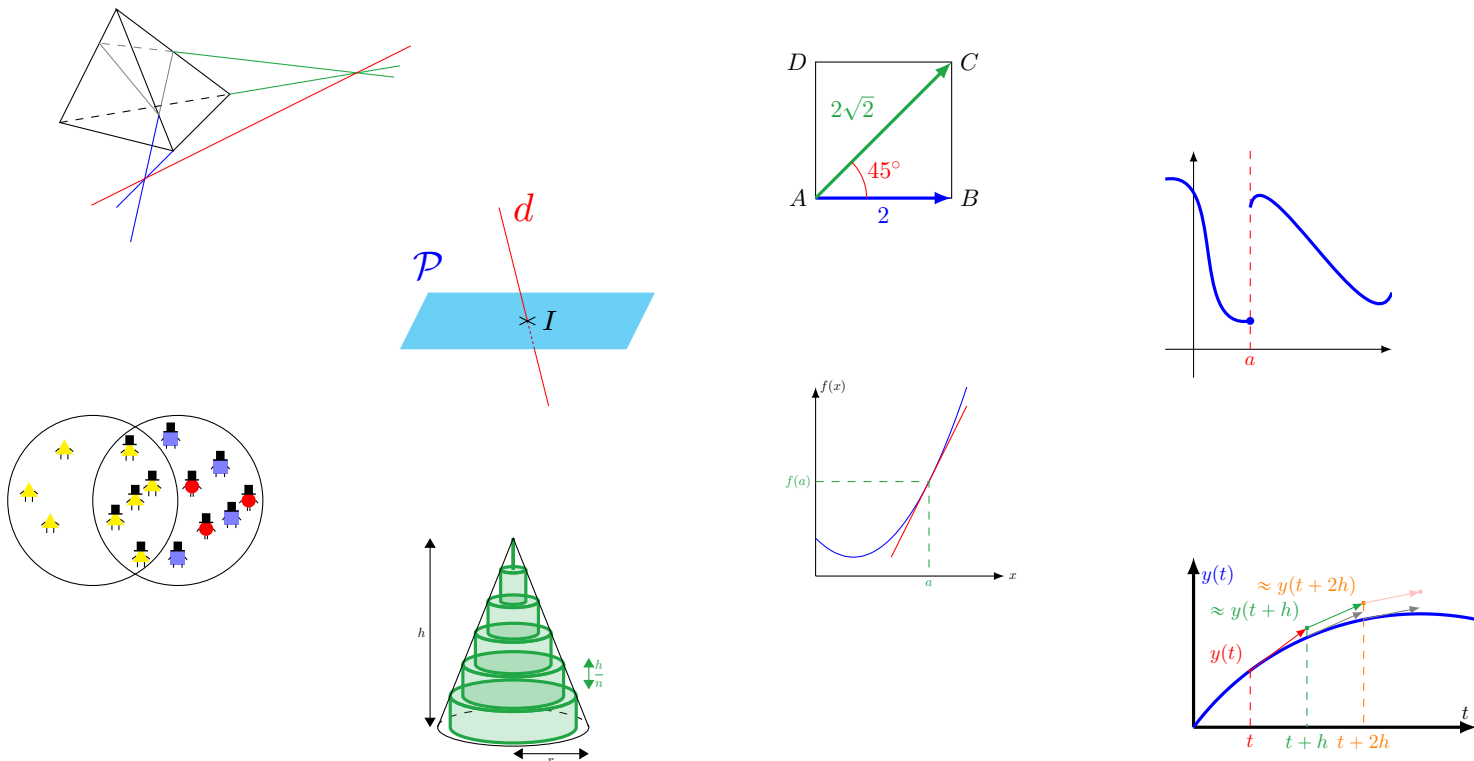
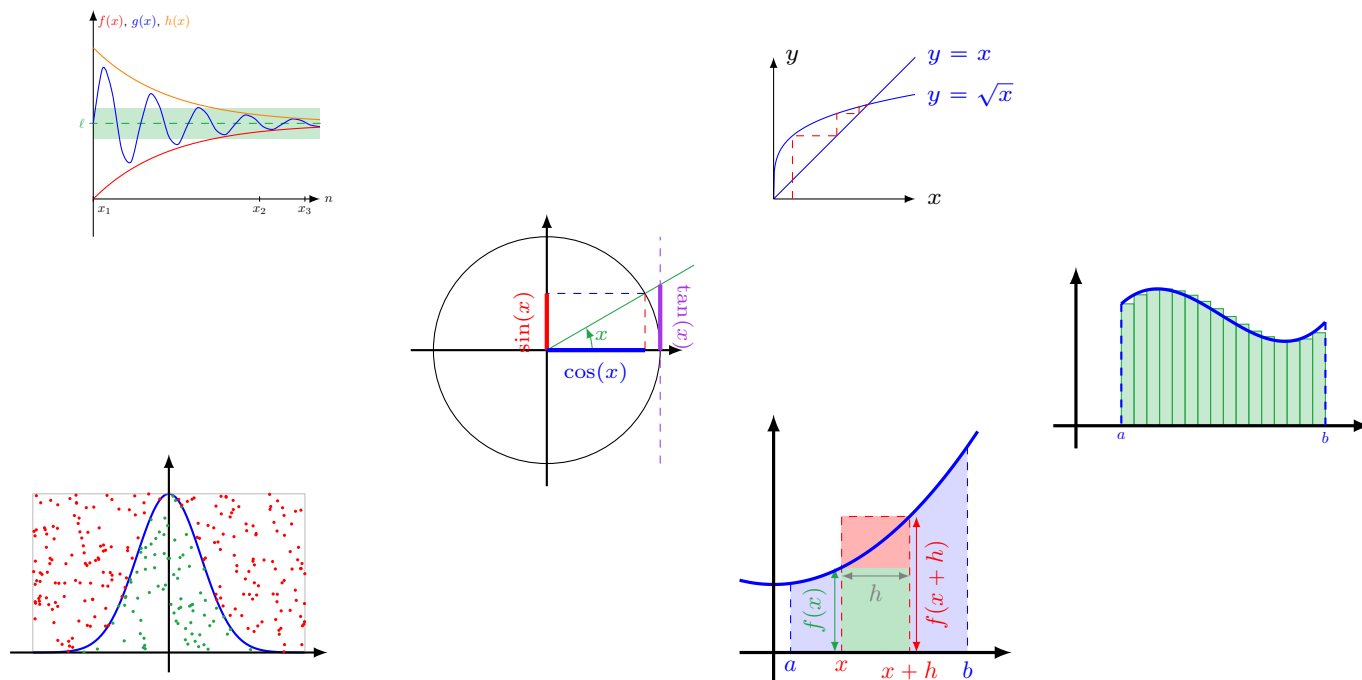




