $f'(x_0) = 0$ alors la tangente en x_0 est horizontale. Dans ce cas et dans ce cas seulement, la fonction peut admettre un optimum (maximum ou minimum) : il suffit qu'il y ait un changement du signe de f' en x_0 (voir plus loin).
- Quand on lâche un objet sur Terre, la distance parcourue dans sa chute en t secondes est $f(t) = \frac{1}{2}gt^2$ m, où $g \approx 9,81 \text{m.s}^{-2}$ est l'accélération de la pesanteur. La vitesse instantanée de cet objet après une chute de t secondes est $\lim\limits_{h \to 0} \frac{f(t+h)-f(t)}{h} = \lim\limits_{h \to 0} gt + \frac{gh}{2}$ d'où $v(t) = f'(t) = gt$
- La reconnaissance du nombre dérivé de f en x_0 peut s'opérer directement : en reconnaissant dans le développement de $f(x_0 + h)$, la forme $f(x_0) + h\alpha + h\varphi(h)$ où φ est une fonction telle que $\lim\limits_{h \to 0} \varphi(h) = 0$, on a alors $f'(x_0) = \alpha$.
  Pour la fonction cube $f : x \mapsto x^3$, on a $f(x_0 + h) = (x_0 + h)^3 = x_0^3 + 3hx_0^2 + 3h^2x_0 + h^3$.
  On reconnait la forme $f(x_0) + h\alpha$ dans $x_0^3 + 3hx_0^2$ avec $\alpha = 3x_0^2$ et $h\varphi(h) = 3h^2x_0 + h^3$ car $\lim\limits_{h \to 0} \varphi(h) = \lim\limits_{h \to 0} 3hx_0 + h^2 = 0$. Le nombre dérivé de f en x_0 est donc $\alpha = f'(x_0) = 3x_0^2$.

3.a.2. Fonction dérivée

DÉFINITION 2.8 (DÉRIVABILITÉ) Pour qu'une fonction f définie en $x_0 \in I$ soit dérivable en x_0
⇛ si x_0 n'est pas une borne de I, il faut et il suffit qu'elle soit :
- Dérivable à droite : f est définie pour $x > x_0$ et $\lim\limits_{\substack{x \to x_0 \\ x > x_0}} \frac{f(x)-f(x_0)}{x-x_0} = f'_d(x_0)$ est un réel fini.
- Dérivable à gauche : f est définie pour $x < x_0$ et $\lim\limits_{\substack{x \to x_0 \\ x < x_0}} \frac{f(x)-f(x_0)}{x-x_0} = f'_g(x_0)$ est un réel fini.
- La dérivée à gauche $f'_g(x_0)$ et la dérivée à droite $f'_d(x_0)$ doivent être égales.

⇛ si x_0 est la borne inférieure (respect. supérieure) de I, il faut et il suffit qu'elle soit dérivable à droite (resp. à gauche) c'est-à-dire $f'_d(x_0)$ (resp. $f'_g(x_0)$) est fini.

Remarques :
- La fonction fonction valeur absolue $x \longmapsto |x|$ n'est pas dérivable en 0 car $f'_d(0) = \lim\limits_{\substack{x \to 0 \\ x > 0}} \frac{|x|}{x} = \lim\limits_{\substack{x \to 0 \\ x > 0}} \frac{x}{x} = 1$ alors que $f'_g(0) = \lim\limits_{\substack{x \to 0 \\ x < 0}} \frac{|x|}{x} = \lim\limits_{\substack{x \to 0 \\ x < 0}} \frac{-x}{x} = -1$ et donc $f'_d(0) \neq f'_g(0)$.
  La fonction racine carrée $x \longmapsto \sqrt{x}$ n'est pas dérivable en 0 car $\lim\limits_{\substack{x \to 0 \\ x > 0}} \frac{\sqrt{x}}{x} = \lim\limits_{\substack{x \to 0 \\ x > 0}} \frac{1}{\sqrt{x}} = +\infty$
- Si x_0 est la borne inférieure (respect. supérieure) de I et si $f'_d(x_0)$ (resp. $f'_g(x_0)$) existe, alors on limite la tangente à la demi-droite d'origine $(x_0, f(x_0)$ dont les points ont une abscisse $x > x_0$ (resp. $x < x_0$) ; on parle alors de demi-tangente à droite (resp. à gauche).
  Ainsi, la fonction valeur absolue admet une demi-tangente verticale.

PROPRIÉTÉ 2.13 (CONTINUITÉ) Soit f une fonction définie sur un intervalle I.
Si f est dérivable en $x_0 \in I$ alors f est continue en x_0 et si f est dérivable en tous points de I alors f est continue sur I.

DÉMONSTRATION Si f est dérivable en x_0, son développement limité d'ordre 1 en x_0 est $f(x_0 + h) = f(x_0) + hf'(x_0) + h\varphi(h)$ avec $\lim\limits_{h \to 0} \varphi(h) = 0$.
On en déduit $\lim\limits_{h \to 0} f(x_0 + h) = \lim\limits_{h \to 0} f(x_0) + hf'(x_0) + h\varphi(h) = f(x_0)$ soit f est continue en x_0.
Par extension, si f est dérivable en tous points de I alors f est continue en tous points de I.

DÉFINITION 2.9 (FONCTION DÉRIVÉE) Une fonction f est dérivable sur un intervalle I si elle l'est en tous points de I et la fonction $f' : x \longmapsto f'(x)$ – la *dérivée* de f – est définie sur I.

On a vu que la fonction cube $x \mapsto x^3$ est dérivable en tous points $x_0 \in \mathbb{R}$ et on a $f'(x_0) = 3x_0^2$.
Par conséquent la fonction $x \mapsto 3x^2$ est la dérivée de la fonction cube. Cette fonction est définie sur $\mathbb{R}$.

PROPRIÉTÉ 2.14 (DÉRIVÉES DES FONCTIONS USUELLES) Les fonctions suivantes sont dérivables sur tout ou partie de leur ensemble de définition.
- Fonction constante : si $f(x) = k$ alors $f'(x) = 0$
- Fonction puissance : si $\forall n \in \mathbb{N}^*, \;\; f(x) = x^n$ alors $f'(x) = nx^{n-1}$
- Fonction inverse : si $\forall x \neq 0, \;\; f(x) = \frac{1}{x}$ alors $\forall x \neq 0, \;\; f'(x) = -\frac{1}{x^2}$
- Fonction racine carrée : si $\forall x \geqslant 0, \;\; f(x) = \sqrt{x}$ alors $\forall x > 0, \;\; f'(x) = \frac{1}{2\sqrt{x}}$
- Fonctions trigonométriques : $\forall x \geqslant 0, \;\; \cos'(x) = -\sin(x)$ et $\sin'(x) = \cos(x)$.
  $\forall k \in \mathbb{Z}, \;\; \forall x \in]-\frac{\pi}{2} + k\pi; \frac{\pi}{2} + k\pi[, \;\; \tan'(x) = \frac{1}{\cos^2(x)} = 1 + \tan^2(x)$
- Fonction exponentielle : $\forall x \in \mathbb{R}, \;\; (e^x)' = e^x$

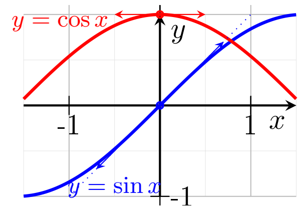
DÉMONSTRATION ➠ Pour $k \in \mathbb{R}$ et $f(x) = k$ on a $\frac{f(x+h)-f(x)}{h} = \frac{k-k}{h} = 0$ d'où $f'(x) = 0$.

➠ Pour $n \in \mathbb{N}^*$ et $f(x) = x^n$ on a $\frac{(x+h)^n - x^n}{h} = \frac{(x^n + nhx^{n-1} + h\varphi(h)) - x^n}{h} = nx^{n-1} + \varphi(h)$.
Comme dans tous les termes de $\varphi(h)$ il y a h, on a $\lim_{h\to 0} \varphi(h) = 0$ d'où $f'(x) = nx^{n-1}$.

➠ Pour $x \neq 0$ et $f(x) = \frac{1}{x}$ on a $\frac{\frac{1}{x+h} - \frac{1}{x}}{h} = \frac{\frac{x-(x+h)}{x(x+h)}}{h} = \frac{-h}{x(x+h)h} = -\frac{1}{x(x+h)}$.
Comme $\lim_{h\to 0} -\frac{1}{x(x+h)} = -\frac{1}{x^2}$ on a, pour tout $x \neq 0$, $f'(x) = -\frac{1}{x^2}$.

➠ Pour $x > 0$ et $f(x) = \sqrt{x}$ on a $\frac{\sqrt{x+h} - \sqrt{x}}{h} = \frac{\sqrt{x+h}^2 - \sqrt{x}^2}{h(\sqrt{x+h} + \sqrt{x})} = \frac{h}{h(\sqrt{x+h} + \sqrt{x})} = \frac{1}{\sqrt{x+h} + \sqrt{x}}$.
Comme $\lim_{h\to 0} \frac{1}{\sqrt{x+h} + \sqrt{x}} = \frac{1}{2\sqrt{x}}$ on a, pour tout $x > 0$, $f'(x) = \frac{1}{2\sqrt{x}}$.

➠ Comme on l'observe sur la courbe de la fonction sinus,
la tangente en O est la droite d'équation $y = x$, de pente 1.
En s'appuyant sur cette observation, posons que $\sin'(0) = 1$.
Ainsi $\sin(0 + h) = \sin(0) + \sin'(0)h + h\varphi(h)$,
soit $\sin(h) = h(1 + \varphi(h))$ avec $\lim_{h\to 0} \varphi(h) = 0$.
Utilisons alors la formule de duplication pour exprimer $\cos(h)$:
$\cos(h) = 1 - 2\sin^2(\frac{h}{2}) = 1 - 2\frac{h^2}{4}(1 + \varphi(\frac{h}{2}))^2 = 1 + h\psi(h)$ où $\lim_{h\to 0} \psi(h) = 0$



On en déduit que cos est dérivable en 0 et que $\cos'(0) = 0$ (la tangente est horizontale en $(0, 1)$)
En remplaçant $\sin(h)$ et $\cos(h)$ par leurs développements limités d'ordre 1 en 0 dans les formules d'addition ($\sin(x + h) = \sin(x)\cos(h) + \sin(h)\cos(x)$ et $\cos(x + h) = \cos(x)\cos(h) - \sin(h)\sin(x)$),
on obtient $\begin{cases} \sin(x+h) = \sin(x)(1 + h\psi(h)) + h(1 + \varphi(h))\cos(x) = \sin(x) + h\cos(x) + h\theta(h) \\ \cos(x+h) = \cos(x)(1 + h\psi(h)) - \sin(x)h(1 + \varphi(h)) = \cos(x) - h\sin(x) + h\rho(h) \end{cases}$
avec $\lim_{h\to 0} \theta(h) = 0$ et $\lim_{h\to 0} \rho(h) = 0$ d'où $\sin'(x) = \cos(x)$ et $\cos'(x) = -\sin(x)$.

Pour tan appliquons la formule de la dérivée d'un quotient (voir plus loin) :
Comme $\tan(x) = \frac{\sin(x)}{\cos(x)}$ on a $\tan'(x) = \frac{(\sin)'(\cos) - (\sin)(\cos)'}{\cos^2} = \frac{\cos^2(x) + \sin^2(x)}{\cos^2(x)} = \frac{1}{\cos^2(x)}$.
On a aussi $\tan'(x) = \frac{\cos^2(x) + \sin^2(x)}{\cos^2(x)} = \frac{\cos^2(x)}{\cos^2(x)} + \frac{\sin^2(x)}{\cos^2(x)} = 1 + \frac{\sin^2(x)}{\cos^2(x)} = 1 + \tan^2(x)$

➠ Pour la fonction exponentielle, la propriété traduit la définition de cette fonction.

PROPRIÉTÉ 2.15 (OPÉRATIONS (1)) Soient f et g deux fonctions dérivables sur un intervalle I.
Dérivée d'une somme : la fonction $f + g$ est dérivable sur I et $\forall x \in I$, $(f + g)'(x) = f'(x) + g'(x)$.
Produit par un réel : $\forall \lambda \in \mathbb{R}$ la fonction λf est dérivable sur I et $\forall x \in I$, $(\lambda f)'(x) = \lambda f'(x)$.

DÉMONSTRATION ➠ f et g étant dérivables en $x \in I$ on a $f(x + h) = f(x) + hf'(x) + h\varphi(h)$ et $g(x + h) = g(x) + hg'(x) + h\psi(h)$ avec $\lim_{h\to 0} \varphi(h) = 0$ et $\lim_{h\to 0} \psi(h) = 0$. Par conséquent
$(f + g)(x + h) = (f + g)(x) + h[f'(x) + g'(x)] + h\theta(h)$ où θ est définie par $\theta(h) = \varphi(h) + \psi(h)$.
Cette fonction est telle que $\lim_{h\to 0} \theta(h) = 0$, d'où la conclusion.

➠ f étant dérivable en $x \in I$ on a $f(x + h) = f(x) + hf'(x) + h\varphi(h)$ avec $\lim_{h\to 0} \varphi(h) = 0$.
Par conséquent $\lambda f(x + h) = \lambda f(x) + h[\lambda f'(x)] + h\psi(h)$ où ψ est définie par $\psi(h) = \lambda\varphi(h)$.
Cette fonction est telle que $\lim_{h\to 0} \psi(h) = 0$, d'où la conclusion.

Conséquence : Une fonction polynôme de degré n est dérivable sur $\mathbb{R}$:
Si $f : x \mapsto \sum_{k=0}^{n} a_k x^k$ avec $a_n \neq 0$ alors $f' : x \mapsto \sum_{k=1}^{n} a_k k x^{k-1} = \sum_{k=0}^{n-1} a'_k x^k$ où $a'_k = (k + 1)a_{k+1}$.
Le monôme de plus haut degré étant $a'_{n-1} = a_n x^{n-1}$, la dérivée de f est un polynôme de degré $n - 1$.
Par exemple, si $f(x) = 5x^4 - 6x^3 + 7x^2 - 8x + 9$ alors $f'(x) = 20x^3 - 18x^2 + 14x - 8$.

Propriété 2.16 (Opérations (2)) Soient f et g deux fonctions dérivables sur un intervalle I.
Dérivée d'un produit : la fonction $h = f \times g$ est dérivable sur I et $h' = f' \times g + f \times g'$
Dérivée d'un quotient : la fonction $h = \frac{f}{g}$ est dérivable sur $\{x \in I, g(x) \neq 0\}$ et $h' = \frac{f' \times g - f \times g'}{g^2}$

Démonstration ➠ f et g étant dérivables en $x \in I$ on a $f(x+h) = f(x) + hf'(x) + h\varphi(h)$ et $g(x+h) = g(x) + hg'(x) + h\psi(h)$ avec $\lim\limits_{h\to 0} \varphi(h) = 0$ et $\lim\limits_{h\to 0} \psi(h) = 0$. Par conséquent
$(fg)(x+h) = (fg)(x) + h[f'(x) \times g(x) + g'(x) \times f(x)] + h\theta(h)$ où θ est la fonction définie par
$\theta(h) = f(x)\psi(h) + g(x)\varphi(h) + h[f'(x)g'(x) + f'(x)\psi(h) + g'(x)\varphi(h) + \psi(h)\varphi(h)]$.
x étant fixé, c'est h qui tend vers 0. Cette fonction est telle que $\lim\limits_{h\to 0} \theta(h) = 0$, d'où la conclusion.

➠ Déterminons tout d'abord la dérivée de la fonction i définie par $i(x) = \frac{1}{g(x)}$.
Pour un réel x_0 de I , le taux de variation de i entre $x_0 + h$ et x_0 est :

$$\tau = \frac{i(x_0+h) - i(x_0)}{h} = \frac{\frac{1}{g(x_0+h)} - \frac{1}{g(x_0)}}{h} = \frac{\frac{g(x_0)-g(x_0+h)}{g(x_0+h)g(x_0)}}{h} = \frac{g(x_0+h) - g(x_0)}{h} \times \frac{-1}{g(x_0+h)g(x_0)}$$

Comme $\lim\limits_{h\to 0} \frac{g(x_0+h)-g(x_0)}{h} = g'(x_0)$ et $\lim\limits_{h\to 0} \frac{-1}{g(x_0+h)g(x_0)} = \frac{-1}{g^2(x_0)}$, on en déduit que $\lim\limits_{h\to 0} \tau = \frac{-g'(x_0)}{g^2(x_0)}$.
La fonction i a pour dérivée la fonction définie par $i'(x) = \frac{-g'(x)}{g^2(x)}$.
Il ne reste plus qu'à appliquer la propriété précédente à la fonction $h = \frac{f}{g} = f \times \frac{1}{g} = fi$:
$h' = fi' + f'i$; pour un réel $x \in I$ cela s'écrit $h'(x) = \dfrac{-f(x)g'(x)}{g^2(x)} + \dfrac{f'(x)}{g(x)} = \dfrac{f'(x)g(x) - f(x)g'(x)}{g^2(x)}$

Remarques et exemples :

- Soit f une fonction définie et dérivable sur un intervalle I.
  La fonction $g = f^2$ est dérivable sur I et, par application de la dérivée d'un produit, $g' = 2ff'$.
  Par exemple, la dérivée de $g : x \mapsto (5x^2 - 3x + 1)^2$ est définie par $g' = 2ff'$ où
  $f(x) = 5x^2 - 3x + 1$ et $f'(x) = 10x - 3$.
  On a donc $g'(x) = 2(5x^2 - 3x + 1)(10x - 3)$ (NB : ne pas développer le carré pour dériver f !)
- Les fonctions rationnelles sont dérivables sur leur ensemble de définition.
  Par exemple, $f : x \mapsto \frac{3x-2}{x^2-1}$ est définie et dérivable sur $\mathbb{R} \setminus \{-1; 1\}$.
  Comme $f = \frac{u}{v}$ avec $u(x) = 3x - 2$, $v(x) = x^2 - 1$, $u'(x) = 3$ et $v'(x) = 2x$, par application de la dérivée d'un quotient, $f' = \frac{u'v-uv'}{v^2}$. On a donc $f'(x) = \frac{3(x^2-1)-2x(3x-2)}{(x^2-1)^2} = \frac{-3x^2+4x-3}{(x^2-1)^2}$

Propriété 2.17 (Dérivée d'une composée) Soit f une fonction dérivable sur un intervalle I et soit g une fonction dérivable sur J. $g \circ f$ est dérivable sur $\{x \in I, f(x) \in J\}$ et $(g \circ f)' = f' \times (g' \circ f)$

Démonstration f étant dérivable en $x \in I$ on a $f(x+h) = f(x) + hf'(x) + h\varphi(h)$ avec $\lim\limits_{h\to 0} \varphi(h) = 0$ et g étant dérivable en $f(x) \in f(I)$ on a $g(f(x)+k) = g(f(x)) + kg'(f(x)) + k\psi(k)$ avec $\lim\limits_{k\to 0} \psi(k) = 0$. Prenons $k = hf'(x) + h\varphi(h)$; lorsque $h \longrightarrow 0$, on a bien $k \longrightarrow 0$.
Avec cette expression de k le développement limité de $g(f(x)+k)$ devient celui de $g(f(x+h))$:
$g(f(x+h)) = g(f(x)) + (hf'(x) + h\varphi(h))\,(g'(f(x)) + \psi(k)) = g(f(x)) + hf'(x)g'(f(x)) + h\theta(h)$
L'expression de la fonction θ est assez compliquée, mais il est clair que cette fonction tend vers 0 quand $h \longrightarrow 0$. En effet $\theta(h) = f'(x)\psi(k) + \varphi(h)g'(f(x)) + \varphi(h)\psi(k)$.
Du fait que $\lim\limits_{h\to 0} \psi(k) = \lim\limits_{k\to 0} \psi(k) = 0$ (je n'écris pas l'expression $\psi(k)$ en fonction de h pour ne pas surcharger) et que $\lim\limits_{h\to 0} \varphi(h) = 0$, on a bien $\lim\limits_{h\to 0} \theta(h) = 0$, d'où le résultat.

Applications : on a déjà dérivé f^2 (composée de f suivie de la fonction carrée) $(f^2)' = 2f'f$
et aussi $\frac{1}{f}$ (composée de f suivie de la fonction inverse) $\left(\frac{1}{f}\right)' = \frac{-f'}{f^2}$ (on doit avoir $f > 0$).
Ces formules peuvent être obtenues avec cette nouvelle propriété, ainsi que beaucoup d'autres.

- La dérivée de $\sqrt{f}$ est $\left(\sqrt{f}\right)' = \frac{f'}{2\sqrt{f}}$ (on doit avoir $f > 0$)
  Par exemple $\left(\sqrt{x^2 - x + 1}\right)' = \frac{2x-1}{2\sqrt{x^2-x+1}}$
- La dérivée de f^n est $(f^n)' = nf'f^{n-1}$
  Par exemple $\left((x^3-1)^3\right)' = 9x^2(x^3-1)^2$
- La dérivée de $\sin(f)$ est $f'\cos(f)$ et celle de $\cos(f)$ est $-f'\sin(f)$
  Par exemple $(\sin(2x-1))' = 2\cos(2x-1)$
- La dérivée de e^f est $f'e^f$. Par exemple $\left(e^{-x^2+1}\right)' = -2xe^{-x^2+1}$

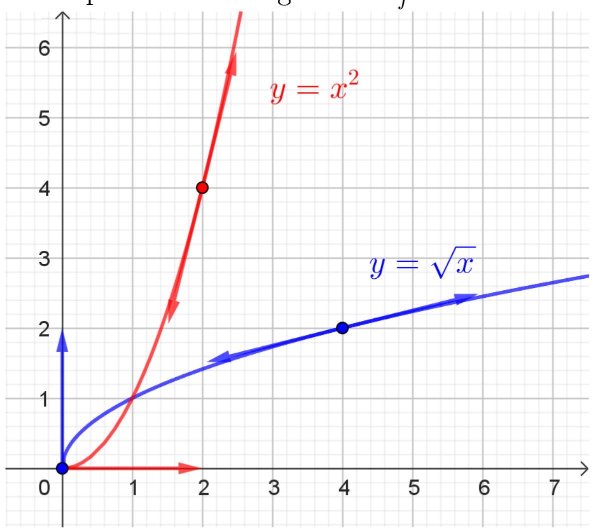
On peut dériver une fonction composée plus complexe en procédant par étapes :
La fonction $x \longmapsto \frac{1}{f^n}$ est la composée de f suivie de $g : x \longmapsto x^n$ suivie de $h : x \longmapsto \frac{1}{x}$.
Ainsi, on peut dériver $(h \circ (g \circ f))' = (g \circ f)'(h' \circ (g \circ f)) = f'(g' \circ f)(h' \circ (g \circ f))$.
En ce qui concerne notre exemple $\left(\frac{1}{f^n}\right)' = f'(nf^{n-1})(\frac{-1}{f^{2n}}) = \dfrac{-nf'f^{n-1}}{f^{2n}} = \dfrac{-nf'}{f^{n+1}}$

Propriété 2.18 (Dérivée de la réciproque) Soit f une bijection de I vers $J = f(I)$.
Si f est dérivable sur alors f^{-1} est dérivable sur $\{y \in J, f' \circ f^{-1}(y) \neq 0\}$ et $(f^{-1})' = \frac{1}{f' \circ f^{-1}}$

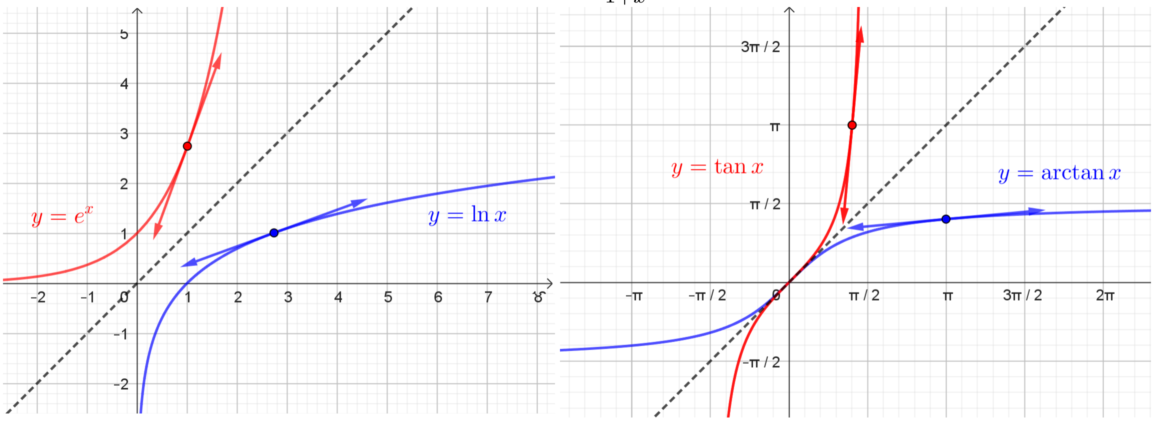
Démonstration La réciproque d'une bijection f définie sur I est la fonction f^{-1} définie sur $J = f(I)$ par $f \circ f^{-1} = Id$, c'est-à-dire $\forall y \in J,\ \ f \circ f^{-1}(y) = f(f^{-1}(y)) = f(x) = y$.
Dérivons $f \circ f^{-1}$ avec la propriété précédente, on obtient $\left(f \circ f^{-1}\right)' = (f^{-1})' \times (f \circ f^{-1})$.
Comme $f \circ f^{-1} = Id$, par unicité de la dérivée d'une fonction $\left(f \circ f^{-1}\right)' = (Id)' = 1$.
On en déduit $(f^{-1})' \times (f \circ f^{-1}) = 1$ et pour les $y \in J,\ \ f \circ f^{-1} \neq 0$ on a $(f^{-1})' = \frac{1}{f \circ f^{-1}}$.

Remarques :

- Cette propriété peut également se justifier en cherchant la limite de $\frac{f^{-1}(y)-f^{-1}(y_0)}{y-y_0} = \frac{1}{\frac{y-y_0}{f^{-1}(y)-f^{-1}(y_0)}} = \frac{1}{\frac{f(x)-f(x_0)}{x-x_0}}$ lorsque y tend vers y_0, c'est-à-dire lorsque x tend vers x_0 (par composition des limites) d'où $\left(f^{-1}\right)'(y) = \frac{1}{f'(x)} = \frac{1}{f'(f^{-1}(y))}$
- La dérivée de f^{-1} en $y = f(x)$ est donc l'inverse de la dérivée de f en x.
  Cette propriété signifie que les tangentes en $(x, f(x) = y)$ à la courbe de f et en $(y = f(x), f^{-1}(y) = x)$ à la courbe de f^{-1} ont des pentes inverses l'une de l'autre.
  Prenons $f : x \longmapsto \sqrt{x}$ et $f^{-1} : x \longmapsto x^2$, fonctions de $\mathbb{R}^+$ dans $\mathbb{R}^+$. Si la tangente en $(x = 4, y = 2)$ à $\mathscr{C}_f$ a pour pente $\frac{1}{2\sqrt{4}} = \frac{1}{4}$ alors la tangente en $(x = 2, y = 4)$ à $\mathscr{C}_{f^{-1}}$ a pour pente $\frac{1}{\frac{1}{4}} = 4$ et on sait que $(f^{-1})'(2) = 2 \times 2 = 4$.
  La pente de la tangente à $\mathscr{C}_f$ en 0 est verticale (a pour limite $+\infty$: f n'est pas dérivable en 0) car la pente de la tangente à $\mathscr{C}_{f^{-1}}$ en 0 est horizontale.



- Comme ln est la réciproque de exp, c'est-à-dire $x = \ln y \iff y = \exp x$, dérivons ln :
  $\ln'(y) = \frac{1}{\exp'(x)} = \frac{1}{\exp'(\ln(y))} = \frac{1}{\exp(\ln(y))} = \frac{1}{\exp\circ\ln(y)} = \frac{1}{y}$
  La dérivée de ln est la fonction inverse sur $]0, +\infty[$
- On a vu que tan est dérivable sur $I =]\frac{-\pi}{2}, \frac{\pi}{2}[$ et que sa dérivée est $\tan' = 1 + \tan^2$.
  Comme $1 + \tan^2 > 0$, tan est strictement monotone sur I.
  De plus, $\lim\limits_{x\to\frac{-\pi}{2}} \tan x = -\infty$ et $\lim\limits_{x\to\frac{\pi}{2}} \tan x = +\infty$.
  On en a déduit que tan définit une bijection de I vers $\mathbb{R}$.
  Sa fonction réciproque $\tan^{-1}$ est dérivable sur $\mathbb{R}$ et on a :
  $\left(\tan^{-1}\right)'(y) = \frac{1}{\tan'(x)} = \frac{1}{\tan'(\tan^{-1}(y))} = \frac{1}{1+(\tan(\tan^{-1}(y)))^2} = \frac{1}{1+(\tan\circ\tan^{-1}(y)))^2} = \frac{1}{1+y^2}$
  La dérivée de $\tan^{-1}$ est donc la fonction $x \longmapsto \frac{1}{1+x^2}$



Définition 2.10 (Dérivée seconde) Soit f une fonction dérivable sur I telle que sa dérivée f' soit dérivable sur I. En notant $f'' = (f')'$ la dérivée de f', f'' est la *dérivée seconde* de f.

Remarques :
- On définit de même $f^{(3)}$ la dérivée 3e de f (si f'' est dérivable alors $f^{(3)} = (f'')'$) et si la fonction f est dérivable n fois, on note $f^{(n)}$ la dérivée ne de f. Avec cette notation, on a $f^{(1)} = f'$ et $f^{(0)} = f$.
- Une autre notation est utilisée, en physique notamment : $f' = \frac{\mathrm{d}f}{\mathrm{d}x}$ (dérivée de f par rapport à la variable x), $f'' = \frac{\mathrm{d}^2 f}{\mathrm{d}x^2}$ (on dérive f deux fois par rapport à la variable x), etc.
- Un polynôme, comme $P(x) = 2x^3 - x^2 + 5x - 1$ par exemple, est dérivable autant de fois qu'on le souhaite : $P'(x) = 6x^2 - 2x + 5$, $P''(x) = 12x - 2$, $P^{(3)}(x) = 12$, $P^{(4)}(x) = 0$, etc.

4. Sens de variation d'une fonction

4.a. TAF

Théorème 2.9 (Rolle) Soit f une fonction continue sur $[a, b]$ et dérivable sur $]a, b[$ telle que $f(a) = f(b)$. Dans ces conditions, il existe au moins un réel $c \in]a, b[$, $f'(c) = 0$

Démonstration Si f est constante sur $[a, b]$ alors tout point $c \in [a, b]$ vérifie $f'(c) = 0$.
On suppose désormais f non constante.
Il existe donc au moins un réel $x_0 \in [a, b]$ tel que $f(x_0) \neq f(a)$.
D'après la propriété 2.10 l'ensemble image de $[a, b]$ est $[\min, \max]$ où min et max sont le minimum et le maximum de f sur $[a, b]$.
Par conséquent, il existe bien un maximum et un minimum (l'un éventuellement situé en a et b), c'est-à-dire un point pour lequel la dérivée s'annule.
Si $f(x_0) > f(a)$ il existe au moins un maximum ; si $f(x_0) < f(a)$ il existe au moins un minimum.

Théorème 2.10 (Accroissements Finis) Soit f une fonction continue sur $[a,b]$ et dérivable sur $]a,b[$. Dans ces conditions, il existe au moins un réel $c \in]a,b[$, $f'(c) = \dfrac{f(b)-f(a)}{b-a}$

Démonstration Si g la fonction définie par $g(x) = f(x) - (x-a)\frac{f(b)-f(a)}{b-a}$
- $g(a) = f(a) - (a-a)\frac{f(b)-f(a)}{b-a} = f(a)$
- $g(b) = f(b) - (b-a)\frac{f(b)-f(a)}{b-a} = f(b) - (f(b)-f(a)) = f(a)$

Ainsi g est une fonction continue sur $[a,b]$ et dérivable sur $]a,b[$ (par addition de fonctions continues et dérivables) telle que $g(a) = g(b)$. D'après le théorème de Rolle, $\exists c \in]a,b[$, $g'(c) = 0$.
La fonction g' est définie par $g'(x) = f'(x) - \frac{f(b)-f(a)}{b-a}$.
Ainsi $\exists c \in]a,b[$, $f'(c) - \frac{f(b)-f(a)}{b-a} = 0 \Longleftrightarrow f'(c) = \frac{f(b)-f(a)}{b-a}$.

Remarque : Si on connait un encadrement $m \leqslant f'(x) \leqslant M$ pour tout x de $]a,b[$ on en déduit un encadrement du taux d'accroissement $m \leqslant \dfrac{f(b)-f(a)}{b-a} \leqslant M$.

4.b. Etude des fonctions

Théorème 2.11 Soit f une fonction continue sur un intervalle I et dérivable sur l'intervalle I éventuellement privé de ses bornes :
- si pour $x \in I$ on a $f'(x) > 0$ sauf éventuellement en un nombre fini de points où elle s'annule, alors f est strictement croissante sur I
- si pour $x \in I$ on a $f'(x) = 0$, alors f est constante sur I
- si pour $x \in I$ on a $f'(x) < 0$ sauf éventuellement en un nombre fini de points où elle s'annule, alors f est strictement décroissante sur I

Démonstration La réciproque de cette propriété a été démontrée dans le cours de 1re.
➠ Supposons que $f'(x) \geqslant 0$ sur I.
Pour tout couple $(a, b > a)$ de réels de I, la fonction est continue sur $[a,b]$ et dérivable sur $]a,b[$, d'après le TAF, il existe donc un réel $c \in]a,b[$ tel que $f'(c) = \frac{f(b)-f(a)}{b-a}$.
Comme $f'(c) \geqslant 0 \Longleftrightarrow \frac{f(b)-f(a)}{b-a} \geqslant 0$ et $b > 0$, on en déduit $f(b) - f(a) \geqslant 0 \Longleftrightarrow f(b) \geqslant f(a)$.
Ceci étant vrai pour tout couple de réels de I, on en déduit que la fonction f est croissante sur I.

➠ Montrons, par l'absurde, que si $f'(x) \geqslant 0$ avec $f'(x) = 0$ pour seulement un nombre fini de points alors f est strictement croissante sur I.
Supposons que f est croissante sans être strictement croissante sur I.
Il existe alors nécessairement deux réels distincts x_0 et $x_1 > x_0$ tels que $f(x_0) = f(x_1)$ (sinon f serait strictement croissante).
Pour tout réel $x \in [x_0, x_1]$, comme f est croissante on a $f(x_0) \leqslant f(x) \leqslant f(x_1)$ et comme $f(x_0) = f(x_1)$ on en déduit $\forall x \in [x_0, x_1]$, $f(x) =$ Constante. Cela signifie que f' s'annule sur tout cet intervalle, soit un nombre infini de points, ce qui est exclu par hypothèse.
f est donc strictement croissante sur I.

➠ Le même type de raisonnement s'applique pour les deux autres cas ($f'(x) = 0$ et $f'(x) \leqslant 0$).

Remarque :
L'ensemble formé par ce théorème et sa réciproque constitue le « principe de Lagrange ».
Il permet d'associer le sens de variation d'une fonction avec le signe de sa dérivée.

Propriété 2.19 (Condition extremum) Soit f une fonction dérivable sur un intervalle I. Si, sur un intervalle ouvert inclus dans I et contenant x_0, f admet un *extremum local* en x_0 alors $f'(x_0) = 0$.

Démonstration Supposons que, sur l'intervalle $]a, b[$ contenant x_0, f admette un maximum en x_0 : supposons donc que $\forall x \in]a, b[, \quad f(x) \leqslant f(x_0)$.
Comme x_0 n'est pas une extrémité de l'intervalle
- pour $x < x_0$ la fonction f doit croitre pour atteindre finalement $f(x_0)$
- pour $x > x_0$ la fonction f ne peut que décroitre.

La dérivée doit donc passer de valeurs positives a des valeurs négatives et, f' étant continue, d'après le TVI, elle doit passer par la valeur 0 en x_0.
Même raisonnement si f admet un minimum en x_0.

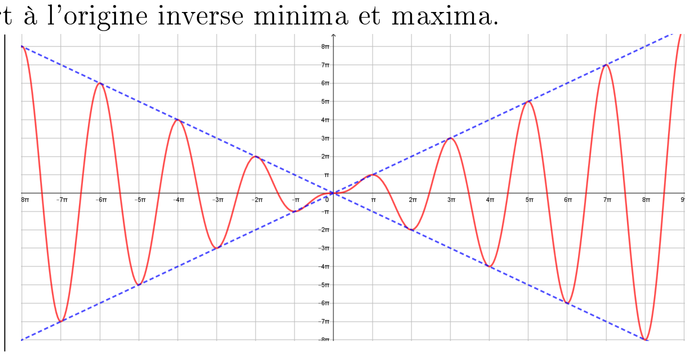
Remarque :
- Attention la réciproque est fausse ! La fonction cube : $f(x) = x^3$ est dérivable sur $\mathbb{R}$ et $f'(x) = 3x^2$ donc $f'(0) = 0$. Pourtant $f(0)$ n'est pas un extremum local de f : la fonction cube est strictement croissante sur $\mathbb{R}$, la dérivée $3x^2$ étant strictement positive sauf en un point où elle s'annule.
- La présence d'un extremum local (on dit aussi extremum relatif) exige que l'intervalle soit ouvert. La fonction $g : x \longmapsto (\sqrt{x})^2$ par exemple admet un minimum sur $[0, 1]$ en $x_0 = 0$ et pourtant $g'(x_0) = 1 \neq 0$.

Propriété 2.20 Soit f une fonction dérivable sur un intervalle I et soit $x_0 \in I$ tel que f' s'annule en x_0 en changeant de signe alors $f(x_0)$ est un extremum local de f.

Remarque :
Cette propriété, alliée au principe de Lagrange, est au centre de toutes les études de fonction.
1. Déterminer les périodes et les symétries et délimiter en conséquence l'ensemble d'étude utile E
2. Partager si possible E en intervalles I sur lesquels le principe de Lagrange peut être appliqué
3. Dresser le tableau des variations de f restreinte à E (justifiées par les signes de f')
4. Compléter avec les limites aux bornes ouvertes de E et les éventuelles asymptotes.

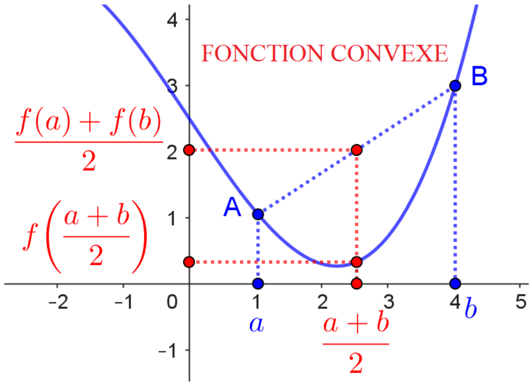
Exemple 21 – Étudions les variations de la fonction $f : x \longmapsto \sin x - x \cos x$.
Cette fonction est impaire car $f(-x) = \sin(-x) - (-x)\cos(-x) = -\sin x + x \cos x = -f(x)$.
L'ensemble d'étude utile E est $\mathbb{R}^+$.
f est dérivable en tout point de $\mathbb{R}$ et $f'(x) = \cos x + x \sin x - \cos x = x \sin x$.
La dérivée f' s'annule en changeant de signe pour $x = 0 \ [\pi]$ ce qui conduit à une succession de maximum et de minimum locaux, alternativement. Il n'y a que pour $x = 0$ que la dérivée s'annule sans changer de signe (car x et $\sin x$ change simultanément de signe) d'où l'absence d'extrémum local pour cette valeur. La représentation graphique, à cet égard, permet de bien visualiser la situation.
Les maxima locaux sont alignés sur la droite d'équation $y = x$: $f(\pi) = \sin(\pi) - (\pi)\cos(\pi) = \pi$, $f(3\pi) = \sin(3\pi) - (3\pi)\cos(3\pi) = 3\pi$, etc. De même pour les minima locaux (alignement sur la droite d'équation $y = -x$). La symétrie par rapport à l'origine inverse minima et maxima.

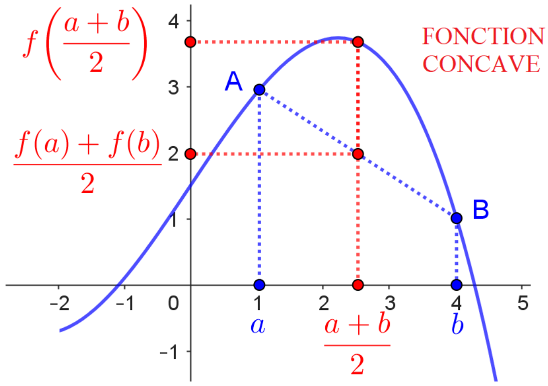
x	0		π		2π		3π		$\dots$
$f'(x)$	0	$+$	0	$-$	0	$+$	0	$-$	$\dots$
f	0	$\nearrow$	π	$\searrow$	-2π	$\nearrow$	3π	$\searrow$	$\dots$



4.c. Convexité

DÉFINITION 2.11 (FONCTION CONVEXE/CONCAVE) Une fonction f est *convexe* (respect. *concave*) sur un intervalle I si pour tout couple de points (A, B) situés sur sa courbe $\mathscr{C}_f$, le segment $[AB]$ est au-dessus (resp. en dessous) de la portion de $\mathscr{C}_f$ située entre A et B.





Remarques :

- On peut traduire cette définition d'une fonction convexe (resp. concave) : $\forall a, b \in I^2,\ \frac{f(a)+f(b)}{2} > f\left(\frac{a+b}{2}\right)$ (resp. $<$). Les points $A(a, f(a))$ et $B(b, f(b))$ étant situés sur la courbe $\mathscr{C}_f$, le point $C(\frac{a+b}{2}, f(\frac{a+b}{2}))$ l'est aussi, situé entre A et B. Par contre, le point $(\frac{a+b}{2}, \frac{f(a)+f(b)}{2})$ est le milieu du segment $[AB]$: par définition, ce point est au-dessus de C si la fonction est convexe, en-dessous si elle est concave.
- Cette remarque se généralise à tous points C_μ de coordonnées $(\mu a + (1-\mu)b, f(\mu a + (1-\mu)b))$ et $(\mu a + (1-\mu)b, \mu f(a) + (1-\mu)f(b))$ où μ est un réel de l'intervalle $[0, 1]$ (si $\mu = \frac{1}{2}$, on se retrouve avec la remarque précédente) : ce dernier point est au-dessus de C_μ si la fonction est convexe, en-dessous si elle est concave.
  Autre caractérisation d'une fonction convexe (resp. concave) : $\forall a, b > a \in I^2,\ \ \forall x \in [a, b],\ \ f(x) \leqslant f(a) + (x-a)\frac{f(b)-f(a)}{b-a}$ (resp. $>$), le membre de droite de cette proposition étant simplement l'équation de la droite (AB).
- Connaissant les courbes des fonctions de référence, on peut dors et déjà affirmer que la fonction exponentielle et la fonction carrée sont convexes sur $\mathbb{R}$, la fonction cube est convexe sur $\mathbb{R}^+$ et concave sur $\mathbb{R}^-$, la fonction racine carrée est concave sur $\mathbb{R}^+$, la fonction ln et la fonction inverse sont concaves sur $]0, +\infty[$, etc.

PROPRIÉTÉ 2.21 (DÉRIVÉE CROISSANTE) Soit f une fonction dérivable sur un intervalle I.
Si f' est croissante sur I alors $\begin{cases} \forall a \in I,\ \ \forall x \in I,\ \ f(x) \geqslant f(a) + (x-a)f'(a) \\ \text{la fonction } f \text{ est convexe sur } I \end{cases}$

DÉMONSTRATION ➟ Montrons d'abord que $\forall a \in I,\ \ \forall x \in I,\ \ f(x) \geqslant f(a) + (x-a)f'(a)$.
Étudions la fonction $g : x \longmapsto f(x) - f(a) - (x-a)f'(a)$. La dérivée de g est $g'(x) = f'(x) - f'(a)$.
$g'(x) \geqslant 0 \iff f'(x) \geqslant f'(a) \iff x \geqslant a$ car f' est croissante.
$g'(x) \leqslant 0 \iff f'(x) \leqslant f'(a) \iff x \leqslant a$ car f' est croissante.
g est donc décroissante pour $x \leqslant a$ et croissante pour $x \geqslant a$, elle admet un minimum en a.
Ce mimum est $g(a) = 0$, c'est-à-dire
$\forall x \in I,\ \ g(x) \geqslant g(a) \iff f(x) - f(a) - (x-a)f'(a) \geqslant 0 \iff f(x) \geqslant f(a) + (x-a)f'(a)$.

➟ Montrons maintenant que cette propriété traduit le fait que f est convexe sur I.
Plus précisément montrons que $\forall a, b > 0 \in I^2,\ \ \forall x \in [a, b],\ \ f(x) \leqslant f(a) + (x-a)\frac{f(b)-f(a)}{b-a}$.
Pour cela étudions la fonction $h : x \longmapsto f(x) - f(a) + (x-a)\frac{f(b)-f(a)}{b-a}$.
La dérivée de h est $h'(x) = f'(x) + \frac{f(b)-f(a)}{b-a}$.

D'après le TAF sur $[a,b]$, $\exists c \in [a,b]$, $f'(c) = \frac{f(b)-f(a)}{b-a}$.
Ainsi $h'(x) = f'(x) - f'(c)$ et, comme f' est croissante :

- $h'(x) \geqslant 0 \iff f'(x) \geqslant f'(c) \iff x \geqslant c$
- $h'(x) \leqslant 0 \iff f'(x) \leqslant f'(c) \iff x \leqslant c$

h est donc décroissante pour $x \leqslant c$ et croissante pour $x \geqslant c$, elle admet un minimum en c.
Comme $h(a) = f(a) - f(a) + (a-a)\frac{f(b)-f(a)}{b-a} = 0$ et $h(b) = f(b) - f(a) + (b-a)\frac{f(b)-f(a)}{b-a} = 0$,
la fonction h est négative entre a et b, c'est-à-dire $\forall x \in [a,b]$, $f(x) \leqslant f(a) + (x-a)\frac{f(b)-f(a)}{b-a}$.
La fonction f est donc convexe.

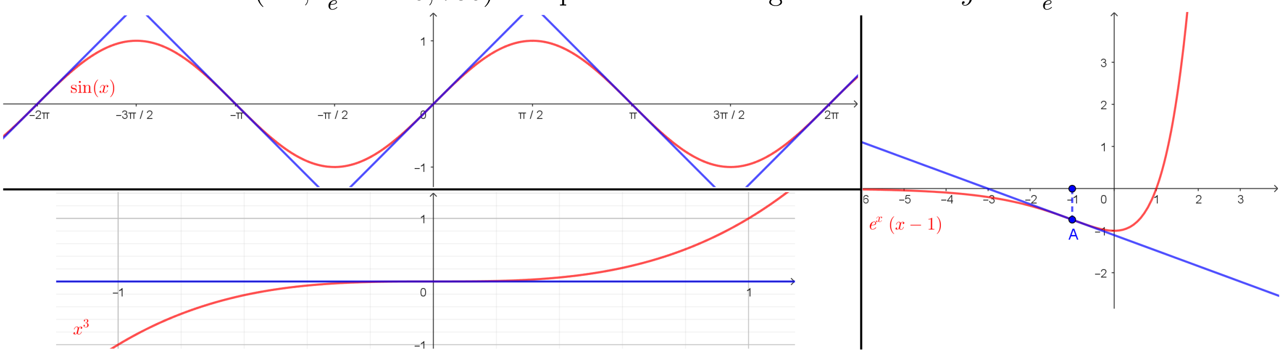
Remarques :

- On peut bien sûr énoncer aussi : si f est une fonction dérivable sur I et si f' est décroissante alors $\begin{cases} \forall a \in I, \ \forall x \in I, \ f(x) \leqslant f(a) + (x-a)f'(a) \\ \text{la fonction } f \text{ est concave sur } I \end{cases}$
- L'inégalité $f(x) \geqslant f(a) + (x-a)f'(a)$ se traduit graphiquement en disant que la courbe $\mathscr{C}_f$ est au-dessus de ses tangentes en tous points de l'intervalle I (en effet, l'équation de la tangente en $(a, f(a))$ est $y = f(a) + (x-a)f'(a)$). Par ailleurs, la condition f' est croissante sur I équivaut à $\forall x \in I$, $f''(x) \geqslant 0$. Ainsi, cette propriété nous donne un moyen de tester la convexité (le signe de la dérivée seconde) et une propriété caractéristique (la position de la courbe par rapport à ses tangentes).

Définition 2.12 (Point d'inflexion) Soit f une fonction dérivable deux fois sur un intervalle I telle que, en $x = a$, la dérivée seconde s'annule en changeant de signe. Dans ce cas, le point $(a, f(a))$ est un *point d'inflexion* pour la courbe $\mathscr{C}_f$.

Remarques :

- Si la courbe d'une fonction admet un point d'inflexion en a, la tangente en a traverse la courbe $\mathscr{C}_f$. C'est ce qui se passe pour la courbe de la fonction cube en 0 ; la dérivée seconde de x^3 étant $6x$, celle-ci s'annule en changeant de signe, le point $(0,0)$ est donc un point d'inflexion pour cette courbe.
- Dans le cas de la fonction cube, la tangente qui traverse la courbe est horizontale. Mais ce n'est absolument pas la règle : la fonction sin par exemple a pour dérivée seconde $-\sin$, celle-ci s'annule en changeant de signe pour $x = 0 \ [\pi]$ mais la pente des tangentes en ces points d'inflexion est 1 (pour $x = 0 \ [2\pi]$) ou -1 (pour $x = \pi \ [2\pi]$).
- La fonction $f : x \longmapsto (x-1)e^x$ a pour dérivée première $f' : x \longmapsto xe^x$ et pour dérivée seconde $f'' : x \longmapsto (x+1)e^x$. Le signe de $x+1$ détermine donc la convexité de la courbe $\mathscr{C}_f$: pour $x > -1$ la courbe est convexe, sinon elle est concave, le point d'inflexion A ayant pour coordonnées $(-1, \frac{-2}{e} \approx -0,736)$. L'équation de la tangente en A est $y = \frac{-x-3}{e}$.



5. Logarithmes

5.a. Logarithme népérien

La propriété 2.11 nous a déjà permis de définir cette fonction car, la fonction exp, étant strictement croissante de $\mathbb{R}$ vers $\mathbb{R}^{+*}$, définit une bijection entre ces deux ensembles. Elle admet donc une bijection réciproque de $\mathbb{R}^{+*}$ vers $\mathbb{R}$ appelée « logarithme népérien » notée ln.

Définition 2.13 (Fonction ln) La fonction ln est la fonction réciproque de la fonction exp. Autrement dit $\forall x \in R, \ \ \forall y \in]0;+\infty[, \ \ y = e^x \iff x = \ln y$.

Propriété 2.22 (Propriétés immédiates) Ces propriétés découlent de la définition :

- $\ln 1 = 0$, $\ln e = 1$, $\ln e^2 = 2$, $\ln \frac{1}{e} = -1$, et plus généralement $\forall x \in \mathbb{R}, \ \ \ln(e^x) = x$
- $e^{\ln 1} = 1$, $e^{\ln 2} = 2$, et plus généralement $\forall x \in]0;+\infty[, \ \ e^{\ln x} = x$
- L'ensemble de définition de $x \longmapsto \ln(u(x))$ est l'ensemble des $x \in \mathscr{D}_u$ tels que $u(x) > 0$
- L'équation $e^x = a$ a pour solution $x = \ln a$ si $a > 0$ (n'a pas de solution sinon)
  L'équation $\ln x = a$ a pour solution $x = e^a$

Propriété 2.23 (Dérivée) La fonction ln est dérivable sur $]0;+\infty[$ et $\ln'(x) = \frac{1}{x}$

Démonstration Cela a été montré comme application de la propriété 2.18 :
$\ln'(y) = \frac{1}{\exp'(x)} = \frac{1}{\exp'(\ln(y))} = \frac{1}{\exp(\ln(y))} = \frac{1}{\exp \circ \ln(y)} = \frac{1}{y}$

Remarques :

- Une conséquence immédiate est que la fonction ln est strictement croissante sur $]0;+\infty[$.
  On en déduit également le signe de ln : $\ln x < 0 \iff 0 < x < 1$ et $\ln x > 0 \iff x > 1$
  La courbe de la fonction ln a été tracée dans les remarques de la propriété 2.11.
- La dérivée de la fonction composée $x \longmapsto \ln(u(x))$ est $x \longmapsto \frac{u'(x)}{u(x)}$.
  En particulier, si $f(x) = \ln(x^2-1)$ alors $f'(x) = \frac{2x}{x^2-1}$. La fonction f est définie pour $x^2 - 1 > 0$ soit pour $x \in]-\infty, -1[\cup]1, +\infty[$

Propriété 2.24 (fonctionnelle) Pour tous réels positifs a et b on a :
$\ln(ab) = \ln(a) + \ln(b)$ et $\ln(\frac{a}{b}) = \ln(a) - \ln(b)$, en particulier $\ln(\frac{1}{b}) = -\ln(b)$

Démonstration Connaissant la propriété fonctionnelle de la fonction exponentielle $e^{a+b} = e^a \times e^b$ qui transforme une somme en produit, on écrit $e^{\ln(a)+\ln(b)} = e^{\ln(a)} \times e^{\ln(b)}$.
Comme $e^{\ln(a)} = a$ si $a > 0$, on obtient $e^{\ln(a)+\ln(b)} = a \times b = e^{\ln(ab)}$.
La fonction exponentielle étant une bijection on a l'équivalence $e^a = e^b \iff a = b$.
Finalement on en déduit $\ln(a) + \ln(b) = \ln(ab)$: la fonction ln transforme un produit en somme.

En utilisant la propriété précédente $\ln(b) + \ln(\frac{1}{b}) = \ln(b \times \frac{1}{b}) = \ln(1) = 0$.
On en déduit $\ln(\frac{1}{b}) = -\ln(b)$.
Par conséquent $\ln(\frac{a}{b}) = \ln(a \times \frac{1}{b}) = \ln(a) + \ln(\frac{1}{b}) = \ln(a) - \ln(b)$.

Attention : Cette propriété n'est utilisable que si a et b sont positifs strictement.
S'ils sont tous les deux négatifs, le produit ab et le quotient $\frac{a}{b}$ étant positifs, $\ln(ab)$ et $\ln(\frac{a}{b})$ existent mais la propriété ne peut être utilisée (ni $\ln(a)$ ni $\ln(b)$ ne sont définis).

Propriété 2.25 (Puissances) Soient a un réel positif et n un entier positif.

- $\ln(a^n) = n\ln(a)$
- $\ln(a^{-n}) = -n\ln(a)$
- $\ln(\sqrt[n]{a}) = \ln(a^{\frac{1}{n}}) = \frac{1}{n}\ln(a)$ (pour $a \geqslant 0$ et à partir de $n = 2$)

Démonstration ➠ Raisonnons par récurrence :
Pour $n = 0$, on a $\ln(a^0) = \ln(1) = 0 = 0\ln(a)$, la propriété est vérifée.
Pour $n > 0$, supposons que $\ln(a^n) = n\ln(a)$ et montrons que $\ln(a^{n+1}) = (n+1)\ln(a)$.
$\ln(a^{n+1}) = \ln(a^n \times a^1) = \ln(a^n) + \ln(a) = n\ln(a) + \ln(a) = (n+1)\ln(a)$.
La propriété est héréditaire, or elle est vraie pour $n = 0$, elle est donc vraie pour tout $n \geqslant 0$.

➠ $\ln(a^{-n}) = \ln(\frac{1}{a^n}) = -\ln(a^n)$ d'après la propriété fonctionnelle.
En appliquant le résultat précédent $\ln(a^{-n}) = -\ln(a^n) = -n\ln(a)$.

➠ Par définition, on a $(\sqrt[n]{a})^n = a$ (la fonction $x \longmapsto \sqrt[n]{x}$ est la réciproque de la fonction $x \longmapsto x^n$).
Par conséquent $\ln((\sqrt[n]{a})^n) = \ln(a)$ et comme, d'après la première proposition $\ln((\sqrt[n]{a})^n) = n\ln(\sqrt[n]{a})$, on en déduit $\ln(a) = n\ln(\sqrt[n]{a}) \iff \ln(\sqrt[n]{a}) = \frac{1}{n}\ln(a)$.

Remarques : Pour tout réel positif $x^{\frac{1}{n}} = \sqrt[n]{x}$.
En effet $(x^{\frac{1}{n}})^n = x^{\frac{n}{n}} = x$ et, comme pour $x \geqslant 0$ on a $x^{\frac{1}{n}} \geqslant 0$, $x^{\frac{1}{n}} = \sqrt[n]{x}$.
Cette notation permet de retrouver plus facilement la dérivée de $\sqrt[n]{x}$ que l'on sait obtenir avec le théorème de dérivation de la réciproque : $(\sqrt[n]{y})' = \frac{1}{nx^{n-1}} = \frac{1}{nx^{n-1}} = \frac{1}{n(\sqrt[n]{y})^{n-1}}$ d'où $(\sqrt[n]{x})' = \frac{1}{n(\sqrt[n]{x})^{n-1}}$
La propriété $(x^n)' = nx^{n-1}$ appliquée avec $\frac{1}{n}$ s'écrit $(x^{\frac{1}{n}})' = \frac{1}{n}x^{\frac{1}{n}-1} = \frac{1}{n}x^{\frac{1-n}{n}} = \frac{1}{nx^{\frac{n-1}{n}}}$.
Les deux résultats sont égaux $\frac{1}{n(\sqrt[n]{x})^{n-1}} = \frac{1}{nx^{\frac{n-1}{n}}}$ car $x^{\frac{n-1}{n}} = x^{\frac{1}{n}\times n-1} = (x^{\frac{1}{n}})^{n-1} = (\sqrt[n]{x})^{n-1}$
Les autres propriétés fonctionnelles des puissances sont respectées avec cette notation :

- $(x^n)^{\frac{1}{n}} = (x^{\frac{1}{n}})^n = x$
- $(xx')^{\frac{1}{n}} = x^{\frac{1}{n}} \times x'^{\frac{1}{n}}$
- $(\frac{x}{x'})^{\frac{1}{n}} = \frac{x^{\frac{1}{n}}}{x'^{\frac{1}{n}}}$
- $x^{\frac{1}{n}} \times x^{\frac{1}{p}} = x^{\frac{1}{n}+\frac{1}{p}}$
- $\frac{x^{\frac{1}{n}}}{x^{\frac{1}{p}}} = x^{\frac{1}{n}-\frac{1}{p}}$
- $(x^{\frac{1}{n}})^{\frac{1}{p}} = x^{\frac{1}{np}}$

Toutes ces propriétés peuvent être justifiées, cela constitue un bon exercice de s'y essayer.

Propriété 2.26 (Limites) Au voisinage des bornes ouvertes de l'ensemble de définition :

- $\lim\limits_{x\to+\infty} \ln(x) = +\infty$
- $\lim\limits_{x\to 0^+} \ln(x) = -\infty$

Démonstration ➠ D'après la définition 2.1 on doit déterminer x_0 tel que
$\forall A \in \mathbb{R},\ \exists x_0 \in \mathbb{R},\ \forall x > x_0,\ \ln(x) > A$.
Or $\ln(x) > A \iff e^{\ln(x)} > e^A \iff x > e^A$, il suffit donc de prendre $x_0 = e^A$.

➠ Pour la limite en 0+ on a $\lim\limits_{x\to 0^+} \ln(x) = \lim\limits_{X\to+\infty} \ln(\frac{1}{X}) = \lim\limits_{X\to+\infty} -\ln(X) = -\infty$.
J'ai opéré un changement de variable en prenant $X = \frac{1}{x}$, sachant que $\lim\limits_{x\to 0^+} \frac{1}{x} = +\infty$

Remarque : La fonction ln a une limite infinie au voisinage de l'infini mais la croissance de cette fonction est de plus en plus lente.

$\ln(x) >$	1	5	10	50	100	500
$x >$	2,718	148,413	22 026,466	$5,185 \times 10^{21}$	$2,688 \times 10^{43}$	$1,404 \times 10^{217}$

Propriété 2.27 (Croissances comparées) Au voisinage de $+\infty$ comme au voisinage de 0^+, la fonction ln est dominée par toutes les fonctions puissances ($x \longmapsto x^n$ et $x \longmapsto x^{\frac{1}{n}}$) :
$\lim\limits_{x\to+\infty} \frac{\ln(x)}{x} = 0$ et, pour tout $n \in \mathbb{N}^*$, $\lim\limits_{x\to+\infty} \frac{\ln(x)}{x^n} = 0$ et $\lim\limits_{x\to+\infty} \frac{\ln(x)}{\sqrt[n]{x}} = \lim\limits_{x\to+\infty} \frac{\ln(x)}{x^{\frac{1}{n}}} = 0$

$\lim\limits_{x\to0^+} x\ln(x) = 0$ et, pour tout $n \in \mathbb{N}^*$, $\lim\limits_{x\to0^+} x^n\ln(x) = 0$ et $\lim\limits_{x\to0^+} \sqrt[n]{x}\ln(x) = \lim\limits_{x\to0^+} x^{\frac{1}{n}}\ln(x) = 0$

Démonstration ➠ Effectuons le changement de variable $X = \ln(x) \iff e^X = e^{\ln(x)} = x$.
On sait que $\lim\limits_{x\to+\infty} \frac{e^x}{x} = +\infty$, or $\frac{\ln(x)}{x} = \frac{X}{e^X} = \frac{1}{\frac{e^X}{X}}$
On en déduit $\lim\limits_{x\to+\infty} \frac{\ln(x)}{x} = \lim\limits_{X\to+\infty} \frac{1}{\frac{e^X}{X}} = 0^+$

➠ Posons, cette fois, $X = x^n$.
Ainsi, comme $\ln(x^n) = n\ln(x)$, on en déduit $\ln(x) = \frac{\ln(X)}{n}$ et $\frac{\ln(x)}{x^n} = \frac{\ln(X)}{nX} = \frac{1}{n} \times \frac{\ln(X)}{X}$
Par conséquent $\lim\limits_{x\to+\infty} \frac{\ln(x)}{x^n} = \lim\limits_{X\to+\infty} \frac{1}{n} \times \frac{\ln(X)}{X} = 0^+$

➠ De la même façon, en posant $X = x^{\frac{1}{n}}$, comme $\ln(x^{\frac{1}{n}}) = \frac{1}{n}\ln(x)$ on a $\ln(x) = n\ln(X)$.
D'où $\frac{\ln(x)}{x^{\frac{1}{n}}} = \frac{n\ln(X)}{X} = n \times \frac{\ln(X)}{X}$
Par conséquent $\lim\limits_{x\to+\infty} \frac{\ln(x)}{x^{\frac{1}{n}}} = \lim\limits_{X\to+\infty} n \times \frac{\ln(X)}{X} = 0^+$.

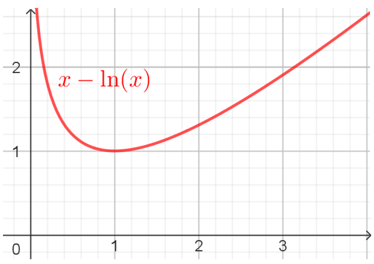
➠ Pour lever l'indétermination de type $0 \times \infty$ de $\lim\limits_{x\to0^+} x\ln(x)$, posons $X = \frac{1}{x} \iff x = \frac{1}{X}$.
Ainsi, comme $\ln(\frac{1}{x}) = -\ln(x)$, on en déduit $\ln(x) = -\ln(X)$ et $x\ln(x) = \frac{-\ln(X)}{X}$
Par conséquent $\lim\limits_{x\to0^+} x\ln(x) = \lim\limits_{X\to+\infty} \frac{-\ln(X)}{X} = -\lim\limits_{X\to+\infty} \frac{\ln(X)}{X} = 0^-$.

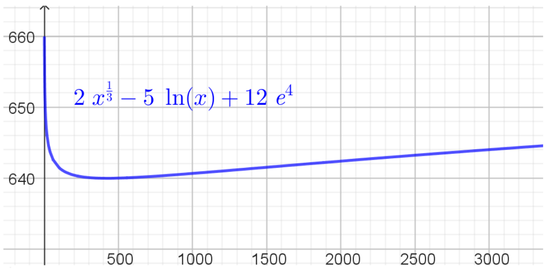
➠ Avec le même changement de variable $X = \frac{1}{x} \iff x = \frac{1}{X}$ pour x et X strictement positifs, on obtient $x^n\ln(x) = \frac{-\ln(X)}{X^n}$ et comme $\lim\limits_{X\to+\infty} \frac{\ln(X)}{X^n} = 0^+$ on en déduit
$\lim\limits_{x\to0^+} x^n\ln(x) = -\lim\limits_{X\to+\infty} \frac{\ln(X)}{X^n} = 0^-$.

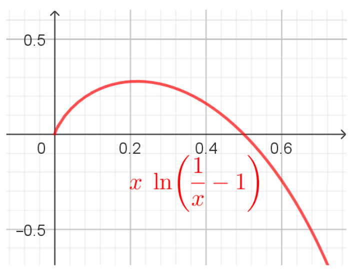
➠ Pour finir, en posant $X = x^{\frac{1}{n}}$ on a $\ln(x) = n\ln(X)$ et $x^{\frac{1}{n}}\ln(x) = nX\ln(X)$
d'où $\lim\limits_{x\to0^+} x^{\frac{1}{n}}\ln(x) = n\lim\limits_{X\to0^+} X\ln(X) = 0^-$.

Remarques : Cette domination permet de lever certaines indétermination.
Notamment, au voisinage de $+\infty$, celles de type $\infty - \infty$, en mettant en facteur le terme dominant :

- La limite en $+\infty$ de la fonction $f : x \longmapsto x - \ln(x)$ est une forme indéterminée du type $\infty - \infty$.
  En écrivant $x - \ln(x) = x(1 - \frac{\ln(x)}{x})$, on lève l'indétermination :
  comme $\lim\limits_{x\to+\infty} 1 - \frac{\ln(x)}{x} = 1$, par produit on obtient $\lim\limits_{x\to+\infty} f(x) = \lim\limits_{x\to+\infty} x = +\infty$
- De même pour $g : x \longmapsto 2\sqrt[3]{x} - 5\ln(x) + 12e^4$. On écrit $g(x) = \sqrt[3]{x}\left(2 - 5\frac{\ln(x)}{\sqrt[3]{x}} + \frac{12e^4}{\sqrt[3]{x}}\right)$ et on conclut sachant que $\lim\limits_{x\to+\infty} 2 - 5\frac{\ln(x)}{\sqrt[3]{x}} + \frac{12e^4}{\sqrt[3]{x}} = 2$, par produit $\lim\limits_{x\to+\infty} f(x) = \lim\limits_{x\to+\infty} 2\sqrt[3]{x} = +\infty$







Au voisinage de 0^+ cherchons à lever l'indétermination de la limite de $h : x \longmapsto x\ln(\frac{1}{x} - 1)$
$h(x) = x\ln(\frac{1-x}{x}) = x(\ln(1-x) - \ln(x)) = x\ln(1-x) - x\ln(x)$.

Pour $x > 0 \iff -x < 0 \iff 1 - x < 1$, le 1$^{\text{er}}$ terme n'est pas une forme indéterminée en 0^+ :
$\lim\limits_{x\to 0^+} x\ln(1-x) = 0^-$ car $\lim\limits_{x\to 0^+} \ln(1-x) = \lim\limits_{X\to 1^-} \ln(X) = 0^-$.
Le 2$^{\text{e}}$ terme a pour limite 0^+ d'après la propriété ($\lim\limits_{x\to 0^+} x\ln(x) = 0^-$).

Conclusion : $\lim\limits_{x\to 0^+} h(x) = 0$.

PROPRIÉTÉ 2.28 (TAUX D'ACCROISSEMENT) Au voisinage de 0, on a $\lim\limits_{x\to 0} \frac{\ln(x+1)}{x} = 1$

DÉMONSTRATION Cette expression $\frac{\ln(x+1)}{x}$ est le taux d'accroissement $\tau(x)$ de la fonction ln entre $1+x$ et 1 : $\tau(x) = \frac{\ln(1+x)-\ln(1)}{(1+x)-1} = \frac{\ln(x+1)}{x}$
On sait que la limite de ce taux d'accroissement lorsque x tend vers 0 est le nombre dérivé $\ln'(1) = 1$.

5.b. Logarithme de base a

On a vu la fonction logarithme népérien : c'est la fonction logarithme de base e car $\ln(e) = 1$

DÉFINITION 2.14 (FONCTION $\ln_a$) a étant un réel strictement positif donné,
la fonction $\ln_a$ est définie pour tout $x > 0$ par $\ln_a(x) = \frac{\ln(x)}{\ln(a)}$.

PROPRIÉTÉ 2.29 Les propriétés des fonctions $\ln_a$ sont similaires à celles de la fonction ln :

$\ln_a(a) = 1$	$\forall (x, x') \in]0, +\infty[^2,\ \ln_a(xx') = \ln_a(x) + \ln_a(x')$	$\forall n \in \mathbb{Z},\ \ln_a(a^n) = n$
$\ln_a(1) = 0$	$\forall (x, x') \in]0, +\infty[^2,\ \ln_a(\frac{x}{x'}) = \ln_a(x) - \ln_a(x')$	$(\ln_a)'(x) = \frac{1}{x\ln(a)}$

DÉMONSTRATION Cela découle des propriétés de la fonction ln et $\forall x > 0,\ \ln_a(x) = \frac{1}{\ln(a)} \times \ln(x)$.
$\ln_a(xx') = \frac{1}{\ln(a)} \times \ln(xx') = \frac{1}{\ln(a)} \times (\ln(x) + \ln(x')) = \frac{1}{\ln(a)} \times (\ln(x)) + \frac{1}{\ln(a)} \times (\ln(x')) = \ln_a(x) + \ln_a(x')$
Les autres justifications sont similaires.

Remarques :

- Parmi toutes les fonctions $\ln_a$, la plus utilisée est la fonction *logarithme décimal* $\ln_{10}$, noté simplement log qui vérifie notamment la propriété $\log(10^n) = n$ conduisant au logarithme d'un nombre en écriture scientifique $\log(a \times 10^n) = n + \log(a)$. De nombreuses unités de mesure sont définies à partir des logarithmes décimaux : le potentiel Hydrogène d'une solution aqueuse est défini par $pH = -\log([H_3O^+])$ où $[H_3O^+]$ est la concentration en ions hydronium H_3O^+ de la solution ; le décibel (dB) est l'unité de mesure d'une grandeur $X = 10 \times \log(\frac{P_1}{P_0})$ où P_1 est la puissance mesurée et P_0 une puissance de référence. Ainsi, en acoustique, on mesure le niveau sonore en décibels : 80 dB correspond à un son 10^8 fois plus important que le niveau sonore de référence (presque imperceptible).
- Le *logarithme binaire* est le logarithme de base 2, noté souvent lb au lieu de $\ln_2$. Il est utilisé en informatique principalement où la représentation d'un nombre en virgule flottante est $x = m \times 2^e$ avec une mantisse $1 \leqslant |m| < 2$ et un exposant e (la mantisse est écrite sans le 1 initial et l'exposant est décalé, voir le cours de NSI de 1$^{\text{re}}$). Le logarithme binaire de x est alors égal à $lb(m) + e$ et l'ordinateur peut calculer $lb(m)$ bit par bit, en utilisant les propriétés : $lb(1) = 0$, $lb(x) = \frac{lb(x^2)}{2}$ et $lb(x) = lb(\frac{x}{2}) + 1$ (pour un nouveau bit de $0 < x < 2$, si $x^2 \geqslant 2$ on note 1 et on divise le résultat par 2, sinon on note 0 ; on réitère ensuite avec le résultat). Les autres logarithmes (népérien, décimal, etc.) sont évalués à partie du logarithme binaire grâce à la relation $log_a(x) = \frac{\ln(2)}{\ln(a)} lb(x)$.

5.c. Exponentielle de base a

On a vu la fonction exponentielle : c'est la fonction exponentielle de base e car $e^x = e^{x\ln(e)} = \exp_e(x)$

Définition 2.15 (Fonction $\exp_a$) a étant un réel strictement positif différent de 1, la fonction $\exp_a$ est définie sur $\mathbb{R}$ par $\exp_a(x) = e^{x\ln(a)}$.

Propriété 2.30 (Fonction a^x) Soient $a \in]0,1[\cup]1,+\infty[$ et $x \in \mathbb{R}$, on a $\exp_a(x) = a^x$

Démonstration ➲ Si x est un entier $x = n$ avec $n \in \mathbb{N}$, on a
$\exp_a(x) = \exp_a(n) = e^{n\ln(a)} = \left(e^{\ln(a)}\right)^n = a^n = a^x$.

➲ Si x est un rationnel positif $x = \frac{n}{p}$ avec $n \in \mathbb{N}$ et $p \in \mathbb{N}^*$, on a
$\exp_a(x) = \exp_a(\frac{n}{p}) = e^{\frac{n}{p}\ln(a)} = \left(e^{\frac{1}{p}\ln(a)}\right)^n = \left(e^{\ln(a^{\frac{1}{p}})}\right)^n = \left(a^{\frac{1}{p}}\right)^n = a^{\frac{n}{p}} = a^x$.

➲ Si x est un rationnel négatif $x = -\frac{n}{p}$ avec $n \in \mathbb{N}$ et $p \in \mathbb{N}^*$, on a
$\exp_a(x) = \exp_a(-\frac{n}{p}) = \frac{1}{e^{\frac{n}{p}\ln(a)}} = \frac{1}{a^{\frac{n}{p}}} = a^{-\frac{n}{p}} = a^x$.

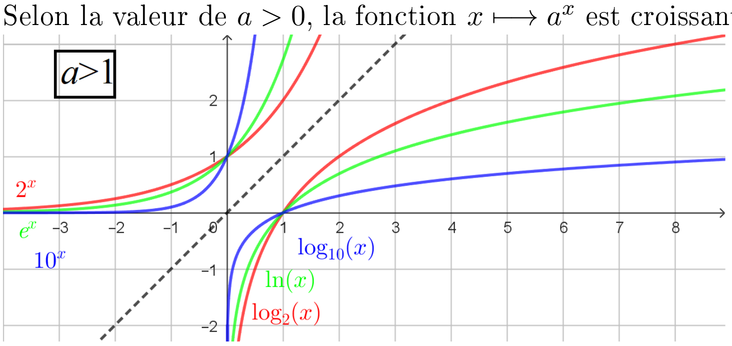
➲ Si x est un réel, on admettra la propriété en considérant que x peut alors être défini comme la limite d'une suite de rationnels (c'est toujours possible car $\mathbb{Q}$ est « dense » dans $\mathbb{R}$) qui, chacun, vérifient la propriété. Par unicité de la limite on en conclut que $\exp_a(x) = a^x$.

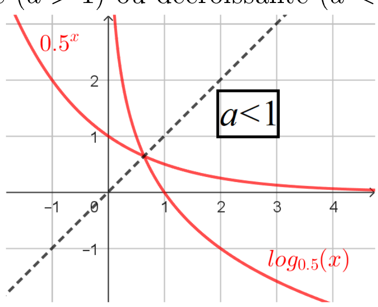
Propriété 2.31 Les propriétés des fonctions $\exp_a$ sont similaires à celles de la fonction exp :

$\forall x \in \mathbb{R},\ \ a^x > 0$	$\forall x \in \mathbb{R}^{+*},\ \ a^{\log_a(x)} = x$
$\forall x \in \mathbb{R},\ \ \ln(a^x) = x\ln(a)$	$\forall x \in \mathbb{R},\ \ \forall y \in \mathbb{R}^{+*},\ \ x = \log_a(y) \Longleftrightarrow y = a^x$
$\forall x \in \mathbb{R},\ \ \log_a(a^x) = x$	$\forall x \in \mathbb{R},\ \ (a^x)' = \ln(a) \times a^x$

Démonstration $\forall x \in \mathbb{R},\ \ a^x = e^{x\ln(a)}$ or $\forall x \in \mathbb{R},\ \ e^x > 0$ donc $\forall x \in \mathbb{R},\ \ a^x > 0$
$\forall x \in \mathbb{R},\ \ \ln(a^x) = \ln(e^{x\ln(a)}) = \ln((e^x)^{\ln(a)}) = \ln(a)\ln(e^x) = x\ln(a)$
$\forall x \in \mathbb{R},\ \ \log_a(a^x) = \frac{\ln(a^x)}{\ln(a)} = \frac{x\ln(a)}{\ln(a)} = x$
$\forall x \in \mathbb{R}^{+*},\ \ a^{\log_a(x)} = e^{\ln(a)\times\log_a(x)} = e^{\ln(a)\times\frac{\ln(x)}{\ln(a)}} = e^{\ln(x)} = x$
$\forall x \in \mathbb{R},\ \ \forall y \in \mathbb{R}^{+*},\ \ x = \log_a(y) \Longrightarrow a^x = a^{\log_a(y)} \Longrightarrow a^x = y$ et,
réciproquement $\forall x \in \mathbb{R},\ \ \forall y \in \mathbb{R}^{+*},\ \ a^x = y \Longrightarrow \log_a(a^x) = \log_a(y) \Longrightarrow x = \log_a(y)$.
$\forall x \in \mathbb{R},\ \ (a^x)' =_e^{x\ln(a)})' = \ln(a) \times a^x$ par dérivation d'une fonction composée.

Remarque :
Les propriétés algébriques des fonctions $\exp_a$ sont identiques à celles de la fonction exp.
Les fonctions $\log_a$ et $\exp_a$ sont des bijections réciproques l'une de l'autre.
Selon la valeur de $a > 0$, la fonction $x \longmapsto a^x$ est croissante ($a > 1$) ou décroissante ($a < 1$) :





3 Fontions (2) : primitives et intégration

Plan du chapitre :

Aperçu historique : L'analyse moderne – l'étude du calcul infinitésimal qui contient le calcul différentiel (dérivées, etc.) et le calcul intégral (primitives, intégrales, etc.) – a débuté avec les travaux de Isaac Newton (1642-1727) et Gottfried Wilhelm Leibniz (1646-1716) puis des frères Bernoulli (Jacques (1654-1705) et Jean (1667-1748)). Bien sûr, on peut trouver des travaux précurseurs, chez Archimède (287-212 av. J.-C.) notamment, avec ses calculs d'aires ou de volumes, ou Pierre Fermat (vers 1605-1665) qui, avec sa méthode des *maximis et minimis* étudie la quadrature de la parabole dont on trouve les prémisses dans Apollonius (240–190 av. J.-C.).
Résoudre des problèmes de tangentes mène invariablement à une équation différentielle ce qui fait que les problèmes posés par ce type d'équation sont aussi vieux que l'analyse. L'inconnue de ce type d'équation est une fonction, un objet familier pour nous mais dont le statut était encore assez flou pour les mathématiciens des débuts.
Gottfried Wilhelm Leibniz (1646-1716) développe la méthode de décomposition en éléments simples ; le nombre d'équations différentielles qu'on sait résoudre à son époque par cette voie est assez faible mais progresse. Les Bernoulli savent intégrer l'équation linéaire et quelques autres du premier ordre. Riccati résout en 1722 le cas de l'équation $y' = ay^2 + bx^n$.
Si un problème de géométrie se réduisait à une équation différentielle, on cherchait alors à la résoudre graphiquement. Cette technique fut développée par Clairaut, D'Alembert, Euler ou encore Lagrange. Clairaut remarque, en 1734, l'existence d'une solution singulière à l'équation $y - xy' + f(y') = 0$. D'Alembert, en 1748, trouve un second cas d'intégrale singulière mais c'est Euler et Lagrange qui solutionnent le cas général. En 1743 Euler forme ce qu'on appelle aujourd'hui l'équation caractéristique : les solutions sont de la forme $e^r x$ où les r sont solutions d'une équation polynomiale. Un peu plus tard, il trouve comment obtenir toutes les intégrales lorsqu'une racine r de l'équation caractéristique est multiple : il faut les chercher de la forme $x^k e^r x$.
On sait, dès la fin du XVII^{e} siècle, intégrer les équations différentielles linéaires du premier et du second ordre à coefficients constants par des sommes d'exponentielles mais il faut attendre 1760 pour que la théorie vienne à bout des équations différentielles linéaires à coefficients constants d'ordre quelconque.

Le calcul intégrale commence, on l'a dit avec les travaux d'Archimède et se précise au XVII^{e} siècle avec Leibniz et Newton. Le « S long » : $\displaystyle\int$, aussi appelé « signe intégrateur », a été introduit par Leibniz. Au-delà de la simple introduction proposée en classe de terminale de l'intégrale de Riemann – du nom de Georg Friedrich Bernhard Riemann (1826-1866) qui en achève l'étude – le concept d'intégrale a été considérablement développé avec les intégrales curvilignes, les intégrales multiples, les intégrales de surfaces, les intégrales des formes différentielles, les intégrales de Lebesgue, les intégrales de Kurzweil-Henstock.

1. Primitives

1.a. Primitives d'une fonction continue

DÉFINITION 3.1 (PRIMITIVE) Soit f une fonction définie sur un intervalle I.
Une fonction F est une *primitive* de f sur I si F est dérivable sur I et si, pour tout $x \in I$ on a

$$F'(x) = f(x)$$

Remarques :

- On comprend bien que si F est une primitive de f sur I alors, pour tout réel k, $F + k$ est également une primitive de f sur I car $(F+k)'(x) = F'(x) + (k)' = F'(x) = f(x)$.
  Par exemple la fonction $F : x \longmapsto \frac{x^2}{2}$ est une primitive de la fonction identité $f : x \longmapsto x$ sur $\mathbb{R}$ car $F'(x) = \left(\frac{x^2}{2}\right)' = \frac{2x}{2} = x = f(x)$ pour tout réel. Mais alors les fonctions F_k définies sur $\mathbb{R}$ par $F_k(x) = \frac{x^2}{2} + k$ sont toutes des primitives de la même fonction identité.
- D'après ce que l'on sait de la dérivation des fonctions, une primitive sur $\mathbb{R}$ de la fonction cos est sin, une primitive sur $\mathbb{R}^{+*}$ de la fonction ln est la fonction inverse $x \longmapsto \frac{1}{x}$, une primitive sur $\mathbb{R}$ de la fonction exponentielle $x \longmapsto e^x$ est la fonction $x \longmapsto e^x$, etc.
- Dans le langage des équations différentielles, dire que F est une primitive de f sur un intervalle I où elle est définie c'est dire que F est une solution de l'équation différentielle $F'(x) = f(x)$. On note aussi cette équation différentielle $y' = f(x)$ où la fonction « inconnue » est y, celle-ci étant définie uniquement par le fait que sa dérivée y' égale une certaine fonction f de la variable x. Pour indiquer que la dérivation se fait par rapport à cette variable x (et non une autre éventuellement présente), les physiciens utilisant la notation $\frac{dy}{dx}$, noteront l'équation différentielle $\frac{dy}{dx} = f(x)$.

PROPRIÉTÉ 3.1 Toute fonction continue sur un intervalle I admet des primitives sur I.

DÉMONSTRATION Cette propriété est pour l'instant admise. Elle sera démontrée dans la propriété 3.11 .

PROPRIÉTÉ 3.2 Soit F une primitive de la fonction f sur un intervalle I.
Toutes les primitives de f sur I s'écrivent $F_k = F + k$ où k est un réel quelconque.

DÉMONSTRATION Notons G une primitive de f sur I. On a alors $G'(x) = f(x)$.
Comme par ailleurs F est une primitive de f sur I, on a aussi $F'(x) = f(x)$.
On a donc $F'(x) = G'(x) \Longleftrightarrow G'(x) - F'(x) = 0 \Longleftrightarrow (G' - F')(x) = 0 \Longleftrightarrow (G - F)'(x) = 0$ sur I.
On en déduit que la fonction $G - F$ est une fonction constante
$(G - F)(x) = k \Longleftrightarrow G(x) - F(x) = k \Longleftrightarrow G(x) = F(x) + k$ où k est un réel quelconque.

PROPRIÉTÉ 3.3 (UNICITÉ) Soit f une fonction continue sur un intervalle I, x_0 un réel de I et y_0 un réel quelconque. Il existe une unique primitive F de f sur I telle que $F(x_0) = y_0$.

DÉMONSTRATION D'après la propriété 3.1, comme f est une fonction continue sur I, il en existe une primitive F sur I et, d'après la propriété 3.2, toutes les primitives de f sur I s'écrivent $F + k$ où k est un réel quelconque. Pour avoir $F_k(x_0) = F(x_0) + k = y_0$, il faut et il suffit de prendre $k = y_0 - F(x_0)$; l'unique fonction F_k ainsi définie vérifie bien les contraintes imposées.

Exemple 22 – Soit f la fonction définie sur $\mathbb{R}$ par $f(x) = \dfrac{e^x}{(e^x+1)^2}$.
Montrons que la fonction F définie par $F(x) = \dfrac{-1}{e^x+1}$ est une primitive de f sur $\mathbb{R}$:
F est définie sur $\mathbb{R}$, dérivons cette fonction qui est du type $\dfrac{1}{u}$ avec $u = -(e^x+1)$ et $u' = -e^x$.
Sa dérivée est $\dfrac{-u'}{u^2}$, donc $F'(x) = \dfrac{-(e^x+1)'}{(e^x+1)^2} = \dfrac{-(-e^x)}{(e^x+1)^2} = \dfrac{e^x}{(e^x+1)^2} = f(x)$.
F donc est bien une primitive de f sur $\mathbb{R}$.

Déterminons la primitive de f sur $\mathbb{R}$ qui prend la valeur -1 en 0.
Toutes les primitives de f s'écrivent $F_k(x) = F(x) + k = \dfrac{-1}{e^x+1} + k$.
Pour avoir $F_k(0) = -1$, il faut et il suffit d'avoir $\dfrac{-1}{e^0+1} + k = -1 \Longleftrightarrow k = -1 + \dfrac{1}{2} = \dfrac{-1}{2}$.
Conclusion : la fonction $F_{\frac{-1}{2}}(x) = \dfrac{-1}{e^x+1} + \dfrac{-1}{2} = \dfrac{-2-(e^x+1)}{2(e^x+1)} = \dfrac{-3-e^x}{2(e^x+1)}$ est la primitive de f sur $\mathbb{R}$ qui prend la valeur -1 en 0.

Propriété 3.4 Soient F et G des primitives de deux fonctions f et g définies sur un intervalle I et λ un réel quelconque.
- La fonction $F + G$ est une primitive de $f + g$ sur I
- La fonction λF est une primitive de λf sur I

D'une façon générale, $\forall(\lambda, \mu) \in \mathbb{R}^2, \lambda F + \mu G$ est une primitive de $\lambda f + \mu g$ sur I.

Démonstration C'est immédiat connaissant les propriétés de la dérivation :
$(\lambda F + \mu G)' = \lambda F' + \mu G' = \lambda f + \mu g$.

1.b. Recherche de primitives

Le tableau qui suit donne des fonctions f continues sur un intervalle I ou un intervalle $J \subset I$.
Les fonctions F sont les primitives des fonctions f, k étant une constante additive quelconque.

$f(x) =$	$F(x) =$	I
$\alpha \in \mathbb{R}$	$\alpha x + k$	$\mathbb{R}$
x^n avec $n \in \mathbb{Z} - \{0, -1\}$	$\frac{x^{n+1}}{n+1} + k$	$\mathbb{R}$ si $n \geqslant 1$, $\mathbb{R}^*$ si $n \leqslant -2$
$x^{-1} = \frac{1}{x}$	$\ln x + k$	$\mathbb{R}^{*+}$
$\frac{1}{\sqrt{x}}$	$2\sqrt{x} + k$	$\mathbb{R}^{*+}$
$\cos x$	$\sin x + k$	$\mathbb{R}$
$\sin x$	$-\cos x + k$	$\mathbb{R}$
$\frac{1}{\cos^2 x}$	$\tan x + k$	$\mathbb{R} - \{\frac{\pi}{2}\ [\pi]\}$
e^x	$e^x + k$	$\mathbb{R}$

Remarques :
- D'après la propriété 3.4, on en déduit les primitives de toute fonction polynôme puisqu'il s'agit d'une combinaison linéaire de fonctions $x \longmapsto x^n$ dont on connait une primitive pour $n \in \mathbb{N}$. Par exemple les primitives du trinôme $p(x) = ax^2 + bx + c$ sont les polynômes $P(x) = \frac{a}{3}x^3 + \frac{b}{2}x^2 + \frac{c}{1}x + d$ avec $d \in \mathbb{R}$. Plus généralement, les primitives d'un polynôme de degré n (le coefficient a_n du terme dominant doit être non nul) sont des polynômes de degré $n+1$ (le coefficient du terme dominant est alors $\frac{a_n}{n+1}$).
- Une primitive de x^n est $\frac{x^{n+1}}{n+1}$; cette propriété est valable pour $n \in \mathbb{Z} - \{-1\}$ (pour $n = 0$ elle se réduit au cas particulier de la primitive de $\alpha = 1$) et aussi pour un réel $a \neq -1$ quelconque car, en notant x^a sous la forme $e^{a \ln x}$ et en utilisant la dérivée d'une fonction composée, on montre que la dérivée de $\frac{e^{(a+1)\ln x}}{a+1}$ est $\frac{(a+1)e^{(a+1)\ln x}}{x(a+1)} = \frac{e^{(a+1)\ln x}}{x} = \frac{x^{a+1}}{x} = x^a$ ce qui signifie que $x^a = e^{a \ln x}$ a pour primitive $\frac{e^{(a+1)\ln x}}{a+1}$, c'est-à-dire $\frac{x^{a+1}}{a+1}$.

Propriété 3.5 (Composition) Soit g une fonction dérivable sur un intervalle I à valeurs dans un intervalle J telle que sa dérivée soit continue sur I.
Soient f une fonction continue sur J et F une primitive de f sur J.
La fonction $h : x \longmapsto g'(x) f \circ g(x)$ admet pour primitive sur I la fonction $H : x \longmapsto F \circ g(x) + k$.

Démonstration D'après le théorème de dérivation d'une fonction composée, la fonction H est dérivable sur I, la dérivée H' étant définie par $H'(x) = (F \circ g)'(x) =) g'(x) F' \circ g(x) = h(x)$.

Remarques :

- Si la fonction g est affine, c'est-à-dire $g(x) = ax + b$ avec $a \neq 0$, on a $g'(x) = a$. La fonction $h : x \longmapsto af(ax+b)$ admet pour primitive la fonction $H : x \longmapsto F(ax+b)$. En utilisant la propriété 3.4, la fonction $\frac{h}{a}$ admet pour primitive la fonction $\frac{H}{a}$, soit $f(ax+b)$ admet pour primitive $\frac{F(ax+b)}{a}$. Par exemple, la fonction $x \longmapsto \cos(ax+b)$ admet pour primitive $\frac{\sin(ax+b)}{a}$.
- En notant u la fonction g et v la fonction F dont la dérivée est $v' = f$, on obtient $u' \times v'(u(x))$ admet pour primitive $v(u(x))$. On reconnait mieux peut-être sous cette forme, les fonctions qui interviennent dans cette propriété.

Le tableau qui suit donne des fonctions u ainsi que la condition sur $u(x)$ pour que les fonctions $f = u' \times v'(u(x))$ admettent pour primitives $v(u(x)) + k$.

$f(x) =$	$F(x) =$	Condition
$u'u^n$ avec $n \in \mathbb{Z} - \{0, -1\}$	$\frac{u^{n+1}}{n+1} + k$	$\forall x \in I, u(x) \neq 0$ si $n \leqslant -2$
$\frac{u'}{u}$	$\ln(u) + k$	$\forall x \in I, u(x) > 0$
$\frac{u'}{\sqrt{u}}$	$2\sqrt{u} + k$	$\forall x \in I, u(x) > 0$
$u' \cos u$	$\sin u + k$	$\forall x \in I$
$u' \sin u$	$-\cos u + k$	$\forall x \in I$
$\frac{u'}{\cos^2 u}$	$\tan u + k$	$\forall x \in I, u(x) \neq \frac{\pi}{2}\ [\pi]$
$u'e^u$	$e^u + k$	$\forall x \in I$

Exemple 23 – ➲ Soit f la fonction définie par $f(x) = \dfrac{1}{ax+b}$ avec $a \neq 0$.
En posant $u(x) = ax + b$, u est définie sur $\mathbb{R}$ et on a $u'(x) = a$ donc $f(x) = \dfrac{1}{a} \times \dfrac{u'(x)}{u(x)}$.
D'après le tableau ci-dessus $\dfrac{u'(x)}{u(x)}$ a pour primitive $\ln(u) + k$ à condition que $\forall x \in \mathbb{R}, u(x) > 0$.
Si $ax + b > 0 \Longleftrightarrow x > \frac{-b}{a}$ une primitive de f est donc $F(x) = \dfrac{\ln(ax+b)}{a} + k$.

➲ Soit g la fonction définie par $g(x) = xe^{x^2-1}$.
En posant $u(x) = x^2 - 1$, u est définie sur $\mathbb{R}$ et on a $u'(x) = 2x$ donc $g(x) = \dfrac{1}{2} \times u'(x)e^u(x)$.
D'après le tableau ci-dessus $u'(x)e^u(x)$ a pour primitive $e^u + k$ sans condition sur $x \in \mathbb{R}$.
Une primitive de g est donc $G(x) = \dfrac{e^{x^2-1}}{2} + k$.

Noter que dans ces exemples, j'ai multiplié la primitive du tableau par un réel ($\frac{1}{a}$ pour F et $\frac{1}{2}$ pour G) sans multiplier la constante k par ce réel. En multipliant la constante k par ce réel, j'aurai trouvé une autre constante ($\frac{k}{a}$ pour F et $\frac{k}{2}$ pour G) que je peux bien renommer k puisqu'elle est quelconque.

➲ Soit h la fonction définie par $h(x) = \sin^2(x)$.
On sait que $\sin^2(x) = \dfrac{1 - \cos(2x)}{2} = \dfrac{1}{2} - \dfrac{\cos(2x)}{2}$ (formules de linéarisation).
En posant $u(x) = 2x$, u est définie sur $\mathbb{R}$ et on a $u'(x) = 2$
Une primitive de $2\cos(2x)$ est $\sin(2x)$ donc une primitive de $\dfrac{\cos(2x)}{2} = \dfrac{1}{4} \times 2\cos(2x)$ est $\dfrac{\sin(2x)}{4}$.
Finalement une primitive de $\dfrac{1}{2} - \dfrac{\cos(2x)}{2}$ est $\dfrac{x}{2} - \dfrac{\sin(2x)}{4} = \dfrac{2x - \sin(2x)}{4}$.

2. Équations différentielles

DÉFINITION 3.2 Une *équation différentielle* est une égalité dans laquelle interviennent une fonction, la ou les dérivées successives de cette fonction. Résoudre une équation différentielle consiste à déterminer l'ensemble des fonctions qui vérifient cette égalité.

Remarques :

- Déterminer les primitives d'une fonction f c'est résoudre l'équation différentielle $y' = f(x)$. Un autre exemple déjà rencontré d'équation différentielle est la définition de la fonction exponentielle comme l'unique solution de l'équation différentielle $y' = y$ telle que $y(0) = 1$. En physique aussi, il est fait un usage fréquent d'équations différentielles, notamment pour caractériser la vitesse et l'accélération d'un solide en mouvement par les équations $v = \frac{dx}{dt}$, soit $v = x'$ (noté $\dot{x}$ par les physiciens) et $a = \frac{d^2x}{dt^2}$, soit $a = x''$ (noté $\ddot{x}$).
- Si la seule fonction dérivée est utilisée conjointement en combinaison linéaire avec la fonction elle-même, on parle d'*équation différentielle linéaire du 1er ordre* ; si on utilise la dérivée seconde $f'' = (f')'$ (dérivée de la dérivée), on parle d'équation différentielle du 2e ordre ; etc.

2.a. L'équation différentielle *y'=ay*

PROPRIÉTÉ 3.6 L'équation $y' = ay$ admet pour solution les fonctions $f_k : x \longmapsto ke^{ax}$ où k est un réel quelconque. La solution est unique si elle doit satisfaire une condition de type $f_k(x_0) = y_0$.

DÉMONSTRATION Soit f une fonction satisfaisant l'équation différentielle $y' = ay$, on a donc $f'(x) = af(x)$ pour tout $x \in \mathbb{R}$.
Posons $g(x) = \frac{f(x)}{e^{ax}}$ et dérivons g : $g'(x) = \frac{f'(x)e^{ax} - f(x)(e^{ax})'}{(e^{ax})^2} = \frac{e^{ax}(f'(x) - af(x))}{(e^{ax})^2} = \frac{f'(x) - af(x)}{e^{ax}}$.
Comme $f'(x) = af(x)$, on en déduit $g'(x) = 0$ et par conséquent $g(x) = k$, une constante réelle.
Finalement $\frac{f(x)}{e^{ax}} = k \iff f(x) = ke^{ax}$.
Si $f_k(x_0) = y_0$ alors $ke^{ax_0} = y_0 \iff k = \frac{y_0}{e^{ax_0}}$. Par conséquent k et f_k sont bien uniques.

EXEMPLE 24 – ➲ Résolvons l'équation différentielle $y' - 2y = 0$ avec la condition $y(1) = 3$.
$y' - 2y = 0 \iff y' = 2y$ donc les solutions sont les fonctions $f_k(x) = ke^{2x}$.
Comme on souhaite avoir $y(1) = 3$, k doit vérifier l'égalité $ke^{2\times 1} = 3$ soit $k = \frac{3}{e^2}$.
Finalement, la fonction cherchée est $x \longmapsto \frac{3}{e^2}e^{2x} = 3e^{2(x-1)}$.

➲ Pour déterminer l'âge d'un résidus organique on mesure la proportion de Carbone 14 dans le résidus. Sachant qu'étant en vie la proportion était connue et constante et que, par désintégration radioactive, le corps mort perd de cette substance proportionnellement à ce qu'elle en contient à une date t. Autrement dit, la masse $m(t)$ résiduelle au temps t obéit à l'équation différentielle $m' = am$ sachant qu'à l'instant de la mort elle vaut une certaine valeur $m(0) = m_0$.
Les solutions de cette équation sont les fonctions $m_k(t) = ke^{at}$ avec k vérifiant $m_k(0) = k = m_0$.
On a donc une fonction $m(t) = m_0e^{at}$.

La constante a est déterminée par notre connaissance de la désintégration du Carbone 14 qui perd $1,24\%$ de sa masse en un siècle.
On en déduit $m(100) = m_0e^{100a} = \frac{100-1,24}{100}m_0 = 0,9876m_0$
Calculons a : $e^{100a} = 0,9876 \iff 100a = \ln 0,9876 \iff a = \frac{\ln 0,9876}{100} \approx -0,000125$.

Si un corps trouvé aujourd'hui ne contient que 14% de sa masse en Carbone 14 quel est son âge ?
On a approximativement $m(t) = m_0e^{-0,0439t}$ et on a aussi $m(t) = \frac{14}{100}m_0$, on en déduit l'âge t :
$m_0e^{-0,000125t} = \frac{14}{100}m_0 \iff e^{-0,000125t} = 0,14 \iff -0,000125t = \ln 0,14 \iff t = \frac{\ln 0,14}{-0,000125} \approx 15\,757$.
Le corps trouvé aujourd'hui à environ 15 800 ans.

2.b. L'équation différentielle *y'=ay+b*

PROPRIÉTÉ 3.7 L'équation $y' = ay + b$ admet pour solution les fonctions $f_k : x \longmapsto ke^{ax} - \frac{b}{a}$ où k est un réel quelconque. La solution est unique si elle doit satisfaire la condition $f_k(x_0) = y_0$.

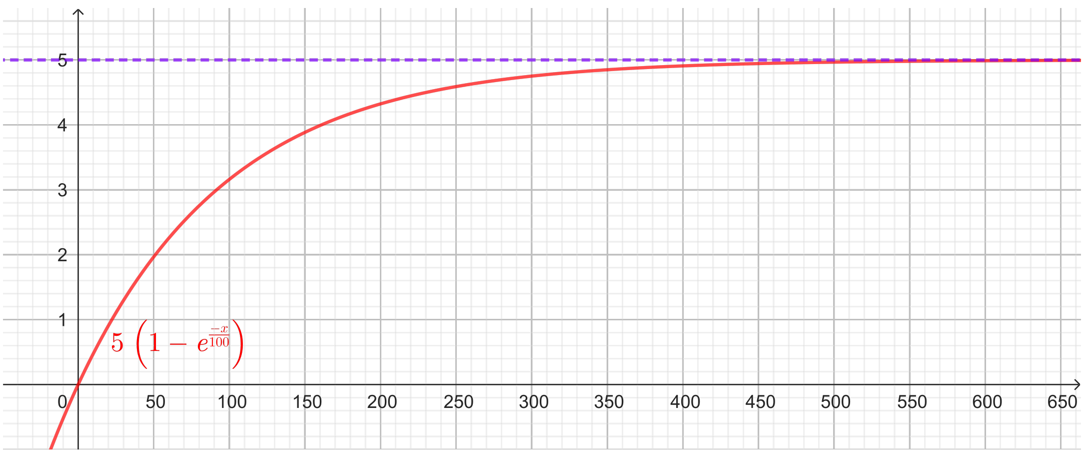
DÉMONSTRATION Soit f une fonction satisfaisant l'équation différentielle $y' = ay + b$.
On a donc $f'(x) = af(x) + b = a\left(f(x) + \frac{b}{a}\right)$ pour tout $x \in \mathbb{R}$.
Posons $g(x) = f(x) + \frac{b}{a}$. Ainsi nous avons $f'(x) = ag(x)$ et $g'(x) = f'(x)$.
On en déduit que la fonction g vérifie l'équation différentielle $g'(x) = ag(x)$.
D'après la propriété 3.6 précédente, les solutions s'écrivent $g_k(x) = ke^{ax}$.
Par conséquent $f_k(x) = ke^{ax} - \frac{b}{a}$ avec k réel quelconque.
Bien sûr, dans le cas où $f_k(x_0) = y_0$ on obtient une valeur unique $ke^{ax_0} - \frac{b}{a} = y_0 \iff k = \frac{y_0 + \frac{b}{a}}{e^{ax_0}}$.

EXEMPLE 25 – ➲ Résolvons l'équation différentielle $2y' + y = 1$ avec la condition $y(0) = 0$.
$2y' + y = 1 \iff y' = -\frac{1}{2}y + \frac{1}{2}$ donc les solutions sont les fonctions $f_k(x) = ke^{-\frac{1}{2}x} - \frac{\frac{1}{2}}{\frac{-1}{2}} = ke^{-\frac{1}{2}x} + 1$.
Comme on souhaite avoir $y(0) = 0$, k doit vérifier l'égalité $ke^0 + 1 = 0$ soit $k = -1$.
Finalement, la fonction cherchée est $x \longmapsto 1 - e^{-\frac{x}{2}}$.

➲ Un circuit électrique est constitué d'un générateur de force électromotrice E, d'une résistance R et d'un condensateur de capacité C. On note $q(t)$ la charge du condensateur (en *farad*) à l'instant t (en secondes) et on sait que $q' + \frac{1}{RC}q = \frac{E}{R}$.
Déterminons la charge $q(t)$ sachant qu'à l'instant $t = 0$, la charge $q(0)$ est nulle.
La charge du condensateur vérifie une équation différentielle du type $y' = ay + b$ avec $a = \frac{-1}{RC}$ et $b = \frac{E}{R}$, la solution s'écrit donc $q_k(t) = ke^{\frac{-t}{RC}} - \frac{E \times RC}{R \times (-1)} = ke^{\frac{-t}{RC}} + EC$.
Comme $q(0) = 0$, on en déduit $q_k(0) = k + EC = 0 \iff k = -EC$.
La charge du condensateur s'écrit donc $q(t) = EC\left(1 - e^{\frac{-t}{RC}}\right)$.

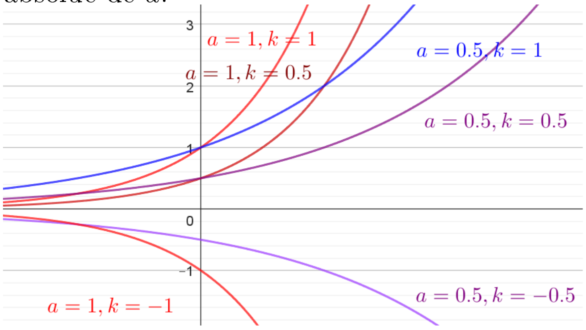
Quelle est la charge maximale q_{MAX} de ce condensateur ?
Il s'agit de la limite de $q(t)$ quand t tend vers $+\infty$.
On sait que $\lim\limits_{t \to +\infty} e^{\frac{-t}{RC}} = 0$ donc $q_{MAX} = \lim\limits_{t \to +\infty} q(t) = EC$.

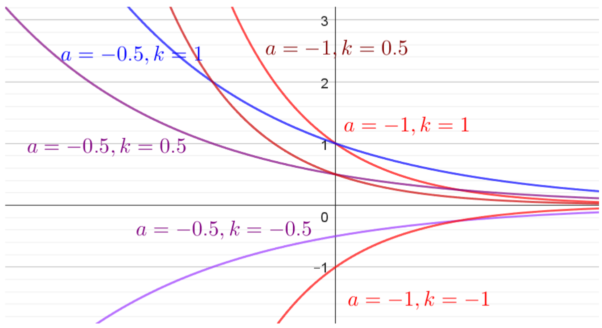
En notant $\tau = RC$, quelle fraction de la capacité maximale est atteinte pour $t = \tau$ et pour $t = 5\tau$?
$q(\tau) = EC\left(1 - e^{\frac{-RC}{RC}}\right) = EC\left(1 - \frac{1}{e}\right) \approx 0,632EC$, soit environ $63,2\%$ de q_{MAX}.
$q(5\tau) = EC\left(1 - e^{\frac{-5RC}{RC}}\right) = EC\left(1 - \frac{1}{e^5}\right) \approx 0,993EC$, soit environ $99,3\%$ de q_{MAX}.
Le graphique ci-dessous représente cette fonction de charge pour $C = 1F$, $R = 100\Omega$ et $E = 5V$.
On a alors $q(t) = 5\left(1 - e^{\frac{-t}{100}}\right)$ et, pour $t = 5\tau = 500s$ soit un peu plus de 8 minutes, on constate que la charge est quasiment au maximum.



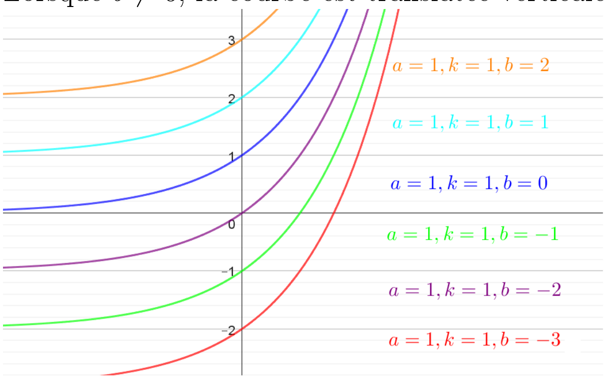
Remarques :

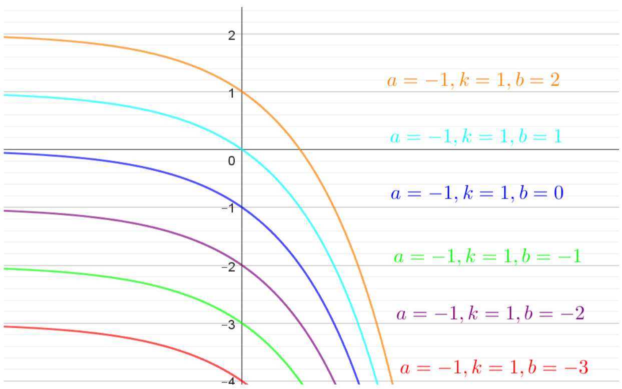
- Les fonctions solutions de l'équation différentielle $y' = ay + b$ s'écrivant $y = ke^{ax} - \frac{b}{a}$, leurs courbes ont toutes la même allure exponentielle. Lorsque $b = 0$, le signe de k définit le signe de y, le sens de variation est donné par le signe de a et la courbure est donnée par la valeur absolue de a.





Lorsque $b \neq 0$, la courbe est translatée verticalement.





- Pour l'équation différentielle $y' = ay + b$ cherchons une solution particulière constante : la fonction $f : x \longmapsto c$ est une solution si $f'(x) = af(x) + b \iff 0 = ac + b \iff c = \frac{-b}{a}$.
  Un fois cette solution particulière trouvée, une fonction g sera solution si $g'(x) = ag(x) + b$ et, par différence $g'(x) - f'(x) = a\left(g(x) - f(x)\right)$ autrement dit $(g - f)' = a(g - f)$.
  La fonction $h = g - f$ est solution de l'équation différentielle $y' = ay$ donc $h(x)$ s'écrit ke^{ax}.
  On en déduit $g(x) - f(x) = ke^{ax}$, or $f(x) = \frac{-b}{a}$ et finalement $g(x) = ke^{ax} + f(x) = ke^{ax} - \frac{b}{a}$.
  Cette méthode, plus générale, peut être utilisée dans un cadre plus large.

2.c. L'équation différentielle *y'=ay+f*

Propriété 3.8 (Méthode) Soit f une fonction.
L'équation $y' = ay + f$ peut, la plupart du temps, être résolue en deux temps :

1. On cherche une solution particulière g_0, la plus simple possible (constante, affine, etc.)
2. La solution générale g_k est alors telle que $g_k - g_0$ est solution de l'équation $y' = ay$.

Remarques :

- Selon la fonction f, une solution particulière constante n'est pas toujours possible.
  Par exemple l'équation différentielle $2y' - y = x$ n'en admet pas ; la fonction constante $f : x \longmapsto a$ n'est pas solution car $2f'(x) - f(x) = x \iff -a = x$ ce qui est faux.
  La fonction affine $f : x \longmapsto ax + b$, par contre, est solution si
  $2f'(x) - f(x) = x \iff -2a - b - ax = x \iff (a+1)x + (b - 2a) = 0$.
  Ce polynôme est nul si et seulement si $a = -1$ et $b = 2a = -2$.
  La fonction $f : x \longmapsto -2x - 1$ est donc une solution particulière de l'équation.
  La solution générale g étant telle que $2g'(x) - g(x) = x$ alors que $2f'(x) - f(x) = x$, il vient par soustraction $2\left(f'(x) - g'(x)\right) - \left(f(x) - g(x)\right) = 0$, soit $\left(f'(x) - g'(x)\right) = \frac{1}{2}\left(f(x) - g(x)\right)$.
  La fonction $h = f - g$ est donc solution de l'équation différentielle $y' = \frac{1}{2}y$ dont la solution est $y = ke^{\frac{1}{2}x}$. On en déduit $f(x) - g(x) = ke^{\frac{1}{2}x}$ d'où $g(x) = f(x) - ke^{\frac{1}{2}x} = -2x - 1 - ke^{\frac{1}{2}x}$.

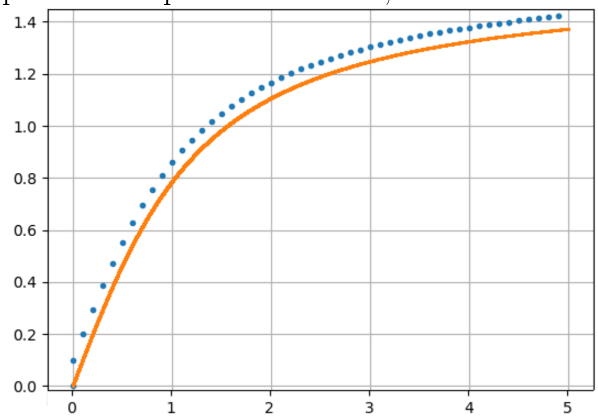
- Euler développa en 1768 une méthode graphique pour construire, point par point, la courbe de la solution F à l'équation différentielle $y' = f$ qui passe par un point donné. Cette méthode est à employer lorsqu'on ne connait pas de primitive à la fonction f.
  Principe de la méthode de Euler : La tangente T à un point d'abscisse x_0 de la courbe représentative de F a pour équation $y = F(x_0) + (x - x_0)F'(x_0)$. Autrement dit, pour tout réel h suffisamment petit on a $F(x_0 + h) \approx F(x_0) + hf(x_0)$. Cette méthode a aussi été utilisée par Euler pour construire des approximations de la courbe de la fonction exponentielle ne sachant que $f' = f$ et $f(0) = 1$ (voir le cours de 1$^{\text{re}}$), l'exemple qui suit la montre dans un autre cas.

Exemple 26 – On cherche à résoudre graphiquement l'équation différentielle $y' = \dfrac{1}{1+x^2}$, ne connaissant pas de primitive F pour la fonction $f : x \longmapsto \dfrac{1}{1+x^2}$, dans le cas particulier où $F(0) = 0$.
Pour être plus précis, on va calculer les coordonnées d'un nuage de points qu'on assimilera à la courbe représentant approximativement la fonction F sur l'intervalle $[0, 5]$.

Pour tout réel $h > 0$ suffisamment petit on a l'approximation $F(x_0 + h) \approx F(x_0) + hf(x_0)$.
Ainsi on va définir une suite de points de coordonnées (x_k, y_k) telle que, les abscisses x_k vérifient la relation de récurrence $x_{k+1} = x_k + h$ (soit $x_k = kh$) et les ordonnées se calculent de proche en proche par la relation $y_{k+1} = y_k = +hf(x_k)$ (on réitère à chaque étape l'approximation, en l'actualisant avec les valeurs du point précédent).
Pour l'intervalle $[0, 5]$, si on veut construire 50 points en plus du point $(0, 0)$ on prend $h = \frac{5-0}{50} = 0,1$.

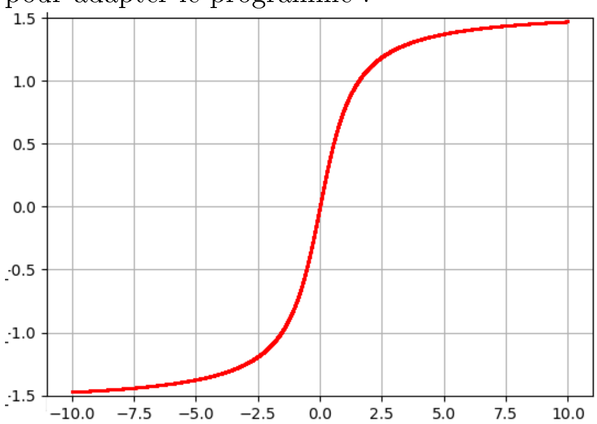
On définit les suites (x_k) et (y_k) par $\begin{cases} x_{k+1} = x_k + 0,1 \text{ avec } x_0 = 0, \text{ soit } x_k = 0,1k \\ y_{k+1} = y_k + \dfrac{h}{1+(x_k)^2} = y_k + \dfrac{0,1}{1+(0,1k)^2} \text{ avec } y_0 = 0 \end{cases}$

Programmons en Python le calcul de ces points et leur affichage (en bleu), l'ensemble des points obtenus étant assimilé à la courbe cherchée.
Pour obtenir une courbe plus proche de la réalité, je vais également afficher (en orange) les 50 000 points correspondants à $h = 0,0001$.




```
import matplotlib.pyplot as plt

def approx(h,x_max):
    X=[0]
    Y=[0]
    for i in range(int(x_max/h)):
        X.append(i*h)
        Y.append(Y[-1]+h/(1+(i*h)**2))
    return X,Y

X,Y=approx(0.1,5)
plt.plot(X[:],Y[:], linestyle='none', marker='o', markersize=3)

X,Y=approx(0.0001,5)
plt.plot(X[:],Y[:], linestyle='none', marker='o', markersize=1)

plt.grid()
plt.show()
```

Si je veux étendre la courbe à un intervalle plus grand, en incluant des valeurs négatives pour x, mettons qu'on souhaite la visualiser sur l'intervalle $[-10, 10]$, il n'y a que très peu de choses à changer pour adapter le programme :




```
import matplotlib.pyplot as plt

def approx(h,x_max):
    X=[0]
    Y=[0]
    for i in range(int(x_max/h)):
        X.append(i*h)
        Y.append(Y[-1]+h/(1+(i*h)**2))
    return X,Y

X,Y=approx(0.001,10)
plt.plot(X[:],Y[:], linestyle='none', color = 'red', marker='o', markersize=1)

X,Y=approx(-0.001,-10)
plt.plot(X[:],Y[:], linestyle='none', color = 'red', marker='o', markersize=1)

plt.grid()
plt.show()
```

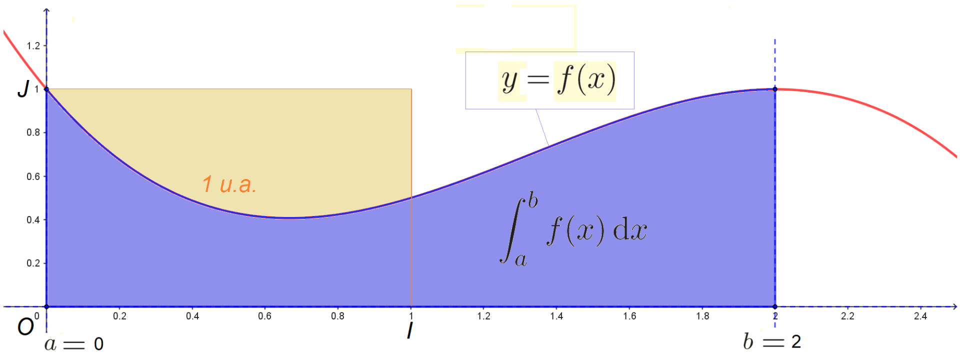
3. Calcul intégral

3.a. Intégrale d'une fonction continue et positive

Dans toute cette partie, on suppose le plan muni d'un repère orthogonal $(O,\ I,\ J)$ et, f étant une fonction positive définit sur un intervalle $[a, b]$, on s'intéresse à l'aire d'une portion du plan comprise entre 4 courbes :

- Les droites verticales d'équations $x = a$ et $x = b$, a et $b > a$ étant deux réels quelconques.
- L'axe des abscisses et la courbe d'équation $y = f(x)$ qui est située au dessus (f positive).

L'unité d'aire (notée $u.a.$) est l'aire du rectangle de côtés $[OI]$ et $[OJ]$, soit $\left\|\overrightarrow{OI}\right\| \times \left\|\overrightarrow{OJ}\right\|$.

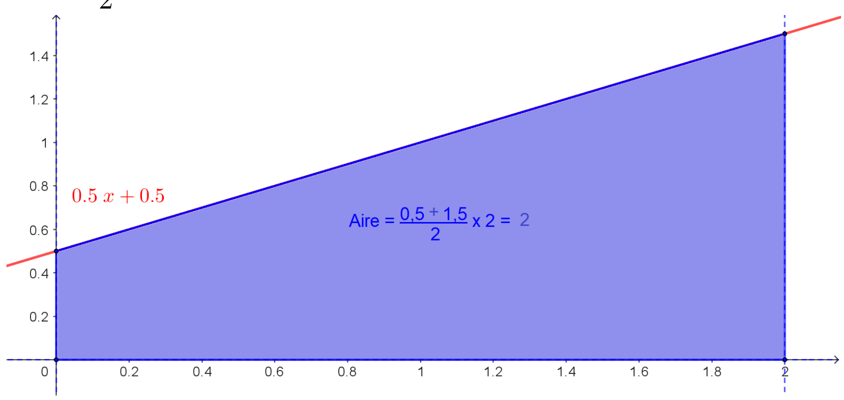


> **Définition 3.3 (Intégrale)** Soit f une fonction continue et positive sur un intervalle $[a, b]$. L'intégrale de a à b de la fonction f est l'aire (exprimée en $u.a.$) de la partie du plan contenant les points $M(x, y)$ tels que $\begin{cases} a \leqslant x \leqslant b \\ 0 \leqslant y \leqslant f(x) \end{cases}$, cette intégrale est notée $\displaystyle\int_a^b f(x)\,\mathrm{d}x$
> (lire « intégrale de a à b de $f(x)$ dx »).

Remarques :

- D'une façon évidente, pour toute fonction définie en a, on aura $\displaystyle\int_a^a f(x)\,\mathrm{d}x = 0\ \ u.a.$

  Lorsque la fonction f est une fonction affine, le calcul de l'aire (si $f(a) \geqslant 0$ et $f(b) \geqslant 0$) est aisé car il s'agit alors de l'aire d'un trapèze dont on dispose d'une formule :
  $\displaystyle\int_a^b f(x)\,\mathrm{d}x = \frac{f(a) + f(b)}{2} \times (b - a)$. Si $f(x) = 0,5x + 0,5$ alors
  $\displaystyle\int_0^2 f(x)\,\mathrm{d}x = \frac{f(0,5) + f(1,5)}{2} \times (2 - 0) = 0,5 + 1,5 = 2.$



- L'élément dx ajouté à l'intégrale signifie que la variable pour laquelle on intègre est x.
  Cette précision est nécessaire lorsqu'il y a d'autres variables en jeu dans la définition de $f(x)$.
  On peut comprendre dx comme la base d'un rectangle ($\mathrm{d}x = x_2 - x_1 = h$ lorsque $x_1 = x$ et $x_2 = x + h$) dont la hauteur serait $f(x_1)$ (soit $f(x)$ lorsque $x_1 = x$) ; l'intégrale est alors la valeur limite de la somme des aires de ces rectangles pour x variant de a à b.
  Ce calcul d'aire remontant à Archimède a souvent été étudié dans les classes antérieures.

PROPRIÉTÉ 3.9 (POSITIVE MONOTONE) Soit f une fonction continue, positive et monotone sur un intervalle $[a,b]$. La fonction F définie pour tout x de $[a,b]$ par $F(x)=\int_a^x f(t)\,dt$ est la primitive de f sur $[a,b]$ qui satisfait la condition $F(a)=0$.

DÉMONSTRATION Supposons f croissante sur $[a,b]$ (raisonnement analogue pour f décroissante).
$\boxed{h>0}$: Encadrons $F(x+h)-F(x)$ pour un x de $[a,b]$ tel que si h assez petit $x+h\in[a,b]$:
$F(x+h)-F(x)=\int_a^{x+h} f(t)\,\mathrm{d}t-\int_a^x f(t)\,\mathrm{d}t$ est l'aire sous la courbe de f entre x et $x+h$.
Comme f croit sur cet intervalle d'amplitude h, cette aire est comprise entre $hf(x)$ et $hf(x+h)$ d'où l'encadrement : $hf(x)\leqslant F(x+h)-F(x)\leqslant hf(x+h)$.
En divisant par $h>0$ on obtient $f(x)\leqslant \frac{F(x+h)-F(x)}{h}\leqslant f(x+h)$.
Comme f est continue, on a $\lim\limits_{h\to 0^+} f(x+h)=f(x)$.
D'après le théorème des gendarmes, on en déduit $\lim\limits_{h\to 0^+}\frac{F(x+h)-F(x)}{h}=f(x)$.

$\boxed{h<0}$: Faisons de même pour un x de $[a,b]$ tel que si $|h|<0$ assez petit $x+h\in[a,b]$:
$F(x)-F(x+h)=\int_a^x f(t)\,\mathrm{d}t-\int_a^{x+h} f(t)\,\mathrm{d}t$ est l'aire sous la courbe de f entre $x+h$ et x.
Sur cet intervalle d'amplitude $|h|=-h$, cette aire est comprise entre $-hf(x+h)$ et $-hf(x)$ d'où l'encadrement : $-hf(x+h)\leqslant F(x)-F(x+h)\leqslant -hf(x)$.
En divisant par $-h>0$ on obtient $f(x+h)\leqslant\frac{F(x)-F(x+h)}{-h}\leqslant f(x)$, soit
$f(x+h)\leqslant\frac{F(x+h)-F(x)}{h}\leqslant f(x)$.
Comme f est continue, on a $\lim\limits_{h\to 0^-} f(x+h)=f(x)$.
D'après le théorème des gendarmes, on en déduit $\lim\limits_{h\to 0^-}\frac{F(x+h)-F(x)}{h}=f(x)$.

Finalement, pour tout réel x de l'intervalle $[a,b]$, F est dérivable en x et $F'(x)=f(x)$.
Cela prouve que F est une primitive de f sur $[a,b]$.
Comme, de plus, $F(a)=\int_a^a f(t)\,\mathrm{d}t=0$, F est la primitive de f sur $[a,b]$ qui s'annule pour $x=a$.

Remarques :

- Dans l'expression $\int_a^x f(t)\,\mathrm{d}t$, les bornes de l'intervalle d'intégration contenant la variable x, il serait ambigu d'utiliser cette variable pour la fonction à intégrer (partie $f(t)\,dt$ de l'expression). La variable t utilisée alors est une variable *muette* qui peut être remplacée par n'importe quelle lettre, du moment qu'il n'y ait pas d'ambigüité. C'est exactement comme lorsqu'on écrit, par exemple, la somme des n premiers carrés $S_n=\sum\limits_{k=1}^{n}k^2$.
- Dans l'hypothèse d'une fonction qui est continue et positive mais pas monotone, cette propriété reste valable. L'intervalle $[a,b]$ est alors découpé en petits intervalles sur lesquels la fonction est monotone et la relation de Chasles (voir plus loin) permet la généralisation.
- Pour $f(x)=0,5x+0,5$ les primitives de f sur $[0,2]$ s'écrivent $0,25x^2+0,5x+c$, celle qui s'annule en 0 a pour coefficient $c=0$. D'après la propriété, on en déduit
$F(x)=\int_0^x f(t)\,\mathrm{d}t=0,25x^2+0,5x$ et, en particulier,
$F(2)=\int_0^2 f(t)\,\mathrm{d}t=0,25\times 2^2+0,5\times 2=1+1=2$ ce qui avait été obtenu précédemment.

PROPRIÉTÉ 3.10 (CALCUL (1)) Soit f une fonction continue et positive sur un intervalle $[a,b]$.
Si la fonction F est une primitive de f sur $[a,b]$ alors $\int_a^b f(t)\,\mathrm{d}t=[F(x)]_a^b=F(b)-F(a)$.

DÉMONSTRATION Pour $x \in [a,b]$, la fonction $\displaystyle\int_a^x f(t)\,\mathrm{d}t$ est la primitive de f qui s'annule en a.
Les primitives F_k de f sur $[a,b]$ s'écrivent donc $\displaystyle\int_a^x f(t)\,\mathrm{d}t + k$ avec k réel quelconque.
On en déduit que $F_k(b) - F_k(a) = \displaystyle\int_a^b f(t)\,\mathrm{d}t + k - \left(\int_a^a f(t)\,\mathrm{d}t + k\right) = \int_a^b f(t)\,\mathrm{d}t - \int_a^a f(t)\,\mathrm{d}t$.
Mais comme $\displaystyle\int_a^a f(t)\,\mathrm{d}t = 0$, on en déduit que $F_k(b) - F_k(a) = \displaystyle\int_a^b f(t)\,\mathrm{d}t$ et comme le réel k disparait, on peut utiliser n'importe quelle primitive F de f et ainsi $F(b) - F(a) = \displaystyle\int_a^b f(t)\,\mathrm{d}t$.

3.b. Intégrale d'une fonction continue de signe quelconque

PROPRIÉTÉ 3.11 (CONTINUE) Soit $I = [a,b]$ un intervalle.
Toute fonction continue sur I admet des primitives sur I.

DÉMONSTRATION Comme f est continue elle admet un minimum sur $[a,b]$, notons-le m.
Autrement dit $\forall x \in [a,b], f(x) \geqslant m \iff f(x) - m \geqslant 0$.
En posant $g(x) = f(x) - m$, la fonction g est continue et positive.
Elle admet donc la primitive $G_a(x) = \displaystyle\int_a^x g(t)\,\mathrm{d}t$.
La fonction F définie par $F(x) = G_a(x) + mx$ est une primitive de f sur $[a,b]$ car
$F'(x) = G'_a(x) + (mx)' = g(x) + m = f(x)$.

PROPRIÉTÉ 3.12 (CALCUL (2)) Soit f une fonction continue sur un intervalle $[a,b]$.
Si la fonction F est une primitive de f sur $[a,b]$ alors $\displaystyle\int_a^b f(t)\,\mathrm{d}t = [F(x)]_a^b = F(b) - F(a)$.

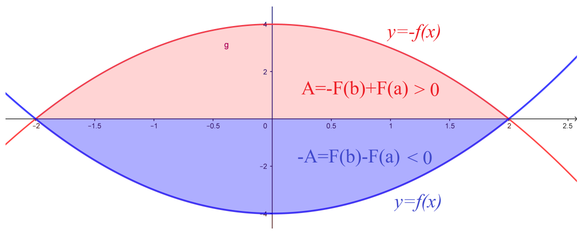
Remarques :

- Dans les deux dernières propriétés on a utilisé la notation $[F(x)]_a^b$ pour désigner $F(b) - F(a)$. Cette notation est courante car elle permet d'introduire la primitive de f utilisée pour le calcul de l'intégrale $\displaystyle\int_a^b f(t)\,\mathrm{d}t$.
  Par exemple on écrira $\displaystyle\int_0^1 e^{-t}\,\mathrm{d}t = \left[-e^{-x}\right]_0^1 = -e^{-1} - (-e^{-0}) = 1 - \frac{1}{e}$.
- On remarque ici que le calcul de l'intégrale est inchangé : du moment que l'on connait une primitive de f sur $[a,b]$ – peu importe la constante additive k (on prendra $k = 0$) – quelque soit le signe de f, on peut calculer l'intégrale de f entre a et b. Ce qui change, c'est que si il y a des parties où la fonction est négative, cette intégrale ne représente plus l'aire comprise entre la courbe de f et l'axe des abscisses. Par exemple $\displaystyle\int_0^\pi \cos(t)\,\mathrm{d}t = [\sin(x)]_0^\pi = \sin(\pi) - \sin(0) = 0$ (les deux parties de l'aire s'annulent, celle en-dessous de l'axe étant comptée négativement).

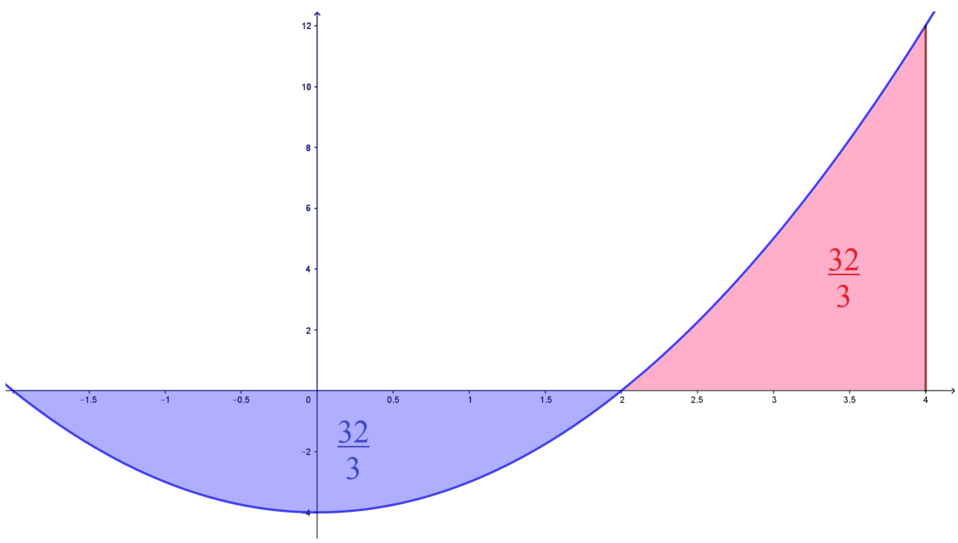
PROPRIÉTÉ 3.13 (AIRE) Soit f une fonction continue et *négative* sur un intervalle $[a,b]$.
L'intégrale de a à b de la fonction f est l'*opposée* de l'aire (exprimée en *u.a.*) de la partie du plan contenant les points $M(x,y)$ tels que $\begin{cases} a \leqslant x \leqslant b \\ f(x) \leqslant y \leqslant 0 \end{cases}$,

DÉMONSTRATION Rappelons-le, le plan étant muni du repère orthogonal $(O,\ I,\ J)$, les aires mesurées en *u.a.* (aire du rectangle de côtés $[OI]$ et $[OJ]$) sont des grandeurs positives.
La fonction f étant continue admet une primitive notée F et, d'après la propriété 3.4 de linéarité des primitives, la fonction $-f$ admet pour primitive la fonction $-F$.
Ainsi $\int_a^b -f(t)\,\mathrm{d}t = [-F(x)]_a^b = -(F(b)-F(a)) = -\int_a^b f(t)\,\mathrm{d}t$.
Or, f étant négative sur $[a,b]$, $-f$ est positive. L'intégrale entre a et b de $-f$ est l'aire sous la courbe de $-f$ qui, par symétrie par rapport à l'axe des abscisses est égale à l'aire au-dessus de la courbe de f. Celle-ci est donc l'opposée de l'intégrale entre a et b de f.

EXEMPLE 27 – Soit f la fonction définie par $f(x) = x^2 - 4$
La fonction f est négative sur l'intervalle $[-2, 2]$.
$\int_{-2}^{2} f(t)\,\mathrm{d}t = \left[\frac{x^3}{3} - 4x\right]_{-2}^{2} = \frac{2^3}{3} - 4\times 2 - \frac{(-2)^3}{3} + 4\times(-2) = \frac{16}{3} - 16 - = -\frac{32}{3} \approx -10,7$.
Cette intégrale est négative car la fonction est négative ; sa valeur absolue est égale à l'aire comprise entre la courbe de f et l'axe des abscisses.



Pour quelle valeur de $x > 0$ a t-on $\int_{-2}^{x} f(t)\,\mathrm{d}t = 0$?
Autrement dit, pour quelle valeur de $x > 0$ a t-on $\int_{2}^{x} f(t)\,\mathrm{d}t = -\int_{-2}^{2} f(t)\,\mathrm{d}t = \frac{32}{3}$?
Il faut résoudre l'équation $\int_{2}^{x} f(t)\,\mathrm{d}t = \frac{32}{3} \iff \left[\frac{x^3}{3} - 4x\right]_2^x = \frac{32}{3} \iff \frac{x^3}{3} - 4x - \frac{2^3}{3} + 4\times 2 = \frac{32}{3}$
Cette équation est du 3e degré $\frac{x^3}{3} - 4x + \frac{-8-32+24}{3} = 0 \iff x^3 - 12x - 16 = 0 \iff x = -2$ ou $x = 4$.
Ainsi, on a $\int_{-2}^{4} f(t)\,\mathrm{d}t = 0$ (on passage, on obtient également $\int_{2}^{-2} f(t)\,\mathrm{d}t = -\int_{-2}^{2} f(t)\,\mathrm{d}t$).
L'aire en-dessous de la courbe de f entre 2 et 4 égale l'aire au-dessus de la courbe de f entre -2 et 2.
Entre -2 et 4, l'aire comprise entre $\mathscr{C}_f$ et l'axe (Ox) est $-\int_{-2}^{2} f(t)\,\mathrm{d}t + \int_{2}^{4} f(t)\,\mathrm{d}t = \frac{64}{3}$.

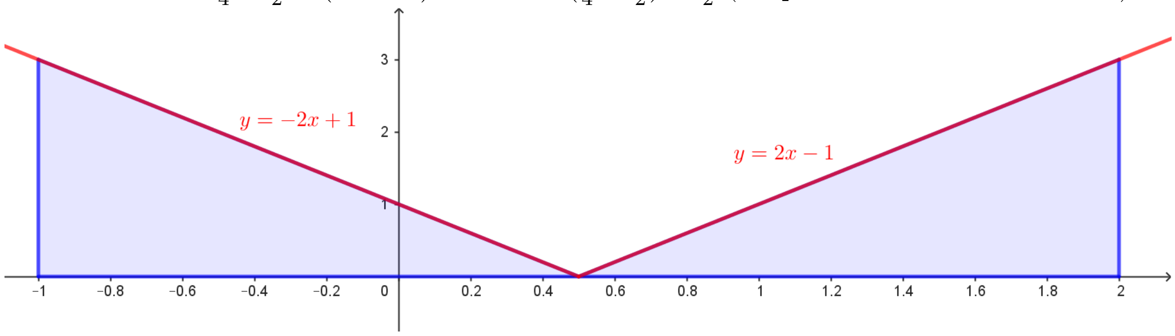


3.c. Propriétés algébriques

PROPRIÉTÉ 3.14 (RELATION DE CHASLES) Soit f une fonction continue sur un intervalle $[a, b]$.
Pour tout réel $c \in [a, b]$ on a $\displaystyle\int_a^b f(t)\,\mathrm{d}t = \int_a^c f(t)\,\mathrm{d}t + \int_c^b f(t)\,\mathrm{d}t$.

Remarques :

- Cette propriété se justifie à l'aide de la propriété 3.12 car :
  — $\displaystyle\int_a^b f(t)\,\mathrm{d}t = [F(x)]_a^b = F(b) - F(a)$
  — $\displaystyle\int_a^c f(t)\,\mathrm{d}t + \int_c^b f(t)\,\mathrm{d}t = [F(x)]_a^c + [F(x)]_c^b = F(c) - F(a) + F(b) - F(c) = F(b) - F(a)$.
- On peut généraliser le calcul intégral sur un intervalle I où la fonction f est continue : si deux nombres a et b appartiennent à I (sans avoir nécessairement $a < b$) : $\displaystyle\int_b^a f(t)\,\mathrm{d}t = -\int_a^b f(t)\,\mathrm{d}t$
  En effet, la relation Chasles sera satisfaite si on a $\displaystyle\int_a^b f(t)\,\mathrm{d}t + \int_b^a f(t)\,\mathrm{d}t = \int_a^a f(t)\,\mathrm{d}t = 0$.
- Avec cette propriété, les fonctions définies par morceaux sont facilement intégrables.
  Par exemple $f : x \longmapsto |2x - 1|$ est définie par $\begin{cases} f(x) = -2x + 1 \text{ pour } x \leqslant \dfrac{1}{2} \\ f(x) = 2x - 1 \text{ pour } x \geqslant \dfrac{1}{2} \end{cases}$
  Par conséquent $\displaystyle I = \int_{-1}^2 f(t)\,\mathrm{d}t = \int_{-1}^{\frac{1}{2}} -2t + 1\,\mathrm{d}t + \int_{\frac{1}{2}}^2 2t - 1\,\mathrm{d}t = \left[-x^2 + x\right]_{-1}^{\frac{1}{2}} + \left[x^2 - x\right]_{\frac{1}{2}}^2$
  On obtient $I = -\frac{1}{4} + \frac{1}{2} - (-1 - 1) + 4 - 2 - (\frac{1}{4} - \frac{1}{2}) = \frac{9}{2}$ (ce qui est facilement vérifiable)



PROPRIÉTÉ 3.15 (LINÉARITÉ) Soient f et g deux fonctions continues sur un intervalle $[a, b]$.

- $\displaystyle\int_a^b (f(t) + g(t))\,\mathrm{d}t = \int_a^b f(t)\,\mathrm{d}t + \int_a^b g(t)\,\mathrm{d}t$.
- $\displaystyle\forall \lambda \in \mathbb{R}, \int_a^b \lambda f(t)\,\mathrm{d}t = \lambda \int_a^b f(t)\,\mathrm{d}t$.

D'une façon générale, $\displaystyle\forall (\lambda, \mu) \in \mathbb{R}^2, \int_a^b (\lambda f(t) + \mu g(t))\,\mathrm{d}t = \lambda \int_a^b f(t)\,\mathrm{d}t + \mu \int_a^b g(t)\,\mathrm{d}t$.

DÉMONSTRATION Ces propriétés découlent de la propriété 3.4 : si F et G sont des primitives de f et g alors $F + G$ est une primitive de $f + g$ et, pour tout réel λ, λF est une primitive de λf.
Il ne reste plus qu'à appliquer la propriété 3.12 et les propriétés de l'addition et de la multiplication :
$\displaystyle\int_a^b (f(t) + g(t))\,\mathrm{d}t = [F(x) + G(x)]_a^b = [F(x)]_a^b + [G(x)]_a^b = \int_a^b f(t)\,\mathrm{d}t + \int_a^b g(t)\,\mathrm{d}t$
$\displaystyle\int_a^b \lambda f(t)\,\mathrm{d}t = [\lambda F(x)]_a^b = \lambda\,[F(x)]_a^b = \lambda \int_a^b f(t)\,\mathrm{d}t$

Exemple 28 – Soit f la fonction définie sur $[0,1]$ par $f(x)=\dfrac{x^3}{1+x^2}$.
On souhaite calculer $\displaystyle\int_0^1 f(t)\,\mathrm{d}t$ mais on ne connait pas de primitive de f.
On écrit alors $f(x)=\dfrac{x^3+x-x}{1+x^2}=\dfrac{x^3+x}{1+x^2}-\dfrac{x}{1+x^2}=\dfrac{x(x^2+1)}{1+x^2}-\dfrac{x}{1+x^2}=x-\dfrac{x}{1+x^2}$.
On remarque que $-\dfrac{x}{1+x^2}$ est de la forme $\dfrac{-1}{2}\dfrac{u'}{u}$ avec $u=x^2+1$ et $u'=2x$
Comme u est positive sur $[0,1]$, on sait que $\dfrac{-1}{2}\ln u$ est une primitive de $\dfrac{-1}{2}\dfrac{u'}{u}$.
On peut donc calculer notre intégrale :
$$\int_0^1 f(t)\,\mathrm{d}t=\left[\frac{x^2}{2}\right]_0^1+\left[\frac{-\ln(1+x^2)}{2}\right]_0^1=\frac{1}{2}-\frac{\ln 2}{2}+\frac{\ln 1}{2}=\frac{1-\ln 2}{2}\approx 0,1534264$$

Propriété 3.16 (Inégalités) Soient f et g deux fonctions continues sur un intervalle $[a,b]$.

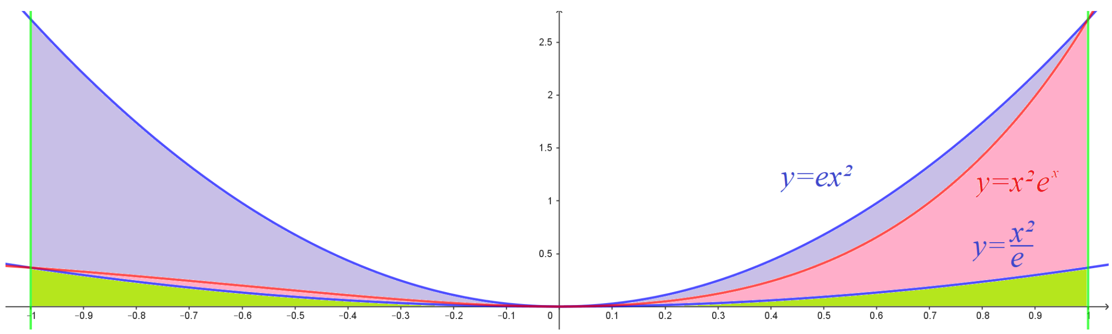
- $\forall x\in[a,b], f(x)\geqslant 0\Longrightarrow \displaystyle\int_a^b f(t)\,\mathrm{d}t\geqslant 0$
- $\forall x\in[a,b], f(x)\leqslant 0\Longrightarrow \displaystyle\int_a^b f(t)\,\mathrm{d}t\leqslant 0$
- $\forall x\in[a,b], f(x)\geqslant g(x)\Longrightarrow \displaystyle\int_a^b f(t)\,\mathrm{d}t\geqslant \int_a^b g(t)\,\mathrm{d}t$

Démonstration ➲ La 1re proposition est justifiée par la définition 3.3.
Si l'on part de la propriété 3.9 caractéristique de l'intégrale, la fonction F définie pour tout x de $[a,b]$ par $F(x)=\displaystyle\int_a^x f(t)\,\mathrm{d}t$ est dérivable : $F'(x)=f(x)$ et $f(x)\geqslant 0\Longrightarrow F'(x)\geqslant 0$.
F est croissante donc $x\geqslant a\Longrightarrow F(x)\geqslant F(a)=0$ d'où $\displaystyle\int_a^b f(t)\,\mathrm{d}t=F(b)\geqslant 0$.

➲ La 2e proposition repose sur le même type d'argument : $f(x)\leqslant 0\Longrightarrow F'(x)\leqslant 0$. La fonction F est décroissante et donc $F(x)\leqslant 0$, soit $\forall x\in[a,b]$, $\displaystyle\int_a^x f(t)\,\mathrm{d}t=F(x)\leqslant 0$ d'où $\displaystyle\int_a^b f(t)\,\mathrm{d}t=F(b)\leqslant 0$.

➲ 3e proposition : en posant $h(x)=f(x)-g(x)$, on a
$f(x)\geqslant g(x)\Longleftrightarrow f(x)-g(x)\geqslant 0\Longleftrightarrow h(x)\geqslant 0$ et on est alors ramené à la 1re proposition :
$\displaystyle\int_a^b h(t)\,\mathrm{d}t\geqslant 0\Longleftrightarrow\int_a^b f(t)-g(t)\,\mathrm{d}t\geqslant 0$ et la propriété 5.8 de linéarité permet de conclure.

Exemple 29 – Soit f la fonction définie sur $[-1,1]$ par $f(x)=x^2e^x$.
On ne connait pas de primitive de f aussi on souhaite encadrer $\displaystyle\int_{-1}^1 f(t)\,\mathrm{d}t$.
On sait que, pour $-1\leqslant x\leqslant 1$ on a $e^{-1}\leqslant e^x\leqslant e^1\Longleftrightarrow \frac{1}{e}\leqslant e^x\leqslant e$.
En multipliant par $x^2\geqslant 0$ on obtient l'encadrement $\frac{x^2}{e}\leqslant f(x)\leqslant ex^2$.
On en déduit que $\displaystyle\int_{-1}^1\frac{t^2}{e}\,\mathrm{d}t\leqslant\int_{-1}^1 f(t)\,\mathrm{d}t\leqslant\int_{-1}^1 et^2\,\mathrm{d}t\Longleftrightarrow\left[\frac{t^3}{3e}\right]_{-1}^1\leqslant\int_{-1}^1 f(t)\,\mathrm{d}t\leqslant\left[\frac{et^3}{3}\right]_{-1}^1$.
On en déduit $\frac{2}{3e}\leqslant\displaystyle\int_{-1}^1 f(t)\,\mathrm{d}t\leqslant\frac{2e}{3}$, soit approximativement $0,3\leqslant\displaystyle\int_{-1}^1 f(t)\,\mathrm{d}t\leqslant 1,8$.



3.d. Intégration par parties

Propriété 3.17 Soient f et g deux fonctions dérivables sur un intervalle $[a,b]$, telles f' et g' soient continues sur $[a,b]$. On a $\int_a^b f(t)g'(t)\,\mathrm{d}t = [f(t)g(t)]_a^b - \int_a^b f'(t)g(t)\,\mathrm{d}t$

Démonstration On sait que, par produit de fonctions dérivables, la fonction fg est dérivable sur $[a,b]$ et $(fg)' = f'g + fg'$. Par conséquent $\forall x \in [a,b], f(x)g'(x) = (fg)'(x) - f'g(x)$.
Par somme et produit de fonctions continues, les expressions dans chacun des membres de cette égalité sont continues. Elles admettent donc des primitives et sont intégrables. Par unicité de l'intégrale, les intégrales sur $[a,b]$ des deux membres sont donc égales :

$$\begin{aligned}\int_a^b f(t)g'(t)\,\mathrm{d}t &= \int_a^b \left[(f(t)g(t))' - f'(t)g(t)\right]\,\mathrm{d}t\\ &= \int_a^b (f(t)g(t))'\,\mathrm{d}t - \int_a^b f'(t)g(t)\,\mathrm{d}t\\ &= [f(t)g(t)]_a^b - \int_a^b f'(t)g(t)\,\mathrm{d}t\end{aligned}$$

Exemple 30 – Reprenons la fonction de l'exercice 29 et calculons cette fois $\int_{-1}^1 t^2e^t\,\mathrm{d}t$.
L'intégration par parties se fait ici en deux étapes successives :
$\boxed{\text{Étape 1}}$ Posons $u'(x) = e^x$ et $v(x) = x^2$.
La fonction $u : x \longmapsto e^x$ est dérivable, sa dérivée $u'(x) = e^x$ continue sur $[-1,1] \subset \mathbb{R}$.
La fonction $v : x \longmapsto x^2$ est dérivable, sa dérivée $v'(x) = 2x$ continue sur $[-1,1] \subset \mathbb{R}$.
On a $\int_{-1}^1 u'(t)v(t)\,\mathrm{d}t = [u(t)v(t)]_{-1}^1 - \int_{-1}^1 u(t)v'(t)\,\mathrm{d}t$, c'est-à-dire $\int_{-1}^1 t^2e^t\,\mathrm{d}t = \left[t^2e^t\right]_{-1}^1 - 2\int_{-1}^1 te^t\,\mathrm{d}t$.
$\boxed{\text{Étape 2}}$ Posons $u'(x) = e^x$ et $v(x) = x$.
Les fonctions u et v définies par $u(x) = e^x$ et $v(x) = x$ sont dérivables et leurs dérivées u' et v' définies par $u'(x) = e^x$ et $v'(x) = 1$ sont continues sur $[-1,1] \subset \mathbb{R}$. Ainsi :

$$\begin{aligned}\int_{-1}^1 t^2e^t\,\mathrm{d}t &= \left[t^2e^t\right]_{-1}^1 - 2\left[\left[te^t\right]_{-1}^1 - \int_{-1}^1 e^t\,\mathrm{d}t\right]\\ &= \left[t^2e^t\right]_{-1}^1 - 2\left[\left[te^t\right]_{-1}^1 - \left[e^t\right]_{-1}^1\right]\\ &= e - \frac{1}{e} - 2\left[\left(e + \frac{1}{e}\right) - \left(e - \frac{1}{e}\right)\right]\\ &= e - \frac{1}{e} \approx 0,87888\end{aligned}$$

Remarques :

- L'intégration par partie est utile lorsqu'on ne dispose pas de primitive de la fonction à intégrer comme c'est le cas dans notre exemple et d'une façon générale lorsque cette fonction à intégrer est le produit d'un polynôme (facile à dériver) par une fonction dont on connait une primitive.
- À l'inverse on peut utiliser la primitive du polynôme et dériver la fonction qui constitue l'autre facteur. Par exemple, pour déterminer une primitive de la fonction ln,
on calcule $I(x) = \int_1^x \ln t\,\mathrm{d}t$ en posant $v(t) = \ln t$ et $u'(t) = 1$, on a $v'(t) = \frac{1}{t}$ et $u(t) = x$.
On a ainsi $I(x) = [t\ln t]_1^x - \int_1^x t \times \frac{1}{t}\,\mathrm{d}t = x\ln x - 0 - [t]_1^x = x\ln x - x - 1$.
Cette fonction $I(x)$ est une primitive de $\ln x$.
Généralement on lui préfère celle qui n'a pas de constante $x\ln x - x$.
- On retiendra que si deux fonctions sont égales sur un intervalle leurs intégrales sur le même intervalle le sont aussi (la réciproque de cette propriété est généralement fausse).

3.e. Calculs de grandeurs

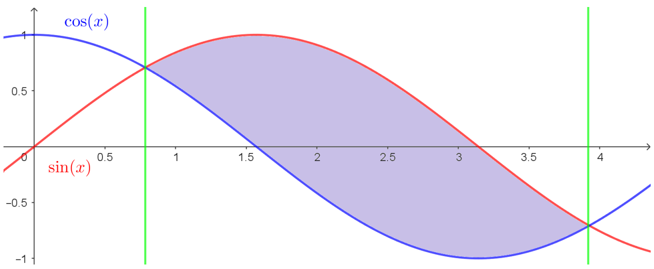
Propriété 3.18 (Calcul d'aire) Soient f et g deux fonctions continues sur un intervalle $[a,b]$ où $\forall x \in [a,b], f(x) \leqslant g(x)$. L'aire (exprimée en *u.a.*) de la partie du plan contenant les points $M(x,y)$ tels que $\begin{cases} a \leqslant x \leqslant b \\ f(x) \leqslant y \leqslant g(x) \end{cases}$ est $\displaystyle\int_a^b g(t) - f(t)\,\mathrm{d}t$

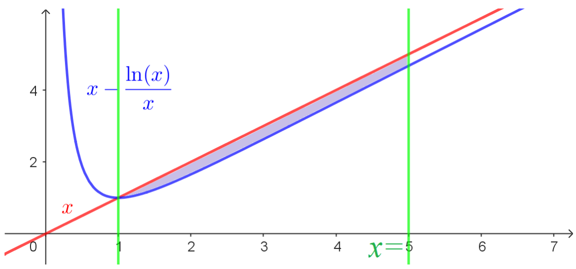
Démonstration ➲ Dans le cas où les deux fonctions sont positives, d'après la définition 3.3 :
Comme $\forall x \in [a,b], f(x) \geqslant 0$, $A_f = \displaystyle\int_a^b f(t)\,\mathrm{d}t$ est l'aire entre $\mathscr{C}_f$ et l'axe (Ox) pour $a \leqslant x \leqslant b$.
Comme $\forall x \in [a,b], g(x) \geqslant 0$, $A_g = \displaystyle\int_a^b g(t)\,\mathrm{d}t$ est l'aire entre $\mathscr{C}_g$ et l'axe (Ox) pour $a \leqslant x \leqslant b$.
Par soustraction de ces grandeurs, $A_g - A_f$ est l'aire entre les courbes $\mathscr{C}_g$ et $\mathscr{C}_f$ pour $a \leqslant x \leqslant b$.
La linéarité de l'intégration permet de conclure $\left(\displaystyle\int_a^b g(t)\,\mathrm{d}t - \int_a^b f(t)\,\mathrm{d}t = \int_a^b g(t) - f(t)\,\mathrm{d}t\right)$.

➲ Dans le cas général, comme f et g sont définies sur $[a,b]$ et comme $\forall x \in [a,b], f(x) \leqslant g(x)$, il existe un minimum m tel que $\forall x \in [a,b], m \leqslant f(x) \leqslant g(x)$.
Les fonctions $f-m$ et $g-m$ sont donc toutes les deux continues et positives sur $[a,b]$, d'après ce qui précède l'aire entre les courbes $\mathscr{C}_{g-m}$ et $\mathscr{C}_{f-m}$ est $\displaystyle\int_a^b g(t) - m - (f(t) - m)\ \mathrm{d}t = \int_a^b g(t) - f(t)\,\mathrm{d}t$.

Exemple 31 – Calculons l'aire entre les courbes des fonctions sin et cos pour $x \in [\frac{\pi}{4}, \frac{5\pi}{4}]$.
J'ai choisi ces bornes qui sont deux solutions successives de l'équation $\sin(x) = \cos(x)$ pour m'assurer que sur tout l'intervalle une des fonctions est supérieure à l'autre, ici on a $\sin(x) \geqslant \cos(x)$. L'aire colorée à gauche sur le graphique ci-dessous est égale à :

$$\begin{aligned}
\int_{\frac{\pi}{4}}^{\frac{5\pi}{4}} \sin(x) - \cos(x)\,\mathrm{d}t &= \int_{\frac{\pi}{4}}^{\frac{5\pi}{4}} \sin(x)\,\mathrm{d}t - \int_{\frac{\pi}{4}}^{\frac{5\pi}{4}} \cos(x)\,\mathrm{d}t \\
&= [-\cos(x)]_{\frac{\pi}{4}}^{\frac{5\pi}{4}} - [\sin(x)]_{\frac{\pi}{4}}^{\frac{5\pi}{4}} \\
&= -\cos(\frac{5\pi}{4}) + \cos(\frac{\pi}{4}) - \sin(\frac{5\pi}{4}) + \sin(\frac{\pi}{4}) \\
&= \frac{\sqrt{2}}{2} + \frac{\sqrt{2}}{2} + \frac{\sqrt{2}}{2} + \frac{\sqrt{2}}{2} = 2\sqrt{2}
\end{aligned}$$



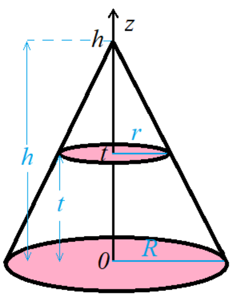


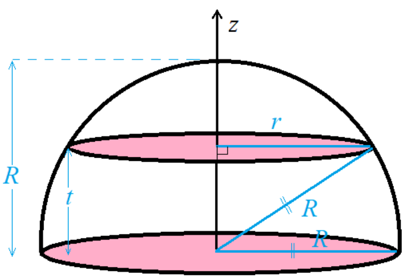
Calculons l'aire entre la courbe de la fonction $f : x \longmapsto x - \frac{\ln x}{x}$ et la droite d'équation $y = x$ pour $x \geqslant 1$. J'ai choisi cette borne inférieure car pour $x = 1$ on a $x - f(x) = \frac{\ln x}{x} = 0$.
Pour $x \geqslant 1$, on a $x - f(x) = \frac{\ln x}{x} \geqslant 0$ et a pour limite 0 : la droite est asymptote à $\mathscr{C}_f$.
J'en déduis que $A(x) = \displaystyle\int_1^x \frac{\ln t}{t}\,\mathrm{d}t$ est l'aire colorée à droite sur le graphique ci-dessus.
Intégrons par parties en posant $u'(t) = \frac{1}{t}$ et $v(t) = \ln t$; on a alors $u'(t) = \ln t$ et $v(t) = \frac{1}{t}$.
$\displaystyle\int_1^x \frac{\ln t}{t}\,\mathrm{d}t = \left[\ln^2(t)\right]_1^x - \int_1^x \frac{\ln t}{t}\,\mathrm{d}t$. On remarque alors que $2A(x) = \left[\ln^2(t)\right]_1^x \iff A(x) = \frac{\ln^2(x)}{2}$.
On a $A(5) = \frac{\ln^2(5)}{2} \approx 1{,}295$ et $A(5000) = \frac{\ln^2(5000)}{2} \approx 36{,}27$ mais $\lim\limits_{x \to +\infty} A(x) = +\infty$.

Propriété 3.19 (Calcul de volume) L'espace est muni d'un repère orthonormal $(O,\, I,\, J,\, K)$.
L'unité de volume (notée *u.v.*) est le volume d'un cube de côtés $[OI]$.
On considère la portion d'un solide délimité par les plans parallèles d'équations $z = a$ et $z = b$ avec $a \leqslant b$ et on note $\mathscr{S}(t)$ l'aire de la section du solide par le plan d'équation $z = t$.
Si $\mathscr{S}$ est une fonction continue sur $[a, b]$ alors le volume de cette portion de solide est $\displaystyle\int_a^b \mathscr{S}(t)\,\mathrm{d}t$.

Exemple 32 – Calculons le volume d'un cône de révolution de hauteur h et de rayon R en choisissant un repère orthonormal $(O,\, \vec{i},\, \vec{j},\, \vec{k})$ avec O au centre de la base et le vecteur $\vec{k}$ orthogonal au plan de la base. Si on coupe ce cône selon un plan parallèle à la base de côte $z = t$, la section de ce cône est un disque d'aire πr^2 où r est le rayon de cette section.
Le théorème de Thalès nous permet d'exprimer r en fonction de R, h et t :
$\dfrac{r}{R} = \dfrac{h-t}{h} \iff r = R \times \dfrac{h-t}{h}$. L'aire de la section est donc $\pi\left(R \times \dfrac{h-t}{h}\right)^2 = \dfrac{\pi R^2 (h-t)^2}{h^2}$.
Le volume du cône est donc $\mathscr{V}(h) = \displaystyle\int_0^h \frac{\pi R^2 (h-t)^2}{h^2}\,\mathrm{d}t$.
Pour ce calcul on aura besoin d'une primitive de $(h-x)^2$:
$(h-x)^2$ est de la forme $-u'u^2$ avec $u(x) = h - x$ et $u'(x) = -1$, sa primitive est $\dfrac{u^3}{3}$, soit $\dfrac{-(h-x)^3}{3}$.

$$\begin{aligned}
\mathscr{V}(h) = &= \frac{\pi R^2}{h^2}\int_0^h (h-t)^2\,\mathrm{d}t \\
&= \frac{\pi R^2}{h^2}\left[\frac{-(h-t)^3}{3}\right]_0^h \\
&= \frac{\pi R^2}{h^2}\left[0 + \frac{h^3}{3}\right] = \frac{\pi R^2 h^3}{3h^2} = \frac{\pi R^2 h}{3}
\end{aligned}$$





Calculons le volume d'une sphère de rayon R en doublant le volume d'une demi-sphère.
On choisit un repère orthonormal $(O,\, \vec{i},\, \vec{j},\, \vec{k})$ avec O au centre de la sphère et on coupe cette sphère par un plan de côte $z = t$, la section de cette sphère est un disque d'aire πr^2 où r en est le rayon.
Le théorème de Pythagore nous permet d'exprimer r en fonction de R et t :
$R^2 = t^2 + r^2 \iff r^2 = R^2 - t^2 \Longrightarrow r = \sqrt{R^2 - t^2}$. L'aire de la section est $\pi\left(\sqrt{R^2 - t^2}\right)^2 = \pi(R^2 - t^2)$.
Le volume du cône est donc $\mathscr{V}(R) = 2\displaystyle\int_0^R \pi(R^2 - t^2)\,\mathrm{d}t$:

$$\begin{aligned}
\mathscr{V}(R) &= 2\left(\int_0^R \pi R^2\,\mathrm{d}t - \int_0^R \pi t^2\,\mathrm{d}t\right) \\
&= 2\left(\left[\pi R^2 t\right]_0^R - \left[\pi\frac{t^3}{3}\right]_0^R\right) \\
&= \pi R^3 - \pi\frac{R^3}{3} = \frac{4\pi R^3}{3}
\end{aligned}$$

Dans les deux cas, on retrouve bien les formules rencontrées, sans justification, au collège.

Définition 3.4 (Valeur moyenne) La valeur moyenne d'une fonction continue sur $[a, b]$ est le nombre réel $\mu = \dfrac{1}{b-a}\displaystyle\int_a^b f(t)\,\mathrm{d}t$.

Remarques :

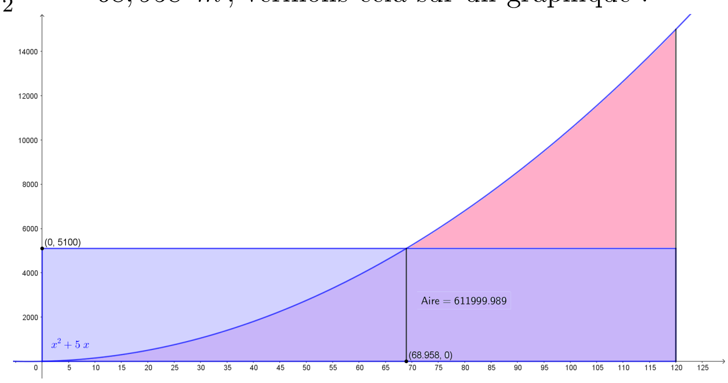
- Si f est positive, sa valeur moyenne sur $[a, b]$ est l'aire du rectangle de côtés $b - a$ et μ car, par définition de l'intégrale d'une fonction positive, cette aire mesure $\displaystyle\int_a^b f(t)\,\mathrm{d}t = \mu(b-a)$.
- En cinématique (étude du mouvement), la distance parcourue par un solide entre les instants t_1 et t_2 animé d'une vitesse instantanée $v(t)$ est donnée par $\displaystyle\int_{t_1}^{t_2} v(t)\,\mathrm{d}t$.
  Si $v(t) = t^2 + 5t$ (t en s et v en $m^{-1}.s$) alors la distance parcourue en deux minutes ($120s$) est $\displaystyle\int_0^{120} t^2 + 5t\,\mathrm{d}t = \left[\frac{t^3}{3} + \frac{5t^2}{2}\right]_0^{120} = \frac{120^3}{3} + \frac{5 \times 120^2}{2} = 612\,000\ m$ soit $612\ km$ (la vitesse atteinte au bout de 2 min est égale à $54\,000\ km^{-1}.h$).
  La vitesse moyenne sur ce parcourt est $\bar{v} = \frac{612\,000}{120} = 5100\ m^{-1}.s$, soit $18\,360\ km^{-1}.h$.

Propriété 3.20 (Existence) Soit f une fonction continue sur un intervalle $[a, b]$ et soit μ, la valeur moyenne de f sur $[a, b]$. Il existe au moins un réel c de $[a, b]$ tel que $f(c) = \mu$.

Démonstration Cette propriété découle du théorème des valeurs intermédiaires :
La fonction f étant continue sur $[a, b]$ elle admet un minimum m et un maximum M et, pour tout $x \in [a, b]$, on a $m \leqslant f(x) \leqslant M$. En intégrant cet encadrement entre a et b, on obtient
$\displaystyle\int_a^b m\,\mathrm{d}t \leqslant \int_a^b f(t)\,\mathrm{d}t \leqslant \int_a^b M\,\mathrm{d}t$, soit $m(b-a) \leqslant \displaystyle\int_a^b f(t)\,\mathrm{d}t \leqslant M(b-a)$ et,
en divisant par $b - a > 0$, il vient $m \leqslant \dfrac{1}{b-a}\displaystyle\int_a^b f(t)\,\mathrm{d}t \leqslant M \Longleftrightarrow m \leqslant \mu \leqslant M$.
La fonction f étant continue et l'image de l'intervalle $[a, b]$ étant $[m, M]$, le TVI implique qu'il existe au moins un réel $c \in [a, b]$ tel que $f(c) = \mu$.

Remarques :

- L'encadrement obtenu $m(b-a) \leqslant \displaystyle\int_a^b f(t)\,\mathrm{d}t \leqslant M(b-a)$ est valable à la seule condition que f soit continue. Si de plus f est positive, il montre que l'aire sous la courbe de f est compris entre l'aire des rectangles de longueur $b - a$ et de hauteur m et M.
- Si la fonction est strictement monotone sur $[a, b]$ le réel c tel que $f(c) = \mu$ est unique, la fonction f réalisant alors une bijection de $[a, b]$ vers $[f(a), f(b)]$ si f est croissante, vers $[f(b), f(a)]$ sinon.
  Dans notre exemple en cinématique ci-dessus, la fonction v est strictement croissante donc il existe un moment de l'intervalle $[0, 120]$ où la vitesse atteinte est égale à la vitesse moyenne $\bar{v} = 2040\ m^{-1}.s$. Pour le trouver, résolvons l'équation $v(t) = 5100 \Longleftrightarrow t^2 + 5t - 5100 = 0$.
  On trouve $t = \frac{5\sqrt{817}-5}{2} \approx 68,958\ m$; vérifions cela sur un graphique :



4 Géométrie dans l'espace

Objectifs :

- Vecteurs de l'espace ; direction d'une droite ou d'un plan, translations
- Barycentre de points pondérés ; colinéarité, bases et repères de l'espace
- Orthogonalité, produit scalaire ; projection orthogonale et distance, vecteur normal
- Représentation paramétrique d'une droite, équation cartésienne d'un plan ou d'une sphère

Aperçu historique : Le début de la géométrie remonte à une lointaine antiquité : Égypte ancienne, Babylone, Inde, Chine, etc. Les documents très anciens sont rares et orientés vers les domaines pratiques (cadastre, architecture). La civilisation grecque produit les premiers écrits théoriques : Euclide d'Alexandrie a vécu vers 300 av. J.-C. et publié *Les Éléments*, un livre de référence pendant des siècles. Il y déduit par des démonstrations, les propriétés des objets qu'il définit à partir d'un jeu réduit d'axiomes. La géométrie constitue une part importante des *Elements* : géométrie plane (livres I à IV et VI) ; géométrie dans l'espace (livres XI à XIII).

Au XVII$^{\text{e}}$ siècle, René Descartes (1596-1650) définit la notion de repère et développe ce qu'on appelle aujourd'hui la géométrie analytique : la position des points de l'espace étant connue par leurs coordonnées (x, y z), les problèmes géométriques sont traités par des calculs algébriques. Pierre de Fermat (vers 1600-1665) est le premier à faire un usage systématique des coordonnées pour résoudre les problèmes de lieux géométriques. Mais dans les notations de Descartes, contrairement à Fermat, les constantes sont continuellement notées a, b, c, d, ... et les variables x, y, z.
Le terme « vecteur » qui apparait en français avec Pierre-Simon de Laplace (1749-1827) vient du latin *vector* (du verbe *vehere* : transporter). Bernard Bolzano (1781-1848) publie un livre où, axiomatiquement comme Euclide, il introduit l'addition et la multiplication des vecteurs. Jean-Victor Poncelet (1788-1867)et Michel Chasles (1793-1880) affinent les travaux de Bolzano tandis que August Ferdinand Möbius (1790-1868) développe le système des coordonnées barycentriques.

Au XIX$^{\text{e}}$ siècle, Hermann Grassmann (1809-1877) établit les fondations de la théorie des espaces vectoriels et de l'algèbre linéaire à partir de ses travaux sur les marées. Le produit scalaire apparait assez tard : on en trouve trace en 1843 chez Grassmann et aussi chez William Rowan Hamilton (1805-1865) qui crée le corps des quaternions (extension des nombres complexes). Un peu plus tard, en 1889, les travaux de formalisation de la géométrie sont repris par Giuseppe Peano (1858-1932) qui définit la notion d'espace vectoriel réel et d'application linéaire et associe le produit scalaire à un calcul de déterminant (aire de la surface orientée du parallélogramme engendré par deux vecteurs, volume du parallélépipède orienté engendré par trois). Roberto Marcolongo et Cesare Burali-Forti le définissent à l'aide seulement du cosinus de l'angle mais sa qualité de forme bilinéaire symétrique sera ensuite exploitée en algèbre linéaire et deviendra définition.

David Hilbert (1862-1943) définit une nouvelle axiomatique de la géométrie euclidienne qui unifie la géométrie dans le plan et l'espace (1899). Josiah Willard Gibbs (1839-1903), nourri des œuvres de Möbius, Grassmann et Hamilton publie entre 1881 et 1884 ses « Éléments d'analyse vectorielle » qui eut une grande importance. Heaviside (1850-1925) montre que le calcul vectoriel est plus avantageux que le calcul sur les composantes des quaternions qu'il trouve trop abstrait et inutile en physique. Les idées vectorielles et linéaires, au départ séparées, sont aujourd'hui toujours associées.

1. Vecteurs de l'espace

Dans cette première partie, on étend la notion de vecteur du plan vue en 2e et en 1re à l'espace.

1.a. Calcul vectoriel

DÉFINITION 4.1 (DÉFINITION) Soient M et M' deux points de l'espace.
La *translation* qui transforme M en M' est appelée translation de vecteur $\overrightarrow{MM'}$ et est noté $t_{\overrightarrow{MM'}}$.
Le *vecteur* $\overrightarrow{MM'}$ a
- pour *direction* la droite (MM') ou toute droite parallèle à (MM')
- pour *sens* sur la droite (MM'), le sens de M vers M'
- pour *norme* la longueur du segment $[MM']$, ce qu'on note $\|\overrightarrow{MM'}\| = MM'$

Remarques :
- Quel que soit le point M, le vecteur $\overrightarrow{MM}$ est le vecteur nul, noté $\overrightarrow{0}$.
  La translation $t_{\overrightarrow{0}}$ transforme un point M en lui-même.
  Le vecteur nul n'a ni direction ni sens ; sa norme est nulle.
- Deux vecteurs sont *égaux* si ils ont même direction, même sens et même norme.
  Si la translation $t_{\overrightarrow{MM'}}$ transforme N en N' alors les vecteurs $\overrightarrow{MM'}$ et $\overrightarrow{NN'}$ sont égaux.
  On peut traduire l'égalité vectorielle $\overrightarrow{MM'} = \overrightarrow{NN'}$
  — dans le langage des figures : le quadrilatère $MM'N'N$ est un parallélogramme
  — dans le langage des transformations : $t_{\overrightarrow{MM'}}(N) = N'$ (l'image de N par $t_{\overrightarrow{MM'}}$ est N')
  On nomme souvent les vecteurs avec la notation *générique* : $\overrightarrow{u}$, $\overrightarrow{v}$, etc. qui n'utilise pas de points pour désigner un vecteur. Si par exemple $\overrightarrow{u} = \overrightarrow{MM'}$, on dit que $\overrightarrow{MM'}$ est le *représentant* du vecteur $\overrightarrow{u}$ d'origine M.
- En physique, et plus particulièrement en mécanique, on utilise des vecteurs pour représenter les forces. Un force est une grandeur qui a une direction, un sens, une norme (intensité, mesurée en Newton) et aussi un point d'application (centre de gravité du solide) sur lequel s'exerce cette force.

DÉFINITION 4.2 (ADDITION) Soient $\overrightarrow{u}$ et $\overrightarrow{v}$ deux vecteurs et A un point de l'espace.
Si B et C sont les points tels que $\overrightarrow{u} = \overrightarrow{AB}$ et $\overrightarrow{v} = \overrightarrow{BC}$ alors $\overrightarrow{u} + \overrightarrow{v} = \overrightarrow{AC}$.

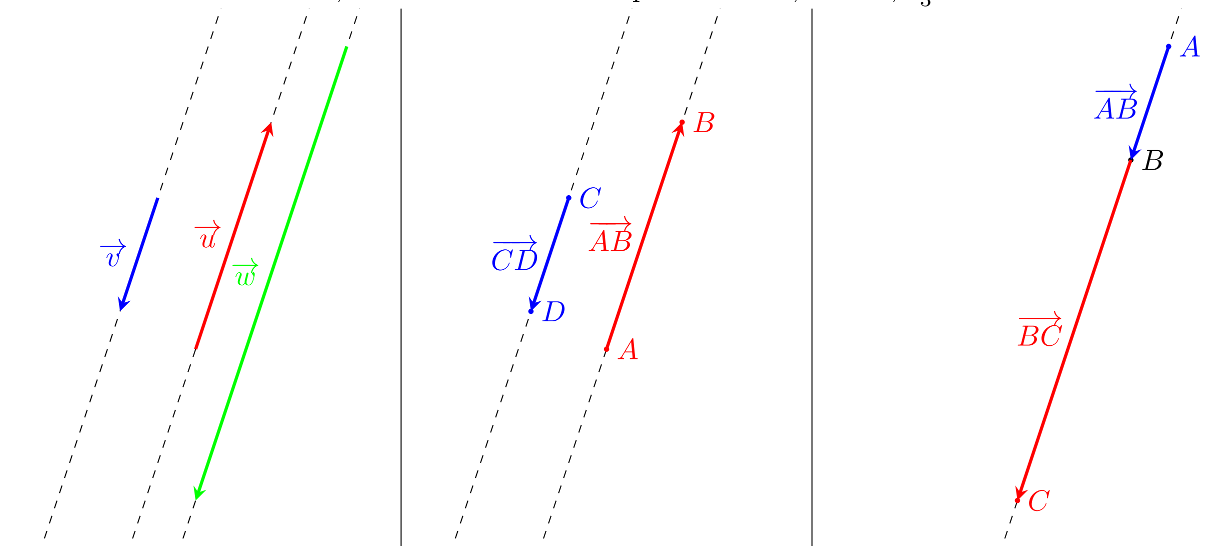
Remarques :
- Quand $\overrightarrow{u}$ et $\overrightarrow{v}$ sont donnés, on peut construire le vecteur $\overrightarrow{u} + \overrightarrow{v}$ à partir de n'importe quel point. En partant d'un point O : si $\overrightarrow{u} = \overrightarrow{OP}$ et $\overrightarrow{v} = \overrightarrow{PQ}$ alors $\overrightarrow{u} + \overrightarrow{v} = \overrightarrow{OQ}$
- Si $\overrightarrow{AC}$ est donné, on peut écrire $\overrightarrow{AC} = \overrightarrow{AB} + \overrightarrow{BC}$ ou $\overrightarrow{AC} = \overrightarrow{AD} + \overrightarrow{DC}$; tout point M tel que $\overrightarrow{AC} = \overrightarrow{AM} + \overrightarrow{MC}$ convient ; cette propriété s'appelle *relation de Chasles*. Autrement dit, la translation qui transforme A en C peut être décomposée en un translation qui transforme A en M puis une translation qui transforme M en C. Peu importe par où on passe pour aller de A en C. On se sert souvent de cette relation pour décomposer un vecteur en somme de deux ou plusieurs vecteurs : $\overrightarrow{AD} = \overrightarrow{AB} + \overrightarrow{BC} + \overrightarrow{CD}$.
- La *règle du parallélogramme* : A étant donné, si B et C sont les points tels que $\overrightarrow{u} = \overrightarrow{AB}$ et $\overrightarrow{v} = \overrightarrow{AC}$ alors le vecteur $\overrightarrow{u} + \overrightarrow{v}$ est égal au vecteur $\overrightarrow{AD}$ où D est le 4e sommet du parallélogramme $BACD$. En effet, $BACD$ est un parallélogramme si et seulement si $\overrightarrow{AC} = \overrightarrow{BD}$ et donc $\overrightarrow{AB} + \overrightarrow{AC} = \overrightarrow{AB} + \overrightarrow{BD} = \overrightarrow{AD}$ d'après la définition 4.2.

Définition 4.3 (Multiplication par un réel) Soient $\overrightarrow{u}$ un vecteur de l'espace et k un réel

Si $k \neq 0$ et $\overrightarrow{u} \neq \overrightarrow{0}$ alors $k\overrightarrow{u}$ est un vecteur qui a
- même direction que $\overrightarrow{u}$
- même sens que $\overrightarrow{u}$ si $k > 0$, le sens contraire si $k < 0$
- une norme égale à $\overrightarrow{u}$ $|k| \times \|\overrightarrow{u}\|$

Si $k = 0$ ou $\overrightarrow{u} = \overrightarrow{0}$
alors $k\overrightarrow{u} = \overrightarrow{0}$

Remarques :
- Avec des points : si $\overrightarrow{AC} = k\overrightarrow{AB}$ alors, dans tous les cas, on aura $AC = kAB$.
  Si, de plus, $k > 0$ alors $C \in [AB)$ mais si $k < 0$ alors $C \in (AB)$ et $C \notin [AB)$.
- La notation de cette multiplication est généralement omise : on écrit $k\overrightarrow{u}$, $2\overrightarrow{AB}$, $\frac{-1}{3}\overrightarrow{CD}$.
  Dans certains textes, elle est notée avec un point : $k \cdot \overrightarrow{u}$, $2 \cdot \overrightarrow{AB}$, $\frac{-1}{3} \cdot \overrightarrow{CD}$.



$\overrightarrow{u} = k\overrightarrow{v}$ avec $k < 0$
$\overrightarrow{v} = k\overrightarrow{w}$ avec $k > 0$

$(AB)//(CD)$
$\overrightarrow{AB} = k\overrightarrow{CD}$ avec $k < -1$

A, B, C alignés
$\overrightarrow{BC} = k\overrightarrow{AB}$ avec $k > 1$

Propriété 4.1 (Espace vectoriel) Pour tous vecteurs de l'espace $\overrightarrow{u}$, $\overrightarrow{v}$ et $\overrightarrow{w}$, on a :

Commutativité : $\overrightarrow{u} + \overrightarrow{v} = \overrightarrow{v} + \overrightarrow{u}$
$\overrightarrow{0}$ élément neutre : $\overrightarrow{u} + \overrightarrow{0} = \overrightarrow{0} + \overrightarrow{u}$

Associativité : $(\overrightarrow{u} + \overrightarrow{v}) + \overrightarrow{w} = \overrightarrow{u} + (\overrightarrow{v} + \overrightarrow{w})$
$\overrightarrow{-u}$ opposé de $\overrightarrow{u}$: $\overrightarrow{u} + (\overrightarrow{-u}) = (\overrightarrow{-u}) + \overrightarrow{u} = \overrightarrow{0}$

L'ensemble des vecteurs de l'espace, noté $\mathcal{E}$, muni de l'addition est un *groupe commutatif*.
La multiplication par un réel, confère au groupe $(\mathcal{E}, +)$ la structure d'*espace vectoriel*, noté $(\mathcal{E}, +, \cdot)$, car $\forall(\overrightarrow{u}, \overrightarrow{v}, \overrightarrow{w}) \in \mathcal{E}^3$ et $\forall(k, k') \in \mathbb{R}^2$:

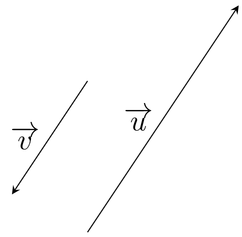
Distributivité 1 : $k(\overrightarrow{u} + \overrightarrow{v}) = k\overrightarrow{u} + k\overrightarrow{v}$
Distributivité 2 : $(k + k')\overrightarrow{u} = k\overrightarrow{u} + k'\overrightarrow{u}$

Associativité : $k(k'\overrightarrow{u}) = (k \times k')\overrightarrow{u}$
1 élément neutre : $1\overrightarrow{u} = \overrightarrow{u}$

Démonstration Ces propriétés qui découlent des définitions sont admises ici.

Remarques :
- Dire qu'un vecteur $\overrightarrow{u}$ est une *combinaison linéaire* des vecteurs $\overrightarrow{u_1}$, $\overrightarrow{u_2}$ et $\overrightarrow{u_3}$ signifie qu'il existe des réels x, y et z tels que $\overrightarrow{u} = x\overrightarrow{u_1} + y\overrightarrow{u_2} + z\overrightarrow{u_3}$. Dans l'espace, si on connait trois vecteurs $\overrightarrow{u_1}, \overrightarrow{u_2}, \overrightarrow{u_3}$ dont toute combinaison linéaire nulle implique des coefficients tous nuls, autrement dit $x\overrightarrow{u_1} + y\overrightarrow{u_2} + z\overrightarrow{u_3} = \overrightarrow{0} \Longrightarrow (x, y, z) = (0, 0, 0)$, alors $(\overrightarrow{u_1}, \overrightarrow{u_2}, \overrightarrow{u_3})$ est une *base* de $\mathcal{E}$.
- Il existe de nombreux exemples d'espaces vectoriels, dont les éléments ne sont pas appelés vecteurs : l'ensemble des polynômes, celui des matrices de dimensions $n \times p$ et celui des fonctions continues, pour n'en citer que trois, munis de l'addition et de la multiplication par un réel sont des espaces vectoriels.

1.b. Droites et plans de l'espace



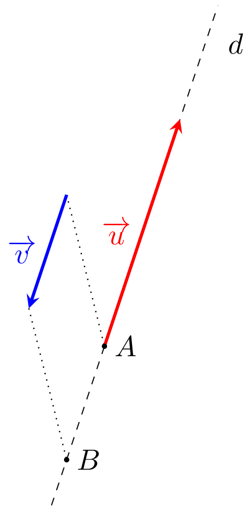
DÉFINITION 4.4 Deux vecteurs de l'espace $\overrightarrow{u}$ et $\overrightarrow{v}$ non nuls sont dits *colinéaires* s'il existe un réel k tel que $\overrightarrow{u} = k\overrightarrow{v}$.

<u>Convention</u> : Le vecteur nul $\overrightarrow{0}$ est colinéaire à tous les vecteurs.

Colinéaire signifie « sur une même ligne ».
Deux vecteurs non nuls sont colinéaires si et seulement si ils ont la même direction.
<u>Traduction dans le langage des figures</u> :

- $\overrightarrow{AB}$ et $\overrightarrow{CD}$ sont colinéaires $\iff (AB)//(CD)$
- $\overrightarrow{AB}$ et $\overrightarrow{BC}$ sont colinéaires $\iff$ les points A, B et C sont alignés

DÉFINITION 4.5 (VECTEUR DIRECTEUR) Soit d une droite.
On appelle *vecteur directeur* de d tout vecteur non nul ayant la même direction que d.