Mathématiques

Terminale spécialité – Extras



Cours de terminale générale

2	Suites et récurrence	7
1	Suites récurrentes linéaires d'ordre 2	7
3	Combinatoire et dénombrement	9
1	Combinaisons avec répétition	9
2	Différents infinis	9
4	Équations différentielles	13
1	Quelques compléments	13
A	Une astuce mnémotechnique	13
B	Méthode de variation de la constante	13
C	Existence et unicité des solutions	14
2	Résolution numérique : la méthode d'Euler explicite	14
3	Les équations différentielles pour modéliser une évolution	15
A	Lien avec les suites	15
B	Désintégration radioactive	15
C	Chute d'un corps	15
D	Oscillateur harmonique	16
4	Modèles démographiques et épidémiques	16
A	Modèles de Malthus et de Verhulst	16
B	Modèle SIR	17
5	Équations différentielles linéaires d'ordre 2 à coefficients constants	19
6	Géométrie et repérage dans l'espace	21
1	Introduction : balances à plateaux et romaines	21
2	Barycentre d'un système pondéré de points	22
3	Isobarycentre et médiane	24
4	Coordonnées barycentriques	25
7	Limites de fonctions	27
	Pour aller (plus) loin : branches infinies	27
8	Continuité	29
1	Résolution numérique d'équations	29
2	Étude d'une relation fonctionnelle	30
9	Loi binomiale	33
1	Problème de la surréservation	33
2	Loi de Poisson	33
3	Les cousines de la loi binomiale	34
A	Loi géométrique	34
B	Loi hypergéométrique	36
C	Loi binomiale négative	37
D	Loi multinomiale	39
10	Logarithme	41
1	La vraie définition de la notation puissance	41
2	Une célèbre limite	42
3	Échelles logarithmiques	42