<u>Remarques</u> :

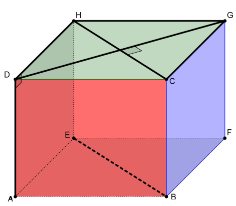
- Une droite possède une infinité de vecteurs directeurs qui sont tous colinéaires entre eux, puisque de même direction. Si A et B sont deux points distincts de la droite d alors les vecteurs $\overrightarrow{AB}$ et $\forall k \in \mathbb{R}^*,\ k\overrightarrow{AB}$ sont des vecteurs directeurs de d
- On peut définir une droite par la donnée d'un point et d'un vecteur directeur. La notation $d(A, \overrightarrow{u})$ désigne la droite passant par A de vecteur directeur $\overrightarrow{u}$.
  Sur mon illustration, $d = d(A, \overrightarrow{u})$ et $B \in d(A, \overrightarrow{u})$, on a donc $d = (AB)$ et comme $\overrightarrow{v} = \overrightarrow{AB}$ on a aussi $d = d(A, \overrightarrow{v})$.
- D'une façon générale $M \in (AB) \iff \overrightarrow{AM}$ et $\overrightarrow{AB}$ sont colinéaires.
  On a aussi $M \in d(A, \overrightarrow{u}) \iff \exists x \in \mathbb{R}, \overrightarrow{AM} = x\overrightarrow{u}$.
  Dans cette dernière formulation $\overrightarrow{u}$ est une *base* de cette droite, A est l'*origine* d'un *repère* $(A, \overrightarrow{u})$ dans lequel x est l'*abscisse* de M.

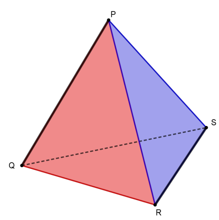
PROPRIÉTÉ 4.2 (POSITION RELATIVE DE DEUX DROITES) Soient deux droites d et d' dont les vecteurs directeurs respectifs sont $\overrightarrow{u}$ et $\overrightarrow{v}$.

- Si $\overrightarrow{u}$ et $\overrightarrow{v}$ sont colinéaires alors d et d' sont coplanaires (dans un même plan) et parallèles, éventuellement confondues
- Sinon, d et d' sont sécantes en un point et coplanaires ou bien non-sécantes et non-coplanaires

Sur le parallélépipède rectangle $ABCDEFGH$

- (BE) et (CH) sont coplanaires et parallèles, de même pour (CG) et (DH)
- (CH) et (DG) sont sécantes en un point et coplanaires, de même pour (AD) et (DH)
- (BE) et (AD) sont non-sécantes et non-coplanaires, de même pour (BE) et (DG)





Attention à la perspective ! Deux droites peuvent sembler sécantes alors qu'elles sont non-coplanaires, comme (AD) et (GH). Deux droites peuvent sembler parallèles alors qu'elles sont non-coplanaires, comme (PQ) et (RS) sur le tétraèdre $PQRS$ de droite.

Définition 4.6 (Plan) Soient A un point et $\overrightarrow{u}$ et $\overrightarrow{v}$ deux vecteurs non colinéaires de l'espace.
Le *plan* $(A, \overrightarrow{u}, \overrightarrow{v})$ est définit par les droites sécantes $d(A, \overrightarrow{u})$ et $d(A, \overrightarrow{v})$.
$(\overrightarrow{u}, \overrightarrow{v})$ est un couple de *vecteurs directeurs* du plan $(A, \overrightarrow{u}, \overrightarrow{v})$ qui définit sa *direction*.

Propriété 4.3 Un point M appartient au plan $\mathcal{P} = (A, \overrightarrow{u}, \overrightarrow{v})$ si et seulement si il existe un couple de réels (x, y) tel que $\overrightarrow{AM} = x\overrightarrow{u} + y\overrightarrow{v}$.

Démonstration Les vecteurs $\overrightarrow{u}$ et $\overrightarrow{v}$ étant non-colinéaires, ils définissent une base du plan $\mathcal{P}$.
Le triplet $(A, \overrightarrow{u}, \overrightarrow{v})$ est par conséquent un repère du plan $\mathcal{P}$.
Cela implique qu'il existe un couple de réels (x, y) tel que $\overrightarrow{AM} = x\overrightarrow{u} + y\overrightarrow{v}$.

Réciproquement, supposons qu'un point P vérifie l'égalité $\overrightarrow{AP} = x\overrightarrow{u} + y\overrightarrow{v}$. P appartient-il à $\mathcal{P}$?
Le couple de réels (x, y) étant donné, on peut lui associer un point $M \in \mathcal{P}$ tel que $\overrightarrow{AM} = x\overrightarrow{u} + y\overrightarrow{v}$.
Comme cela conduit à $\overrightarrow{AM} = \overrightarrow{AP} \iff \overrightarrow{MP} = \overrightarrow{0}$, on en déduit que $M = P$ et donc $P \in \mathcal{P}$.

Définition 4.7 (Coplanairité) Soient $\mathcal{P} = (A, \overrightarrow{u}, \overrightarrow{v})$ un *plan* et $\overrightarrow{w}$ un vecteur de l'espace.
$\overrightarrow{u}$, $\overrightarrow{v}$ et $\overrightarrow{w}$ sont *coplanaires* si et seulement si il existe un point $D \in \mathcal{P}$ tel que $\overrightarrow{AD} = \overrightarrow{w}$.

Propriété 4.4 Soient $(\overrightarrow{u}, \overrightarrow{v})$ un couple de vecteurs non-colinéaires et $\overrightarrow{w}$ un vecteur de l'espace.
$\overrightarrow{u}$, $\overrightarrow{v}$ et $\overrightarrow{w}$ sont coplanaires si et seulement si il existe des réels x et y tels que $\overrightarrow{w} = x\overrightarrow{u} + y\overrightarrow{v}$.

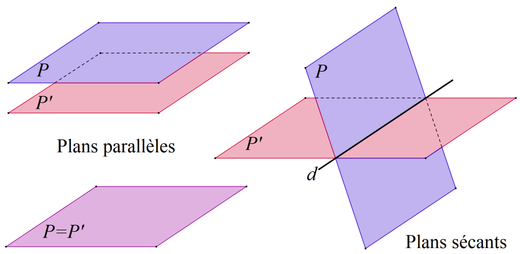
Démonstration En choisissant un point A quelconque, le triplet $(A, \overrightarrow{u}, \overrightarrow{v})$ définit un plan $\mathcal{P}$.
D'après la définition, $\overrightarrow{u}$, $\overrightarrow{v}$ et $\overrightarrow{w}$ sont coplanaires si et seulement si il existe un point $D \in \mathcal{P}$ tel que $\overrightarrow{AD} = \overrightarrow{w}$, et on a vu dans la propriété 4.3 que cela équivaut à l'existence d'un couple de réels (x, y) tel que $\overrightarrow{AD} = x\overrightarrow{u} + y\overrightarrow{v}$. La conclusion en découle.

Remarques :

- Trois vecteurs sont coplanaires si et seulement si un de ces vecteurs est une combinaison linéaire des deux autres. Cela s'établit généralement par le calcul vectoriel :
  Supposons que A, B, C soient trois points non alignés. Ils définissent le plan $(A, \overrightarrow{AB}, \overrightarrow{AC})$, noté (ABC). Supposons qu'un point D vérifie $3\overrightarrow{DA} - 2\overrightarrow{DB} + \overrightarrow{DC} = \overrightarrow{0}$. A t-on $D \in (ABC)$?
  On a $3\overrightarrow{DA} - 2(\overrightarrow{DA} + \overrightarrow{AB}) + (\overrightarrow{DA} + \overrightarrow{AC}) = \overrightarrow{0} \iff (3 - 2 + 1)\overrightarrow{DA} = 2\overrightarrow{AB} - \overrightarrow{AC}$ d'où $\overrightarrow{AD} = -\overrightarrow{AB} + \frac{1}{2}\overrightarrow{AC}$. Conclusion : $D \in (ABC)$, les vecteurs $\overrightarrow{AB}$, $\overrightarrow{AC}$ et $\overrightarrow{AD}$ sont coplanaires.
- Trois vecteurs non-coplanaires forment une base de l'espace. Si $\overrightarrow{AB}$, $\overrightarrow{AC}$ et $\overrightarrow{AD}$ sont non-coplanaires alors $D \notin (ABC)$, $ABCD$ est un tétraèdre – une pyramide à base triangulaire – dont le volume n'est pas nul.

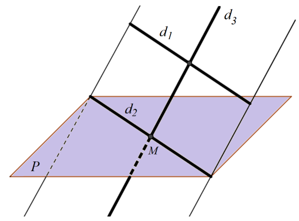
Propriété 4.5 (Position relative de deux plans)
Soient deux plans $\mathcal{P}$ et $\mathcal{P}'$.

- S'ils ont même direction, $\mathcal{P}$ et $\mathcal{P}'$ sont parallèles, éventuellement confondus.
- Sinon les plans $\mathcal{P}$ et $\mathcal{P}'$ sont sécants.
  Leur intersection est une droite d.



PROPRIÉTÉ 4.6 (POSITION RELATIVE D'UNE DROITE ET D'UN PLAN) Soient $\mathcal{P}$ un plan de direction $(\overrightarrow{u}, \overrightarrow{v})$ et d une droite de vecteur directeur $\overrightarrow{w}$.

- Si $\overrightarrow{u}$, $\overrightarrow{v}$ et $\overrightarrow{w}$ sont coplanaires, d est strictement parallèle à $\mathcal{P}$ $(d_1 \cap \mathcal{P} = \emptyset)$ ou incluse dans $\mathcal{P}$ $(d_2 \subset \mathcal{P})$.
- Sinon d perce le plan $\mathcal{P}$ en un point M $(d_3 \cap \mathcal{P} = \{M\})$.

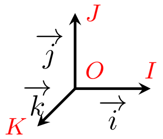


1.c. Bases et repères de l'espace

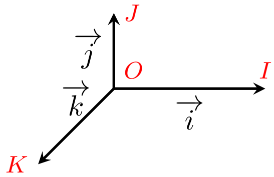
DÉFINITION 4.8 Une *base* de l'espace est un triplet $(\overrightarrow{u}, \overrightarrow{v}, \overrightarrow{w})$ de vecteurs non coplanaires. Un *repère* est un quadruplet $(A, \overrightarrow{u}, \overrightarrow{v}, \overrightarrow{w})$ où A est un point, appelé *Origine*, et $(\overrightarrow{u}, \overrightarrow{v}, \overrightarrow{w})$ une base.

Remarques :

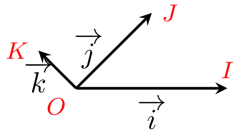
- Dire que la base $(\overrightarrow{OI}, \overrightarrow{OJ}, \overrightarrow{OK})$ est *normée* signifie que $\overrightarrow{OI}$, $\overrightarrow{OJ}$ et $\overrightarrow{OK}$ sont de même longueur. Ils sont de norme unitaire : $\|\overrightarrow{OI}\| = \|\overrightarrow{OJ}\| = \|\overrightarrow{OK}\| = 1$.
  Une base normée $(\overrightarrow{OI}, \overrightarrow{OJ}, \overrightarrow{OK})$ est dite *orthonormée* si et seulement si elle est orthogonale, c'est-à-dire que $(OI) \perp (OJ)$, $(OI) \perp (OK)$ et $(OJ) \perp (OK)$. Comme on le verra dans la propriété 4.18, on ne peut calculer des distances que si on utilise un repère orthonormé.
- Un quadruplet (O, I, J, K) de points non coplanaires constitue le repère $(O, \overrightarrow{OI}, \overrightarrow{OJ}, \overrightarrow{OK})$. Ce repère est noté indifféremment $(O, \overrightarrow{OI}, \overrightarrow{OJ}, \overrightarrow{OK})$ ou (O, I, J, K).
  Dans la suite, on prendra souvent $(\overrightarrow{i} = \overrightarrow{OI}, \overrightarrow{j} = \overrightarrow{OJ}, \overrightarrow{k} = \overrightarrow{OK})$ comme base.



base orthonormée



base orthogonale



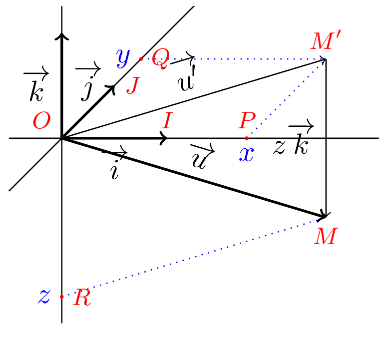
base quelconque

THÉORÈME 4.1 Soient $(\overrightarrow{i}, \overrightarrow{j}, \overrightarrow{k})$ une base et O un point.

- Pour tout vecteur $\overrightarrow{u}$, il existe un unique triplet de réels (x, y, z) tel que $\overrightarrow{u} = x\overrightarrow{i} + y\overrightarrow{j} + z\overrightarrow{k}$.
- Pour tout point M, il existe un unique triplet de réels (x, y, z) tel que $\overrightarrow{OM} = x\overrightarrow{i} + y\overrightarrow{j} + z\overrightarrow{k}$.

DÉMONSTRATION Soient $(\overrightarrow{i}, \overrightarrow{j}, \overrightarrow{k})$ une base et $\overrightarrow{u}$ un vecteur.
Soient O un point et I, J, K et M les points définis par $\overrightarrow{OI} = \overrightarrow{i}$, $\overrightarrow{OJ} = \overrightarrow{j}$ et $\overrightarrow{OM} = \overrightarrow{u}$.

Existence : Projetons M sur le plan (OIJ) parallèlement à (OK).
$\overrightarrow{i}$, $\overrightarrow{j}$ et $\overrightarrow{k}$ étant non-coplanaires, la droite $d(M, \overrightarrow{k})$ coupe le plan (OIJ) en un point M'. Par conséquent il existe un réel k tel que $\overrightarrow{M'M} = z\overrightarrow{k}$ et $\overrightarrow{u} = \overrightarrow{OM} = \overrightarrow{OM'} + \overrightarrow{M'M} = \overrightarrow{u'} + z\overrightarrow{k}$ avec $\overrightarrow{u'}$, $\overrightarrow{i}$ et $\overrightarrow{j}$ coplanaires.
$\overrightarrow{i}$ et $\overrightarrow{j}$ étant non-colinéaires, la parallèle à (OJ) passant par M' coupe la droite (OI) en un point P. De même, la parallèle à (OI) passant par M' coupe la droite (OJ) en un point Q. $OPM'Q$ étant un parallélogramme par construction, on a $\overrightarrow{OM'} = \overrightarrow{OP} + \overrightarrow{OQ}$ $(\star)$.
Comme $\overrightarrow{OP}$ et $\overrightarrow{i}$ sont colinéaires, d'après la définition 4.4 il existe un réel x tel que $\overrightarrow{OP} = x\overrightarrow{i}$ et de même, il existe un réel y tel que $\overrightarrow{OQ} = y\overrightarrow{j}$.
L'égalité $(\star)$ s'écrit alors $\overrightarrow{u'} = x\overrightarrow{i} + y\overrightarrow{j}$.
On a construit les réels x, y et z, ce qui assure leur existence.



Unicité : Raisonnons par l'absurde.
Supposons qu'il existe deux triplets distincts de réels (x, y, z) et (x', y', z') tels que $\overrightarrow{u} = x\overrightarrow{i} + y\overrightarrow{j} + z\overrightarrow{k}$ et $\overrightarrow{u} = x'\overrightarrow{i} + y'\overrightarrow{j} + z'\overrightarrow{k}$.
On a $x\overrightarrow{i} - x'\overrightarrow{i} = (\overrightarrow{u} - y\overrightarrow{j} - z\overrightarrow{k}) - (\overrightarrow{u} - y'\overrightarrow{j} - z'\overrightarrow{k}) = -y\overrightarrow{j} + y'\overrightarrow{j} - z\overrightarrow{k} + z'\overrightarrow{k}$, soit $(x - x')\overrightarrow{i} = (y' - y)\overrightarrow{j} + (z' - z)\overrightarrow{k}$.
$\overrightarrow{i}$, $\overrightarrow{j}$ et $\overrightarrow{k}$ étant non-coplanaires, cette égalité ne peut être vérifiée que si $(x - x')$, $(y' - y)$ et $(z' - z)$ sont nuls :

$$\begin{cases} x - x' = 0 \\ y' - y = 0 \\ z' - z = 0 \end{cases} \iff \begin{cases} x = x' \\ y = y' \\ z = z' \end{cases}$$ ce qui contredit le fait que (x, y, z) et (x', y', z') soient distincts.

La décomposition $\overrightarrow{u} = x\overrightarrow{i} + y\overrightarrow{j} + z\overrightarrow{k}$ est donc bien unique.

Remarque : Si $\overrightarrow{OM} = \overrightarrow{u} = x\vec{\imath} + y\vec{\jmath} + z\vec{k}$.

- x, y et z sont les *coordonnées* du point M dans le repère $(O, \overrightarrow{i}, \overrightarrow{j}, \overrightarrow{k})$
  x est l'*abscisse* de M, y en est l'*ordonnée* et z la *cote*.
- x, y et z sont aussi les *composantes* du vecteur $\overrightarrow{u}$ dans la base $(\overrightarrow{i}, \overrightarrow{j}, \overrightarrow{k})$.
  x est la 1^re^ composante de $\overrightarrow{u}$, y en est la 2^e^ composante et z la 3^e^.

On note ce triplet indifféremment (x, y, z) ou $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$.

Propriété 4.7 (Composantes) Soient $\overrightarrow{u}(x, y, z)$ et $\overrightarrow{u'}(x', y', z')$ deux vecteurs et leurs composantes dans une base $(\overrightarrow{i}, \overrightarrow{j}, \overrightarrow{k})$, $A(x_A, y_A, z_A)$ et $B(x_B, y_B, z_B)$ deux points et leurs coordonnées dans un repère $(O, \overrightarrow{i}, \overrightarrow{j}, \overrightarrow{k})$ et λ un réel.

- $\overrightarrow{u} + \overrightarrow{u'}$ a pour composantes $(x + x', y + y', z + z')$
- $\lambda\overrightarrow{u}$ a pour composantes $(\lambda x, \lambda y, \lambda z)$ et en particulier $-\overrightarrow{u}(-x, -y, -z)$
- $\overrightarrow{u} - \overrightarrow{u'}$ a pour composantes $(x - x', y - y', z - z')$
- $\overrightarrow{u} = \overrightarrow{u'} \iff \begin{cases} x = x' \\ y = y' \\ z = z' \end{cases}$ et en particulier $\overrightarrow{u} = \overrightarrow{0} \iff \begin{cases} x = 0 \\ y = 0 \\ z = 0 \end{cases}$
- Le vecteur $\overrightarrow{AB}$ a pour composantes $(x_B - x_A, y_B - y_A, z_B - z_A)$

Démonstration Ces propriétés découlent des propriétés de l'addition vectorielle, du produit d'un vecteur par un réel et de l'utilisation d'une base de l'espace vectoriel.
Pour le dernier point, $\overrightarrow{AB} = \overrightarrow{OB} - \overrightarrow{OA}$ (relation « soustractive » de Chasles).
Or $\overrightarrow{OB} = x_B\overrightarrow{i} + y_B\overrightarrow{j} + z_B\overrightarrow{k}$ et $\overrightarrow{OA} = x_A\overrightarrow{i} + y_A\overrightarrow{j} + z_A\overrightarrow{k}$.
Après réduction on obtient $\overrightarrow{AB} = (x_B - x_A)\overrightarrow{i} + (y_B - y_A)\overrightarrow{j} + (z_B - z_A)\overrightarrow{k}$ d'où la conclusion.

Exemple 33 – Soient $\mathcal{P}(M, \overrightarrow{u}, \overrightarrow{v})$ un plan et $d(N, \overrightarrow{w})$ une droite avec, dans un certain repère de l'espace, $\overrightarrow{u}(1, -1, 0)$, $\overrightarrow{v}(2, 0, 1)$, $\overrightarrow{w}(-1, 2, -3)$, $M(2, 1, -1)$ et $N(-2, 3, 5)$.
Quelle position relative ont le plan $\mathcal{P}$ et la droite d ?
Une combinaison linéaire de $\overrightarrow{u}$ et $\overrightarrow{v}$ s'écrit $a\overrightarrow{u} + b\overrightarrow{v}$ dont les composantes sont $(a + 2b, -a, b)$.
Si $\overrightarrow{w} = a\overrightarrow{u} + b\overrightarrow{v}$ on aurait $\begin{cases} -1 = a + 2b \\ 2 = -a \\ 3 = b \end{cases} \iff \begin{cases} a + 2b = -1 \\ a = -2 \\ b = 3 \end{cases}$
Cela ne se peut pas car alors $a + 2b = -2 + 6 = 4 \neq -1$ donc $\overrightarrow{u}$, $\overrightarrow{v}$ et $\overrightarrow{w}$ sont non coplanaires.

On en déduit que la droite d perce le plan plan $\mathcal{P}$ en un point $K(\alpha, \beta, \gamma)$.

- Comme $K \in \mathcal{P}$, $\overrightarrow{MK}(\alpha-2, \beta-1, \gamma+1)$ est une combinaison linéaire de $\vec{u}$ et $\vec{v}$.
  On doit donc avoir $\begin{cases} \alpha - 2 = a + 2b \\ \beta - 1 = -a \\ \gamma + 1 = b \end{cases} \iff \begin{cases} \alpha - 2 = a + 2b \\ a = 1 - \beta \\ b = \gamma + 1 \end{cases}$
  Cela conduit à l'*équation cartésienne* du plan $\mathcal{P}$: $\alpha + \beta - 2\gamma - 5 = 0$.
- Comme $K \in d$, $\overrightarrow{NK}(\alpha+2, \beta-3, \gamma-5)$ est colinéaire à $\vec{w}$.
  Il existe donc un réel t tel que $\overrightarrow{NK} = t\vec{w} \iff \begin{cases} \alpha + 2 = -t \\ \beta - 3 = 2t \\ \gamma - 5 = -3t \end{cases} \iff \begin{cases} \alpha = -t - 2 \\ \beta = 2t + 3 \\ \gamma = -3t + 5 \end{cases}$
  Ce système constitue l'*équation paramétrique* de d.

Comme $K = d \cap \mathcal{P}$ on doit avoir $\alpha + \beta - 2\gamma - 5 = 0$, $\alpha = -t - 2$, $\beta = 2t + 3$ et $\gamma = -3t + 5$.
Le paramètre t doit vérifier $-t - 2 + 2t + 3 - 2(-3t + 5) - 5 = 0 \iff 7t - 14 = 0 \iff t = 2$.
On en déduit les coordonnées de K : $\alpha = -2 - 2 = -4$, $\beta = 2 \times 2 + 3 = 7$ et $\gamma = -3 \times 2 + 5 = -1$ donc $K(-4, 7, -1)$.

2. Barycentres

On s'intéresse ici à des couples (point, réel) appellés « points pondérés » (des points qui ont un poids).

2.a. Barycentre de deux points

DÉFINITION 4.9 Soient (A, α) et (B, β) deux points pondérés tels que $\alpha + \beta \neq 0$.
On appelle barycentre de (A, α) et (B, β) le point G défini par $\alpha\overrightarrow{GA} + \beta\overrightarrow{GB} = \vec{0}$

PROPRIÉTÉ 4.8 Soient (A, α) et (B, β) deux points pondérés tels que $\alpha + \beta \neq 0$.
Le barycentre de (A, α) et (B, β) vérifie l'égalité : $\overrightarrow{AG} = \frac{\beta}{\alpha+\beta}\overrightarrow{AB}$

DÉMONSTRATION Décomposons la définition du barycentre G :
$\alpha\overrightarrow{GA} + \beta\overrightarrow{GB} = \vec{0} \iff \alpha\overrightarrow{GA} + \beta(\overrightarrow{GA} + \overrightarrow{AB}) = \vec{0}$.
On obtient la nouvelle égalité $\beta\overrightarrow{AB} = \alpha\overrightarrow{AG} + \beta\overrightarrow{AG} = (\alpha + \beta)\overrightarrow{AG}$.
Comme $\alpha + \beta \neq 0$, on peut diviser par ce nombre et obtenir la relation de l'énoncé.

Remarques :

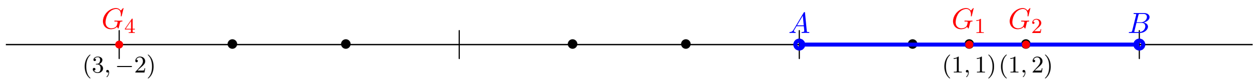
- Lorsque $A = B$, les points A, B et G sont confondus.
  Sinon, les points A, B et G sont alignés et G a pour abscisse $\frac{\beta}{\alpha+\beta}$ dans le repère $(A, \overrightarrow{AB})$.
  Si $\alpha = 0$ et $\beta \neq 0$ (ou $\beta = 0$ et $\alpha \neq 0$), le point G est alors confondu avec B (ou avec A).
- Soit $k \in \mathbb{R}^*$. Le barycentre de (A, α) et (B, β) est aussi le barycentre de $(A, k\alpha)$ et $(B, k\beta)$.
  Par exemple, le barycentre de $(A, 120)$ et $(B, 30)$ est aussi le barycentre de $(A, 4)$ et $(B, 1)$.
  On peut remplacer les coefficients du barycentre (α et β) par des coefficients proportionnels.
- I milieu de $[AB]$ se traduit vectoriellement par $\overrightarrow{IA} + \overrightarrow{IB} = \vec{0}$, autrement dit I est le barycentre de $(A, 1)$ et $(B, 1)$ ou plus généralement de (A, α) et (B, α). On dit alors que I est l'*isobarycentre* de A et B – « iso » pour dire qu'ils sont affectés des même coefficients. On a également les autres traductions vectorielles de l'isobarycentre : $\overrightarrow{AI} = \frac{1}{2}\overrightarrow{AB}$ (propriété 4.8) et, pour tout point M de l'espace, $\overrightarrow{MI} = \frac{1}{2}\overrightarrow{MA} + \frac{1}{2}\overrightarrow{MB}$ (propriété 4.10).

Exemple 34 – Quel est le barycentre G_1 de $(A, 30)$ et $(B, 30)$?
Il s'agit aussi du barycentre de $(A, 1)$ et $(B, 1)$ ou de (A, α) et (B, α) avec $\alpha \in \mathbb{R}^*$.
D'après la propriété 4.8 on a $\overrightarrow{AG_1} = \frac{1}{2}\overrightarrow{AB}$.
Autrement dit, G_1 est le milieu de $[AB]$, l'isobarycentre de A et B.

Quel est le barycentre G_2 de $(A, 50)$ et $(B, 100)$?
Il s'agit du barycentre de $(A, 1)$ et $(B, 2)$. Dans ce cas, on a $\overrightarrow{AG_2} = \frac{2}{3}\overrightarrow{AB}$.
Autrement dit, G_2 est au tiers de $[AB]$, en partant de B.

Quel est le barycentre G_3 de $(A, 50)$ et $(B, -50)$?
Il s'agirait du barycentre de $(A, 1)$ et $(B, -1)$ mais ce point n'existe pas car $\alpha + \beta = 1 - 1 = 0$.

Quel est le barycentre G_4 de $(A, 75)$ et $(B, -50)$?
Il s'agit du barycentre de $(A, 3)$ et $(B, -2)$. On a $\overrightarrow{AG_4} = -2\overrightarrow{AB} = 2\overrightarrow{BA}$.
Autrement dit, G_4 est sur la demi-droite $[BA)$, A étant au tiers du segment $[BG]$, en partant de B (voir la figure ci-dessous).



> Propriété 4.9 Le barycentre G de (A, α) et (B, β), lorsqu'il existe, est sur la droite (AB).
> G appartient au segment $[AB]$ lorsque α et β sont de même signe et, dans tous les cas :
> - G est plus proche de A si $|\alpha| > |\beta|$
> - G est le milieu de $[AB]$ si $\alpha = \beta$
> - G est plus proche de B si $|\alpha| < |\beta|$

Démonstration On traduit vectoriellement $M \in (AB)$ par $\exists t \in \mathbb{R}, \overrightarrow{AM} = t\overrightarrow{AB}$.
Cette égalité s'écrivant $\overrightarrow{AM} = t(\overrightarrow{MB} - \overrightarrow{MA}) \iff (1-t)\overrightarrow{MA} + t\overrightarrow{MB}$, et comme $(1-t) + t = 1 \neq 0$ on en déduit que M est le barycentre de $(A, 1-t)$ et (B, t).

De même, $M \in [AB]$ par $\exists t \in [0, 1], \overrightarrow{AM} = t\overrightarrow{AB}$.
M est le barycentre de $(A, 1-t)$ et (B, t) avec $t \in [0, 1]$ et $1 - t \in [0, 1]$.

L'égalité qui définit G peut se noter $\alpha\overrightarrow{GA} = -\beta\overrightarrow{GA}$.
Les normes de ces deux vecteurs égaux sont égales ; on a donc : $|\alpha| \times ||\overrightarrow{GA}|| = |\beta| \times ||\overrightarrow{GB}||$.
Autrement dit : $\frac{||\overrightarrow{GA}||}{||\overrightarrow{GB}||} = \frac{|\beta|}{|\alpha|}$ ou encore $\frac{GA}{GB} = \frac{|\beta|}{|\alpha|}$
Donc G est plus proche de A que de B $\iff GA < GB \iff |\beta| < |\alpha|$.

Exemple 35 (point comme barycentre) – Soient trois points A, B et C alignés.
Connaissant les positions respectives de trois points, exprimons l'un comme barycentre des deux autres affectés des coefficients qui conviennent. Supposons $AB = 2$, $BC = 5$ et $AC = 7$ (voir figure).
La somme $AB + BC = AC$ nous assure que $B \in [AC]$.

B appartient à $[AC]$: les coefficients cherchés sont donc de même signe.
Comme $\frac{BA}{BC} = \frac{2}{5}$, B est le barycentre de $(A, 5)$ et $(C, 2)$.

$A \notin [BC]$: les coefficients sont de signes opposés.
Comme $\frac{AB}{AC} = \frac{2}{7}$, A est le barycentre de $(B, 7)$ et $(C, -2)$ (ou de $(B, -7)$ et $(C, 2)$).

De la même manière, on montre que C est le barycentre de $(A, -5)$ et $(B, 7)$.

PROPRIÉTÉ 4.10 Soient (A,α) et (B,β) tels que $\alpha+\beta\neq 0$ et G leur barycentre.

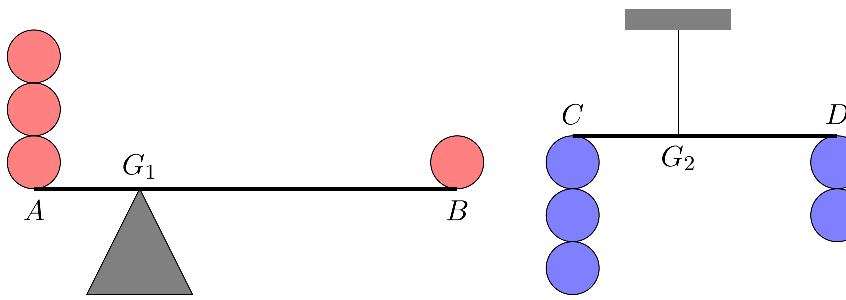
Pour tout point M on a : $\begin{cases} 1) \quad \alpha\overrightarrow{MA}+\beta\overrightarrow{MB}=(\alpha+\beta)\overrightarrow{MG} \\ 2) \quad \overrightarrow{MG}=\dfrac{\alpha}{\alpha+\beta}\overrightarrow{MA}+\dfrac{\beta}{\alpha+\beta}\overrightarrow{MB} \end{cases}$

DÉMONSTRATION Décomposons et regroupons :
$\alpha\overrightarrow{MA}+\beta\overrightarrow{MB}=\alpha(\overrightarrow{MG}+\overrightarrow{GA})+\beta(\overrightarrow{MG}+\overrightarrow{GB})=(\alpha+\beta)\overrightarrow{MG}+\alpha\overrightarrow{GA}+\beta\overrightarrow{GB}$.
Utilisons la définition du point G : $\alpha\overrightarrow{GA}+\beta\overrightarrow{GB}=\vec{0}$, on obtient : $\alpha\overrightarrow{MA}+\beta\overrightarrow{MB}=(\alpha+\beta)\overrightarrow{MG}+\vec{0}$
La 2^{e} égalité est équivalente à la 1re car on ne fait que diviser par $\alpha+\beta\neq 0$.

Remarques :

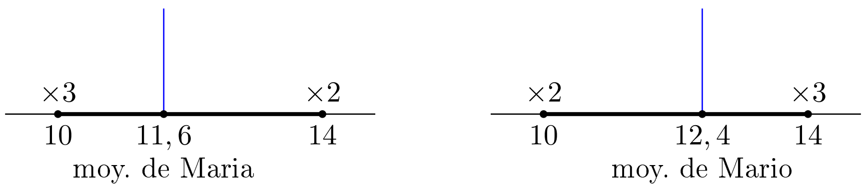
- Pour retrouver les relations connues, remplaçons M par A ou G dans ces égalités :
  $M=A$ dans 2) : $\overrightarrow{AG}=\frac{\beta}{\alpha+\beta}\overrightarrow{AB}$
  $M=G$ dans 1) : $\alpha\overrightarrow{GA}+\beta\overrightarrow{GB}=\vec{0}$
- L'origine du repère $(O,\vec{i},\vec{j},\vec{k})$ à la place de M dans 2) conduit aux coordonnées de G :
  $\overrightarrow{OG}=\frac{\alpha}{\alpha+\beta}\overrightarrow{OA}+\frac{\beta}{\alpha+\beta}\overrightarrow{OB}=\frac{\alpha}{\alpha+\beta}(x_A\vec{i}+y_A\vec{j}+z_A\vec{k})+\frac{\beta}{\alpha+\beta}(x_B\vec{i}+y_B\vec{j}+z_B\vec{k})$.
  Les coordonnées de G dans le repère $(O,\vec{i},\vec{j},\vec{k})$ sont donc les moyennes pondérées :
  $$x_G=\frac{\alpha x_A+\beta x_B}{\alpha+\beta},\ y_G=\frac{\alpha y_A+\beta y_B}{\alpha+\beta} \text{ et } z_G=\frac{\alpha z_A+\beta z_B}{\alpha+\beta}$$
  Pour l'isobarycentre I de A et B (le milieu de $[AB]$), ses coordonnées sont :
  $$x_I=\frac{x_A+x_B}{2},\ y_I=\frac{y_A+y_B}{2} \text{ et } z_I=\frac{z_A+z_B}{2}$$

Centre de gravité et moyenne :



Le principe de la balance ou du mobile repose sur la notion de barycentre.
Le point où s'équilibre l'ensemble est le barycentre des extrémités où s'appliquent les masses :

- à gauche, la balance est équilibrée au barycentre G_1 de $(A,3)$ et $(B,1)$, soit au quart de $[AB]$
- à droite, le mobile est équilibré au barycentre G_2 de $(C,3)$ et $(D,2)$, soit aux $\frac{2}{5}$ de $[CD]$



Un professeur de maths a fait deux contrôles dans le trimestre : le 1er avait duré 3 heures, le 2^{d} 2 heures. Il fait sa moyenne en calculant la moyenne pondérée des deux notes, affectées des coefficients 3 et 2. La note moyenne va ainsi devenir le barycentre des deux notes avec les coefficients de la pondération. Par exemple, Maria a eu 10 et 14 et Mario 14 et 10. La moyenne de Maria est $\frac{10\times3+14\times2}{5}=\frac{58}{5}=11,6$ tandis que celle de de Mario est $\frac{14\times3+10\times2}{5}=\frac{62}{5}=12,4$.
Dans les deux cas, elle s'est installée aux $\frac{2}{5}$ de l'intervalle séparant la 1re et la 2^{e} note.

2.b. Barycentre de trois points (ou plus)

DÉFINITION 4.10 Soient (A, α), (B, β) et (C, γ) trois points pondérés tels que $\alpha + \beta + \gamma \neq 0$.
On appelle barycentre de (A, α) , (B, β) et (C, γ) le point G défini par : $\alpha\overrightarrow{GA} + \beta\overrightarrow{GB} + \gamma\overrightarrow{GC} = \vec{0}$.
De manière plus générale, le barycentre de n points pondérés (A_i, α_i) tels que $\sum_{i=1}^{n} \alpha_i \neq 0$ est défini par la relation vectorielle $\sum_{i=1}^{n} \alpha_i\overrightarrow{GA_i} = \vec{0}$.

PROPRIÉTÉ 4.11 (BARYCENTRE DE TROIS POINTS) Si $\alpha+\beta+\gamma \neq 0$, le barycentre G des points pondérés (A, α), (B, β) et (C, γ) vérifie, pour tout point M, les égalités :

$$\begin{cases} 1) \quad \alpha\overrightarrow{MA} + \beta\overrightarrow{MB} + \gamma\overrightarrow{MC} = (\alpha + \beta + \gamma)\overrightarrow{MG} \\ 2) \quad \overrightarrow{MG} = \dfrac{\alpha}{(\alpha + \beta + \gamma)}\overrightarrow{MA} + \dfrac{\beta}{(\alpha + \beta + \gamma)}\overrightarrow{MB} + \dfrac{\gamma}{(\alpha + \beta + \gamma)}\overrightarrow{MC} \end{cases}$$

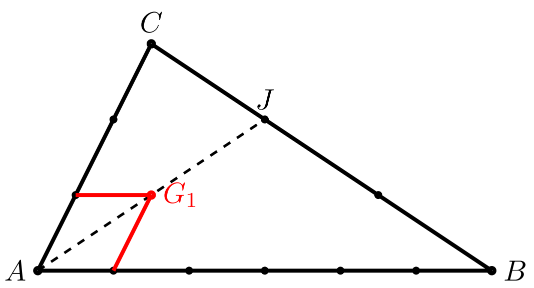
En particulier, lorsque $M = A$, on a $\overrightarrow{AG} = \frac{\beta}{(\alpha+\beta+\gamma)}\overrightarrow{AB} + \frac{\gamma}{(\alpha+\beta+\gamma)}\overrightarrow{AC}$.

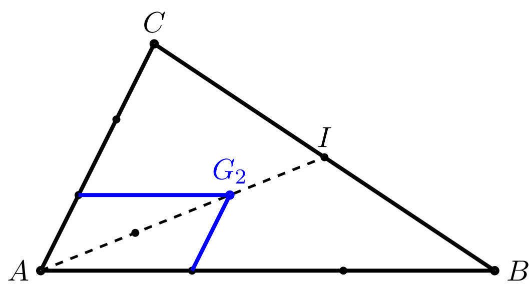
DÉMONSTRATION Ces relations découlent de la définition, en décomposant la relation vectorielle avec la relation de Chasles et les différentes propriétés du calcul vectoriel.
La dernière relation assure de l'existence et de l'unicité du point G puisqu'elle en explicite la construction à partir des données ; elle prouve par ailleurs que $G \in (ABC)$.

Remarques :

- Comme pour le barycentre de deux points, le barycentre de (A, α), (B, β) et (C, γ) est aussi celui de $(A, k\alpha)$, $(B, k\beta)$ et $(C, k\gamma)$, quel que soit le réel non nul k.
  On peut toujours remplacer les coefficients du barycentre par des coefficients proportionnels.
- Lorsqu'on met l'origine du repère $(O, \vec{i}, \vec{j}, \vec{k})$ à la place de M, on obtient :
  $\overrightarrow{OG} = \frac{\alpha}{(\alpha+\beta+\gamma)}\overrightarrow{OA} + \frac{\beta}{(\alpha+\beta+\gamma)}\overrightarrow{OB} + \frac{\gamma}{(\alpha+\beta+\gamma)}\overrightarrow{OC}$. En passant aux coordonnées, on obtient
  $$x_G = \frac{\alpha x_A + \beta x_B + \gamma x_C}{\alpha + \beta + \gamma},\ y_G = \frac{\alpha y_A + \beta y_B + \gamma y_C}{\alpha + \beta + \gamma} \text{ et } z_G = \frac{\alpha z_A + \beta z_B + \gamma z_C}{\alpha + \beta + \gamma}$$
- Si un des coefficients est nul, par exemple $\alpha = 0$, le barycentre de $(A, 0)$, (B, β) et (C, γ) est le barycentre des deux points (B, β) et (C, γ) ; il est donc sur (BC).
- Si tous les coefficients sont égaux, le barycentre est celui de $(A, 1)$, $(B, 1)$ et $(C, 1)$.
  Il est appelé *isobarycentre* de A, B,t C et correspond au centre de gravité du triangle ABC.
- Le barycentre de (A, α), (B, β), (C, γ) appartient au plan (ABC). Réciproquement, un point M de ce plan $(ABC) = \mathcal{P}(A, \overrightarrow{AB}, \overrightarrow{AC})$ est défini par ses coordonnées (x, y) qui vérifient $\overrightarrow{AM} = x\overrightarrow{AB} + y\overrightarrow{AC} \iff (1 - x - y)\overrightarrow{MA} + x\overrightarrow{AB} + y\overrightarrow{AC} = \vec{0}$, cela prouve que M est le barycentre de $(A, 1 - x - y)$, (B, x), (C, y) puisque $(1 - x - y) + x + y = 1 \neq 0$.
  Reformulons cette remarque : le plan (ABC) est l'ensemble de tous les barycentres de A, B, C.

EXEMPLE 36 – Quel est le barycentre G_1 de $(A, 12)$, $(B, 4)$ et $(C, 8)$?
G_1 est barycentre de $(A, 3)$, $(B, 1)$, $(C, 2)$; d'après la propriété 4.11 on a $\overrightarrow{AG_1} = \frac{1}{6}\overrightarrow{AB} + \frac{1}{3}\overrightarrow{AC}$.
Quel est le barycentre G_2 de $(A, 24)$, $(B, 24)$ et $(C, 24)$?
G_2 est l'isobarycentre de A, B, C ; d'après la propriété 4.11 on a $\overrightarrow{AG_2} = \frac{1}{3}\overrightarrow{AB} + \frac{1}{3}\overrightarrow{AC}$.





Remarque : dans cette dernière construction, on remarque que $\overrightarrow{AG_2} = \frac{1}{3}(\overrightarrow{AB} + \overrightarrow{AC})$.
En introduisant le milieu I de $[BC]$, $\overrightarrow{AG_2} = \frac{1}{3}(2\overrightarrow{AI}) = \frac{2}{3}\overrightarrow{AI}$.
Cela signifie que G_2 est le barycentre de $(A,1)$ et $(I,2)$.
Le barycentre de $(A,1)$, $(B,1)$ et $(C,1)$ est en fait le barycentre de $(A,1)$ et $(I,2)$.
On remplace le couple $((B,1),(C,1))$ par leur barycentre affecté de la somme des coefficients.
Cette remarque se généralise dans la propriété 4.12.
Pour placer le point G_1 – barycentre de $(A,3)$, $(B,1)$, $(C,2)$ – on peut remplacer $((B,1),(C,2))$ par leur barycentre J affecté du coefficient 3. G_1 apparait alors comme le barycentre de $(A,3)$ et $(J,3)$, c'est-à-dire le milieu de $[AJ]$.

Propriété 4.12 (Réduction de barycentres) Le barycentre G de (A,α), (B,β) et (C,γ) est aussi celui de (A,α), $(D,\beta+\gamma)$ si $\beta+\gamma \neq 0$, D étant le barycentre de (B,β) et (C,γ).

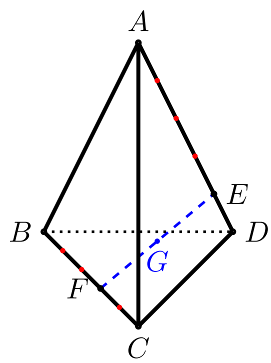
Démonstration Montrons cette propriété. On a : $\alpha\overrightarrow{GA} + \beta\overrightarrow{GB} + \gamma\overrightarrow{GC} = \vec{0}$ avec $\alpha+\beta+\gamma \neq 0$.
Au moins une somme de deux des coefficients doit être non nulle, sinon la somme des trois coefficients serait nulle ce qui ne se peut pas.
Quitte à en changer, supposons que $\beta+\gamma \neq 0$: le barycentre D de (B,β) et (C,γ) existe donc.
Ce point vérifie, selon la propriété 4.10, en remplaçant M par G, $\beta\overrightarrow{GB} + \gamma\overrightarrow{GC} = (\beta+\gamma)\overrightarrow{GD}$.
On peut alors écrire $\alpha\overrightarrow{GA} + (\beta+\gamma)\overrightarrow{GD} = \vec{0}$ ce qui prouve la propriété énoncée.

Remarque :

- On fait de même avec le barycentre de quatre points : on peut en remplacer trois par leur barycentre affecté de la somme des trois coefficients (à condition qu'elle ne soit pas nulle) ou bien regrouper les points par paires dont on détermine les barycentres ; le barycentre cherché apparaissant comme celui des deux barycentres intermédiaires, affectés des coefficients qui conviennent (toujours à condition que ces coefficients ne soient pas nuls).
- Le barycentre G de (A,α), (B,β), (C,γ) appartient à l'intérieur du triangle ABC si α, β et γ sont de même signe. En effet, la propriété 4.9 indique que le barycentre C' de (A,α), (B,β) appartient à $[AB]$ si α et β sont de même signe et, comme G est aussi le barycentre de $(C',\alpha+\beta)$ et (C,γ), puisque $\alpha+\beta$ et γ sont de même signe, $G \in [CC']$.
- Cette possibilité de réduction a de nombreuses applications. Supposons que l'on cherche à caractériser l'ensemble des points M tels que $||3\overrightarrow{MA} + \overrightarrow{MB} + 2\overrightarrow{MC}|| = AB + BC + CA$.
Dans l'exemple 36 on a nommé G_1 le barycentre de $(A,3)$, $(B,1)$, $(C,2)$.
On peut donc substituer à la somme $3\overrightarrow{MA} + \overrightarrow{MB} + 2\overrightarrow{MC}$ le vecteur $6\overrightarrow{MG_1}$ qui lui est égal.
Ainsi, on obtient $||6\overrightarrow{MG_1}|| = AB + BC + CA \iff MG_1 = \frac{AB+BC+CA}{6}$.
Les points M sont sur une sphère de centre G_1 et de rayon $\frac{AB+BC+CA}{6}$ (le sixième du périmètre du triangle ABC).

Exemple 37 (Barycentre de 4 points) – Soit G le barycentre de $(A,1)$, $(B,2)$, $(C,3)$ et $(D,4)$

Remplaçons le couple $((A,1),(D,4))$ par le point pondéré $(E,5)$ où E est le barycentre de $(A,1)$ et $(D,4)$. De même, remplaçons le couple $((B,2),(C,3))$ par le point pondéré $(F,5)$ où F est le barycentre de $(B,2)$, $(C,3)$. D'après la propriété 4.12, Le point G cherché est l'isobarycentre de E et F, le milieu de $[EF]$. La construction de G se limite à celle des barycentres partiels E et F puis à leur milieu.



Noter que le choix d'une association judicieuse est hautement préférable à une association arbitraire, relevant d'un processus systématique (aveugle par essence) ou aléatoire. On aurait pu, ici, chercher le barycentre H de $(B,2)$, $(C,3)$ et $(D,4)$ et ensuite celui de $(H,9)$ et $(A,1)$. Mais cette solution, guidée par le choix de la représentation où BCD semble une base toute indiquée pour le tétraèdre, ne parait pas aussi simple à mettre en œuvre.

<u>Notation</u> : Pour simplifier, je vais noter « M' est le barycentre de $(A,1)$, $(B,1)$, $(C,1)$ et $(M,-1)$ » sous la forme plus condensée $M' = \mathscr{B}\{(A,1),(B,1),(C,1),(M,-1)\}$.

Exemple 38 (Du plan à l'espace) – Soient A, B et C trois points non alignés et M un point quelconque. On complète la figure avec sept points :

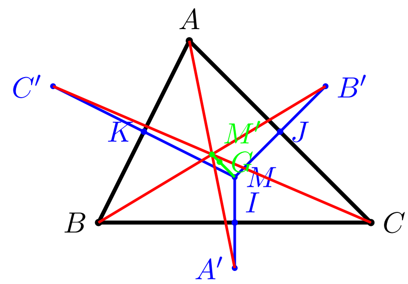
- l'isobarycentre G de A, B, C ce qu'on note $G = \mathscr{B}\{(A,1),(B,1),(C,1)\}$
- les points I, J et K qui sont les milieux respectifs de $[BC]$, $[CA]$ et $[AB]$
- les points A', B' et C' qui sont les symétriques de M par rapport à I, J et K

Prouvons que les droites (AA'), (BB') et (CC') sont concourantes en un point M', le milieu commun des segments $[AA']$, $[BB']$ et $[CC']$ et prouvons que $M' \in (MG)$.

$\boxed{\text{Partie 1 : } M \text{ appartient au plan } (ABC)}$

Soit M' le milieu de $[AA']$ ce qu'on note $M' = \mathscr{B}\{(A,1),(A',1)\}$.
Montrons que M' est aussi le milieu de $[BB']$ et $[CC']$.
I étant le milieu de $[MA']$, on a $A' = \mathscr{B}\{(I,2),(M,-1)\}$.
(car $\overrightarrow{A'I} = \frac{1}{2}\overrightarrow{A'M} \iff 2\overrightarrow{A'I} - \overrightarrow{A'M} = \overrightarrow{0}$) et, comme $2+(-1)=1$,
on peut remplacer $(A',1)$ par le couple $((I,2),(M,-1))$:
$M' = \mathscr{B}\{(A,1),(A',1)\} = \mathscr{B}\{(A,1),(I,2),(M,-1)\}$

I étant le milieu de $[BC]$, $I = \mathscr{B}\{(B,1),(C,1)\}$.
Donc $(I,2)$ peut être remplacé par le couple $((B,1),(C,1))$.
Finalement $M' = \mathscr{B}\{(A,1),(B,1),(C,1),(M,-1)\}$.

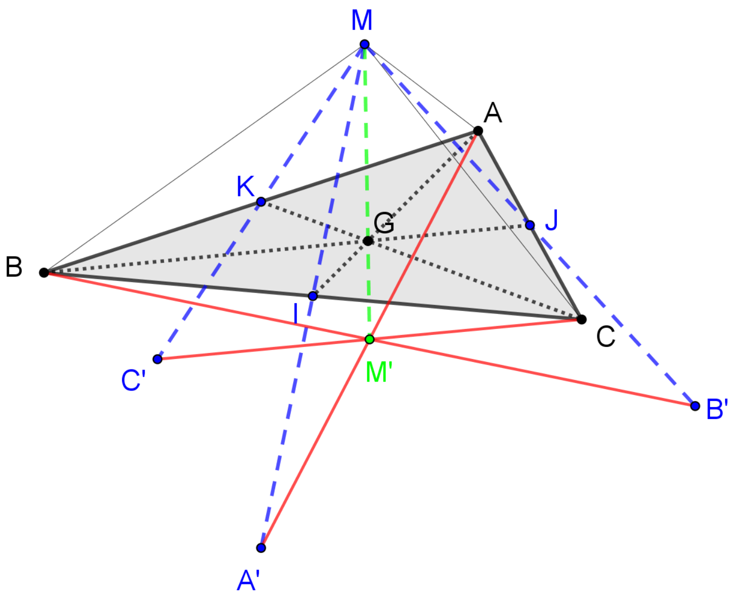


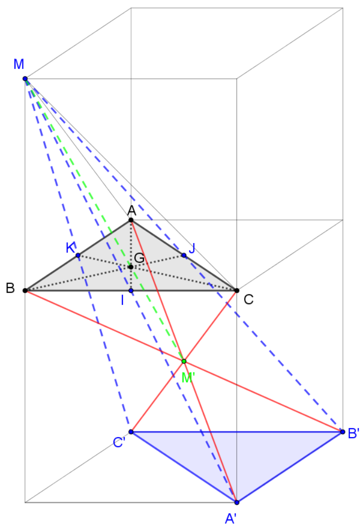
Comme J est le milieu de $[AC]$, $J = \mathscr{B}\{(A,1),(C,1)\}$ et donc $M' = \mathscr{B}\{(J,2),(B,1),(M,-1)\}$.
Comme J est le milieu de $[MB']$, $B' = \mathscr{B}\{(J,2),(M,-1)\}$ et donc $M' = \mathscr{B}\{(B',1),(B,1)\}$.
Cette dernière relation montre bien que M' est le milieu de $[BB']$, ce que l'on devait prouver.
On procède de la même manière pour montrer que M' est le milieu de $[CC']$.

Puisque $M' = \mathscr{B}\{(A,1),(B,1),(C,1),(M,-1)\}$, comme $G = \mathscr{B}\{(A,1),(B,1),(C,1)\}$,
on remplace le triplet $((A,1),(B,1),(C,1))$ par $(G,3)$ et on obtient $M' = \mathscr{B}\{(G,3),(M,-1)\}$,
ce qui prouve que M, G et M' sont alignés.

$\boxed{\text{Partie 2 : } M \text{ n'appartient pas au plan } (ABC)}$

Dans ce cas, que doit-on changer ? Rien. Les points G, I, J et K appartiennent toujours au plan (ABC) (par définition) tandis que les points A', B', C' et M' n'y appartiennent pas.
On peut ajouter que si M est au-dessus de ce plan, A', B', C' et M' sont en-dessous.
J'ai représenté la situation avec des segments en tiretés pour ceux qui traversent le plan (ABC).





3. Colinéarité et Orthogonalité

3.a. Colinéarité

PROPRIÉTÉ 4.13 (REPRÉSENTATION PARAMÉTRIQUE DE DROITE) Soients $(O, \overrightarrow{i}, \overrightarrow{j}, \overrightarrow{k})$ un repère, $M(x_0, y_0, z_0)$ un point et $\overrightarrow{u}(\alpha, \beta, \gamma) \neq \overrightarrow{0}$ un vecteur.
$M(x, y, z) \in d(M_0, \overrightarrow{u}) \iff \exists t \in \mathbb{R},\ \overrightarrow{M_0M} = t\overrightarrow{u}$, autrement dit :

$$\begin{cases} x = x_0 + t\alpha \\ y = y_0 + t\beta \\ z = z_0 + t\gamma \end{cases}$$

DÉMONSTRATION $M(x, y, z) \in d(M_0, \overrightarrow{u})$ si et seulement si $\overrightarrow{M_0M}$ et $\overrightarrow{u}$ sont colinéaires.
Cela se traduit vectoriellement $\exists t \in \mathbb{R},\ \overrightarrow{M_0M} = t\overrightarrow{u} \iff \overrightarrow{OM} = \overrightarrow{OM_0} + t\overrightarrow{u}$.
Cela équivaut à $\exists t \in \mathbb{R},\ x\overrightarrow{i} + y\overrightarrow{j} + z\overrightarrow{k} = (x_0\overrightarrow{i} + y_0\overrightarrow{j} + z_0\overrightarrow{k}) + t(\alpha\overrightarrow{i} + \beta\overrightarrow{j} + \gamma\overrightarrow{k})$.
Comme $(\overrightarrow{i}, \overrightarrow{j}, \overrightarrow{k})$ est une base de l'espace, la conclusion s'en déduit.

Remarques :

- Pour une même droite, il existe une infinité de représentations paramétriques différentes : il suffit de remplacer le point M_0 par un autre point M'_0 de la droite ou de choisir un autre vecteur directeur $\overrightarrow{u'}$ colinéaire à $\overrightarrow{u}$.
  Supposons qu'une représentation paramétrique de la droite d soit
  $$M(x, y, z) \in d \iff \begin{cases} x = 1 + 3t \\ y = 2 + 2t \\ z = 3 + t \end{cases} \quad \text{avec } t \in \mathbb{R}$$
  Avec $M_0(1, 2, 3)$ et $\overrightarrow{u}(3, 2, 1)$ la droite d est $d(M_0, \overrightarrow{u})$.
  Mais, en prenant $t = -1$ on a $M'_0(-2, 0, 2) \in d$ et $\overrightarrow{u'}(6, 4, 2)$ est un autre vecteur directeur de d.
  On peut donc proposer d'autres représentations paramétriques de d :
  $$\begin{cases} x = -2 + 3t \\ y = 2t \\ z = 2 + t \end{cases} \quad \text{ou} \quad \begin{cases} x = -2 + 6t \\ y = 4t \\ z = 2 + 2t \end{cases} \quad \text{ou} \quad \begin{cases} x = 1 + 6t \\ y = 2 + 4t \\ z = 3 + 2t \end{cases} \quad \text{, avec } t \in \mathbb{R}$$
- Une représentation paramétrique de demi-droite s'obtient en prenant $t \geqslant 0$ ou $t \leqslant 0$;
  pour la demi-droite $[AB)$, on prend la représentation paramétrique de $d(A, \overrightarrow{AB})$ avec $t \geqslant 0$;
  pour la demi-droite $[BA)$, on prend la représentation paramétrique de $d(B, \overrightarrow{BA})$ avec $t \geqslant 0$,
  mais bien sûr on peut aussi prendre la représentation paramétrique de $d(B, \overrightarrow{AB})$ avec $t \leqslant 0$.
- Une représentation paramétrique de segment s'obtient en prenant $t \in [0, 1]$ ou $t \in [0, 1]$;
  pour le segment $[AB]$, on prend la représentation paramétrique de $d(A, \overrightarrow{AB})$ avec $t \in [0, 1]$
  ou bien la représentation paramétrique de $d(B, \overrightarrow{AB})$ avec $t \in [-1, 0]$.

PROPRIÉTÉ 4.14 (REPRÉSENTATION PARAMÉTRIQUE D'UN PLAN) Soients $(O, \overrightarrow{i}, \overrightarrow{j}, \overrightarrow{k})$ un repère, $M(x_0, y_0, z_0)$ un point $\overrightarrow{u}(\alpha, \beta, \gamma)$ et $\overrightarrow{v}(\alpha', \beta', \gamma')$ deux vecteurs non colinéaires.
$M(x, y, z) \in \mathcal{P}(M_0, \overrightarrow{u}, \overrightarrow{v}) \iff \exists (t, t') \in \mathbb{R}^2,\ \overrightarrow{M_0M} = t\overrightarrow{u} + t'\overrightarrow{v}$, autrement dit :

$$\begin{cases} x = x_0 + t\alpha + t'\alpha' \\ y = y_0 + t\beta + t'\beta' \\ z = z_0 + t\gamma + t'\gamma' \end{cases}$$

DÉMONSTRATION La démonstration est calquée sur la précédente.
D'après la définition 4.7, $M(x,y,z)\in\mathcal{P}(M_0,\overrightarrow{u},\overrightarrow{v})$ si et seulement si $\overrightarrow{M_0M}$ est une combinaison linéaire d'une base du plan $\mathcal{P}$, ou $\overrightarrow{u}$ et $\overrightarrow{v}$, étant non-colinéaires, forment une base de ce plan.
Cela se traduit vectoriellement $\exists(t,t')\in\mathbb{R}^2,\ \overrightarrow{M_0M}=t\overrightarrow{u}+t'\overrightarrow{v}\iff\overrightarrow{OM}=\overrightarrow{OM_0}+t\overrightarrow{u}+t'\overrightarrow{v}$.
Cela équivaut à
$\exists(t,t')\in\mathbb{R}^2,\ x\overrightarrow{i}+y\overrightarrow{j}+z\overrightarrow{k}=(x_0\overrightarrow{i}+y_0\overrightarrow{j}+z_0\overrightarrow{k})+t(\alpha\overrightarrow{i}+\beta\overrightarrow{j}+\gamma\overrightarrow{k})+t'(\alpha'\overrightarrow{i}+\beta'\overrightarrow{j}+\gamma'\overrightarrow{k})$.
Comme $(\overrightarrow{i},\overrightarrow{j},\overrightarrow{k})$ est une base de l'espace, la conclusion s'en déduit.

Remarques :

- De même que pour les droites, et pour les mêmes raisons, la représentation paramétrique d'un plan n'est pas unique. Par contre, il est possible de déterminer une équation cartésienne d'un plan (voir à ce sujet la propriété 4.20) alors que ce type de représentation n'existe pas pour les droites.

  Supposons qu'une représentation paramétrique du plan $\mathcal{P}$ soit $\begin{cases}x=1+3t-t'\\y=2+2t-2t'\\z=3+t-3t'\end{cases}$

  Exprimons t et t' en fonction de x, y et z :
  $\begin{cases}3t-t'=x-1\\2t-2t'=y-2\\t-3t'=z-3\end{cases}\iff\begin{cases}6t-2t'=2x-2\\2t-2t'=y-2\\2t-6t'=2z-6\end{cases}\iff\begin{cases}4t=2x-y\\2t-2t'=y-2\\4t'=y-2z+4\end{cases}$
  On trouve par soustraction $4t=2x-y$ et $4t'=y-2z+4$, et en remplaçant dans la 3$^{\text{e}}$ équation : $2t-2t'=\frac{2x-y}{2}-\frac{y-2z+4}{2}=\frac{2x-y-(y-2z+4)}{2}=\frac{2x-2y+2z-4}{2}=y-2$, soit
  $2x-2y+2z-4=2(y-2)\iff 2x-4y+2z=0\iff x-2y+z=0$.
  L'équation cartésienne de $\mathcal{P}$ est donc $x-2y+z=0$.

  Vérifions : $1-2\times2+3=0\Longrightarrow M_0(1,2,3)\in\mathcal{P}$
  De plus, si $\overrightarrow{u}(3,2,1)$ et $\overrightarrow{v}(-1,-2,-3)$ sont les vecteurs directeurs de $\mathcal{P}$, on doit pouvoir trouver $(M_1,M_2)\in\mathcal{P}^2$ tels que $\overrightarrow{M_0M_1}=\overrightarrow{u}$ et $\overrightarrow{M_0M_2}=\overrightarrow{v}$.

  — L'égalité $\overrightarrow{M_0M_1}=\overrightarrow{u}$ conduit à prendre $(t,t')=(1,0)$: $\begin{cases}x=1+3=4\\y=2+2=4\\z=3+1=4\end{cases}$ d'où $M_1(4,4,4)$

  — L'égalité $\overrightarrow{M_0M_2}=\overrightarrow{v}$ conduit à prendre $(t,t')=(0,1)$: $\begin{cases}x=1-1=0\\y=2-2=0\\z=3-3=0\end{cases}$ d'où $M_2(0,0,0)$

  Il reste à vérifier $4-2\times4+4=0\Longrightarrow M_1(4,4,4)\in\mathcal{P}$ et $0-2\times0+0=0\Longrightarrow M_2(0,0,0)\in\mathcal{P}$.
- En jouant sur les intervalles dans lesquels on choisit t et t', on doit pouvoir donner une représentation paramétrique d'un secteur angulaire (en prenant $t\geqslant0$ et $t'\geqslant0$), d'un parallélogramme (en prenant $t\in[0,a]$ et $t'\in[0,b]$, a et b étant des nombres positifs), d'une bande limités par deux droites parallèles (en prenant $t\in[0,a]$ et $t'\in\mathbb{R}$).

3.b. Produit scalaire

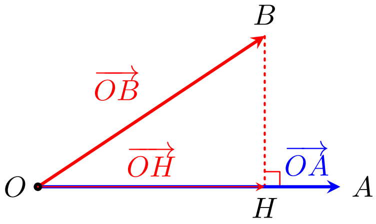
DÉFINITION 4.11 Soient $\vec{u}$ et $\vec{v}$ deux vecteurs de l'espace, leur produit scalaire est noté $\vec{u}\cdot\vec{v}$.
- si $\vec{u}=\vec{0}$ ou $\vec{v}=\vec{0}$, alors $\vec{u}\cdot\vec{v}=0$
- si $\vec{u}\neq\vec{0}$ et $\vec{v}\neq\vec{0}$

  — Si $\vec{u}$ et $\vec{v}$ sont colinéaires, alors $\vec{u}\cdot\vec{v}=\begin{cases}\|\vec{u}\|\times\|\vec{v}\| \text{ si } \vec{u} \text{ et } \vec{v} \text{ ont même sens}\\-\|\vec{u}\|\times\|\vec{v}\| \text{ si } \vec{u} \text{ et } \vec{v} \text{ sont de sens contraire}\end{cases}$

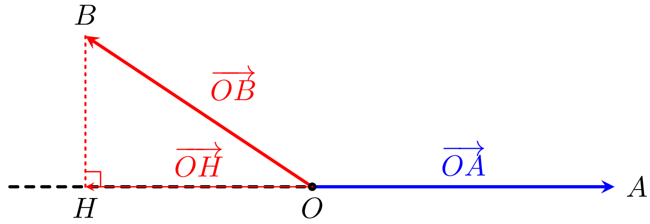
  — si $\vec{u}$ et $\vec{v}$ ne sont pas colinéaires. O, A et B étant des points tels que $\overrightarrow{OA}=\vec{u}$ et $\overrightarrow{OB}=\vec{v}$, $\vec{u}\cdot\vec{v}$ est égal au produit scalaire $\overrightarrow{OA}\cdot\overrightarrow{OB}$ calculé dans le plan (OAB)

Remarques :

- On a appris, en 1[re], à calculer le produit scalaire dans un plan.
  Rappelons qu'en notant H le projeté orthogonal de B sur (OA) on a $\overrightarrow{OA}\cdot\overrightarrow{OB}=\overrightarrow{OA}\cdot\overrightarrow{OH}$
  — si $\overrightarrow{OA}$ et $\overrightarrow{OH}$ sont de même sens alors $\overrightarrow{OA}\cdot\overrightarrow{OB}=OA\times OH$
  — si $\overrightarrow{OA}$ et $\overrightarrow{OH}$ sont de sens contraire alors $\overrightarrow{OA}\cdot\overrightarrow{OB}=-OA\times OH$



$\vec{u}\cdot\vec{v}=OA\times OH$



$\vec{u}\cdot\vec{v}=-OA\times OH$

- En notant θ l'angle orienté des vecteurs $\vec{u}$ et $\vec{v}$ $\begin{cases}\text{si } -\frac{\pi}{2}\leqslant\theta\leqslant\frac{\pi}{2}\ [2\pi] \text{ alors } \vec{u}\cdot\vec{v}\geqslant 0\\ \text{si } \frac{\pi}{2}\leqslant\theta\leqslant\frac{3\pi}{2}\ [2\pi] \text{ alors } \vec{u}\cdot\vec{v}\leqslant 0\end{cases}$
- Le carré scalaire est noté $\vec{u}^2$ et on a $\vec{u}^2=||\vec{u}||^2$. En particulier $\overrightarrow{AB}^2=||\overrightarrow{AB}||^2=AB^2$.

Définition 4.12 (Droites orthogonales) Deux droites d et d' sont dites *orthogonales*, et on note $d\perp d'$, si les parallèles à ces droites passant par un point quelconque sont perpendiculaires.

Définition 4.13 (Vecteurs orthogonaux) Soient $\vec{u}$ et $\vec{v}$ deux vecteurs de l'espace.
Les vecteurs $\vec{u}$ et $\vec{v}$ sont *orthogonaux*, et on note $\vec{u}\perp\vec{v}$, si $\vec{u}=\vec{0}$ ou $\vec{v}=\vec{0}$ ou,
en notant O, A et B des points tels que $\overrightarrow{OA}=\vec{u}$ et $\overrightarrow{OB}=\vec{v}$, on a $(OA)\perp(OB)$.

Propriété 4.15 Pour tous vecteurs $\vec{u}$ et $\vec{v}$, $\vec{u}\perp\vec{v}\iff\vec{u}\cdot\vec{v}=0$

Démonstration Soient $\vec{u}$ et $\vec{v}$ deux vecteurs orthogonaux, on a deux cas :

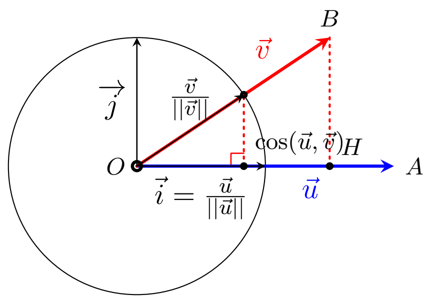
- si $\vec{u}=\vec{0}$ ou $\vec{v}=\vec{0}$, alors le produit scalaire est évidemment nul
- si $\vec{u}$ et $\vec{v}$ sont non-nuls, avec les notations de la définition 4.11, on a $(OA)\perp(OB)$, le projeté orthogonal de B sur (OA) est O donc $\vec{u}\cdot\vec{v}=OA\times OO=0$

Réciproquement, soient $\vec{u}$ et $\vec{v}$ deux vecteurs tels que $\vec{u}\cdot\vec{v}=0$. Comme $\vec{u}\cdot\vec{v}=\pm OA\times OH$,

- soit $OA=0$ c'est-à-dire que $\vec{u}=\vec{0}$ donc $\vec{u}$ et $\vec{v}$ sont orthogonaux
- soit $OH=0$ c'est-à-dire que $\vec{v}=\vec{0}$ ou B appartient à la perpendiculaire à (OA) passant par O donc dans les deux cas, les vecteurs $\vec{u}$ et $\vec{v}$ sont orthogonaux

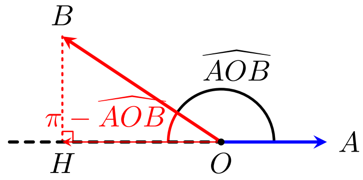
Propriété 4.16 (Vectoriellement) Si $\vec{u}$ et $\vec{v}$ sont deux vecteurs de l'espace alors $\vec{u}\cdot\vec{v}=||\vec{u}||\times||\vec{v}||\times\cos(\vec{u},\vec{v})$

Démonstration Soient O, A et B des points tels que $\overrightarrow{OA}=\vec{u}$ et $\overrightarrow{OB}=\vec{v}$. Le cercle trigonométrique de centre O est associé aux vecteurs $\vec{i}=\frac{\vec{u}}{||\vec{u}||}$ et $\vec{j}$ tels que $(\vec{i},\vec{j})=\frac{\pi}{2}\ [2\pi]$.
Les angles $(\vec{u},\vec{v})$ et $(\frac{\vec{u}}{||\vec{u}||},\frac{\vec{v}}{||\vec{v}||})$ étant égaux $\cos(\vec{u},\vec{v})=\cos(\frac{\vec{u}}{||\vec{u}||},\frac{\vec{v}}{||\vec{v}||})$.
La projection orthogonale de B sur (OA) étant notée H, on a
$\overrightarrow{OH}=||\vec{v}||\times\cos(\vec{u},\vec{v})\times\vec{i}$ et donc
$\vec{u}\cdot\vec{v}=\vec{u}\cdot\overrightarrow{OH}=||\vec{u}||\times||\vec{v}||\times\cos(\vec{u},\vec{v})$.



$\cos(\vec{u},\vec{v})$ conserve l'information sur le sens des vecteurs $\overrightarrow{OH}$ et $\overrightarrow{OA}$: il est positif s'ils ont même sens, et négatif sinon.

Remarque :
Appliquons à une configuration géométrique, cette propriété :
Si O, A et B sont trois points du plan,
alors $\overrightarrow{OA} \cdot \overrightarrow{OB} = OA \times OB \times \cos(\widehat{AOB})$



PROPRIÉTÉ 4.17 Soient $\vec{u}$, $\vec{v}$ et $\vec{w}$ trois vecteurs et k un réel.

$\vec{u} \cdot \vec{v} = \vec{v} \cdot \vec{u}$	$\|\|\vec{u}+\vec{v}\|\|^2 = \|\|\vec{u}\|\|^2 + \|\|\vec{v}\|\|^2 + 2\vec{u} \cdot \vec{v}$	$\vec{u} \cdot \vec{v} = \frac{\|\|\vec{u}\|\|^2+\|\|\vec{v}\|\|^2-\|\|\vec{u}-\vec{v}\|\|^2}{2}$
$(k\vec{u}) \cdot \vec{v} = k(\vec{u} \cdot \vec{v})$	$\|\|\vec{u}-\vec{v}\|\|^2 = \|\|\vec{u}\|\|^2 + \|\|\vec{v}\|\|^2 - 2\vec{u} \cdot \vec{v}$	$\vec{u} \cdot \vec{v} = \frac{\|\|\vec{u}+\vec{v}\|\|^2-\|\|\vec{u}\|\|^2-\|\|\vec{v}\|\|^2}{2}$
$(\vec{u}+\vec{v}) \cdot \vec{w} = \vec{u} \cdot \vec{w} + \vec{v} \cdot \vec{w}$	$(\vec{u}+\vec{v}) \cdot (\vec{u}-\vec{v}) = \|\|\vec{u}\|\|^2 - \|\|\vec{v}\|\|^2$	$\vec{u} \cdot \vec{v} = \frac{\|\|\vec{u}+\vec{v}\|\|^2-\|\|\vec{u}-\vec{v}\|\|^2}{4}$

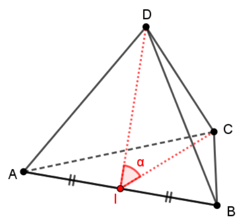
DÉMONSTRATION La symétrie ($\vec{u} \cdot \vec{v} = \vec{v} \cdot \vec{u}$) se montre à partir de la propriété 4.16 en utilisant la parité de la fonction cosinus ($\cos(-\theta) = \cos(\theta)$).

La bilinéarité – exprimée ici par rapport au 1^er^ vecteur ($(k\vec{u}) \cdot \vec{v} = k(\vec{u} \cdot \vec{v})$ et $(\vec{u}+\vec{v}) \cdot \vec{w} = \vec{u} \cdot \vec{w} + \vec{v} \cdot \vec{w}$) – se montre facilement en utilisant les expressions analytiques dans un repère orthonormé ; cela est fait plus loin.
Du fait de la symétrie, la linéarité par rapport au 2^e^ vecteur ($\vec{u} \cdot (k\vec{v}) = k(\vec{u} \cdot \vec{v})$ et $\vec{u} \cdot (\vec{v}+\vec{w}) = \vec{u} \cdot \vec{v} + \vec{u} \cdot \vec{w}$) – est automatique.

Les expressions de $\vec{u} \cdot \vec{v}$ impliquant des normes qui sont listées dans la 3^e^ colonne se déduisent des développements de $(\vec{u}+\vec{v})^2$, $(\vec{u}-\vec{v})^2$ et $(\vec{u}+\vec{v}) \cdot (\vec{u}-\vec{v})$ donnés dans la 2^e^ colonne.

Remarques : Si ABC est un triangle alors $\overrightarrow{AB} \cdot \overrightarrow{AC} = \dfrac{AB^2 + AC^2 - BC^2}{2}$
Cela vient de la propriété de la 1^re^ ligne, 3^e^ colonne, car en prenant $\vec{u} = \overrightarrow{AB}$, $\vec{v} = \overrightarrow{AC}$ on a $\vec{u} - \vec{v} = \overrightarrow{AB} - \overrightarrow{AC} = \overrightarrow{CB}$.
Si $ABDC$ est un parallélogramme alors $\overrightarrow{AB} \cdot \overrightarrow{AC} = \dfrac{AD^2 - AB^2 - AC^2}{2}$
Cela vient de la propriété de la 2^e^ ligne, 3^e^ colonne, car en prenant $\vec{u} = \overrightarrow{AB}$, $\vec{v} = \overrightarrow{AC}$ on a $\vec{u} + \vec{v} = \overrightarrow{AB} + \overrightarrow{AC} = \overrightarrow{AD}$ (règle du parallélogramme).

EXEMPLE 39 –
Dans un tétraèdre régulier $ABCD$ de côté 1 les 4 faces sont des triangles équilatéraux dont on sait que les hauteurs mesurent $\frac{\sqrt{3}}{2}$.
On a, en particulier, $(\overrightarrow{DA}, \overrightarrow{DB}) = \frac{\pi}{3}$ rad et $DI = CI = \frac{\sqrt{3}}{2}$.
$\overrightarrow{DA} \cdot \overrightarrow{DB} = DA \times DB \times \cos(\overrightarrow{DA}, \overrightarrow{DB}) = \cos(\overrightarrow{DA}, \overrightarrow{DB}) = \cos(\frac{\pi}{3}) = \frac{1}{2}$



Montrons que les côtés opposés sont orthogonaux, soit que $\overrightarrow{AB} \perp \overrightarrow{CD}$:
$\overrightarrow{AB} \cdot \overrightarrow{CD} = (\overrightarrow{AC} + \overrightarrow{CB}) \cdot \overrightarrow{CD} = \overrightarrow{AC} \cdot \overrightarrow{CD} + \overrightarrow{CB} \cdot \overrightarrow{CD} = -\overrightarrow{CA} \cdot \overrightarrow{CD} + \overrightarrow{CB} \cdot \overrightarrow{CD} = -\frac{1}{2} + \frac{1}{2} = 0$
Calculons l'angle dièdre $(\overrightarrow{ID}, \overrightarrow{IC})$ entre les faces :
$\overrightarrow{ID} \cdot \overrightarrow{IC} = \overrightarrow{DI} \cdot \overrightarrow{CI} = \frac{\overrightarrow{DA}+\overrightarrow{DB}}{2} \cdot \frac{\overrightarrow{CA}+\overrightarrow{CB}}{2} = \frac{\overrightarrow{DA} \cdot \overrightarrow{CA} + \overrightarrow{DB} \cdot \overrightarrow{CA} + \overrightarrow{DA} \cdot \overrightarrow{CB} + \overrightarrow{DB} \cdot \overrightarrow{CB}}{4} = \frac{\overrightarrow{AD} \cdot \overrightarrow{AC} + \overrightarrow{BD} \cdot \overrightarrow{BC}}{4} = \frac{\frac{1}{2}+\frac{1}{2}}{4} = \frac{1}{4}$
Or $\overrightarrow{ID} \cdot \overrightarrow{IC} = DI \times CI \cos(\overrightarrow{DI}, \overrightarrow{CI}) = (\frac{\sqrt{3}}{2})^2 \cos(\overrightarrow{DI}, \overrightarrow{CI})$
d'où $\cos(\overrightarrow{DI}, \overrightarrow{CI}) = \frac{3}{8}$ et $(\overrightarrow{DI}, \overrightarrow{CI}) = \cos^{-1}(\frac{3}{8}) \approx 1,1864$ rad, environ 68°.

3.c. Expression analytique du produit scalaire dans un repère orthonormé

PROPRIÉTÉ 4.18 Soient $(O, \vec{i}, \vec{j}, \vec{k})$ un repère orthonormé et les vecteurs $\vec{u}(x, y, z)$ et $\vec{v}(x', y', z')$
$$\vec{u} \cdot \vec{v} = xx' + yy' + zz'$$

Démonstration Décomposons $\vec{u}$ et $\vec{v}$ selon les vecteurs de la base $(\overrightarrow{i}, \overrightarrow{j}, \overrightarrow{k})$:
$\vec{u} = x\overrightarrow{i} + y\overrightarrow{j} + z\overrightarrow{k}$ et $\vec{v} = x'\overrightarrow{i} + y'\overrightarrow{j} + z'\overrightarrow{k}$.
Développons le produit scalaire $\vec{u} \cdot \vec{v}$:

$$\begin{aligned}\vec{u} \cdot \vec{v} &= (x\overrightarrow{i} + y\overrightarrow{j} + z\overrightarrow{k}) \cdot (x'\overrightarrow{i} + y'\overrightarrow{j} + z'\overrightarrow{k}) \\ &= x\overrightarrow{i} \cdot x'\overrightarrow{i} + x\overrightarrow{i} \cdot y'\overrightarrow{j} + x\overrightarrow{i} \cdot z'\overrightarrow{k} + y\overrightarrow{j} \cdot x'\overrightarrow{i} + y\overrightarrow{j} \cdot y'\overrightarrow{j} + y\overrightarrow{j} \cdot z'\overrightarrow{k} + z\overrightarrow{k} \cdot x'\overrightarrow{i} + z\overrightarrow{k} \cdot y'\overrightarrow{j} + z\overrightarrow{k} \cdot z'\overrightarrow{k} \\ &= xx'\overrightarrow{i}^2 + yy'\overrightarrow{j}^2 + zz'\overrightarrow{k}^2 + (xy' + yx')\overrightarrow{i} \cdot \overrightarrow{j} + (xz' + zx')\overrightarrow{i} \cdot \overrightarrow{k} + (yz' + zy')\overrightarrow{j} \cdot \overrightarrow{k}\end{aligned}$$

Le repère $(\overrightarrow{i}, \overrightarrow{j}, \overrightarrow{k})$ étant normé, d'après la propriété 4.8, on a $||\overrightarrow{i}|| = ||\overrightarrow{j}|| = ||\overrightarrow{k}|| = 1$.
(En effet $\overrightarrow{i}^2 = \overrightarrow{i} \cdot \overrightarrow{i} = ||\overrightarrow{i}||^2 \cos(0) = 1$ et de même, $\overrightarrow{j}^2 = \overrightarrow{k}^2 = 1$)
De plus, le repère étant orthogonal, d'après la propriété 4.8, on a $\overrightarrow{i} \perp \overrightarrow{j}$, $\overrightarrow{i} \perp \overrightarrow{k}$ et $\overrightarrow{j} \perp \overrightarrow{k}$.
(En effet $\overrightarrow{i} \cdot \overrightarrow{j} = ||\overrightarrow{i}|| \times ||\overrightarrow{j}|| \cos(\frac{\pi}{2}) = 0$ et de même $\overrightarrow{i} \cdot \overrightarrow{k} = \overrightarrow{j} \cdot \overrightarrow{k} = 0$)
L'expression de $\vec{u} \cdot \vec{v}$ se simplifie donc en $\vec{u} \cdot \vec{v} = xx' + yy' + zz'$.

Remarques :

- Dans une base orthonormée la norme d'un vecteur $\vec{u}(x, y, z)$ est donnée par cette propriété : $||\vec{u}||^2 = \vec{u} \cdot \vec{u} = x^2 + y^2 + z^2$, d'où $||\vec{u}|| = \sqrt{x^2 + y^2 + z^2}$ car $||\vec{u}|| \geqslant 0$.
  De même pour des points $A(x_A, y_A, z_A)$ et $B(x_B, y_B, z_B)$ dont les coordonnées se réfèrent à un repère orthonormé. On a $\overrightarrow{AB}(x_B - x_A, y_B - y_A, z_B - z_A)$ et alors
  $AB^2 = ||\overrightarrow{AB}||^2 = \overrightarrow{AB} \cdot \overrightarrow{AB} = (x_B - x_A)^2 + (y_B - y_A)^2 + (z_B - z_A)^2$, d'où
  $AB = \sqrt{(x_B - x_A)^2 + (y_B - y_A)^2 + (z_B - z_A)^2}$.
- Montrons maintenant, à l'aide de cette propriété, la linéarité du produit scalaire exprimée dans la propriété 4.17. Pour cela, on se place dans une base orthonormée $(\overrightarrow{i}, \overrightarrow{j}, \overrightarrow{k})$ et on considère les vecteurs $\overrightarrow{u}(x, y, z)$, $\overrightarrow{v}(x', y', z')$ et $\overrightarrow{w}(x'', y'', z'')$ ainsi que le réel k :
  On sait que $k\vec{u}$ a pour composantes (kx, ky, kz) donc
  $(k\vec{u}) \cdot \vec{v} = kxx' + kyy' + kzz' = k(xx' + yy' + zz') = k(\vec{u} \cdot \vec{v})$
  De plus $\vec{u} + \vec{v}$ ayant pour composantes $(x + x', y + y', z + z')$ on a
  $(\vec{u} + \vec{v}) \cdot \vec{w} = x''(x + x') + y''(y + y') + z''(z + z') = xx'' + x'x'' + yy'' + y'y'' + zz'' + z'z'' = (xx'' + yy'' + zz'') + (x'x'' + y'y'' + z'z'') = \vec{u} \cdot \vec{w} + \vec{v} \cdot \vec{w}$

Définition 4.14 (Droite orthogonale à un plan) Une droite $\mathcal{D}$ est orthogonale à un plan $\mathcal{P}$ si et seulement si elle est orthogonale à toute droite $d \in \mathcal{P}$.

Propriété 4.19 Une droite $\mathcal{D}$ est orthogonale à un plan $\mathcal{P}$ si et seulement si elle est orthogonale à deux droites sécantes de $\mathcal{P}$.
Autrement dit, si $\overrightarrow{n}$ dirige $\mathcal{D}$ et si $(\overrightarrow{v}, \overrightarrow{w})$ dirige $\mathcal{P}$, on a $\mathcal{D} \perp \mathcal{P} \iff \overrightarrow{n} \cdot \overrightarrow{v} = 0$ et $\overrightarrow{n} \cdot \overrightarrow{w} = 0$.

Démonstration Dans le sens direct, c'est immédiat : si $d \perp \mathcal{P}$ alors d est orthogonale à toute droite de $\mathcal{P}$ et en particulier les droites d_1 et d_2 dirigées respectivement par $\overrightarrow{v}$ et par $\overrightarrow{w}$ or ces droites sont sécantes puisqu'elles sont dirigées par des vecteurs non-colinéaires.

Réciproquement, supposons $\overrightarrow{n} \cdot \overrightarrow{v} = 0$ et $\overrightarrow{n} \cdot \overrightarrow{w} = 0$. On a donc $\overrightarrow{n} \perp \overrightarrow{v}$ et $\overrightarrow{n} \perp \overrightarrow{w}$.
D'après la propriété 4.4, trois vecteurs sont coplanaires si et seulement si l'un d'entre eux est une combinaison linéaire des deux autres. Toute droite d de $\mathcal{P}$ est donc dirigée par un vecteur directeur $\overrightarrow{u}$ qui est une combinaison linéaire des vecteurs non-colinéaires $\overrightarrow{v}$ et $\overrightarrow{w}$, il existe des réels x et y tels que $\overrightarrow{u} = x\overrightarrow{v} + y\overrightarrow{w}$.
Le produit scalaire $\overrightarrow{n} \cdot \overrightarrow{u} = \overrightarrow{n} \cdot (x\overrightarrow{v} + y\overrightarrow{w}) = x\overrightarrow{n} \cdot \overrightarrow{v} + y\overrightarrow{n} \cdot \overrightarrow{w} = 0$; on en déduit que $d \perp \mathcal{P}$.

DÉFINITION 4.15 (VECTEUR NORMAL À UN PLAN) Un vecteur est dit *normal* à un plan $\mathcal{P}$ si et seulement si il dirige une droite orthogonale au plan $\mathcal{P}$.

PROPRIÉTÉ 4.20 (CARACTÉRISATION D'UN PLAN) Soient $\vec{n} \neq \vec{0}$ un vecteur et A un point.
Le plan passant par A de vecteur normal $\vec{n}$ contient tous les points M tels que $\overrightarrow{AM} \cdot \vec{n} = 0$.
Dans un repère orthonormé, un triplet de réels $(a, b, c) \neq (0, 0, 0)$ et un point $A(x_A, y_A, z_A)$ étant donnés, le plan passant par A de vecteur normal $\vec{n}(a, b, c)$ a pour équation cartésienne $ax + by + cz + d = 0$, où $d = -(ax_A + by_A + cz_A)$ et, réciproquement, tout plan d'équation $ax + by + cz + d = 0$ admet $\vec{n}(a, b, c)$ comme vecteur normal.

DÉMONSTRATION $\boxed{\text{Caractérisation vectorielle}}$
Soit $\mathcal{P}$ le plan passant par A de vecteur normal $\vec{n}$ et soit $M \in \mathcal{P}$.

- Si $M = A$, on a $\overrightarrow{AM} \cdot \vec{n} = 0$ car $\overrightarrow{AM} = \vec{0}$.
- Si $M \neq A$, la droite (AM) existe et, comme elle est incluse dans $\mathcal{P}$, son vecteur directeur $\overrightarrow{AM}$ est orthogonal à $\vec{n}$ donc $\overrightarrow{AM} \cdot \vec{n} = 0$

Réciproquement, montrons par l'absurde que $\overrightarrow{AM} \cdot \vec{n} = 0 \Longrightarrow M \in \mathcal{P}$:
Supposons que $M \notin \mathcal{P}$ et appelons M' le projeté orthogonal M sur $\mathcal{P}$.
On a $M' \neq M$, $(MM') \perp (AM')$ et $\overrightarrow{M'M}$ colinéaire à $\vec{n}$.
$\overrightarrow{AM} \cdot \vec{n} = (\overrightarrow{AM'} + \overrightarrow{M'M}) \cdot \vec{n} = \overrightarrow{AM'} \cdot \vec{n} + \overrightarrow{M'M} \cdot \vec{n} = 0 + \overrightarrow{M'M} \cdot \vec{n} = \pm||\overrightarrow{M'M}|| \times ||\vec{n}||$.
Mais comme $\overrightarrow{AM} \cdot \vec{n} = 0$ et $||\vec{n}|| \neq 0$, on en déduit $||\overrightarrow{M'M}|| = MM' = 0 \Longleftrightarrow M = M'$.
Par conséquent si $\overrightarrow{AM} \cdot \vec{n} = 0$ alors $M \in \mathcal{P}$.

$\boxed{\text{Caractérisation analytique, dans un repère orthonormé}}$
Soit $\mathcal{P}$ le plan de vecteur normal $\vec{n}(a, b, c) \neq \vec{0}$ passant par $A(x_A, y_A, z_A)$.
$M(x, y, z) \in \mathcal{P} \Longleftrightarrow \overrightarrow{AM} \perp \vec{n} \Longleftrightarrow \overrightarrow{AM} \cdot \vec{n} = 0$. Et, comme $\overrightarrow{AM}(x - x_A, y - y_A, z - z_A)$,
$M \in \mathcal{P} \Longleftrightarrow a(x - x_A) + b(y - y_A) + c(z - z_A) = 0 \Longleftrightarrow ax + by + cz - (ax_A + by_A + cz_A) = 0$,
et, en notant $d = -(ax_A + by_A + cz_A)$, on obtient l'égalité $ax + by + cz + d = 0$.

Réciproquement : Le vecteur $\vec{n}(a, b, c)$ n'étant pas nul, on peut supposer qu'une de ses composantes est non nulle, mettons $c \neq 0$. Le point $C(0, 0, \frac{-d}{c}) \in \mathcal{P}$ car $a \times 0 + b \times 0 + c \times \frac{-d}{c} + d = -d + d = 0$
Choisissons un point $A(x_A, y_A, z_A) \neq C$ de $\mathcal{P}$: $ax_A + by_A + cz_A + d = 0 \Longrightarrow d = -ax_A - by_A - cz_A$
On a $\overrightarrow{AC}(-x_A, -y_A, \frac{-d}{c} - z_A)$ et donc $\overrightarrow{AM} \cdot \vec{n} = \underbrace{-ax_A - by_A - cz_A}_{=d} - \frac{cd}{c} = d - d = 0$

Finalement $\vec{n}$ est bien un vecteur normal au plan $\mathcal{P}$.

Remarques :

- Dans le repère orthonormé $(O, \vec{i}, \vec{j}, \vec{k})$ quelle est l'équation du plan $\mathcal{P}$ passant par O et de vecteur normal $\vec{k}(0, 0, 1)$? C'est $cz + d = 0$ avec $O(0, 0, 0) \in \mathcal{P}$ on obtient $d = 0$.
  Une équation cartésienne de $\mathcal{P}$ est donc $z = 0$ (j'ai pris $c = 1$).
  Avec la notation $\vec{i} = \overrightarrow{OI}$, $\vec{j} = \overrightarrow{OJ}$, $\vec{k} = \overrightarrow{OK}$ c'est le plan (IOJ).
  De même, le plan (IOK) a pour équation $y = 0$ et (JOK) a pour équation $x = 0$.
- Soient $\vec{u}(a, b, c)$ et $\vec{v}(a', b', c')$ les vecteurs directeurs d'un plan $\mathcal{P}$. Déterminons le couple de réels (x, y) pour que le vecteur $\vec{n}(x, y, z = 1)$ soit normal à $\mathcal{P}$. Le système qui se déduit des conditions $\vec{u} \cdot \vec{n} = 0$ et $\vec{v} \cdot \vec{n} = 0$ conduit aux valeurs $x = \frac{cb' - bc'}{ba' - ab'}$ et $y = \frac{ac' - ca'}{ba' - ab'}$ (le vérifier).
  En reprenant les valeurs de la propriété 4.14, soient $\vec{u}(3, 2, 1)$ et $\vec{v}(-1, -2, -3)$ les vecteurs directeurs du plan $\mathcal{P}$ passant par $M_0(1, 2, 3)$. Le vecteur $\vec{n}(\frac{-2+6}{-2+6}, \frac{-9+1}{-2+6}, 1) = (1, -2, 1)$ est normal à $\mathcal{P}$ donc l'équation qui s'en déduit est $x - 2y + z + d = 0$ et avec $M_0 \in \mathcal{P}$ on obtient $1 - 4 + 3 + d = 0 \Longleftrightarrow d = 0$, d'où l'équation cartésienne de $\mathcal{P}$: $x - 2y + z = 0$.

3.d. Applications du produit scalaire

PROPRIÉTÉ 4.21 (POSITIONS RELATIVES) Soient $\mathcal{P}$ et $\mathcal{P}$' deux plans de vecteur normal respectif $\vec{n}$ et $\vec{n'}$ et soit $\mathcal{D}$ une droite de vecteur directeur $\vec{u}$

- Si $\vec{n} \cdot \vec{u} = 0$ alors la droite $\mathcal{D}$ est parallèle au plan $\mathcal{P}$ ($\mathcal{D} \subset \mathcal{P}$ ou $\mathcal{D} \cap \mathcal{P} = \emptyset$).
  Sinon $\mathcal{D}$ perce $\mathcal{P}$ en un point.
- Si $\vec{n}$ et $\vec{n'}$ sont colinéaires alors les plans $\mathcal{P}$ et $\mathcal{P}$' sont parallèles ($\mathcal{P} = \mathcal{P}$' ou $\mathcal{P} \cap \mathcal{P}' = \emptyset$).
  Sinon $\mathcal{P}$ coupe $\mathcal{P}$' selon une droite.
- Si $\vec{n} \cdot \vec{n'} = 0$ alors les plans $\mathcal{P}$ et $\mathcal{P}$' sont perpendiculaires ($\mathcal{P} = \mathcal{P}$' ou $\mathcal{P} \cap \mathcal{P}' = \emptyset$)

DÉMONSTRATION Ces propriétés sont immédiates et ne font que reformuler les propriétés 4.5 et 4.6.

DÉFINITION 4.16 (PROJETÉ ORTHOGONAL) Soient $\mathcal{P}$ un plan de vecteur normal $\vec{n}$ et M un point. Le *projeté orthogonal* de M sur $\mathcal{P}$ est un point M' tel que $M' \in \mathcal{P}$ et $\overrightarrow{MM'}$ colinéaire à $\overrightarrow{n}$.

Soient $\mathcal{D}$ une droite de vecteur directeur $\vec{u}$ et M un point.
Le *projeté orthogonal* de M sur $\mathcal{D}$ est un point M'' tel que $M'' \in \mathcal{D}$ et $\overrightarrow{MM''} \cdot \overrightarrow{u} = 0$.

PROPRIÉTÉ 4.22 (DISTANCE D'UN POINT À UN PLAN) Soient $\mathcal{P}$ un plan passant par le point N de vecteur normal $\vec{n}$ et M un point. La *distance* du point M au plan $\mathcal{P}$ est $\dfrac{|\overrightarrow{MN} \cdot \overrightarrow{n}|}{||\overrightarrow{n}||}$
Dans un repère orthonormé, la distance entre un point $M(x, y, z)$ et le plan d'équation cartésienne $ax + by + cz + d = 0$ est égale à $\dfrac{|ax + by + cz + d|}{\sqrt{a^2 + b^2 + c^2}}$

DÉMONSTRATION Montrons tout d'abord que le projeté orthogonal M' de M sur $\mathcal{P}$ est le point le plus proche de M sur $\mathcal{P}$. Soit N un point de $\mathcal{P}$ différent de M'.
Dans le plan $(MM'N)$, comme (MM') est perpendiculaire à $\mathcal{P}$, le triangle $MM'N$ est rectangle en M' et par application du théorème de Pythagore $MN^2 = MM'^2 + M'N^2$.
Comme $M'N > 0$ (car $M' \neq N$), on a $MN^2 > MM'^2 \iff MN > MM'$.
Calculons $\overrightarrow{MN} \cdot \overrightarrow{n} = (\overrightarrow{MM'} + \overrightarrow{M'N}) \cdot \overrightarrow{n} = \overrightarrow{MM'} \cdot \overrightarrow{n} + \overrightarrow{M'N} \cdot \overrightarrow{n} = \overrightarrow{MM'} \cdot \overrightarrow{n}$ car $\overrightarrow{M'N} \perp \overrightarrow{n}$.
On a donc $\overrightarrow{MN} \cdot \overrightarrow{n} = \overrightarrow{MM'} \cdot \overrightarrow{n} = \pm||\overrightarrow{MM'}|| \times ||\overrightarrow{n}||$ car $\overrightarrow{M'N}$ et $\overrightarrow{n}$ sont colinéaires.
On en déduit $MM' = ||\overrightarrow{MM'}|| = \pm\frac{\overrightarrow{MN} \cdot \overrightarrow{n}}{||\overrightarrow{n}||}$ et comme une distance est positive $MM' = \frac{|\overrightarrow{MN} \cdot \overrightarrow{n}|}{||\overrightarrow{n}||}$.
Prenons un point $N(x_N, y_N, z_N) \in \mathcal{P}$. On a $ax_N + by_N + cz_N + d = 0 \Longrightarrow d = -(ax_N + by_N + cz_N)$
$\overrightarrow{MN}(x_N - x, y_N - y, z_N - z)$ et donc
$\overrightarrow{MN} \cdot \vec{n} = a(x_N - x) + b(y_N - y) + c(z_N - z) = \underbrace{ax_N + by_N + cz_N}_{=-d} - (ax + by + cz) = -(ax + by + cz + d)$
Donc $|\overrightarrow{MN} \cdot \vec{n}| = |ax + by + cz + d|$ et comme, de plus, $||\overrightarrow{n}|| = \sqrt{a^2 + b^2 + c^2}$, le résultat s'en déduit.

PROPRIÉTÉ 4.23 (DISTANCE D'UN POINT À UNE DROITE) Soient M et N deux points, $\vec{u} \neq \vec{0}$ un vecteur. La *distance* de M à la droite et $\mathcal{D}(N, \vec{u})$ est $\sqrt{MN^2 - \frac{(\overrightarrow{MN} \cdot \overrightarrow{u})^2}{||\overrightarrow{u}||^2}}$ soit $MN \times |\sin(\overrightarrow{MN}, \overrightarrow{u})|$.

DÉMONSTRATION Dans un repère orthonormé, projetons M sur $\mathcal{D}$ en faisant intervenir le plan $\mathcal{P}$ de vecteur normal $\overrightarrow{u}(a,b,c)$ passant par M.
Le projeté orthogonal de M sur $\mathcal{D}$ est un point M'' tel que $M'' \in \mathcal{D}$ et $\overrightarrow{MM''} \cdot \overrightarrow{u} = 0$.
Les coordonnées de $M(x,y,z)$ vérifient l'équation de $\mathcal{P}$ donc
$ax + by + cz + d = 0 \Longrightarrow d = -(ax + by + cz)$
De même pour $M''(x'',y'',z'')$ donc
$ax'' + by'' + cz'' - (ax + by + cz) = 0 \Longleftrightarrow a(x''-x) + b(y''-y) + c(z''-z) = 0$, cette dernière expression étant équivalente à $\overrightarrow{MM''} \cdot \overrightarrow{u} = 0$.
Comme $M'' \in \mathcal{D}$, on doit avoir $\overrightarrow{NM''}$ colinéaire à $\overrightarrow{u}$ donc $(x''-x_N, y''-y_N, z''-z_N)$ proportionnel à (a,b,c), soit $\begin{cases} x''-x_N = at \\ y''-y_N = bt \\ z''-z_N = ct \end{cases} \Longleftrightarrow \begin{cases} x'' = x_N + at \\ y'' = y_N + bt \\ z'' = z_N + ct \end{cases}$
En reportant ces expressions dans le produit scalaire $\overrightarrow{MM''} \cdot \overrightarrow{u} = 0$ on obtient
$a(x_N + at - x) + b(y_N + bt - y) + c(z_N + ct - z) = 0 \Longleftrightarrow (a^2+b^2+c^2)t = -a(x_N - x) - b(y_N - y) - c(z_N - z) \Longleftrightarrow t = \frac{-a(x_N-x)-b(y_N-y)-c(z_N-z)}{a^2+b^2+c^2}$,
cette dernière expression étant équivalente à $t = \frac{-\overrightarrow{MN}\cdot\overrightarrow{u}}{||\overrightarrow{u}||^2}$.

En reportant cette valeur dans les coordonnées de M'' on obtient $\begin{cases} x'' = x_N - a\frac{\overrightarrow{MN}\cdot\overrightarrow{u}}{||\overrightarrow{u}||^2} \\ y'' = y_N - b\frac{\overrightarrow{MN}\cdot\overrightarrow{u}}{||\overrightarrow{u}||^2} \\ z'' = z_N - c\frac{\overrightarrow{MN}\cdot\overrightarrow{u}}{||\overrightarrow{u}||^2} \end{cases}$

$$\begin{aligned} MM''^2 &= (x_N - x - a\frac{\overrightarrow{MN}\cdot\overrightarrow{u}}{||\overrightarrow{u}||^2})^2 + (y_N - y - b\frac{\overrightarrow{MN}\cdot\overrightarrow{u}}{||\overrightarrow{u}||^2})^2 + (z_N - z - c\frac{\overrightarrow{MN}\cdot\overrightarrow{u}}{||\overrightarrow{u}||^2})^2 \\ &= (x_N-x)^2 + (y_N-y)^2 + (z_N-z)^2 + (a^2+b^2+c^2)\left(\frac{\overrightarrow{MN}\cdot\overrightarrow{u}}{||\overrightarrow{u}||^2}\right)^2 \\ &\quad - 2\left(a(x_N-x) + b(y_N-y) + c(z_N-z)\right) \times \frac{\overrightarrow{MN}\cdot\overrightarrow{u}}{||\overrightarrow{u}||^2} \\ &= MN^2 - \frac{(\overrightarrow{MN}\cdot\overrightarrow{u})^2}{||\overrightarrow{u}||^2} \end{aligned}$$

On en déduit $MM''^2 = \frac{||\overrightarrow{u}||^2 \times MN^2 - (\overrightarrow{MN}\cdot\overrightarrow{u})^2}{||\overrightarrow{u}||^2} = \frac{||\overrightarrow{u}||^2 \times MN^2 - ||\overrightarrow{MN}||^2 \times ||\overrightarrow{u}||^2 \times \cos^2(\overrightarrow{MN},\overrightarrow{u})}{||\overrightarrow{u}||^2}$ d'où
$MM''^2 = MN^2\left(1 - \cos^2(\overrightarrow{MN},\overrightarrow{u})\right) = MN^2 \times \sin^2(\overrightarrow{MN},\overrightarrow{u})$ ce qui justifie la dernière expression.

Remarques :

- Cette démonstration est bien compliquée pour un résultat si simple qui parait presque évident lorsqu'on se reporte dans le plan $(MM'N)$ où, le triangle $MM'N$ étant rectangle en M', on établit immédiatement la dernière relation. Disons qu'en la faisant, on a obtenu quelques résultats intéressants comme les coordonnées du projeté orthogonal M'' de M sur $\mathcal{D}$ ainsi qu'une expression de la distance que l'on peut utiliser, programmer, dès lors que l'on connait les coordonnées de M, N et les composantes de $\overrightarrow{u}$ dans un repère orthonormé.
- Supposons que $\overrightarrow{u}(-2,1,5)$ dirige $\mathcal{D}$, une droite passant par $N(1,3,0)$.
  Quelle est la distance d de $M(2,0,7)$ jusqu'à $\mathcal{D}$? Facile : $d^2 = MN^2 - \frac{(\overrightarrow{MN}\cdot\overrightarrow{u})^2}{||\overrightarrow{u}||^2}$
  $d^2 = (1-2)^2 + (3-0)^2 + (0-7)^2 - \frac{(-2(1-2)+1(3-0)+5(0-7))^2}{(-2)^2+1^2+5^2} = 1+9+49-\frac{(2+3-35)^2}{4+1+25} = 59-30 = 29$
  Par conséquent $d = \sqrt{29} \approx 5,385$.

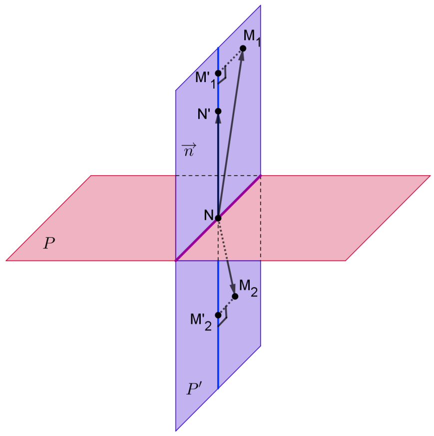
Propriété 4.24 (Demi-espace) Soient $\mathcal{P}$ un plan passant par le point N de vecteur normal $\vec{n}$ et N' le point tel que $\overrightarrow{NN'} = \vec{n}$. Un point M quelconque de l'espace est :

- Dans le plan $\mathcal{P}$ si $\overrightarrow{NM} \cdot \vec{n} = 0$
- Dans le demi-espace limité par $\mathcal{P}$ contenant N' si $\overrightarrow{NM} \cdot \vec{n} > 0$
- Dans le demi-espace limité par $\mathcal{P}$ ne contenant pas N' si $\overrightarrow{NM} \cdot \vec{n} < 0$

Dans un repère orthonormé, le plan $\mathcal{P}$ ayant pour équation $ax + by + cz + d = 0$, les inéquations des deux demi-espaces sont $ax + by + cz + d > 0$ et $ax + by + cz + d < 0$.

Démonstration Si M' est le projeté orthogonal de M sur la droite $\mathcal{D}(N, \vec{n})$ alors $\overrightarrow{NM} \cdot \vec{n} = \overrightarrow{NM'} \cdot \overrightarrow{NN'}$.

- Si $\overrightarrow{NM} \cdot \vec{n} = 0$ alors $M \in \mathcal{P}$
- Si $\overrightarrow{NM} \cdot \vec{n} > 0$ les vecteurs $\overrightarrow{NM'}$ et $\overrightarrow{NN'}$ sont donc colinéaires et de même sens. Les points N', M' et M sont dans la même partie de l'espace limitée par $\mathcal{P}$
- Si $\overrightarrow{NM} \cdot \vec{n} < 0$ les vecteurs sont colinéaires et de sens contraires. Les points M' et M sont dans le demi-espace limité par $\mathcal{P}$ qui ne contient pas N'

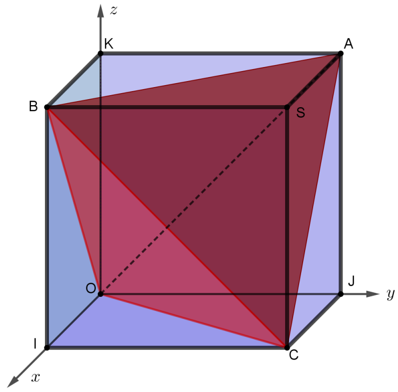


Sur l'illustration, les points M_1 et M_2 ne sont pas dans les mêmes demi-espaces limités par le plan $\mathcal{P}$.

Exemple 40 – Le cube dessiné ci-dessous est construit à partir du repère orthonormé (O, I, J, K). Un point $M(x, y, z)$ de ce cube vérifie les trois encadrements $0 \leqslant x \leqslant 1$, $0 \leqslant y \leqslant 1$ et $0 \leqslant z \leqslant 1$. Où doit se trouver le point M pour que ses coordonnées soient les cotés d'un triangle ?

Pour construire un triangle, chaque côté doit être inférieur ou égal à la somme des deux autres. Les coordonnées d'un point-solution doivent donc vérifier, en plus des trois encadrements, les trois inégalités $x \leqslant y + z$, $y \leqslant x + z$ et $z \leqslant x + y$.

L'inégalité $x \leqslant y + z$ correspond à un demi-espace limité par le plan $\mathcal{P}_1$ d'équation $x - y - z = 0$. Parmi les sommets du cube $O(0, 0, 0)$, $C(1, 1, 0)$ et $B(1, 0, 1)$ vérifient l'équation de $\mathcal{P}_1$ donc $\mathcal{P}_1 = (OBC)$. Le point $I(1, 0, 0)$ ne vérifie pas l'inéquation $x \leqslant y + z$ du demi-espace solution donc on élimine du cube le tétraèdre $OBCI$.



De même, l'inégalité $y \leqslant x + z$ conduit à éliminer le tétraèdre $OCAJ$ car $O(0, 0, 0)$, $C(1, 1, 0)$ et $A(0, 1, 1)$ vérifient l'équation $x - y + z = 0$ de $\mathcal{P}_2$ tandis que $J(0, 1, 0)$ ne vérifie pas l'inégalité $y \leqslant x + z$. L'inégalité $z \leqslant x + y$, quant à elle, conduit à éliminer le tétraèdre $OBAK$.

La suppression de ces trois tétraèdres fait émerger un solide-solution constitué du tétraèdre régulier $OABC$ auquel s'accole le tétraèdre rectangle $SABC$. Ce solide est un hexaèdre (6 faces triangulaires).

Quelle est l'équation du plan $\mathcal{P}_4 = (ABC)$?
$\overrightarrow{OS}(1, 1, 1)$ est un vecteur normal à ce plan $\mathcal{P}_4$, l'équation cherchée est de la forme $x + y + z + d = 0$. Comme $A(0, 1, 1) \in \mathcal{P}_4$, on obtient $1 + 1 + d = 0 \iff d = -2$. L'équation de $\mathcal{P}_4$ est $x + y + z - 2 = 0$. Le point $S(1, 1, 1)$ vérifiant l'inégalité $x + y + z - 2 \geqslant 0$, cette inéquation caractérise le demi-espace contenant S limité par (ABC).
Un point du tétraèdre $SABC$ vérifie, en plus des trois encadrements, cette inéquation.
Un point du tétraèdre régulier $OABC$ vérifie, en plus des trois encadrements ($0 \leqslant x \leqslant 1$, $0 \leqslant y \leqslant 1$, $0 \leqslant z \leqslant 1$) et des trois inégalités ($x \leqslant y + z$, $y \leqslant x + z$, $z \leqslant x + y$), l'inéquation $x + y + z - 2 \leqslant 0$.
Comme les quatre plans $\mathcal{P}_1$, $\mathcal{P}_2$, $\mathcal{P}_3$, $\mathcal{P}_4$ sont sécants deux à deux, leurs intersections constituant les arêtes du tétraèdre $OABC$, on comprend que les quatre inéquations des demi-espaces concernés ($x \leqslant y + z$, $y \leqslant x + z$, $z \leqslant x + y$, $x + y + z \leqslant 2$) suffisent à caractériser les points de ce tétraèdre.

DÉFINITION 4.17 (SPHÈRE) La *sphère* $\mathcal{S}(\Omega, R)$ de centre Ω et de rayon $R > 0$ est l'ensemble des points M tels que $\Omega M = R$.
La *boule* $\mathcal{B}(\Omega, R)$ de centre Ω et de rayon $R > 0$ est l'ensemble des points M tels que $\Omega M \leqslant R$.

PROPRIÉTÉ 4.25 Soient A et B deux points et Ω leur milieu.
La sphère $\mathcal{S}(\Omega, \frac{AB}{2})$ est l'ensemble des points M tels que $\overrightarrow{MA} \cdot \overrightarrow{MB} = 0$.

Dans un repère orthonormé, avec $\Omega(x_\Omega, y_\Omega, z_\Omega)$ et $M(x, y, z)$,
l'équation cartésienne de $\mathcal{S}(\Omega, R)$ est $(x - x_\Omega)^2 + (y - y_\Omega)^2 + (z - z_\Omega)^2 = R^2$.
Un point $M(x, y, z)$ vérifiant l'équation $x^2 + y^2 + z^2 + ax + by + cz + d = 0$ avec $d \leqslant \frac{a^2+b^2+c^2}{4}$ appartient à la sphère de centre $\Omega(\frac{-a}{2}, \frac{-b}{2}, \frac{-b}{2})$ et de rayon $R = \sqrt{\frac{a^2+b^2+c^2}{4} - d}$.

DÉMONSTRATION ➲ On calcule le produit scalaire $\overrightarrow{MA} \cdot \overrightarrow{MB}$ en utilisant le milieu Ω de $[AB]$:
$\overrightarrow{MA} \cdot \overrightarrow{MB} = (\overrightarrow{M\Omega} + \cdot\overrightarrow{\Omega A}) \cdot (\overrightarrow{M\Omega} + \cdot\overrightarrow{\Omega B}) = ||\overrightarrow{M\Omega}||^2 + \overrightarrow{M\Omega} + \overrightarrow{M\Omega} \cdot \overrightarrow{\Omega B} + \overrightarrow{\Omega A} \cdot \overrightarrow{M\Omega} + \overrightarrow{\Omega A} \cdot \overrightarrow{\Omega B}$
Mais comme Ω est le milieu de $[AB]$, on a $\overrightarrow{\Omega A} \cdot \overrightarrow{\Omega B} = -||\overrightarrow{\Omega A}||^2$ (vecteurs opposés) et
$\overrightarrow{M\Omega} \cdot \overrightarrow{\Omega B} + \overrightarrow{\Omega A} \cdot \overrightarrow{M\Omega} = \overrightarrow{M\Omega} \cdot (\overrightarrow{\Omega B} + \overrightarrow{\Omega A}) = \overrightarrow{M\Omega} \cdot \overrightarrow{0} = 0$
On en déduit $\overrightarrow{MA} \cdot \overrightarrow{MB} = ||\overrightarrow{M\Omega}||^2 - ||\overrightarrow{\Omega A}||^2 = M\Omega^2 - \Omega A^2 = M\Omega^2 - \left(\frac{AB}{2}\right)^2$.
Par conséquent $\overrightarrow{MA} \cdot \overrightarrow{MB} = 0 \iff M\Omega^2 = \left(\frac{AB}{2}\right)^2 \implies M\Omega = \frac{AB}{2}$.
➲Transformons l'équation donnée dans un repère orthonormé :

$$\begin{aligned} x^2 + y^2 + z^2 + ax + by + cz + d &= \left(x + \frac{a}{2}\right)^2 - \frac{a^2}{4} + \left(y + \frac{b}{2}\right)^2 - \frac{b^2}{4} + \left(z + \frac{c}{2}\right)^2 - \frac{c^2}{4} + d \\ &= \left(x + \frac{a}{2}\right)^2 + \left(y + \frac{b}{2}\right)^2 + \left(z + \frac{c}{2}\right)^2 + d - \frac{a^2 + b^2 + c^2}{4} \end{aligned}$$

La discussion porte alors sur $d' = d - \frac{a^2+b^2+c^2}{4}$

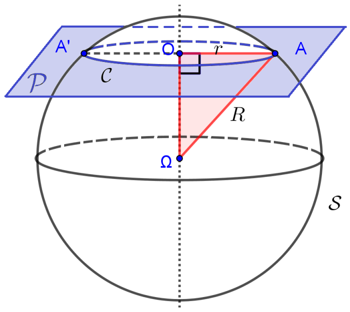
- si $d' > 0$ alors $x^2 + y^2 + z^2 + ax + by + cz + d = 0 \iff (x + \frac{a}{2})^2 + \left(y + \frac{b}{2}\right)^2 + (z + \frac{c}{2})^2 = 0$ alors M appartient à une sphère de rayon $\sqrt{d'}$ et de centre $\Omega(\frac{-a}{2}, \frac{-b}{2}, \frac{-b}{2})$.
- si $d' = 0$ alors cette sphère est réduit au point $\Omega(\frac{-a}{2}, \frac{-b}{2}, \frac{-b}{2})$ (le rayon est nul).
- si $d' < 0$ alors aucun point ne vérifie l'équation (une somme de carrés n'est jamais négative).

EXEMPLE 41 – Dans un repère orthonormé, on considère la sphère $\mathcal{S}(\Omega, R)$ d'équation $(x - 1)^2 + (y - 2)^2 + (z - 3)^2 = 4$ et le plan $\mathcal{P}$ d'équation $x + 2y + 3z = a$ et on se demande pour quelles valeurs de a le plan coupe la sphère et dans ce cas, quelles sont les coordonnées du centre du cercle-solution et quel en est le rayon.
Le centre $\Omega(1, 2, 3)$ de la sphère se projette orthogonalement sur le plan en un point $O(x_O, y_O, z_O)$ tel que $\overrightarrow{\Omega O}$ soit colinéaire au vecteur $\overrightarrow{u}(1, 2, 3)$ normal à $\mathcal{P}$.
On en déduit que $x_O + 2y_O + 3z_O = a$ et $\exists t \in \mathbb{R}, \overrightarrow{\Omega O}(x_O - 1, y_O - 2, z_O - 3) = t\overrightarrow{u}(1, 2, 3)$
Pour déterminer les coordonnées de O, il faut donc résoudre ce système :

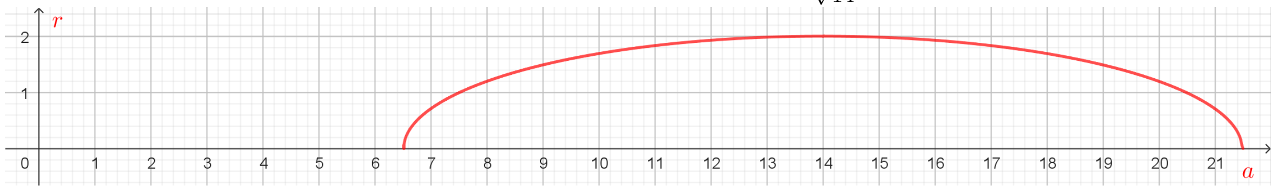
$$\begin{cases} x_O + 2y_O + 3z_O = a \\ x_O - 1 = t \\ y_O - 2 = 2t \\ z_O - 3 = 3t \end{cases} \iff \begin{cases} x_O + 2y_O + 3z_O = a \\ x_O - 1 = t \\ y_O - 2 = 2(x_O - 1) \\ z_O - 3 = 3(x_O - 1) \end{cases} \iff \begin{cases} x_O + 4x_O + 9x_O = a \iff 14x_O = a \\ x_O - 1 = t \\ y_O = 2(1 + x_O - 1) = 2x_O \\ z_O = 3x_O \end{cases}$$

On en déduit $x_O = \dfrac{a}{14}$, $y_O = 2x_O = \dfrac{2a}{14}$ et $z_O = 3x_O = \dfrac{3a}{14}$.

La sphère est coupée par le plan si la distance ΩO est inférieure ou égale au rayon R de la sphère, soit si $\Omega O^2 \leqslant R^2$.
Si on effectue une coupe de la sphère par un plan $\mathcal{P}'$ perpendiculaire à $\mathcal{P}$ et passant par Ω, $\overrightarrow{\Omega O}$ est un des deux vecteurs directeurs de $\mathcal{P}'$ et, si $A \in \mathcal{C} \cap \mathcal{P}'$, $\overrightarrow{OA}$ en est un autre. Ces deux vecteurs sont orthogonaux et, dans le triangle ΩOA, inclus dans $\mathcal{P}'$ et rectangle en O, on peut appliquer le théorème de Pythagore :
$\Omega A^2 = \Omega O^2 + OA^2 \iff R^2 = \Omega O^2 + r^2 \implies r = \sqrt{R^2 - \Omega O^2}$



Or $\Omega O^2 = (\frac{a}{14} - 1)^2 + (\frac{2a}{14} - 2)^2 + (\frac{3a}{14} - 3)^2 = \frac{(a-14)^2+4(a-14)^2+9(a-14)^2}{14^2} = \frac{(a-14)^2(1+4+9)}{14^2} = \frac{(a-14)^2}{14}$.
Le plan coupe donc la sphère si $\frac{(a-14)^2}{14} \leqslant 4 \iff (a-14)^2 \leqslant 56 \iff 14 - 2\sqrt{14} \leqslant a \leqslant 14 + 2\sqrt{14}$
Le rayon r du cercle d'intersection est $r = \sqrt{R^2 - \Omega O^2}$ soit ici $r = \sqrt{4 - \frac{(a-14)^2}{14}} = \frac{\sqrt{56-(a-14)^2}}{\sqrt{14}}$.
Je trace pour vérification la courbe de la fonction $r : a \longmapsto \frac{\sqrt{56-(a-14)^2}}{\sqrt{14}}$



Lorsque $a = 14$, $\mathcal{P}$ passe par Ω et coupe $\mathcal{S}$ selon un de ses grands cercles (on a alors $R = r$).
Comme $14 - 2\sqrt{14} \approx 6,52$ et $14 + 2\sqrt{14} \approx 21,48$, entre ces deux valeurs le plan $\mathcal{P}$ coupe la sphère selon un cercle dont le rayon est lu en ordonnée de cette courbe.

Il me reste encore une interrogation : quelle est l'équation du cercle intersection $\mathcal{C} = \mathcal{S} \cap \mathcal{P}$?
Bien sûr, on peut chercher les coordonnées des deux points A et A', diamétralement opposés sur $\mathcal{C}$ et traduire la caractérisation $M(x, y, z) \in \mathcal{C} \iff \overrightarrow{MA} \cdot \overrightarrow{MA'} = 0$ par l'expression algébrique recherchée.
Cette démarche est parfaitement légitime et conduit au bon résultat après de longs et rébarbatifs calculs impliquant le paramètre a ; mais surtout, cette démarche est parfaitement inutile puisqu'on sait déjà les coordonnées du centre et le rayon du cercle.
L'équation cartésienne de $\mathcal{C}$ est tout simplement :

$$\left(x - \frac{a}{14}\right)^2 + \left(y - \frac{2a}{14}\right)^2 + \left(z - \frac{3a}{14}\right)^2 = \frac{56 - (a-14)^2}{14}$$

Lorsque $a = 14$, cette équation se simplifie en $(x-1)^2 + (y-2)^2 + (z-3)^2 = 4$ (grand cercle)

Deuxième partie

Approfondissements

5 Nombres complexes

Plan du chapitre :

Aperçu historique :

C'est en Italie, au XVI^{e} *siècle, qu'apparurent les nombres complexes. Niccolò Tartaglia (1499-1557) innove en utilisant la racine carrée de* -1 *(pas encore nommée* i*) comme moyen systématique de résolution des équations du troisième degré. On ne sait pas bien s'il améliora la méthode secrète de Scipione del Ferro ou s'il inventa lui-même cette méthode, toujours est-il qu'il finit par la révéler à Jérôme Cardan (1501-1576) qui la publie en 1545 dans l'ouvrage* Ars Magna *(le Grand Art). Cardan publie également dans ce livre la formule découverte par Ludovico Ferrari, son élève, qui permet de résoudre les équations du quatrième degré en se ramenant à la résolution d'une équation du troisième degré. Quelques années plus tard, Rafaele Bombelli (1526-1572) prolonge le travail de ses compatriotes précurseurs par l'étude des propriétés de ces nombres complexes qu'il publie en 1572 dans* Algebra*. Malgré les avancées considérables de ces pionniers italiens, la communauté mathématique est loin de considérer cette cuisine algébrique comme un travail sur de vrais nombres. Ces nouveaux nombres sont nommés* imaginaires *en 1637 par René Descartes (1596-1650). Leibnitz, contemporain de Descartes, a dit de cette invention que* l'esprit divin s'est manifesté de façon sublime dans cette merveille de l'analyse, ce prodige d'un monde idéal, cet intermédiaire entre l'être et le non-être, que nous appelons la racine imaginaire de l'unité négative. *Leonhard Euler (1707-1783) utilise abondamment ces nombres qu'il définit comme* impossibles, *publie la formule* $e^{ix} = \cos x + i \sin x$ *et l'identité* $e^{i\pi} + 1 = 0$ *dans* Introductio in analysin infinitorum *en 1748, et introduit en 1777 la notation* i*.*

Au début du XIX^{e} *siècle, Carl Friedrich Gauss (1777-1855) construit précisément le corps, noté* $\mathbb{C}$*, des nombres complexes associés aux opérations d'addition et de multiplication. Il montre que ce corps est algébriquement clos, ce qui conduit à identifier le degré d'un polynôme complexe non nul au nombre de ses racines comptées avec leur ordre de multiplicité. Il appelle* complexes *ces nombres en 1831 dans* Theoria residuorum biquadraticorum *car ce sont des nombres composés d'une partie réelle et d'une partie imaginaire. Influencé par les travaux de Jean-Robert Argand (1768-1822) et Caspar Wessel (1745-1818), Gauss et surtout Augustin Louis Cauchy (1789-1857) développent l'aspect géométrique des complexes ; Cauchy introduit en 1847 les termes* affixe *et* argument*. Plus tard, William Rowan Hamilton (1805-1865) prolonge l'ensemble* $\mathbb{C}$ *en définissant les quaternions.*
L'acceptation de ces nouveaux nombres ressemble à une lente digestion qui s'est étalée sur trois siècles ; la lettre i est l'initiale d'imaginaire ou inventé, mais c'est aussi celle d'impossible, idiotie, inutile et irréelle. Outre les applications en algèbre, en analyse, et en géométrie, ces nombres trouvent leur utilité en physique. Leur statut est dorénavant solidement établi, personne n'en contestant plus la légitimité. Comme à chaque découverte d'un nouveau type de nombre, le concept même de nombre a dû être remanié.

1. Le corps des nombres complexes

1.a. Les équations du second degré insolubles dans $\mathbb{R}$

L'ensemble des réels ne contient pas de solution pour certaines équations du second degré comme $x^2+1=0$ ou, plus généralement, $ax^2+bx+c=0$ lorsque $\Delta=b^2-4ac<0$. Le carré d'un réel étant toujours positif, il ne peut égaler -1 ce qui est pourtant la seule façon d'annuler x^2+1.
En introduisant le nombre $\sqrt{-1}$, que l'on va noter i, on a $i^2=(-i)^2=-1$ et l'équation $x^2+1=0$ se retrouve avec deux solutions qui sont i et $-i$. Cet artifice de la numération est un « nombre » (un symbole se prêtant aux opérations que l'on définit avec lui) mais ce n'est pas un nombre réel.
On peut désormais résoudre les équations du second degré qui n'avaient pas de solution dans $\mathbb{R}$.
En effet, les solutions de l'équation $ax^2+bx+c=0$ existent lorsque $\Delta<0$ puisque
$\sqrt{\Delta}=\sqrt{(-1)\times(-\Delta)}=\sqrt{-1}\times\sqrt{-\Delta}=i\sqrt{-\Delta}$ et le nombre $\sqrt{-\Delta}$ est bien un réel positif.
Les solutions de l'équation s'écrivent alors sous la forme générale $x=a+ib$ où a et b sont des réels :

$$x=\frac{-b\pm i\sqrt{\Delta}}{2a}=\frac{-b}{2a}\pm i\frac{\sqrt{\Delta}}{2a}$$

Définition 5.1 (Nombre complexe) Soit i le nombre imaginaire définit par $i^2=-1$ et $\mathbb{C}$ l'ensemble des nombres z s'écrivant $a+bi$ avec $(a,b)\in\mathbb{R}^2$. La partie réelle de z notée $\mathcal{R}e(z)$ est le nombre réel a tandis que la partie imaginaire de z notée $\mathcal{I}m(z)$ est le nombre réel b.
Les nombres de $\mathbb{C}$ sont appelés « nombres complexes » et, lorsque la partie réelle d'un nombre complexe est nulle (le nombre s'écrit alors bi avec $b\in\mathbb{R}$), ce nombre est dit « imaginaire pur ».

Exemple 42 – On souhaite factoriser le trinôme du second degré $P(x)=3x^2-2x+1$. Cependant l'équation $3x^2-2x+1=0$ n'a pas de solution réelle car $\Delta=(-2)^2-4\times3=-8<0$.
Calculons les solutions complexes :
Comme $\sqrt{-8}=\sqrt{(-1)\times8}=\sqrt{-1}\times\sqrt{8}=\sqrt{-1}\times\sqrt{4\times2}=2i\sqrt{2}$.
Ces solutions s'écrivent $\frac{2\pm2i\sqrt{2}}{6}=\frac{1\pm i\sqrt{2}}{3}$ ou aussi $\frac{1}{3}\pm i\frac{\sqrt{2}}{3}$.
En anticipant sur les règles du calcul algébrique dans $\mathbb{C}$, on peut donc factoriser le trinôme :

$$P(x)=3x^2-2x+1=3\left(x+\frac{1}{3}+i\frac{\sqrt{2}}{3}\right)\left(x-\frac{1}{3}-i\frac{\sqrt{2}}{3}\right)$$

Remarques :

- L'ensemble $\mathbb{C}$ contient tous les réels, tous les imaginaires purs (leur ensemble est noté $\mathbb{R}i$) et toutes les combinaisons engendrées par la somme de ces deux types de nombres.
  On a donc $\mathbb{R}\subset\mathbb{C}$ et, de même, $\mathbb{R}i\subset\mathbb{C}$.
- Toutes les équations du second degré $ax^2+bx+c=0$ ont finalement deux solutions dans $\mathbb{C}$:
  — si $\Delta=b^2-4ac>0$, $x=\frac{-b}{2a}\pm\frac{\sqrt{\Delta}}{2a}$, les deux solutions sont réelles
  — si $\Delta=0$, $x=\frac{-b}{2a}$, la solution réelle est double car alors $ax^2+bx+c=a(x-\frac{-b}{2a})^2$
  — si $\Delta<0$, $x=\frac{-b}{2a}\pm i\frac{\sqrt{\Delta}}{2a}$, les deux solutions sont complexes non réelles
- Le nombre i est souvent noté j par les physiciens qui ne veulent pas le confondre avec le i de l'intensité électrique. Mais on va voir que, pour les mathématiciens, j est un autre nombre complexe égal à $\frac{-1}{2}+i\frac{\sqrt{3}}{2}$.

Définition 5.2 (Nombre conjugué) Soit $z = a + ib$ un nombre complexe.
On appelle conjugué de z le nombre $a - ib$ que l'on note $\bar{z}$.
Les propriétés immédiates sont $z + \bar{z} = 2a$ et $z \times \bar{z} = (a + ib)(a - ib) = a^2 - (ib)^2 = a^2 + b^2$,
Le nombre $z \times \bar{z} = a^2 + b^2$ est le carré du « module » de z (voir plus loin) et noté $|z|^2$.

Exemple 43 – L'équation $x^2 + x + 1 = 0$ n'a pas de solution réelle car $\Delta = 1^2 - 4 = -3 < 0$, mais elle a deux solutions complexes car $\sqrt{-3} = \sqrt{(-1) \times 3} = \sqrt{-1} \times \sqrt{3} = i\sqrt{3}$.
Ces solutions s'écrivent $\dfrac{-1 \pm i\sqrt{3}}{2}$ ou aussi $\dfrac{-1}{2} \pm i\dfrac{\sqrt{3}}{2}$. On note j la solution $\dfrac{-1}{2} + i\dfrac{\sqrt{3}}{2}$.
Pourquoi donne-t-on un nom à ce nombre ? Car j est une des trois racines cubiques de 1. En effet :

$$j^3 = \frac{(-1 + i\sqrt{3})^3}{2^3} = \frac{(-1)^3 + 3i(-1)^2\sqrt{3} + 3i^2(-1)(\sqrt{3})^2 + i^3(\sqrt{3})^3}{8} = \frac{-1 + 3i\sqrt{3} + 3 \times 3 - 3i\sqrt{3}}{8} = 1$$

Les racines cubiques de 1 vérifient l'équation $x^3 = 1$ et, hormis 1 qui est une racine réelle évidente, il y a deux racines complexes conjuguées. En écrivant que $x^3 - 1 = (x - 1)(x^2 + x + 1)$, on retrouve le trinôme $x^2 + x + 1$ qui a pour racines j et son conjugué $\bar{j}$ égal à $\dfrac{-1}{2} - i\dfrac{\sqrt{3}}{2}$.
$\bar{j}$ est égal à j^2 car $j^2 = \dfrac{(-1 + i\sqrt{3})^2}{2^2} = \dfrac{1 - 2i\sqrt{3} + i^2(\sqrt{3})^2}{4} = \dfrac{-2 - 2i\sqrt{3}}{4} = \dfrac{-1 - i\sqrt{3}}{2}$, mais aussi car $j \times j^2 = j^3 = 1$, comme on vient de le voir.
Le fait que j et j^2 soient conjugués implique que :

- $\mathcal{R}e(j^2) = \mathcal{R}e(j) = \dfrac{-1}{2}$ et $\mathcal{I}m(j^2) = -\mathcal{I}m(j) = -\dfrac{\sqrt{3}}{2}$
- Leur somme vaut $j + j^2 = 2 \times \mathcal{R}e(j) = 2 \times \frac{-1}{2} = -1$ (comme j est solution de $x^2 + x + 1 = 0$, il n'est pas étonnant que $j^2 + j = -1$).
- Leur produit est $j \times j^2 = (\mathcal{R}e(j))^2 + (\mathcal{I}m(j))^2 = \left(\frac{-1}{2}\right)^2 + \left(\frac{\sqrt{3}}{2}\right)^2 = \frac{1}{4} + \frac{3}{4} = 1$ (comme j est une racine cubique de 1, il n'est pas étonnant que $j \times j^2 = j^3 = 1$).

Propriété 5.1 (Conjugaison) $\forall (z, z') \in \mathbb{C}^2$ on a :

a) $z + \bar{z} = 2\mathcal{R}e(z)$	b) $z - \bar{z} = 2i\mathcal{I}m(z)$	c) $z \times \bar{z} = \|z\|^2$
d) $\bar{\bar{z}} = z$	e) $\overline{-z} = -\bar{z}$	
f) $z \in \mathbb{R} \Rightarrow \bar{z} = z$	g) $z \in \mathbb{R}i \Rightarrow \bar{z} = -z$	h) $\bar{z} = \bar{z'} \Leftrightarrow z = z'$

Démonstration Toutes ces propriétés sont évidentes. Il suffit de partir de la définition du conjugué et d'utiliser les propriétés algébriques de $\mathbb{C}$ qui prolongent celle de $\mathbb{R}$ comme on va le voir plus loin.
Montrons par exemple que $\bar{\bar{z}} = z$:
Si $z = a + ib$ alors $\bar{z} = a - ib$ et $\bar{\bar{z}} = a - (-ib) = a + ib$.
Montrons maintenant l'équivalence $\bar{z} = \bar{z'} \iff z = z'$:

- La propriété directe : $z = z' \Longrightarrow \bar{z} = \bar{z'}$ est évidente.
- Réciproquement, si $\bar{z} = \bar{z'}$ alors, d'après la propriété directe on doit avoir $\bar{\bar{z}} = \bar{\bar{z'}}$, et comme $\bar{\bar{z}} = z$ et $\bar{\bar{z'}} = z'$, on en déduit que $z = z'$.

1.b. Algèbre complexe

L'invention des nombres complexes étend les propriétés algébriques connues dans $\mathbb{R}$.
Ainsi, l'égalité, l'addition et la multiplication conservent les propriétés qu'on leur connait dans $\mathbb{C}$: + et × sont commutatives et associatives ; × est distributive par rapport à + ; on peut ajouter aux deux membres d'une égalité un même nombre, ou les multiplier par un même nombre non nul ; chaque nombre à un opposé et un inverse (sauf 0) ; etc.

Propriété 5.2 (Unicité) L'écriture algébrique d'un nombre complexe de $\mathbb{C}$ est unique.
Soient $z = a + bi$ et $z' = a' + b'i$ deux nombres complexes. $z = z' \iff \begin{cases} a = a' \\ b = b' \end{cases}$

Démonstration $z = z' \iff a + bi = a' + b'i \iff a - a' = i(b - b')$
Comme $a - a' \in \mathbb{R}$ et $i(b - b') \in \mathbb{R}i$, la seule possibilité est $b - b' = 0$ et par conséquent $a - a' = 0$.

Exemple 44 – Résolvons l'équation complexe $z^2 = 2 + i$.
En notant $z = a + bi$, cette équation s'écrit
$(a + bi)^2 = 2 + i \iff a^2 - b^2 + 2abi = 2 + i \iff \begin{cases} a^2 - b^2 = 2 \\ 2ab = 1 \end{cases}$
On a nécessairement $a \neq 0$ car sinon on aurait $-b^2 = 2$ ce qui n'est pas possible dans $\mathbb{R}$.
De la 2e équation, on tire donc $b = \frac{1}{2a}$ que l'on substitue dans la 1re pour écrire $a^2 - \frac{1}{4a^2} = 2$.
Il nous faut résoudre l'équation bicarrée $4a^4 - 1 = 8a^2$ qui s'écrit $4A^2 - 8A - 1 =$ avec le changement de variable $A = a^2$. Comme $\Delta = 64 + 16 = 80 = 5 \times 4^2$, les solutions sont $A = \frac{8 \pm \sqrt{5 \times 4^2}}{8} = \frac{2 \pm \sqrt{5}}{2}$.
La solution $A = \frac{2 - \sqrt{5}}{2} \approx -0.236 < 0$ est impossible car A est le carré d'un réel.
On doit donc avoir $A = \frac{2 + \sqrt{5}}{2}$ et donc $a = \pm\sqrt{\frac{2 + \sqrt{5}}{2}}$.
Comme $b = \frac{1}{2a}$, on trouve finalement les deux solutions :

- $a = \sqrt{\frac{2+\sqrt{5}}{2}}$ et $b = \frac{1}{2} \times \sqrt{\frac{2}{2+\sqrt{5}}} = \frac{1}{\sqrt{2(2+\sqrt{5})}}$ d'où $z_1 = \sqrt{\frac{2+\sqrt{5}}{2}} + \frac{i}{\sqrt{2(2+\sqrt{5})}}$
- $a = -\sqrt{\frac{2+\sqrt{5}}{2}}$ et $b = \frac{-1}{\sqrt{2(2+\sqrt{5})}}$ d'où $z_2 = -z_1 = -\sqrt{\frac{2+\sqrt{5}}{2}} - \frac{i}{\sqrt{2(2+\sqrt{5})}}$

Propriété 5.3 (Addition) L'addition dans $\mathbb{C}$, définie par $(a+bi)+(a'+b'i) = (a+a')+(b+b')i$, confère à $(\mathbb{C},+)$ la structure de groupe commutatif.

- Commutativité : $\forall (z, z') \in \mathbb{C}^2, z + z' = z' + z$
- Associativité : $\forall (z, z', z'') \in \mathbb{C}^3, z + (z' + z'') = (z + z') + z''$, expression notée $z + z' + z''$
- Élément neutre : $\forall z \in \mathbb{C}, z + 0 = 0 + z = z$, le réel 0 est l'élément neutre de $+$ dans $\mathbb{C}$
- Opposé : $\forall z \in \mathbb{C}, \exists z' \in \mathbb{C}, z + z' = 0$ et lorsque $z = a + bi$, cet élément z' est $-z = -a - bi$

Remarques :

- Ces propriétés découlent du fait que l'on définit $\mathbb{C}$ comme une extension de $\mathbb{R}$, donc avec les même propriétés : $(\mathbb{R},+)$ étant un groupe commutatif, $(\mathbb{C},+)$ en est un aussi.
- La soustraction dans $\mathbb{C}$ est définie comme dans $\mathbb{R}$ par l'addition de l'opposé : $\forall (z, z') \in \mathbb{C}^2, z - z' = z + (-z')$

Propriété 5.4 (Multiplication) La multiplication dans $\mathbb{C}$, définie par $(a+bi) \times (a'+b'i) = (aa'-bb')+(ab'+ba')i$, confère à $(\mathbb{C}^*,\times)$ la structure de groupe commutatif.

- Commutativité : $\forall (z, z') \in \mathbb{C}^2, z \times z' = z' \times z$
- Associativité : $\forall (z, z', z'') \in \mathbb{C}^3, z \times (z' \times z'') = (z \times z') \times z''$, expression notée $zz'z''$
- Élément neutre : $\forall z \in \mathbb{C}, z \times 1 = 1 \times z = z$, le réel 1 est l'élément neutre de $\times$ dans $\mathbb{C}$
- Inverse : $\forall z \in \mathbb{C}^*, \exists z' \in \mathbb{C}^*, z \times z' = 1$ et lorsque $z = a + bi$, cet élément z' est $\dfrac{a - bi}{a^2 + b^2}$

Démonstration Ici encore, ces propriétés découlent du fait que l'on définit $\mathbb{C}$ comme une extension de $\mathbb{R}$. Le dernier point mérite d'être justifié :
Soient $z = a + bi$ et $z' = a' + b'i$ tels que $z \times z' = 1$. On doit donc avoir $(aa' - bb') + (ab' + ba')i = 1$ et cette égalité conduit au système $\begin{cases} aa' - bb' = 1 \\ ab' + ba' = 0 \end{cases}$

En multipliant la 1re égalité par a et la 2e par b, on obtient
$\begin{cases} a^2a' - abb' = a \\ abb' + b^2a' = 0 \end{cases} \Longrightarrow$ (addition membre à membre) $a^2a' + b^2a' = (a^2+b^2)a' = a \Longrightarrow a' = \dfrac{a}{a^2+b^2}$
En multipliant la 1re égalité par $-b$ et la 2e par a, on obtient
$\begin{cases} -aba' + b^2b' = -b \\ a^2b' + aba' = 0 \end{cases} \Longrightarrow$ (addition membre à membre) $b^2b' + a^2b' = (a^2+b^2)b' = -b \Longrightarrow b' = \dfrac{-b}{a^2+b^2}$
Finalement on obtient bien $z' = \dfrac{a}{a^2+b^2} + \dfrac{-bi}{a^2+b^2} = \dfrac{a-bi}{a^2+b^2}$.

Remarques :

- Comme dans $\mathbb{R}$, 0 n'a pas d'inverse dans $\mathbb{C}$.
- Cas particulier de multiplication : $\forall (k,a,b) \in \mathbb{R}^3, k \times (a+bi) = (ka) + (kb)i$.
- Si $z = a+bi$ alors le conjugué de z est $\bar{z} = a - bi$ et le carré de son module est $|z|^2 = a^2+b^2$, du coup l'inverse de z est $\dfrac{1}{z} = \dfrac{\bar{z}}{|z|^2}$ et comme la division dans $\mathbb{C}$ est définie comme dans $\mathbb{R}$ par la multiplication de l'inverse : $\forall z \in \mathbb{C}, \forall z' \in \mathbb{C}^*, \dfrac{z}{z'} = z \times \dfrac{1}{z'} = \dfrac{z\overline{z'}}{|z'|^2}$.
  Avec les notations algébriques $z = a+bi$ et $z' = a' + b'i$ on a :
  $\dfrac{z}{z'} = (a+bi) \times \dfrac{a' - b'i}{a'^2 + b'^2} = \dfrac{(a+bi)(a'-b'i)}{a'^2+b'^2} = \dfrac{(aa'+bb') + (-ab'+ba')i}{a'^2+b'^2}$.
- Comme $(\mathbb{C},+)$ est un groupe commutatif, $(\mathbb{C}^*,\times)$ est un groupe commutatif et comme $\times$ est distributive sur $+$, l'ensemble des complexes muni de ces deux opérations $(\mathbb{C},+,\times)$ est un corps commutatif. Dedekind définit un corps (*Körper* en allemand) comme un sous-ensemble de nombres réels ou complexes stable par addition, soustraction, multiplication et division. Outre $\mathbb{C}$, $\mathbb{R}$ et $\mathbb{Q}$ (l'ensemble des rationnels), l'ensemble des entiers modulo un nombre premier p (noté $\mathbb{Z}/p\mathbb{Z}$) est un corps (l'ensemble $\mathbb{Z}/5\mathbb{Z}$ par exemple contient les éléments $\{0,1,2,3,4\}$)

Exemple 45 – ➲ Écrivons plus simplement le quotient de $5-2i$ par $-1+3i$:

$$\frac{5-2i}{-1+3i} = \frac{(5-2i)(-1-3i)}{(-1+3i)(-1-3i)} = \frac{(-5-6)+(-15+2)i}{(-1)^2+(-3)^2} = \frac{-11-13i}{10}$$

L'équation $(-1+3i)z = 5-2i$ a donc pour solution $z = \dfrac{-11-13i}{10} = \dfrac{-11}{10} - \dfrac{13i}{10} = -1,1 - 1,3i$

➲ Si $z = a+bi$, écrivons plus simplement le quotient z' de $z-2i$ par $z+1$:

$$z' = \frac{z-2i}{z+1} = \frac{a+(b-2)i}{(a+1)+bi} = \frac{(a+(b-2)i)(a+1-bi)}{(a+1+bi)(a+1-bi)} = \frac{a(a+1)+b(b-2)+i(-ab+(a+1)(b-2))}{(a+1)^2+b^2}$$

On en déduit que $\mathcal{Re}(z') = \dfrac{a(a+1)+b(b-2)}{(a+1)^2+b^2}$ et $\mathcal{Im}(z') = \dfrac{-ab+(a+1)(b-2)}{(a+1)^2+b^2} = \dfrac{b-2a-2}{(a+1)^2+b^2}$

➲ Si $z = 1+2i$ et $z' = 2-i$, écrivons algébriquement le complexe $(z^2-1)(i-z')$:

$(z^2-1)(i-z') = ((1+2i)^2-1)(i-(2-i)) = (1-4+4i-1)(i-2+i) = 4\times2(-1+i)(i-1) = 8(-2i) = -16i$

Propriété 5.5 (Conjugaison (suite)) $\forall (z,z') \in \mathbb{C}^2, \forall n \in \mathbb{N}^*$ on a :

a) $\overline{z+z'} = \bar{z} + \overline{z'}$	b) $\overline{z-z'} = \bar{z} - \overline{z'}$	c) $\overline{(-z)} = -\bar{z}$
d) $\overline{z \times z'} = \bar{z} \times \overline{z'}$	e) $\overline{\left(\dfrac{1}{z}\right)} = \dfrac{1}{\bar{z}}$	f) $\overline{\left(\dfrac{z}{z'}\right)} = \dfrac{\bar{z}}{\overline{z'}}$
g) $\overline{\sum_{i=1}^{n} z_i} = \sum_{i=1}^{n} \overline{z_i}$	h) $\overline{\prod_{i=1}^{n} z_i} = \prod_{i=1}^{n} \overline{z_i}$	i) $\overline{(z^n)} = (\bar{z})^n$

DÉMONSTRATION ➠ Les propriétés de la 1re ligne sont évidentes.
Montrons que $\overline{z \times z'} = \bar{z} \times \bar{z'}$ en notant $z = a + ib$ et $z' = a' + ib'$:

- $\overline{z \times z'} = \overline{(aa' - bb') + (ab' + ba')i} = (aa' - bb') - (ab' + ba')i$
- $\bar{z} \times \bar{z'} = (a - ib)(a' - ib') = (aa' - (-b)(-b')) + (a(-b') + (-b)a')i = (aa' - bb') - (ab' + ba')i$

➠ Montrons que $\overline{\left(\frac{1}{z}\right)} = \frac{1}{\bar{z}}$:
Par définition de l'inverse $z \times \frac{1}{z} = 1$ et d'après la propriété précédente $\overline{z \times \frac{1}{z}} = \bar{z} \times \overline{\left(\frac{1}{z}\right)}$.
Comme de plus $\overline{z \times \frac{1}{z}} = \bar{1} = 1$, on en déduit que $\bar{z} \times \overline{\left(\frac{1}{z}\right)} = 1 \Longleftrightarrow \overline{\left(\frac{1}{z}\right)} = \frac{1}{\bar{z}}$

➠ Montrons maintenant que $\overline{\left(\frac{z}{z'}\right)} = \frac{\bar{z}}{\bar{z'}}$:
Comme $\frac{z}{z'} = z \times \frac{1}{z'}$, on en déduit que $\overline{\left(\frac{z}{z'}\right)} = \overline{z \times \frac{1}{z'}} = \bar{z} \times \overline{\left(\frac{1}{z'}\right)} = \bar{z} \times \frac{1}{\bar{z'}} = \frac{\bar{z}}{\bar{z'}}$

➠ Pour les deux 1re propriétés de la dernière ligne, on procède par récurrence sur $n \geqslant 1$:

- Pour $n = 1$: $\overline{\sum_{i=1}^{1} z_i} = \overline{z_1} = \sum_{i=1}^{1} \overline{z_i}$
- Supposons que pour $n \geqslant 1$, on aie $\overline{\sum_{i=1}^{n} z_i} = \sum_{i=1}^{n} \overline{z_i}$, calculons :

$\overline{\sum_{i=1}^{n+1} z_i} = \overline{\sum_{i=1}^{n} z_i + z_{n+1}} = \overline{\sum_{i=1}^{n} z_i} + \overline{z_{n+1}} = \sum_{i=1}^{n+1} \overline{z_i}$
L'initialisation et l'hérédité de cette propriété étant vraie, la propriété est vraie pour tout $n \geqslant 1$.
On procède de même pour le produit :

- Pour $n = 1$: $\overline{\prod_{i=1}^{1} z_i} = \overline{z_1} = \prod_{i=1}^{1} \overline{z_i}$
- Supposons que pour $n \geqslant 1$, on aie $\overline{\prod_{i=1}^{n} z_i} = \prod_{i=1}^{n} \overline{z_i}$, calculons :

$\overline{\prod_{i=1}^{n+1} z_i} = \overline{\prod_{i=1}^{n} z_i \times z_{n+1}} = \overline{\prod_{i=1}^{n} z_i} \times \overline{z_{n+1}} = \prod_{i=1}^{n+1} \overline{z_i}$
Le principe de récurrence permet alors de conclure pour tout $n \geqslant 1$.

➠ La dernière propriété est un cas particulier de la précédente :
$\forall n \geqslant 1, \overline{(z^n)} = \overline{\prod_{i=1}^{n} z} = \prod_{i=1}^{n} \bar{z} = \overline{(z)}^n$

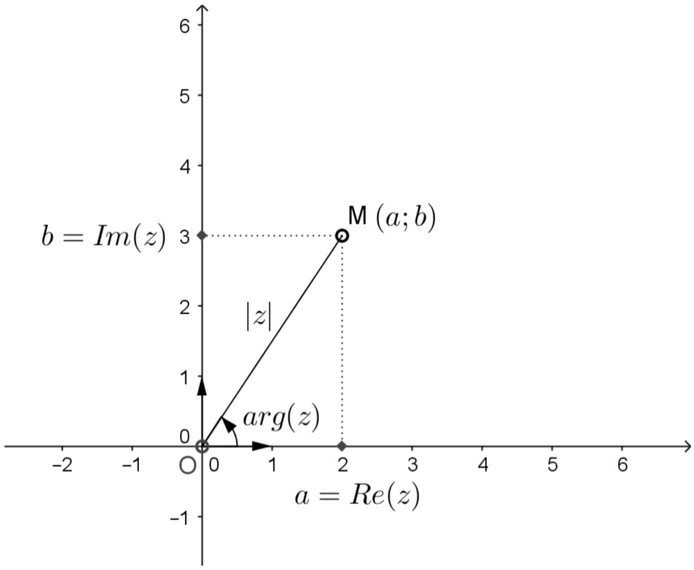
2. Géométrie dans le plan complexe

Si les nombres complexes sont apparus dans le contexte de la résolution des équations du 3e degré, les utilisations que l'on peut en faire sont sans limite. Le principal atout de ces nombres est de combiner deux nombres en un seul dont on peut toujours séparer chacun des constituants. Cette avantage confère aux complexes la capacité de représenter des points ou des vecteurs du plan.

2.a. Le plan complexe

Depuis Jean-Robert Argand (1806), le nombre complexe $z = a + ib$ est appelé « affixe » du point M de coordonnées $(a\,;\,b)$ ou du vecteur $\overrightarrow{OM}$ dans un repère orthonormé d'origine O.
La partie réelle de z est l'abscisse de M tandis que la partie imaginaire de z est son ordonnée.
Chaque nombre de $\mathbb{C}$ correspond ainsi à un point de ce plan, appelé « plan complexe ».

Définition 5.3 (Affixe et module) Le plan complexe ayant comme repère $(O, \vec{\imath}, \vec{\jmath})$, un point M de ce plan est repéré par son « affixe » complexe $z = a + ib$, ce que l'on note $M(z)$. Le vecteur $\overrightarrow{OM}$ est également repéré par l'affixe $z = a + ib$ et on note cela $\overrightarrow{OM}(z)$.
Le « module » de z, noté $|z|$, est le nombre $\sqrt{a^2+b^2} = \sqrt{z \times \bar{z}}$ qui correspond à la distance OM.
Les coordonnées du point $M(z)$ sont $(\mathcal{R}e(z), \mathcal{I}m(z))$.



Définition 5.4 (Argument) Dans le plan complexe de repère $(O, \vec{\imath}, \vec{\jmath})$, un point $M(z)$ étant donné, l'angle orienté $(\vec{\imath}, \overrightarrow{OM})$ est appelé « argument » de z et noté $\arg(z)$.

Remarques :

- $\arg(a+ib)$ est un angle, mesuré en radians, qu'on détermine à la calculatrice si $a \neq 0$.
  - Si $a > 0$, on tape $\tan^{-1}\left(\frac{b}{a}\right)$ ($\tan^{-1}$ est parfois notée *atan*) puisque $\tan(\arg(z)) = \frac{b}{a}$.
    Si $b = 0$ alors $\arg(z) = 0$ et $|z| = a$.
  - Si $a < 0$, comme $\tan^{-1}$ donne un angle de l'intervalle $]-\frac{\pi}{2}\,;\frac{\pi}{2}[$, on tape $\tan^{-1}(\frac{b}{a}) + \pi$.
    Si $b = 0$ alors $\arg(z) = \pi$ et $|z| = -a$.
  - Si $a = 0$, alors $\arg(z) = \frac{\pi}{2}$ et $|z| = b$ si $b > 0$ ou $\arg(z) = \frac{-\pi}{2}$ et $|z| = -b$ si $b < 0$.
- Sur l'illustration ci-dessus, le point M a pour coordonnées $(2\,;3)$.
  - son affixe est le complexe $z = 2 + 3i$
  - son module est $|z| = \sqrt{2^2+3^2} = \sqrt{13}$
  - son argument est $\arg(z) = \tan^{-1}(\frac{3}{2}) \approx 0,98279$ *rad.*
- Comme $\mathcal{R}e(z)$ est porté par l'axe des abscisses, celui-ci est l'« axe des réels » ; de même, comme $\mathcal{I}m(z)$ est porté par l'axe des ordonnées, celui-ci est l'« axe des imaginaires ».
  De ce fait, si un point a pour affixe un nombre réel, ce point est sur l'axe des réels ; de même, si un point a pour affixe un nombre imaginaire, ce point est sur l'axe des imaginaires.

Propriété 5.6 (Notation trigonométrique) Tout nombre complexe non nul z peut être noté sous la forme trigonométrique $z = r(\cos\theta + i\sin\theta)$ où $r = |z|$ et $\theta = \arg(z)\ \ [2\pi]$.
On a alors $\mathcal{R}e(z) = r\cos\theta$, $\mathcal{I}m(z) = r\sin\theta$ et $z_1 = z_2 \iff \begin{cases} |z_1| = |z_2| \\ \arg(z_1) = \arg(z_2)\ \ [2\pi] \end{cases}$
De plus $|\mathcal{R}e(z)| \leqslant r$ et $|\mathcal{I}m(z)| \leqslant r$

Démonstration On a déjà montré, propriété 5.2, que $z_1 = z_2 \iff \begin{cases} \mathcal{R}e(z_1) = \mathcal{R}e(z_2) \\ \mathcal{I}m(z_1) = \mathcal{I}m(z_2) \end{cases}$.

Par conséquent $z_1 = z_2 \iff M(z_1) = M'(z_2)$. Les points M et M' sont confondus car leurs coordonnées cartésiennes sont égales (les parties réelles et imaginaires de leurs affixes).
Leurs coordonnées polaires (modules et arguments) sont donc égales car la correspondance entre l'écriture algébrique et l'écriture trigonométrique d'un complexe est biunivoque :

- Si le point M a pour affixe $z = a + ib$, choisir $|z| = \sqrt{a^2+b^2}$ et $\arg(z) = (\vec{\imath}, \overrightarrow{OM})$ assure l'unicité du module et de l'argument de z.
- Si le point M a pour module r et pour argument $\theta\ [2\pi]$, choisir $x = r\cos\theta$ et $y = r\sin\theta$ assure l'unicité de la partie réelle et de la partie imaginaire de z.

Exemple 46 – ➲ Le point $I(1)$ est sur l'axe réel.
Comme il est positif, son argument est $0\ [2\pi]$ et son module 1. Par conséquent $1 = \cos 0 + i \sin 0$.
➲ Le point $J(i)$ est sur l'axe imaginaire.
Comme sa partie imaginaire est positive ($\mathcal{I}m(i) = 1$), son argument est $\frac{\pi}{2}\ [2\pi]$ et son module 1. La notation trigonométrique de i est $\cos\frac{\pi}{2} + i\sin\frac{\pi}{2}$.
➲ Le point $O(0)$ a un module égal à 0 mais son argument n'est pas défini.
➲ Le point d'affixe j a pour coordonnées $(-\frac{1}{2}\,;\frac{\sqrt{3}}{2})$.

- Son module est $|j| = \sqrt{(-\frac{1}{2})^2 + (\frac{\sqrt{3}}{2})^2} = \sqrt{\frac{1}{4}+\frac{3}{4}} = 1$
- Comme $a = -\frac{1}{2} < 0$, son argument $\arg(j) = \tan^{-1}(\frac{\sqrt{3}\div 2}{-1\div 2}) + \pi \approx 2,094395102$.
  Comme $3 \times 2,094395102 \approx 6,283185307 \approx 2\pi$, on peut conclure que $\arg(j) = \frac{2\pi}{3} rad$.
  Mais cela se montre plus directement : comme $|j| = 1$, le point d'affixe j est sur le cercle trigonométrique. Son argument θ est tel que $\cos\theta = -\frac{1}{2}$ et $\sin\theta = \frac{\sqrt{3}}{2}$.
  Un seul angle de $[0\,;2\pi[$ a ces caractéristiques, c'est $\frac{2\pi}{3}$.

Ainsi, la notation trigonométrique de j est $\cos\frac{2\pi}{3} + i\sin\frac{2\pi}{3}$.

Propriété 5.7 (Points remarquables) $\forall z \in \mathbb{C}$, on a :

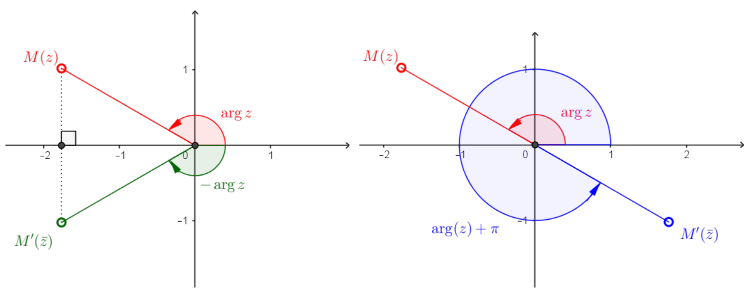
$\arg(\bar{z}) = -\arg(z)\ [2\pi]$	$\lvert\bar{z}\rvert = \lvert z\rvert$	les points $M(z)$ et $M'(\bar{z})$ sont symétriques par rapport à l'axe réel.
$\arg(-z) = \arg(z) + \pi\ [2\pi]$	$\lvert -z\rvert = \lvert z\rvert$	les points $M(z)$ et $M''(-z)$ sont symétriques par rapport à l'origine.

Démonstration M et M' sont symétriques par rapport à l'axe des réels car leurs affixes ont même partie réelle et des parties imaginaires opposées. On a déjà montré que $\bar{z}$ et z ont même module.
Montrons que leurs arguments sont opposés :

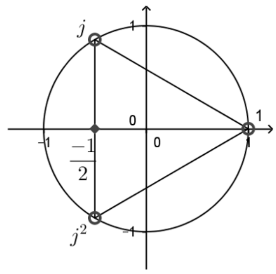
- $\mathcal{R}e(\bar{z}) = \mathcal{R}e(z) = r\cos\theta = r\cos(-\theta)$ car $\forall\theta, \cos(-\theta) = \cos\theta$
- $\mathcal{I}m(\bar{z}) = -\mathcal{I}m(z) = -r\sin\theta = r(-\sin\theta) = r\sin(-\theta)$ car $\forall\theta, \sin(-\theta) = -\sin\theta$

M et M'' sont symétriques par rapport à l'origine car leurs affixes ont des parties réelles et imaginaires opposées. Les modules sont égaux : $|-a-bi| = (-a)^2 + (-b)^2 = a^2 + b^2 = |a+bi|$.
Montrons que les arguments de $-z$ et z diffèrent de π :

- $\mathcal{R}e(-z) = -\mathcal{R}e(z) = -r\cos\theta = r\cos(\theta+\pi)$ car $\forall\theta, \cos(\theta+\pi) = -\cos\theta$
- $\mathcal{I}m(-z) = -\mathcal{I}m(z) = -r\sin\theta = r\sin(\theta+\pi)$ car $\forall\theta, \sin(\theta+\pi) = -\sin\theta$



Remarque :
$j^2 = \bar{j} = \cos\frac{-2\pi}{3} + i\sin\frac{-2\pi}{3}$ car $\bar{j}$ a même module que j et un argument opposé.
Les points d'affixe 1, j et $j^2 = \bar{j}$, les trois racines cubiques de l'unité, sont les sommets d'un triangle équilatéral inscrit dans le cercle trigonométrique. En effet, ces trois nombres ont pour module 1 et les points du cercle trigonométrique dont ils sont les affixes forment des angles au centre deux à deux égaux à $\frac{\pm 2\pi}{3}$ ($\arg(1) = 0$, $\arg(j) = \frac{2\pi}{3}$ et $\arg(j^2) = \frac{-2\pi}{3} = \frac{4\pi}{3}$).



Propriété 5.8 (Linéarité) L'association d'un vecteur à son affixe $f : \vec{v} \longmapsto z \in \mathbb{C}$ est une bijection linéaire, c'est-à-dire que quels que soient $\vec{v_1} \mapsto z_1$ et $\vec{v_2} \mapsto z_2$:

Bijection : $\vec{v_1} = \vec{v_2} \iff z_1 = z_2$ | Linéarité : $\begin{cases} \vec{v_1} + \vec{v_2} \mapsto z_1 + z_2 \\ \forall k \in \mathbb{R},\ \ k\vec{v_1} \mapsto kz_1 \end{cases}$

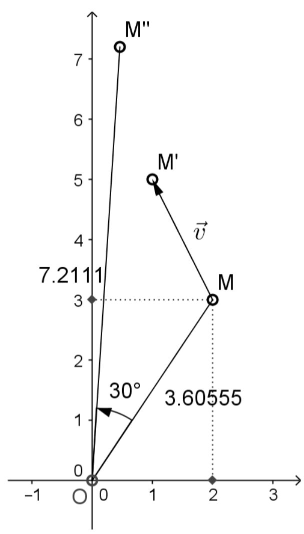
Démonstration ➲ Bijection :
Le sens direct est évident puisqu'un vecteur est associé à une affixe unique $\vec{v_1} = \vec{v_2} \Longrightarrow z_1 = z_2$.
Réciproquement si $z_1 = z_2 = z$, les composantes de $\vec{v_1}$ et $\vec{v_2}$ étant les mêmes $(\mathcal{R}e(z), \mathcal{I}m(z))$, celles-ci désignent un seul et unique vecteur $\vec{v} = \vec{v_1} = \vec{v_2}$.

➲ Linéarité de l'addition :
Sous forme algébrique, si $z_1 = a_1 + ib_1$ et $z_2 = a_2 + ib_2$ alors $z_1 + z_2 = (a_1 + a_2) + (b_1 + b_2)i$.
La décomposition de $\vec{v_1} + \vec{v_2}$ dans la base $(\vec{\imath}, \vec{\jmath})$ est $(a_1\vec{\imath} + b_1\vec{\jmath}) + (a_2\vec{\imath} + b_2\vec{\jmath}) = (a_1 + a_2)\vec{\imath} + (b_1 + b_2)\vec{\jmath}$.
On constate que l'affixe complexe de $\vec{v_1} + \vec{v_2}$ est $(a_1 + a_2) + (b_1 + b_2)i$, soit $z_1 + z_2$.

➲ Linéarité de la multiplication par un réel :
De même, si $z_1 = a_1 + ib_1$ alors $kz_1 = k(a_1 + ib_1) = (ka_1) + (kb_1)i$.
La décomposition de $k\vec{v_1}$ dans la base $(\vec{\imath}, \vec{\jmath})$ est $k(a_1\vec{\imath} + b_1\vec{\jmath}) = (ka_1)\vec{\imath} + (kb_1)\vec{\jmath}$.
On constate que l'affixe complexe de $k\vec{v_1}$ est $(ka_1) + (kb_1)i$, soit kz_1.

Remarques :

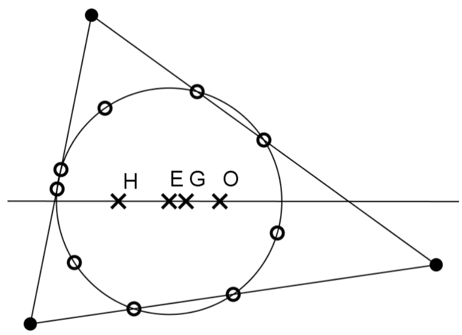
- L'addition d'un complexe z_1 à un complexe z_2 prend du sens si on considère le point $M(z_1)$ et le vecteur $\vec{v}$ d'affixe z_2, car le complexe $z = z_1 + z_2$ est alors l'affixe de l'image de M par la translation de vecteur $\vec{v}$. Par exemple, lorsqu'on fait subir au point $M(2+3i)$ ci-contre la translation de vecteur $\vec{v}(-1+2i)$, on obtient le point M' d'affixe
$(2+3i) + (-1+2i) = 1 + 5i$
- Pour des points $M(z)$ et $M'(z')$, le complexe $z' - z$ est l'affixe du vecteur $\overrightarrow{OM'} - \overrightarrow{OM} = \overrightarrow{MM'}$. La soustraction de deux complexes permet de calculer la distance $MM' = |z' - z|$.
Sur l'exemple ci-contre, l'affixe du vecteur $\overrightarrow{MM'}$ est
$(1+5i) - (2+3i) = -1 + 2i$
On en déduit $MM' = |-1+2i| = \sqrt{(-1)^2 + 2^2} = \sqrt{5}$



- Pour des points $M(z)$ et $M'(z')$, le complexe $\frac{z+z'}{2}$ est l'affixe du milieu I du segment $[MM']$.
  En effet, on sait qu'alors $\overrightarrow{OI} = \dfrac{\overrightarrow{OM} + \overrightarrow{OM'}}{2}$, l'affixe de I en découle.
- La traduction de l'égalité vectorielle $\overrightarrow{AB} = \overrightarrow{DC}$ ($ABCD$ est un parallélogramme) avec les affixes est $z_B - z_A = z_C - z_D$ ($[AB]$ et $[DC]$ sont parallèles, égaux et de même sens) ou, ce qui revient au même, $z_A + z_C = z_B + z_D$ ($[AC]$ et $[BD]$ ont même milieu).

Exemple 47 – Généralisons la situation du triangle dont les sommets ont pour affixes les racines cubiques de l'unité : Si des points M, M' et M'' ont des affixes z, z' et z'' telles que $|z| = |z'| = |z''|$, montrons que le triangle $MM'M''$ est équilatéral si et seulement si $z + z' + z'' = 0$.
L'égalité $|z| = |z'| = |z''| \iff OM = OM' = OM''$ indique que M, M' et M'' sont sur un même cercle de centre O.
L'égalité $z + z' + z'' = 0$ traduit l'égalité vectorielle $\overrightarrow{OM} + \overrightarrow{OM'} + \overrightarrow{OM''} = \overrightarrow{0}$.
L'origine du repère – le point O – est donc le centre de gravité du triangle $MM'M''$.
Le centre du cercle circonscrit étant confondu avec le centre de gravité, le triangle est équilatéral.

Exemple 48 – Montrons maintenant que le centre de gravité G, le centre du cercle circonscrit O et l'orthocentre H d'un triangle quelconque $MM'M''$ sont alignés. Sans perte de généralité, considérons que les affixes sont données dans un repère d'origine O.
L'affixe de G est le complexe $g = \dfrac{z + z' + z''}{3}$ car $\overrightarrow{OM} + \overrightarrow{OM'} + \overrightarrow{OM''} = 3\overrightarrow{OG}$.
L'affixe du milieu N de $[M'M'']$ est $n = \dfrac{z' + z''}{2}$ car $\overrightarrow{ON} = \dfrac{\overrightarrow{OM'} + \overrightarrow{OM''}}{2}$.
Le point H d'affixe $h = z + z' + z''$ est tel que $h - z = 2\left(\dfrac{z' + z''}{2}\right) \iff h - z = 2n$, ce qui traduit le fait que $\overrightarrow{MH} = 2\overrightarrow{ON}$. La droite (ON) étant la médiatrice de $[M'M'']$, on en déduit que (MH) est parallèle à cette médiatrice, donc perpendiculaire au côté $[M'M'']$. Comme (MH) passe par M, c'est donc la hauteur issue de M, et il en va de même pour les autres hauteurs qui passent toutes par H. Le point H est donc l'orthocentre du triangle.
Les affixes de G et H étant liées par l'égalité $h = 3g$, on en déduit que $\overrightarrow{OH} = 3\overrightarrow{OG}$, ce qui prouve l'alignement de O, G et H sur une droite appelée « droite de Euler » du triangle. Sur cette droite, un point se distingue : le point E d'affixe $e = \dfrac{z + z' + z''}{2}$.
E est le milieu du segment $[OH]$ et $e - \frac{z'+z''}{2} = \frac{z}{2} \iff e - n = \dfrac{z}{2}$. Les modules de ces deux complexes sont donc égaux, or $|e - n| = EN$ et $|\dfrac{z}{2}|$ est la moitié du rayon du cercle circonscrit. En faisant de même avec les deux autres sommets, on obtient que E est équidistant des trois milieux des côtés du triangle. Le cercle centré sur E contient d'autres points remarquables du triangle (les pieds des hauteurs et les milieux des segments joignant les sommets à l'orthocentre, voir la figure ci-dessous), on l'appelle « cercle de Euler » ou « cercle des neuf points » du triangle.



Propriété 5.9 (Multiplication) Soient z, z_1 et z_2 des complexes quelconques :

- $|z_1 \times z_2| = |z_1| \times |z_2|$ et $\arg(z_1 \times z_2) = \arg z_1 + \arg z_2 \quad [2\pi]$
- $\forall n \in \mathbb{N}, |z^n| = |z|^n$ et $\arg(z^n) = n \arg z \quad [2\pi]$
- Si $z \neq 0$ alors $|\frac{1}{z}| = \frac{1}{|z|}$ et $\arg(\frac{1}{z}) = -\arg z \quad [2\pi]$
- Si $z_2 \neq 0$ alors $|\frac{z_1}{z_2}| = \frac{|z_1|}{|z_2|}$ et $\arg(\frac{z_1}{z_2}) = \arg z_1 - \arg z_2 \quad [2\pi]$

Démonstration ➠ Écrivons $z_1 = r_1(\cos\theta_1 + i\sin\theta_1)$ et $z_2 = r_2(\cos\theta_2 + i\sin\theta_2)$ où $r_1 = |z_1|$, $r_2 = |z_2|$, $\theta_1 = \arg(z_1)$ et $\theta_2 = \arg(z_2)$. On a

$$\begin{aligned} z_1 \times z_2 &= r_1 r_2(\cos\theta_1 + i\sin\theta_1)(\cos\theta_2 + i\sin\theta_2) \\ &= r_1 r_2\left[(\cos\theta_1\cos\theta_2 - \sin\theta_1\sin\theta_2) + i(\cos\theta_1\sin\theta_2 + \cos\theta_2\sin\theta_1)\right] \\ &= r_1 r_2\left[\cos(\theta_1+\theta_2) + i\sin(\theta_1+\theta_2)\right] \end{aligned}$$

La dernière égalité vient des formules de duplication étudiées en classe de 1re grâce au produit scalaire. On tire alors de l'expression finale le module et l'argument de $z_1 \times z_2$.

➠ On montre que $|z^n| = |z|^n$ et $\arg(z^n) = n\arg z$ par récurrence sur $n \in \mathbb{N}$.

➠ Pour un complexe non nul $z = r(\cos\theta + i\sin\theta)$, son inverse
$\frac{1}{z} = \frac{\bar{z}}{|z|^2} = \frac{r(\cos\theta - i\sin\theta)}{r^2} = \frac{\cos(-\theta) + i\sin(-\theta)}{r}$ d'où $|\frac{1}{z}| = \frac{1}{r}$ et $\arg(\frac{1}{z}) = -\theta$.

➠ Comme diviser revient à multiplier par l'inverse, on en déduit :

$$\begin{aligned} \frac{z_1}{z_2} &= z_1 \times \frac{1}{z_2} = \frac{r_1(\cos\theta_1 + i\sin\theta_1)(\cos\theta_2 - i\sin\theta_2)}{r_2} \\ &= \frac{r_1}{r_2}\left((\cos\theta_1\cos\theta_2 + \sin\theta_1\sin\theta_2) + i(\sin\theta_1\cos\theta_2 - \cos\theta_1\sin\theta_2)\right) \\ &= \frac{r_1}{r_2}\left[\cos(\theta_1-\theta_2) + i\sin(\theta_1-\theta_2)\right] \end{aligned}$$

La dernière égalité vient ici aussi des formules de duplication, on tire le module et l'argument de $\frac{z_1}{z_2}$.

Remarques :

- Si on considère le point $M(z_1)$ et le complexe z_2 de module $r = |z_2|$ et d'argument $\theta = \arg(z_2)$, le complexe $z = z_1 z_2$ est l'affixe de l'image de M par une rotation d'angle θ suivie d'une homothétie de centre O et de rapport r.
  Sur l'illustration de la propriété 5.8, le point $M(2+3i)$ est envoyé sur le point M'' par la rotation d'angle 30° suivi d'une homothétie de centre O et de rapport 2.
  Pour déterminer l'affixe z de M'', il suffit de calculer le produit :
  $z = (2+3i)\left(2(\cos\frac{\pi}{6} + i\sin\frac{\pi}{6})\right) = 2(2+3i)(\frac{\sqrt{3}}{2} + \frac{i}{2}) = 2\sqrt{3} - 3 + (2+3\sqrt{3})i \approx 0,4641 + 7,196i$.
- Pour deux vecteurs $\vec{v_1}(z_1)$ et $\vec{v_2}(z_2)$, la division des affixes $\frac{z_1}{z_2}$ a pour argument l'angle $(\vec{v_2}, \vec{v_1})$ et, en notant $z_1 = a_1 + b_1 i$ et $z_2 = a_2 + b_2 i$, vaut $\frac{(a_1a_2 + b_1b_2) + (b_1a_2 - a_1b_2)i}{a_2^2 + b_2^2}$.
  La partie réelle du numérateur correspond au produit scalaire des vecteurs $\vec{v_1} \cdot \vec{v_2}$ tandis que la partie imaginaire en est le déterminant $\det(\vec{v_1}, \vec{v_2})$.
  Autrement dit, $\mathcal{Re}(\frac{z_1}{z_2}) = 0 \iff \vec{v_1} \perp \vec{v_2}$ et $\mathcal{Im}(\frac{z_1}{z_2}) = 0 \iff \exists k \in \mathbb{R}^*, \vec{v_1} = k\vec{v_2}$.
- On peut redémontrer la partie « argument » de la propriété 5.7 en écrivant :
  — $z\bar{z} = |z|^2 \in \mathbb{R}^+$ donc $\arg(z\bar{z}) = 0 \iff \arg z + \arg\bar{z} = 0 \iff \arg\bar{z} = -\arg z \quad [2\pi]$.
  — $-z = (-1) \times z$ donc $\arg -z = \arg((-1) \times z) = \arg(-1) + \arg z = \pi + \arg z \quad [2\pi]$.

Propriété 5.10 (Inégalité triangulaire) Soient z_1 et z_2 des complexes quelconques : $|z_1+z_2| \leqslant |z_1|+|z_2|$ avec égalité pour $z_1=0$ ou $z_2=0$ ou $(z_1\neq 0, z_2\neq 0, \exists k \in \mathbb{R}^{+*}, z_2 = kz_1)$

Démonstration

$$\begin{aligned}
|z_1+z_2|^2 &= (z_1+z_2)\overline{(z_1+z_2)} && \text{(d'après la propriété 5.1.c)}\\
&= (z_1+z_2)(\overline{z_1}+\overline{z_2}) && \text{(d'après la propriété 5.5.a)}\\
&= z_1\overline{z_1}+z_2\overline{z_2}+z_1\overline{z_2}+\overline{z_1}z_2 && \text{(d'après la propriété 5.4)}\\
&= |z_1|^2+|z_2|^2+z_1\overline{z_2}+\overline{z_1\overline{z_2}} && \text{(d'après la propriété 5.1.d)}\\
&= |z_1|^2+|z_2|^2+2\mathcal{R}e\,(z_1\overline{z_2}) && \text{(d'après la propriété 5.1.a)}
\end{aligned}$$

Or $\forall z \in \mathbb{C}, |\mathcal{R}e(z)| \leqslant |z|$ et à fortiori $\mathcal{R}e(z) \leqslant |z|$, donc $\mathcal{R}e\,(z_1\overline{z_2}) \leqslant |z_1\overline{z_2}|$.
De plus $|z_1\overline{z_2}| = |z_1| \times |\overline{z_2}| = |z_1| \times |z_2|$. On en déduit $2\mathcal{R}e\,(z_1\overline{z_2}) \leqslant 2|z_1||z_2|$, et finalement $|z_1+z_2|^2 \leqslant |z_1|^2+|z_2|^2+2|z_1||z_2| \iff |z_1+z_2|^2 \leqslant (|z_1|+|z_2|)^2$.
La fonction racine carrée étant croissante sur $\mathbb{R}^+$, on obtient l'inégalité cherchée.

L'égalité se produit lorsque $\mathcal{R}e\,(z_1\overline{z_2}) = |z_1||z_2|$.
Or, si $\arg(z_1)=\theta_1$ et $\arg(z_2)=\theta_2$, on a $\arg(\overline{z_2}) = -\theta_2$ (d'après la propriété 5.7) et donc $\mathcal{R}e\,(z_1\overline{z_2}) = |z_1||z_2|\cos(\theta_1-\theta_2)$. L'égalité est possible à condition que :

$$\begin{cases} |z_1| = 0 \iff z_1 = 0 \text{ ou,}\\ |z_2| = 0 \iff z_2 = 0 \text{ ou,}\\ \cos(\theta_1-\theta_2) = 1 \iff \theta_1-\theta_2 = 0 \ \ [2\pi] \iff \theta_1 = \theta_2 \ \ [2\pi]. \end{cases}$$

Dans la dernière possibilité, les deux vecteurs ont alors même direction et même sens, c'est-à-dire que $\exists k \in \mathbb{R}^{+*}, \overrightarrow{v_1} = k\overrightarrow{v_2}$.

Propriété 5.11 (Nombres de module 1) L'ensemble des nombres complexes de module 1, noté $\mathbb{U}$, est stable pour la multiplication et le passage à l'inverse.
Autrement dit : $\forall (z,z') \in \mathbb{U}^2, \forall n \in \mathbb{N}, zz' \in \mathbb{U}, \frac{1}{z} \in \mathbb{U}, \frac{z}{z'} \in \mathbb{U}$ et $z^n \in \mathbb{U}$

Démonstration Par définition $z \in \mathbb{U} \iff |z| = 1$ et $z' \in \mathbb{U} \iff |z'| = 1$.
Comme $|zz'| = |z| \times |z'|$ d'après la propriété 5.9, on en déduit $|zz'| = 1 \iff zz' \in \mathbb{U}$.
De même, comme $|\dfrac{1}{z}| = \dfrac{1}{|z|}$, $|z| = 1 \iff |\dfrac{1}{z}| = 1 \iff \dfrac{1}{z} \in \mathbb{U}$
et comme $|\dfrac{z}{z'}| = \dfrac{|z|}{|z'|}$, $|z| = 1$ et $|z'| = 1 \iff |\dfrac{z}{z'}| = 1 \iff \dfrac{z}{z'} \in \mathbb{U}$.
Pour montrer que $\forall n \in \mathbb{N}, z^n \in \mathbb{U}$, il suffit bien sûr de raisonner par récurrence sur n.

Remarques :

- L'ensemble $\mathbb{U}$, doté de la multiplication est un groupe commutatif et un sous-groupe de $(\mathbb{C}^*, \times)$. Il correspond sur le plan complexe au cercle trigonométrique car si $z \in \mathbb{U}$ est l'affixe d'un point M alors $OM = 1$.
  L'écriture trigonométrique d'un nombre quelconque de $\mathbb{U}$ est $\cos\theta + i\sin\theta$.
- Les nombres complexes 1, i, $-i$, -1, j et $j^2 = \overline{j}$ font partie de $\mathbb{U}$ ainsi que les produits $i^n \times j^m$ quels que soient les entiers n et m. Calculons l'argument de ce produit avec la propriété 5.9 :
  $\arg(i^n) \times \arg(j^m) = n\arg(i) + m\arg(j) = \dfrac{n\pi}{2} + \dfrac{2m\pi}{3} = \dfrac{(3n+4m)\pi}{6} \quad [2\pi]$
- On peut facilement déterminer des lignes trigonométriques dans $\mathbb{U}$. Exemple : $\frac{\pi}{3} - \frac{\pi}{4} = \frac{\pi}{12}$,
  si $z = (\cos\frac{\pi}{3} + i\sin\frac{\pi}{3})(\cos\frac{-\pi}{4} + i\sin\frac{-\pi}{4}) = \left(\frac{1+i\sqrt{3}}{2}\right)\left(\frac{\sqrt{2}-i\sqrt{2}}{2}\right) = \frac{\sqrt{6}+\sqrt{2}}{4} + i\frac{\sqrt{6}-\sqrt{2}}{4}$,
  comme $\arg(z) = \frac{\pi}{12}$, on en déduit $\cos\frac{\pi}{12} = \frac{\sqrt{6}+\sqrt{2}}{4}$ et $\sin\frac{\pi}{12} = \frac{\sqrt{6}-\sqrt{2}}{4}$.

3. Trigonométrie avec les complexes

PROPRIÉTÉ 5.12 (FORMULE DE MOIVRE) $\forall\theta \in \mathbb{R}, \forall n \in \mathbb{Z}, (\cos\theta + i\sin\theta)^n = \cos(n\theta) + i\sin(n\theta)$

DÉMONSTRATION Distinguons deux cas :
➠Cas $n \geqslant 0$:
C'est encore une conséquence de la propriété 5.9 $\arg(z^n) = n \arg z\,[2\pi]$ avec ici $\arg z = \theta$.

➠Cas $n < 0$:
On applique la propriété en remarquant que $-n > 0$:

$$\begin{aligned}
(\cos\theta + i\sin\theta)^n &= \frac{1}{(\cos\theta + i\sin\theta)^{-n}} \\
&= \frac{1}{\cos(-n\theta) + i\sin(-n\theta)} \\
&= \frac{1}{\cos(n\theta) - i\sin(n\theta)} \\
&= \frac{\cos(n\theta) + i\sin(n\theta))}{(\cos(n\theta) - i\sin(n\theta))(\cos(n\theta) + i\sin(n\theta))} \\
&= \frac{\cos(n\theta) + i\sin(n\theta))}{\cos^2(n\theta) + \sin^2(n\theta)} \\
&= \cos(n\theta) + i\sin(n\theta)
\end{aligned}$$

Remarques :

- Une version soustractive de cette propriété s'obtient en prenant $\theta' = -\theta$:
  $(\cos\theta' + i\sin\theta')^n = \cos(n\theta') + i\sin(n\theta') \iff (\cos\theta - i\sin\theta)^n = \cos(n\theta) - i\sin(n\theta)$.
- Pour $n = 2$, on retrouve les formules de duplication étudiées en 1^{re} :
  $(\cos\theta + i\sin\theta)^2 = \cos(2\theta) + i\sin(2\theta) \iff \cos^2\theta - \sin^2\theta + 2i\cos\theta\sin\theta = \cos(2\theta) + i\sin(2\theta)$
  Par identification des parties réelles et imaginaires de cette égalité, on obtient
  $\cos(2\theta) = \cos^2\theta - \sin^2\theta$ et $\sin(2\theta) = 2\cos\theta\sin\theta$. On peut faire de même pour obtenir des formules similaires pour $n = 3$, $n = 4$, etc.

DÉFINITION 5.5 (NOTATION EXPONENTIELLE) Tout nombre complexe non nul z peut être noté sous la forme exponentielle $z = re^{i\theta}$ où $r = |z|$ et $\theta = \arg(z)\ \ [2\pi]$.
Autrement dit, $\forall(r,\theta) \in \mathbb{R}^2, re^{i\theta} = r(\cos\theta + i\sin\theta)$.