4	Histoire du logarithme	42
11	Orthogonalité dans l'espace	43
1	Produit vectoriel et déterminant	43
A	L'idée du déterminant dans le plan, transposée à l'espace	43
B	Produit vectoriel de deux vecteurs	44
2	Autour des sphères	46
A	Rappel : cercles du plan	46
B	Équations de sphères	47
C	Sections de sphères	47
3	Pour la culture : mais à quoi ça sert ? !	49
4	Bonus : démonstration des propriétés du produit scalaire plan	51
12	Intégration	55
1	Série harmonique	55
2	Intégration numérique : méthode de Monte-Carlo	57
3	Intégrales, longueurs et volumes	59
A	Intro-digression : l'approximation de π par Archimède	59
B	Intégrales curvilignes	60
C	Intégrales et volumes	61
4	Changement de variable	62
13	Variables aléatoires	65
1	Covariance et corrélation	65
2	Intervalles de confiance	68
14	Convexité	71
1	Dérivée de la dérivée de la dérivée de la...	71
2	Inégalité arithmético-géométrique	73
3	Cercle osculateur et étude qualitative	73
15	Trigonométrie	75
1	Pour aller (plus) loin : la fonction tangente	75
2	Pour la culture : mais à quoi ça sert ? !	76
3	La trigonométrie post-bac	77
A	Fonctions inverses	77
B	Relations trigonométriques	78
C	Écriture trigonométrique des nombres complexes – fonctions hyperboliques	79
A	Rappels de probabilités	81
1	Rappels : vocabulaire et formulaire de probabilités	81
2	Succession d'épreuves indépendantes ou non	83
	Exercices	84
B	Second degré : algèbre et géométrie	87
1	Définition et forme développée	87
2	Forme canonique, aspects graphiques et variations	88
A	Représentation graphique : parabole	88
B	Forme canonique, conséquences graphiques	90
C	Variations	91
3	(In)équations de degré 2 exploitant la symétrie	92
4	Pour aller (plus) loin	93
C	Second degré : analyse	95
1	Racines d'une fonction polynôme	95
2	Discriminant d'une fonction polynôme du second degré	96

3	Forme factorisée et racines	97
4	Racines et représentation graphique	99
5	Étude du signe	100
D	Logique mathématique	101
1	Pour la culture : la méthode axiomatique	102
A	Un « peu » d'histoire : l'exemple des axiomes d'Euclide	102
B	À quoi ça sert, la logique ?	103
2	Logique propositionnelle	104
A	Propositions mathématiques	104
B	Connecteurs logiques et quantificateurs	105
i	Les opérateurs logiques	105
ii	Les quantificateurs	106
iii	Les connecteurs logiques	106
3	Les différents types de raisonnement	108
A	Les contre-exemples	108
B	Le raisonnement par déduction	109
C	Le raisonnement par contraposée	109
D	Le raisonnement par l'absurde	111
E	Le raisonnement par disjonction de cas	111
E	Compléments sur le logarithme	113
1	Propriétés analytiques du logarithme : démonstration	113
A	Continuité	113
B	Dérivabilité et dérivée	115
2	Définition alternative du logarithme et de l'exponentielle	115
F	Des lois de probabilité continues	117
1	Notion de loi continue	117
2	Loi uniforme	118
3	Loi exponentielle	118
4	Lois de puissance	118
5	Loi normale	118
G	La géométrie des courbes de fonctions	119

2 Suites et récurrence

1 – Suites récurrentes linéaires d'ordre 2

Soit (u_n) une suite. Une relation de récurrence n'est pas qu'une expression de u_{n+1} en fonction de u_n , elle autorise plus généralement l'intervention de tous les termes de rangs inférieurs.

Les suites récurrentes linéaires d'ordre 2 sont, en quelque sorte, une généralisation des suites géométriques, que l'on peut appeler suites récurrentes linéaires d'ordre 1. Plus précisément, on dit que (u_n) est récurrente linéaire d'ordre 2 s'il existe a et b deux réels tels que, pour tout n entier naturel, $u_{n+2} = au_{n+1} + bu_n$. Plutôt qu'un terme initial, il en faut deux : par exemple, soit (u_n) définie par $u_0 = 2$, $u_1 = 5$, et pour tout $n \in \mathbb{N}$, $u_{n+2} = -3u_{n+1} + u_n$. Dans ce cas, $u_2 = -3u_1 + u_0 = -13$, puis $u_3 = -3u_2 + u_1 = 44$, etc.

La plus célèbre d'entre toutes est la suite de Fibonacci, définie par $F_0 = 0$, $F_1 = 1$, et pour tout $n \in \mathbb{N}$, $F_{n+1} = F_{n+1} + F_n$ (avec donc $a = b = 1$). Autrement dit, chaque terme est la somme des deux précédents. On a donc, dans l'ordre, 0, 1, 1, 2, 3, 5, 8, 13, ...

Quel est le centième terme de la suite de Fibonacci ? On pourrait calculer tous les termes un par un jusqu'à F_{99} , mais ce serait long. Pour les suites arithmétiques ou géométriques, voire arithmético-géométrique, on peut trouver une expression explicite permettant de calculer un terme de rang quelconque sans connaître les précédents. Existe-t-il alors une formule permettant d'expliciter une suite récurrente linéaire d'ordre 2 ?

La réponse est : une formule, non, mais une méthode, oui. Elle repose sur la notion d'équation caractéristique, $X^2 = aX + b$. C'est une équation polynomiale du second degré, notons Δ son discriminant.

- ▶ Si $\Delta > 0$, alors l'équation caractéristique admet deux solutions réelles distinctes $r_1 \neq r_2$. Il existe λ et μ deux réels tels que pour tout $n \in \mathbb{N}$, $u_n = \lambda r_1^n + \mu r_2^n$.
- ▶ Si $\Delta = 0$, alors l'équation caractéristique admet une unique solution réelle r_0 . Il existe λ et μ deux réels tels que pour tout $n \in \mathbb{N}$, $u_n = (\lambda + \mu n)r_0^n$.
- ▶ Si $\Delta < 0$, alors l'équation caractéristique admet deux solutions complexes conjuguées $\rho e^{\pm i\theta}$. Il existe λ et μ deux réels tels que pour tout $n \in \mathbb{N}$, $u_n = \rho^n [\lambda \cos(n\theta) + \mu \sin(n\theta)]$.

Prenons pour exemple la suite de Fibonacci. Son équation caractéristique est $X^2 = X + 1$, soit $X^2 - X - 1 = 0$. Son discriminant est $\Delta = (-1)^2 - 4 \times 1 \times (-1) = 5$, donc l'équation admet deux solutions réelles distinctes $\varphi = \frac{1 + \sqrt{5}}{2}$ et $\psi = \frac{1 - \sqrt{5}}{2}$. Les yeux aiguisés reconnaîtront le nombre d'or dans ces racines. Ainsi, il existe λ et μ deux réels tels que pour tout $n \in \mathbb{N}$, $F_n = \lambda \varphi^n + \mu \psi^n$. Or $F_0 = 0$, donc $\lambda \varphi^0 + \mu \psi^0 = 0$, c'est-à-dire $\lambda + \mu = 0$, donc $\mu = -\lambda$. D'autre part, $F_1 = 1$, or $F_1 = \lambda \varphi^1 + \mu \psi^1$, c'est-à-dire $\lambda \varphi - \lambda \psi = 1$, donc $\lambda(\varphi - \psi) = 1$, or $\varphi - \psi = \frac{1 + \sqrt{5}}{2} - \frac{1 - \sqrt{5}}{2} = \sqrt{5}$, donc $\lambda = \frac{1}{\sqrt{5}}$. Finalement, pour tout $n \in \mathbb{N}$, $F_n = \frac{1}{\sqrt{5}}(\varphi^n - \psi^n)$. En remarquant que $\psi = -\frac{1}{\varphi}$, cette explicitation est parfois réécrite : $F_n = \frac{1}{\sqrt{5}}(\varphi^n - (-\varphi)^{-n})$.