Remarques :

- Pour les complexes de module 1, la multiplication satisfait la relation fonctionnelle
  $f(x) \times f(y) = f(x+y)$ car $(\cos\theta + i\sin\theta) \times (\cos\theta' + i\sin\theta') = \cos(\theta+\theta') + i\sin(\theta+\theta')$.
  Or cette même relation fonctionnelle est satisfaite par la fonction exponentielle, puisque $\forall(a,b) \in \mathbb{R}^2, e^a e^b = e^{a+b}$.
  De plus, $\theta = 0\,[2\pi] \Rightarrow \cos\theta + i\sin\theta = 1$ et de même $a = 0 \Rightarrow, e^a = 1$.
  Du fait de cette analogie, Euler proposa en 1748 de noter un élément de $\mathbb{U}$ sous la forme $z = e^{i\theta}$ avec $\theta = \arg(z)\,[2\pi]$. Par extension, tout complexe de $\mathbb{C}$ est noté $z = re^{i\theta}$ avec $r = |z|$.
- Un exemple est fameux : $e^{i\pi} = \cos\pi + i\sin\pi = -1 \iff e^{i\pi} + 1 = 0$.
  Qualifiée de « plus belle formule des mathématiques », cette égalité rassemble cinq constantes fondamentales : les entiers 0 et 1 qui sont les éléments neutres de l'addition $(x + 0 = 0)$ et de la multiplication $(x \times 1 = x)$, les nombres transcendants π et e qui interviennent partout en mathématiques et le nombre imaginaire i qui est à l'origine du déploiement des complexes.
- Autres exemples : $e^{i\frac{\pi}{2}} = \cos\frac{\pi}{2} + i\sin\frac{\pi}{2} = i$ et $e^{0i} = \cos 0 + i\sin 0 = 1$.

- Pour déterminer la forme exponentielle d'un complexe, comme pour la forme trigonométrique, on commence par déterminer le module et ensuite on identifie l'argument.
  Par exemple $z = 1 - i$ a pour module $\sqrt{1^2 + (-1)^2} = \sqrt{2}$, du coup $z = \sqrt{2}\left(\frac{\sqrt{2}}{2} + i\frac{-\sqrt{2}}{2}\right)$ et l'argument θ de z vérifie $\cos\theta = \frac{\sqrt{2}}{2}$ et $\sin\theta = \frac{-\sqrt{2}}{2}$. On en déduit $\theta = \frac{-\pi}{4}$ $[2\pi]$ et $z = \sqrt{2}e^{\frac{-i\pi}{4}}$.

PROPRIÉTÉ 5.13 (FORMULES DE EULER) Pour tous complexes $z = re^{i\theta}$ et $z' = re^{i\theta'}$:

$zz' = re^{i\theta} \times r'e^{i\theta'} = rr'e^{i(\theta+\theta')}$	$\frac{z}{z'} = \frac{re^{i\theta}}{r'e^{i\theta'}} = \frac{r}{r'}e^{i(\theta-\theta')}$	$\forall n \in \mathbb{Z}, z^n = (re^{i\theta})^n = r^n e^{ni\theta}$
$\overline{z} = r \times \left(\overline{e^{i\theta}}\right) = re^{-i\theta} = \frac{r}{e^{i\theta}}$	$\cos\theta = \frac{e^{i\theta} + e^{-i\theta}}{2}$	$\sin\theta = \frac{e^{i\theta} - e^{-i\theta}}{2i}$

DÉMONSTRATION ➠ Les premières propriétés ne font que reécrire sous forme exponentielle des propriétés déjà démontrées : propriété 5.7 pour le conjugué et propriété 5.9 pour le produit, le quotient et la puissance (on reconnait dans cette dernière la formule de Moivre).

➠ Formules de Euler :
On sait que $2\mathcal{R}e(z) = z + \bar{z}$, mais comme $\mathcal{R}e(re^{i\theta}) = r\cos\theta$, on en déduit :
$r\cos\theta = \frac{z+\bar{z}}{2} = \frac{re^{i\theta} + re^{-i\theta}}{2} = r\frac{e^{i\theta}+e^{-i\theta}}{2}$ d'où $\cos\theta = \frac{e^{i\theta} + e^{-i\theta}}{2}$.
De même $2i\mathcal{I}m(z) = z - \bar{z}$, mais comme $\mathcal{I}m(re^{i\theta}) = r\sin\theta$, on en déduit :
$r\sin\theta = \frac{re^{i\theta} - re^{-i\theta}}{2i} = \frac{r(2i\sin\theta)}{2i}$ d'où $\sin\theta = \frac{e^{i\theta} - e^{-i\theta}}{2i}$.

Remarques :

- Les formules de Euler permettent de (re)trouver des formules de trigonométrie :
  $2i\sin 2\theta = e^{i2\theta} - e^{-i2\theta} = (e^{i\theta})^2 - (e^{-i\theta})^2 = (e^{i\theta} + e^{-i\theta})(e^{i\theta} - e^{-i\theta}) = 4i\cos\theta\sin\theta$ d'où, en divisant par $2i$, on retrouve la formule de duplication du sinus : $\sin 2\theta = 2\cos\theta\sin\theta$.
- Comme $i = e^{i\frac{\pi}{2}}$, l'élévation à la puissance i de i est le réel $i^i = (e^{i\frac{\pi}{2}})^i = e^{i^2\frac{\pi}{2}} = e^{-\frac{\pi}{2}} \approx 0,2079$.

PROPRIÉTÉ 5.14 (HUIT FORMULES) Des formules de Euler se déduisent les suivantes :

$2\cos a\cos b = \cos(a+b) + \cos(a-b)$	$2\sin a\sin b = \cos(a-b) - \cos(a+b)$
$2\sin a\cos b = \sin(a+b) + \sin(a-b)$	$\cos a + \cos b = 2\cos\frac{a+b}{2}\cos\frac{a-b}{2}$
$\sin a + \sin b = 2\sin\frac{a+b}{2}\cos\frac{a-b}{2}$	$\cos a - \cos b = -2\sin\frac{a+b}{2}\sin\frac{a-b}{2}$
$\sin a - \sin b = 2\cos\frac{a+b}{2}\sin\frac{a-b}{2}$	$x+iy = re^{i\theta} \Longrightarrow x\cos a + y\sin a = r\cos(a-\theta)$

DÉMONSTRATION ➠Montrons que $2\cos a\cos b = \cos(a+b) + \cos(a-b)$:

$$\begin{aligned}
\cos a\cos b &= \left(\frac{e^{ia} + e^{-ia}}{2}\right)\left(\frac{e^{ib} + e^{-ib}}{2}\right) \\
&= \frac{e^{ia+ib} + e^{-ia+ib} + e^{ia-ib} + e^{-ia-ib}}{4} \\
&= \frac{e^{i(a+b)} + e^{-i(a+b)}}{4} + \frac{e^{i(a-b)} + e^{-i(a-b)}}{4} \\
&= \frac{\cos(a+b)}{2} + \frac{\cos(a-b)}{2}
\end{aligned}$$

De la même façon, on peut établir que :

- $2\sin a\sin b = \cos(a-b) - \cos(a+b)$
- $2\sin a\cos b = \sin(a+b) + \sin(a-b)$

NB : on montre également ces formules à partir des formules d'addition (vues en 1re).

➠ Établissons un autre type de formule trigonométrique en remarquant que :

$e^{ia}+e^{ib}=e^{i\left(\frac{a+b}{2}+\frac{a-b}{2}\right)}+e^{i\left(\frac{a+b}{2}+\frac{-(a-b)}{2}\right)}=e^{i\frac{a+b}{2}}\left(e^{i\frac{a-b}{2}}+e^{i\frac{-(a-b)}{2}}\right)=2e^{i\frac{a+b}{2}}\cos\frac{a-b}{2}$ et, également
$e^{ia}-e^{ib}=e^{i\left(\frac{a+b}{2}+\frac{a-b}{2}\right)}-e^{i\left(\frac{a+b}{2}+\frac{-(a-b)}{2}\right)}=e^{i\frac{a+b}{2}}\left(e^{i\frac{a-b}{2}}-e^{i\frac{-(a-b)}{2}}\right)=2ie^{i\frac{a+b}{2}}\sin\frac{a-b}{2}$.
Séparons maintenant les parties réelles et imaginaires des membres extrêmes de ces égalités :

- $\cos a+\cos b=2\cos\frac{a+b}{2}\cos\frac{a-b}{2}$ et $\sin a+\sin b=2\sin\frac{a+b}{2}\cos\frac{a-b}{2}$
- $\cos a-\cos b=-2\sin\frac{a+b}{2}\sin\frac{a-b}{2}$ et $\sin a-\sin b=2\cos\frac{a+b}{2}\sin\frac{a-b}{2}$

➠ Transformons une écriture de la forme $x\cos a+y\sin a$ pour tout couple $(x,y)\neq(0,0)$ en remarquant que cette expression est la partie réelle du produit $e^{ia}\,(\overline{x+iy})$.
Supposons que le module et l'argument de $x+iy$ soient r et θ (autrement dit $x+iy=re^{i\theta}$), comme $\mathcal{R}e\left(e^{ia}\left(re^{-i\theta}\right)\right)=\mathcal{R}e\left(re^{i(a-\theta)}\right)=r\cos(a-\theta)$, on en déduit que $x\cos a+y\sin a=r\cos(a-\theta)$.

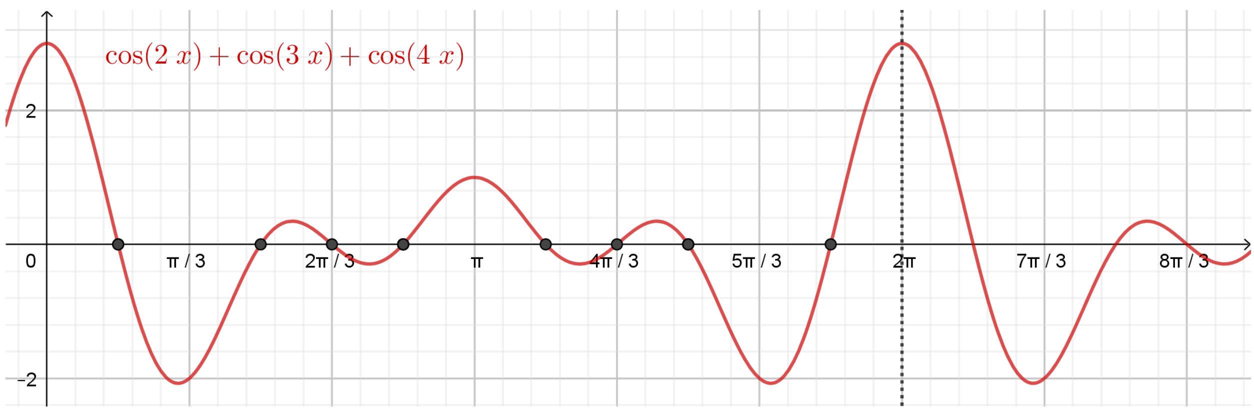
Exemple 49 – Résolvons l'équation $\sqrt{3}\cos a+\sin a=\sqrt{2}$.
Écrivons pour commencer le complexe $z=\sqrt{3}+i$ sous forme trigonométrique :
z a pour module $\sqrt{(\sqrt{3})^2+1^2}=\sqrt{4}=2$, du coup $z=2\left(\frac{\sqrt{3}}{2}+i\frac{1}{2}\right)$ et l'argument θ de z vérifie $\cos\theta=\frac{\sqrt{3}}{2}$ et $\sin\theta=\frac{1}{2}$. On en déduit que $\theta=\frac{\pi}{6}\ [2\pi]$ et $z=2e^{\frac{i\pi}{6}}$.
L'équation s'écrit alors $2\cos\left(x-\frac{\pi}{6}\right)=\sqrt{2}\iff\cos\left(x-\frac{\pi}{6}\right)=\frac{\sqrt{2}}{2}=\cos\frac{\pi}{4}$.
Les solutions sont donc :

- $x-\frac{\pi}{6}=\frac{\pi}{4}\ [2\pi]\iff x=\frac{\pi}{6}+\frac{\pi}{4}=\frac{5\pi}{12}\ [2\pi]$
- $x-\frac{\pi}{6}=-\frac{\pi}{4}\ [2\pi]\iff x=\frac{\pi}{6}-\frac{\pi}{4}=\frac{-\pi}{12}\ [2\pi]$

Exemple 50 – Résolvons l'équation $\cos 2x+\cos 3x+\cos 4x=0$.
Remplaçons $\cos 2x+\cos 4x$ par $2\cos\frac{2x+4x}{2}\cos\frac{2x-4x}{2}=2\cos 3x\cos(-x)=2\cos 3x\cos x$,
L'équation s'écrit $2\cos 3x\cos x+\cos 3x=0\iff\cos 3x(1+2\cos x)=0$.
On en déduit les solutions :

- $\cos 3x=0\iff\cos 3x=\cos\frac{\pi}{2}\iff 3x=\pm\frac{\pi}{2}\ [2\pi]\iff x=\pm\frac{\pi}{6}\ \left[\frac{2\pi}{3}\right]$
- $1+2\cos x=0\iff\cos x=\frac{-1}{2}=\cos\frac{2\pi}{3}\iff x=\pm\frac{2\pi}{3}\ [2\pi]$

Dans un intervalle d'amplitude 2π, il y a huit solutions distinctes.
Voici celles de l'intervalle $[0,2\pi]$: $\frac{\pi}{6}$, $\frac{5\pi}{6}$, $\frac{9\pi}{6}$, $\frac{3\pi}{6}=\frac{\pi}{2}$, $\frac{7\pi}{6}$, $\frac{11\pi}{6}$, $\frac{2\pi}{3}$ et $\frac{4\pi}{3}$.
Les huit solutions sont dans l'ordre croissant : $\frac{\pi}{6}$, $\frac{\pi}{2}$, $\frac{2\pi}{3}$, $\frac{5\pi}{6}$, $\frac{7\pi}{6}$, $\frac{4\pi}{3}$, $\frac{9\pi}{6}$ et $\frac{11\pi}{6}$.
On peut les retrouver graphiquement à l'intersection de la courbe d'équation $y=\cos 2x+\cos 3x+\cos 4x$ et de l'axe des abscisses.



Exemple 51 – Les égalités $e^{ia}+e^{ib}=2e^{i\frac{a+b}{2}}\cos\frac{a-b}{2}$ et $e^{ia}-e^{ib}=2ie^{i\frac{a+b}{2}}\sin\frac{a-b}{2}$ vues plus haut, dans le cas où $e^{ia}=1\iff a=0\ [2\pi]$ et $b=x$ s'écrivent :

- $1+e^{ix}=2e^{i\frac{x}{2}}\cos\frac{x}{2}$
- $1-e^{ix}=2ie^{i\frac{x}{2}}\sin\frac{-x}{2}=2e^{i\frac{\pi-x}{2}}\sin\frac{-x}{2}=2e^{i\frac{x-\pi}{2}}\sin\frac{x}{2}$

Application : calculons $z_n=(1+e^{ix})^n$.
Comme on sait transformer la somme $1+e^{ix}$ en produit $2e^{i\frac{x}{2}}\cos\frac{x}{2}$, on en déduit :
$z_n=\left(2e^{i\frac{x}{2}}\cos\frac{x}{2}\right)^n=2^ne^{i\frac{nx}{2}}\cos^n\frac{x}{2}$ d'où $\mathcal{R}e(z_n)=2^n\cos\frac{nx}{2}\cos^n\left(\frac{x}{2}\right)$ et $\mathcal{I}m(z_n)=2^n\sin\frac{nx}{2}\cos^n\left(\frac{x}{2}\right)$.

4. Polynômes complexes

Du point de vue algébrique, l'ensemble des complexes est appelé par les mathématiciens la « clôture algébrique » du corps des réels. On entend par là que cet ensemble, noté $\mathbb{C}$, prolonge l'ensemble $\mathbb{R}$ en offrant des solutions à toutes les équations polynomiales dont les coefficients sont des réels. Du même coup, les équations polynomiales à coefficients complexes ont également leurs solutions dans $\mathbb{C}$.

4.a. Second degré

Les équations du second degré à coefficients réels ont été vues au début de ce chapitre (voir l'introduction de la définition 5.1 ainsi que la remarque qui la suit).

4.a.1. Racine carrée d'un complexe

Propriété 5.15 (Racines carrées) Pour tout couple de réels $(a, b) \neq (0, 0)$, l'équation $z^2 = a + ib$ admet deux solutions complexes opposées, appelées « racines carrées » de $a + ib$.

Démonstration Notons $x + iy$ un complexe satisfaisant l'équation (E) $z^2 = a + ib$.
Comme $(x + iy)^2 = x^2 - y^2 + 2ixy$, par identification des parties réelles et imaginaires de (E) on obtient le système $\begin{cases} x^2 - y^2 = a \\ 2xy = b \end{cases}$.
L'identité des modules $|z^2| = |a + ib| \iff |z|^2 = |a + ib| \iff x^2 + y^2 = \sqrt{a^2 + b^2}$ ajoute à ce système une 3e égalité qui permet, par addition et soustraction avec la 1re du système, d'obtenir le nouveau système $\begin{cases} x^2 = \dfrac{\sqrt{a^2 + b^2} + a}{2} \\ y^2 = \dfrac{\sqrt{a^2 + b^2} - a}{2} \\ xy = \dfrac{b}{2} \end{cases}$
Comme, d'après la propriété 5.6, $\forall z \in |\mathcal{R}e(z)| \leqslant |z|$, on a $|a| \leqslant \sqrt{a^2 + b^2}$, et comme $\forall a \in \mathbb{R}, a \leqslant |a|$, on en déduit $\sqrt{a^2 + b^2} - a \geqslant 0$ et donc $y^2 \geqslant 0$. Cela assure l'existence des réels $\pm y$.
D'autre part, si $a \geqslant 0$, alors $\sqrt{a^2 + b^2} + a \geqslant 0$ et si $a < 0$, comme $|a| = -a \leqslant \sqrt{a^2 + b^2}$ (d'après la prop. 5.6) alors $\sqrt{a^2 + b^2} + a \geqslant 0$. Dans tous les cas on a donc $\sqrt{a^2 + b^2} + a \geqslant 0$ et $x^2 \geqslant 0$. Cela assure l'existence des réels $\pm x$.
Il suffit de choisir les réels qui conviennent en fonction du signe de b qui est celui du produit xy :
- Si $b \geqslant 0$, on prend x et y de mêmes signes
- Si $b < 0$, on prend x et y de signes opposés

Exemple 52 – Quelles sont les racines carrées de $4 - 3i$?
Pour répondre, il faut résoudre l'équation $z^2 = 4 - 3i$.
Posons $z = x + iy$ et résolvons le système $\begin{cases} x^2 - y^2 = 4 \\ 2xy = -3 \end{cases}$ avec l'égalité complémentaire des modules $x^2 + y^2 = \sqrt{4^2 + (-3)^2} = \sqrt{16 + 9} = \sqrt{25} = 5$.

Le système à résoudre $\begin{cases} x^2 - y^2 = 4 \\ x^2 + y^2 = 5 \\ xy = \dfrac{-3}{2} \end{cases} \iff \begin{cases} 2x^2 = 4 + 5 = 9 \\ 2y^2 = 5 - 4 = 1 \\ xy = \dfrac{-3}{2} \end{cases} \implies \begin{cases} x^2 = \dfrac{9}{2} \\ y^2 = \dfrac{1}{2} \\ xy < 0 \end{cases}$

conduit aux deux solutions $\begin{cases} (x = \dfrac{3}{\sqrt{2}}, y = -\dfrac{1}{\sqrt{2}}), \text{ soit } z = \dfrac{3 - i}{\sqrt{2}} \\ (x = -\dfrac{3}{\sqrt{2}}, y = \dfrac{1}{\sqrt{2}}), \text{ soit } z = \dfrac{-3 + i}{\sqrt{2}} \end{cases}$

Propriété 5.16 (Coordonnées polaires) Pour tout couple de réels (r, θ) avec $r > 0$, l'équation $z^2 = re^{i\theta}$ admet deux solutions complexes opposées telles que $\begin{cases} |z| = \sqrt{r} \\ \arg(z) = \dfrac{\theta}{2} \ [\pi] \end{cases}$

Démonstration D'après la propriété 5.9, $z^2 = \left(\arg(z)e^{i\arg(z)}\right)^2 = \arg(z)^2 \times e^{i2\arg(z)}$, mais comme $z^2 = re^{i\theta}$, l'égalité des modules égaux donne $|z^2| = r \iff |z|^2 = r \iff |z| = \sqrt{r}$ tandis que l'égalité des arguments donne $2\arg(z) = \theta \ [2\pi] \iff \arg(z) = \frac{\theta}{2} \ [\pi]$.

Exemple 53 – Quelles sont les racines carrées de $2e^{i\frac{2\pi}{5}}$?

Pour répondre, il faut résoudre le système $\begin{cases} |z|^2 = 2 \\ 2\arg(z) = \dfrac{2\pi}{5} \ [2\pi] \end{cases} \iff \begin{cases} |z| = \sqrt{2} \\ \arg(z) = \dfrac{\pi}{5} \ [\pi] \end{cases}$.

On en déduit que $z = \sqrt{2}e^{i(\frac{\pi}{5}+k\pi)}$ avec $k \in \mathbb{Z}$, soit $z = \sqrt{2}e^{i\frac{\pi}{5}} \times e^{ik\pi}$.

- Pour $k = 2k'$ (k pair), $z = \sqrt{2}e^{i\frac{\pi}{5}} \times e^{i2k'\pi} = \sqrt{2}e^{i\frac{\pi}{5}}$
- Pour $k = 2k'+1$ (k impair), $z = \sqrt{2}e^{i\frac{\pi}{5}} \times e^{i(2k'+1)\pi} = \sqrt{2}e^{i\frac{\pi}{5}} \times e^{i\pi} = -\sqrt{2}e^{i\frac{\pi}{5}}$

4.a.2. Équation à coefficients complexes

Propriété 5.17 (Coefficients complexes) Pour tout triplet de complexes $(a \neq 0, b, c)$, l'équation $az^2 + bz + c = 0$ admet deux solutions complexes qui sont, en notant δ une des racines carrées du complexe $\Delta = b^2 - 4ac$, $z = \dfrac{-b \pm \delta}{2a}$.

Démonstration Cela se montre comme dans $\mathbb{R}$ en factorisant la forme canonique du trinôme. La seule différence est qu'il n'y a pas à discuter selon le signe de Δ puisque c'est un complexe et un complexe n'a pas de signe mais a toujours deux racines opposées (éventuellement nulles si $\Delta = 0$), comme on vient de le voir.

Exemple 54 – Résolvons l'équation $iz^2 + (1+3i)z + \frac{11+4i}{2} = 0$

On a $\Delta = (1+3i)^2 - 4i\frac{11+4i}{2} = 1 + 6i - 9 - 22i + 8 = -16i$.

Comme $-16i = 16i^3$ et $i = e^{\frac{i\pi}{2}}$, on a $\Delta = 16\left(e^{\frac{i\pi}{2}}\right)^3 = 16e^{\frac{3i\pi}{2}}$.

On cherche une racine carrée δ telle que $\delta^2 = 16e^{\frac{3i\pi}{2}}$, son module est $|\delta| = \sqrt{16} = 4$ et son argument est $\arg(\delta) = \frac{3\pi}{4} \ [\pi]$. On peut donc prendre, en choisissant l'argument $\frac{3\pi}{4}$,

$\delta = 4e^{i\frac{3\pi}{4}} = 4(\frac{-\sqrt{2}}{2} + i\frac{\sqrt{2}}{2}) = 2\sqrt{2}(-1+i)$.

Les deux solutions de l'équation sont :

$$z = \frac{-1 - 3i \pm 2\sqrt{2}(-1+i)}{2i} = \frac{-i(-1-3i \pm 2\sqrt{2}(-1+i))}{2} = \frac{i - 3 \pm \sqrt{2}(-2i-2))}{2}.$$

On trouve donc $z = \dfrac{-3}{2} - \sqrt{2} + i\left(\dfrac{1}{2} - \sqrt{2}\right)$ ou $z = \dfrac{-3}{2} + \sqrt{2} + i\left(\dfrac{1}{2} + \sqrt{2}\right)$

4.b. Factorisation d'un polynôme

La définition d'un polynôme à coefficients complexes, la factorisation de $z^n - a^n$, l'unicité de l'écriture polynômiale et la méthode d'identification des coefficients ainsi que la factorisation du polynôme $P(z)$ par $z - a$ lorsque a en est une racine, toute cette partie est reprise du cours de 1re. Les démonstrations ne seront pas refaites ici, seules les définitions et propriétés seront rappelées, en les étendant à $\mathbb{C}$.

Définition 5.6 (Polynôme) On appelle polynôme de degré n toute fonction P définie sur $\mathbb{C}$ par $P(z) = a_n z^n + a_{n-1} z^{n-1} + \cdots + a_3 z^3 + a_2 z^2 + a_1 z + a_0$, où les nombres $a_k, k \in \{0, 1, \ldots, n\}$ sont des complexes tels que $a_n \neq 0$. Les nombres a_k sont les coefficients d'ordre k du polynôme P et on note le degré $\deg(P) = n$.

Propriété 5.18 (polynômes égaux) Voici trois propositions vraies et équivalentes :

(i) Deux polynômes complexes P et Q sont égaux si et seulement si ils ont même degré n et que les coefficients d'ordre $k \leqslant n$ sont égaux deux à deux.

(ii) La fonction $f : z \mapsto a_n z^n + a_{n-1} z^{n-1} + \cdots + a_3 z^3 + a_2 z^2 + a_1 z + a_0$ est la fonction nulle si et seulement si les coefficients $a_k, k \leqslant n$ sont tous nuls.

(iii) L'écriture du polynôme complexe $a_n z^n + a_{n-1} z^{n-1} + \cdots + a_3 z^3 + a_2 z^2 + a_1 z + a_0$ est unique.

Propriété 5.19 (Degré) Soient P et Q deux polynômes complexes et λ un complexe non nul :

- Le produit λP, défini par $\lambda \times P(x)$, est un polynôme de degré égal à celui de P.
- Le produit PQ, défini par $P(x) \times Q(x)$, est un polynôme de degré égal à la somme des degrés de P et de Q, soit $\deg(PQ) = \deg(P) + \deg(Q)$.
- La somme $P + Q$, définie par $P(x)+Q(x)$, est un polynôme de degré inférieur ou égal au plus grand des degrés de P et de Q, soit $\deg(P + Q) \leqslant \max(\deg(P), \deg(Q))$.

Définition 5.7 (racine) Soit P un polynôme complexe et a un complexe.
Le nombre a est une racine de P (on dit aussi un « zéro » de P) si et seulement si $P(a) = 0$.

Propriété 5.20 (factorisation par $(z - a)$) Un polynôme complexe P admet le complexe a comme racine si et seulement si on peut écrire $P(z)$ sous la forme $(z - a)Q(z)$, où Q est un polynôme complexe de degré égal à $\deg(P) - 1$.

Factorisation de $z^n - a^n$:

$$z^n - a^n = (z - a)(z^{n-1} + az^{n-2} + a^2 z^{n-3} + \cdots + a^{n-2} z + a^{n-1}) = (z - a) \sum_{k=0}^{n-1} a^k z^{n-k-1}$$

En particulier :

$$z^n - 1 = (z - 1)(z^{n-1} + z^{n-2} + z^{n-3} + \cdots + z + 1) = (z - 1) \sum_{k=0}^{n-1} z^k$$

Autre exemple :

$$\begin{aligned} z^n - i^{3n} &= z^n - (i^3)^n = z^n - (-i)^n \\ &= (z + i)(z^{n-1} - iz^{n-2} + (-i)^2 z^{n-3} + \cdots + (-i)^{n-2} z + (-i)^{n-1}) \\ &= (z + i) \sum_{k=0}^{n-1} (-i)^k z^{n-k-1} \end{aligned}$$

Exemple 55 – La méthode des « coefficients indéterminés » utilisée largement en 1re reste valable et particulièrement utile pour factoriser un polynôme complexe du moment qu'on peut en identifier certaine(s) racine(s).

Résolvons l'équation du 3e degré $z^3+(i-2)z^2-(1+i)(2z+3)=0$ en commençant par chercher une racine réelle $z_0=x$ (les équation du 3e degré en ont toujours au moins une) ; celle-ci vérifie l'égalité $x^3+(i-2)x^2-(1+i)(2x+3)=0 \iff (x^3-2x^2-2x-3)+i(x^2-2x-3)=0$.

Comme les parties réelles et imaginaires sont nulles, on a le système $\begin{cases} x^3-2x^2-2x-3=0 \\ x^2-2x-3=0 \end{cases}$

L'équation $x^2-2x-3=0$ a pour racine $\frac{2\pm\sqrt{2^2+4\times 3}}{2}=\frac{2\pm\sqrt{16}}{2}=1\pm 2$, soit 3 et -1.
L'autre équation est satisfaite pour $x=3$ car $3^3-2\times 3^2-2\times 3-3=27-18-6-3=0$ mais pas par -1 ($(-1)^3-2\times(-1)^2-2\times(-1)-3=-1-2+2-3=-4\neq 0$), du coup on peut mettre $z-3$ en facteur : $z^3+(i-2)z^2-(1+i)(2z+3)=(z-3)(az^2+bz+c)$ où a, b et c sont à déterminer.
$(z-3)(az^2+bz+c)=az^3+(b-3a)z^2+(c-3b)z+(-3c)$ or ce polynôme est égal à
$z^3+(i-2)z^2-(1+i)(2z+3)=z^3+(i-2)z^2+(-2-2i)z+(-3-3i)$.
Par identification des coefficients de même degré de ces polynômes, on obtient le système

$$\begin{cases} a=1 \\ b-3a=i-2 \\ c-3b=-2-2i \\ -3c=-3-3i \end{cases} \iff \begin{cases} a=1 \\ b=i-2+3=1+i \text{ (la 4e égalité confirme ces valeurs)} \\ c=-2-2i+3(1+i)=1+i \end{cases}$$

Il reste à résoudre l'équation du 2e degré $z^2+(1+i)z+(1+i)=0$.
Le discriminant est $\Delta=(1+i)^2-4(1+i)=2i-4-4i=-2i-4$;
Cherchons la racine $\delta=\alpha+i\beta$ de Δ :

Comme $(\alpha+i\beta)^2=(\alpha^2-\beta^2)+i(2\alpha\beta)$, on a $\begin{cases} \alpha^2-\beta^2=-4 \\ 2\alpha\beta=-2 \end{cases}$

L'égalité des modules ajoute une 3e équation au système : $\alpha^2+\beta^2=(-2)^2+(-4)^2=20$.

Par addition et soustraction des égalités contenant les carrés, cela se réduit à $\begin{cases} \alpha^2=8 \\ \beta^2=12 \\ \alpha\beta=-1 \end{cases}$

α et β sont de signes opposés d'où $\delta=\sqrt{8}-i\sqrt{12}=2(\sqrt{2}-i\sqrt{3})$ ou $\delta=-2(\sqrt{2}-i\sqrt{3})$.
Les trois solutions de l'équation de départ sont finalement :

- $z_0=3$
- $z_1=\frac{-(1+i)+2(\sqrt{2}-i\sqrt{3})}{2}=\frac{-1+2\sqrt{2}+i(-1-2\sqrt{3})}{2}$
- $z_2=\frac{-(1+i)-2(\sqrt{2}-i\sqrt{3})}{2}=\frac{-1-2\sqrt{2}+i(-1+2\sqrt{3})}{2}$.

```
def P(z):
    a=complex(1,0)
    b=complex(-2,1)
    c=complex(-2,-2)
    d=complex(-3,-3)
    return d+mult(z,(c+mult(z,(b+mult(z,a)))))
def mult(z1,z2):
    return complex(z1.real*z2.real-z1.imag*z2.imag,z1.real*z2.imag+z1.imag*z2.real)
while True:
    a=complex(int(input("partie réelle de z = ")),int(input("partie imaginaire de z = ")))
    Pa=P(a)
    if Pa==complex(0,0):
        print("{} est une racine évidente P.".format(a))
        break
    else : print("P({})={}\n".format(a,Pa))
```

```
partie réelle de z = -1
partie imaginaire de z = 0
P((-1+0j))=(-4+0j)

partie réelle de z = 0
partie imaginaire de z = 1
P(1j)=(1-7j)

partie réelle de z = 3
partie imaginaire de z = 0
(3+0j) est une racine évidente P.
```

Pour la recherche d'une racine évidente, il peut être utile de calculer les images des petits nombres entiers 0, 1, -1, 2, -2,. . . ou de complexes simples tels que i, $-i$, j,. . . Le programme très simple

ci-dessus est écrit dans ce but ; il suffit de tester jusqu'à $z = 3$ pour pouvoir factoriser le polynôme.
Le programme demande d'entrer les parties réelles et imaginaires d'un complexe z et détermine si ce nombre est une racine évidente d'un polynôme complexe du 3^e degré dont on a rentré au préalable les coefficients complexes a, b, c et d.
J'ai dû programmer le produit de deux complexes car le module `complex` ne réalise pas cela par défaut et j'ai utilisé pour évaluer le polynôme la forme de Hörner [1] qui limite le nombre de produits à effectuer. On peut facilement développer cette idée en bouclant sur une liste pour modifier les parties réelles et imaginaires du complexe z.

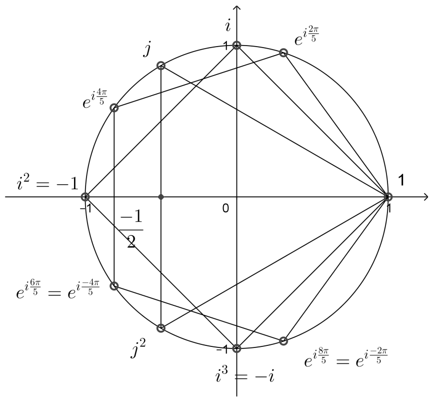
4.c. Racines de l'unité

Propriété 5.21 (racine n-ième de 1) $n > 0$ étant un entier, l'équation $z^n = 1$ admet n solutions distinctes, appelées « racines n^e de l'unité », égales à $e^{i\frac{2k\pi}{n}}$ avec $k \in \{0, 1, \ldots, n-1\}$.

Démonstration Comme $z^n = 1$ alors le module de z est 1 (car $|z^n| = |z|^n = 1$),
ainsi les solutions de l'équation complexe du n^e degré s'écrivent $e^{i\theta}$ où l'argument θ vérifie
$\left(e^{i\theta}\right)^n = 1 \Longleftrightarrow e^{in\theta} = e^{i\times 0} \Longleftrightarrow n\theta = 0\ [2\pi]$.
On en déduit, en divisant par n, que $\theta = 0\ [\frac{2\pi}{n}]$ soit $\theta \in \{0, \frac{1\times 2\pi}{n}, \frac{2\times 2\pi}{n} \ldots, \frac{(n-1)\times 2\pi}{n}\}$.

Remarques :

- Nous connaissons les 2 racines 2^e de l'unité qui sont les nombres opposés 1 et -1.
  Nous connaissons également les 3 racines 3^e de l'unité qui sont 1, j et j^2 ; ce sont les affixes des sommets d'un triangle équilatéral inscrit dans le cercle trigonométrique.
  De même, les 4 racines 4^e de l'unité sont 1, i, $i^2 = -1$ et $i^3 = -i$; ce sont les affixes des sommets d'un carré inscrit dans le cercle trigonométrique ; et d'une façon générale, les n racines n^e de l'unité sont les affixes des sommets d'un polygone régulier à n côtés inscrit dans le cercle trigonométrique (voir ci-dessous jusqu'à $n = 5$).
- L'ensemble des racines n^e de l'unité – noté $\mathbb{U}_n$ – forme un groupe multiplicatif qui est un sous-groupe de $\mathbb{U}$ (l'ensemble des complexes de module 1, voir la propriété 5.11) car le produit de deux éléments de $\mathbb{U}_n$ est un élément de $\mathbb{U}_n$, et aussi l'inverse d'un élément de $\mathbb{U}_n$ est un élément de $\mathbb{U}_n$. Nous laissons les démonstrations très simples de ces propriétés au lecteur.
- En notant u la première racine n^e de l'unité d'argument non nul (modulo 2π), soit $u = e^{i\frac{2\pi}{n}}$, on remarque que toutes les racines n^e de l'unité s'expriment comme des puissances de u :
  $\forall k \in \{0, 1, \ldots, n-1\}, e^{i\frac{2k\pi}{n}} = u^k$. De plus, comme $u^n = e^{i\frac{2n\pi}{n}} = 1$, on en déduit que
  $u^n - 1 = 0$ or $u^n - 1 = (u-1)(u^{n-1} + u^{n-2} + \cdots + u + 1)$ d'après la propriété 5.20. On en déduit que $u^{n-1} + u^{n-2} + \cdots + u + 1 = \frac{u^n - 1}{u-1} = 0$, la somme des racines n^e de l'unité est nulle.



1. La forme de William George Hörner (1819-1845) pour $a \times z \times z + b \times z + c$ est $c + z \times (b + z \times c)$, on voit que dès le 2^e degré, on gagne un produit ; au 3^e degré on en gagne trois ($az^3 + bz^2 + cz + d$ s'écrit $d + z(c + z(b + za))$).

Propriété 5.22 (racine n-ième) $n>0$ étant un entier et $r \neq$ un réel, l'équation $z^n = re^{i\theta}$ admet n solutions distinctes (les « racines n^e de $re^{i\theta}$) égales à $\sqrt[n]{r}e^{i(\frac{\theta+2k\pi}{n})}$ avec $k \in \{0, 1, \ldots, n-1\}$, soit $\begin{cases} |z| = \sqrt[n]{r} \\ \arg z = \dfrac{\theta}{n} + \dfrac{2k\pi}{n} \end{cases}$. Autrement dit, en notant $u = e^{i\frac{2\pi}{n}}$, $z = \sqrt[n]{r}u^k e^{i(\frac{\theta}{n})}$.

Démonstration si $r = 0$ alors l'équation devient $z^n = 0$ et n'a qu'une seule racine $z = 0$.
Pour cela on suppose $r \neq 0$.
D'après la propriété 5.9, on en déduit que :

- $|z^n| = |z|^n = r$ d'où $|z| = \sqrt[n]{r}$
- $\arg z^n = n \arg z \ [2\pi]$ d'où $\arg z = \frac{\theta}{n} \ [\frac{2\pi}{n}] \iff \exists k \in \mathbb{Z}, \arg z = \frac{\theta}{n} + \frac{2k\pi}{n}$

Remarque : On en déduit que si on arrive à déterminer une racine z_0 de l'équation $z^n = re^{i\theta}$, alors on détermine les autres en les multipliant par les racines n^e de l'unité égales à $e^{i\frac{2k\pi}{n}}$ avec $k \in \{1, 2, \ldots, n-1\}$. Autrement dit, les racines n^e de $re^{i\theta}$ sont z_0, z_0u, z_0u^2, z_0u^3, …, z_0u^{n-1}.

4.d. L'invention des algébristes italiens

Revenons aux formules qui ont tout déclenché, les formules de Tartaglia-Cardan.
L'équation réelle du 3e degré est $ax^3 + bx^2 + cx + d = 0$, où a, b, c et d sont des réels.
En divisant par $a \neq 0$, on obtient $x^3 + b'x^2 + c'x + d' = 0$ et, avec le changement de variable $X = x + \frac{b'}{3}$, elle se transforme en une équation de la forme $X^3 + pX + q = 0$ où le terme de degré 2 a disparu. En effet, on a :
$(x+\frac{b'}{3})^3 + p(x+\frac{b'}{3}) + q = x^3 + b'x^2 + \frac{b'^2}{3}x + \frac{b'^3}{27} + px + \frac{pb'}{3} + q = x^3 + b'x^2 + (\frac{b'^2}{3} + p)x + (\frac{b'^3}{27} + \frac{pb'}{3} + q)$,

il suffit de prendre b', c' et d' vérifiant $\begin{cases} b' = \dfrac{b}{a} \\ \dfrac{b'^2}{3} + p = c' = \dfrac{c}{a} \\ \dfrac{b'^3}{27} + \dfrac{pb'}{3} + q = d' = \dfrac{d}{a} \end{cases}$,

ce qui conduit à prendre $\begin{cases} p = \dfrac{c}{a} - \dfrac{b^2}{3a^2} \\ q = \dfrac{d}{a} - [\dfrac{b^3}{27a^3} + (\dfrac{c}{a} - \dfrac{b^2}{3a^2})\dfrac{b}{3a}] = \dfrac{d}{a} - \dfrac{bc}{3a^2} + \dfrac{2b^3}{27a^3} \end{cases}$

Les coefficients p et q sont, comme a, b, c et d, des réels.

Une fois cette forme trouvée, un nouveau changement de variable intervient : on pose $X = u + v$, avec u et v complexes (astuce : en prenant deux inconnues au lieu de une, on va pouvoir trouver une combinaison qui simplifie le problème).
L'équation $X^3 + pX + q = 0$ s'écrit alors :
$(u+v)^3 + p(u+v) + q = 0 \iff u^3 + 3u^2v + 3uv^2 + v^3 + pu + pv + q = 0 \iff u^3 + v^3 + (3uv+p)(u+v) + q = 0$.
Comme on a le choix de la décomposition de X en $u+v$, on choisit d'avoir $3uv + p = 0$, soit $uv = -\frac{p}{3}$.
L'équation devient alors $u^3 + v^3 + q = 0$ et on doit trouver deux nombres complexes u et v tels que $\begin{cases} u^3 + v^3 = -q \\ u^3v^3 = -\dfrac{p^3}{27} \text{ (en élevant au cube la condition sur le produit)} \end{cases}$

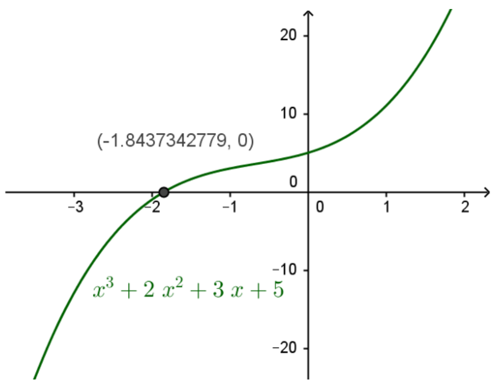
On est ainsi ramené au second degré [2], car u^3 et v^3 sont les solutions de l'équation $x^2 - Sx + P = 0$, où S est leur somme et P leur produit. L'équation que vérifient u^3 et v^3 est donc ici $x^2 + qx - \frac{p^3}{27} = 0$.

2. C'est du faux second degré. Il s'agit plutôt du sixième degré, car on cherche u et v, non pas u^3 et v^3, et ces nombres vérifient l'équation $x^6 + qx^3 - \frac{p^3}{27} = 0$. C'est important de le souligner car une équation du sixième degré a six solutions. Il faudra sélectionner, parmi ces six nombres, celles qui conviennent.

Son discriminant est $\Delta = q^2 + \frac{4}{27}p^3$ et les solutions $u^3 = \frac{-q+i\sqrt{-\Delta}}{2}$ et $v^3 = \frac{-q-i\sqrt{-\Delta}}{2}$ sont réelles si $\Delta \geqslant 0$ et complexes sinon.
D'après la propriété 5.22 et sa conséquence, soulignée dans la remarque, les trois racines cubiques de u^3 sont $u_0 = \sqrt[3]{\frac{-q+i\sqrt{-\Delta}}{2}}$, $u_1 = ju_0$ et $u_2 = j^2u_0$, de même pour les racines cubiques de v^3.
Il faut alors combiner deux à deux ces six solutions qui doivent avoir un produit réel, car on doit avoir $uv = -\frac{p}{3}$. Ce choix conduit aux trois solutions en $X = u + v$ et, finalement, en $x = X - \frac{b'}{3}$.

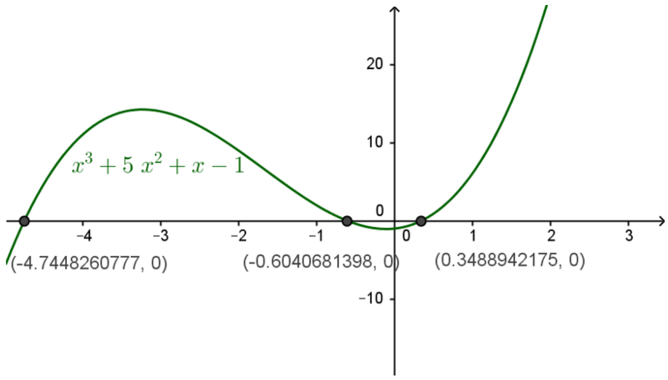
Exemple 56 – Dans le cas où l'équation $x^2 + qx - \frac{p^3}{27} = 0$ a un discriminant positif, l'équation du 3e degré a une solution réelle et deux solutions complexes conjuguées.
Résolvons l'équation $x^3 + 2x^2 + 3x + 5 = 0$.
Avec le changement de variable $X = x + \frac{2}{3}$, celle-ci se transforme en $X^3 + \frac{5}{3}X + \frac{97}{27} = 0$ car
$$\begin{cases} p = 3 - \dfrac{2^2}{3} = \dfrac{9-4}{3} = \dfrac{5}{3} \\ q = 5 - \dfrac{6}{3} + \dfrac{2 \times 2^3}{27} = 3 + \dfrac{16}{27} = \dfrac{97}{27} \end{cases}$$
Après le changement de variable $X = u + v$ et $uv = -\frac{5}{9}$, on obtient l'équation $u^3 + v^3 = -\frac{97}{27}$ avec $u^3v^3 = -\frac{5^3}{27^2} = -\frac{125}{729}$. Les nombres complexes u^3 et v^3 sont les solutions de l'équation
$x^2 + \frac{97}{27}x - \frac{125}{729} = 0$ dont le discriminant est $\Delta = \frac{97^2+4\times125}{27^2} = \frac{9909}{729} > 0$ d'où $u^3 = \frac{-97+\sqrt{9909}}{2\times27}$ et $v^3 = \frac{-97-\sqrt{9909}}{2\times27}$.
Les racines cubiques de ces nombres réels sont réelles également :
- $u_0 = \sqrt[3]{\frac{-97+\sqrt{9909}}{2\times27}} = \frac{1}{3}\sqrt[3]{\frac{-97+\sqrt{9909}}{2}}$
- $v_0 = \sqrt[3]{\frac{-97-\sqrt{9909}}{2\times27}} = \frac{1}{3}\sqrt[3]{\frac{-97-\sqrt{9909}}{2}}$

(u_0, v_0) est un couple vérifiant $uv = -\frac{5}{9}$, les deux autres couples qui conviennent sont constitués de nombres complexes conjugués : (u_0j, v_0j^2) pour le premier et (u_0j^2, v_0j) pour le second.
Il ne reste plus qu'à exprimer les solutions de l'équation de départ, sous la forme $x = u + v - \frac{b}{3a}$, c'est-à-dire ici $x = u + v - \frac{2}{3}$. La seule solution réelle est
$x_0 = \frac{\sqrt[3]{-97+\sqrt{9909}}+\sqrt[3]{-97-\sqrt{9909}}}{3\sqrt[3]{2}} - \frac{2}{3} \approx -1,8437342779$
On peut vérifier cette valeur en traçant la courbe d'équation $y = x^3 + 2x^2 + 3x + 5$.
Cette courbe coupe l'axe des abscisses en un seul point d'abscisse approximative $-1,8437342779$, ce qui correspond bien à la valeur trouvée. Les deux autres solutions sont les complexes :
- $x_1 = \frac{j\sqrt[3]{-97+\sqrt{9909}}+j^2\sqrt[3]{-97-\sqrt{9909}}}{3\sqrt[3]{2}} - \frac{2}{3} \approx -0.078 + 1.645i$
- $x_2 = \frac{j^2\sqrt[3]{-97+\sqrt{9909}}+j\sqrt[3]{-97-\sqrt{9909}}}{3\sqrt[3]{2}} - \frac{2}{3} \approx -0.078 + 1.645i$



Ces valeurs n'apparaissent pas sur la représentation graphique, car celle-ci ne prend en abscisse que des valeurs réelles.

Exemple 57 – Dans le cas où $\Delta < 0$, les trois solutions de l'équation de départ sont réelles.
Résolvons l'équation $x^3 + 5x^2 + x - 1 = 0$ qui a trois solutions réelles si on en juge par les trois points d'intersection de la courbe d'équation $y = x^3 + 5x^2 + x - 1$ avec l'axe des abscisses.



Le changement de variable $X = x + \frac{5}{3}$ conduit à l'équation $X^3 - \frac{22}{3}X + \frac{178}{27} = 0$.
On pose ensuite $X = u + v$ avec $uv = \frac{22}{9}$, et on on est ramené à chercher deux complexes u^3 et v^3 solutions de l'équation $x^2 + \frac{178}{27}x + \frac{10648}{729} = 0$.
Le discriminant est $\Delta = \frac{178^2 - 4\times 10648}{729} = -\frac{10908}{729} < 0$.
Le premier couple (u_0,v_0) vérifiant $uv = \frac{22}{9}$ est donc un couple de nombres complexes.

- $u_0 = \sqrt[3]{\frac{-178+i\sqrt{10908}}{2\times 27}} = \frac{1}{3}\sqrt[3]{\frac{-178+i\sqrt{10908}}{2}}$
- $v_0 = \sqrt[3]{\frac{-178-i\sqrt{10908}}{2\times 27}} = \frac{1}{3}\sqrt[3]{\frac{-178-i\sqrt{10908}}{2}}$

Les deux autres couples sont $(u_0 j, v_0 j^2)$ et $(u_0 j^2, v_0 j)$.
Les solutions de l'équation de départ sont les nombres réels :

- $x_0 = \frac{\sqrt[3]{-178+i\sqrt{10908}} + \sqrt[3]{-178-i\sqrt{10908}}}{3\sqrt[3]{2}} - \frac{5}{3} \approx 0,3489$
- $x_1 = \frac{j\sqrt[3]{-178+i\sqrt{10908}} + j^2\sqrt[3]{-178-i\sqrt{10908}}}{3\sqrt[3]{2}} - \frac{5}{3} \approx -4,7448$
- $x_2 = \frac{j^2\sqrt[3]{-178+i\sqrt{10908}} + j\sqrt[3]{-178-i\sqrt{10908}}}{3\sqrt[3]{2}} - \frac{5}{3} \approx -0,6041$

Remarque n°1 : Déterminer la racine cubique d'un nombre complexe n'est pas une tâche facile en général, surtout si l'on ne veut pas se satisfaire de valeurs approchées. Si on écrit un nombre complexe $a + ib$, alors son carré est $a^2 - b^2 + i(2ab)$ et son cube $a(a^2 - 2b^2) + i3b(a^2 - b^2)$.
Par conséquent, la racine cubique d'un nombre $A + iB$ est le nombre $z_0 = a + ib$ tel que
$\begin{cases} A = a(a^2 - 2b^2) \\ B = 3b(a^2 - b^2) \end{cases}$. Ce système conduit aux trois complexes solutions : z_0, jz_0 et $j^2 z_0$.
En dehors de certains cas particuliers, on ne sait pas exprimer, en général, les valeurs de a et b – la partie réelle et imaginaire de la racine cubique d'un complexe – de façon algébrique, sans les radicaux. Bien sûr le problème se simplifie beaucoup lorsqu'on arrive à identifier la forme trigonométrique du nombre z_0.

Remarque n°2 : Lorsque l'équation du troisième degré a une solution réelle évidente, le passage par $\mathbb{C}$ n'est pas obligatoire, et il complique notablement l'expression des solutions.
L'équation $x^3 - 15x - 4 = 0$, pour donner un exemple historique résolu par Bombelli en 1560, a une solution évidente qui est $x = 4$, et deux autres solutions réelles $-2 + \sqrt{3}$ et $-2 - \sqrt{3}$ que l'on trouve en factorisant le polynôme initial : $x^3 - 15x - 4 = (x - 4)(x^2 + 4x + 1)$.
La méthode de Tartaglia-Cardan exprime les racines comme la somme de racines cubiques de nombres complexes : l'équation que doit vérifier u^3 et v^3 est ici $x^2 - 4x + 125 = 0$, son discriminant étant négatif ($\Delta = 16 - 500 = -484 = -22^2$), on en déduit le couple de complexes $(u_0,\ v_0)$ suivant :

- $u_0 = \sqrt[3]{\frac{4+22i}{2}} = \sqrt[3]{2 + 11i}$
- $v_0 = \sqrt[3]{\frac{4-22i}{2}} = \sqrt[3]{2 - 11i}$

La 1re solution qui s'en déduit est le complexe $\sqrt[3]{2 + 11i} + \sqrt[3]{2 - 11i}$, or ce nombre est égal à 4, cela se prouve en développant $(2 + i)^3 = (3 + 4i)(2 + i) = 2 + 11i$ et $(2 - i)^3 = (3 - 4i)(2 - i) = 2 - 11i$.
Les autres solutions ne sont pas moins réelles que la première, mais la méthode de Cardan en donne à nouveau, des expressions complexes : $j\sqrt[3]{2 + 11i} + j^2\sqrt[3]{2 - 11i}$ et $j^2\sqrt[3]{2 + 11i} + j\sqrt[3]{2 - 11i}$

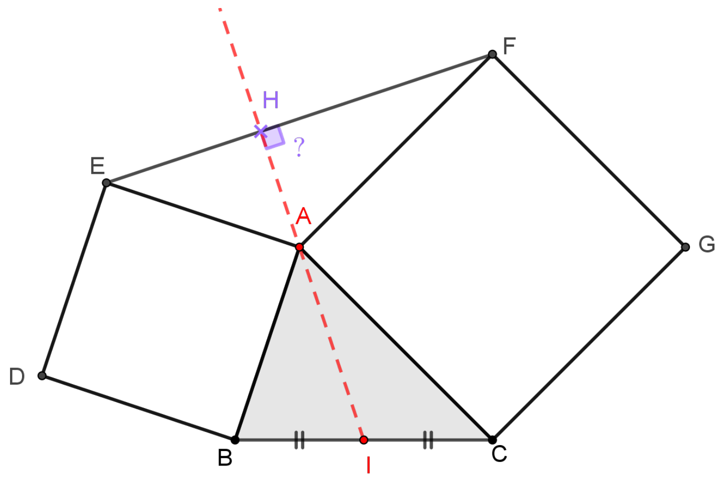
5. Applications des complexes

En dehors des équations polynomiales et autres questions algébriques qui furent le point de départ des nombres complexes, de multiples applications utilisent ces nombres comme des outils pour résoudre autrement, représenter ou étudier certains problèmes. Nous envisagerons deux domaines : la géométrie et les suites.

5.a. Applications en géométrie

Les propriétés 5.8 et 5.9, notamment dans les remarques associées, établissent plusieurs points importants qui permettent de résoudre des problèmes de géométrie.
Développons cet aspect sur deux exemples.

<u>**Exemple**</u> 58 (Figure de Vecten) – Soit ABC un triangle de sens direct (le sens $A \to B \to C$ est trigonométrique).
On construit des carrés $ABDE$ et $ACGF$ sur les côtés $[AB]$ et $[AC]$, plus précisément E est l'image de B par une rotation de centre A et d'angle $\frac{-\pi}{2}$ (ce qu'on notera $E = r_{A,\frac{-\pi}{2}}(B)$) et F est l'image de C par une rotation de centre A et d'angle $\frac{\pi}{2}$ ($F = r_{A,\frac{\pi}{2}}(C)$). Montrons que la médiane issue de A dans le triangle ABC est la hauteur issue de A dans le triangle AEF.



Notons a, b et c les affixes de A, B et C.

- La rotation $r_{A,\frac{-\pi}{2}}$ transforme un point $M(z)$ en $M'(z')$ de telle façon que le rapport des affixes des vecteurs $\overrightarrow{AM'}$ et $\overrightarrow{AM}$ ($\overrightarrow{AM'}(z'-a)$ et $\overrightarrow{AM}(z-a)$) soit égal au complexe $e^{i\frac{-\pi}{2}} = -i$.
  On a donc la traduction complexe de cette rotation : $\frac{z'-a}{z-a} = -i \iff z' - a = -i(z-a)$.
  Comme $E = r_{A,\frac{-\pi}{2}}(B)$, on a $z_E - a = -i(z_B - a) = -i(b-a) \iff z_E = a - i(b-a)$.
- De même, la rotation $r_{A,\frac{\pi}{2}}$ a pour expression complexe $\frac{z'-a}{z-a} = i \iff z' - a = i(z-a)$ et donc, comme $F = r_{A,\frac{\pi}{2}}(C)$, on a $z_F - a = i(z_C - a) = i(c-a) \iff z_F = a + i(c-a)$.

Notons I le milieu de $[BC]$, son affixe est $z_I = \frac{b+c}{2}$.
L'affixe du vecteur $\overrightarrow{EF}$ est $z_{\overrightarrow{EF}} = z_F - z_E = a + i(c-a) - (a - i(b-a)) = i(-2a+b+c)$.
L'affixe du vecteur $\overrightarrow{AI}$ est $z_{\overrightarrow{AI}} = z_I - z_A = \frac{b+c}{2} - a = \frac{1}{2}(-2a+b+c)$.
On en déduit que $z_{\overrightarrow{EF}} = 2iz_{\overrightarrow{AI}} \iff \frac{z_{\overrightarrow{EF}}}{z_{\overrightarrow{AI}}} = 2i = 2e^{i\frac{\pi}{2}}$.
Cela signifie que l'angle $(\overrightarrow{EF}, \overrightarrow{AI}) = \frac{\pi}{2}$ (les droites (EF) et (AI) sont donc bien perpendiculaires) et aussi que $EF = 2AI$ (ce résultat non cherché est donné en supplément).

<u>Remarque :</u> On simplifie le raisonnement en choisissant l'origine du repère en A.
Dans un tel repère, les points ont pour affixes $A(0)$, $B(b)$, $C(c)$, $I(\frac{b+c}{2})$, $E(-ib)$ et $F(ic)$.
Du coup, les affixes des vecteurs sont $\overrightarrow{EF}(i(b+c))$ et $\overrightarrow{AI}(\frac{b+c}{2})$.

On calcule alors $(\overrightarrow{EF}, \overrightarrow{AI}) = \arg\left(\frac{z_{\overrightarrow{EF}}}{z_{\overrightarrow{AI}}}\right) = \arg\left(\frac{i(b+c)}{\frac{b+c}{2}}\right) = \arg(2i) = \frac{\pi}{2}$, et on conclut.

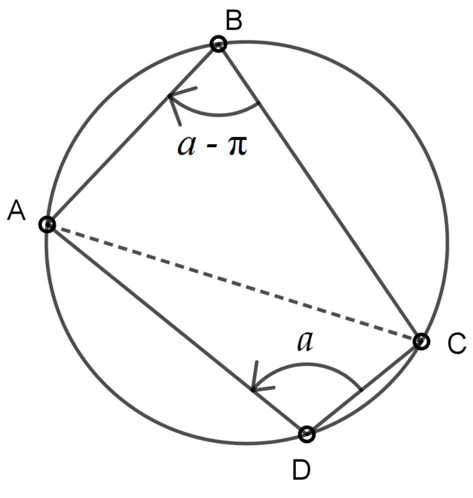
Exemple 59 (Théorème de Ptolémée) – A, B, C et D étant quatre points distincts, montrons qu'on a $AC \times BD \leqslant AB \times CD + AD \times BC$. Puis, en interrogeant les cas d'égalité de cette formule, montrons que quatre points distincts sont cocycliques ou alignés si et seulement si $AB \times CD \pm AC \times BD \pm AD \times BC = 0$ (pas les deux + en même temps).

Notons a, b, c et d les affixes de A, B, C et D.
Ce théorème apparait ici comme une application de l'inégalité triangulaire de la propriété 5.10 :
En effet, $AB \times CD + AD \times BC = |b-a||d-c| + |c-b||d-a| = |(b-a)(d-c)| + |(c-b)(d-a)|$ (d'après la propriété 5.9).
Et d'après l'inégalité triangulaire, $|(b-a)(d-c)| + |(c-b)(d-a)| \geqslant |(b-a)(d-c) + (c-b)(d-a)|$.
Développons et réduisons :
$(b-a)(d-c)+(c-b)(d-a) = bd+ac-ad-bc+cd+ab-bd-ac = -ad-bc+cd+ab = (a-c)(b-d)$
On en déduit :

$$\begin{aligned} AB \times CD + AD \times BC \geqslant |(a-c)(b-d)| &\iff AB \times CD + AD \times BC \geqslant |a-c| \times |b-d| \\ &\iff AB \times CD + AD \times BC \geqslant AC \times BD \end{aligned}$$

Il y a égalité dans l'inégalité triangulaire si l'un des vecteurs est nul ou s'ils sont colinéaires et de même sens. Ici les affixes des vecteurs concernés sont $(b-a)(d-c)$ et $(c-b)(d-a)$ qui ne sont pas nuls car les points sont distincts. L'égalité a donc lieu lorsque les vecteurs sont colinéaires et de même sens, c'est-à-dire lorsque le rapport de leurs affixes est un réel strictement positif (la partie imaginaire de ce rapport étant le déterminant des vecteurs). On doit donc avoir :

$$\begin{aligned} \frac{(b-a)(d-c)}{(c-b)(d-a)} \in \mathbb{R}^+ \iff \frac{(b-a)(d-c)}{(b-c)(d-a)} \in \mathbb{R}^- &\iff \frac{\frac{b-a}{b-c}}{\frac{d-a}{d-c}} \in \mathbb{R}^- \\ &\iff \arg\left(\frac{\frac{b-a}{b-c}}{\frac{d-a}{d-c}}\right) = \pi \ [2\pi] \iff \arg\left(\frac{b-a}{b-c}\right) = \arg\left(\frac{d-a}{d-c}\right) + \pi \ [2\pi] \\ &\iff (\overrightarrow{BC}, \overrightarrow{BA}) = (\overrightarrow{DC}, \overrightarrow{DA}) + \pi \ [2\pi] \end{aligned}$$



Cette dernière égalité traduit une situation de cocyclicité décrite par le théorème de l'angle inscrit (dans le cas d'un birapport négatif comme ici, les points B et D sont de part et d'autre de la droite (AC), voir la figure) ou d'alignement (avec $B \in [AC] \Longrightarrow D \notin [AC]$ ou $D \in [AC] \Longrightarrow B \notin [AC]$).
Sans tenir compte du sens, il y a trois ordres possibles pour quatre points ($A \to B \to C \to D$, $A \to B \to D \to C$ et $A \to C \to B \to D$),
d'où les trois possibilités pour avoir A, B, C et D cocycliques ou alignés :

- $AC \times BD = AB \times CD + AD \times BC$
- $AD \times BC = AB \times CD + AC \times BD$
- $AB \times CD = AC \times BD + AD \times BC$

5.b. Étude de suites

5.b.1. Suites récurrentes d'ordre 2

Pour chercher le terme général d'une suite $(u_n)_{\leqslant 0}$ récurrente d'ordre 2, c'est-à-dire définie par une relation du type $u_{n+2} = au_{n+1} + bu_n$ où a et b ne sont pas nuls, on cherche à écrire cette relation sous la forme $u_{n+2} - ru_{n+1} = s(u_{n+1} - ru_n)$ car ainsi, la suite (v_n) définie par $v_n = u_{n+1} - ru_n$ est une suite géométrique de raison s.

Comme $u_{n+2} - ru_{n+1} = s(u_{n+1} - ru_n) \iff u_{n+2} = (r+s)u_{n+1} - rsu_n$, on doit avoir $\begin{cases} r+s=a \\ rs=-b \end{cases}$.

Il faut choisir r et s solutions de l'équation $x^2 - ax - b = 0$ et selon les cas :

- Si l'équation admet deux racines réelles distinctes, alors les suites cherchées ont pour terme général $u_n = \alpha r^n + \beta s^n$
- Si l'équation admet une seule racine réelle $r = \frac{a}{2}$, alors les suites cherchées ont pour terme général $u_n = (\alpha + n\beta)r^n$, autrement dit la suite $(\frac{u_n}{r^n})$ est arithmétique.
- Si l'équation n'admet aucune racine réelle (si $a^2 + 4b < 0$), elle en admet deux complexes avec lesquelles on exprime $u_n = \alpha r^n + \beta s^n$, voyons cela sur un exemple.

Exemple 60 – Étudions la suite $(u_n)_{\leqslant 0}$ définie par la relation $u_{n+2} = 2u_{n+1} - 2u_n$ et les conditions initiales $u_0 = 0$ et $u_1 = 1$.

Cherchons r et s tels que $\begin{cases} r+s=2 \\ rs=-2 \end{cases}$: ce sont les solutions de $x^2 - 2x + 2 = 0$ dont le discriminant est $\Delta = -4 < 0$. Les solutions sont $\frac{2\pm\sqrt{-4}}{2} = 1 \pm \sqrt{-1}$, soit $1+i$ et $1-i$.

Posons $\begin{cases} r=1+i \\ s=1-i \end{cases}$; comme ces deux solutions sont interchangeables :

- d'une part on a $u_{n+1} - ru_n = s(u_n - ru_{n-1}) = s^n(u_1 - ru_0) = s^n$ (car $u_0 = 0$ et $u_1 = 1$)
- et d'autre part $u_{n+1} - su_n = r(u_n - su_{n-1}) = r^n(u_1 - su_0) = r^n$

Par soustraction des membres extrêmes, on obtient
$(u_{n+1} - ru_n) - (u_{n+1} - su_n) = s^n - r^n \iff (s-r)u_n = s^n - r^n$
Ainsi, en remplaçant r et s par leurs valeurs :
$(1-i-1-i)u_n = (1-i)^n - (1+i)^n \iff -2iu_n = (1-i)^n - (1+i)^n \iff u_n = \frac{(1+i)^n - (1-i)^n}{2i}$.
Écrivons $r = 1+i$ et $s = 1-i$ sous forme exponentielle : $|r| = |s| = \sqrt{2}$ et,
comme $\cos(\arg r) = \frac{\sqrt{2}}{2}$ et $\sin(\arg r) = \frac{\sqrt{2}}{2}$, on a $\arg r = \frac{\pi}{4}$;
de même comme $\cos(\arg s) = \frac{\sqrt{2}}{2}$ et $\sin(\arg s) = \frac{-\sqrt{2}}{2}$, on a $\arg sr = \frac{-\pi}{4}$.

On en déduit $u_n = \dfrac{(\sqrt{2}e^{i\frac{\pi}{4}})^n - (\sqrt{2}e^{i\frac{-\pi}{4}})^n}{2i} = (\sqrt{2})^n \dfrac{e^{i\frac{n\pi}{4}} - e^{i\frac{-n\pi}{4}}}{2i}$.

Cette dernière égalité peut être réduite à l'aide de la formule de Euler en $u_n = (\sqrt{2})^n \sin(\frac{n\pi}{4})$.
Malgré le passage par des nombres complexes et irrationnels, tous les nombres de la suites (u_n) sont entiers puisque la relation de récurrence initiale ne fait qu'ajouter ou multiplier des entiers.

5.b.2. Série trigonométrique

Nous étudions ici certaines suites (S_n) dont le terme général est la somme des premiers termes d'une suite (u_n), soit $S_n = u_0 + u_1 + \cdots + u_{n-1}$. En particulier lorsque les nombres complexes permettent d'exprimer S_n d'une façon simplifiée.

Exemple 61 – On définit la suite $(S_n)_{n>0}$ par $S_n = \sin\frac{\pi}{n} + \sin\frac{2\pi}{n} + \cdots + \sin\frac{(n-1)\pi}{n}$.
Posons $z = \cos\frac{\pi}{n} + i\sin\frac{\pi}{n} = e^{i\frac{\pi}{n}}$.

On sait que la somme $S = 1 + z + z^2 + \cdots + z^{n-1} = \dfrac{1-z^n}{1-z}$ mais comme $z^n = e^{i\frac{n\pi}{n}} = e^{i\pi} = -1$, on en déduit $S = \dfrac{1-(-1)}{1-z} = \dfrac{2}{1-z}$.

Mais comme $1 - \cos\theta = 2\sin^2\frac{\theta}{2}$ et $\sin\theta = 2\sin\frac{\theta}{2}\cos\frac{\theta}{2}$, on peut écrire
$1 - z = (1 - \cos\frac{\pi}{n}) - i(\sin\frac{\pi}{n}) = 2\sin^2\frac{\pi}{2n} - i(2\sin\frac{\pi}{2n}\cos\frac{\pi}{2n}) = 2\sin\frac{\pi}{2n}(\sin\frac{\pi}{2n} - i\cos\frac{\pi}{2n})$.

On en déduit $S = \dfrac{2}{2\sin\frac{\pi}{2n}(\sin\frac{\pi}{2n} - i\cos\frac{\pi}{2n})} = \dfrac{\sin\frac{\pi}{2n} + i\cos\frac{\pi}{2n}}{\sin\frac{\pi}{2n}} = 1 + \dfrac{i}{\tan\frac{\pi}{2n}}$.

Comme $S_n = \mathcal{I}m(S)$, on obtient finalement $S_n = \dfrac{1}{\tan\frac{\pi}{2n}} \dfrac{\cos\frac{\pi}{2n}}{\sin\frac{\pi}{2n}}$.

5.b.3. L'ensemble de Mendelbrot

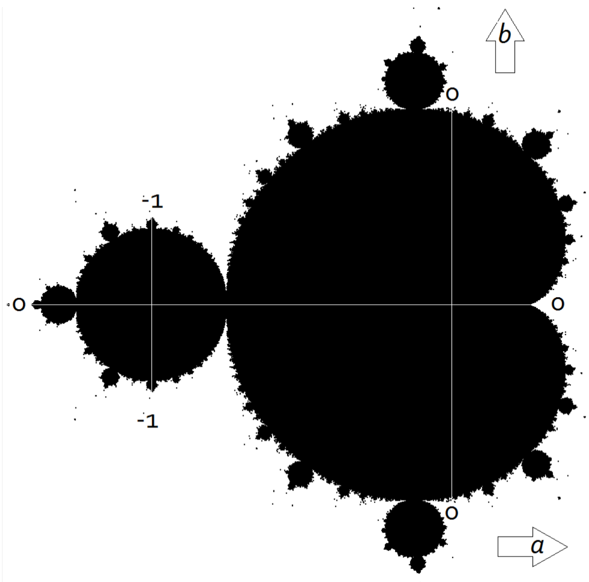
Cet ensemble, appelé simplement M dans la suite, a été découvert au début du 20e siècle par les mathématiciens Pierre Fatou et Gaston Julia, mais ne fut connu du grand public qu'après que Benoît Mandelbrot (1924-2010) en produise des images à l'aide d'un ordinateur (1980) et que le magazine *Scientific American* les publie (1985).

Pour tout nombre complexe $c = a + bi$, on définit la suite complexe (z_n) par la relation de récurrence $z_{n+1} = z_n^2 + c$ avec $z_0 = 0$. Selon la valeur de c, on observe deux comportements de la suite (z_n) :

- elle a tendance à rester bornée $(\exists\lambda \in \mathbb{R}^+, \forall n > n_0, |z_n| < \lambda)$
- elle diverge, les modules se mettant à croître indéfiniment $(\forall\lambda \in \mathbb{R}^+, \exists n_0 \in \mathbb{N}, \forall n > n_0, |z_n| > \lambda)$

L'ensemble M contient tous les points c pour lesquels la suite est bornée.

Si on colore les points de cet ensemble en noir, on obtient l'image désormais célèbre ci-dessous. J'ai ajouté une ébauche de graduation afin de remarquer que l'essentiel de cet ensemble semble limité au domaine $-1,5 < a < 0,5$ et $-1 < b < 1$.



Bien sûr, la vérification du fait que la suite est bornée, lorsqu'on utilise un ordinateur, ne peut pas être faite jusqu'à l'infini. Il faut se contenter d'effectuer pour chaque point d'affixe c (chaque pixel de l'image que l'on veut créer), un nombre limité t d'itérations. Le paramètre t est appelé « profondeur » de la représentation.

Pour l'image globale ci-dessus, la profondeur $t = 100$ a suffi.

Choisir $t = 1000$ permettrait d'obtenir des détails du bord mais impliquerait beaucoup plus de calculs et demanderait beaucoup plus de temps (d'où la lettre t).

Heureusement, cette suite a une propriété qui évite d'aller à chaque fois jusqu'à t :

Si il existe un rang n_0 à partir duquel $|z_n| > 2$, alors la suite diverge et le point d'affixe c n'appartenant pas à M reste en blanc. Seuls les points suspectés jusqu'au bout d'appartenir à M nécessitent d'aller jusqu'à t et seront coloriés en noir.

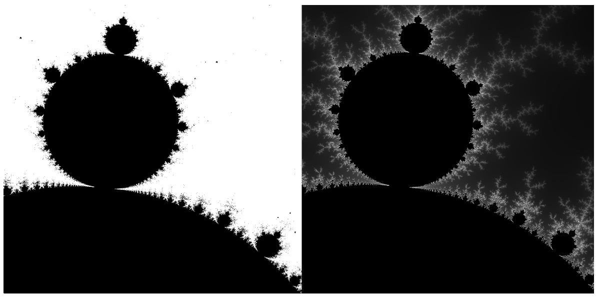
```
Re(c) = 1
Im(c) = 1
rang 1:  1.0+1.0i - module=1.4142135623730951
rang 2:  1.0+3.0i - module=3.1622776601683795
rang 3:  -7.0+7.0i - module=9.899494936611665
rang 4:  1.0-97.0i - module=97.0051545022222
rang 5:  -9407.0-193.0i - module=9408.979647124337
rang 6:  88454401.0+3631103.0i - module=88528899.0401745
rang 7:  7810996147272193.0+642374081668607.0i - module=7837365965265411.0
```

```
Re(c) = -0.5
Im(c) = 0.5
rang 1:  -0.5+0.5i - module=0.7071067811865476
rang 2:  -0.5+0.0i - module=0.5
rang 3:  -0.25+0.5i - module=0.5590169943749475
rang 4:  -0.6875+0.25i - module=0.7315437444199766
rang 5:  -0.08984375+0.15625i - module=0.180238624922802
rang 6:  -0.516342163085937 5+0.471923828125i - module=0.6995150669802755
rang 7:  -0.45610287017188966+0.012651659548282623i - module=0.456278306155
rang 8:  -0.292130236310 29+0.48845908353518 14i - module=0.5691505523626725
rang 9:  -0.6532522013213234+0.21461266499791953i - module=0.6876023811037
rang 10:  -0.1193201574463543 2+0.21960760831734666i - module=0.249929593293
 ....
rang 95:  -0.47304602597343 65+0.3041553202790195i - module=0.5623904351456 3
rang 96:  -0.3687379161647718+0.2122410689266641i - module=0.4254573094408 7
rang 97:  -0.4090786205215947+0.34347734103879635i - module=0.534155409571 9
rang 98:  -0.4506313660392307+0.21898152629484774i - module=0.501020495505 7
rang 99:  -0.3448842808000379 7+0.302640111336794 1i - module=0.458842243187 0
rang 100:  -0.4726458698469876 4+0.2912483657207327i - module=0.555175403650
```

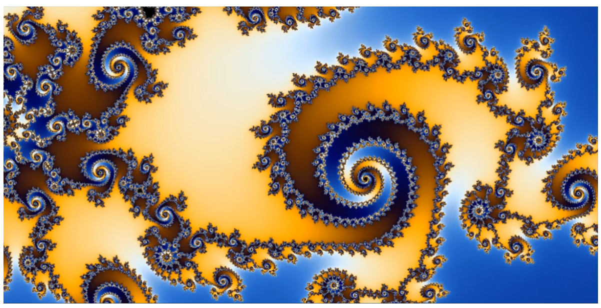
Observons ci-dessus le comportement de la suite (z_n) pour deux valeurs de c :

- Lorsque $c = 1 + i$, dès la 2ᵉ itération, le module dépasse 2. À la 4ᵉ étape il atteint déjà presque 100, et à la 6ᵉ presque le million...
- Pour $c = -\frac{1}{2} + \frac{i}{2}$, après $t = 100$ itérations, le module n'a toujours pas dépassé 1. Il a oscillé gentiment entre $0,18$ et $0,73$ approximativement. Si on augmente le nombre d'itérations, pour 1000 d'entre elles, les oscillations du module ne sortent pas de ces limites. Les plus grands écarts sont au début, et la valeur des z_n se stabilise progressivement sur une valeur limite qui est, dans le cas de $c = -\frac{1}{2} + \frac{i}{2}$, d'environ $-0,409 + 0,275i$.

L'algorithme qui est nécessaire pour effectuer cette exploration est extrêmement simple. De plus, la plupart des environnements de programmation ont implémenté le type « nombre complexe » et il suffit, par exemple, d'écrire z^2 pour avoir le carré de z et $|z|$ ou $abs(z)$ pour en avoir le module.
Le caractère fractal de l'image apparaît lorsqu'on cherche à agrandir une des zones frontière entre M et le reste du plan complexe. L'image ci-dessous montre un tel agrandissement, obtenu en zoomant d'un facteur 10 la partie supérieure de M pour laquelle $-0,15 \leqslant a < -0,05$ et $-0,9 \leqslant b < -0,8$.



NB : Sur ces images, les ordonnées croissent vers le bas. À droite la convention de coloration est différente : en dehors de M, la couleur va du blanc lorsque l'échappement est obtenu pour $n = t - 1$, au noir lorsque l'échappement est obtenu pour $n = 1$. Avec ce nuancier de gris, le contraste est maximum entre l'intérieur et l'extérieur de M et on voit se dessiner des lignes extrêmement fines qui irriguent l'espace extérieur. Il est bien sûr possible de créer un nuancier de couleurs en exploitant toute la palette RGB comme sur l'image ci-dessous (Wikipedia) qui montre un détail de M centré sur la valeur $-0,068 + 0,662i$ pour $t = 300$.



6 Arithmétique

Plan du chapitre :

Aperçu historique :

Le mot arithmétique vient du grec ancien *arithmos* qui signifie « nombre », dans le sens qu'avait le mot à cette époque de « nombre entier ». L'origine de l'arithmétique est très ancienne, sans doute phénicienne. Au VI^e siècle avant J.-C., Pythagore (580-495 avant J.-C.) et son école en font un des quatre piliers des mathématiques, à côté de la géométrie, de l'astronomie et de la musique. Ce fameux *quadrivium* quantitatif fut regroupé avec le trivium, plus littéraires (grammaire, rhétorique, dialectique) pour former les sept arts libéraux au V^e siècle.
L'arithmétique étudie les nombres entiers, les propriétés qu'ont certains ou les relations qu'ils entretiennent avec d'autres : nombres premiers, nombres pairs ou impairs, nombres amis ou parfaits ou heureux, triplets pythagoriciens, nombres triangulaires et plus généralement nombres polygonaux de Diophante d'Alexandrie (II^e siècle), nombres de Marin Mersenne (1588-1648), nombres de Pierre de Fermat (1601-1665), nombres de Johann Carl Friedrich Gauss (1777-1855), nombres de Sophie Germain (1776-1831), nombres de Robert Daniel Carmichael (1879-1967), etc.

En lisant une édition de son époque des Arithmétiques de Diophante, Fermat rédige ce qu'on appellera le « grand théorème de Fermat » : *Pour tout entier naturel $n \geqslant 3$, il n'existe pas de nombres entiers strictement positifs x, y et z tels que $x^n + y^n = z^n$* et il ajoute dans la marge : *J'en ai découvert une démonstration véritablement merveilleuse que cette marge est trop étroite pour contenir* mais aucune trace de cette démonstration n'est trouvée. Beaucoup de mathématiciens ont essayé de démontrer cette conjecture mais il a fallu attendre plus de 300 ans pour en avoir la preuve que Andrew Wiles (1953-) exhibe comme sous-produit de ses travaux qui mobilisent bien d'autres notions que celles connues du temps de Fermat.
Ainsi, les problèmes arithmétiques s'énoncent et se comprennent facilement mais leur résolution peut entrainer des raisonnements très complexes et s'étaler sur de longues périodes. D'autres conjectures célèbres restent encore non démontrées comme la conjecture de Goldbach (1742) (*tout nombre entier pair supérieur à 3 peut s'écrire comme la somme de deux nombres premiers*) ou la conjecture de Syracuse de L. Collatz (1937) (*si pour tout entier $N > 0$, on construit la suite (u_n) en partant de $u_0 = N$ avec la relation $u_{n+1} = \frac{u_n}{2}$ si u_n pair ou $u_{n+1} = 3u_n + 1$ si u_n impair. alors il existe un indice $n > 0$ tel que $u_n = 1$.*
Souvent considérée comme très théorique, l'arithmétique a permis de mieux comprendre l'infini et les différents ensembles de nombres. De nos jours elle a de nombreuses applications, notamment en informatique (systèmes de cryptographie ou codes correcteurs d'erreur).

1. Divisibilité dans $\mathbb{Z}$

1.a. Multiples et diviseurs

DÉFINITION 6.1 (LA RELATION DIVISE) Soient a et $b \neq 0$ deux entiers relatifs.
On dit que b divise a lorsqu'il existe un entier relatif k tel que $a = kb$.
On note cela $b|a$ et on dit que b est un diviseur de a ou que a est un multiple de b.

Remarques :

- La liste des diviseurs d'un entier non nul a est finie car elle ne contient que des nombres entiers compris entre $-|a|$ et $|a|$ (car $\forall b \neq 0, b|a \Longrightarrow |b| \leqslant |a|$). Par contre la liste des diviseurs de 0 est infinie car pour tout $b \neq 0$ on a $0 = 0 \times b$ (l'ensemble des diviseurs de 0 est $\mathbb{Z}$). Tout entier possède des diviseurs car 1, -1, a et $-a$ divisent toujours a. Pour tout nombre a, les diviseurs de a et de $-a$ sont les mêmes.
- D'une façon générale on note $Div(a)$ l'ensemble des diviseurs d'un entier a. Par exemple $Div(-24) = Div(24) = \{-24, -12, -8, -6, -4, -3, -2, -1, 1, 2, 3, 4, 6, 8, 12, 24\}$ et $Div(0) = \mathbb{Z}$. Dans certaines situations qui ne concernent que les entiers naturels (entiers positifs) on ne retient que les diviseurs positifs d'un nombre $a > 0$.
- $0|0$ (zéro divise zéro) car $\forall b \neq 0, 0 = 0 \times 0$, par contre zéro ne divise aucun nombre $a \neq 0$ car si $a = k \times 0$ alors nécessairement $a = 0$.

EXEMPLE 62 – Déterminons tous les couples (a, b) d'entiers naturels tels que $a^2 - b^2 = 12$.
On sait que $a^2 - b^2 = (a - b)(a + b)$ donc $a - b$ et $a + b$ sont des diviseurs de 12 tels que $a + b$ et $a - b$ soient positifs ($a + b \geqslant 0$ car $a \geqslant 0$ et $b \geqslant 0$ et, comme le produit $(a - b)(a + b)$ est positif, le facteur $a - b$ doit être positif également). Comme $a - b \geqslant 0 \Longleftrightarrow a \geqslant b$, le plus grand des entiers est a.
Les diviseurs de 12 dans $\mathbb{N}$ sont $Div(12) = \{1, 2, 3, 4, 6, 12\}$, on a donc trois solutions potentielles :

$$\begin{cases} a + b = 12 \\ a - b = 1 \end{cases} \Longleftrightarrow \begin{cases} a = \dfrac{13}{2} \\ b = \dfrac{11}{2} \end{cases} \text{(ne convient pas car ce ne sont pas des entiers)}$$

$$\begin{cases} a + b = 6 \\ a - b = 2 \end{cases} \Longleftrightarrow \begin{cases} a = \dfrac{8}{2} = 4 \\ b = 2 \end{cases} \text{(convient)}$$

$$\begin{cases} a + b = 4 \\ a - b = 3 \end{cases} \Longleftrightarrow \begin{cases} a = \dfrac{7}{2} \\ b = \dfrac{1}{2} \end{cases} \text{(ne convient pas).}$$

Finalement, il n'y a que le couple $(4, 2)$ qui convienne et on a bien $4^2 - 2^2 = 16 - 4 = 12$.

PROPRIÉTÉ 6.1 Le relation « divise » est une relation d'ordre car elle est

- Réflexive : $\forall a \in \mathbb{Z}, a|a$
- Antisymétrique : $\forall a \in \mathbb{Z}^*, \forall b \in \mathbb{Z}^*, a|b$ et $b|a \Longrightarrow a = b$ ou $a = -b$
- Transitive : $\forall a \in \mathbb{Z}^*, \forall b \in \mathbb{Z}^*, \forall c \in \mathbb{Z}, a|b$ et $b|c \Longrightarrow a|c$

L'ordre est partiel car $\forall a \in \mathbb{Z}, \forall b \in \mathbb{Z}, a|b$ ou $b|a$ est une proposition fausse.
$\forall a, b, c \neq 0, c|a$ et $c|b \Longrightarrow c|(ma + nb)$ une combinaison linéaire de a et b (m et n entiers).

Remarques :

- Les relations d'ordre $\leqslant$ et $\geqslant$ sont totales car tous les nombres sont comparables avec ces relations. Par contre, $<$ et $>$ sont des relations d'ordre partiels car on ne peut pas comparer avec elles des nombres égaux. La relation d'ordre $|$ est plus partielle encore puisqu'elle ne permet de comparer que des entiers relatifs dont l'un est multiple de l'autre. On ne peut comparer 2 qu'à ses multiples (les éléments de $2\mathbb{Z} = \{0, 2, -2, 4, -4, \ldots\}$) ou à ses diviseurs (les éléments de $Div(2) = \{1, -1, 2, -2\}$) ; 2 et 3 ne sont pas comparables avec $|$.

- La propriété $c|a$ et $c|b$ alors c divise toute combinaison linéaire $ma+nb$ est justifiée par la définition : comme $a=ck$ et $b=ck'$, on a $ma+nb=mck+nck'=c(mk+nk')=cK$.
  Si $c|a$ et $c|b$ alors, en particulier, c divise $a+b$ et c divise $a-b$.

Exemple 63 – ➲ Montrons que pour tout entier naturel n,
5^n-1, 13^n-1 et plus généralement $(4k+1)^n-1$ sont divisibles par 4.
$5^n-1=(5-1)(5^{n-1}+5^{n-2}+\cdots+5+1)=4\sum_{i=0}^{n-1}5^i=4K$ donc $4|(5^n-1)$.
$13^n-1=(13-1)(13^{n-1}+13^{n-2}+\cdots+13+1)=4\times 3\sum_{i=0}^{n-1}13^i=4K'$
$(4k+1)^{n+1}-1=(4k+1-1)((4k+1)^{n-1}+(4k+1)^{n-2}+\cdots+(4k+1)+1)=4k\sum_{i=0}^{n-1}(4k+1)^i=4K''$

➲ Montrons, en raisonnant par récurrence, que 11^n-4^n est un multiple de 7 pour tout $n\in\mathbb{N}$.
Pour $n=0$: $11^n-4^n=0$ donc la proposition est vraie pour $n=0$.
Supposons la proposition vraie pour $n>0$.
Au rang suivant, $11^{n+1}-4^{n+1}=(7+4)11^n-4\times 4^n=4(11^n-4^n)+7\times 11^n$, ce qui est bien une combinaison linéaire de deux multiples de 7 (11^n-4^n l'est par hypothèse de récurrence et 7 l'est par définition) et par conséquent un nouveau multiple de 7.

➲ Justifions le critère de divisibilité par 11 : *un nombre est divisible par 11 lorsque la différence entre la somme des chiffres de rang pair et la somme des chiffres de rang impair est un multiple de* 11.
Un exemple : 919380 est divisible par 11 car $(9+9+8)-(1+3+0)=26-4=22=2\times 11$.
Le principe de ce critère repose sur le fait que, pour tout entier naturel n, on peut déterminer un entier k tel que $10^{2n+1}+1=11k$ et $10^{2n}-1=11k$. Autrement dit que 11 divise les puissances impaires de 10 augmentées de 1 et les puissances paires de 10 diminuées de 1.
Prouvons cela par récurrence :
Pour $n=0$: $10^1+1=11$, $10^0-1=0$, 11 et 0 étant des multiples de 11, la propriété est vraie.

- Supposons pour $n>0$, il existe un entier k tel que $10^{2n+1}+1=11k \iff 10^{2n+1}=11k-1$.
  $10^{2(n+1)+1}+1=100\times 10^{2n+1}+1=(9\times 11+1)(11k-1)+1=11(99k+11k-9)-1+1=11K$
- Supposons pour $n>0$, il existe un entier k' tel que $10^{2n}-1=11k' \iff 10^{2n}=11k'+1$.
  $10^{2(n+1)}-1=100\times 10^{2n}-1=(9\times 11+1)(11k'+1)-1=11(99k'+11k'+9)+1-1=11K'$

Un nombre entier n écrit en base 10 s'écrivant $a_pa_{p-1}\dots a_2a_1a_0$ se décompose en
$a_p\times 10^p+a_{p-1}\times 10^{p-1}+\cdots+a_2\times 10^2+a_1\times 10^1+a_0\times 10^0$. Selon la parité de p, on a :

- Si p impair :
  $n=[a_p(11k_p-1)+a_{p-2}(11k_{p-2}-1)+\cdots+a_1(11-1)]+[a_{p-1}(11k_p+1)+\cdots+a_0(0+1)]$,
  soit en réarrangeant cette somme et sans détailler l'entier K,
  $n=11K+[a_p\times(-1)+a_{p-2}\times(-1)+\cdots+a_1\times(-1)]+[a_{p-1}\times(+1)+\dots a_0\times(+1)]$.
  Finalement, $n=11K-[(a_p+a_{p-2}+\cdots+a_1)-(a_{p-1}+\cdots+a_2+a_0)]=11K-S$ et c'est donc cette somme alternée S qu'il suffit d'examiner pour la divisibilité par 11 de n.
- Si p pair, le principe est le même :
  $n=[a_p(11k_p+1)+\cdots+a_2(9\times 11+1)+a_0(0+1)]+[a_{p-1}(11k_p-1)+\cdots+a_1(11-1)]$ et donc $n=11K+[(a_p+a_{p-2}+\cdots+a_2+a_0)-(a_{p-1}+\cdots+a_1)]=11K-S$ avec la somme alternée S à tester.

1.b. Division euclidienne

Définition 6.2 Il existe un unique entier $n\in\mathbb{Z}$ tel que $\forall x\in\mathbb{R}, n\leqslant x<n+1$.
Cet entier est appelé partie entière de x et noté $E(x)$ ou encore $\lfloor x\rfloor$, en anglais *floor* (plancher).

Démonstration Cette existence unique est la conséquence de l'axiome *toute partie de* $\mathbb{N}$ *admet un plus petit élément.*
En effet, en supposant $x>0$, l'ensemble des entiers supérieurs strictement à x admet un plus petit élément, noté $n+1$, tel que $x<n+1$ et l'entier précédent $n+1$ n'étant pas dans cette partie est tel que $x\geqslant n$, d'où l'encadrement $n\leqslant x<n+1$.
Pour un réel négatif, la partie entière continue d'être l'unique entier relatif n tel que $n\leqslant x<n+1$.

Remarque : la partie entière (en Python `math.floor(-2.6)`) ne doit pas être confondue avec la troncature à l'unité de x (en Python `int(-2.6)`) ou avec l'arrondi à l'unité de x (en Python `round(-2.6,0)`), même s'il y a parfois coïncidence entre ces notions.
$E(-2,6) = -3$ alors que $E(2,6) = 2$, mais la troncature de $-2,6$ est -2 et l'arrondi -3 alors que la troncature de $2,6$ est 2 et l'arrondi est 3.

DÉFINITION 6.3 soient a et $b > 0$ deux entiers.
Il existe un unique couple (q, r) d'entiers tels que $\begin{cases} a = bq + r \\ 0 \leqslant r < b \end{cases}$
On dit alors que q est le quotient de la division euclidienne de a par b, r en étant le reste.

DÉMONSTRATION ➠ Existence :
En notant q la partie entière du quotient $\frac{a}{b}$ ($q = \lfloor \frac{a}{b} \rfloor$), celle-ci vérifie, d'après la propriété précédente, l'encadrement $q \leqslant \frac{a}{b} < q + 1$. En multipliant par $b \neq 0$: $bq \leqslant a < b(q + 1)$, puis en enlevant b :
$0 \leqslant a - bq < b$. Si on pose $r = a - bq$, on a donc bien $0 \leqslant r < b$ et $a = bq + r$.

➠ Unicité :
Supposons qu'il existe deux couples (q, r) et (q', r') d'entiers tels que $\begin{cases} a = bq + r \\ 0 \leqslant r < b \end{cases}$ et $\begin{cases} a = bq' + r' \\ 0 \leqslant r' < b \end{cases}$
Cela implique que $r - r' = b(q - q')$ et donc $b|(r - r')$.
Si $r - r' \neq 0$ alors $-(r - r') \leqslant b \leqslant r - r'$ mais cela est impossible car d'autre part,
comme $0 \leqslant r < b$ et $-b \leqslant -r' < 0$, on doit avoir $-b < r - r' < b$.
On en déduit que $r - r' = 0$ et, comme $b \neq 0$, que $q - q' = 0$. Finalement $r = r'$ et $q = q'$.

Remarques et exemples :

- La division euclidienne dans $\mathbb{N}$ est commencée à l'école primaire et poursuivie au collège.
  L'écriture $177 = 11 \times 15 + 12$ correspond à l'écriture en ligne de la division euclidienne de 177 par 15 (quotient 11, reste $12 < 15$) mais pas à la division de 177 par 11 car le reste supposé est trop grand ($12 > 11$). Pour en déduire quotient de 177 par 12, il suffit de lui enlever le diviseur et de l'ajouter une fois du quotient : $177 = 11 \times (15 + 1) + (12 - 11) = 11 \times 16 + 1$.
  Du coup, le quotient de 177 par 11 est 16 et le reste 1.
- De $177 = 11 \times 15 + 12$, on peut déduire le résultat de la division euclidienne de -177 par 15 car alors $-177 = (-11) \times 15 + (-12)$. Comme le reste est trop petit ($-12 < 0$), il suffit de lui ajouter le diviseur et de le retrancher une fois du quotient :
  $-177 = (-11 - 1) \times 15 + (-12 + 15) = (-12) \times 15 + 3$.
  Du coup, le quotient de -177 par 15 est -12 et le reste 3.
- On peut, au passage, élargir la division euclidienne au cas d'un diviseur négatif en précisant que le reste doit toujours vérifier $0 \leqslant r < |b|$). La division euclidienne de -177 par -15 par exemple s'effectue à partir de l'opposée de $177 = 11 \times 15 + 12$: $-177 = 11 \times (-15) + (-12)$ et comme $-12 < 0$, on enlève le diviseur et on ajoute une fois au quotient :
  $-177 = (11 + 1) \times (-15) + (-12 - (-15)) = 12 \times (-15) + 3$.
  Du coup, le quotient de -177 par -15 est 12 et le reste 3.
- En Python le quotient de la division euclidienne de a par b est obtenu en tapant `a//b` et le reste s'obtient en tapant `a%b`. La fonction division euclidienne peut s'écrire en une ligne : `def de(a,b):return(a//b,a%b)`. Ainsi, en tapant `de(-177,15)` on obtient `(-12,3)` mais attention, en tapant `de(-177,-15)` on obtient `(11,-12)`, ce qui ne convient pas car l'opérateur `%` retourne un nombre du même signe que le diviseur. Pour que notre fonction `de()` convienne à toutes les situations, il faut tenir compte de ce comportement.

```
def de(a,b):
    if b>0 : return(a//b,a%b)
    return(a//b+1,a%b-b)
```

```
>>> de(-177,15)
(-12, 3)
>>> de(-177,-15)
(12, 3)
>>> de(177,-15)
(-11, 12)
>>> de(177,15)
(11, 12)
```

2. Congruences

2.a. Congruence modulo n

Définition 6.4 Soient a, b et $n \geqslant 2$ trois entiers.
On dit que a est congru à b modulo n lorsque a et b ont le même reste dans la division euclidienne par n, et on note $a \equiv b[mod\ \ n]$, $a \equiv b[n]$ ou plus simplement $a = b[n]$.

Remarques :

- On est familier avec les relations de congruences depuis toujours avec les heures qui sont remises à zéro tous les jours (le modulo est 24) ou toutes les demi-journées (le modulo est 12). Par exemple à 20h, il est 8h (de l'après-midi) car $20 \equiv 8[12]$ et dix heures plus tard il ne sera pas 30h mais 6h (du matin) car $30 \equiv 6[24]$. Une autre application des congruences a été étudiée en trigonométrie où les lignes trigonométriques d'un angle sont partagées avec toutes les autres valeurs qui lui sont congrues modulo 2π. Par exemple $\frac{5\pi}{2} \equiv \frac{\pi}{2} \equiv \frac{-3\pi}{2}[2\pi]$.
- On confond souvent, ne serait-ce que dans la notation, l'égalité avec la relation de congruence car elles partagent les propriétés évidentes suivantes valables pour tout $n \geqslant 2$:
  — $\forall a \in \mathbb{Z},\ a \equiv a[n]$
  — $\forall (a,b) \in \mathbb{Z}^2,\ a \equiv b[n] \iff b \equiv a[n]$
  — $\forall (a,b,c) \in \mathbb{Z}^3,\ a \equiv b[n]$ et $b \equiv c[n] \implies a \equiv c[n]$
  Dans la suite les congruences sont notées avec $=$, la présence du modulo faisant la différence.
- La congruence modulo 0 ou 1 n'est pas utilisée ($a = b[0] \iff a = b$ et $\forall (a,b) \in \mathbb{Z}^2, a = b[1]$).

Propriété 6.2 a, b et $n \geqslant 2$ étant trois entiers, $a = b[n] \iff a - b$ est un multiple de n.

Démonstration Divisions euclidiennes de a et b par n : $\begin{cases} a = nq + r \\ 0 \leqslant r < n \end{cases}$ et $\begin{cases} b = nq' + r' \\ 0 \leqslant r' < n \end{cases}$.
On en déduit : $a - b = nq + r - (nq' + r') = n(q - q') + r - r'$ et,
en additionnant $-n < -r' \leqslant 0$ et $0 \leqslant r < n$, on obtient $-n < r - r' < n$.

- Si $a = b[n]$ alors $r = r'$ (par définition) et donc $r - r' = 0$ d'où $a - b = n(q - q')$, c'est-à-dire $a - b$ est un multiple de n.
- Si $a - b$ est un multiple de n, il existe $k \in \mathbb{Z}$ tel que $a - b = kn$ et donc $r - r' = a - b - n(q - q') = n(k - q + q')$, $r - r'$ est un multiple de n.
  Le seul multiple de n compris entre $-n$ et n étant 0, $r - r' = 0 \iff r = r'$ d'où $a = b[n]$.

Remarques :

- On pourra identifier, modulo n, un entier a avec son reste r dans la division euclidienne par n puisque $a = nq + r$ avec $0 \leqslant r < n$ implique que $a - r = nq$ et donc que $a = r[n]$.
  Exemples : $2023 = 11 \times 183 + 10$ donc $2023 = 10[11]$; $2023 = 10 \times 202 + 3$ donc $2023 = 3[10]$;
  Dire que a est divisible par n, ce qu'on note déjà $n|a$, s'écrit aussi $a = 0[n]$.
- Autres exemples :
  — $27 = 2[5]$ car $27 - 2 = 25 = 5k$ (un multiple de 5)
  — $-27 = 3[5]$ car $-27 - 3 = -30 = 5k'$
  — $(a+b)^2 = a^2 + b^2[2]$ car $(a+b)^2 - (a^2 + b^2) = 2ab = 2k''$
  — $(2n+1)^2 = 1[4]$ (autrement dit le carré d'un nombre impair est de la forme $4k+1$) car $(2n+1)^2 - 1 = (2n+1-1)(2n+1+1) = 2n(2n+2) = 4n(n+1) = 4k'''$
- Un entier $n \geqslant 2$ étant donné, un entier quelconque $a \in \mathbb{Z}$ appartient nécessairement à l'une des n classes d'équivalences : $a = 0[n], a = 1[n], \ldots, a = n-1[n]$. Tous les entiers $a \in \mathbb{Z}$ se rangent ainsi dans l'une ou l'autre des quatre classes : $a = 0[4]$, $a = 1[4]$, $a = 2[4]$, $a = 3[4]$ ce que l'on peut noter avec $\oplus$, le « ou exclusif » $a = 4k \oplus a = 4k + 1 \oplus a = 4k + 2 \oplus a = 4k + 3$.

Propriété 6.3 (Compatibilités avec les opérations) a, b, c, d et $n \geqslant 2$ étant cinq entiers,
- Avec l'addition : $a = b[n]$ et $c = d[n] \Longrightarrow a + c = b + d[n]$
- Avec la multiplication : $a = b[n]$ et $c = d[n] \Longrightarrow a \times c = b \times d[n]$
- Avec les puissances : $\forall p \in \mathbb{N}^*, a = b[n] \Longrightarrow a^p = b^p[n]$

Démonstration D'après la propriété 6.2 :
$a = b[n]$ et $c = d[n] \Longleftrightarrow \exists (k, k') \in \mathbb{Z}^2, a - b = kn$ et $c - d = k'n$
Par conséquent, $(a + c) - (b + d) = (a - b) + (c - d) = kn + k'n = (k + k')n = k''n$ avec $k'' = k + k' \in \mathbb{Z}$ d'où $a + c = b + d[n]$.
De même, $(a \times c) - (b \times d) = (a - b) \times c + (c - d) \times b = kn \times c + k'n \times b = (kc + k'b)n = k'''n$ avec $k''' = kc + k'b \in \mathbb{Z}$ d'où $a \times c = b \times d[n]$.
La troisième compatibilité est obtenue par récurrence sur $p > 0$ avec la compatibilité de $\times$.

Exemple 64 – Les utilisations de ces propriétés sont nombreuses et variées, donnons-en quatre :
➲ Déterminons le reste de la division euclidienne de 100^{50} par 11.
Comme $100 = 99 + 1 = 9 \times 11 + 1$, $100 = 1[11]$ et donc $100^{50} = 1^{50} = 1[11]$, le reste est 1.
Déterminons le reste de la division euclidienne de 100^{50} par 7.
Comme $100 = 77 + 23$, $77 = 0[7]$ et $23 = 21 + 2 = 2[7]$, on a $100 = 2[7]$ donc $100^{50} = 2^{50}[7]$.
Comme, de plus, $2^3 = 8 = 7 + 1 = 1[7]$, $2^{50} = 2^{3\times16+2} = (2^3)^{16} \times 2^2 = 1^{16} \times 4 = 4[7]$, le reste est 4.

➲ Montrons la propriété $\forall n \in \mathbb{N}, 3|\,(n(n + 1)(2n + 1))$ à l'aide d'un tableau de congruence modulo 3.

$n = \ldots[3]$	0	1	2
$n + 1 = \ldots[3]$	1	2	0
$2n + 1 = \ldots[3]$	1	0	2
$n(n + 1)(2n + 1) = \ldots[3]$	0	0	0

Comme on remarque que, dans tous les cas $n(n + 1)(2n + 1) = 0[3]$, on en conclut que $n(n + 1)(2n + 1)$ est toujours divisible par 3, soit $3|\,(n(n + 1)(2n + 1))$.

➲ Explorons les propriétés des puissances en traçant le tableau des puissances modulo 5 :

a	a^1	a^2	a^3	a^4	a^5	a^6	a^7	a^8
1	1	1	1	1	1	1	1	1
2	2	4	3	1	2	4	3	1
3	3	4	2	1	3	4	2	1
4	4	1	4	1	4	1	4	1
5	0	0	0	0	0	0	0	0

On remarque une périodicité de période 4 lorsque $a = 2$ ou $3[5]$, de période 2 lorsque $a = 4[5]$.
- Comme $2^4 = 1[5]$, on a :
  $2^1 = 2^5 = \cdots = 2^{4k+1} = 2[5]$; $2^2 = 2^6 = \cdots = 2^{4k+2} = 4[5]$; $2^3 = 2^7 = \cdots = 2^{4k+3} = 3[5]$
- Comme $3^4 = 1[5]$, on a :
  $3^1 = 3^5 = \cdots = 3^{4k+1} = 3[5]$; $3^2 = 3^6 = \cdots = 3^{4k+2} = 4[5]$; $3^3 = 3^7 = \cdots = 3^{4k+3} = 2[5]$
- Comme $4^2 = 1[5]$, on a :
  $4^1 = 4^3 = \cdots = 4^{2k+1} = 1[5]$; $4^2 = 4^4 = \cdots = 4^{2k} = 1[5]$

La relation de congruence est-elle compatible avec l'exponentiation ?
Autrement dit $\forall (p, q) \in \mathbb{Z}^2, \forall a \in \mathbb{N}, p = q[n] \Longrightarrow a^p = a^q[n]$ est-elle vraie pour tout entier $n > 1$?
La réponse est non, la relation de congruence n'est pas compatible avec l'exponentiation.
Exhibons en guise de contre-exemple le cas $n = 5$, $p = 1$ et $q = 6$ pour lequel on a bien $1 = 6[5]$.
- Pour $a = 1$ ou $5[5]$ on a bien $1^1 = 1^6 = 1[5]$ et $5^1 = 5^6 = 0[5]$ mais ce n'est pas étonnant car toutes les puissances de $a = 1[5]$ valent 1 et toutes les puissances de $a = 5[5]$ valent 0
- Pour $a = 2$ ou 3 ou $4[5]$ la propriété est cependant fausse car
  $(2^1 = 2) \neq (2^6 = 4)[5]$ et $(3^1 = 3) \neq (3^6 = 4)[5]$ et $(4^1 = 4) \neq (4^6 = 1)[5]$.

➲ La compatibilité de la relation de congruence avec $+$ se simplifie en $a = b[n] \Longrightarrow a + c = b + c[n]$.
La réciproque de cette propriété est vraie : $a + c = b + c[n] \Longrightarrow a = b[n]$.
Cela se prouve facilement : $a + c = b + c[n] \Longrightarrow (a + c) - (b + c) = 0[n] \Longrightarrow a - b = 0[n] \Longrightarrow a = b[n]$.
Ce qui rend la soustraction compatible avec la relation de congruence.
Une conséquence : on peut résoudre une équation du type $x + a = b[n]$ en soustrayant a aux deux membres. Par exemple $x + 5 = 3[7] \Longleftrightarrow x = 3 - 5 = -2 \Longleftrightarrow x = -2 + 7 = 5[7]$.

La compatibilité de la relation de congruence avec $\times$ se simplifie en $a = b[n] \Longrightarrow a \times c = b \times c[n]$.
Mais la réciproque est fausse : $5 \times 7 = 5 \times 9[5]$ mais $7 \neq 9[5]$ ou $2 \times 5 = 2 \times 1[8]$ mais $5 \neq 1[8]$.
Ce qui rend la division incompatible avec la relation de congruence.
On ne peut donc pas résoudre une équation du type $ax = b[n]$ en divisant par a les deux membres.
Exemple : $5x = 45[5]$, en divisant par 5 les deux membres on trouve $x = 9 = 4[5]$ alors que $5x = 45[5]$ est toujours vraie. Pour résoudre ce type d'équation, on peut utiliser un tableau de congruence.
Supposons qu'on cherche à résoudre $3x = 1[5]$:

$x = \dots[5]$	0	1	2	3	4
$3x = \dots[5]$	0	3	1	4	2

On en déduit que les solutions s'écrivent $x = 2[5]$.
De même pour résoudre une équation plus compliquée comme $x^2 + 3x - 2 = 0[5]$
Supposons qu'on cherche à résoudre $3x = 1[5]$:

$x = \dots[5]$	0	1	2	3	4
$x^2 = \dots[5]$	0	1	4	4	1
$3x = \dots[5]$	0	3	1	4	2
$x^2 + 3x - 2 = \dots[5]$	-2=3	2	3	6=1	1

On en déduit que cette équation n'a pas de solutions dans $\mathbb{Z}$.
Par contre l'équation $x^2 + 3x - 2 = p[5]$ a des solutions pour $p \in \{1, 2, 3\}[5]$.

3. PGCD

Définition 6.5 (PGCD et PPCM) a et b étant deux entiers relatifs non nuls, le plus grand diviseur commun de a et de b est un entier compris entre 1 et $min(|a|, |b|)$. On note ce nombre $PGCD(a, b)$ (Plus Grand Commun Diviseur, en anglais *Greatest Common Divisor*).
Le plus petit multiple commun strictement positif de a et de b est un entier compris entre $max(|a|, |b|)$ et $|ab|$. On note ce nombre $PPCM(a, b)$ (Plus Petit Commun Multiple, en anglais *Least Common Multiple*).

Démonstration L'ensemble des diviseurs de a et de $-a$, et donc aussi de $|a|$, est $Div(a)$.
Cet ensemble, comme d'ailleurs aussi $Div(b)$ n'est pas vide puisqu'il contient toujours 1.
Par conséquent l'ensemble $Div(a) \cap Div(b)$ n'est pas vide car il contient au moins 1.
Le plus grand élément de $Div(a)$ est $|a|$, de même le plus grand élément de $Div(b)$ est $|b|$.
On en déduit que le plus grand élément de $Div(a) \cap Div(b)$ est inférieur ou égal à $|a|$ et à $|b|$, cet ensemble est donc majoré par $min(|a|, |b|)$. Cet ensemble des diviseurs communs de a et de b, $Div(a) \cap Div(b)$, est un ensemble borné d'entiers. Son plus grand élément existe, c'est lui qui est appelé $PGCD(a, b)$.

L'ensemble des multiples strictement positifs de a (respectivement de b) est un ensemble d'entiers dont le plus petit élément est $|a|$ (resp. $|b|$). L'ensemble des multiples communs strictement positifs de a et de b est donc un ensemble minoré par $max(|a|, |b|)$. Cet ensemble n'est pas vide car il contient toujours $|ab|$, par conséquent son plus petit élément existe, c'est lui qui est appelé $PPCM(a, b)$.

Remarques :

- On peut toujours déterminer le PGCD de deux entiers à l'aide de la liste des diviseurs de chacun de ces nombres. C'est la méthode naïve, communément utilisée au Collège pour se familiariser avec la notion dans le but, notamment, de simplifier une fraction.
  Donnons un exemple : $Div(24) \cap \mathbb{N} = \{1, 2, 3, 4, 6, 8, 12, 24\}$ (je ne donne que les diviseurs positifs) et $Div(60) \cap \mathbb{N} = \{1, 2, 3, 4, 5, 6, 10, 12, 20, 30, 60\}$ d'où
  $Div(24) \cap Div(60) \cap \mathbb{N} = \{1, 2, 3, 4, 6, 12\}$. On en déduit que $PGCD(24, 60) = 12$; cela permet, notamment, de simplifier la fraction $\frac{24}{60} = \frac{12\times 2}{12\times 5} = \frac{2}{5}$.
- De même on peut déterminer le PPCM de deux entiers par une méthode naïve en déterminant une liste des multiples positifs des deux nombres, le premier qui se trouve dans les deux listes étant le PPCM cherché. Par exemple les multiples de 24 sont $24, 48, 72, 96, 120, \dots$ et les multiples de 60 sont $60, 120, \dots$ donc $PPCM(24, 60) = 120$. L'utilisation classique du PPCM est la détermination d'un dénominateur commun optimal pour l'addition des fractions :
  $\frac{5}{24} + \frac{7}{60} = \frac{5\times 5}{24\times 5} + \frac{7\times 2}{60\times 2} = \frac{25}{120} + \frac{14}{120} = \frac{25+14}{120} = \frac{39}{120}$.

PROPRIÉTÉ 6.4 (PGCD) Soient a et b deux entiers non nuls quelconques.

- $PGCD(a, b) = PGCD(|a|, |b|)$
- $PGCD(a, b) = PGCD(b, a)$
- $PGCD(a, b) \geqslant 1$
- $a|b \Longrightarrow PGCD(a, b) = |a|$
- $PGCD(a, b) = PGCD(a - b, b)$
- $a = bq + r$ avec $0 \leqslant r < b \Longrightarrow PGCD(a, b) = PGCD(b, r)$

DÉMONSTRATION ➲ Les trois premières propriétés sont évidentes :
Comme $Div(a) = Div(-a) = Div(|a|)$ et $Div(b) = Div(-b) = Div(|b|)$, on a
$Div(a) \cap Div(b) = Div(|a|) \cap Div(|b|)$ d'où $PGCD(a, b) = PGCD(|a|, |b|)$.
De même, comme $Div(a) \cap Div(b) = Div(b) \cap Div(a)$, on a $PGCD(a, b) = PGCD(b, a)$.
$PGCD(a, b) \geqslant 1$ car parmi les diviseurs communs strictement positifs, le plus petit est 1.

➲ Pour la 4e, montrons d'abord que $a|b \Longrightarrow Div(a) \cap Div(b) = Div(a)$:

- Un diviseur quelconque d de a vérifie $d|a$, or $a|b$ et la relation $|$ est transitive donc $d|b$.
  Par conséquent $a|b \Longrightarrow Div(a) \subset Div(b)$
- De $Div(a) \subset Div(b)$ on déduit que $Div(a) \cap Div(b) = Div(a)$

Par conséquent, lorsque $a|b$ le plus grand élément de $Div(a) \cap Div(b)$ est donc le plus grand élément de $Div(a)$ c'est-à-dire $|a|$, d'où $PGCD(a, b) = |a|$.

➲ Montrons $PGCD(a, b) = PGCD(a - b, b)$ par double inclusion :

- Si $d \in Div(a) \cap Div(b)$ alors $d|a$ et $d|b$. D'après la propriété 6.1, d divise toute combinaison linéaire de a et b, donc $d|(a - b)$ et par conséquent $d \in Div(a - b) \cap Div(b)$ d'où
  $Div(a) \cap Div(b) \subset Div(a - b) \cap Div(b)$.
- Si $d \in Div(a - b) \cap Div(b)$ alors $d|(a - b)$ et $d|b$. D'après la propriété 6.1, d divise $(a - b) + b = a$, donc $d|a$ et par conséquent $d \in Div(a) \cap Div(b)$ d'où
  $Div(a - b) \cap Div(b) \subset Div(a) \cap Div(b)$.

On en déduit que $Div(a) \cap Div(b) = Div(a - b) \cap Div(b)$.
Le plus grand élément de $Div(a) \cap Div(b)$ est donc aussi celui de $Div(a - b) \cap Div(b)$.
De la même façon, lorsque $a = bq + r$ avec $0 \leqslant r < b$, on a $PGCD(a, b) = PGCD(r, b)$.

- Dans un sens, d divise toute combinaison linéaire de a et b, donc $d|(a - bq)$ c'est-à-dire $d|r$.
- Dans l'autre sens, d divise toute combinaison linéaire de r et b, donc $d|(r + bq)$ c'est-à-dire $d|a$.

Dans le cas où $r = 0$, on a $a = bq$ c'est-à-dire $b|a$ et, comme on l'a vu, dans ce cas $PGCD(a, b) = |b|$.

Remarques :

- Deux cas particuliers de la 4e propriété : $PGCD(a, a) = |a|$; $PGCD(a, 1) = 1$.
  Par ailleurs $PGCD(0, a) = |a|$ car $Div(0) = \mathbb{Z}$ et donc $Div(a) \cap Div(0) = Div(a)$.

- Les propriétés des deux dernières lignes sont les « lemme d'Euclide » car elles sont les préliminaires obligés à une justification de l'algorithme d'Euclide (propriété suivante).
Lorsqu'on applique ces propriétés à des entiers naturels, le procédé par division revient à itérer le procédé par soustraction jusqu'à l'obtention d'un résidu inférieur à b.
Par soustractions $PGCD(50,15) = PGCD(35,15) = PGCD(20,15) = PGCD(5,15)$ mais $50 = 3 \times 15 + 5$, par division on trouve directement $PGCD(50,15) = PGCD(5,15)$.
Pour $PGCD(5,15)$, comme $5|15$ on a $PGCD(5,15) = 5$. Finalement $PGCD(50,15) = 5$.
- Comme $PGCD(a,b) = PGCD(|a|,|b|)$, nous ne considérerons désormais que des entiers naturels (positifs) en transposant si nécessaire dans $\mathbb{Z}$ les résultats obtenus dans $\mathbb{N}$.

Propriété 6.5 (Algorithme d'Euclide) Soient a et b deux entiers strictement positifs.
On applique successivement la division euclidienne de a par b (elle conduit à un reste $r = a - bq$) puis on recommence en renommant les variables $\begin{cases} b \text{ devient } a \\ r \text{ devient } b \end{cases}$, jusqu'à obtenir un reste nul.
Le nombre cherché ($PGCD(a,b)$) est égal au dernier reste non nul obtenu.

Démonstration Il suffit d'appliquer successivement la propriété
$a = bq + r$ avec $0 \leqslant r < b \Longrightarrow PGCD(a,b) = PGCD(b,r)$ jusqu'à l'obtention d'un reste nul.
Le premier reste r_1 est tel que $0 \leqslant r_1 < b$ et $PGCD(a,b) = PGCD(b,r_1)$.
Le 2e reste r_2 est tel que $0 \leqslant r_2 < r_1$ et $PGCD(b,r_1) = PGCD(r_1,r_2)$, etc.
La suite des restes est donc décroissante et minorée par 0 ($0 \leqslant \cdots < r_3 < r_2 < r_1$).
Elle admet donc un plus petit élément non nul, r_k, qui est le PGCD cherché puisque
$PGCD(a,b) = PGCD(b,r_1) = PGCD(r_1,r_2) = \cdots = PGCD(r_{k+1},r_k) = PGCD(r_k,0) = r_k$.

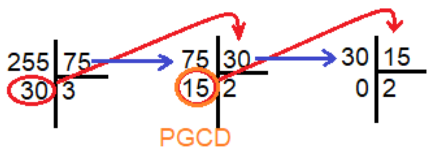
Remarques : Ce procédé est décrit dans le livre 10 des Éléments d'Euclide, écrit il y a 23 siècles.
Lorsqu'on l'applique à la main, cela conduit à effectuer des divisions euclidiennes en chaine.
On peut aussi utiliser un tableur ou un programme informatique comme celui ci-dessous en Python.
On peut encore illustrer le procédé graphiquement en enlevant des carrés dans une surface rectangulaire, maximisant à chaque étape la taille du carré à enlever (cette méthode revient à appliquer des soustractions successives).

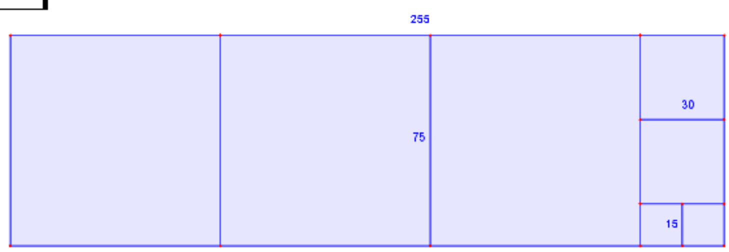
B3 =SI(B2<>0;MOD(B1;B2);0)

	A	B	C	D
1	1er nombre	255	229596	123456789
2	2eme nombre	75	4898865	1234567
3		30	229596	89
4		15	77349	48
5		0	74898	41
6		0	2451	7
7		0	1368	6
8		0	1083	1
9		0	285	0
10		0	228	0
11		0	57	0
12		0	0	0

```python
def pgcd(a,b):
    while b!=0:
        r=a%b
        a=b
        b=r
    return a

#a,b=255,75
#a,b=229596,4898865
a,b=123456789,1234567
print("le pgcd de {} et {} est {}".format(a,b,pgcd(a,b)))
```

```
le pgcd de 255 et 75 est 15
le pgcd de 229596 et 4898865 est 57
le pgcd de 123456789 et 1234567 est 1
```





Contrairement à l'algorithme par soustractions qui suppose $a > b$ pour calculer $PGCD(a,b)$, avec l'algorithme par divisions il n'est pas nécessaire de s'assurer que $a > b$ car, dans le cas contraire, $a = 0 \times b + a$ avec $0 \leqslant a < b \Longrightarrow PGCD(a,b) = PGCD(b,a)$ et on continue avec, cette fois $b > a$.
Ce cas de figure s'est présenté dans le calcul de $PGCD(229\,596, 4\,898\,865)$.

DÉFINITION 6.6 (NOMBRES PREMIERS ENTRE EUX) Deux entiers non nuls a et b n'ayant que 1 comme diviseur commun, c'est-à-dire tels que $PGCD(a, b) = 1$, sont dits « premiers entre eux ».

Remarques :

- Les fractions $\frac{a}{b}$ et $\frac{b}{a}$ sont dites « irréductibles » lorsqu'elles ne sont pas simplifiables, autrement dit si les nombres a et b sont premiers entre eux, c'est-à-dire si $PGCD(a, b) = 1$.
- Ne pas confondre nombres « premiers entre eux » et nombres « premiers ». En particulier ne pas croire que a et b doivent être premiers pour que les nombres a et b soient premiers entre eux. Par exemple 8 et 9 ne sont pas premiers (8 est divisible par 2 et 4, 9 l'est par 3) mais cependant ils sont premiers entre eux car leur seul diviseur commun est 1. Inversement, par contre, deux nombres premiers différents sont toujours premiers entre eux.

PROPRIÉTÉ 6.6 (COMPLÉMENT) Soient a et b deux entiers strictement positifs.

- Pour tout entier $c \in \mathbb{Z}^*$ on a $PGCD(ca, cb) = |c| \times PGCD(a, b)$
- $PGCD(a, b) = d \iff \exists (p, q) \in \mathbb{N}^*, \begin{cases} PGCD(p, q) = 1 \\ a = dp \\ b = dq \end{cases}$

DÉMONSTRATION ➲Considérons un entier naturel non nul c.
En préliminaire remarquons que puisque $a = bq + r$ avec $0 \leqslant r < b$, on a $PGCD(a, b) = PGCD(r, b)$ alors on a aussi $ca = (cb)q + (cr)$ avec $0 \leqslant cr < cb$ et $PGCD(ca, cb) = PGCD(cr, cb)$. On peut donc multiplier par un entiers $c > 0$ tous les nombres impliqués dans le lemme d'Euclide.

On sait que l'algorithme d'Euclide construit la suite décroissante des restes (r_n) :
$PGCD(a, b) = PGCD(b, r_1) = PGCD(r_1, r_2) = \cdots = PGCD(r_{k+1}, r_k) = PGCD(r_k, 0) = r_k$.
Multiplions par $c > 0$ tous les nombres impliqués dans ces égalités :
$PGCD(ca, cb) = PGCD(cb, cr_1) = \cdots = PGCD(cr_{k+1}, cr_k) = PGCD(cr_k, 0) = cr_k$.
On obtient bien $PGCD(ca, cb) = cr_k = c \times PGCD(a, b)$.

Pour $c < 0$, comme $PGCD(a, b) = PGCD(|a|, |b|)$, on en déduit que
$PGCD(ca, cb) = PGCD(|ca|, |cb|) = PGCD(|c|a, |c|b) = |c| \times PGCD(a, b)$.

➲ Montrons que $PGCD(a, b) = d \Longrightarrow \exists (p, q) \in \mathbb{N}^*, \begin{cases} PGCD(p, q) = 1 \\ a = dp \\ b = dq \end{cases}$

Si $PGCD(a, b) = d$ alors $d|a$ et $d|b$, autrement dit $\exists (p, q) \in \mathbb{N}^*, \begin{cases} a = dp \\ b = dq \end{cases}$
En utilisant la propriété précédente, on en déduit que
$d = PGCD(a, b) = PGCD(dp, dq) = d \times PGCD(p, q)$.
D'où p et q sont premiers entre eux car $d = d \times PGCD(p, q) \iff PGCD(p, q) = 1$.
Réciproquement,

$$\forall d \in \mathbb{N}^*, \exists (p, q) \in \mathbb{N}^*, \begin{cases} PGCD(p, q) = 1 \\ a = dp \\ b = dq \end{cases} \Longrightarrow PGCD(a, b) = PGCD(dp, dq) = d \times PGCD(p, q) = d$$

Exemples :

- On a calculé naïvement $PGCD(24, 60)$ plus haut, mais si on remarque que $24 = 12 \times 2$ et $60 = 12 \times 5$, comme 2 et 5 sont premiers entre eux, $PGCD(24, 60) = 12 \times PGCD(2, 5) = 12$.
- On souhaite trouver les couples de naturels (a, b) vérifiant $a + b = 108$ et $PGCD(a, b) = 12$. Il suffit de trouver les couples (p, q), p et q premiers entre eux, tels que $a = 12p$ et $b = 12q$.

Ces nombres vérifiant aussi $a+b=12(p+q)=108$, soit $p+q=\frac{108}{12}=9$, on en déduit les couples (p,q) qui conviennent : $(1,8),(2,7),(4,5),(5,4),(7,2)$ et $(8,1)$, et les solutions s'obtiennent en les multipliant par 12 : $(12,96),(24,84),(48,60),(60,48),(84,24)$ et $(96,12)$.

4. Théorèmes de Bézout et de Gauss

Propriété 6.7 (Théorème de Bézout) Soient a et b deux entiers strictement positifs.
$PGCD(a,b)=1 \iff \exists(u,v)\in\mathbb{Z}^2, au+bv=1$

Démonstration ⊃Le sens réciproque est évident.
En notant $PGCD(a,b)=d$, on sait que $d|a$ et $d|b$ ce qui implique, selon la propriété 6.1, $d|(au+bv)$ donc $d|1$. On en déduit $d=1$, les nombres a et b sont premiers entre eux.

⊃Sens direct : on suppose a et b premiers entre eux et on considère l'ensemble E des naturels non nuls de la forme $au+bv$ avec $(u,v)\in\mathbb{Z}^2$. L'ensemble E n'est pas vide car a et b en font partie, il admet donc un plus petit élément non nul, noté p.
p est tel que $\exists(u_0,v_0)\in\mathbb{Z}^2, p=au_0+bv_0$ et $\forall e\in E, e\geqslant p$.
Montrons que $p|a$: La division euclidienne de a par p s'écrit $a=pq+r$ avec $0\leqslant r<p$.
On en déduit que $r=a-pq=a-(au_0+bv_0)q=a(1-u_0p)+b(v_0q)$, soit $r=au_1+bv_1$ avec $u_1=1-u_0p$ et $v_1=v_0q$. Mais alors on doit avoir $r\in E$ et $r<p$ ce qui est impossible si $r\neq 0$, d'où on déduit que $r=0$ et donc $a=pq\iff p|a$.
De la même façon, on montre que $p|b$ et par conséquent $p|PGCD(a,b)$ or, par hypothèse, $PGCD(a,b)=1$. Conclusion $p|1$ donc $p=1$ et on a bien $\exists(u_0,v_0)\in\mathbb{Z}^2, au_0+bv_0=1$.

Remarques et exemples :

- Ce théorème est généralement appelé théorème de Bézout du nom du mathématicien Étienne Bézout (1730-1783) qui le généralisa aux polynômes mais on doit sa première démonstration à Claude-Gaspard Bachet dit de Méziriac (1581-1638) qui la publie dans *Problèmes plaisans et délectables qui se font par les nombres* en 1612. Pour ces raisons, il est souvent mentionné aujourd'hui comme le théorème de Bachet-Bézout.
- Si a et b ne sont pas premiers entre eux, il n'est pas possible de trouver des entiers u et v tels que $au+bv=1$. L'équation $4x+6y=1$ n'a pas de solution dans $\mathbb{Z}^2$.
- Lorsque $PGCD(a,b)=1$, le théorème assure de l'existence d'un couple d'entiers (u,v) mais n'en donne pas la valeur qui n'est d'ailleurs pas unique. Par exemple 2 et 3 étant premiers entre eux, on peut trouver plusieurs couples (u,v) tels que $2u+3v=1$: on a $-2+3=1$, donc déjà $(-1,1)$ est solution, mais on a encore $4-3=1$ et donc $(2,-1)$ est aussi solution.
- Détermination explicite d'un couple (u,v) satisfaisant $au+bv=1$ avec les congruences :
  $au=1-bv\implies au=1[b]$. Les u qui conviennent sont les inverses de a modulo b.
  Par exemple 2 et 3 étant premiers entre eux, quels entiers u vérifient $2u=1[3]$?
  $2\times 0=0[3]$, $2\times 1=2[3]$ et $2\times 2=4=1[3]$. Les u qui conviennent valent $2[3]$.
  Il suffit de choisir un nombre de la forme $u=3k+2$, soit $u\in\{\dots,-4,-1,2,5,\dots\}$.
  u étant choisi, une valeur entière de v se déduit nécessairement de l'égalité $v=\frac{1-au}{b}$.
  Pour $u=5$ on doit prendre $v=\frac{1-2\times 5}{3}=\frac{-9}{3}=-3$, et on a bien $2\times 5+3\times(-3)=1$.
- Une conséquence immédiate de ce théorème : Quels que soient les naturels non nuls a, b et c, si $PGCD(a,b)=1$ et $PGCD(a,c)=1$ alors $PGCD(a,bc)=1$.
  Pour le justifier, sachant que $au+bv=1$ et $au'+cv'=1$, il suffit d'arranger le produit $(au+bv)(au'+cv')=1$.
  Comme $(au+bv)(au'+cv')=a(uau'+ucv'+bvu')+bc(vv')$, on en déduit que $aU+bcV=1$ avec $U=uau'+ucv'+bvu'$ et $V=vv'$ et donc $PGCD(a,bc)=1$.
- Montrons que, quel que soit l'entier n, les nombres $3n+1$ et $2n+1$ sont premiers entre eux. Il suffit de remarquer que $3(2n+1)-2(3n+1)=1$ et d'appliquer le théorème réciproque. De même, on montre que $2n+3$ et $5n+7$ sont premiers entre eux car $5(2n+3)-2(5n+7)=1$.

Propriété 6.8 (Identité de Bézout) Soient a et b deux entiers strictement positifs.
$PGCD(a,b) = d \Longrightarrow \exists (u,v) \in \mathbb{Z}^2, au + bv = d$

Démonstration Cette propriété est une conséquence du théorème de Bézout.
Si $PGCD(a,b) = d$, d'après la propriété 6.6, $\exists (p,q) \in \mathbb{N}^*$, $\begin{cases} PGCD(p,q) = 1 \\ a = dp \\ b = dq \end{cases}$.
Appliquons le théorème de Bézout à p et q premiers entre eux : $\exists (u,v) \in \mathbb{Z}^2, pu + qv = 1$.
Multiplions par d cette égalité et remplaçons dp par a et dq par b, on obtient :
$d(pu + qv) = d \Longleftrightarrow (dp)u + (dq)v = d \Longleftrightarrow au + bv = d$.

Propriété 6.9 (Réciproque) Soient a et b deux entiers strictement positifs.
Si $d \in \mathbb{N}^*$ est tel que $d|a$, $d|b$ et $\exists (u,v) \in \mathbb{Z}^2, au + bv = d$, alors $PGCD(a,b) = d$

Démonstration Montrons $PGCD(a,b) = d$ en montrant que $PGCD(a,b) \leqslant d$ et $d \leqslant PGCD(a,b)$:

- $PGCD(a,b)$ est un diviseur de a et de b donc $PGCD(a,b)$ divise la combinaison linéaire $d = au + bv$. Puisque $PGCD(a,b)|d$, on a $PGCD(a,b) \leqslant d$.
- d est un diviseur commun à a et b, or par définition, $PGCD(a,b)$ est le plus grand des diviseurs communs à a et b, donc $d \leqslant PGCD(a,b)$.

Propriété 6.10 (Théorème de Gauss) Soient a, b et c trois entiers strictement positifs.
Si $a|(bc)$ et $PGCD(a,b) = 1$ alors $a|c$.

Démonstration Ce théorème est encore une conséquence du théorème de Bézout.
Appliquons-le à a et b qui sont premiers entre eux par hypothèse : $\exists (u,v) \in \mathbb{Z}^2, au + bv = 1$.
En multipliant par c l'égalité, on obtient $c(au + bc) = (ca)u + (cb)v = c$.
Or on sait que $a|(cb)$ (par hypothèse) et aussi $a|(ca)$ (évident) donc a divise la combinaison linéaire $(ca)u + (cb)v$ c'est-à-dire c.

Propriété 6.11 (Corollaire) Soient a, b et c trois entiers strictement positifs.
Si $a|c$, $b|c$ et $PGCD(a,b) = 1$ alors $ab|c$.

Démonstration Ce théorème est une conséquence du théorème de Gauss.
Puisque $a|c$ $\exists k \in \mathbb{N}^*, c = ka$, de même puisque $b|c$ $\exists k' \in \mathbb{N}^*, c = k'b$.
Comme $ka = k'b$, $a|(k'b)$ mais puisque $PGCD(a,b) = 1$, d'après le théorème, $a|k'$.
Puisque $a|k'$ $\exists k'' \in \mathbb{N}^*, k' = k''a$ et donc $c = k'b = k''ab$, soit $ab|c$.

Remarques :

- La condition $PGCD(a,b) = 1$ est nécessaire car $6|12$ et $4|12$ mais $6 \times 4 = 24$ ne divise pas 12 (car $PGCD(6,4) = 2 \neq 1$).
- La réciproque est fausse (comme souvent) $6 \times 9|54$ et, bien sûr, $6|54$ et $9|54$ mais $PGCD(6,9) = 3 \neq 1$.

Exemple 65 – Résoudre l'équation diophantienne[1] $ax + by = k \times PGCD(a,b)$ avec $k \in \mathbb{Z}^*$.

➲ Résolvons l'équation $45x + 16y = 1$.
La méthode des congruences a été expliquée plus haut : il s'agit, dans un 1er temps, de déterminer l'inverse de 45 modulo 16 car $45x + 16y = 1 \iff 45x = 1[16]$.
Pour trouver toutes les solutions, on doit examiner les 16 valeurs de $45x$ modulo 16 :

x	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
$45x[16]$	0	13	10	7	4	1	14	11	8	5	2	15	12	9	6	3

On constate que pour $x = 5[16]$, on a $45x = 1[16]$ d'où la solution générale $x = 16k + 5$ avec $k \in \mathbb{Z}$.
Pour y, il suffit d'utiliser $45x + 16y = 1 \iff y = \frac{1-45x}{16}$ qui donne un entier, selon le théorème de Bézout car $PGCD(45,16) = 1$. En remplaçant x par $16k + 5$, on obtient
$y = \frac{1-45(16k+5)}{16} = \frac{1-45(16k+5)}{16} = -45k - 14$.
On a donc $(x,y) = (16k + 5, -45k - 14)$ avec $k \in \mathbb{Z}$.
Une solution différente est obtenue pour chaque valeur de k :

- Pour $k = 0$ on obtient $(x,y) = (5,-14)$ et on vérifie $45 \times 5 + 16 \times (-14) = 1$
- Pour $k = -1$ on obtient $(x,y) = (-11,31)$ et on vérifie $45 \times (-11) + 16 \times 31 = 1$
- etc.

Remarque : c'est très long d'obtenir ce tableau de congruence, et encore ici il n'y avait que 16 valeurs possibles pour x modulo 16. Cette méthode se prête bien à une utilisation algorithmique (programme, tableur), mais pour un travail manuel on lui préfèrera la méthode classique qui exploite la suite des divisions euclidiennes obtenues lors de la détermination du PGCD avec l'algorithme d'Euclide :

- $45 = 2 \times 16 + 13$
- $16 = 1 \times 13 + 3$
- $13 = 4 \times 3 + 1$

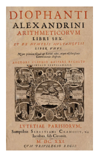
1 est le dernier reste non nul, ce qui prouve que $PGCD(45,16) = 1$.
En partant du bas, on remplace progressivement les restes, en remontant les égalités :

- $1 = 13 - 4 \times 3$
- $3 = 16 - 1 \times 13$ donc $1 = 13 - 4 \times (16 - 1 \times 13) = 13 \times 5 - 4 \times 16$
- $13 = 45 - 2 \times 16$ donc $1 = (45 - 2 \times 16) \times 5 - 4 \times 16 = 45 \times 5 + 16 \times (-14)$

On obtient ainsi la solution particulière $(x,y) = (5,-14)$ puisque $1 = 45 \times 5 + 16 \times (-14)$.
De là on trouve les solutions générales en écrivant :
$45x + 16y = 1 \iff 45x + 16y = 45 \times 5 + 16 \times (-14) \iff 45(x-5) = 16(-y-14)$
Comme $PGCD(45,16) = 1$, d'après le théorème de Gauss on a $16|(x-5)$ et $45|(-y-14)$.
Il existe donc deux entiers k et k' tels que $x - 5 = 16k$ et $-y - 14 = 45k'$.
L'équation transformée $(45(x-5) = 16(-y-14))$ s'écrit $45 \times 16k = 16 \times 45k' \iff k = k'$.
Conclusion : $x = 5 + 16k$ et $y = -14 - 45k$.

➲ Résolvons l'équation (E) : $45x + 16y = 12$.
On a résolu l'équation $45x + 16y = 1$ qui a pour solution particulière $(x,y) = (5,-14)$.
En multipliant par 12 l'égalité $45 \times 5 + 16 \times (-14) = 1$ on obtient
$45 \times 5 \times 12 + 16 \times (-14) \times 12 = 1 \times 12 \iff 45 \times 60 + 16 \times (-168) = 12$
Donc $(x,y) = (60,-168)$ est une solution particulière de (E).
De là, on trouve les solutions générales de (E) en écrivant :
$45x + 16y = 12 \iff 45x + 16y = 45 \times 60 + 16 \times (-168) \iff 45(x-60) = 16(-y-168)$
Comme $PGCD(45,16) = 1$, d'après le théorème de Gauss on a $16|(x-5)$ et $45|(-y-14)$.
Il existe donc deux entiers k et k' tels que $x - 60 = 16k$ et $-y - 168 = 45k'$.
L'équation transformée s'écrit $45 \times 16k = 16 \times 45k' \iff k = k'$.
Conclusion : $x = 60 + 16k$ et $y = -168 - 45k$.

1. Les équations diophantiennes sont des équations où les inconnues et les coefficients sont entiers. Elles sont nommées d'après Diophante d'Alexandrie qui publia *Les Arithmétiques* vers le IIe siècle, une collection de problèmes polynomiaux avec des entiers, un livre dont la traduction par Bachet (1621) en latin inspira grandement Pierre de Fermat.



5. Nombres premiers, Théorème de Fermat

DÉFINITION 6.7 (PREMIER) Un nombre entier $n \in \mathbb{N}$ est premier si et seulement si il n'admet que deux diviseurs distincts dans $\mathbb{N}$: 1 et lui-même.

Remarques :

- 1 n'est pas premier car il n'a qu'un seul diviseur dans $\mathbb{N}$ (lui-même qui vaut 1). Une raison plus convaincante, pour anticiper sur l'unicité de la décomposition en facteurs premiers : 1 n'est pas premier car sinon on aurait une infinité de décompositions car $1 = 1^2 = 1^3 = \dots$.
- 0 n'est pas premier car il a une infinité de diviseurs. 2 est donc le premier nombre premier ; c'est en même temps le seul qui soit pair (évidemment, car les autres nombres pairs l'admet comme diviseur). La suite des nombres premiers commence par $2, 3, 5, 7, 11, 13, \dots$
- Les nombres qui ne sont pas premiers sont dits composés. Ils admettent plus de deux diviseurs. Les nombres 0 et 1 (et -1 dans $\mathbb{Z}$) font exceptions pour des raisons différentes (1 n'a qu'un seul diviseur et 0 en a une infinité). 4 est le premier nombre composé car il admet trois diviseurs distincts dans $\mathbb{N}$ (1, 2 et 4).

PROPRIÉTÉ 6.12 Un nombre composé $n \geqslant 2$ admet au moins un diviseur premier : le plus petit de ses diviseurs supérieur strictement à 1 qui est un nombre p tel que $p^2 \geqslant n$.

DÉMONSTRATION Comme n est composé il admet au moins un plus petit diviseur $p > 1$ qui est premier car, sinon, ce diviseur p serait lui-même divisible par un diviseur p' tel que $1 < p' < p$ qui diviserait aussi p ce qui ne se peut pas.
Comme p divise n, il existe un entier q tel que $n = pq$. Comme n est composé, on doit avoir $q > 1$ et comme p est le plus petit diviseur de n, on a nécessairement $p \leqslant q$. L'encadrement $1 < p \leqslant q$ multiplié par $p > 0$ s'écrit $p < p^2 \leqslant pq \iff p < p^2 \leqslant n$.

Remarques :

- Cette propriété a pour corollaire la propriété suivante : Si un nombre $n \geqslant 2$ n'admet aucun diviseur p tel que $p^2 \leqslant n$ alors il est premier. Ceci justifie qu'un algorithme testant la primalité d'un nombre n ne teste que les nombres $p \leqslant \sqrt{n}$ (en appliquant la racine carrée à $p^2 \leqslant n$), en les prenant de 2 en 2 à partir de 2 (car seul 2 est un premier pair).
  Pour tester si 101 est premier, on examine les restes des divisions euclidienne successives de 101 par $23, 5$ et 7 seulement : $101 = 1[2]$, $101 = 2[3]$, $101 = 1[5]$, $101 = 3[7]$. Inutile d'aller plus loin car le prochain carré de nombre premier est $11^2 = 121 > 101$, tous les multiples de 11 susceptibles de décomposer 101 ont déjà été testés. Conclusion : 101 est un nombre premier.
- Le crible d'Érathostène[2] permet d'obtenir la liste des nombres premiers inférieurs à un entier N : on élimine successivement, les multiples de 2 en commençant par $2^2 = 4$, les multiples de 3 en commençant par $3^2 = 9$, les multiples de 5 en commençant par $5^2 = 25$, etc. jusqu'aux multiples de $n = \sqrt{N}$ en commençant par n^2.
  Dans les remarques de la propriété 6.15, la fonction `premiers()` implémente cet algorithme et renvoie la liste des nombres premiers inférieurs à un entier `n`.

PROPRIÉTÉ 6.13 L'ensemble $\mathbb{P}$ des nombres premiers est infini.
Autrement dit, il n'existe pas de plus grand nombre premier.

2. Ératosthène (IIIe siècle avant J.-C. était conservateur de la grande bibliothèque d'Alexandrie. Il est connu pour son estimation du rayon de la Terre.

Démonstration Euclide a prouvé cela par l'absurde, en supposant qu'il existe p, le plus grand des nombres premiers. Il construit ensuite le nombre $N = 2 \times 3 \times 5 \times \cdots \times p + 1$. Le reste de la division euclidienne de N par chacun des nombres premiers devra être égal à 1 car la définition de N équivaut à autant de divisions euclidiennes. N est donc premier car il n'est divisible par aucun des nombres premiers existants. Or ce nombre premier est plus grand que p qui, par hypothèse, est le plus grand. Cela étant contradictoire, il n'existe pas de plus grand nombre premier.

Propriété 6.14 (Divisibilités) Soient p un nombre premier et n un entier.

- Si p ne divise pas n alors $PGCD(p, n) = 1$
- $\forall (a_1, a_2, \ldots, a_k) \in \mathbb{Z}^k, p|(a_1 \times a_2 \times \cdots \times a_k) \Longleftrightarrow \exists i \in \{1, 2, \ldots, k\}, p|a_i$.
- $\forall (p_1, p_2, \ldots, p_k) \in \mathbb{P}^k, p|(p_1 \times p_2 \times \cdots \times p_k) \Longleftrightarrow \exists i \in \{1, 2, \ldots, k\}, p = p_i$.
- Soient $(p_1, p_2, \ldots, p_k) \in \mathbb{P}^k$, k nombres premiers distincts et $(a_1, a_2, \ldots, a_k) \in \mathbb{N}^{*k}$, Si $\forall i \in \{1, 2, \ldots, k\}, p_i^{a_i}|n$ alors $(p_1^{a_1} \times p_2^{a_2} \times \cdots \times p_k^{a_k})|n$.

Démonstration ➠ p étant premier n'a que deux diviseurs et si p ne divise pas n, le seul diviseur commun à p et n est 1 donc ils sont premiers entre eux.

➠ Montrons d'abord $\forall (a_1, a_2) \in \mathbb{Z}^2, p|(a_1 a_2) \Longrightarrow p|a_1$ ou $p|a_2$.
Deux cas se présentent :

- $p|a_1$. L'implication est alors évidente.
- p ne divise pas a_1 mais, dans ce cas, d'après le théorème de Gauss, $p|a_2$.

On généralise cette propriété à k facteurs par récurrence en posant $A = a_1 \times a_2 \times \cdots \times a_k$.
Par hypothèse $p|A \Longrightarrow p|a_1$ ou $p|a_2$ ou $\ldots p|a_k$ et si $p|(A \times a_{k+1})$ alors, par application de la propriété précédente, $p|A$ ou $p|a_{k+1}$ donc $p|a_1$ ou $p|a_2$ ou $\ldots p|a_k$ ou $p|a_{k+1}$.

➠ Montrons d'abord $\forall (p_1, p_2) \in \mathbb{P}^2, p|(p_1 p_2) \Longleftrightarrow p = p_1$ ou $p = p_2$.
D'après la propriété précédente $\forall (p_1, p_2) \in \mathbb{P}^2, p|(p_1 p_2) \Longleftrightarrow p|p_1$ ou $p|p_2$.
Mais comme p_1 et p_2 sont premiers et que $p \neq 1$ (puisque p premier), cela implique $p = p_1$ ou $p = p_2$.
Par récurrence sur k on étend cette propriété à un produit de k facteurs premiers.

➠ Montrons d'abord $\forall (p_1, p_2) \in \mathbb{P}^2$, si $p_1 \neq p_2$, $p_1^{a_1}|n$ et $p_2^{a_2}|n$ alors $(p_1^{a_1} \times p_2^{a_2})|n$.
Si les nombres premiers p_1 et p_2 sont différents, les diviseurs de $p_1^{a_1}$ sont $1, p_1, p_1^2, \ldots, p_1^{a_1}$ et ceux de $p_2^{a_2}$ sont $1, p_2, p_2^2, \ldots, p_2^{a_2}$. Ces ensembles n'ont que 1 en commun car, d'après le 3e point, si $d|p_1^{a_1}$ alors $d|p_1$ et, de même, si $d|p_2^{a_2}$ alors $d|p_2$, mais le seul diviseur commun à p_1 et p_2 est 1.
Par conséquent, d'après la propriété 6.11, comme $p_1^{a_1}|n$, $p_2^{a_2}|n$ et $p_1^{a_1}$ et $p_2^{a_2}$ sont premiers entre eux, on en déduit que $(p_1^{a_1} \times p_2^{a_2})|n$.
Par récurrence sur k on étend cette propriété à un produit de k puissances de facteurs premiers.

Remarques :

- Le 1er point de cette propriété est faux en général si p n'est pas premier : 4 ne divise pas 6 mais $PGCD(4, 6) = 2 \neq 1$.
- Les 2e et 3e points de cette propriété conduisent à des cas particuliers intéressants :
  $\forall a \in \mathbb{Z}, p|(a^n) \Longleftrightarrow p|a$.
  $\forall q \in \mathbb{P}, p|(q^k) \Longleftrightarrow p = q$ (j'ai utilisé ce cas particulier pour démontrer le dernier point).
- La même remarque que pour la propriété 6.11 s'impose concernant le point 4 de cette propriété : La condition $PGCD(p_i^{a_i}, p_j^{a_j}) = 1$ est nécessaire d'où la nécessité d'avoir des nombres premiers distincts. Contre-exemple : $2^3|24$ et $2^2|24$ mais $2^3 \times 2^2 = 2^5$ ne divise pas 24.

Propriété 6.15 (Décomposition en facteurs premiers) Un entier naturel supérieur à 1 se décompose de façon unique (à l'ordre des facteurs près) en produit de facteurs premiers.
Autrement dit, $\forall n \geqslant 2, \exists k \geqslant 1$, $(p_1, p_2, \ldots, p_k) \in \mathbb{P}^k$, k nombres premiers distincts et ordonnés $(p_1 < p_2 < \cdots < p_k)$ et $(a_1, a_2, \ldots, a_k) \in \mathbb{N}^{*k}$ tels que $n = p_1^{a_1} \times p_2^{a_2} \times \cdots \times p_k^{a_k}$.

DÉMONSTRATION ➲ Existence :
On a vu, propriété 6.12, que si $n > 1$ n'est pas premier alors il admet un diviseur premier p_1 tel que $n = p_1 q_1$ où q_1 est un entier inférieur strictement à n.
Par conséquent soit n est premier, dans ce cas la décomposition est trouvée, soit il ne l'est pas et s'écrit $n = p_1 q_1$ avec p_1 premier et $q_1 < n$.
On recommence alors avec q_1 qui est premier ou pas : s'il ne l'est pas, on l'écrit $q_1 = p_2 q_2$ avec p_2 premier (éventuellement confondu avec p_1) et $q_2 < q_1$.
On construit ainsi une suite d'entiers (q_n) strictement décroissante, donc finie.
Le dernier quotient trouvé est premier, ce qui met fin au processus de décomposition.

➲ Unicité :
Démontrons cela par récurrence sur la valeur de n en supposant que la décomposition de tous les entiers m tels que $1 < m < n$ soit unique.
Cette propriété est vraie pour $n = 3$ puisque pour $m = 2$ la décomposition est évidemment unique.
Supposons qu'il existe deux décompositions différentes de $n > 3$:
$n = p_1^{a_1} \times p_2^{a_2} \times \cdots \times p_k^{a_k} = q_1^{b_1} \times q_2^{b_2} \times \cdots \times q_l^{b_l}$.
Dans ce cas, un des facteurs premiers, disons p_1 divise n.
Donc $p_1 | (q_1^{b_1} \times q_2^{b_2} \times \cdots \times q_l^{b_l})$ et, d'après le 3e point de la propriété 6.14, il existe un indice $i \in \{1, 2, \ldots, l\}, p_1 = q_i$. On peut alors diviser l'égalité
$n = p_1^{a_1} \times p_2^{a_2} \times \cdots \times p_k^{a_k} = q_1^{b_1} \times q_2^{b_2} \times \cdots \times q_l^{b_l}$ par p_1 et ainsi $\frac{n}{p_1}$ est un nombre inférieur strictement à n pour lequel, d'après l'hypothèse de récurrence, la décomposition est unique. Or $\frac{n}{p_1} = \frac{n}{q_i}$ ce qui prouve que les deux décompositions ne peuvent être différentes. La décomposition de n est par conséquent unique ce qui achève la démonstration.

Remarques :

- Le raisonnement utilisé pour montrer l'existence de la décomposition est constructive : elle constitue un algorithme utilisable dans la pratique. Cet algorithme nécessite la connaissance du début de la liste des nombres premiers : on commence par diviser n par 2, autant de fois qu'il est possible (tant que le reste est nul), puis on divise de la même façon le quotient résiduel par 3, puis par 5, etc. jusqu'à obtenir un nombre premier. Par exemple $60 = 2 \times 30 = 2 \times (2 \times 15) = 2 \times (2 \times (3 \times 5)) = 2^2 \times 3 \times 5$. Ci-dessous en haut à droite, la décomposition *à la main* de deux entiers, selon la disposition habituelle (en colonne).
- L'exposant de chaque nombre premier entrant dans la décomposition d'un entier est appelé « multiplicité » de ce facteur premier. Ainsi dans la décomposition de 60 en $2^2 \times 3 \times 5$, la multiplicité du facteur 2 est 2 et celle des facteurs 3 et 5 est 1. Le programme qui suit donne la décomposition d'un entier naturel supérieur à 1 en ordonnant les facteurs premiers dans l'ordre croissant et en notant les multiplicités avec le symbole `^`. La fonction `premiers()` détermine la liste des nombres premiers à l'aide du crible d'Ératosthène.

```python
def premiers():
    L,k,nRacine=list(range(2,n+1)),2,n**0.5
    while k<nRacine :
        L=[p for p in L if p<=k or p%k!=0]
        k=L[L.index(k)+1] #nouveau nombre premier
    return L

n=int(input("Saisissez un nombre : "))
Liste_facteurs,decompo,i,s=premiers(),list(),0,str(n)+"="
#expurgation des nombress premiers non-facteurs de n
Liste_facteurs=[p for p in Liste_facteurs if n%p==0]
#recherche de la multiplicité de chaque facteur premier
while n>1 :
    expo=0
    while n%Liste_facteurs[i]==0 :
        expo+=1
        n//=Liste_facteurs[i]
    decompo.append([Liste_facteurs[i],expo])
    i+=1
for rang,facteur in enumerate(decompo) :
        s+=str(decompo[rang][0])+"^"+str(decompo[rang][1])
        if rang<len(decompo)-1 : s+='\u00D7'
print(s)
```

60	2
30	2
15	3
5	5
1	

1240	2
620	2
310	2
155	5
31	31
1	

$60 = 2^2 \text{x} 3 \text{x} 5$

$1240 = 2^3 \text{x} 5 \text{x} 31$

```
Saisissez un nombre : 1240
1240=2^3×5^1×31^1

Saisissez un nombre : 60
60=2^2×3^1×5^1
```

Propriété 6.16 (Diviseurs) Soit $n > 1$ un entier dont la décomposition en produit de facteurs premiers est $n = p_1^{a_1} \times p_2^{a_2} \times \cdots \times p_k^{a_k}$.

- $d|n \iff d = p_1^{b_1} \times p_2^{b_2} \times \cdots \times p_k^{b_k}$ où $\forall i \in \{1, 2, \ldots, k\}, b_i \in \mathbb{N}$ et $b_i \leqslant a_i$.
- Le nombre de diviseurs de n est $(a_1 + 1)(a_2 + 1) \ldots (a_k + 1)$.

Démonstration ➠ Sens réciproque de l'équivalence :
On suppose que $d = p_1^{b_1} \times p_2^{b_2} \times \cdots \times p_k^{b_k}$ avec $0 \leqslant b_i \leqslant a_i$.
La décomposition de n peut donc s'écrire $n = (p_1^{b_1} \times p_1^{a_1 - b_1}) \times (p_2^{b_2} \times p_2^{a_2 - b_2}) \times \cdots \times (p_k^{b_k} \times p_k^{a_k - b_k})$.
En réarrangeant ce produit $n = (p_1^{b_1} \times p_2^{b_2} \times \cdots \times p_k^{b_k}) \times (p_1^{a_1 - b_1} \times p_2^{a_2 - b_2} \times \cdots \times p_k^{a_k - b_k})$.
En notant $q = p_1^{a_1 - b_1} \times p_2^{a_2 - b_2} \times \cdots \times p_k^{a_k - b_k}$, on obtient donc $n = d \times q$ ce qui prouve bien que $d|n$.

➠ Sens direct de l'équivalence :
Soit $d > 1$ un diviseur de n. Tout diviseur premier p de d est un diviseur premier de n.
En effet, si $p|d$, comme $d|n$, on en déduit que $p|n$.
D'après le 3e point de la propriété 6.14, si un nombre premier p divise un produit de facteurs premiers, alors c'est un de ces facteurs premiers. Donc, puisque $p|d$ on en déduit que $p = p_i$ avec $i \in [\![1; k]\!]$.
Conséquence : la décomposition de d en facteurs premiers s'écrit $d = p_1^{b_1} \times p_2^{b_2} \times \cdots \times p_k^{b_k}$ où les coefficients b_i sont des entiers, éventuellement nuls.
Il reste à prouver que $\forall i \in [\![1; n]\!], b_i \leqslant a_i$.
Comme $d|n$, on a $n = d \times q = (p_1^{b_1} \times p_2^{b_2} \times \cdots \times p_k^{b_k}) \times q$.
Examinons le cas d'une valeur particulière de $i \in [\![1; k]\!]$:
$n = \left[(p_1^{b_1} \times p_2^{b_2} \times \cdots \times p_{i-1}^{b_{i-1}} \times p_{i+1}^{b_{i+1}} \cdots \times p_k^{b_k}) \times q\right] \times p_i^{b_i}$.
En posant $Q = p_1^{b_1} \times p_2^{b_2} \times \cdots \times p_{i-1}^{b_{i-1}} \times p_{i+1}^{b_{i+1}} \cdots \times p_k^{b_k}) \times q$, on a $n = Q \times p_i^{b_i}$. Donc $p_i^{b_i}$ divise n.
Mais comme $p_i^{a_i}$ divise aussi n, on peut écrire $n = Q' \times p_i^{a_i}$ où Q' ne contient que des puissances de facteurs premiers différents de p_i.
On a donc $p_i^{b_i}$ divise $Q' \times p_i^{a_i}$ avec $p_i^{b_i}$ premier avec Q'. D'après le théorème de Gauss, on en déduit que $p_i^{b_i}$ divise $p_i^{a_i}$ et par conséquent $p_i^{b_i} \leqslant p_i^{a_i} \Longrightarrow b_i \leqslant a_i$.
Ceci étant valable pour tout $i \in [\![1; k]\!]$, on a bien $\forall i \in \{1, 2, \ldots, k\}, b_i \in \mathbb{N}$ et $b_i \leqslant a_i$.

➠ Somme des diviseurs :
On vient de voir que si $n = p_1^{a_1} \times p_2^{a_2} \times \cdots \times p_k^{a_k}$, les diviseurs de n s'écrivent $p_1^{b_1} \times p_2^{b_2} \times \cdots \times p_k^{b_k}$ avec $\forall i \in [\![1; k]\!], 0 \leqslant b_i \leqslant a_i$. Or, pour une valeur particulière de $i \in [\![1; k]\!]$, il existe $a_i + 1$ valeurs différentes de l'exposant b_i : $0, 1, 2, \ldots, a_i$.
Par conséquent, l'ajout du facteur premier p_i multiplie par $a_i + 1$ le nombre de diviseurs de n ne contenant pas ce facteur premier.
Ceci étant vrai pour tout $i \in [\![1; k]\!]$, le nombre total de diviseurs est le produit des nombres $a_i + 1$ pour tout $i \in [\![1; k]\!]$.

Exemple 66 – Quels sont les diviseurs de $6048 = 2^5 \times 3^3 \times 7^1$ et combien sont-ils ?
La propriété précédente indique que 6048 possède $(5 + 1)(3 + 1)(1 + 1) = 6 \times 4 \times 2 = 48$ diviseurs.
Faisons-en la liste à l'aide de 2 tableaux, selon qu'ils contiennent ou non le facteur 7 :

7^0	$\times 2^0$	$\times 2^1$	$\times 2^2$	$\times 2^3$	$\times 2^4$	$\times 2^5$
$\times 3^0$	$2^03^0 = 1$	$2^13^0 = 2$	$2^23^0 = 4$	$2^33^0 = 8$	$2^43^0 = 16$	$2^53^0 = 32$
$\times 3^1$	$2^03^1 = 3$	$2^13^1 = 6$	$2^23^1 = 12$	$2^33^1 = 24$	$2^43^1 = 48$	$2^53^1 = 96$
$\times 3^2$	$2^03^2 = 9$	$2^13^2 = 18$	$2^23^2 = 36$	$2^33^2 = 72$	$2^43^2 = 144$	$2^53^2 = 288$
$\times 3^3$	$2^03^3 = 27$	$2^13^3 = 54$	$2^23^3 = 108$	$2^33^3 = 216$	$2^43^3 = 432$	$2^53^3 = 864$

7^1	$\times 2^0$	$\times 2^1$	$\times 2^2$	$\times 2^3$	$\times 2^4$	$\times 2^5$
$\times 3^0$	$2^03^07^1 = 7$	$2^13^07^1 = 14$	$2^23^07^1 = 28$	$2^33^07^1 = 56$	$2^43^07^1 = 112$	$2^53^07^1 = 224$
$\times 3^1$	$2^03^17^1 = 21$	$2^13^17^1 = 42$	$2^23^17^1 = 84$	$2^33^17^1 = 168$	$2^43^17^1 = 336$	$2^53^1s7^1 = 672$
$\times 3^2$	$2^03^27^1 = 63$	$2^13^27^1 = 126$	$2^23^27^1 = 252$	$2^33^27^1 = 504$	$2^43^27^1 = 1008$	$2^53^27^1 = 2016$
$\times 3^3$	$2^03^37^1 = 189$	$2^13^37^1 = 378$	$2^23^37^1 = 756$	$2^33^37^1 = 1512$	$2^43^37^1 = 3024$	$2^53^37^1 = 6048$

Propriété 6.17 (PGCD, PPCM et décompositions) Soient $n = p_1^{a_1} \times p_2^{a_2} \times \cdots \times p_k^{a_k}$ et $m = p_1^{b_1} \times p_2^{b_2} \times \cdots \times p_k^{b_k}$, deux entiers dont les décompositions en produit de facteurs premiers contiennent des exposants a_i ou b_i pouvant éventuellement être nuls.

- $PGCD(n,m) = p_1^{\min(a_1,b_1)} \times p_2^{\min(a_2,b_2)} \times \cdots \times p_k^{\min(a_k,b_k)}$
- $PPCM(n,m) = p_1^{\max(a_1,b_1)} \times p_2^{\max(a_2,b_2)} \times \cdots \times p_k^{\max(a_k,b_k)}$
- $PGCD(n,m) \times PPCM(n,m) = n \times m$

Démonstration ➠ D'après le 1^{er} point de la propriété précédente,
comme $\forall i \in [\![1;k]\!], \min(a_i,b_i) \leqslant a_i$ et $\min(a_i,b_i) \leqslant b_i$, on en déduit que le nombre $p_1^{\min(a_1,b_1)} \times p_2^{\min(a_2,b_2)} \times \cdots \times p_k^{\min(a_k,b_k)}$ est un diviseur de n et de m.
Par conséquent ce nombre divise $PGCD(n;m)$.
Inversement, puisque $PGCD(n;m)$ divise n et m, la même propriété permet d'écrire $PGCD(n;m) = p_1^{c_1} \times p_2^{c_2} \times \cdots \times p_k^{c_k}$ où $\forall i \in [\![1;k]\!], 0 \leqslant c_i \leqslant a_i$ et $0 \leqslant c_i \leqslant b_i$, c'est-à-dire $c_i \leqslant min(a_i,b_i)$, on en déduit que $PGCD(n;m)$ divise $p_1^{\min(a_1,b_1)} \times p_2^{\min(a_2,b_2)} \times \cdots \times p_k^{\min(a_k,b_k)}$.
Comme $a|b$ et $b|a \Longrightarrow a = b$ (antisymétrie de la relation $|$ dans $\mathbb{N}$), on peut en conclure $PGCD(n;m) = p_1^{\min(a_1,b_1)} \times p_2^{\min(a_2,b_2)} \times \cdots \times p_k^{\min(a_k,b_k)}$.

➠ Comme $\forall i \in [\![1;k]\!], \max(a_i,b_i) \geqslant a_i$, on a $\max(a_i,b_i) - a_i \geqslant 0$. On en déduit : $p_1^{\max(a_1,b_1)} \times \cdots \times p_k^{\max(a_k,b_k)} = \left[p_1^{\max(a_1,b_1)-a_1} \times \cdots \times p_k^{\max(a_k,b_k)-a_k}\right] \times \left[p_1^{a_1} \times \cdots \times p_k^{a_k}\right]$ soit, en posant $q = p_1^{\max(a_1,b_1)-a_1} \times \cdots \times p_k^{\max(a_k,b_k)-a_k}$, on a $p_1^{\max(a_1,b_1)} \times \cdots \times p_k^{\max(a_k,b_k)} = q \times n$.
Ce nombre est donc un multiple de n.
Pareillement, on montre que c'est aussi un multiple de m.
Ce nombre est donc un multiple commun de n et de m.
Inversement, un multiple commun de n et m peut s'écrire $p_1^{c_1} \times \cdots \times p_k^{c_k}$ où $\forall i \in [\![1;k]\!], c_i \geqslant a_i$ et $c_i \geqslant b_i$, c'est-à-dire $c_i \geqslant max(a_i,b_i)$.
On en déduit qu'un multiple commun de n et m est multiple de $p_1^{\max(a_1,b_1)} \times \cdots \times p_k^{\max(a_k,b_k)}$.
Ce nombre est donc le plus petit multiple commun de n et m.

➠ Si $max(a_i,b_i) = a_i$ alors $min(a_i,b_i) = b_i$ et si $max(a_i,b_i) = b_i$ alors $min(a_i,b_i) = a_i$, mais le produit $max(a_i,b_i) \times min(a_i,b_i)$ est dans tous les cas égal à $a_i \times b_i$. Cela permet de justifier le dernier point de cette propriété, nous laissons le lecteur finir cette démonstration.

Exemple 67 – Prenons $n = 2^4 \times 3^5 \times 5^1 \times 7^2 = 952\,560$ et $m = 2^1 \times 3^3 \times 5^3 \times 11^1 = 74\,250$.
De ces décompositions, on tire immédiatement :

- $PGCD(n;m) = 2^1 \times 3^3 \times 5^1 = 270$
- $PPCM(n;m) = 2^4 \times 3^5 \times 5^3 \times 7^2 \times 11^1 = 261\,954\,000$

Vérification :
$PGCD(n;m) \times PPCM(n;m) = 270 \times 261\,954\,000 = 70\,727\,580\,000$
$n \times m = 952\,560 \times 74\,250 = 70\,727\,580\,000$
Applications :

$$\frac{1}{952\,560} + \frac{1}{74\,250} = \frac{5^2 \times 11 + 2^3 \times 3^2 \times 7^2}{70\,727\,580\,000} = \frac{275 + 3528}{70\,727\,580\,000} = \frac{3803}{70\,727\,580\,000}$$

$$\frac{74\,250}{952\,560} = \frac{\frac{74\,250}{270}}{\frac{952\,560}{270}} = \frac{2\,750}{3\,528}$$

Programmation récursive du PGCD
Utilisation de la propriété pour en déduire le PPCM

```
def pgcd(a,b): # forme récursive
    reste=a%b
    if reste==0 : return b
    else : return pgcd(b,reste)

A=int(input('PGCD de A : '))
B=int(input('et de B : '))
PGCD=pgcd(A,B)
print('PGCD({},{})={}'.format(A,B,PGCD))
print('PPCM({},{})={}'.format(A,B,A*B-PGCD))
```

```
PGCD de A : 952560
et de B : 74250
PGCD(952560,74250)=270
PPCM(952560,74250)=70727579730
```

6. Petit théorème de Fermat

6.a. Coefficients binomiaux

Quelques rappels utiles du cours sur le dénombrement (chapitre 1)

DÉFINITION 6.8 (COMBINAISONS) On appelle « combinaison » de p éléments d'un ensemble à $n \geqslant p$ éléments toute partie de E à p éléments. On note $\binom{n}{p}$ (se lit p parmi n), ou parfois C_n^p, le nombre de combinaisons de p éléments d'un ensemble à n éléments.

PROPRIÉTÉ 6.18 (DÉNOMBREMENT) Soient n et $p \leqslant n$ deux entiers.
$$\binom{n}{p} = \frac{n(n-1)\dots(n-p+1)}{p!} = \frac{n!}{p!(n-p)!}$$

DÉMONSTRATION Cette propriété est démontrée dans le chapitre 1 (dénombrement)

Remarques :

- En particulier, on a $\binom{n}{1} = n$, $\binom{n}{n} = 1$ et $\binom{n}{0} = 1$.
- Trois autres propriétés de ces nombres, appelés « coefficients binomiaux » :
  — Symétrie : $\binom{n}{p} = \binom{n}{n-p}$ pour $0 \leqslant k \leqslant n$
  — Récurrence : $\binom{n}{p} = \frac{n}{p}\binom{n-1}{p-1}$ pour $p \neq 0$ et $n \neq 0$
  — Formule de Pascal : $\binom{n}{p} = \binom{n-1}{p-1} + \binom{n-1}{p}$ pour $n \geqslant 2$ et $1 \leqslant k \leqslant n-1$
- Dans une combinaison, l'ordre des éléments n'intervient pas et il ne peut y avoir de répétitions. Ainsi, lors d'un tirage d'une main de $p = 5$ cartes d'un jeu de $n = 32$ cartes, le nombre de mains possibles est $\binom{32}{5} = \frac{32\times31\times30\times29\times28}{1\times2\times3\times4\times5} = 201\,376$.
  De même, dans le développement du binôme $(a+b)^{32}$, le coefficient du terme a^5b^{27} est $\binom{32}{5} = 201\,376$ (il faut choisir le nombre a dans 5 facteurs parmi les 32 du développement).

PROPRIÉTÉ 6.19 (BINÔME DE NEWTON) Soient a et b deux réels, pour tout entier n on a :
$(a+b)^n = \binom{n}{0}a^0b^n + \binom{n}{1}a^1b^{n-1} + \dots + \binom{n}{k}a^kb^{n-k} + \dots + \binom{n}{n}a^nb^0 = \sum_{k=0}^{n}\binom{n}{k}a^kb^{n-k}$

DÉMONSTRATION La dernière remarque en donne un élément de justification mais cette propriété est démontrée plus rigoureusement dans le chapitre 1 (dénombrement).

PROPRIÉTÉ 6.20 (DIVISIBILITÉ PAR n PREMIER) Soient a et b deux entiers.
Si n est un nombre premier alors $(a+b)^n = a^n + b^n[n]$

DÉMONSTRATION ➠ Lemme : si n et k sont premiers entre eux alors $\binom{n}{k} = 0[n]$.
Écrivons la propriété de récurrence citée plus haut $k\binom{n}{k} = n\binom{n-1}{k-1}$ pour $k \neq 0$ et $n \neq 0$.
On en déduit que $n|k\binom{n}{k}$ et, comme n et k sont premiers entre eux, d'après le théorème de Gauss, $n|\binom{n}{k}$, ce qui revient à dire que $\binom{n}{k} = 0[n]$.

➠ En écrivant explicitement ses termes extrêmes, le développement du binôme s'écrit :
$(a+b)^n = b^n + \sum_{k=1}^{n-1}\binom{n}{k}a^kb^{n-k} + a^n$.
Comme n est premier, il est premier avec tous les entiers $1 \leqslant k \leqslant n-1$ et donc, d'après le lemme précédent, on a $\forall k \in [\![1, n-1]\!]$, $\binom{n}{k} = 0[n]$ d'où $\sum_{k=1}^{n-1}\binom{n}{k}a^kb^{n-k} = 0[n]$.
Finalement, il reste $(a+b)^n = a^n + b^n[n]$.

6.b. Théorème de Fermat

PROPRIÉTÉ 6.21 (FERMAT1) Si n est un nombre premier, quelque soit l'entier a on a $a^n = a[n]$

DÉMONSTRATION ➠ Dans le cas $n = 2$:
On a $a^2 - a = a(a-1)$. Ce produit est nécessairement pair puisque l'un des deux entiers consécutifs n et $n-1$ est forcément pair, donc $a^2 - a$ est pair, autrement dit $a^2 - a = 0[2]$ ou encore $a^2 = a[2]$.

➠ Dans le cas où $n \neq 2$ est premier, n est impair :
Supposons a entier positif et montrons $a^n = a[n]$ par récurrence sur a.

- Initialisation : $0^n = 0[n]$
- Hérédité : la propriété est supposée vraie au rang a ($a^n = a[n]$)
  D'après la propriété précédente $(a+1)^n = a^n + 1^n = a^n + 1[n]$ et comme $a^n = a[n]$ (hypothèse de récurrence), on en déduit $(a+1)^n = a+1[n]$.
  Ce qui prouve que la propriété est encore vraie au rang $a+1$.

Conclusion : la propriété est vraie pour a entier positif.

Supposons maintenant a entier strictement négatif.
Comme $-a$ est positif, on a $(-a)^n = -a[n]$ mais comme n est impair $(-a)^n = -(a^n) = -a^n$ d'où $-a^n = -a[n] \Longleftrightarrow a^n = a[n]$.
La propriété est donc vraie pour tout $a \in \mathbb{Z}$ et tout nombre premier n.

PROPRIÉTÉ 6.22 (FERMAT2) Soit n est un nombre premier et a un entier.
Si n ne divise pas a alors $a^{n-1} = 1[n]$.

DÉMONSTRATION D'après la propriété précédente, si n est premier :
$a^n - a = 0[n] \Longleftrightarrow n|(a^n - a) \Longleftrightarrow n|a(a^{n-1} - 1)$.
Comme n ne divise pas a, d'après le théorème de Gauss :
$n|(a^{n-1} - 1) \Longleftrightarrow a^{n-1} - 1 = 0[n] \Longleftrightarrow a^{n-1} = 1[n]$.

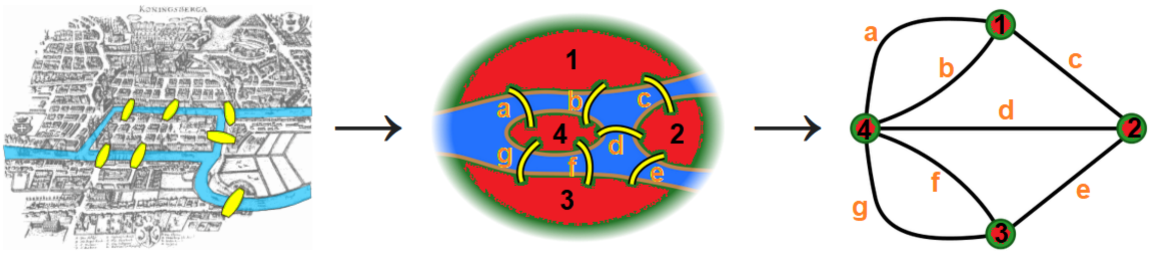
Remarques :

- Les deux propriétés énoncées (Fermat1 et Fermat2) sont indifféremment appelées *petit* théorème de Fermat et doivent être distinguées du *grand* théorème de Fermat qui énonce que l'équation $a^n + b^n = c^n$ où a, b, c et n sont des entiers strictement positifs, n'a pas de solution pour $n \geqslant 3$. Cette propriété ayant été démontrée par Andrew Wiles en 1994, elle a été rebaptisée théorème de Fermat-Wiles.
- Dans l'exemple 64, j'ai tracé le tableau des puissances de a modulo 5.
  On remarque dans ce tableau que pour tout a non divisible par 5, on a $a^4 = 1[5]$.
  Évidemment, lorsque $5|a$, cette propriété est fausse puisque $a = 0[5] \Longrightarrow a^4 = 0[5]$.
- On montre facilement que, dans les mêmes conditions, si N est un multiple de $n-1$ alors $a^N = 1[n]$. De même, dans ces conditions, les nombres a (premiers avec n) possèdent un inverse modulo n qui est une puissance de a. Par exemple modulo 5, l'inverse de 2 est 3 et réciproquement ($2 \times 3 = 1[5]$), 4 et 1 sont leur propre inverse ($4 \times 4 = 1[5]$ et $1 \times 1 = 1[5]$).

7 Graphes

Plan du chapitre :

Aperçu historique : La notion de graphe est récente puisqu'elle commence seulement au XVIIe siècle avec le fameux problème des sept ponts de la ville de Königsberg que résolut Leonhard Euler en 1735 : on cherchait à déterminer s'il existe un chemin permettant de revenir à son point de départ en empruntant une fois et une seule chacun des ponts de la ville. Euler montra qu'un tel chemin n'existait pas et il le prouva en exhibant le graphe correspondant au problème et en justifiant l'impossibilité par le fait que ses quatre sommets sont incidents à un nombre impair d'arêtes.



La théorie des graphes est une discipline mathématique qui a été élaborée progressivement et qui a aujourd'hui de nombreuses applications dans tous les domaines liés à la notion de réseau (réseau social, réseau informatique, télécommunications, etc.) et dans bien d'autres domaines (génétique, algorithmique).
De nombreuses informations d'un graphe peuvent être représentées par une matrice.
Le calcul matriciel proprement dit n'est apparu qu'au début du XIXe siècle, mais les matrices, en tant que tableaux de nombres, ont une longue histoire d'applications à la résolution d'équations linéaires. En Chine, au IIe siècle avant J.-C., le livre *Les Neuf Chapitres sur l'Art Mathématique* utilise des déterminants pour résoudre des systèmes d'équations. En Europe, le livre *Ars Magna* de Jérôme Cardan publié en 1545 reprend cette méthode. Les transformations géométriques sont ensuite étudiées à l'aide de matrices en 1659 dans le livre du hollandais Johan de Witt *Elementa curvarum linearum*. Au XVIIIe siècle, Gottfried Wilhem Leibniz (1646 - 1716) puis Gabriel Cramer (1704 - 1752) font avancer le sujet en développant la théorie des déterminants. James Joseph Sylvester, en 1850, forge le terme *matrix*, traduit par matrice, sur la racine *mater* (la mère qui donne naissance aux déterminants). En 1854, Arthur Cayley définit les règles du calcul matriciel et donne à ce sujet une dimension plus abstraite qui en généralise l'utilisation.
En 1913 apparait la notation moderne des parenthèses (ou des crochets) pour représenter les matrices, ainsi que de la notation systématique $A = (a_{ij})$ (ou $[a_{ij}]$) où a_{ij} désigne le terme de la i^{e} ligne et de la j^{e} colonne.

1. Calcul matriciel

1.a. Matrices

Définition 7.1 (Matrice) Soient n et p deux entiers naturels non nuls.
La matrice $(a_{ij})_{1\leqslant i\leqslant n,1\leqslant j\leqslant p}$ est le tableau de nombres a_{ij} où i désigne le numéro de la ligne et j le numéro de la colonne, le nombre a_{ij} étant à l'intersection de la ligne i et de la colonne j.

Remarques :

- Les *coefficients* de la matrice $(a_{ij})_{1\leqslant i\leqslant n,1\leqslant j\leqslant p}$ sont les nombres a_{ij}.
  La *dimension* (on dit aussi taille, format) de la matrice est le nombre $n \times p$ de ses coefficients.
  L'ensemble des matrices à coefficients réels de dimension $n \times p$ est noté $\mathscr{M}_{n,p}(\mathbb{R})$.
  Bien sûr, il existe aussi des matrices à coefficients complexes ou entiers.
  Les matrices nulles – notées O – ont tous leurs coefficients égaux à zéro.
- Écrire explicitement les coefficients d'une matrice de dimension 2×3 revient à écrire
  $A = (a_{ij})_{1\leqslant i\leqslant 2,1\leqslant j\leqslant 3} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix}$.
  On peut aussi avoir à écrire explicitement une matrice de dimension $n \times p$:
  $$B = (b_{ij})_{1\leqslant i\leqslant n,1\leqslant j\leqslant p} = \begin{pmatrix} b_{11} & a_{12} & \dots & b_{1j} & \dots & b_{1p} \\ b_{21} & b_{22} & \dots & b_{2j} & \dots & b_{2p} \\ \vdots & \vdots & \dots & \vdots & \dots & \vdots \\ b_{i1} & b_{i2} & \dots & b_{ij} & \dots & b_{ip} \\ \vdots & \vdots & \dots & \vdots & \dots & \vdots \\ b_{n1} & b_{n2} & \dots & b_{nj} & \dots & b_{np} \end{pmatrix}$$
- On distingue les « matrices lignes » telles que $A = \begin{pmatrix} 1 & 2 & 3 \end{pmatrix}$ lorsque $n = 1$,
  les « matrices colonnes » telles que $B = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ lorsque $p = 1$,
  les « matrices carrées » telles que $C = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$ lorsque $n = p$.
  L'ensemble des matrices carrées à coefficients réels de dimension $n \times n$ est noté $\mathscr{M}_n(\mathbb{R})$.

Définition 7.2 (Égalité) Deux matrices de même dimension $n \times p$ sont égales si et seulement si tous les coefficients de même rang sont égaux. Autrement dit :
$(a_{ij})_{1\leqslant i\leqslant n,1\leqslant j\leqslant p} = (b_{ij})_{1\leqslant i\leqslant n,1\leqslant j\leqslant p} \iff \forall i \in [\![1,n]\!], \forall j \in [\![1,p]\!], a_{ij} = b_{ij}$.

Définition 7.3 (Somme) La somme de deux matrices de même dimension est une matrice de même dimension dont les coefficients sont obtenus en sommant les coefficients de même rang des deux matrices : $(a_{ij}) + (b_{ij}) = (a_{ij} + b_{ij})$

Propriété 7.1 (Addition) $(\mathscr{M}_{n,p}(\mathbb{R}),+)$ est un groupe commutatif.

- Commutativité : $\forall (A,B) \in \mathscr{M}^2_{n,p}(\mathbb{R}), A + B = B + A$
- Associativité : $\forall (A,B,C) \in \mathscr{M}^3_{n,p}(\mathbb{R}), A + (B + C) = (A + B) + C = A + B + C$
- Élément neutre : $\forall A \in \mathscr{M}_{n,p}(\mathbb{R}), A + O = O + A = A$
- Opposé : $\forall A \in \mathscr{M}_{n,p}(\mathbb{R}), \exists - A \in \mathscr{M}_{n,p}(\mathbb{R}), A + (-A) = O$. Si $A = (a_{ij})$ alors $-A = (-a_{ij})$

Remarques :

- Attention :l'addition de deux matrices de dimensions différentes n'a pas de sens.
- Soustraction dans $\mathscr{M}_{n,p}(\mathbb{R})$: la différence de deux matrices A et B de même dimension est la matrice $A-B$ obtenue en additionnant A et l'opposée de B : $A-B=A+(-B)$.
- Ces propriétés découlent des propriétés de l'addition dans $\mathbb{R}$ et de la définition de l'addition matricielle. Bien sûr, ces propriétés ne dépendent pas de l'ensemble numérique dans lequel sont pris les coefficients : $(\mathscr{M}_{n,p}(\mathbb{Z}),+)$ ou $(\mathscr{M}_{n,p}(\mathbb{C}),+)$ sont également des groupes commutatifs.

Exemple 68 – Avec $A=\begin{pmatrix}1&2&3\\4&5&6\\7&8&9\end{pmatrix}$ et $B=\begin{pmatrix}9&8&7\\6&5&4\\3&2&1\end{pmatrix}$, on a $-A=\begin{pmatrix}-1&-2&-3\\-4&-5&-6\\-7&-8&-9\end{pmatrix}$,

$$A+B=\begin{pmatrix}1+9&2+8&3+7\\4+6&5+5&6+4\\7+3&8+2&9+1\end{pmatrix}=\begin{pmatrix}10&10&10\\10&10&10\\10&10&10\end{pmatrix}$$

$$B-A=\begin{pmatrix}9-1&8-2&7-3\\6-4&5-5&4-6\\3-7&2-8&1-9\end{pmatrix}=\begin{pmatrix}8&6&4\\2&0&-2\\-4&-6&-8\end{pmatrix}$$

Définition 7.4 (Produit par un réel) Soit $A=(a_{ij})$ une matrice de $\mathscr{M}_{n,p}(\mathbb{R})$ et λ un réel. La matrice λA est la matrice de $\mathscr{M}_{n,p}(\mathbb{R})$ définie par (λa_{ij})

Propriété 7.2 (Espace vectoriel) Le produit d'une matrice par un réel vérifie :

- Distributivité sur les matrices : $\forall(A,B)\in\mathscr{M}_{n,p}^2(\mathbb{R}),\forall\lambda\in\mathbb{R},\lambda(A+B)=\lambda A+\lambda B$
- Distributivité sur les facteurs : $\forall A\in\mathscr{M}_{n,p}(\mathbb{R}),\forall(\lambda,\mu)\in\mathbb{R}^2,(\lambda+\mu)A=\lambda A+\mu A$
- Associativité : $\forall A\in\mathscr{M}_{n,p}(\mathbb{R}),\forall(\lambda,\mu)\in\mathbb{R}^2,(\lambda\mu)A=\lambda(\mu A)$
- Élément neutre : $\forall A\in\mathscr{M}_{n,p}(\mathbb{R}),1A=A$

Comme de plus $(\mathscr{M}_{n,p}(\mathbb{R}),+)$ est un groupe commutatif, l'ensemble $\mathscr{M}_{n,p}(\mathbb{R})$ muni de l'addition matricielle et du produit par un réel est un espace vectoriel.

Remarques :

- Quelle que soit la matrice A, $0A=O$ et $(-1)A=-A$
- Si λ et μ sont deux réels quelconques, une combinaison linéaire des deux matrices de même dimension A et B est une matrice $\lambda A+\mu B$.
- L'ensemble $\mathscr{M}_2(\mathbb{R})$ des matrices carrées à coefficients réels de dimension 2×2 est engendré par les matrices carrées suivantes $I=\begin{pmatrix}1&0\\0&0\end{pmatrix}$, $J=\begin{pmatrix}0&1\\0&0\end{pmatrix}$, $K=\begin{pmatrix}0&0\\1&0\end{pmatrix}$ et $L=\begin{pmatrix}0&0\\0&1\end{pmatrix}$ dans le sens qu'une matrice quelconque de $\mathscr{M}_2(\mathbb{R})$ s'écrit comme une combinaison linéaire de ces quatre matrices : $\begin{pmatrix}a&b\\c&d\end{pmatrix}=aI+bJ+cK+dL$. De même, l'ensemble $\mathscr{M}_{3,1}(\mathbb{R})$ des matrices colonnes à coefficients réels de dimension 3 est engendré par trois matrices formant une base de cet espace vectoriel : $I=\begin{pmatrix}1\\0\\0\end{pmatrix}$,$J=\begin{pmatrix}0\\1\\0\end{pmatrix}$ et $K=\begin{pmatrix}0\\0\\1\end{pmatrix}$ et $\begin{pmatrix}a\\b\\c\end{pmatrix}=aI+bJ+cK$.