3 Combinatoire et dénombrement

1 – Combinaisons avec répétition

Un homme louche a offert 100 bonbons à Alice au détour d'une ruelle sombre. Ne sachant que faire d'autant de bonbons, Alice décide de les partager avec ses camarades de classe. Combien de manières a-t-elle de distribuer au moins 1 bonbon à chacun des 20 élèves de sa classe ?

Alice a reçu 100 bonbons qu'elle souhaite distribuer au sein de sa classe, en s'assurant d'en donner au moins 1 à chacun des 20 élèves de sa classe (elle incluse), sans pour autant en donner la même quantité à tout le monde. Elle se demande combien de distributions possibles s'offrent à elle.

Comme presque tous les problèmes de dénombrement, la première étape est de trouver une manière adéquate de modéliser le problème. Ici, le problème correspond à une notion simplifiée de *combinaison avec répétition* : on note Γ_n^k le nombre de façons de choisir k éléments *avec répétition* sans ordre. Autrement dit, c'est le nombre de sous-ensembles d'un ensemble E à n éléments construits de la manière suivant : on choisit k éléments de E *successivement et avec remise*, et on forme un ensemble avec les éléments tirés. La subtilité de cette notion se voit peut-être mieux avec un exemple : étudions les combinaisons avec répétition à 2 éléments de l'ensemble $E = \{1, 2, 3\}$. Les tirages avec remise possibles de 2 éléments dans E sont $(1, 1)$, $(1, 2)$, $(1, 3)$, $(2, 2)$, $(2, 3)$, $(3, 3)$ et tous les tirages dans l'ordre inverse. Comme $\{1, 1\} = \{1\}$, les ensembles résultants sont $\{1\}$, $\{1, 2\}$, $\{1, 3\}$, $\{2\}$, $\{2, 3\}$, $\{3\}$. Ainsi, $\Gamma_3^2 = 6$.

Il est temps de résoudre le problème. Pour rappel, Alice doit distribuer 100 bonbons à sa classe de 20 élèves, en donnant au moins 1 bonbon à chacun. Alice numérote les élèves de 1 à 20. Elle a l'idée suivante : elle commence par aligner les 100 bonbons sur une table. Elle se munit ensuite de 19 bâtons, et les dispose entre les bonbons comme sur le schéma suivant. La distribution suit ensuite le principe suivant : l'élève 1 reçoit les bonbons avant le premier bâton, l'élève 2 reçoit les bonbons entre le premier bâton et le deuxième, etc. On ne peut pas placer plus d'un bâton entre deux bonbons, autrement dit il doit toujours y avoir un bonbon entre deux bâtons, correspondant au fait que chaque élève doit recevoir au moins 1 bonbon. Ainsi, Alice doit placer 19 bâtons dans les 99 espaces entre les bonbons, autrement dit choisir 19 emplacements parmi 99. Elle a donc $\binom{99}{19} = 107196674080761936594$ possibilités.