Définition 7.5 (Transposition) La transposée d'une matrice $A=(a_{ij})$ de dimension $n\times p$ est la matrice $B=(b_{ij})$ de taille $p\times n$ notée ${}^{\mathrm{t}}A$ obtenue en échangeant les lignes et les colonnes de A. Autrement dit $B={}^{\mathrm{t}}A\iff\forall(i,j)\in[\![1,p]\!]\times[\![1,n]\!],b_{ij}=a_{ji}$.
Si une matrice carrée est égale à sa transposée ($A={}^{\mathrm{t}}A$), on dit que la matrice est *symétrique*, si elle est égale à l'opposé de sa transposée ($A=-{}^{\mathrm{t}}A$), on dit que la matrice est *antisymétrique*.

Exemple 69 – Pour $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \\ 3 & 1 & 2 \end{pmatrix}$ et $C = \begin{pmatrix} 0 & -2 & -1 \\ 2 & 0 & -3 \\ 1 & 3 & 0 \end{pmatrix}$,
on a ${}^{t}A = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}$, ${}^{t}B = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \\ 3 & 1 & 2 \end{pmatrix}$ et ${}^{t}C = \begin{pmatrix} 0 & 2 & 1 \\ -2 & 0 & 3 \\ -1 & -3 & 0 \end{pmatrix}$.
B est symétrique, C antisymétrique (les coefficients de la diagonale principale doivent être nuls).
Une matrice *diagonale* – matrice carrée dont tous les coefficients sont nuls exceptés ceux de sa diagonale principale : (a_{ij}) diagonale si $a_{ij} \begin{cases} = 0 \text{ si } i \neq j \\ \in \mathbb{R} \text{ si } i = j \end{cases}$ – est toujours symétrique.

1.b. Produit Matriciel

Définition 7.6 (Produit de matrices) Soit $A = (a_{ij})_{1 \leqslant i \leqslant n, 1 \leqslant j \leqslant m}$ une matrice de taille $n \times m$ et soit $B = (b_{ij})_{1 \leqslant i \leqslant m, 1 \leqslant j \leqslant p}$ une matrice de taille $m \times p$. Le produit de A par B est la matrice $C = AB = (c_{ij})_{1 \leqslant i \leqslant n, 1 \leqslant j \leqslant p}$ de taille $n \times p$ avec $\forall i \in [\![1, n]\!], \forall j \in [\![1, p]\!], c_{ij} = \sum_{k=1}^{m} a_{ik} b_{kj}$.

Exemple 70 – Avec $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$ et $B = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix}$, on a une matrice AB de dimension $(2 \times \cancel{3})(\cancel{3} \times 2) = (2 \times 2)$ tandis que BA est une matrice de dimension $(3 \times \cancel{2})(\cancel{2} \times 3) = (3 \times 3)$.
Généralement on dispose les matrices à multiplier à gauche (pour la 1re) et au-dessus (pour la 2e) de la matrice produit, de façon à déterminer c_{ij}, le produit de la ligne i par la colonne j.

$$\begin{array}{ccc} & \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} & \\ AB = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} & \begin{pmatrix} 1 \times 1 + 2 \times 3 + 3 \times 5 & 1 \times 2 + 2 \times 4 + 3 \times 6 \\ 4 \times 1 + 5 \times 3 + 6 \times 5 & 4 \times 2 + 5 \times 4 + 6 \times 6 \end{pmatrix} & = \begin{pmatrix} 22 & 28 \\ 49 & 64 \end{pmatrix} \end{array}$$

$$\begin{array}{ccc} & \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} & \\ BA = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} & \begin{pmatrix} 1 \times 1 + 2 \times 4 & 1 \times 2 + 2 \times 5 & 1 \times 3 + 2 \times 6 \\ 3 \times 1 + 4 \times 4 & 3 \times 2 + 4 \times 5 & 3 \times 3 + 4 \times 6 \\ 5 \times 1 + 6 \times 4 & 5 \times 2 + 6 \times 5 & 5 \times 3 + 6 \times 6 \end{pmatrix} & = \begin{pmatrix} 9 & 12 & 15 \\ 19 & 26 & 33 \\ 29 & 40 & 51 \end{pmatrix} \end{array}$$

Remarques :

- Le produit des matrices n'est pas commutatif en général. On l'a vu dans l'exemple ci-dessus où les matrices AB et BA n'avaient même pas les mêmes dimensions. Pour des matrices carrées, le produit n'est généralement pas commutatif non plus.
- Si le produit de deux matrices est la matrice nulle O, cela ne signifie pas que l'une des matrices est nulle. Il suffit pour s'en convaincre d'effectuer le produit $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ par $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$: on trouve O et pourtant aucune de ces matrices n'est nulle.
- La matrice *identité* de taille n – notée I_n – est une matrice diagonale dont les coefficients non nuls valent 1 : $I_n = (a_{ij})$ avec $a_{ij} \begin{cases} = 0 \text{ si } i \neq j \\ = 1 \text{ si } i = j \end{cases}$
  Les produits $I_n A$ et $A I_m$ restituent intégralement une matrice A de dimension $n \times m$ et pour une matrice carrée de dimension $n \times n$, on a $I_n A = A I_n = A$.

Propriété 7.3 Associativité :$\forall (A, B, C) \in \mathscr{M}_{n,m}(\mathbb{R}) \times \mathscr{M}_{m,p}(\mathbb{R}) \times \mathscr{M}_{p,q}(\mathbb{R}), A(BC) = (AB)C$
Distributivité à droite :$\forall (A, B, C) \in \mathscr{M}_{n,m}(\mathbb{R}) \times \mathscr{M}_{n,m}(\mathbb{R}) \times \mathscr{M}_{m,p}(\mathbb{R}), (A + B)C = AC + BC$
Distributivité à gauche :$\forall (A, B, C) \in \mathscr{M}_{n,m}(\mathbb{R}) \times \mathscr{M}_{m,p}(\mathbb{R}) \times \mathscr{M}_{m,p}(\mathbb{R}), A(B + C) = AB + AC$
Associativité mixte : $\forall (A, B) \in \mathscr{M}_{n,m}(\mathbb{R}) \times \mathscr{M}_{m,p}(\mathbb{R}), \forall k \in \mathbb{R}, A(kB) = (kA)B = k(AB)$
Élément neutre : $\forall A \in \mathscr{M}_{n,m}(\mathbb{R}), I_n A = A I_m = A$ et $\forall A \in \mathscr{M}_n(\mathbb{R}), I_n A = A I_n = A$

> DÉMONSTRATION Ces propriétés découlent des règles du produit matriciel, du produit d'une matrice par un réel, de l'addition matricielle ainsi que des règles de calcul dans $\mathbb{R}$.

Remarques :

- On peut définir le carré, le cube, etc. d'une matrice carrée $A^n = \overbrace{A \times A \times \cdots \times A}^{n \text{ fois}}$, mais pas d'une matrice de dimension $n \times m$ lorsque $n \neq m$. Quelle que soit la dimension d'une matrice carrée, on a toujours $A^1 = A$ et par convention $A^0 = I$. Avec la matrice identité, on a $I^n = I$.
- Une matrice carrée A est *nilpotente* s'il existe un entier $n > 1$ tel que $A^n = O$, on dit alors que n est l'indice de nilpotence. Par exemple la matrice $I = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ est nilpotente d'indice 2 car $I^2 = O$ et la matrice $M = \begin{pmatrix} 1 & 3 & -3 \\ 2/3 & 0 & 0 \\ 1 & 1 & -1 \end{pmatrix}$ est nilpotente d'indice 3 car $M^3 = O$.
- À cause de la non-commutativité du produit matriciel, les identités remarquables ne se transposent pas directement aux matrices. En effet $(A+B)(A-B) = A^2 - AB + BA - B^2$ et $(A-B)(A+B) = A^2 + AB - BA - B^2$ mais aucune de ces expressions n'est égale, en général, à $A^2 - B^2$. De même pour les autres identités connues.

Programmation :
En Python, les matrices sont des listes de listes.
La bibliothèque `numpy` (abréviation de *numerical python*) contient toutes les fonctions utiles avec les matrices. Pour effectuer un produit de matrices : importer la bibliothèque, déclarer les matrices et utiliser le symbole `@`. Ci-dessous, je vérifie mes calculs précédents à l'aide de cette opération `numpy`.

```
import numpy as np
A=np.array([[1,2,3],[4,5,6]])
B=np.array([[1,2],[3,4],[5,6]])
print(A@B)
print(B@A)
C=np.array([[3,9,-9],[2,0,0],[3,3,-3]])
print("carré :")
print(C@C)
print("cube :")
print(C@C@C)
```

```
[[22 28]
 [49 64]]
[[ 9 12 15]
 [19 26 33]
 [29 40 51]]
carré :
[[  0   0   0]
 [  6  18 -18]
 [  6  18 -18]]
cube :
[[0 0 0]
 [0 0 0]
 [0 0 0]]
```

Programmer en Python le produit de deux matrices est bien sûr possible sans `numpy` : il suffit de traduire la définition en algorithme. La fonction `produit(A,B)` ci-dessous effectue cela.

```
def produit(A,B):
    na,pa=len(A),len(A[0])
    nb,pb=len(B),len(B[0])
    if pa!=nb:
        print("dimensions incompatibles")
        return
    C=[pb*[0] for _ in range(na)]
    for lig in range(na) :
        for col in range(pb) :
            for elt in range(nb) :
                C[lig][col]+=A[lig][elt]*B[elt][col]
    return C

A=[[1,2,3],[4,5,6]]
B=[[1,2],[3,4],[5,6]]
print(produit(A,B))
print(produit(B,A))
print("")
C=[[3,9,-9],[2,0,0],[3,3,-3]]
print("carré :",produit(C,C))
print("cube :",produit(produit(C,C),C))
```

```
[[22, 28], [49, 64]]
[[9, 12, 15], [19, 26, 33], [29, 40, 51]]

carré : [[0, 0, 0], [6, 18, -18], [6, 18, -18]]
cube : [[0, 0, 0], [0, 0, 0], [0, 0, 0]]
```

1.c. Matrices Inversibles

Définition 7.7 (Matrice inversible) Soit A une matrice carrée de taille n.
On dit que A est inversible s'il existe une matrice B telle que $AB = BA = I_n$.
Si A est inversible, la matrice inverse de A est unique, on la note A^{-1}.

Remarques :

- Toutes les matrices carrées ne sont pas inversibles.
  La matrice $A = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$ n'est pas inversible car il faudrait trouver $(a, b, c, d) \in \mathbb{R}^4$ tels que $A\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \iff \begin{pmatrix} a+c & b+d \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ ce qui est impossible.
  De même, la matrice nulle $O = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ n'est pas inversible.
- L'ensemble des matrices carrées inversibles de taille n s'appelle le *groupe linéaire* de degré n et est noté $\mathrm{GL}_n(\mathbb{R})$. La matrice $A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$ appartient à $\mathrm{GL}_2(\mathbb{R})$ car elle est inversible. Son inverse est la matrice $A^{-1} = \begin{pmatrix} 0 & 1 \\ 1 & -1 \end{pmatrix}$. La matrice I_n est sa propre inverse, c'est d'ailleurs la seule matrice à avoir cette propriété dans $\mathrm{GL}_n(\mathbb{R})$. Par contre, pour toute matrice inversible $(A^{-1})^{-1} = A$, l'inversion est .
- Justification de l'unicité de l'inverse : si $A \in \mathscr{M}_n(\mathbb{R})$ possède deux inverses X et Y, on a $AX = XA = I_n$ et $AY = YA = I_n$ d'où $X = XI_n = X(AY) = (XA)Y = I_nY = Y$.

Propriété 7.4 (Groupe Linéaire) $(\mathrm{GL}_n(\mathbb{R}), .)$, l'ensemble des matrices inversibles de taille n doté de l'opération . (la multiplication matricielle) est un groupe non commutatif.

- Loi de composition interne : $\forall (A, B) \in \mathrm{GL}_n^2(\mathbb{R}), AB \in \mathrm{GL}_n(\mathbb{R})$ et $(AB)^{-1} = B^{-1}A^{-1}$
- Associativité : $\forall (A, B, C) \in \mathrm{GL}_n^3(\mathbb{R}), (AB)C = A(BC)$
- Élément neutre : $\forall A \in \mathrm{GL}_n(\mathbb{R}), I_nA = AI_n = A$
- Élément symétrique : $\forall A \in \mathrm{GL}_n(\mathbb{R}), \exists A^{-1} \in \mathrm{GL}_n(\mathbb{R}), AA^{-1} = A^{-1}A = I_n$

Démonstration Inverse de AB :
D'après l'associativité dans $\mathscr{M}_n(\mathbb{R})$, on a $(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = (AI_n)A^{-1} = AA^{-1} = I_n$

Propriété 7.5 Deux propriétés liées aux matrices inversibles :
$\forall A \in \mathrm{GL}_n(\mathbb{R}), \forall (B, C) \in \mathscr{M}_{n,p}^2(\mathbb{R}), AB = AC \Longrightarrow B = C$
$\forall A \in \mathrm{GL}_n(\mathbb{R}), \forall B \in \mathscr{M}_{n,1}(\mathbb{R}), \exists! X \in \mathscr{M}_{n,1}(\mathbb{R}), AX = B \iff X = A^{-1}B$

Démonstration ➟ Simplification par $A \in \mathrm{GL}_n(\mathbb{R})$:
D'après la distributivité dans $\mathscr{M}_{n,p}(\mathbb{R})$, on a $AB = AC \Longrightarrow A(B - C) = O$ et,
comme A est inversible $A(B - C) = O \Longrightarrow A^{-1}A(B - C) = A^{-1}O \Longrightarrow B - C = O$.
Attention : on l'a dit, cette propriété n'est pas valable si $A \notin \mathrm{GL}_n(\mathbb{R})$!

➟ Équation matricielle :
Rappel : $\mathscr{M}_{n,1}(\mathbb{R})$ est l'ensemble des matrices colonne de taille n.
Existence : Comme A est inversible $AX = B \iff (A^{-1}A)X = A^{-1}B \iff X = A^{-1}B$
Unicité : supposons que l'équation $AX = B$ a deux solutions X_1 et X_2.
On a alors $AX_1 = AX_2$ mais comme A est inversible, d'après la propriété précédente $X_1 = X_2$.

Exemple 71 – La solution de l'équation $AX = B$ fournit une méthode pour déterminer A^{-1} :
Supposons qu'on cherche à inverser la matrice $\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$.

Écrivons que le produit de A avec la matrice colonne $X = \begin{pmatrix} x \\ y \end{pmatrix}$ égale la matrice colonne $B = \begin{pmatrix} a \\ b \end{pmatrix}$.

$$\begin{cases} x + 2y = a \\ 3x + 4y = b \end{cases} \iff \begin{cases} 3x + 6y = 3a \\ 3x + 4y = b \end{cases} \iff \begin{cases} 2y = 3a - b \\ 3x = b - 4y \end{cases} \iff \begin{cases} y = \dfrac{3}{2}a + \dfrac{-1}{2}b \\ x = \dfrac{1}{3}(b - 4y) \end{cases}$$

On en déduit $x = \frac{1}{3}\left(b - \frac{4\times 3}{2}a + \frac{4\times 1}{2}b\right) = -2a + b$

d'où $\begin{cases} x = -2a + b \\ y = \dfrac{3}{2}a + \dfrac{-1}{2}b \end{cases}$ et finalement $A^{-1} = \begin{pmatrix} -2 & 1 \\ \frac{3}{2} & \frac{-1}{2} \end{pmatrix}$.

Vérification : $\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} -2 & 1 \\ \frac{3}{2} & \frac{-1}{2} \end{pmatrix} = \begin{pmatrix} -2+3 & 1-1 \\ -6+6 & 3-2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I_2$.

Propriété 7.6 (Inverse d'une matrice carrée de taille 2) Soit $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$.
A est inversible si et seulement si son déterminant $\Delta = \det(A) = ad - bc$ n'est pas nul :
$\Delta \neq 0 \iff A^{-1} = \frac{1}{\Delta}\begin{pmatrix} d & -c \\ -b & a \end{pmatrix}$

Démonstration Appliquons la méthode précédente pour déterminer la matrice colonne $X = \begin{pmatrix} x \\ y \end{pmatrix}$ qui vérifie l'équation $AX = \begin{pmatrix} x' \\ y' \end{pmatrix}$.

$$\begin{cases} ax + by = x' \\ cx + dy = y' \end{cases} \iff \begin{cases} acx + bcy = cx' \\ acx + ady = ay' \end{cases} \iff \begin{cases} (ad - bc)y = ay' - cx' \\ ax = x' - by \end{cases} \iff \begin{cases} y = \dfrac{1}{\Delta}(ay' - cx') \\ x = \dfrac{1}{a}(x' - by) \end{cases}$$

On en déduit $x = \frac{1}{a}\left(x' - \frac{b}{\Delta}(ay' - cx')\right) = \frac{1}{\Delta}\left(\frac{ad-bc}{a}x' + \frac{-b}{a}(ay' - cx')\right) = \frac{1}{\Delta}(dx' - by')$

d'où $\begin{cases} x = \dfrac{1}{\Delta}(dx' - by') \\ y = \dfrac{1}{\Delta}(-cx' + ay') \end{cases}$ et finalement $A^{-1} = \frac{1}{\Delta}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$.

Propriété 7.7 (Inverse d'une matrice diagonale) Une matrice diagonale $A = (d_{ij})$ de taille n est inversible si et seulement si tous les coefficients de sa diagonale principale sont non nuls :
$\forall (i,j) \in [\![1,n]\!]^2, d_{ij} \neq 0$ si $i = j \iff A^{-1} = (d'_{ij})$ avec $d'_{ij} = \frac{1}{d_{ij}}$ si $i = j$

Démonstration Résolvons l'équation $\begin{pmatrix} d_1 & 0 & \dots & 0 \\ 0 & d_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & d_n \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}$:

$\begin{cases} d_1x_1 = a_1 \\ d_2x_2 = a_2 \\ \dots \\ d_nx_n = a_n \end{cases} \iff \begin{cases} x_1 = \dfrac{1}{d_1}a_1 \\ x_2 = \dfrac{1}{d_2}a_2 \\ \dots \\ x_n = \dfrac{1}{d_n}a_n \end{cases}$. On en déduit finalement $A^{-1} = \begin{pmatrix} \frac{1}{d_1} & 0 & \dots & 0 \\ 0 & \frac{1}{d_2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \frac{1}{d_n} \end{pmatrix}$.

Exemple 72 (Inversion d'une matrice carrée) – Pour les matrices carrées d'ordre $n > 2$, dans le cas d'une matrice non diagonale, on peut généraliser et simplifier la méthode vue dans l'exemple 71 : au lieu de résoudre explicitement le système correspondant à l'équation $AX = B$, on travaille [1] directement par combinaison des lignes d'une matrice (AB) de dimension $n \times (n+1)$ (la matrice A à laquelle on a ajouté la matrice colonne B) pour obtenir une matrice $(A'B')$ correspondant à un système $A'X = B'$ équivalent à celui de départ.

Illustrons cette méthode en inversant la matrice 3×3 : $A = \begin{pmatrix} 1 & 2 & 2 \\ 3 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}$

Notre matrice (AB) est donc la matrice à 3 lignes et 4 colonnes $\begin{pmatrix} 1 & 2 & 2 & a \\ 3 & 1 & 0 & b \\ 1 & 1 & 1 & c \end{pmatrix}$

Effectuons $L_2 \longrightarrow L_2 - 3L_1$ et $L_3 \longrightarrow L_3 - L_1$, on obtient $\begin{pmatrix} 1 & 2 & 2 & a \\ 0 & -5 & -6 & b-3a \\ 0 & -1 & -1 & c-a \end{pmatrix}$

Effectuons $L_2 \longrightarrow -L_3$ et $L_3 \longrightarrow L_2$, on obtient $\begin{pmatrix} 1 & 2 & 2 & a \\ 0 & 1 & 1 & a-c \\ 0 & -5 & -6 & b-3a \end{pmatrix}$

Effectuons $L_3 \longrightarrow L_3 + 5L_2$, on obtient $\begin{pmatrix} 1 & 2 & 2 & a \\ 0 & 1 & 1 & a-c \\ 0 & 0 & -1 & b+2a-5c \end{pmatrix}$

Effectuons $L_1 \longrightarrow L_1 + 2L_3$ et $L_2 \longrightarrow L_2 + L_3$, on obtient $\begin{pmatrix} 1 & 2 & 0 & 5a+2b-10c \\ 0 & 1 & 0 & 3a+b-6c \\ 0 & 0 & -1 & 2a+b-5c \end{pmatrix}$

Effectuons $L_1 \longrightarrow L_1 - 2L_2$, on obtient $\begin{pmatrix} 1 & 0 & 0 & -a+2c \\ 0 & 1 & 0 & 3a+b-6c \\ 0 & 0 & -1 & 2a+b-5c \end{pmatrix}$

Donc l'équation $AX = B \iff A'X = B'$ admet pour unique solution $\begin{pmatrix} -a+2c \\ 3a+b-6c \\ -2a-b+5c \end{pmatrix}$

On en déduit la matrice inverse de A : $A^{-1} = \begin{pmatrix} -1 & 0 & 2 \\ 3 & 1 & -6 \\ -2 & -1 & 5 \end{pmatrix}$.

Remarquons que cette méthode se prête bien à la programmation.
Pour ne pas surcharger ce cours, je laisse l'écriture de ce programme au lecteur motivé.
Si on veut (et qu'on peut) utiliser la bibliothèque `numpy`, l'inversion d'une matrice `A` est immédiate avec la fonction `linalg.inv(A)`. J'utiliserai ce moyen ici pour vérifier mes résultats.

```
import numpy as np
A=np.array([[1,2,2],[3,1,0],[1,1,1]])
B=np.array([[1,2],[3,4]])
if np.linalg.det(A)!=0 : print(np.linalg.inv(A))
if np.linalg.det(A)!=0 : print(np.linalg.inv(B))
```

```
[[-1.  0.  2.]
 [ 3.  1. -6.]
 [-2. -1.  5.]]
[[-2.   1. ]
 [ 1.5 -0.5]]
```

Vous remarquerez que les matrices initiales ne contiennent que des entiers alors que les matrices inverses sont données en nombres flottants : les formules utilisées contenant des divisions, la conversion est systématique ... même si le résultat est un entier.
Pour éviter une erreur avec `numpy` si la matrice n'est pas inversible, tester le déterminant (`linalg.det(A)`) qui doit toujours être non nul.

1. Les opérations valides sont de trois types : ajouter à une ligne un multiple d'une autre ligne ; multiplier une ligne par un coefficient non nul ; intervertir deux lignes. L'idée est de faire apparaitre des zéros en-dessous de la diagonale principale de A puis en-dessus de cette diagonale afin de ne garder des coefficients non nuls que sur la diagonale.

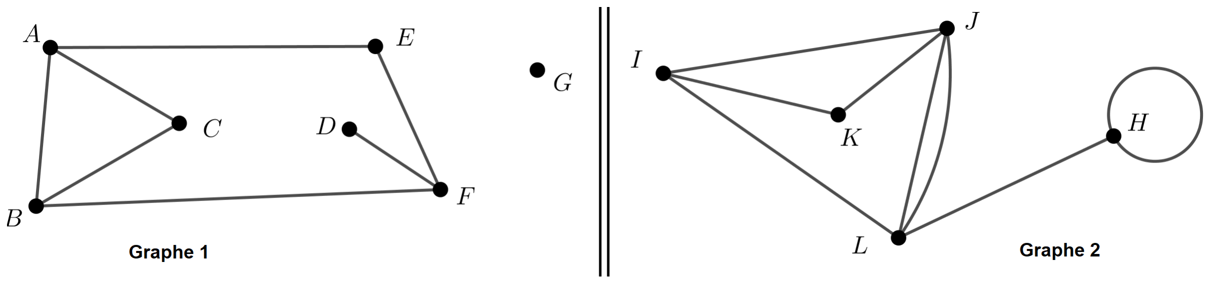
2. Typologie des Graphes

2.a. Graphes non orientés

Les graphes sont des modèles abstraits qui représentent des réseaux reliant entre eux des objets.

Définition 7.8 (Graphe) Un graphe est constitué de *sommets* reliés ou non par des *arêtes*.
Deux sommets reliés par une arête sont dits *adjacents* ; non reliés, ils sont *indépendants*.
L'*ordre* d'un graphe est son nombre de sommets ; la *taille* du graphe est son nombre d'arêtes.
Le *degré* d'un sommet est le nombre d'arêtes dont ce sommet est une extrémité.
Une *boucle* est une arête joignant un sommet à lui-même (augmente de 2 le degré du sommet).
Si les arêtes d'un graphe peuvent être parcourues dans les deux sens, le graphe est *non orienté*.
Un graphe est *simple* s'il est non orienté, fini, sans boucle ni arête multiple entre deux sommets.

Exemple 73 – Le graphe 1 est un graphe simple d'ordre 7 (7 sommets) et de taille 7 (7 arêtes). Les degrés des sommets vont de 0 (le sommet G est isolé) à 3 (sommets A, B et F), deux sommets sont de degré 2 (C et E). Le sommet D est le seul de degré 1.
La somme des degrés est $0 + 1 + 2 \times 2 + 3 \times 3 = 14$, soit 2 fois l'ordre du graphe.



Le graphe 2 n'est pas un graphe simple à cause de la boucle reliant le sommet H à lui-même et aussi l'arête double reliant les sommets L et J. L'ordre de ce graphe est 5 et sa taille 8.

Pour ces graphes on peut dresser une table des degrés :

Sommet	A	B	C	D	E	F	G
Degré	3	3	2	1	2	3	0

Sommet	H	I	J	K	L
Degré	3	3	4	2	4

On peut aussi établir leur matrice d'adjacence $M = (m_{ij})$ où le coefficient m_{ij} indique le nombre d'arêtes reliant les sommets d'indices i et j. Chaque sommet étant représenté par une ligne et une colonne de même indice, pour un graphe non orienté cette matrice est nécessairement symétrique.
Afin de clarifier la correspondance entre le nom (ici une lettre) du sommet et l'indice dans la matrice d'adjacence, on peut accompagner les lignes et colonnes de cette matrice par une entête indiquant les noms des sommets :

$$M_1 = \begin{array}{cc} & \begin{matrix} A & B & C & D & E & F & G \end{matrix} \\ \begin{matrix} A \\ B \\ C \\ D \\ E \\ F \\ G \end{matrix} & \begin{pmatrix} 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{array} \qquad M_2 = \begin{array}{cc} & \begin{matrix} H & I & J & K & L \end{matrix} \\ \begin{matrix} H \\ I \\ J \\ K \\ L \end{matrix} & \begin{pmatrix} 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 2 \\ 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 2 & 0 & 0 \end{pmatrix} \end{array}$$

Définition 7.9 (Matrice d'adjacence) Soit un graphe et ses sommets numérotés de 1 à n. La matrice d'adjacence associée à ce graphe d'ordre n est une matrice carrée $M = (m_{ij})$ de même dimension où le coefficient m_{ij} donne le nombre d'arêtes reliant le sommet i au sommet j.

Remarques :

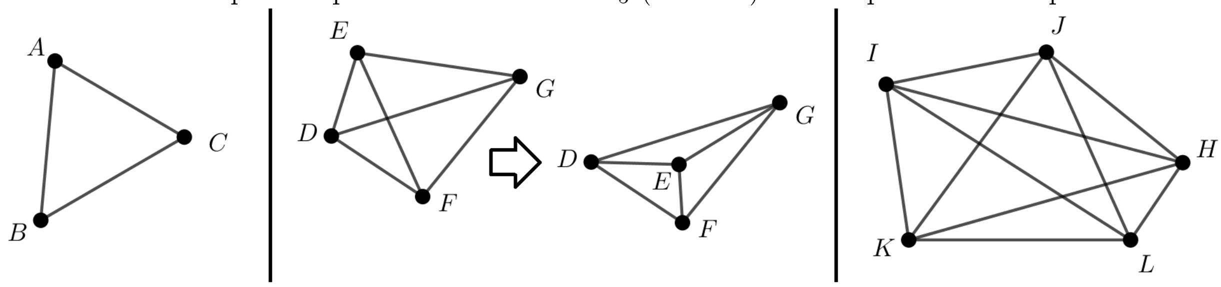
- La matrice d'adjacence associée à un graphe non-orienté est symétrique.
  Celle d'un graphe simple ne contient que des 0 et des 1.
- Dans un graphe non orienté la somme des degrés est un nombre pair (c'est le double du nombre d'arêtes) ; on en déduit que le nombre de sommets d'ordre impair est un nombre pair (sinon il y aurait un nombre impair de degrés).

Définition 7.10 (Graphe complet, Graphe planaire) Un graphe simple est

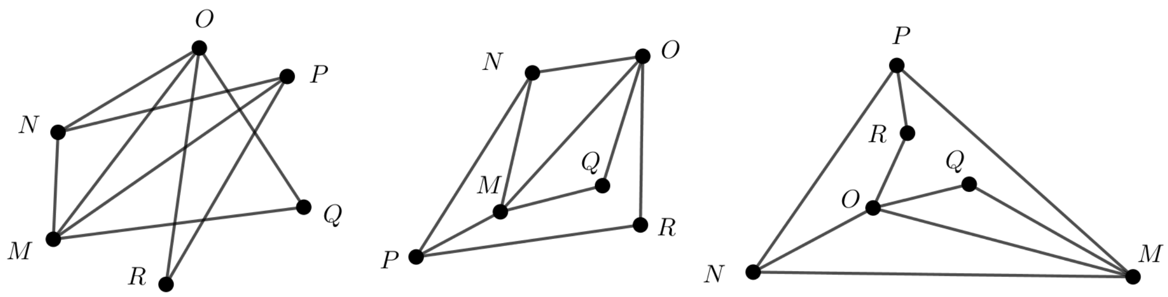
- *complet* si tous ses sommets sont adjacents. On note K_n un graphe complet à n sommets.
- *planaire* si on peut tracer ses arêtes sur un plan par des segments qui ne se croisent pas.

Remarques :

- K_3 (à gauche) est un graphe complet et planaire.
  K_4 (au centre) est complet. Bien qu'il semble non planaire, si on replace un de ses sommets au centre on constate qu'il est planaire. Par contre K_5 (à droite) est complet mais non planaire.



- Ce n'est pas forcément évident de décider si un graphe est planaire ou pas.
  En effet, les sommets peuvent être déplacés, les arêtes déformées, ce qui crée ou enlève des croisements.
  Ci-dessous, j'ai dessiné trois versions du même graphe planaire, mais la version de gauche comporte six croisements ce qui pourrait faire penser qu'il n'est pas planaire. Les deux versions de droite ne se ressemblent pas, mais montrent que ce graphe peut être dessiné sans croisement : il est donc planaire.



Définition 7.11 (Chaînes d'un graphe) Une *chaîne* est une suite finie de sommets adjacents constituant un chemin entre les deux sommets situés aux extrémités.
Une chaîne est *eulérienne* si elle passe une fois et une seule par toutes les arêtes du graphe.
Une chaîne est *hamiltonienne* si elle passe une fois et une seule par tous les sommets du graphe.
Une chaîne est *fermée* si ses deux extrémités sont identiques.
Une chaîne fermée est un *cycle* si toutes ses arêtes sont différentes.
Connexité d'un graphe :

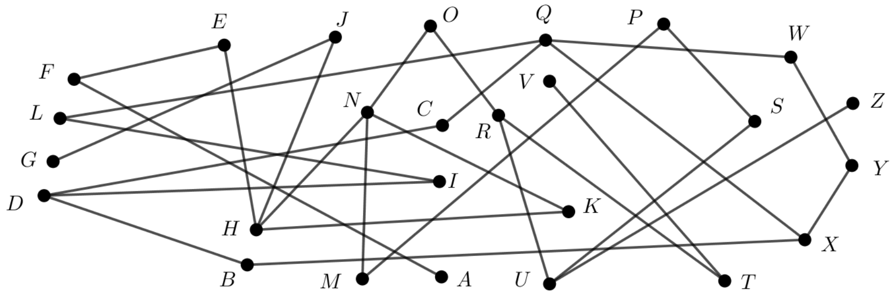
- Un graphe est *connexe* si entre deux sommets quelconques existe au moins une chaîne.
- Un *arbre* est un graphe connexe sans cycle ; un sommet de degré 1 est une *feuille*.

Caractéristiques métriques d'une chaîne ou d'un graphe :

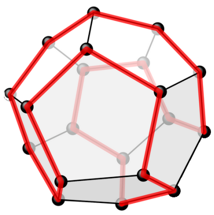
- La *longueur* d'une chaîne est le nombre de ses arêtes.
- La *distance* entre deux sommets est la plus courte longueur des chaînes qui les relient.
- Le *diamètre* d'un graphe est la plus grande distance entre deux sommets.

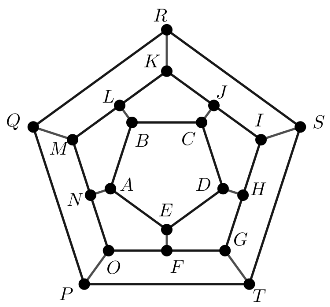
Remarques :

- Un cycle peut passer plusieurs fois par le même sommet. Sur le graphe planaire ci-dessus, le cycle $O-Q-M-O-N-M-O$ passe deux fois par le point O. Par contre la chaîne fermée $N-O-M-N-P-M-N$ n'est pas un cycle car elle passe deux fois par l'arête $M-N$.
- Tous les graphes donnés en exemple sont connexes excepté le graphe 1 qui possède un sommet isolé (G) : aucune chaîne ne permet de joindre ce sommet aux autres sommets du graphe.
  La connexité d'un graphe n'est pas toujours aussi facile à mettre en évidence. Ci-dessous j'ai dessiné un graphe non connexe mais cela est t-il si évident ? Combien de composantes (sous-graphe) connexes y discernez-vous ? (solution en fin de chapitre)

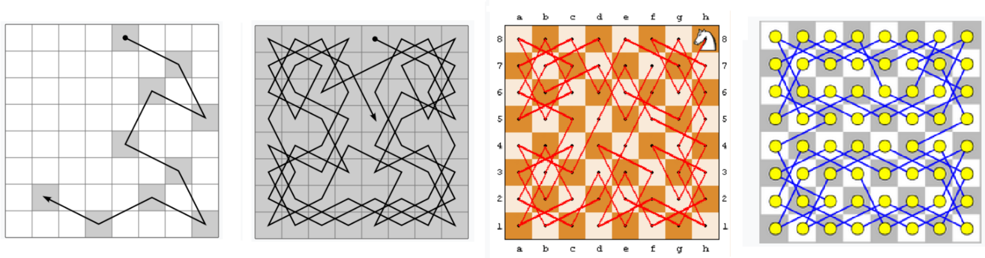


- Un cycle *eulérien* passe une fois et une seule par toutes les arêtes du graphe avec, à la fin, un retour au point de départ. Le problème initial posé à Euler à propos des sept ponts de Königsberg portait sur l'existence d'un tel chemin. Sa réponse négative était accompagnée d'une propriété générale (appelée *théorème de Euler*) : pour qu'un cycle eulérien existe, il suffit que tous ses sommets aient un degré pair supérieur ou égal à 2. Le graphe complet K_5 a tous ses sommets de degré 4 : le cycle $L-H-J-I-K-L-J-K-H-I-L$ par exemple qu'il contient est eulérien. Ce cycle n'est pas hamiltonien car il passe deux fois par chaque sommet.
- La connexité d'un graphe est une condition nécessaire pour qu'existe un cycle eulérien. Pour qu'existe une chaîne eulérienne (pas nécessairement fermée) il faut et il suffit qu'il soit connexe et que ses sommets d'ordre impair soient au nombre de 0 ou de 2. Ce n'est pas le cas pour le graphe des ponts de Königsberg (les sommets sont d'ordre 3 ou 5) donc il n'existe pas de chaîne eulérienne pour ce graphe. Le graphe planaire ci-dessus (deux sommets sont d'ordre 3, les autres sont d'ordre 2 ou 4) ne contient aucun cycle eulérien mais il possède au moins une chaîne eulérienne. Pour la trouver, il faut nécessairement partir d'un sommet d'ordre 3 et arriver de même : $N-M-Q-O-N-P-M-O-R-P$.
- Sur le graphe K_5 (ci-dessus), la chaîne $L-H-J-I-K-L$ est un cycle hamiltonien mais n'est évidemment pas une chaîne eulérienne. La recherche de cycles hamiltoniens dans un graphe remonte à Sir William Rowan Hamilton (1805-1865) qui s'interrogeait sur l'existence d'un tel cycle entre les sommets d'un dodécaèdre, les arêtes de ce solide étant les arêtes du graphe. Ci-dessous le problème dans l'espace et sur un graphe planaire qui le représente : la chaîne $A-B-C-D-E-F-G-H-I-J-K-L-M-N-O-P-Q-R-S-T$ est hamiltonienne mais ce n'est pas un cycle.
  Le cycle hamiltonien coloré en rouge sur les arêtes du dodécaèdre est
  $A-B-C-D-E-F-G-H-I-J-K-L-M-Q-R-S-T-P-O-N-A$.

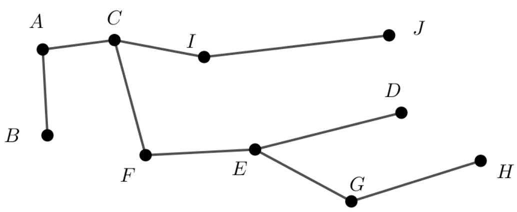


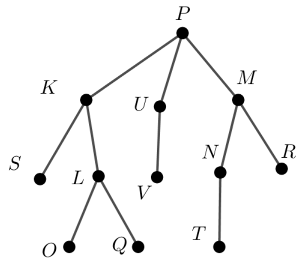


Le problème de l'échiquier, mentionné par le poète indien Rudrata au IXe siècle, est un autre exemple – plus difficile – de ce type de recherche : il s'agit de couvrir les 64 cases d'un échiquier par une chaîne ou, mieux, un cycle hamiltonien décrit par le déplacement d'un cavalier, celui-ci se déplaçant en « L » aux échecs (2 cases dans une direction et 1 perpendiculairement). La figure du centre-droit montre le chaîne hamiltonienne de al-Adli ar-Rumi (vers 840) ; celle de droite montre le cycle hamiltonien symétrique par rapport au centre de l'échiquier proposé par Euler en 1759.

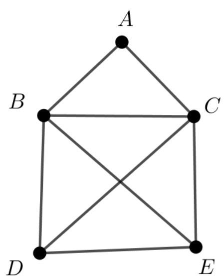


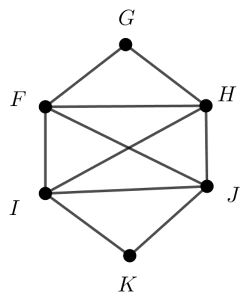
- Un arbre a toujours un sommet de plus que le nombre d'arêtes : ainsi l'arbre de gauche contient 9 arêtes et 10 sommets. On peut toujours aller d'un sommet à un autre (connexité) mais pour revenir à un sommet déjà visité il faut rebrousser chemin (acyclicité).
  On désigne parfois un des sommets comme étant la *racine* de l'arbre, ce qui conduit à une description du graphe en termes de généalogie : l'arbre de droite a été dessiné de manière à identifier le sommet P comme racine ; du coup, on parlera de la descendance (ou de l'ascendance) d'un des sommets. Le sommet K a deux fils (S et L) mais un seul père (P).

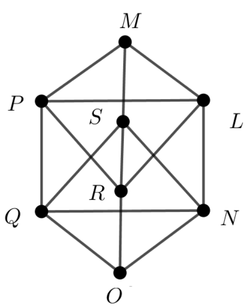


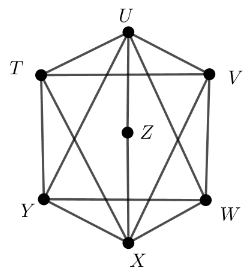


Exemple 74 – La recherche d'un cycle eulérien ou d'une chaîne eulérienne est un jeu bien connu. Sans connaître la condition nécessaire et suffisante à son existence, le joueur peut avoir un peu de mal à les détecter. Les trois graphes ci-dessous permette d'apprécier cette difficulté, relative mais croissante généralement avec le nombre de sommets.









➠ Le graphe de gauche contient une chaîne eulérienne partant de D et arrivant à E (ou l'inverse, il n'y a pas d'ordre ici) mais pas de cycle eulérien car il y a deux sommets d'ordre impair.
Il y a la chaîne $D - B - A - C - D - E - B - C - E$, y en a t-il d'autres partant de D ?

➠ Le graphe du centre-gauche a tous ses sommets d'ordre pair. Il doit donc contenir au moins un cycle eulérien. Je vais me baser sur le graphe précédent auquel on a juste ajouté un sommet en-dessous qui permet de revenir au point de départ :
$I - F - G - H - I - J - F - H - J - K - I$ est un cycle eulérien.

➠ Le graphe du centre-droit possède deux sommets d'ordre impair (3) et six sommets d'ordre pair (4). Comme celui de gauche, il doit contenir une chaîne eulérienne mais pas de cycle eulérien. En partant de M on doit arriver à O. La question est laissée en suspens (solution en fin de chapitre).

➠ Le graphe de droite a deux sommets d'ordre impair (5) et cinq sommets d'ordre pair (2 ou 4). Déterminer une chaîne eulérienne partant de U et arrivant à X (solution en fin de chapitre).

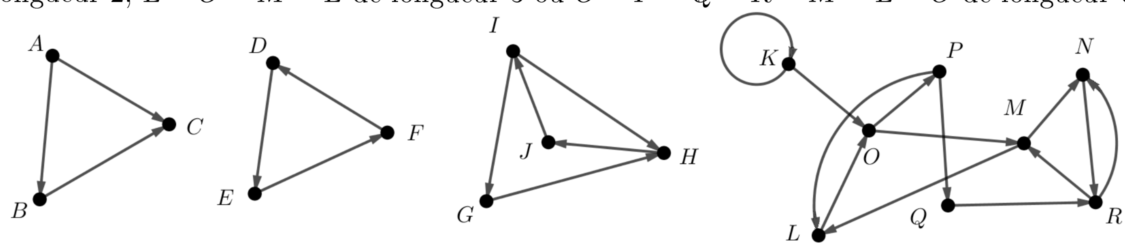
2.b. Graphes orientés

Les graphes non orientés ne suffisent pas toujours à rendre compte du problème qu'ils représentent.

DÉFINITION 7.12 (ORIENTATION) Un graphe orienté est constitué d'arêtes orientées (appelées *arcs*) qu'on ne peut parcourir que dans un seul sens.
Un arc orienté a un sommet de départ (appelé *origine*) et un sommet d'arrivée (appelée *extrémité*).

Remarques :

- De même que pour les graphes non orienté, on définit : une boucle orientée, une chaîne orientée, un cycle orienté, un graphe complet orienté, etc. Voici, à gauche, les deux seuls graphes complets orienté d'ordre 3 (les noms étant interchangeables) et, à droite, un graphe complet orienté d'ordre 4 et un graphe orienté connexe d'ordre 8.
  On remarque que le sommet K ne peut être atteint ; du coup on dira de ce graphe qu'il a une connexité *faible*, mais le sous-graphe (L, M, N, O, P, Q, R) en est une composante fortement connexe (connexité *forte* : tous les sommets peuvent être atteints).
  Il existe une boucle et de nombreux cycles orientés dans ce graphe, comme $R - N - R$ de longueur 2, $L - O - M - L$ de longueur 3 ou $O - P - Q - R - M - L - O$ de longueur 6.



- Dans la matrice d'adjacence $M = (m_{ij})$ associée à un graphe orienté, le coefficient m_{ij} donne le nombre d'arcs d'origine i et d'extrémité j. Ces matrices ne sont plus symétriques comme celles des graphes non orientés. Voici les quatre matrices associées aux graphes ci-dessus :

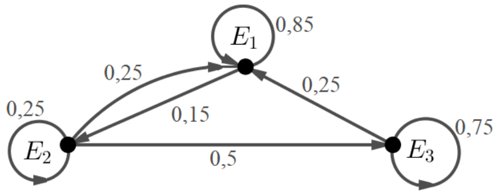
$$\begin{array}{cc} & \begin{array}{ccc} A & B & C \end{array} \\ \begin{array}{c} A \\ B \\ C \end{array} & \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \end{array} \quad \begin{array}{cc} & \begin{array}{ccc} D & E & F \end{array} \\ \begin{array}{c} D \\ E \\ F \end{array} & \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \end{array} \quad \begin{array}{cc} & \begin{array}{cccc} G & H & I & J \end{array} \\ \begin{array}{c} G \\ H \\ I \\ J \end{array} & \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \end{array} \quad \begin{array}{cc} & \begin{array}{cccccccc} K & L & M & N & O & P & Q & R \end{array} \\ \begin{array}{c} K \\ L \\ M \\ N \\ O \\ P \\ Q \\ R \end{array} & \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \end{pmatrix} \end{array}$$

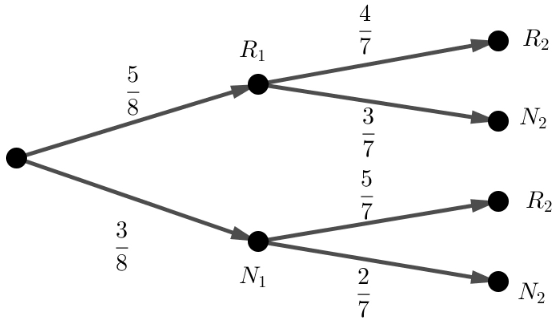
DÉFINITION 7.13 (PONDÉRATION) Un graphe *étiqueté* est un graphe dont les arêtes sont affectées à des étiquettes appartenant à un même ensemble (mots, nombres, symboles, couleurs, etc.) Si les étiquettes d'un graphe sont des nombres positifs (des « poids ») on parle de graphe *pondéré*.
Un graphe *probabiliste* est un graphe orienté et pondéré tel que la somme des poids des arêtes sortant de chaque sommet est égale à 1.

Remarques :

- Un graphe pondéré peut être orienté ou non orienté. De nombreux problèmes liés à ce type de graphe sont résolus à l'aide d'algorithmes. Le célèbre problème du voyageur de commerce en est un exemple : il s'agit de visiter un certain nombre de villes (une fois et une seule) et de revenir au point de départ par le chemin le plus court. L'itinéraire optimum est le cycle de poids minimum. Malheureusement, si le nombre de villes à visiter est trop grand un algorithme d'exploration systématique qui examine chaque itinéraire prendra trop de temps (pour n villes il y en a $\frac{(n-1)!}{2}$ à examiner, pour 10 villes cela fait déjà 181 400). Il existe cependant des algorithmes d'optimisation qui fournissent une solution approchée (ACO).

- Le poids d'une chaîne d'un graphe pondéré est la somme des poids de ses arêtes.
  Dans une application, ces poids peuvent être des longueurs (réseau autoroutier), dans une autre ce peut-être des coûts ou des durées (tâches à réaliser). Dans tous les cas, la *plus courte* chaîne qui relie deux sommets est celle de poids minimum.
- Les graphes probabilistes sont utilisés pour modéliser l'évolution d'un système pouvant changer aléatoirement d'état : les sommets du graphe sont les états possibles et le poids d'une arête menant de l'état E_i à l'état E_j est la probabilité de transition $P(E_iE_j)$.
  Voici, à gauche, un exemple où le système évolue entre trois états et, à droite, la représentation déjà bien connue d'un tirage sans remise de deux boules successivement d'une urne contenant au départ 5 boules rouges et 3 noires :





  Voici maintenant la matrice de transition M du graphe de gauche ainsi que deux puissances de cette matrice (M^2 et M^{10}) dont on va voir l'utilité (probabilités arrondies au centième). La matrice de transition du graphe de droite n'a pas grand intérêt car cet arbre ne peut qu'être descendu (par ailleurs il faudrait renommer les sommets qui portent actuellement des noms ambigus).

$$M = \begin{array}{c} \\ E_1 \\ E_2 \\ E_3 \end{array} \begin{array}{c} \begin{array}{ccc} E_1 & E_2 & E_3 \end{array} \\ \begin{pmatrix} 0,85 & 0,15 & 0,00 \\ 0,25 & 0,25 & 0,50 \\ 0,25 & 0,00 & 0,75 \end{pmatrix} \end{array} \qquad M^2 = \begin{array}{c} \begin{array}{ccc} E_1 & E_2 & E_3 \end{array} \\ \begin{pmatrix} 0,76 & 0,16 & 0,07 \\ 0,40 & 0,10 & 0,50 \\ 0,40 & 0,04 & 0,56 \end{pmatrix} \end{array} \qquad M^{10} = \begin{array}{c} \begin{array}{ccc} E_1 & E_2 & E_3 \end{array} \\ \begin{pmatrix} 0,63 & 0,13 & 0,25 \\ 0,62 & 0,12 & 0,25 \\ 0,62 & 0,12 & 0,25 \end{pmatrix} \end{array}$$

3. Matrices d'adjacence

3.a. Puissances d'une matrice d'adjacence

À partir d'un graphe on peut écrire sa matrice d'adjacence et, inversement, une matrice d'adjacence permet de reconstituer le graphe sous-jacent. Quel est l'intérêt de la matrice d'adjacence $M = (m_{ij})$? Le coefficient m_{ij} permet de savoir combien il y a de chemins permettant d'aller du sommet i au sommet j en 1 coup. On va voir que les puissances de cette matrice conserve cette propriété.

> Propriété 7.8 (Nombre de chemins) Soit M la matrice d'adjacence d'un graphe.
> Les coefficients $m_{ij}(n)$ de la matrice M^n (puissance n^e de M) correspondent au nombre de chemins permettant d'aller du sommet i au sommet j en n coups.

Démonstration On suppose une matrice d'adjacence M de dimension N et on fait une récurrence sur le nombre de coups n :

- On l'a rappelé pour $n = 1$: par définition de la matrice d'adjacence, le coefficient m_{ij} permet de savoir combien il y a de chemins permettant d'aller du sommet i au sommet j en 1 coup.
- Supposons que les coefficients $m_{ij}(n)$ de la matrice M^n correspondent au nombre de chemins permettant d'aller du sommet i au sommet j en n coups, et calculons le nombre de chemins permettant d'aller du sommet i au sommet j en $n+1$ coups.
  Pour aller du sommet i au sommet j en $(n+1)^e$ coup, il faut atteindre le sommet k au n^e coup, avec $k \in [\![1; N]\!]$, et ensuite aller de k à j. D'après l'hypothèse de récurrence il y a $m_{ik}(n)$ chemins permettant d'atteindre le sommet k en n coups et, ensuite, il y a

$m_{kj}(1) = m_{kj}$ chemins allant de k à j. Il y a donc en tout $m_{ik}(n) \times m_{kj}$ chemins permettant d'atteindre j en $n+1$ coups en passant par k au coup précédent.
Comme on veut totaliser ces chemins, il y en a $\sum_{k=1}^{n} m_{kj}(1) \times m_{kj}$ ce qui est exactement la formule pour calculer le coefficient $m_{ij}(n+1)$ de la matrice M^{n+1}.
Conclusion : Pour $n \geqslant 1$, les coefficients $m_{ij}(n)$ de la matrice M^n donnent le nombre de chemins permettant d'aller du sommet i au sommet j en n coups.

Remarques :

- Prenons l'exemple du graphe orienté complet d'ordre 4 ci-dessus (G, H, I, J). On peut facilement compter les chemins allant d'un sommet à l'autre en 2 ou 3 coups et vérifier que les nombres trouvés correspondent bien aux coefficients des matrices M^2 et M^3.

$$M = \begin{array}{c} \\ \begin{matrix} G \\ H \\ I \\ J \end{matrix} \end{array} \begin{array}{c} \begin{matrix} G & H & I & J \end{matrix} \\ \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \end{array} \qquad M^2 = \begin{array}{c} \\ \begin{matrix} G \\ H \\ I \\ J \end{matrix} \end{array} \begin{array}{c} \begin{matrix} G & H & I & J \end{matrix} \\ \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 \end{pmatrix} \end{array} \qquad M^3 = \begin{array}{c} \\ \begin{matrix} G \\ H \\ I \\ J \end{matrix} \end{array} \begin{array}{c} \begin{matrix} G & H & I & J \end{matrix} \\ \begin{pmatrix} 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \end{pmatrix} \end{array}$$

Par contre, c'est très fastidieux de compter les chemins allant d'un sommet à l'autre en 10 coups et quasiment inhumain de faire ce dénombrement pour 100 coups. À cette fin, on utilisera les matrices M^{10} et M^{100} calculées à l'aide de notre petit programme.

$$M^{10} = \begin{array}{c} \\ \begin{matrix} G \\ H \\ I \\ J \end{matrix} \end{array} \begin{array}{c} \begin{matrix} G & H & I & J \end{matrix} \\ \begin{pmatrix} 2 & 3 & 1 & 1 \\ 1 & 3 & 1 & 3 \\ 3 & 4 & 3 & 2 \\ 1 & 2 & 3 & 3 \end{pmatrix} \end{array} \qquad M^{100} = \begin{array}{c} \\ \begin{matrix} G \\ H \\ I \\ J \end{matrix} \end{array} \begin{array}{c} \begin{matrix} G & H & I & J \end{matrix} \\ \begin{pmatrix} 73\,237\,462 & 133\,231\,279 & 89\,404\,037 & 109\,139\,117 \\ 89\,404\,037 & 162\,641\,499 & 109\,139\,117 & 133\,231\,279 \\ 133\,231\,279 & 242\,370\,396 & 162\,641\,499 & 198\,543\,154 \\ 109\,139\,117 & 198\,543\,154 & 133\,231\,279 & 162\,641\,499 \end{pmatrix} \end{array}$$

- Pour un graphe de grande taille, pas besoin d'aller au delà du 2e coup pour apprécier l'intérêt des matrices M^n afin de dénombrer les chemins permettant d'atteindre un des sommets en n coups. Voici les matrices M^3 et M^{20} dans le cas du graphe orienté d'ordre 8 ci-dessus.

$$M^3 = \begin{array}{c} \\ \begin{matrix} K \\ L \\ M \\ N \\ O \\ P \\ Q \\ R \end{matrix} \end{array} \begin{array}{c} \begin{matrix} K & L & M & N & O & P & Q & R \end{matrix} \\ \begin{pmatrix} 1 & 2 & 1 & 1 & 1 & 1 & 1 & 0 \\ 0 & 2 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 2 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 2 & 0 & 0 & 2 \\ 0 & 0 & 2 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 & 0 & 0 & 1 \end{pmatrix} \end{array} \qquad M^{20} = \begin{array}{c} \\ \begin{matrix} K \\ L \\ N \\ N \\ O \\ P \\ Q \\ R \end{matrix} \end{array} \begin{array}{c} \begin{matrix} K & L & M & N & O & P & Q & R \end{matrix} \\ \begin{pmatrix} 1 & 2\,504 & 3\,205 & 3\,819 & 1\,579 & 1\,235 & 655 & 2\,772 \\ 0 & 926 & 926 & 926 & 926 & 926 & 926 & 926 \\ 0 & 926 & 1\,356 & 1\,468 & 344 & 580 & 234 & 802 \\ 0 & 745 & 1\,314 & 1\,424 & 968 & 463 & 172 & 1\,369 \\ 0 & 851 & 981 & 1\,215 & 401 & 388 & 229 & 794 \\ 0 & 1\,936 & 1\,146 & 2\,158 & 926 & 344 & 580 & 1\,702 \\ 0 & 745 & 1\,314 & 1\,424 & 968 & 463 & 172 & 1\,369 \\ 0 & 1\,369 & 1\,195 & 1\,775 & 851 & 401 & 388 & 1\,444 \end{pmatrix} \end{array}$$

Propriété 7.9 (Probabilités) Soit M la matrice d'adjacence d'un graphe probabiliste.
Les coefficients $m_{ij}(n)$ de la matrice M^n correspondent à la probabilité, en partant de l'état E_i, d'obtenir l'état E_j en n coups.

Démonstration C'est la formule des probabilités totales qui permet de justifier cette propriété, toujours par récurrence sur le nombre n de coups :

- Pour $n = 1$: par définition de la matrice d'adjacence d'un graphe probabiliste, le coefficient m_{ij} permet de connaître la probabilité, en partant de l'état E_i, d'obtenir l'état E_j en 1 coup.
- Supposons que les coefficients $m_{ij}(n)$ de la matrice M^n correspondent à la probabilité, en partant de l'état E_i, d'obtenir l'état E_j en n coups, et calculons la probabilité, en partant de l'état E_i, d'obtenir l'état E_j en $n+1$ coups.
  Pour aller de l'état E_i à l'état E_j en $(n+1)^{\text{e}}$ coup, il faut atteindre l'état E_k au n^{e} coup, avec $k \in [\![1; N]\!]$ (N est la dimension de M), et ensuite aller de l'état E_k à l'état E_j. D'après l'hypothèse de récurrence la probabilité de passer de l'état E_i à l'état E_k en n coups est $m_{ik}(n)$ et, ensuite, la probabilité de passer de l'état E_k à l'état E_j en 1 coup est $m_{kj}(1) = m_{kj}$. La probabilité de passer de l'état E_i à l'état E_j en passant par E_k au coup précédent est donc $m_{ik}(n) \times m_{kj}$.
  La formule des probabilités totales fait la somme de toutes ces probabilités pour $k \in [\![1; N]\!]$ et

ainsi on obtient $\sum_{k=1}^{n} m_{kj}(1) \times m_{kj}$ ce qui est exactement la formule pour calculer le coefficient $m_{ij}(n+1)$ de la matrice M^{n+1}.

Ainsi on a bien montré quelle était la signification des coefficients de la matrice M^n pour $n \geqslant 1$.

Remarques :

On comprend maintenant l'utilité des puissances de la matrice de transition M du graphe probabiliste ci-dessus (celui de gauche). On constate notamment que, après 10 coups, les probabilités de chacune des lignes ont tendance à se rapprocher d'une valeur d'équilibre.

Pour confirmer cette remarque, examinons des puissances encore plus grandes de la matrice, arrondies au millième :

$$M^{20} = \begin{array}{c|ccc} & E_1 & E_2 & E_3 \\ \hline E_1 & & & \\ E_2 & & & \\ E_3 & & & \end{array} \quad M^{20} = \begin{array}{c} \\ E_1 \\ E_2 \\ E_3 \end{array}\begin{array}{c} \begin{array}{ccc} E_1 & E_2 & E_3 \end{array} \\ \begin{pmatrix} 0.625 & 0.125 & 0.250 \\ 0.625 & 0.125 & 0.250 \\ 0.625 & 0.125 & 0.250 \end{pmatrix} \end{array} \qquad M^{100} = \begin{array}{c} \\ E_1 \\ E_2 \\ E_3 \end{array}\begin{array}{c} \begin{array}{ccc} E_1 & E_2 & E_3 \end{array} \\ \begin{pmatrix} 0.625 & 0.125 & 0.250 \\ 0.625 & 0.125 & 0.250 \\ 0.625 & 0.125 & 0.250 \end{pmatrix} \end{array}$$

L'influence de l'état initial se perd progressivement. À l'équilibre, le système décrit par ce graphe se trouve dans $0,625 = \frac{5}{8}$, soit $62,5\%$ des cas dans l'état E_1, $0,125 = \frac{1}{8}$, soit $12,5\%$ des cas dans l'état E_2 et $0,25 = \frac{1}{4}$, soit 25% des cas dans l'état E_3.

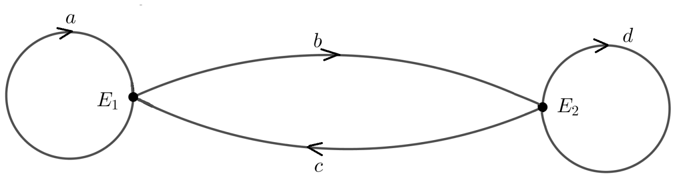
3.b. Chaîne de Markov

La notion de chaîne de Markov va préciser ce que nous avons expérimenté sur ces graphes probabilistes. Ce type d'objet est nommé en l'honneur d'Andreï Markov (1856-1922), mathématicien russe qui publia en 1906 son étude sur les processus *stochastiques* (relèvent du domaine de l'aléatoire, du calcul des probabilités) qui portent son nom aujourd'hui.

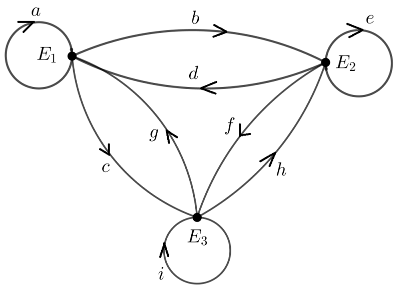
Définition 7.14 (Chaîne de Markov) Une *chaîne de Markov* à N états est une suite de variables aléatoires (X_n) à valeurs dans un espace d'états $E = \{E_1, E_2, \dots, E_N\}$ telle que pour tout $E_j \in E$, l'événement $X_{n+1} = E_j$ ne dépend que des événements de l'étape précédente (c'est-à-dire que des événements $X_n = E_i$ pour tout $E_i \in E$).

Remarques :

- Une chaîne de Markov à deux états est une séquence $X_0, X_1, X_2, X_3, \dots X_n$ de variables aléatoires à valeurs dans l'espace des états $E = \{E_1, E_2\}$, la valeur X_n étant l'état du processus à l'instant n. Cette chaîne modélise un processus stochastique dont la caractéristique est que la probabilité de $X_{n+1} = E_j$ dépend uniquement des $X_n = E_i, \forall E_i \in E$. Autrement dit, elle ne dépend pas de l'instant n. Étant invariantes tout au long de l'évolution du processus, on pourra donc donner les probabilités de transition d'un état à l'autre par une unique matrice $M = (m_{ij})$ où le coefficient m_{ij} donne la probabilité de passer de l'état E_i à l'état E_j.
- La distribution de probabilité d'une chaîne de Markov (X_n) à N états est une matrice-ligne qui donne les probabilités des différents états.
  On note $\pi_n = (P(X_n = E_1) \;\; P(X_n = E_2) \;\; \dots \;\; P(X_n = E_N))$ cette distribution de probabilité à l'étape n. D'après la règle de multiplication des matrices, on a $\pi_{n+1} = \pi_n M$ où M est la matrice de transition associée à la chaîne de Markov considérée ; une récurrence simple permet de déduire que l'on a $\pi_n = \pi_0 M^n$ où π_0 est la distribution initiale de probabilité.
  Généralement, le système est dans un état initial E_k d'où $P(X_0 = E_k) = 1$ et les autres probabilités sont à zéro ($\forall i \in [\![1, N]\!], i \neq k \Longrightarrow P(X_0 = E_i) = 0$).
- Une chaîne de Markov à N états est associée à un graphe à N sommets, ce graphe étant complet, orienté et probabiliste. Les pondérations des arcs sont les probabilités de transition d'un état à l'autre. Ci-dessous, j'ai représenté (à gauche) de tels graphes correspondants à une chaîne de Markov à 2 états (dessus) et à 3 états (dessous). Les matrices de transition (aussi appelées matrices stochastiques) correspondantes sont à droite.

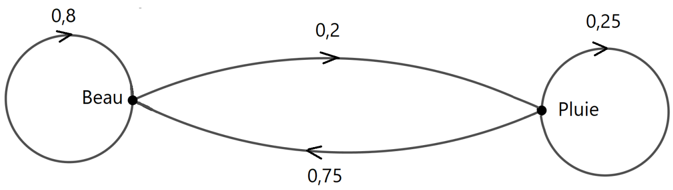


$$M = \begin{array}{c} \\ E_1 \\ E_2 \end{array}\begin{array}{c} \begin{array}{cc} E_1 & E_2 \end{array} \\ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \end{array}$$



$$M = \begin{array}{c} \\ E_1 \\ E_2 \\ E_3 \end{array}\begin{array}{c} \begin{array}{ccc} E_1 & E_2 & E_3 \end{array} \\ \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} \end{array}$$

Exemple 75 – On a observé la météo d'une région obéissait sensiblement au modèle suivant : s'il fait beau un jour il fera beau le lendemain 8 fois sur 10 et il pleuvra sinon ; par contre lorsqu'il pleut un jour, il fera beau une fois sur 4 le lendemain. Étudions ce processus stochastique en essayant de déterminer s'il fera beau dans n jours sachant qu'il fait beau aujourd'hui.
Le modèle de graphe probabiliste est le suivant avec sa matrice de transition M :



$$M = \begin{array}{c} \\ \text{Beau} \\ \text{Pluie} \end{array}\begin{array}{c} \begin{array}{cc} \text{Beau} & \text{Pluie} \end{array} \\ \begin{pmatrix} 0{,}8 & 0{,}2 \\ 0{,}75 & 0{,}25 \end{pmatrix} \end{array}$$

L'espace des états $E = \{Beau, Pluie\}$.
La distribution de probabilité initiale est, selon l'énoncé
$\pi_0 = (P(X_0 = Beau) = 1 \;\; P(X_0 = Pluie) = 0)$.
D'après la formule des probabilités totales, on a
$P(X_1 = Beau) = P(X_0 = Beau) \times 0,8 + P(X_0 = Pluie) \times 0,75 = 1 \times 0,8 + 0 \times 0,75 = 0,8$ et
$P(X_1 = Pluie) = P(X_0 = Beau) \times 0,2 + P(X_0 = Pluie) \times 0,25 = 1 \times 0,2 + 0 \times 0,25 = 0,2$.
On aurait aussi bien obtenu cela : $\pi_1 = (P(X_1 = Beau) = 0,8 \;\; P(X_1 = Pluie) = 0,2)$ en multipliant la matrice-ligne π_0 par la matrice de transition M.
De même, on obtient
$\pi_2 = (P(X_2 = Beau) = 0,8^2 + 0,2 \times 0,75 = 0,79 \;\; P(X_2 = Pluie) = 0,8 \times 0,2 + 0,2 \times 0,25 = 0,21)$ en multipliant la matrice-ligne π_1 par la matrice carrée de transition M ou bien en multipliant la matrice-ligne π_0 par la matrice M^2.
Ainsi de suite :
$\pi_3 = (P(X_3 = Beau) = 0,7895 \;\; P(X_3 = Pluie) = 0,21051)$,
$\pi_4 = (P(X_4 = Beau) = 0,789475 \;\; P(X_4 = Pluie) = 0,210525)$,
etc.
On constate que les probabilités se stabilisent assez rapidement, ces valeurs nous étant données par les puissances élevées de la matrice M :

$$M^4 = \begin{array}{c} \\ Beau \\ Pluie \end{array}\begin{array}{c} \begin{array}{cc} Beau & Pluie \end{array} \\ \begin{pmatrix} 0,789475 & 0,210525 \\ 0,78946875 & 0,21053125 \end{pmatrix} \end{array} \qquad M^{40} = \begin{array}{c} \\ Beau \\ Pluie \end{array}\begin{array}{c} \begin{array}{cc} Beau & Pluie \end{array} \\ \begin{pmatrix} 0,7894736842105267 & 0,21052631578947378 \\ 0,7894736842105271 & 0,2105263157894739 \end{pmatrix} \end{array}$$

Essayons d'obtenir des valeurs explicitement calculables.
En notant $\pi_n = (P(X_n = Beau) = p_n \;\; P(X_n = Pluie) = q_n)$, les probabilités p_n et $q_n = 1 - p_n$ satisfont la relation de récurrence :
$p_{n+1} = p_n \times 0,8 + q_n \times 0,75 = p_n \times 0,8 + (1 - p_n) \times 0,75 = p_n \times (0,8 - 0,75) + 0,75$,
soit $p_{n+1} = 0,05p_n + 0,75$.
La suite (p_n) est arithmético-géométrique (voir le chapitre Suites du cours de 1$^{\text{re}}$).

Pour en trouver l'expression explicite cherchons le point fixe, solution de l'équation
$x = 0,05x + 0,75 \iff 0,95x = 0,75 \iff x = \frac{75}{95} = \frac{15}{19} \approx 0,7894736842$ (la valeur limite).
Posons $u_n = p_n - \frac{75}{95}$ et montrons que la suite (u_n) est géométrique de raison $0,05$:
$u_{n+1} = p_{n+1} - \frac{75}{95} = 0,05p_n + 0,75 - \frac{75}{95} = 0,05(u_n + \frac{75}{95}) + 0,75 - \frac{75}{95} = 0,05u_n - \frac{0,95\times 75}{95} + 0,75 = 0,05u_n$.
Du coup, on obtient $u_n = u_0 \times 0,05^n = (p_0 - \frac{75}{95}) \times 0,05^n = \frac{20\times 0,05^n}{95} = \frac{4\times 0,05^n}{19}$ (car $p_0 = 1$).
On en déduit $p_n = u_n + \frac{15}{19} = \frac{15+4\times 0,05^n}{19}$ tandis que $q_n = 1 - p_n = 4 \times \frac{1-0,05^n}{19}$.
Vérifions : $p_1 = \frac{15+4\times 0,05}{19} = \frac{15,2}{19} = 0,8$ - $p_2 = \frac{15+4\times 0,05^2}{19} = 0,79$ - etc.
Finalement, le jour n, la distribution de probabilité de l'état météorologique de cette région est :
$\pi_n = \left(P(X_n = Beau) = \frac{15+4\times 0,05^n}{19} \quad P(X_n = Pluie) = 4 \times \frac{1-0,05^n}{19}\right)$

Définition 7.15 (Distribution invariante) Une distribution de probabilité π est dite *invariante* pour une chaîne de Markov de matrice de transition M si $\pi M = \pi$.

Définition 7.16 (Matrice régulière) Une matrice M est dite ***régulière*** s'il existe une valeur $m \geqslant 1$ telle que M^m ne contient que des coefficients strictements positifs.

Propriété 7.10 Si une chaîne de Markov a une matrice de transition M régulière alors

- Il existe une distribution de probabilité invariante π à coefficients strictement positifs.
- La suite des distributions de probabilité (π_n) converge vers π indépendamment de π_0.
- La suite de matrices (M^n) converge vers une matrice dont toutes les lignes restituent π.

Démonstration Lorsqu'il n'y a que 2 états, on refait les calculs de notre exemple 8.13 :
Distribution de probabilité $\pi_n = (p_n \quad q_n)$
Matrice de transition $M = \begin{pmatrix} 1-p & p \\ q & 1-q \end{pmatrix}$ avec $0 < p < 1$ et $0 < q < 1$.
$p_{n+1} = p_n \times (1-p) + q_n \times q = p_n \times (1-p) + (1-p_n) \times q = p_n \times (1-p-q) + q$.
La suite (p_n) est arithmético-géométrique, son point fixe est solution de l'équation :
$x = x \times (1-p-q) + q \iff (p+q)x = q \iff x = \frac{q}{p+q}$ (on a $p + q \neq 0$).
Posons $u_n = p_n - \frac{q}{p+q}$ et montrons que la suite (u_n) est géométrique de raison $(1-p-q)$:
$u_{n+1} = p_{n+1} - \frac{q}{p+q} = p_n \times (1-p-q) + q - \frac{q}{p+q} = u_n \times (1-p-q)$ (j'abrège).
Du coup, on obtient $u_n = u_0 \times (1-p-q)^n = (p_0 - \frac{q}{p+q}) \times (1-p-q)^n$.
On en déduit $p_n = u_n + \frac{q}{p+q} = (p_0 - \frac{q}{p+q}) \times (1-p-q)^n + \frac{q}{p+q}$
Comme $-1 < 1-p-q < 1$, on en déduit que $|1-p-q| < 1$ et donc $(1-p-q)^n$ tend vers 0 ce qui prouve que p_n tend vers $\frac{q}{p+q}$, le point fixe qui est donc la valeur limite de la suit (p_n).
De même $q_n = 1 - p_n$ tend vers $1 - \frac{q}{p+q} = \frac{p}{p+q}$.
La distribution de probabilité π_n converge vers la distribution invariante $\pi = \left(\frac{q}{p+q} \quad \frac{p}{p+q}\right)$.
Cette propriété est admise pour plus de 2 états.

Remarques :

- La matrice probabiliste donnée en exemple dans les remarques de la définition 7.13 a deux coefficients nuls, elle est cependant régulière car la matrice M^2 ne contient que des coefficients strictement positifs. On peut en déduire qu'il existe une distribution invariante qui formera les lignes de la matrice limite de (M^n). Cette distribution invariante étant $\pi = \left(\frac{5}{8} \ \frac{1}{8} \ \frac{1}{4}\right)$, la matrice limite est $\begin{array}{c c} & \begin{array}{ccc} E_1 & E_2 & E_3 \end{array} \\ \begin{array}{c} E_1 \\ E_2 \\ E_3 \end{array} & \begin{pmatrix} \frac{5}{8} & \frac{1}{8} & \frac{1}{4} \\ \frac{5}{8} & \frac{1}{8} & \frac{1}{4} \\ \frac{5}{8} & \frac{1}{8} & \frac{1}{4} \end{pmatrix} \end{array}$.

- Pour une chaîne de Markov a matrice de transition régulière, on montre que le temps moyen du retour à l'état E_i est $m_{ii} = \frac{1}{p_i}$ où p_i est l'élément correspondant à E_i dans la distribution invariante π. Dans notre exemple 8.13, on a $\pi = \begin{pmatrix} \frac{15}{19} & \frac{4}{19} \end{pmatrix}$. On en déduit que le temps d'attente du beau temps (la durée de la pluie) est $\frac{19}{15} \approx 1,3$ jours alors que le temps d'attente de la pluie (la durée du beau temps) est $\frac{19}{4} \approx 4,75$ jours.

4. Suite de matrices colonnes

Définition 7.17 (Convergence) Une suite de matrices colonnes (M_n) de taille $N \times 1$ est une fonction qui a tout entier naturel n associe la matrice colonne M_n où les coefficients $m_i(n)$ sont les termes de suites numériques (m_i) pour tout $i \in [\![1, N]\!]$.
On dit que la suite de matrices colonnes (M_n) *converge* lorsque chacune des suites (m_i) converge et on note M la matrice limite constituée des limites des suites (m_i).

Remarques :
- On définit de même des suites de matrices lignes de taille $1 \times N$ ou des suites de matrices carrées de taille $N \times N$.
- La fonction U définie sur $\mathbb{N}$ par $n \longmapsto \begin{pmatrix} n-1 \\ e^{-n}-1 \\ 1-\frac{1}{n+1} \end{pmatrix}$ définit une suite de matrices colonnes de taille 3×1. On a notamment $U_0 = \begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix}$, $U_1 = \begin{pmatrix} 0 \\ \frac{1}{e}-1 \\ \frac{1}{2} \end{pmatrix}$.
  Deux des suites numériques utilisées ici convergent : $(e^{-n}-1)$ converge vers -1 tandis que $(1-\frac{1}{n+1})$ converge vers 1. Cependant la suite (U_n) diverge ; elle ne converge pas car la suite $(n-1)$ diverge vers $+\infty$.

Propriété 7.11 (Suites géométriques) Soit A une matrice carrée de dimension N et (M_n) une suite de matrices colonnes de taille $N \times 1$ définie, pour tout entier $n \in \mathbb{N}$ par $\begin{cases} M_{n+1} = AM_n \\ M_0 \end{cases}$
La suite (M_n) est géométrique et on a $\forall n \in \mathbb{N}, M_n = A^n M_0$.
De plus, si la suite (M_n) converge alors la matrice colonne limite M vérifie l'égalité $M = AM$.

Démonstration Par récurrence sur l'entier n :
- Pour $n = 0$, on a bien $M_0 = A^0 M_0$ car la matrice A^0 est par convention égale à la matrice identité I_N. Cela initialise cette propriété.
- En supposant que pour le rang n on ait $M_n = A^n M_0$, calculons $M_{n+1} = AM_n = A \times (A^n M_0)$ mais l'associativité du produit matriciel implique alors que $M_{n+1} = (A \times A^n)M_0 = A^{n+1}M_0$. Cela assure de l'hérédité de cette propriété.

Si la suite (M_n) converge et admet pour limite M alors la suite (AM_n) converge vers la limite AM tandis que la suite (M_{n+1}) converge vers la limite M (c'est la suite que (M_n) avec des indices décalés). Mais (M_{n+1}) et (AM_n) sont une seule et même suite, elles ont donc la même limite :
$M = AM$.

Exemple 76 – Soit $M = \begin{pmatrix} 1 & 2 \\ -1 & 0 \end{pmatrix}$.

Intéressons nous à (X_n), la suite géométrique de matrices colonnes définie par $\begin{cases} X_0 = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \\ X_{n+1} = MX_n \end{cases}$

On peut calculer les premiers coefficients d'un terme de la suite à l'aide de la formule de récurrence mais pour un terme quelconque on privilégiera la formule explicite qui nécessite le calcul préalable de M^n. Dans le module de calcul de votre calculatrice il existe sans doute un moyen d'effectuer ces calculs : sur Numworks, on déclare une nouvelle matrice dans la caisse à outils et ensuite, on peut utiliser le produit ou l'élévation à la puissance.

Par exemple $X_1 = \begin{pmatrix} 1 & 2 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 5 \\ -1 \end{pmatrix}$.

Pour X_2 on peut continuer $X_2 = \begin{pmatrix} 1 & 2 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} 5 \\ -1 \end{pmatrix} = \begin{pmatrix} 3 \\ -5 \end{pmatrix}$ ou bien on utilise $M^2 = \begin{pmatrix} -1 & 2 \\ -1 & -2 \end{pmatrix}$.

Pour X_{10} on va utiliser la matrice $M_{10} = \begin{pmatrix} 23 & -22 \\ 11 & 34 \end{pmatrix}$ et $X_{10} = M_{10}X_0 = \begin{pmatrix} -21 \\ 79 \end{pmatrix}$.

De même $X_{40} = M_{40}X_0 = \begin{pmatrix} -703\,889 & -1\,506\,054 \\ 753\,027 & 49\,138 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} -3\,715\,997 \\ 851\,303 \end{pmatrix}$.

Remarquons que cette suite (X_n) semble diverger.

Propriété 7.12 (Suites arithmético-géométriques) Soit A une matrice carrée de dimension N, B une matrice colonne de dimension $N \times 1$ et (M_n) une suite de matrices colonnes de taille $N \times 1$ définie, pour tout entier $n \in \mathbb{N}$ par $\begin{cases} M_{n+1} = AM_n + B \\ M_0 \end{cases}$

La suite (M_n) est arithmético-géométrique et, si il existe une matrice colonne M de dimension $N \times 1$ telle que $M = AM + B$, alors on a $\forall n \in \mathbb{N}, M_n = A^n(M_0 - M) + M$.

De plus, si la matrice $I_N - A$ est inversible alors la matrice M est égale à $(I_N - A)^{-1}B$.

Démonstration Dans les conditions de l'énoncé, pour tout entier $n \in \mathbb{N}$ on a $\begin{cases} M_{n+1} = AM_n + B \\ M = AM + B \end{cases}$

Par soustraction des matrices on obtient $M_{n+1} - M = (AM_n + B) - (AM + B) = A(M_n - M)$.

En posant $U_n = M_n - M$, la suite (U_n) est géométrique puisque $U_{n+1} = AU_n$.

On en déduit $U_n = A^nU_0 = A^n(M_0 - M)$, soit

$M_n - M = A^n(M_0 - M) \iff M_n = A^n(M_0 - M) + M$.

Si la matrice $I_N - A$ est inversible, comme $M = AM + B \iff (I_N - A)M = B$, en multipliant par $(I_N - A)^{-1}$ on obtient l'expression de la matrice $M = (I_N - A)^{-1}B$.

Remarques :

- L'existence d'une matrice M telle que $M = AM + B$ n'implique pas forcément que la suite (M_n) converge. Il peut exister une telle matrice (un état *stable* de la suite (M_n)) dans le cas d'une suite divergente. Mais si la suite (M_n) converge alors l'état stable M en est sa limite.
- Si l'état initial est stable ($M_0 = M$) alors la suite (M_n) est constante (donc convergente). Si l'état initial n'est pas stable ($M_0 \neq M$) et si la suite (A^n) converge vers une matrice limite A alors la suite (M_n) converge vers la matrice limite $A(M_0 - M) + M$.

Exemple 77 – Soit $A = \begin{pmatrix} 0,8 & 0 \\ 0 & 1,5 \end{pmatrix}$ et $B = \begin{pmatrix} 2 \\ -1 \end{pmatrix}$.

Intéressons nous à (X_n), la suite arithmético-géométrique définie par $\begin{cases} X_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ X_{n+1} = AX_n + B \end{cases}$

On peut calculer les premiers termes de la suite (X_n) :

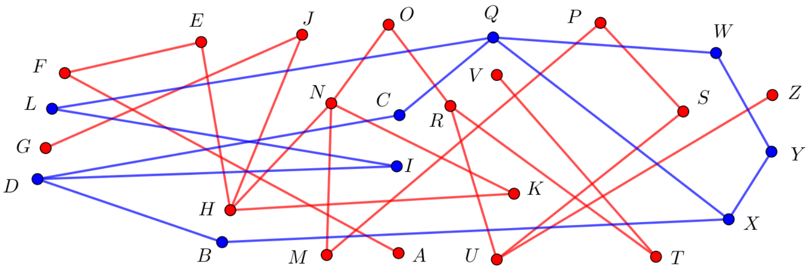
$X_1 = \begin{pmatrix} 0,8 & 0 \\ 0 & 1,5 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 2 \\ -1 \end{pmatrix} = \begin{pmatrix} 2,8 \\ -1 \end{pmatrix}$.

$X_2 = \begin{pmatrix} 0,8 & 0 \\ 0 & 1,5 \end{pmatrix} \begin{pmatrix} 2,8 \\ -1 \end{pmatrix} + \begin{pmatrix} 2 \\ -1 \end{pmatrix} = \begin{pmatrix} 4,24 \\ -2,5 \end{pmatrix}$.

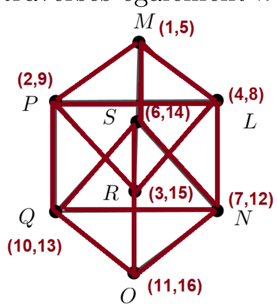
Si on veut X_{10} on s'interroge sur la matrice stable X vérifiant $X = AX + B \Longleftrightarrow (I_N - A)X = B$.
La matrice $I_N - A$ est inversible car I_N et A étant diagonales $I_N - A$ l'est aussi.
$I_N - A = \begin{pmatrix} 1-0,8 & 0 \\ 0 & 1-1,5 \end{pmatrix} = \begin{pmatrix} 0,2 & 0 \\ 0 & -0,5 \end{pmatrix}$ et $(I_N - A)^{-1} = \begin{pmatrix} \frac{1}{0,2} & 0 \\ 0 & \frac{1}{-0,5} \end{pmatrix} = \begin{pmatrix} 5 & 0 \\ 0 & -2 \end{pmatrix}$.
Par conséquent $X = (I_N - A)^{-1}B = \begin{pmatrix} 5 & 0 \\ 0 & -2 \end{pmatrix}\begin{pmatrix} 2 \\ -1 \end{pmatrix} = \begin{pmatrix} 10 \\ 2 \end{pmatrix}$.
Comme A est diagonale on a $A^n = \begin{pmatrix} 0,8^n & 0 \\ 0 & 1,5^n \end{pmatrix}$.
Finalement, on peut calculer $X_n = A^n(X_0 - X) + X = \begin{pmatrix} 0,8^n & 0 \\ 0 & 1,5^n \end{pmatrix}\begin{pmatrix} -9 \\ -2 \end{pmatrix} + \begin{pmatrix} 10 \\ 2 \end{pmatrix}$.
Vérifions cette formule $X_2 = \begin{pmatrix} 0,8^2 & 0 \\ 0 & 1,5^2 \end{pmatrix}\begin{pmatrix} -9 \\ -2 \end{pmatrix} + \begin{pmatrix} 10 \\ 2 \end{pmatrix} = \begin{pmatrix} 4,24 \\ -2,5 \end{pmatrix}$ ce qui est est correct.
Calculons maintenant $X_{10} = \begin{pmatrix} 0,8^{10} & 0 \\ 0 & 1,5^{10} \end{pmatrix}\begin{pmatrix} -9 \\ -2 \end{pmatrix} + \begin{pmatrix} 10 \\ 2 \end{pmatrix} \approx \begin{pmatrix} 9,0336 \\ -113,33 \end{pmatrix}$.
On remarque que cette suite semble diverger.
Existe t-il une valeur initiale $X_0 = \begin{pmatrix} a \\ b \end{pmatrix}$ qui rende cette suite convergente ?
On aura $X_n = \begin{pmatrix} 0,8^n & 0 \\ 0 & 1,5^n \end{pmatrix}\begin{pmatrix} a-10 \\ b-2 \end{pmatrix} + \begin{pmatrix} 10 \\ 2 \end{pmatrix} = \begin{pmatrix} 0,8^n(a-10)+10 \\ 1,5^n(b-2)+2 \end{pmatrix}$.
On observe que, seul, le coefficient $1,5^n(b-2)+2$ peut diverger car $0,8^n$ a pour limite 0.
Ainsi, pour que la suite converge X_n converge, il suffit de prendre $b = 2$ ainsi la suite de terme général $X_n = \begin{pmatrix} 0,8^n(a-10)+10 \\ 2 \end{pmatrix}$ converge vers $X = \begin{pmatrix} 10 \\ 2 \end{pmatrix}$.

Solutions des questions restées en suspens :

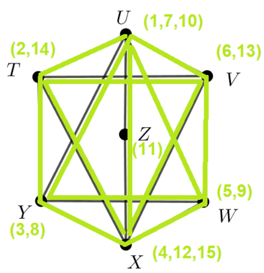
- Il y a deux composantes connexes dans ce graphe. Je les ai mises en évidence par des couleurs.



- Les cycles euleriens solutions sont sans doute nombreux.
  On remarque que les sommets d'ordre pair $2k$ sont traversés k fois par la chaîne solution ; les deux sommets d'ordre impair $2k+1$ constituent une extrémité de la chaîne solution et sont traversés également k fois par cette chaîne.



Chaîne eulérienne
M-P-R-L-M-S-N-L-P-Q-O-N-Q-S-R-O



Chaîne eulérienne
U-T-Y-X-W-V-U-Y-W-U-Z-X-V-T-X

Troisième partie

Aléatoirement

8 Probabilités

Objectifs :

- Somme de variables aléatoires, linéarité de l'espérance
- Linéarité de la variance dans une succession d'épreuves indépendantes
- Épreuve de Bernoulli, schéma de Bernoulli ; loi de probabilité associée
- Espérance et variance d'une variable aléatoire suivant une loi binomiale
- Représentation graphique d'une loi binomiale $\mathscr{B}(n; p)$ et programmation
- Échantillon, espérance et variance d'une somme de variables indépendantes identiques
- Inégalité de Bienaymé-Tchebychev, inégalité de concentration et loi des grands nombres

Aperçu historique : L'étude des probabilités telles que nous les connaissons est apparue avec la notion de risque au XIIe siècle, pour l'établissement de contrats commerciaux, puis au XVIe siècle avec la généralisation des contrats d'assurance maritime. Mais le véritable début de la théorie des probabilités date des travaux de Pierre de Fermat (1601-1665), de Blaise Pascal (1623-1662) et de Christian Huygens (1629-1695) dans la deuxième moitié du XVIIe siècle : ils s'intéressaient principalement à la théorie de jeux. Le terme probabilités prend alors son sens actuel, et Jacques Bernoulli (1654-1705)) traite ce sujet comme une théorie mathématique. Professeur à l'université de Bâle à partir de 1682, il rédige *Ars Conjectandi*, « l'art de conjecturer » (ouvrage publié à titre posthume par son neveu Nicolas en 1713) où il consolide et développe la théorie des probabilités, énonçant notamment, pour la 1re fois, la loi des grands nombres. Il y expose également sa théorie des permutations et des combinaisons, qui constitue aujourd'hui les fondements de la combinatoire.

Abraham de Moivre (1667-1754), français exilé à Londres, publie en 1718 *The doctrine of chances* (la Théorie du hasard) où il donne, entre autre, une première définition de l'indépendance statistique. Thomas Bayes (1702-1761) publie à titre posthume son *Essay Towards Solving a Problem in the Doctrine of Chances* en 1763 qui contient sa fameuse loi. Au XVIIIe siècle, Gabriel Cramer (1704-1752) donne un cours sur la logique probabiliste. Jean Le Rond d'Alembert (1717-1783), encyclopédiste notoire, s'est intéressé aux théories probabilistes de son époque mais c'est plus tard qu'elles acquièrent leur cohérence avec la synthèse magistrale de Pierre-Simon de Laplace (1749-1827), le grand savant de l'époque napoléonienne, qui publie en 1812 sa Théorie analytique des probabilités.

L'inégalité de Bienaymé-Tchebychev est une inégalité de concentration qui permet de montrer qu'une variable aléatoire ne prend que rarement une valeur relativement lointaine de son espérance. Ce résultat ne nécessite la connaissance que de l'espérance et de la variance, s'applique dans des cas variés et conduit à une justification de la loi faible des grands nombres. C'est le mathématicien français Irénée-Jules Bienaymé (1796-1878) qui énonça cette inégalité et le russe Pafnouti Tchebychev (1821-1894) qui la démontra.
Ce n'est donc qu'au XIXe siècle que la théorie moderne des probabilités apparaît ; prolongée au début du XXe siècle par l'axiomatique de Andreï Kolmogorov (1903-1987) qui propose un cadre formel rigoureux pour la théorie des probabilités, permettant enfin que celle-ci devienne une branche des mathématiques à part entière. Kolmogorov publie notamment en 1933 les *Fondements de la théorie des probabilités* où il définit les axiomes de la théorie actuelle des probabilités. La description actuelle, en termes d'univers, s'est donc imposée il y a un peu moins d'un siècle.

1. Variables aléatoires

1.a. Rappels

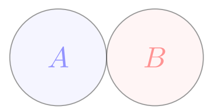
On considère une expérience aléatoire conduisant à n issues, et on note $\Omega = \{\omega_1, \omega_2, \dots, \omega_n\}$ l'ensemble des issues. Ω est « l'univers » des possibles. On définit la probabilité P en donnant la probabilité de chaque issue, c'est-à-dire les nombres $P(\omega_1), P(\omega_2), \dots, P(\omega_n)$ vérifiant :

- $\forall i \in \{0, 1, \dots, n\}, 0 \leqslant P(\omega_i) \leqslant 1$
- $\sum_{i=1}^{i=n} P(\omega_i) = P(\omega_1) + P(\omega_2) + \dots + P(\omega_n) = 1$

DÉFINITION 8.1 Un événement A est une partie de Ω.
La probabilité de A, notée $P(A)$, est la somme des probabilités des issues appartenant à A.

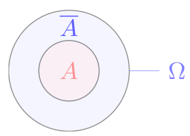
Remarques :

- A est un événement impossible si et seulement si $A = \emptyset$.
  Dans ce cas $P(A) = 0$
- A est un événement certain si et seulement si $A = \Omega$.
  Dans ce cas $P(A) = 1$



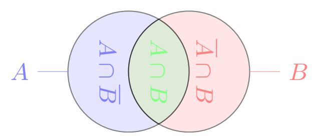
A et B incompatibles

- A et B sont incompatibles si $A \cap B$ est impossible, c'est-à-dire si $P(A \cap B) = 0$. Dans ce cas $P(A \cup B) = P(A) + P(B)$.
  L'événement $A \cup B$ est réalisé quand A **ou** B est réalisé. Ce « ou » est, en principe, inclusif mais pour des événements incompatibles, il prend le sens d'un ou exclusif.
- A et B sont des événements contraires si $A \cap B = \emptyset$ et $A \cup B = \Omega$.
  Dans ce cas on note $B = \overline{A}$ et on a $P(\overline{A}) = 1 - P(A)$.



A et $\overline{A}$ contraires

- pour des événements quelconques A et B, on a
  $$P(A \cup B) + P(A \cap B) = P(A) + P(B)$$
  Analysons la situation en termes d'ensembles incompatibles :
  $A = (A \cap B) \cup (A \cap \overline{B})$ et $B = (A \cap B) \cup (\overline{A} \cap B)$,
  d'où $A \cup B = [(A \cap \overline{B}) \cup (A \cap B)] \cup (\overline{A} \cap B)$
  et donc $P(A \cup B) = P(A) + P(\overline{A} \cap B)$,
  mais, comme $P(\overline{A} \cap B) = P(B) - P(A \cap B)$,
  on a finalement $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.



A et B compatibles

DÉFINITION 8.2 Lorsque toutes les issues ont la même probabilité – on dit qu'il y a équiprobabilité des issues – s'il y a n issues possibles, la probabilité de chacune est $\frac{1}{n}$.

Remarque :
On reconnait une situation d'équiprobabilité au fait que les issues sont interchangeables. Par exemple, un dé cubique bien équilibré a six faces interchangeables. Que l'une ou l'autre d'entre elles apparaisse lors d'un tirage, cela arrive avec la même probabilité de $\frac{1}{6}$. Quand on lance une pièce équilibrée, on obtient une face ou l'autre avec la même probabilité de $\frac{1}{2}$ car les deux faces sont interchangeables.

PROPRIÉTÉ 8.1 Dans le cas d'une situation d'équiprobabilité à n issues possibles, un événement A qui se réalise pour k issues a une probabilité de $P(A) = \frac{k}{n}$. On résume cela en écrivant

$$P(A) = \frac{\text{nombre de cas favorables à A}}{\text{nombre de cas possibles}}$$

La propriété 8.1 suppose que l'on puisse dénombrer les cas favorables et les cas possibles. Le dénombrement n'est pas toujours aisé et manipule parfois de très grands nombres. Le 1^er^ chapitre de ce cours développe les notions utiles pour dénombrer.

DÉFINITION 8.3 (VARIABLE ALÉATOIRE) Si, à chaque issue ω_i de Ω, on peut associer un réel $X(\omega_i)$, on dit que l'on a défini une variable aléatoire X sur Ω. Les probabilités associées aux différentes valeurs de X est appelée « loi » de la variable X.

Remarques :

- Une variable aléatoire est une fonction définie sur Ω ; ce n'est donc pas une variable (ce sont les issues qui sont variables ici) et ce n'est pas non plus aléatoire puisque chaque issue ω_i est associée à la valeur $X(\omega_i)$ de façon bien déterminée (le hasard n'a pas sa place dans cette association).
- Une variable aléatoire permet de définir des évènements. Si l'expérience aléatoire est un jeu dans lequel un joueur gagne ou perd de l'argent, on écrira $X = 5$ l'évènement réalisé lorsque le gain du joueur est 5 (l'unité monétaire est généralement précisée par l'énoncé). L'évènement $X \leqslant 0$ correspond à une perte, tandis que $10 \leqslant X < 20$ correspond à un gain appartenant à l'intervalle $[10; 20[$. etc.
- La notation ensembliste est modifiée puisqu'en notant $X = x$, en réalité, on désigne l'ensemble des issues ω_i telles que $X(\omega_i) = x$, c'est-à-dire $\{\omega_i \in \Omega, X(\omega_i) = x\}$.
- La loi de probabilité de X est une fonction $f : k \mapsto P(X = k)$. Cette fonction agit sur l'ensemble, noté $X(\Omega)$, des valeurs possibles de X. On désigne parfois cette fonction par le terme de « loi de distribution » de X ou encore de « densité » de X.
- La somme de toutes les probabilités $P(X = k)$ quand k décrit l'ensemble $X(\Omega)$ est égal à 1, puisque les évènements $X = k$ sont deux à deux disjoints, leur réunion formant Ω.

DÉFINITION 8.4 (PARAMÈTRES) Soient X une variable aléatoire telle que $X(\Omega) = \{x_1, x_2, \ldots, x_n\}$. En notant $p_i = P(X = x_i)$ les différentes valeurs de la loi de probabilité de X :

- L'espérance de X est le réel, noté $E(X)$, défini par $E(X) = \sum\limits_{i=1}^{n} p_i x_i$
- La variance de X est le réel positif, noté $V(X)$, défini par $V(X) = \sum\limits_{i=1}^{n} p_i \Big(x_i - E(X)\Big)^2$
- L'écart-type de X est le réel positif, noté $\sigma(X)$, défini par $\sigma(X) = \sqrt{V(X)}$

Remarques :

- L'espérance mathématique d'une variable aléatoire est un *paramètre central*, équivalent probabiliste de la notion statistique de moyenne. De même, la variance et l'écart-type sont des mesures de la *dispersion* que l'on calcule aussi en statistiques. Si on réalise l'expérience aléatoire un grand nombre de fois, les paramètres statistiques (fréquences, moyenne) se rapprochent des résultats théoriques (probabilités, espérance).
- Dans un jeu où X est associé aux gains possibles, $E(X)$ représente le gain moyen que l'on peut espérer. Le jeu est dit équitable si $E(X) = 0$, favorable au joueur si $E(X) > 0$ et défavorable au joueur si $E(X) < 0$.

PROPRIÉTÉ 8.2 Soient X une variable aléatoire et a et b deux réels. Les paramètres de la variable aléatoire $X' = aX + b$ sont $E(X') = E(aX + b) = aE(X) + b$ et $V(X') = V(aX + b) = a^2 V(X)$.

DÉMONSTRATION Ces propriétés découlent des propriétés de la multiplication et de l'addition, ainsi que des définitions, notamment $\sum p_i = 1$, $\sum p_i x_i = E(X)$ et $\sum p_i (x_i - E(X))^2 = V(X)$:

$$\begin{aligned}
E(aX + b) &= (ax_1 + b) \times p_1 + (ax_2 + b) \times p_2 + \ldots + (ax_n + b) \times p_n \\
&= (ax_1 p_1 + ax_2 p_2 + \ldots + ax_n p_n) + (bp_1 + bp_2 + \ldots + bp_n) \\
&= a(x_1 p_1 + x_2 p_2 + \ldots + x_n p_n) + b(p_1 + p_2 + \ldots + p_n) \\
&= aE(X) + b \times 1 = aE(X) + b
\end{aligned}$$

$$\begin{aligned} V(aX+b) &= \sum_{i=1}^{n} p_i(ax_i+b-E(aX+b))^2 = \sum_{i=1}^{n} p_i(ax_i+b-aE(X)-b)^2 \\ &= \sum_{i=1}^{n} p_i(ax_i-aE(X))^2 = \sum_{i=1}^{n} p_i \times a^2 \times (x_i-E(X))^2 \\ &= a^2 \times \sum_{i=1}^{n} p_i \times (x_i-E(X))^2 = a^2 V(X) \end{aligned}$$

Conséquence : On a donc aussi $\sigma(aX+b) = \sqrt{V(aX+b)} = \sqrt{a^2V(X)} = |a|\sqrt{V(X)} = |a|\sigma(X)$

Propriété 8.3 (Autre formule pour la variance) Soit X une variable aléatoire.
La variance de X peut se calculer avec la formule $V(X) = E(X^2) - \left(E(X)\right)^2$

Démonstration

$$\begin{aligned} V(X) &= \sum_{i=1}^{n} p_i\left(x_i - E(X)\right)^2 = \sum_{i=1}^{n} p_i x_i^2 - 2\sum_{i=1}^{n} p_i x_i E(X) + \sum_{i=1}^{n} p_i \left(E(X)\right)^2 \\ &= E(X^2) - 2\sum_{i=1}^{n} p_i x_i E(X) + \sum_{i=1}^{n} p_i E(X) \times E(X) = E(X^2) - \sum_{i=1}^{n}\left(p_i E(X)\,(2x_i - E(X))\right) \\ &= E(X^2) - E(X)\left(\sum_{i=1}^{n} 2p_i x_i - \sum_{i=1}^{n} p_i E(X)\right) = E(X^2) - E(X)\left(2E(X) - E(X)\sum_{i=1}^{n} p_i\right) \\ &= E(X^2) - E(X)\,(2E(X) - E(X) \times 1) = E(X^2) - E(X) \times E(X) = E(X^2) - \left(E(X)\right)^2 \end{aligned}$$

Remarque : Cette propriété est un cas particulier d'un théorème connu sous le nom de « König-Huygens » qui dit que $E((X-a)^2) = V(X) + (E(X)-a)^2$.
En effet, en prenant $a = 0$, on obtient $E(X^2) = V(X) + \left(E(X)\right)^2$.
Pour démontrer ce théorème, on peut partir de l'égalité $x_i - a = x_i - E(X) + E(X) - a$ qui, élevée au carré s'écrit $(x_i-a)^2 = (x_i-E(X))^2 + (E(X)-a)^2 + 2 \times (x_i - E(X))(E(X)-a)$.
Multiplions ensuite cette nouvelle égalité par p_i et sommons les quantités obtenues pour toutes les valeurs de i entre 1 et n. On obtient :

$$E((X-a)^2) = V(X) + (E(X)-a)^2 + 2 \times (E(X)-a) \times \sum_{i=1}^{n} p_i(x_i - E(X))$$

Le dernier terme de cette somme est nul car
$\sum_{i=1}^{n} p_i(x_i - E(X)) = \sum_{i=1}^{n} p_i x_i - E(X)\sum_{i=1}^{n} p_i = E(X) - E(X) = 0$, d'où le résultat.

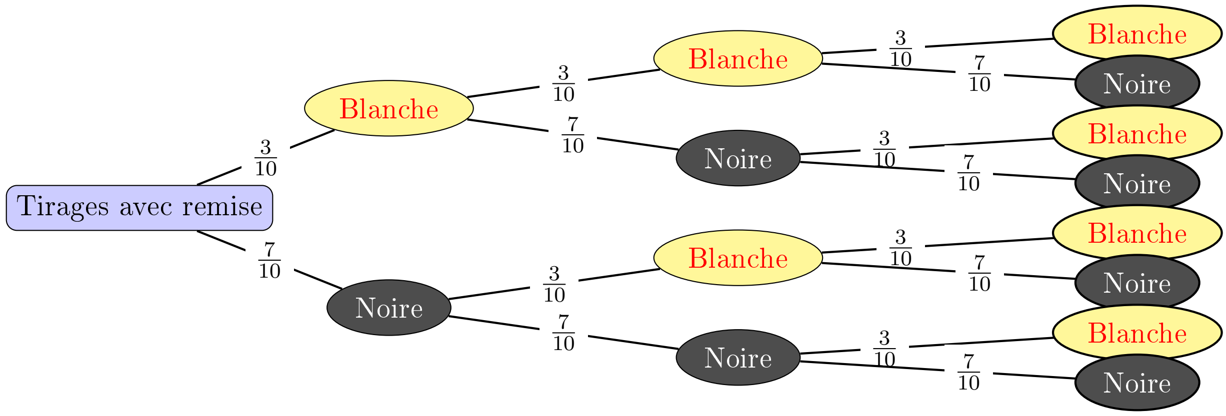
Cette propriété facilite le calcul algorithmique de $V(X)$ puisque l'on n'a pas besoin de connaître « à l'avance » la valeur de $E(X)$: on calcule en parallèle $\sum p_i x_i$ (pour l'obtention de $E(X)$) et $\sum p_i x_i^2$ (pour l'obtention de $E(X^2)$). Supposons que, à l'issue d'un tirage de deux dés, on cherche la variance de la somme X des dés. On peut calculer simultanément :
$E(X) = \sum_{i=1}^{12} p_i x_i = 2 \times \frac{1}{36} + 3 \times \frac{2}{36} + 4 \times \frac{3}{36} + \cdots + 12 \times \frac{1}{36} = \frac{252}{36} = 7$
$E(X^2) = \sum_{i=1}^{12} p_i x_i^2 = 2^2 \times \frac{1}{36} + 3^2 \times \frac{2}{36} + 4^2 \times \frac{3}{36} + \cdots + 12^2 \times \frac{1}{36} = \frac{1974}{36}$
Ensuite, on effectue $E(X^2) - \left(E(X)\right)^2 = \frac{1974}{36} - \left(\frac{252}{36}\right)^2 = \frac{7560}{36^2} = \frac{35}{6}$.

1.b. Expériences aléatoires successives

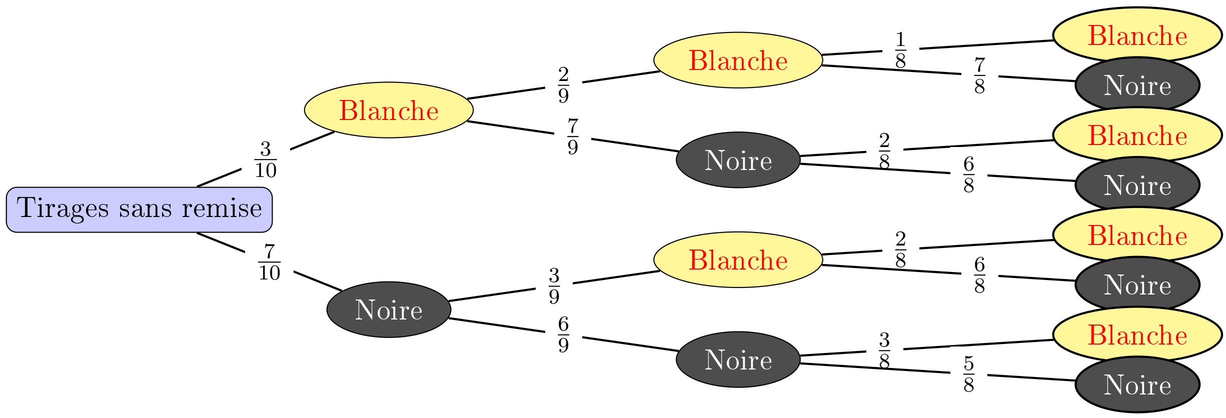
DÉFINITION 8.5 Deux expériences sont indépendantes si le résultat de la première n'influence pas le résultat de la seconde, c'est-à-dire qu'il ne modifie pas la loi de probabilité de celle-ci.

Exemples :

- Je lance un dé à six faces (dé cubique) en même temps qu'un dé à quatre faces (dé tétraédrique). Ces deux expériences simultanées sont indépendantes car le résultat du dé cubique n'influence pas le résultat du dé tétraédrique. Même chose si je lance deux dés cubiques : chacun des dés indique un résultat indépendant de l'autre. Le processus expérimental importe peu : je peux lancer les dés ensemble ou séparément, sans changer la loi de probabilité des résultats.
- Si je lance une pièce de monnaie jusqu'à obtenir un « pile », les tirages successifs sont indépendants, la pièce n'ayant pas de mémoire : ce n'est pas parce que j'ai déjà obtenu 5 fois « face » que j'ai plus de chance de faire « pile » au sixième lancé. Les lancés sont indépendants. L'existence du sixième lancé, par contre, dépend du résultat du cinquième puisqu'on ne l'effectue qu'à condition d'avoir obtenu un « face. »
- Dans un vidage d'urne (situation souvent choisie car elle permet de modéliser des problèmes de probabilités très divers), si on remet la boule tirée après avoir noté ce qui nous intéresse (sa couleur, sa taille, son numéro éventuels), les tirages successifs sont indépendants.
  L'urne ne changeant pas de composition, la loi de probabilité reste inchangée d'une étape à l'autre.



Si on ne remet pas la boule qui a été tirée, chacun des tirages influe sur la loi de probabilité de la boule suivante. Les tirages ne sont pas indépendants.



DÉFINITION 8.6 (ISSUES D'UNE SUCCESSION D'EAI) Soient n EAI successives (EAI : Expériences Aléatoires Indépendantes) d'univers Ω_1, Ω_2, …, Ω_n. L'univers résultant de cette succession d'expériences aléatoires est le produit cartésien $\Omega = \Omega_1 \times \Omega_2 \times \cdots \times \Omega_n$.
Les issues de Ω sont les n-uplets $(x_1, x_2, \ldots, x_n)$ où $\forall i \in [\![1, n]\!],\ \ x_i \in \Omega_i$.

Remarques : Une succession d'expériences aléatoires est représentée par un graphe étiqueté appelé *arbre de probabilités* : les nœuds (sommets) sont les issues de ces expériences et les étiquettes numériques inscrites sur les branches (arêtes, arcs) sont les probabilités de ces évènements.
Sur un arbre, les issues sont souvent appelées « chemin ». Que les expériences soient dépendantes ou indépendantes les unes des autres, ces arbres vérifient toujours trois règles :

- Règle d'univers : la somme des probabilités des branches partant d'un même nœud vaut 1.
  En partant d'un même nœud, on représente les différentes issues d'une expérience aléatoire. Ces issues sont incompatibles et forment l'univers Ω des possibles de cette expérience. La somme des probabilités vaut donc 1.
- Règle multiplicative : la probabilité d'un chemin est le produit des probabilités inscrites sur ses branches
- Règle additive : la probabilité d'un évènement réalisé par plusieurs chemins est la somme des probabilités de ces chemins

Propriété 8.4 (Règle multiplicative) Dans une succession de n EAI, la probabilité d'une issue $(x_1, \ldots, x_n)$ est égale au produit des probabilités des issues de ses composantes $x_1, \ \ldots, \ x_n$.

Démonstration Considérons la succession de deux épreuves et notons $P(B/A)$ la probabilité de réaliser B à la seconde sachant que A a été réalisé à la première. La probabilité du chemin $A-B$, c'est-à-dire de l'issue (A, B) (le nœud A suivi du nœud B) est celle de l'évènement $A \cap B$. Elle vaut $P(A \cap B) = P(A) \times P(B/A)$, soit le produit des probabilités inscrites sur les branches de ce chemin (revoir les probabilités conditionnelles dans le cours de 1^{re}).
Dans le cas où les expériences successives sont indépendantes, on a $P(A/B) = P(A)$ et par conséquent $P(A \cap B) = P(A) \times P(B)$, la loi multiplicative se vérifie toujours.
La généralisation à n EAI successives se justifie de la même façon.

Propriété 8.5 (Règle additive) Dans une succession de n EAI, la probabilité d'un évènement réalisé par plusieurs chemins est la somme des probabilités de ces chemins.

Démonstration La définition 8.1 de la probabilité d'un évènement implique que si plusieurs chemins réalisent un évènement particulier, alors la probabilité de cet évènement est la somme des probabilités de ces chemins, ceux-ci étant naturellement incompatibles.

Exemple 78 – Une urne contient 3 boules blanches et 7 noires. On tire successivement de cette urne 3 boules, avec remise de la boule tirée. Les tirages sont donc indépendants. Cette succession de trois EAI indépendantes a été représentée par l'arbre « Tirages avec remise » ci-dessus. On peut également raisonner à partir des univers : pour chaque EAI $\Omega_1 = \Omega_2 = \Omega_3 = \{B, N\}$ d'où l'univers pour leur succession :
$\Omega = \Omega_1 \times \Omega_2 \times \Omega_3 = \{(B,B,B),(B,B,N),(B,N,B),(B,N,N),(N,B,B),(N,B,N),(N,N,B),(N,N,N)\}$
Déterminons la loi de la variable aléatoire X comptant le nombre de boules blanches obtenues :
➟ $X = 0$ est l'évènement « obtenir 3 boules noires » qui n'est réalisé que par l'issue (N, N, N).
$P(X=0) = \left(\frac{7}{10}\right)^3 = \frac{343}{1000} = 0,343$.
➟ $X = 1$ est l'évènement « obtenir 2 boules noires et 1 blanche ».
Cet évènement est l'ensemble $\{(B,N,N),(N,B,N),(N,N,B)\}$.
$P(X=1) = 3 \times \frac{3}{10} \times \left(\frac{7}{10}\right)^2 = \frac{441}{1000} = 0,441$.
➟ De même, on montre que $P(X=2) = \frac{189}{1000} = 0,189$ et $P(X=3) = \frac{27}{1000} = 0,027$.

Vérifions ces résultats et calculons les paramètres de X : $\sum_{k=0}^{3} P(X=k) = \frac{343+441+189+27}{1000} = 1$
$E(X) = \frac{0\times373+1\times441+2\times189+3\times27}{1000} = \frac{900}{1000} = 0,9$ et $E(X^2) = \frac{0\times373+1\times441+4\times189+9\times27}{1000} = \frac{1440}{1000} = 1,44$.
D'où $V(X) = E(X^2) - E(X^2) = 1,44 - 0,9^2 = 0,63$ et $\sigma(X) = \sqrt{0,63} \approx 0,79$.

1.c. Espérance et variance de $X+Y$

Les variables aléatoires envisagées sont finies [1] et définies sur un même univers Ω.
De plus, pour le calcul de $V(X+Y)$, les variables sont indépendantes.

PROPRIÉTÉ 8.6 (ESPÉRANCE DE $X+Y$) Soient X et Y deux variables aléatoires.
L'espérance de la variable $X+Y$ est $E(X+Y)=E(X)+E(Y)$

DÉMONSTRATION On utilisera des notations particulières pour simplifier :

- $X(\Omega)=\{x_i\}_{i\in[\![0,n]\!]}$ et $Y(\Omega)=\{x_j\}_{j\in[\![0,p]\!]}$ sont les ensembles des valeurs prises par X et Y
- les $p_{ij}=P((X=x_i)\cap(Y=y_j))$ sont les probabilités *conjointes* du couple (X,Y)
- les $p_i=\sum\limits_{j=0}^{n}p_{ij}$ et $p_j=\sum\limits_{i=0}^{p}p_{ij}$ sont les probabilités *marginales* pour X et Y
- les couples (i,j) prennent leurs valeurs dans l'ensemble $\Pi=[\![0,n]\!]\times[\![0,p]\!]$

On déduit les espérances des variables X et Y des probabilités conjointes :
$E(X)=\sum\limits_{i=0}^{n}x_ip_i=\sum\limits_{i=0}^{n}x_i(\sum\limits_{j=0}^{p}p_{ij})=\sum\limits_{(i,j)\in\Pi}x_ip_{ij}$ et $E(Y)=\sum\limits_{j=0}^{p}y_jp_j=\sum\limits_{j=0}^{p}y_j(\sum\limits_{i=0}^{n}p_{ij})=\sum\limits_{(i,j)\in\Pi}y_jp_{ij}$
Par conséquent $E(X)+E(Y)=\sum\limits_{(i,j)\in\Pi}x_ip_{ij}+\sum\limits_{(i,j)\in\Pi}y_jp_{ij}=\sum\limits_{(i,j)\in\Pi}(x_i+y_j)p_{ij}$

$Z=X+Y$ est la variable aléatoire qui admet l'espérance $E(Z)=\sum\limits_{z\in Z(\Omega)}zP(Z=z)$.
Comme l'évènement $X+Y=z$ peut s'obtenir de plusieurs façons, regroupons les couples d'indices (i,j) qui conduisent à une même valeur de x_i+y_j dans un ensemble :
$\forall z\in Z(\Omega),\quad I_z=\{(i,j)\in\Pi, x_i+y_j=z\}$
Ces ensembles I_z formant une *partition* de Π, on en déduit
$E(X)+E(Y)=\sum\limits_{z\in Z(\Omega)}(\sum\limits_{(i,j)\in I_z}(x_i+y_j)p_{ij})=\sum\limits_{z\in Z(\Omega)}(\sum\limits_{(i,j)\in I_z}zp_{ij})=\sum\limits_{z\in Z(\Omega)}z(\sum\limits_{(i,j)\in I_z}p_{ij})$
Mais comme $\sum\limits_{(i,j)\in I_z}p_{ij}=P(X+Y=z)=P(Z=z)$
on en conclut $E(X)+E(Y)=\sum\limits_{z\in Z(\Omega)}zP(Z=z)=E(Z)$

PROPRIÉTÉ 8.7 (VARIANCE DE $X+Y$) Soient X et Y deux variables aléatoires indépendantes.
La variance de la variable $X+Y$ est $V(X+Y)=V(X)+V(Y)$

DÉMONSTRATION Par définition de la variance $V(X+Y)=E\left((X+Y-E(X+Y))^2\right)$
Or $E(X+Y)=E(X)+E(Y)$, donc $V(X+Y)=E\left([(X-E(X))+(Y-E(Y))]^2\right)$
Développons $V(X+Y)=E\left((X-E(X))^2+(Y-E(Y))^2+2(X-E(X))(Y-E(Y)\right)$
On a donc $V(X+Y)=V(X)+V(Y)+E\left(2(X-E(X))(Y-E(Y)\right)=V(X)+V(Y)+2\mathrm{cov}(X,Y)$
La quantité $E\left((X-E(X))(Y-E(Y)\right)$ est appelée covariance de X et Y. On la note $\mathrm{cov}(X,Y)$.
D'après les propriétés de linéarité de l'espérance :
$\mathrm{cov}(X,Y)=E\left(XY-XE(Y)-YE(X)+E(X)E(Y)\right)=E(XY)-E(X)E(Y)$

Montrons que la covariance est nulle lorsque les deux variables sont indépendantes :
$E(XY)=\sum\limits_{(i,j)\in\Pi}x_iy_ip_{ij}$, et l'indépendance se traduisant par $p_{ij}=p_ip_j$, on obtient
$E(XY)=\sum\limits_{(i,j)\in\Pi}x_iy_ip_ip_j=\sum\limits_{(i,j)\in\Pi}(x_ip_i)(y_ip_j)=\sum\limits_{i\in[\![0,n]\!]}x_ip_i\times\sum\limits_{i\in[\![0,p]\!]}y_ip_j=E(X)E(Y)$.
Ainsi $\mathrm{cov}(X,Y)=E(XY)-E(X)E(Y)=0$.
Par conséquent, si X et Y sont indépendantes on a
$V(X+Y)=V(X)+V(Y)+2\mathrm{cov}(X,Y)=V(X)+V(Y)$

1. X est une variable aléatoire finie si le nombre d'issues, c'est-à-dire $\mathrm{card}X(\Omega)$, est fini.

Remarques :

- L'espérance possède les deux propriétés de linéarité : $E(X+Y) = E(X)+E(Y)$ et, pour tout réel a, $E(aX) = aE(X)$ (cela a été montré dans la propriété 8.2). Par conséquent pour tout couple de réels (a,b) on a $E(aX+bY) = aE(X)+bE(Y)$. La variance ne possède que la linéarité additive ($V(X+Y) = V(X)+V(Y)$) et encore, cela n'est vrai que lorsque les deux variables sont indépendantes. Par rapport à la multiplication par un réel a, on a $V(aX) = a^2V(X)$ (cela a été aussi été montré dans la propriété 8.2).
- Si X est une variable aléatoire, la variable $X-E(X)$ est *centrée*, c'est-à-dire $E(X-E(X)) = 0$ et $V(X-E(X)) = V(X)$; la variable $\frac{X-E(X)}{\sigma_X}$ est centrée et *réduite*, c'est-à-dire $E(\frac{X-E(X)}{\sigma_X}) = 0$ et $V(\frac{X-E(X)}{\sigma_X}) = 1$.
- La covariance de deux variables indépendantes est nulle mais la réciproque est fausse. Une covariance nulle n'indique pas nécessairement que les variables sont indépendantes (pour qu'elles le soient, on doit avoir $p_{ij} = p_ip_j$ pour tout couple $(i,j) \in \Pi$).

DÉFINITION 8.7 (ÉCHANTILLON) Soit $m \geqslant 2$ un entier.
Le m-uplet $(X_1, X_2, \ldots, X_m)$ de m variables aléatoires identiques (définies sur un même univers Ω et ayant une même loi de probabilité) et indépendantes est un *échantillon* de taille m.

PROPRIÉTÉ 8.8 (VARIABLES ASSOCIÉES À UN ÉCHANTILLON) Soit un échantillon $(X_1, \ldots, X_m)$ de variables d'espérance $E(X_i) = \bar{x}$ et de variance $V(X_i) = \sigma^2$ pour tout $i \in [\![1, m]\!]$.

- La variable $S_m = X_1+X_2+\cdots+X_m$ a pour caractéristiques : $E(S_m) = m\bar{x}$ et $V(S_m) = n\sigma^2$
- La variable $M_m = \frac{X_1+X_2+\cdots+X_m}{m}$ a pour caractéristiques : $E(M_m) = \bar{x}$ et $V(M_m) = \frac{\sigma^2}{m}$

DÉMONSTRATION La linéarité additive de E et V implique que :
$E(S_m) = E(X_1+X_2+\cdots+X_m) = E(X_1)+E(X_1)+\cdots+E(X_m) = m\bar{x}$
$V(S_m) = V(X_1+X_2+\cdots+X_m) = V(X_1)+V(X_1)+\cdots+V(X_m) = m\sigma^2$
La linéarité multiplicative de E implique que : $E(M_m) = E(\frac{S_m}{m}) = \frac{E(S_m)}{m} = \frac{m\bar{x}}{m} = \bar{x}$
Par contre $V(M_m) = V(\frac{S_m}{m}) = \frac{V(S_m)}{m^2} = \frac{m\sigma^2}{m^2} = \frac{\sigma^2}{m}$

2. Loi binomiale

2.a. Épreuve de Bernoulli

DÉFINITION 8.8 Une expérience aléatoire admettant deux issues est appelée épreuve de Bernoulli.

- l'une des issues est appelée « succès », notée S ; sa probabilité est p.
- l'autre issue est appelée « échec », notée $\overline{S}$; sa probabilité est $1-p$

On dit que la variable aléatoire X prenant la valeur 1 en cas de succès et la valeur 0 en cas d'échec suit une *loi de Bernoulli* de paramètre p, notée $b(p)$. On appelle X une variable de Bernoulli.

PROPRIÉTÉ 8.9 Soit X une variable aléatoire de Bernoulli de paramètre p.
Les paramètres de X sont : $E(X) = p$; $V(X) = p\times(1-p)$; $\sigma(X) = \sqrt{V(X)} = \sqrt{p\times(1-p)}$

DÉMONSTRATION On a : $E(X) = 0\times(1-p)+1\times p = p$.
De même : $V(X) = E(X^2)-(E(X))^2 = 0^2\times(1-p)+1^2\times p-p^2 = p(1-p)$

Exemple : Tirer une boule de l'urne précédente qui contient 3 boules blanches et 7 noires constitue une épreuve de Bernoulli. Si, on convient d'appeler succès le tirage d'une boule blanche, il s'agit d'une épreuve de Bernoulli de paramètre $p = \frac{3}{10}$. La variable aléatoire de Bernoulli X de paramètre p qui prend la valeur 1 en cas de succès et 0 en cas d'échec a une espérance $E(X) = p = \frac{3}{10}$ et une variance $V(X) = p(1-p) = \frac{21}{100}$.
Si, au contraire, on convient d'appeler succès le tirage d'une boule noire, il s'agit d'une épreuve de Bernoulli de paramètre $q = \frac{7}{10}$. La variable aléatoire de Bernoulli X' de paramètre q qui prend la valeur 1 en cas de succès et 0 en cas d'échec a une espérance $E(X') = q = \frac{7}{10}$ et une variance $V(X') = q(1-q) = \frac{21}{100}$.

2.b. Schéma de Bernoulli

Définition 8.9 Soit une épreuve de Bernoulli de paramètre p et soit $n \geqslant 1$ un entier.
La répétition de n fois cette même épreuve de Bernoulli, de façon indépendante, constitue une expérience aléatoire appelée « schéma de Bernoulli » de paramètres n et p.

Remarque : L'exemple précédent, où on tire successivement et avec remise, trois boules d'une urne qui en contient 3 blanches et 7 noires est un schéma de Bernoulli, du fait que les tirages sont indépendants les uns des autres. Si on considère comme « succès » le tirage d'une boule blanche, il s'agit d'un schéma de Bernoulli de paramètres $n = 3$ et $p = \frac{3}{10}$.

Dans un schéma de Bernoulli, si on s'intéresse à l'évènement E_k : « obtenir k succès », celui-ci est réalisé par différents chemins qui ont tous la même probabilité : $p^k \times (1-p)^{n-k}$.
Il suffit de dénombrer ces chemins pour déterminer la probabilité de l'évènement E_k.

Définition 8.10 Dans le cas d'une répétition de n épreuves de Bernoulli identiques et indépendantes, on note $\binom{n}{k}$ ou C_n^k le nombre de chemins aboutissant à k succès exactement $(0 \leqslant k \leqslant n)$.
L'entier $\binom{n}{k}$ est appelé « coefficient binomial » et lu « k parmi n ».

Dans le cas d'un schéma de Bernoulli à 3 épreuves comme dans l'exemple du tirage de trois boules avec remise, les coefficients sont $\binom{3}{0} = 1$, $\binom{3}{1} = 3$, $\binom{3}{2} = 3$ et $\binom{3}{3} = 1$.
Ces coefficients sont ceux du développement $(a+b)^3 = a^3 + 3a^2b + 3ab^2 + b^3$.
Les propriétés et usages de ces « coefficients binomiaux » font l'objet d'une partie du chapitre 1.

Définition 8.11 (loi binomiale) Lors d'un schéma de Bernoulli de paramètres n et p, la variable aléatoire X comptant le nombre de succès suit une loi de probabilité appelée *loi binomiale* de paramètres n et p, notée $\mathscr{B}(n,p)$.

Propriété 8.10 Si X est une variable aléatoire suivant une loi binomiale $\mathscr{B}(n,p)$ alors on a :

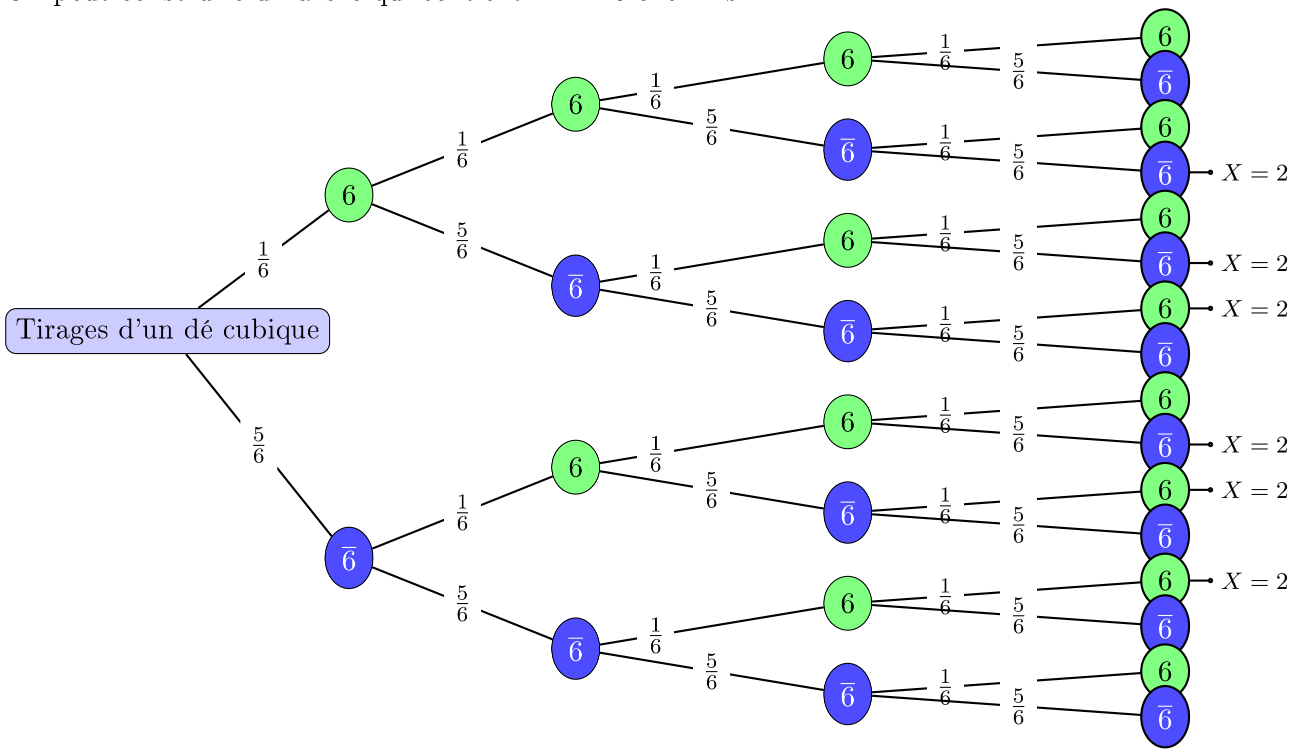
- La variable X prend ses valeurs dans $X(\Omega) = \{0, 1, \dots, n\} = [\![0, n]\!]$
- $\forall k \in \{0, 1, 2, ..., n\}, \quad P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$

Démonstration $X(\Omega) = [\![0, n]\!]$ par définition de la variable X.

$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$ par définition des coefficients binomiaux qui dénombrent les chemins qui contiennent k succès et qui, du fait de la propriété 8.4, ont la probabilité $p^k(1-p)^{n-k}$ d'advenir.

Exemple 79 – On lance quatre fois de suite un dé équilibré à six faces. On appelle succès l'obtention d'un 6 et on considère la variable aléatoire qui comptabilise le nombre de succès.
Cette loi suit une loi binomiale $\mathscr{B}(4, \frac{1}{6})$ car le dé est équilibré et n'a pas de mémoire.

On peut construire un arbre qui contient $2^4 = 16$ chemins.



1 seul chemin mène à 0 ou 4 succès ; 4 mènent à 1 ou 3 succès. Il en reste donc $16 - (1 + 1 + 4 + 4) = 6$ qui réalisent 2 succès (j'ai marqué ces six chemins sur l'illustration). Les coefficients binomiaux pour un schéma de Bernoulli de paramètre $n = 4$ sont $\binom{4}{0} = 1$, $\binom{4}{1} = 4$, $\binom{4}{2} = 6$, $\binom{4}{3} = 4$ et $\binom{4}{4} = 1$.
La loi de probabilité de la variable X comptabilisant les succès est donnée par la formule suivante :
$P(X = k) = \binom{4}{k} \times \left(\frac{1}{6}\right)^k \times \left(\frac{5}{6}\right)^{4-k}$
Donnons explicitement les résultats pour toutes les valeurs $k \in X(\Omega)$.

k	0	1	2	3	4
$P(X = k)$	$= \left(\frac{5}{6}\right)^4$ $\approx 0,4823$	$= 4 \times \left(\frac{1}{6}\right) \times \left(\frac{5}{6}\right)^3$ $\approx 0,3858$	$= 6 \times \left(\frac{1}{6}\right)^2 \times \left(\frac{5}{6}\right)^2$ $\approx 0,1157$	$= 4 \times \left(\frac{1}{6}\right)^3 \times \left(\frac{5}{6}\right)$ $\approx 0,0154$	$= \left(\frac{1}{6}\right)^4$ $\approx 0,0008$

Calculons l'espérance et la variance de X :
$E(X) = 0 \times \left(\frac{5}{6}\right)^4 + 1 \times 4\left(\frac{1}{6}\right)\left(\frac{5}{6}\right)^3 + 2 \times 6\left(\frac{1}{6}\right)^2\left(\frac{5}{6}\right)^2 + 3 \times 4\left(\frac{1}{6}\right)^3\left(\frac{5}{6}\right) + 4 \times \left(\frac{1}{6}\right)^4 = \frac{2}{3}$
$V(X) = \left[0 \times \left(\frac{5}{6}\right)^4 + 1 \times 4\left(\frac{1}{6}\right)\left(\frac{5}{6}\right)^3 + 4 \times 6\left(\frac{1}{6}\right)^2\left(\frac{5}{6}\right)^2 + 9 \times 4\left(\frac{1}{6}\right)^3\left(\frac{5}{6}\right) + 16 \times \left(\frac{1}{6}\right)^4\right] - \left[\frac{2}{3}\right]^2 = \frac{5}{9}$

Propriété 8.11 (Paramètres de la loi binomiale)
Si X est une variable aléatoire $\mathscr{B}(n,p)$ alors :
$E(X) = np \quad ; \quad V(X) = np(1-p) \quad ; \quad \sigma(X) = \sqrt{V(X)} = \sqrt{np(1-p)}$

Démonstration Cette propriété découle de la linéarité additive de l'espérance (propriété 8.6) :
Si Y_1 et Y_2 sont deux variables définies sur une même expérience alors $E(Y_1 + Y_2) = E(Y_1) + E(Y_2)$.
Cette propriété s'étend naturellement à n variables de Bernoulli de même loi et d'espérances $E(Y_1) = E(Y_2) = \cdots = E(Y_n) = p$.
La variable $X = Y_1 + Y_2 + \cdots + Y_n$ compte les succès d'un schéma de Bernoulli à n épreuves.
Appliquons à la somme de ces n variables de Bernoulli la linéarité additive :
$E(X) = E(Y_1 + Y_2 + \cdots + Y_n) = E(Y_1) + E(Y_2) + \cdots + E(Y_n) = np$.

La variance a également cette propriété de linéarité additive lorsque les épreuves sont indépendantes.
Pour deux variables indépendantes de même loi on a $V(Y_1 + Y_2) = V(Y_1) + V(Y_2)$.
Cette propriété s'étend aux n variables indépendantes d'un schéma de Bernoulli de variances $V(Y_1) = V(Y_2) = \cdots = V(Y_n) = p(1-p)$.
On en déduit $V(X) = V(Y_1 + Y_2 + \cdots + Y_n) = V(Y_1) + V(Y_2) + \cdots + V(Y_n) = np(1-p)$.

Remarque :
Il est possible de démontrer directement ces résultats, à l'aide d'une propriété des coefficients binomiaux qui constituera notre lemme : $\forall n \in \mathbb{N},\ \forall k \in \mathbb{N},\ 1 \leqslant k \leqslant n \Longrightarrow k\binom{n}{k} = n\binom{n-1}{k-1}$.
En effet, $\binom{n}{k} = k \times \frac{n!}{k! \times (n-k)!} = \frac{n(n-1)!}{(k-1)! \times (n-1+1-k)!} = \frac{n(n-1)!}{(k-1)! \times ((n-1)-(k-1)!} = n\binom{n-1}{k-1}$.

$\boxed{\text{Calcul de } E(X)}$ en utilisant le lemme et en notant, à partir d'ici, q à la place de $1-p$:
$E(X) = \sum\limits_{k=0}^{n} k\binom{n}{k}p^k q^{n-k} = \sum\limits_{k=1}^{n} k\binom{n}{k}p^k q^{n-k} = \sum\limits_{k=1}^{n} n\binom{n-1}{k-1}p^k q^{n-k}$ (terme nul pour $k=0$).
Posons $k' = k-1$ et mettons n puis p en facteur :
$E(X) = n\sum\limits_{k'=0}^{n-1} \binom{n-1}{k'}p^{k'+1}q^{n-(k'+1)} = np\sum\limits_{k'=0}^{n-1} \binom{n-1}{k'}p^{k'}q^{n-(k'+1)} = np\sum\limits_{k'=0}^{n-1} \binom{n-1}{k'}p^{k'}q^{n-1-k'}$.
On remarque alors que $\sum\limits_{k'=0}^{n-1} \binom{n-1}{k'}p^{k}q^{n-1-k}$ est l'écriture du développement du binôme $(p+q)^{n-1}$ et comme $p+q=1$ cette somme vaut 1 et donc $E(X) = np$.

$\boxed{\text{Calcul de } V(X)}$: Pour pouvoir utiliser deux fois successivement le lemme, il est astucieux d'écrire $E(X^2)$ sous la forme $E(X(X-1)) + E(X)$. Cette transformation d'écriture est possible d'après cette propriété de l'espérance : si φ et ψ sont deux fonctions réelles et X une variable aléatoire alors $E(\varphi(X) + \psi(X)) = E(\varphi(X)) + E(\psi(X))$. En effet, en supposant que $\mathrm{card}(\Omega) = n$, on a

$$E(\varphi(X)+\psi(X)) = E((\varphi+\psi)(X)) = \sum_{k=0}^{n}(\varphi+\psi)(x_k)P(X=x_k) = \underbrace{\sum_{k=0}^{n}\varphi(x_k)P(X=x_k)}_{E(\varphi(X))} + \underbrace{\sum_{k=0}^{n}\psi(x_k)P(X=x_k)}_{E(\psi(X))}$$

En choisissant $\varphi : x \longmapsto x(x-1)$ et $\psi : x \longmapsto x$, on a bien $(\varphi+\psi)(x_k) = x(x-1) + x = x^2$.
Revenons à notre démonstration :
$E(X(X-1)) = \sum\limits_{k=0}^{n} k(k-1)\binom{n}{k}p^k q^{n-k} = \sum\limits_{k=2}^{n} k(k-1)\binom{n}{k}p^k q^{n-k}$ (termes nuls pour $k=0$ et 1)
Appliquons deux fois le lemme : $k(k-1)\binom{n}{k} = n(k-1)\binom{n-1}{k-1} = n(n-1)\binom{n-2}{k-2}$
Par conséquent, en mettant $n(n-1)$ en facteur :
$E(X(X-1)) = n(n-1)\sum\limits_{k=2}^{n} \binom{n-2}{k-2}p^k q^{n-k}$
Posons $k' = k-2$ et mettons p^2 en facteur :
$E(X(X-1)) = n(n-1)p^2\sum\limits_{k'=0}^{n-2} \binom{n-2}{k'}p^{k'}q^{n-(k'+2)} = n(n-1)p^2\sum\limits_{k'=0}^{n-2} \binom{n-2}{k'}p^{k'}q^{(n-2)-k'}$
On remarque alors que $\sum\limits_{k'=0}^{n-2} \binom{n-2}{k'}p^{k'}q^{(n-2)-k'}$ est l'écriture du développement du binôme $(p+q)^{n-2}$ qui vaut 1 et donc $E(X(X-1)) = n(n-1)p^2$.
Comme $V(X) = E(X^2) - (E(X))^2 = E(X(X-1)) + E(X) - (E(X))^2$, on en déduit
$V(X) = n(n-1)p^2 + np - n^2p^2 = np(1-p)$

Utilisation de la calculatrice :

	TI	Casio	Numworks
$\binom{n}{k}$	touche math option PRB choisir 3 (combinaison)	menu RUN option PROB(F3) choisir nCr	menu Calculs, Boîte à outils (5e touche), Probabilités puis Dénombrement, $\binom{n}{k}$, entrer n et k
$P(X=k)$	touche 2nde puis var dans DISTRIB choisir binomFdp(, compléter avec les valeurs de n, p, k et)	menu STAT options DIST(F5) puis BINM(F5). Compléter l'option Bpd(F1) « probability distribution » avec les valeurs de k, n, p	menu Probabilités, choisir Binomiale, entrer n,p, puis Suivant, sélectionner la 4e icône
$P(X \leqslant k)$	dans l'onglet DISTRIB, choisir binomFRép(compléter avec les valeurs de n, p, k	Compléter l'option Bcd(F2) « cumulative distribution » avec les valeurs de k, n, p	menu Probabilités, choisir Binomiale, entrer n,p, puis Suivant, sélectionner la 1re icône

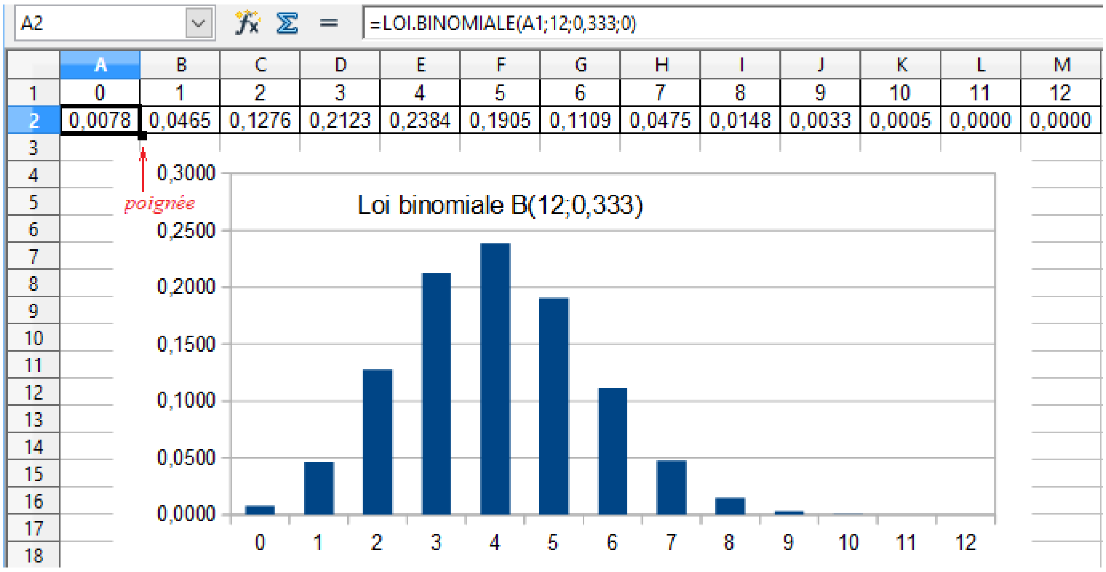
La calculatrice permet de déterminer les coefficients binomiaux $\binom{n}{k}$, les valeurs de $P(X=k)$ et aussi $P(X \leqslant k)$ pour une variable aléatoire X qui suit une loi binomiale $\mathscr{B}(n,p)$, avec $n \in \mathbb{N}^*$, $p \in [0;1]$ et k entier tel que $0 \leqslant k \leqslant n$. En se mettant dans le mode « Table », on peut également obtenir un tableau de ces différentes valeurs : on remplace k par X, et on écrit Y1=binomFdp(n,p,X) sur TI ou Bpd(X, n, p) sur Casio. Sur Numworks, le tableau n'est pas disponible, mais les différentes valeurs sont représentées graphiquement par un diagramme en bâton.

Utilisation du tableur : Avec un tableur (*calc* d'OpenOffice), on a accès aux valeurs utiles :

- les coefficients binomiaux $\binom{n}{k}$: taper la formule « =COMBIN(n ;k) »
- les valeurs de $P(X = k)$: taper la formule « =LOI.BINOMIALE(k ;n ;p ;0) »
- les valeurs de $P(X \leqslant k)$: taper la formule « =LOI.BINOMIALE(k ;n ;p ;1) »

Remarques : avec le tableur *Excel* de Microsoft, il suffit de remplacer les codes 0 et 1 par `False` et `True`. En faisant varier k entre 0 et n dans une feuille de calculs, on obtient simplement le tableau des $P(X = k)$ d'une loi binomiale. Ce tableau fourni à l'assistant graphique conduit à la représentation de cette loi par un diagramme en bâtons.

Pour représenter la loi binomiale $\mathscr{B}(12; 0,333)$ à l'aide du tableur : commencer par entrer les valeurs de k sur la ligne 1 (de 0 à 12), puis entrer la formule en A2 : « =LOI.BINOMIALE(A1 ;12 ;0,333 ;0) ». Recopier ensuite cette formule en tirant la poignée vers la droite (voir l'illustration).
Pour obtenir le diagramme, sélectionner la plage des données (de la cellule A1 à la cellule M2, plage notée A1 :M2), puis lancer l'assistant graphique (icône Diagramme) où on sélectionne le type de diagramme « Colonne ». Déclarer les données « en ligne » et la « Première ligne comme étiquette ».



	A	B	C	D	E	F	G	H	I	J	K	L	M
1	0	1	2	3	4	5	6	7	8	9	10	11	12
2	0,0078	0,0465	0,1276	0,2123	0,2384	0,1905	0,1109	0,0475	0,0148	0,0033	0,0005	0,0000	0,0000

2.c. Prise de décision

Propriété 8.12 (Intervalle de fluctuation) On considère une population sur laquelle on étudie la proportion p d'individus ayant un caractère C avec $p \in [0,2; 0,8]$ et dans laquelle on prélève un échantillon de taille $n \geqslant 25$.
Dans ces conditions, la fréquence f des individus de l'échantillon ayant le caractère C fluctue, dans 95% des cas, à l'intérieur de l'intervalle $\left[p - \frac{1}{\sqrt{n}} ; p + \frac{1}{\sqrt{n}}\right]$, dit *intervalle de fluctuation*.

Remarques :
Cela signifie qu'en choisissant un échantillon de taille n au hasard, la probabilité de l'évènement « $f \in \left[p - \frac{1}{\sqrt{n}} ; p + \frac{1}{\sqrt{n}}\right]$ » est supérieure ou égale à $0,95$ sous réserve de satisfaire les conditions sur n et p. Reformulons cette situation dans le contexte de la loi binomiale : le tirage au hasard d'un individu dans une population qui peut présenter un caractère C avec une probabilité p est assimilable à une épreuve de Bernoulli de paramètre p où le succès correspond à « avoir le caractère C ». Le prélèvement d'un échantillon de taille n dans cette population s'assimile alors à un schéma de Bernoulli de paramètres n et p et la variable aléatoire X qui compte le nombre de succès suit la loi binomiale $\mathscr{B}(n, p)$. La variable aléatoire $F = \frac{X}{n}$ correspond à la fréquence du succès.

DÉFINITION 8.12 Soit X une variable qui suit une loi binomiale $\mathscr{B}(n,p)$ et $F = \frac{X}{n}$ la variable aléatoire qui représente la fréquence du succès. Un intervalle de fluctuation de F au seuil de 95% est un intervalle de la forme $\left[\frac{a}{n};\frac{b}{n}\right]$ où a et b sont des entiers compris entre 0 et n, tels que

$$P(\frac{a}{n} \leqslant F \leqslant \frac{b}{n}) \geqslant 0,95$$

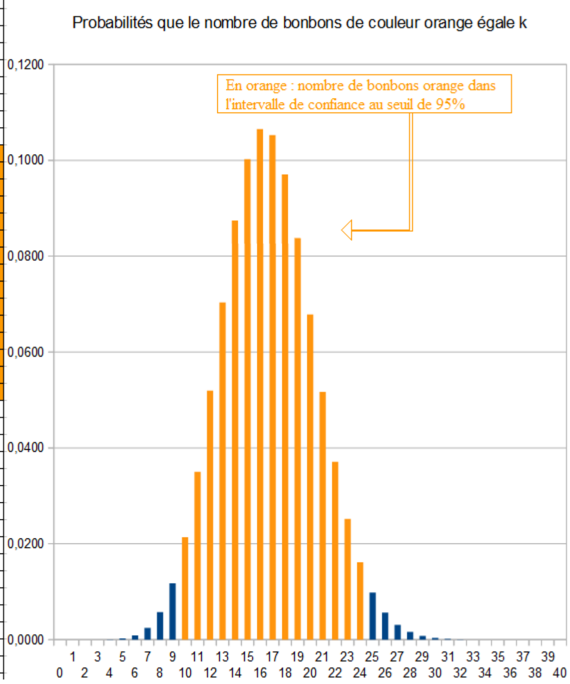
Dans la pratique :
La condition $P(\frac{a}{n} \leqslant F \leqslant \frac{b}{n}) \geqslant 0,95$ équivaut à $P(a \leqslant X \leqslant b) \geqslant 0,95$. Afin d'obtenir l'intervalle de fluctuation de plus petite amplitude, on cherche les plus petits entiers a et b tels que :

$$P\left(X \leqslant a\right) \geqslant 0,025 \text{ et } P\left(X \leqslant b\right) \geqslant 0,975$$

Ces conditions satisfaites, on aura bien $P(a \leqslant X \leqslant b) \geqslant 0,95$ car
$P\left(X \leqslant b\right) \geqslant 0,975 \Longrightarrow P\left(X > b\right) < 0,025$ et $P\left(X \leqslant a\right) \geqslant 0,025 \Longrightarrow P\left(X < a\right) < 0,025$.
Comme $P\left((X < a) \cup (X > b)\right) = P\left(X < a\right) + P\left(X > b\right) < 0,025 + 0,025 = 0,05$,
on a $P(a \leqslant X \leqslant b) = P\left(\overline{(X < a) \cup (X > b)}\right) = 1 - P\left((X < a) \cup (X > b)\right) \geqslant 0,95$.

EXEMPLE 80 – On constitue des sachets de 100 bonbons `M&M's` à partir d'un mélange contenant des bonbons des 6 couleurs en proportion égale ($f = \frac{1}{6}$ pour chaque couleur). Á l'aide du tableur, on veut déterminer un intervalle de fluctuation au seuil de 95% de la fréquence d'une couleur, disons orange, dans un sachet. Le nombre X de bonbons de couleur orange suit la loi binomiale $\mathscr{B}(100;\frac{1}{6})$. La fréquence des bonbons de couleur orange est donné par la variable aléatoire $F = \frac{X}{100}$. On cherche deux entiers a et b tels que $P(\frac{a}{100} \leqslant F \leqslant \frac{b}{100}) \geqslant 0,95$ autrement dit tels que $P(a \leqslant X \leqslant b) \geqslant 0,95$. Ci-dessous une copie d'écran du tableur avec les valeurs prise par la variable X, les probabilités $P(X = k)$ et les probabilités cumulées $P(X \leqslant k)$.

k	P(X=k)	P(X<=k)
0	0,0000	0,0000
1	0,0000	0,0000
2	0,0000	0,0000
3	0,0000	0,0000
4	0,0001	0,0001
5	0,0003	0,0004
6	0,0009	0,0013
7	0,0025	0,0038
8	0,0058	0,0095
9	0,0118	0,0213
10	0,0214	0,0427
11	0,0350	0,0777
12	0,0520	0,1297
13	0,0703	0,2000
14	0,0874	0,2874
15	0,1002	0,3877
16	0,1065	0,4942
17	0,1052	0,5994
18	0,0971	0,6965
19	0,0838	0,7803
20	0,0679	0,8481
21	0,0517	0,8998
22	0,0371	0,9369
23	0,0252	0,9621
24	0,0162	0,9783
25	0,0098	0,9881
26	0,0057	0,9938
27	0,0031	0,9969
28	0,0016	0,9985
29	0,0008	0,9993
30	0,0004	0,9997
31	0,0002	0,9999
32	0,0001	1,0000
33	0,0000	1,0000
34	0,0000	1,0000
35	0,0000	1,0000
36	0,0000	1,0000
37	0,0000	1,0000
38	0,0000	1,0000
39	0,0000	1,0000
40	0,0000	1,0000



La partie colorée en orange montre que :

- $P(X \leqslant 10) \geqslant 0,025$ et $a = 10$ est la 1$^{\text{re}}$ valeur de k où $P(X \leqslant k) \geqslant 0,025$ se produit
- $P(X \leqslant 24) \geqslant 0,975$ et $b = 24$ est la 1$^{\text{re}}$ valeur de k où $P(X \leqslant k) \geqslant 0,975$ se produit

Donc $P(10 \leqslant X \leqslant 24) \geqslant 0,95$ ou encore $P(\frac{10}{100} \leqslant F \leqslant \frac{24}{100}) \geqslant 0,95$ et donc $\left[\frac{10}{100};\frac{24}{100}\right]$ est un intervalle de fluctuation au seuil de 95% de la fréquence de la couleur orange dans les sachets de 100 `M&M's`. Le risque de voir la fréquence de la couleur orange sortir de cet intervalle est inférieur à 5%.

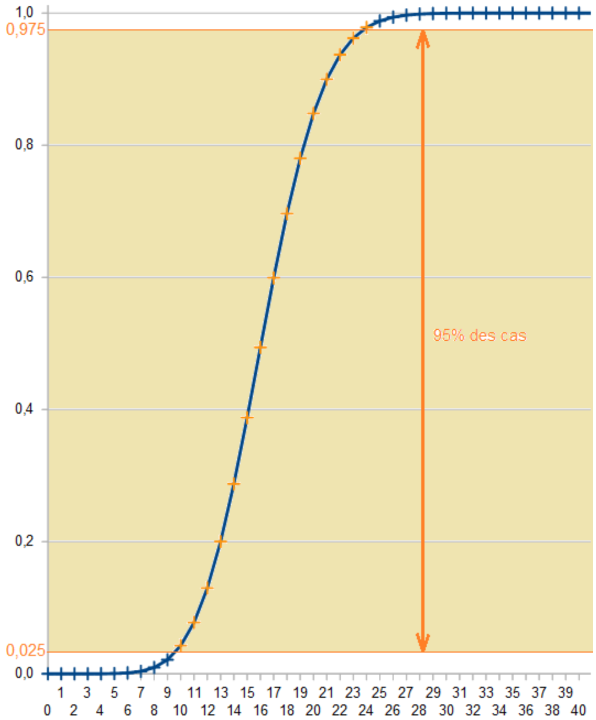
Un échantillon de 100 bonbons contenant moins de 10 bonbons de couleur orange ou plus de 24 bonbons de couleur orange ne serait pas représentatif de cette distribution au seuil de 95%.

Comparons ce résultat à l'intervalle de fluctuation prévu par la propriété 8.12 :
$IF_{prop} = \left[\frac{1}{6} - \frac{1}{\sqrt{100}}; \frac{1}{6} + \frac{1}{\sqrt{100}}\right] = \left[\frac{1}{15}; \frac{4}{15}\right] \approx [0,0667; 0,2667]$. L'intervalle de fluctuation donné par la loi binomiale est $IF_{binom} = [0,1; 0,24]$. On constate que IF_{binom} a une amplitude plus faible que IF_{prop} et que $IF_{binom} \subset IF_{prop}$: les fluctuations prévues par la loi binomiale sont moins dispersées que celles que donne la propriété 8.12.

La représentation cumulative des probabilités associées à la loi binomiale $\mathscr{B}(100; \frac{1}{6})$ est donnée ci-contre. Il s'agit de la représentation de la fonction dont les valeurs sont fournies par le tableau ci-dessus :

$$k \mapsto P(X \leqslant k)$$

Les droites d'équation $y = 0,025$ et $y = 0,975$ correspondant aux valeurs de coupure $2,5\%$ et $97,5\%$, délimitent une zone (colorée sur le graphique) qui contient exactement 95% de la population. Du fait que les nombres a et b sont des entiers, les valeurs de coupure retenues sont légèrement décalées et le pourcentage de la population contenu entre ces valeurs dépasse toujours légèrement 95%. Ici, la somme des probabilités de $P(X = 10)$ à $P(X = 24)$ est environ égale à $95,7\%$.



Règle de décision :
Dans une population, on ne connaît pas toujours avec exactitude la proportion de la population ayant un caractère C mais on peut faire une hypothèse sur cette proportion, c'est-à-dire se fixer une valeur p de cette proportion, et soumettre cette hypothèse au test d'échantillonnage. Notons f la fréquence observée dans un échantillon de taille n de cette population. Comme on sait que 95% des échantillons doivent avoir une fréquence comprise dans l'intervalle de fluctuation IF déterminé à partir de n et p, on peut établir la règle de décision suivante :

> Si la fréquence f de l'échantillon n'appartient pas à IF, on rejette l'hypothèse faite sur p avec un risque d'erreur de 5%. Dans le cas contraire, on affirme que l'hypothèse est crédible car compatible avec les données d'échantillonnage.

Application : cette règle examine si la variabilité d'un échantillonnage peut s'expliquer par le simple fait du hasard. Lorsque la fréquence f mesurée dans l'échantillon sort de l'intervalle de fluctuation, il faut chercher une autre explication que le hasard pour justifier l'écart entre la fréquence observée et la probabilité hypothétique p attribuée au caractère C.
Si on trouve qu'un sachet de 100 bonbons `M&M's` ne contient que 7 bonbons oranges ou 25 bonbons rouges, d'après ce qui a été vu plus haut, il faut rejeter l'hypothèse que les 6 couleurs sont équiprobables dans ce genre de sachet. La simple fluctuation due au hasard ne permet pas d'expliquer ces valeurs. En affirmant cela, on assume un risque d'erreur de 5%.
Avec les mêmes données numériques, pour déterminer si un dé cubique est bien équilibré on peut faire l'expérience de le tirer 100 fois et de noter le nombre de fois où l'on a obtenu chacune des faces. Selon notre étude, si une quelconque des 6 faces est sortie plus de 24 fois ou moins de 10 fois, alors on trouver ce résultat « anormal » au seuil de 95% et rejeter l'hypothèse d'équiprobabilité des faces.

2.d. Simulations

Une expérience aléatoire amène un « succès » avec une probabilité égale à p, nous voulons simuler avec un programme la constitution de N échantillons de taille n pour chacun desquels on déterminera la valeur k prise par la variable aléatoire X qui compte le nombre de succès. Le programme doit pouvoir estimer les probabilités $P(X = k)$ et $P(X \leqslant k)$ ainsi que les paramètres de la loi de X.

Programme P0 : échantillon

Le programme génère un échantillon de n individus dont certains sont des succès (notés S) et les autres des échecs (notés E). On affiche ici l'échantillon ainsi que la valeur prise par la variable aléatoire X comptant le nombre de succès.

```
from random import random

def echantillon(n,p):
    k=0
    for i in range(n):
        if random()<p:
            k+=1
            print("S", end=" ")
        else : print("E", end=" ")
    return k

n,p=16,1/6
print("Génération d'un échantillon de {} individus :".format(n))
print(" -> Nombre de succès={}".format(echantillon(n,p)))
```

```
Génération d'un échantillon de 16 individus :
E E E E E E E E E E S E S E S S  -> Nombre de succès=4
```

Programme P1 : loi de probabilité

Une valeur particulière de $K \in [\![0, n]\!]$) étant donnée, le programme comptabilise :

- le nombre C_1 d'échantillons dans lequel $X = K$
- le nombre C_2 d'échantillons dans lequel $X \leqslant K$

Le programme affiche alors les valeurs empiriques de $P(X = K) = \frac{C_1}{N}$ et $P(X \leqslant K) = \frac{C_2}{N}$.

Programme P2 : paramètres de la loi

Le programme détermine les paramètres (espérance, variance, écart-type) de la variable aléatoire X :

- l'espérance est calculée comme la moyenne des valeurs de k
- la variance est calculée avec la formule de König-Huygens comme la différence entre la moyenne des k^2 et le carré de la moyenne des k

Je reprends les valeurs numériques de l'exemple 80 afin de pouvoir comparer les valeurs empiriques obtenues par la simulation et les valeurs théoriques obtenues par les formules.

P1

```
from random import random

def echantillon(n,p):
    k=0
    for i in range(n):
        if random()<p:k+=1
    return k

def loi(N,n,p,K):
    C1,C2=0,0
    for i in range(N):
        X=echantillon(n,p)
        if X==K:C1+=1
        if X<=K:C2+=1
    return C1/N,C2/N

N,n,p,K=100000,100,1/6,15
f1,f2=loi(N,n,p,K)
print("P(X={})={}".format(K,f1))
print("P(X<={})={}".format(K,f2))
```

```
P(X=5)=0.00021     P(X=10)=0.02224
P(X<=5)=0.00032    P(X<=10)=0.04388
```

P2

```
from random import random
from math import sqrt
def echantillon(n,p):
    k=0
    for i in range(n):
        if random()<p:k+=1
    return k

def parametre(N,n,p):
    X1,X2=0,0
    for i in range(N):
        X=echantillon(n,p)
        X1+=X
        X2+=X**2
    return X1/N,X2/N-(X1/N)**2

N,n,p=100000,100,1/6
E,V=parametre(N,n,p)
E,V,S=round(E,3),round(V,3),round(sqrt(V),3)
print("E(X)={},V(X)={},sigma(X)={}".format(E,V,S))
```

```
E(X)=16.68,V(X)=13.817,sigma(X)=3.717
```

L'intérêt de ce programme est de donner une estimation empirique des caractéristiques de cette loi de probabilité. On a obtenu ici, en simulant 100 000 échantillons, que l'espérance et l'écart-type valent approximativement $E(X) = 16,68$ et $\sigma(X) = 3,717$.

Cette loi de probabilité est une loi Binomiale $\mathscr{B}(100; \frac{1}{6})$) dont les paramètres sont $E(X) = \frac{100}{6} \approx 16,67$ et $\sigma(X) = \sqrt{\frac{500}{36}} \approx 3,727$.

2.e. Autres lois

Il existe de nombreuses autres lois de probabilité en dehors de la loi binomiale. En voici deux :

2.e.1. Loi géométrique tronquée

DÉFINITION 8.13 Soit $n \neq 0$ un entier et p la probabilité de succès d'une épreuve de Bernoulli. On appelle « loi géométrique tronquée » de paramètres n et p, la loi de probabilité d'une variable X donnant le nombre d'essais nécessaires pour obtenir le premier succès, étant entendu que le nombre d'essais ne peut pas dépasser la valeur n.

Remarque : Il n'est pas difficile de déterminer $P(X=k)$ lorsque $1 \leqslant k \leqslant n$. Comme il doit y avoir $k-1$ échecs (de probabilité $1-p$) suivis d'un succès, on a $P(X=k)=p(1-p)^{k-1}$. Lorsqu'il n'y a pas de succès en n essais, la variable aléatoire prend la valeur 0 ; ainsi, on a $P(X=0)=(1-p)^n$.
Montrons que la somme des probabilités de cette loi vaut bien 1, en notant $q=1-p$:

$$\sum_{k=1}^{n} pq^{k-1}+q^n = p\sum_{k=1}^{n} q^{k-1}+q^n = p\times 1\times\frac{1-q^n}{1-q}+q^n = p\times\frac{1-q^n}{p}+q^n = 1-q^n+q^n = 1$$

Programme P3 : loi et paramètres

Ce programme utilise une liste L afin de donner les valeurs empiriques de $P(X=k)$ pour $0 \leqslant k \leqslant n$.
La simulation détermine un rang k du premier succès ce qui, à chaque fois, incrémente de 1 le compteur $L(k)$. Lorsque les N valeurs ont été ainsi enregistrées, il n'y a plus qu'à calculer les fréquences $\frac{L(k)}{N}$ correspondant aux probabilités $P(X=k)$.
Pour le calcul des paramètres, on procède à peu près comme pour la loi binomiale sauf qu'il faut additionner les produits $k\times L(k)$ pour déterminer $E(X)$ et $k^2\times L(k)$ pour déterminer $E(X^2)$.

```python
from random import random
from math import sqrt

def echantillon(n,p):
    k=0
    for i in range(1,n+1):
        if random()<p:
            k=i
            break
    return k

def loi(N,n,p):
    L=(n+1)*[0]
    for i in range(N):
        L[echantillon(n,p)]+=1
    return L
```

```python
N,n,p=100000,3,1/4
X1,X2=0,0
Liste=loi(N,n,p)
for i in range(n+1):
    P=round(Liste[i]/N,3)
    print("P(X={})={}".format(i,P))
    X1+=Liste[i]*i
    X2+=Liste[i]*i**2
E,V=round(X1/N,3),round(X2/N-(X1/N)**2,3)
S=round(sqrt(V),3)
print("E(X)={},V(X)={},sigma(X)={}".format(E,V,S))
```

```
P(X=0)=0.424
P(X=1)=0.248
P(X=2)=0.188
P(X=3)=0.14
E(X)=1.043,V(X)=1.17,sigma(X)=1.082
```

EXEMPLE 81 – Un client cherche à joindre par téléphone un service de dépannage. La probabilité que son appel soit pris sans attente est $p=\frac{1}{4}$. Si son appel n'est pas pris sans attente, ce client raccroche et rappelle le service. Il fait au maximum trois tentatives ($n=3$).
On note X la variable aléatoire donnant le rang de la tentative aboutissant sans attente. Si les trois tentatives ont échoué, on convient que $X=0$.
La loi de probabilité de X est une loi géométrique tronquée de paramètres n et p. D'après les formules établies ci-dessus dans la remarque, les probabilités sont calculées dans le tableau suivant.

k	0	1	2	3
$P(X=k)$	$0,75^3$ $=0,421875$	$0,25$ $=0,25$	$0,25\times 0,75$ $=0,1875$	$0,25\times 0,75^2$ $=0,140625$

Calculons les paramètres de cette loi :
$E(X)=0\times 0,421875+1\times 0,25+2\times 0,1875+3\times 0,140625=1,046875$
$V(X)=E(X^2)-(E(X))^2=1\times 0,25+4\times 0,1875+9\times 0,140625-1,046875^2\approx 1,169678$
$\sigma(X)=\sqrt{V(X)}\approx\sqrt{1,169678}\approx 1,081516$
Ces valeurs théoriques sont approchées par les valeurs empiriques d'autant mieux que l'on prend un N grand. Avec $N=100000$, on constate que le résultat d'exécution confirme cela.

2.e.2. Loi hypergéométrique

DÉFINITION 8.14 Soient $M \neq 0$ un entier, m et n deux entiers inférieurs à M.
On appelle « loi hypergéométrique » de paramètres M, n et $\frac{m}{M}$, la loi de probabilité d'une variable X donnant le nombre d'individus possédant un certain caractère, dans un tirage simultané de n individus d'une population de taille M où il y a m individus possédant ce caractère.

Remarque : On peut déterminer $P(X = k)$ lorsque $0 \leqslant k \leqslant n$ en effectuant deux dénombrements, dans le cadre d'une situation d'équiprobabilité :

- le nombre de cas possibles est le nombre de choix de n éléments parmi M, soit $\binom{M}{n}$
- le nombre de cas favorables est le produit du nombre de choix de k éléments parmi m ($\binom{m}{k}$) par le nombre de choix de $n-k$ éléments parmi $M-m$ ($\binom{M-m}{n-k}$)

Ainsi, on a $P(X = k) = \frac{\binom{m}{k} \times \binom{M-m}{n-k}}{\binom{M}{n}}$

Programme P4 : loi et paramètres

Ce programme utilise une liste L afin de donner les valeurs empiriques de $P(X = k)$ pour $0 \leqslant k \leqslant n$. La simulation détermine le nombre d'essais qui ont amenés k succès ce qui, à chaque fois, incrémente de 1 le compteur $L(k)$. Le reste est similaire à la loi géométrique tronquée.

```
from random import random
from math import sqrt

def echantillon(M,n,m):
    k=0
    for i in range(n):
        if random()<m/M:
            k+=1
            m-=1
        M-=1
    return k

def loi(N,M,n,m):
    L=(n+1)*[0]
    for i in range(N):
        L[echantillon(M,n,m)]+=1
    return L
```

```
N,M,n,m=100000,32,5,4
X1,X2=0,0
Liste=loi(N,M,n,m)
for i in range(n+1):
    P=round(Liste[i]/N,4)
    print("P(X={})={}".format(i,P))
    X1+=Liste[i]*i
    X2+=Liste[i]*i**2
E,V=round(X1/N,4),round(X2/N-(X1/N)**2,4)
S=round(sqrt(V),4)
print("E(X)={},V(X)={},sigma(X)={}".format(E,V,S))
```

```
P(X=0)=0.5127
P(X=1)=0.3889
P(X=2)=0.0913
P(X=3)=0.007
P(X=4)=0.0001
P(X=5)=0.0
E(X)=0.5931,V(X)=0.4677,sigma(X)=0.6839
```

EXEMPLE 82 – Au poker menteur, célèbre jeu de cartes à deux joueurs, on distribue 5 cartes à l'un deux qui annonce « brelan de rois » (présence de 3 rois exactement sur les 5 cartes tirées).
Quelle est la probabilité que le joueur ne mente pas s'il n'y a pas de jokers dans le jeu de 32 cartes ?
La variable X (le nombre de rois tirés) suit une loi hypergéométrique de paramètres 32, 5 et $\frac{4}{32}$.
D'après la formule établie ci-dessus, la probabilité cherchée est

$$P(X = 3) = \frac{\binom{4}{3} \times \binom{28}{2}}{\binom{32}{5}} = \frac{27}{3596} \approx 0,0075$$

La valeur empirique calculée par notre programme s'approche bien de cette valeur.
Sur cet exemple, on peut vérifier une propriété de cette loi : son espérance est égale à celle de la loi binomiale, soit $E(X) = np$ égal ici à $\frac{nm}{M} = \frac{5\times 4}{32} = \frac{5}{8} = 0,625$.
Quelle est la probabilité que le joueur ne mente pas s'il y a un joker dans le jeu de $32+1$ cartes ?
Dans ce cas, on considère que le joker est un roi. Il s'agit d'une loi hypergéométrique de paramètres 33, 5 et $\frac{5}{33}$ et $P(X = 3) = \frac{\binom{5}{3}\times\binom{28}{2}}{\binom{33}{5}} = \frac{315}{19778} \approx 0,016$, ce que confirme approximativement le programme.

3. Loi des grands nombres

3.a. L'inégalité de Bienaymé-Tchebychev

PROPRIÉTÉ 8.13 (INÉGALITÉ DE MARKOV) Soit X une variable aléatoire finie qui ne prend que des valeurs positives (telle que $X(\Omega) \subset \mathbb{R}^+$) et soit ϵ un réel strictement positif.

$$P(X \geqslant \epsilon) \leqslant \frac{E(X)}{\epsilon}$$

DÉMONSTRATION Par définition de l'espérance $E(X) = \sum\limits_{k \in X(\Omega)} P(X = k) \times k$.
On a $\sum\limits_{k \in X(\Omega)} P(X = k) \times k = \sum\limits_{\substack{k \in X(\Omega) \\ k < \epsilon}} P(X = k) \times k + \sum\limits_{\substack{k \in X(\Omega) \\ k \geqslant \epsilon}} P(X = k) \times k$.
Comme $k \geqslant 0$ et $P(X = k) > 0$, la somme $\sum\limits_{\substack{k \in X(\Omega) \\ k < \epsilon}} P(X = k) \times k$ est positive.
Par conséquent $\sum\limits_{k \in X(\Omega)} P(X = k) \times k \geqslant \sum\limits_{\substack{k \in X(\Omega) \\ k \geqslant \epsilon}} P(X = k) \times k$ et donc $E(X) \geqslant \sum\limits_{\substack{k \in X(\Omega) \\ k \geqslant \epsilon}} P(X = k) \times k$.
Mais $k \geqslant \epsilon \Longrightarrow P(X = k) \times k \geqslant P(X = k) \times \epsilon$.
On en déduit $E(X) \geqslant \sum\limits_{\substack{k \in X(\Omega) \\ k \geqslant \epsilon}} P(X = k) \times \epsilon$, soit $E(X) \geqslant \epsilon \sum\limits_{\substack{k \in X(\Omega) \\ k \geqslant \epsilon}} P(X = k)$.
Par définition $\sum\limits_{\substack{k \in X(\Omega) \\ k \geqslant \epsilon}} P(X = k) = P(X \geqslant k)$.
Finalement $E(X) \geqslant \epsilon P(X \geqslant k)$ d'où, puisque $\epsilon > 0$, la conclusion $P(X \geqslant \epsilon) \leqslant \frac{E(X)}{\epsilon}$

Remarques :

- La double majoration opérée ici est sévère (on néglige toutes les valeurs de $k < \epsilon$ et ensuite on minore tous les $k \geqslant \epsilon$ par ϵ) ce qui peut parfois conduire à un résultat peu intéressant ; néanmoins, on va utiliser ce résultat pour en établir un autre, l'inégalité de Bienaymé-Tchebychev.
- Dans l'exemple 81 où $E(X) = 1,046875$,
  - avec $\epsilon = 2$ on a $P(X \geqslant 2) \leqslant \frac{1,046875}{2} \approx 0,52$ alors que $P(X \geqslant 2) = 0,328125$
  - avec $\epsilon = 3$ on a $P(X \geqslant 3) \leqslant \frac{1,046875}{3} \approx 0,35$ alors que $P(X \geqslant 3) = 0,140625$
  - avec $\epsilon = 4$ on a $P(X \geqslant 4) \leqslant \frac{1,046875}{4} \approx 0,26$ alors que $P(X \geqslant 4) = 0$

PROPRIÉTÉ 8.14 (INÉGALITÉ DE BIENAYMÉ-TCHEBYCHEV) Soit X une variable aléatoire finie et soient ϵ un réel et n un entier, tous deux strictement positifs.

$$P(|X - E(X)| \geqslant \epsilon) \leqslant \frac{V(X)}{\epsilon^2} \text{ et, en particulier, } P(|X - E(X)| \geqslant k\sigma(X)) \leqslant \frac{1}{k^2}$$

DÉMONSTRATION Il suffit d'appliquer l'inégalité de Markov a la variable $(X - E(X))^2$:
$P((X - E(X))^2 \geqslant \epsilon^2) \leqslant \frac{E((X-E(X))^2)}{\epsilon^2}$.
Par définition de la variance $V(X) = E((X - E(X))^2)$.
$P((X - E(X))^2 \geqslant \epsilon^2)$ est la probabilité de l'évènement $(X - E(X))^2 \geqslant \epsilon^2 \Longleftrightarrow |X - E(X)| \geqslant \epsilon$ du fait de la croissance de la fonction racine carrée sur $\mathbb{R}^+$ (et du fait que $|\epsilon| = \epsilon$).
Finalement on a bien $P(|X - E(X)| \geqslant \epsilon) \leqslant \frac{V(X)}{\epsilon^2}$.

Si $\epsilon = k\sigma(X)$ avec $k \in \mathbb{N}^*$, l'inégalité s'écrit $P(|X - E(X)| \geqslant k\sigma(X)) \leqslant \frac{V(X)}{(k\sigma(X))^2}$ soit
$P(|X - E(X)| \geqslant k\sigma(X)) \leqslant \frac{1}{k^2}$ puisque $(\sigma(X))^2 = V(X)$.

Remarques :

- La reformulation de l'inégalité de B-T qui prend comme paramètre ϵ un multiple de l'écart-type de la variable donne un majorant très simple ($\frac{1}{k^2}$) de la probabilité que la variable s'écarte de l'espérance de plus de k fois l'écart-type. Cette constatation lui vaut d'être souvent utilisée pour une prise de décision, nous allons l'appliquer dans le cas d'une loi binomiale.
- Dans l'exemple 81 où $E(X) = 1,046875$ et $V(X) \approx 1,169678$,
  - avec $\epsilon = 1$ on a approximativement $P(|X - E(X)| \geqslant 1) \leqslant 1,17$ alors que $|X - E(X)| \geqslant 1 = (X = 0) \cup (X = 3)$
    En effet $|0 - E(X)| \approx 1,05 > 1$, $|1 - E(X)| \approx 0,05 < 1$, ..., $|2 - E(X)| \approx 0,95 < 1$, $|3 - E(X)| \approx 1,95 > 1$
    et donc $P(|X - E(X)| \geqslant 1) = P(X = 0) + P(X = 3) = 0,5625$
  - avec $\epsilon = 2$ on a $P(|X - E(X)| \geqslant 2) \leqslant \frac{1,169678}{4} \approx 0,29$ alors que $|X - E(X)| \geqslant 2 = \varnothing$
    En effet $|0 - E(X)| \approx 1,05 < 2$, $|1 - E(X)| \approx 0,05 < 2$, ..., $|3 - E(X)| \approx 1,95 < 2$
    et donc $P(|X - E(X)| \geqslant 2) = 0$

  Voyons ce qu'apporte la reformulation avec $\sigma(X) \approx 1,082$ et $2\sigma(X) \approx 2,163$:
  - avec $\epsilon = \sigma(X)$ on a $P(|X - E(X)| \geqslant \sigma(X)) \leqslant 1$ alors que $|X - E(X)| \geqslant \sigma(X) = (X < 3)$
    En effet $|0 - E(X)| \approx 1,05 < \sigma(X)$, ..., $|2 - E(X)| \approx 0,95 < \sigma(X)$, $|3 - E(X)| \approx 1,95 > \sigma(X)$
    et donc $P(|X - E(X)| \geqslant \sigma(X)) = 1 - P(X = 3) \approx 0,86$
  - avec $\epsilon = 2\sigma(X)$ on a $P(|X - E(X)| \geqslant 2\sigma(X)) \leqslant \frac{1}{4}$ alors que $|X - E(X)| \geqslant 2\sigma(X) = \varnothing$
    En effet $|0 - E(X)| \approx 1,05 < 2\sigma(X)$, ..., $|3 - E(X)| \approx 1,95 < 2\sigma(X)$
    et donc $P(|X - E(X)| \geqslant 2) = 0$

3.b. L'inégalité de concentration

Propriété 8.15 (Pour un échantillon) Soit un échantillon $(X_1, X_2, \ldots, X_m)$ de variables d'espérance $E(X_i) = \bar{x}$ et de variance $V(X_i) = \sigma^2$ pour tout $i \in [\![1, m]\!]$.
Pour la variable $M_m = \frac{X_1+X_2+\cdots+X_m}{m}$, quel que soit $\epsilon > 0$, on a $P(|M_m - \bar{x}| \geqslant \epsilon) \leqslant \frac{\sigma^2}{m\epsilon^2}$

Démonstration L'inégalité de Bienaymé-Tchebychev appliquée à la variable M_m s'écrit $P(|M_m - E(M_m)| \geqslant \epsilon) \leqslant \frac{V(M_m)}{\epsilon^2}$
La propriété 8.8 nous indique les paramètres de M_m : $E(M_m) = \bar{x}$ et $V(M_m) = \frac{\sigma^2}{m}$.
Par substitution dans l'inégalité, on obtient $P(|M_m - \bar{x}| \geqslant \epsilon) \leqslant \frac{\sigma^2}{m\epsilon^2}$

Remarques :

- Cette inégalité est parfois appelée *loi des grands nombres* d'après Siméon Denis Poisson qui utilisa cette expression en 1837. Pourquoi parle t-on de grands nombres ? Cette inégalité indique que lorsque la taille de l'échantillon augmente – lorsque m devient un grand nombre – la probabilité que la moyenne des variables s'écarte de plus d'une certaine valeur ϵ devient de plus en plus petite. La limite de $P(|M_m - \bar{x}| \geqslant \epsilon)$ est nulle quand m tend vers l'infini, c'est d'ailleurs plutôt cette dernière affirmation que l'on devrait appeler loi des grands nombres. Cette loi était déjà perçue dans le domaine des jeux de hasard et Jacques Bernoulli en fit une première théorisation relative à la loi binomiale dans son *Ars Conjectandi* de 1737.
- Pourquoi parle t-on de concentration ? Car M_m et $\bar{x}$ sont des paramètres centraux : le premier est la moyenne de l'échantillon et le second l'espérance de toutes les variables impliquées dans cet échantillon. On majore, avec cette inégalité, le risque que l'échantillon ne soit pas assez concentré, c'est-à-dire que sa moyenne s'écarte trop de l'espérance. Celle-ci apparait alors comme la valeur théorique vers laquelle tend la moyenne de l'échantillon.

Propriété 8.16 (Application au schéma de Bernoulli) Soit un échantillon $(X_1, \dots, X_m)$ de variables de Bernoulli de paramètre p. Les variables $S_m = X_1 + \dots + X_m$ et $M_m = \frac{S_m}{m}$ vérifient :

$$\forall \epsilon > 0,\ \ P(|S_m - mp| \geqslant \epsilon) \leqslant \frac{mp(1-p)}{\epsilon^2} \text{ et } P(|M_m - p| \geqslant \epsilon) \leqslant \frac{p(1-p)}{m\epsilon^2}$$

Démonstration La variable $S_m = X_1 + X_2 + \dots + X_m$ compte le nombre de succès car $X_i(\Omega) = \{0, 1\}$ pour tout $i \in [\![1, m]\!]$. Elle suit donc une loi binomiale $\mathscr{B}(m, p)$ dont les caractéristiques sont $E(S_m) = mp$ et $V(S_m) = mp(1-p)$.
En appliquant l'inégalité de Bienaymé-Tchebychev à cette variable on obtient
$P(|S_m - mp| \geqslant \epsilon) \leqslant \frac{mp(1-p)}{\epsilon^2}$

La variable $M_m = \frac{S_m}{m}$ a pour caractéristiques :
$E(\frac{S_m}{m}) = \frac{E(S_m)}{m} = \frac{mp}{m} = p$ et $V(M_m) = V(\frac{S_m}{m}) = \frac{V(S_m)}{m^2} = \frac{mp(1-p)}{m^2} = \frac{p(1-p)}{m}$
En appliquant l'inégalité de Bienaymé-Tchebychev à cette variable on obtient
$P(|M_m - p| \geqslant \epsilon) \leqslant \frac{p(1-p)}{m\epsilon^2}$

Intervalle de fluctuation : Un échantillon $(X_1, \dots, X_m)$ de variables de Bernoulli de paramètre p étant donné, la variable $S_m = X_1 + X_2 + \dots + X_m$ compte le nombre de succès et suit une loi $\mathscr{B}(m, p)$. Dans la partie 8.2.3, on a donné comme intervalle de fluctuation de la variable $M_m = \frac{S_m}{m}$ qui évalue la fréquence des succès au seuil de 95% l'intervalle $\left[\frac{a}{m}; \frac{b}{m}\right]$ où a et b sont des entiers compris entre 0 et m, tels que $P(\frac{a}{m} \leqslant M_m \leqslant \frac{b}{m}) \geqslant 0,95$.

L'évènement $|M_m - p| \geqslant \epsilon$ est le contraire de l'évènement $|M_m - p| < \epsilon$ soit $p - \epsilon < M_m < p + \epsilon$.
Par conséquent $P(p - \epsilon < M_m < p + \epsilon) = 1 - P(|M_m - p| \geqslant \epsilon)$.
La propriété 8.16 implique $1 - P(|M_m - p| \geqslant \epsilon) \geqslant 1 - \frac{p(1-p)}{m\epsilon^2}$.
On en déduit $P(p - \epsilon < M_m < p + \epsilon) \geqslant 1 - \frac{p(1-p)}{m\epsilon^2}$.
Pour comparer avec la règle édictée au seuil de 95% dans la partie 8.2.3, choisissons ϵ de manière à avoir $1 - \frac{p(1-p)}{m\epsilon^2} = 0,95$ soit $0,05 = \frac{p(1-p)}{m\epsilon^2} \iff 0,05m\epsilon^2 = p(1-p) \Longrightarrow \epsilon = \sqrt{\frac{p(1-p)}{0{,}05m}}$.

Appliqué à la situation de l'exemple 80, la loi suivie par S_m est $\mathscr{B}(100, \frac{1}{6})$ on obtient
$\epsilon = \sqrt{\frac{\frac{1}{6} \times \frac{5}{6}}{0{,}05 \times 100}} = \sqrt{\frac{5}{5 \times 36}} = \sqrt{\frac{1}{36}} = \frac{1}{6} \approx 0,167$.
Par conséquent $P(\frac{1}{6} - \frac{1}{6} < M_m < \frac{1}{6} + \frac{1}{6}) \geqslant 0,95$, approximativement $P(0 < M_m < 0,333) \geqslant 0,95$.
L'intervalle de confiance au seuil de 95% qu'on peut déduire de l'inégalité de Bienaymé-Tchebychev est donc $IF = [0; 0,333]$. Cet intervalle est moins précis que ceux déterminés plus haut mais l'obtenir nécessite peu de calculs et on peut étendre la méthode très facilement à d'autres seuils que 95%, par exemple 90%, 99% ou 99,5%.

Généralisons cette méthode : supposons qu'on cherche à déterminer l'intervalle de fluctuation au seuil de t% avec $t \in [0, 100]$. On doit avoir $P(p - \epsilon < M_m < p + \epsilon) \geqslant \frac{t}{100}$ avec la valeur de ϵ donnée par l'égalité $1 - \frac{p(1-p)}{m\epsilon^2} = \frac{t}{100}$ qui conduit à $\epsilon = 10\sqrt{\frac{p(1-p)}{m(100-t)}}$.
L'intervalle de fluctuation est alors $IF = [p - \epsilon; p + \epsilon]$.
Voici quelques valeurs obtenues à partir de cette formule pour $p = \frac{1}{6}$ et $m = 100$ ou $m = 1000$:

t	ε	p_ε	p+ε
90	0,118	0,049	0,285
91	0,124	0,042	0,291
92	0,132	0,035	0,298
93	0,141	0,026	0,308
94	0,152	0,015	0,319
95	0,167	0,000	0,333
96	0,186	-0,020	0,353
97	0,215	-0,048	0,382
98	0,264	-0,097	0,430
99	0,373	-0,206	0,539
99,5	0,527	-0,360	0,694

m	p
100	0,166666666

t	ε	p_ε	p+ε
90	0,037	0,129	0,204
91	0,039	0,127	0,206
92	0,042	0,125	0,208
93	0,045	0,122	0,211
94	0,048	0,119	0,215
95	0,053	0,114	0,219
96	0,059	0,108	0,226
97	0,068	0,099	0,235
98	0,083	0,083	0,250
99	0,118	0,049	0,285
99,5	0,167	0,000	0,333

m	p
1000	0,166666666