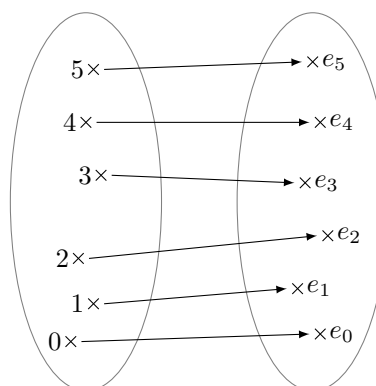
* * | * * * | * * | * ... * | * *

Exemple du placement de bâtons

De là, il est aisé de généraliser le raisonnement : le nombre de manières de distribuer n bonbons à p personnes, en donnant au moins 1 bonbon à chaque personne, est $\binom{n-1}{p-1}$. Notez que la condition "chaque personne reçoit au moins 1 bonbon" est capitale : en effet, sans celle-ci, on pourrait placer plusieurs bâtons entre deux bonbons, signifiant que la ou les personnes correspondantes reçoivent 0 bonbons. Le raisonnement n'est plus valide. On peut cependant montrer à l'aide d'une récurrence assez compliquée qu'il y a en toute généralité $\binom{n+p-1}{n}$ manières de distribuer n bonbons à p personnes.

2 – Différents infinis

Numéroter les éléments d'un ensemble, c'est attribuer à chaque élément de cet ensemble un unique entier naturel, autrement dit dire "toi tu es le numéro 0, toi le numéro 1, toi le numéro 2, ...". Décrit comme ça, ce procédé n'est pas très mathématique. Formellement, numéroter les éléments d'un ensemble, c'est trouver une fonction qui, à chaque entier naturel associe un élément de l'ensemble sans répétition, et de sorte que tout élément de l'ensemble sans exception est représenté. C'est une correspondance "un à un", les éléments de l'ensemble et les entiers naturels peuvent être associés en paires



sans conflit et sans oubli. Une telle fonction est dite *bijjective* (ou que c'est une *bijection*... ça vous rappelle quelque chose ? Celle-ci n'est pas continue, pour autant !).

Mais pour raisonner, on n'a pas besoin de trouver une expression universelle d'une telle fonction, une explication, une description du processus d'association est suffisante.

Armés de cette notion, nous pouvons définir plus rigoureusement ce qu'est un ensemble fini. On dit qu'un ensemble est fini s'il existe un entier n et une "fonction de numérotation" (une association unique et sans oubli, une bijection donc) entre les entiers de 1 à n et les éléments de l'ensemble. Cet entier n est unique, et est appelé cardinal de cet ensemble. Dans le cours, nous définissons un ensemble fini comme un ensemble ayant un nombre fini d'éléments... mais "nombre d'éléments" signifie "cardinal", et nous définissons le cardinal d'un ensemble comme son nombre d'éléments ! Cette définition est circulaire.

Dans la grande famille des ensembles qui ne sont pas finis, on distingue deux groupes : les ensembles dénombrables, et ceux qui ne le sont pas. On dit qu'un ensemble est *dénombrable* si on peut numéroter ses éléments avec tous les entiers naturels $\mathbb{N}$, et qu'il est *indénombrable* dans le cas contraire. Autrement dit, un ensemble dénombrable est un ensemble qui a autant d'éléments que $\mathbb{N}$. L'ensemble dénombrable le plus évident est $\mathbb{N}$ lui-même, puisque chaque entier peut se numéroter par lui-même. La question que l'on se pose alors est la suivante : de tous les ensembles élémentaires vus au lycée, lesquels sont dénombrables, lesquels ne le sont pas ?

Commençons par observer un phénomène typique des ensembles infinis. Si on enlève l'entier 0 pour former l'ensemble $\mathbb{N}^*$, a-t-on moins d'éléments que dans $\mathbb{N}$? La réponse est non, car on peut dire que 1 est le numéro 0, puis 2 est le numéro 1, etc. en décalant la numérotation de 1. Ainsi, chaque élément de $\mathbb{N}^*$ peut être numéroté par un entier, donc il y a autant d'éléments dans $\mathbb{N}^*$ que dans $\mathbb{N}$. Pire encore, ce raisonnement tient encore pour n'importe quelle suppression finie : si on retire le premier milliard d'éléments de $\mathbb{N}$, alors il suffit de décaler la numérotation de 1 milliard pour se rendre compte qu'il y a toujours autant d'éléments que dans $\mathbb{N}$.

$\emptyset$	$\cancel{1}$	$\cancel{2}$	$\cancel{3}$	$\cancel{4}$	5	6	7	8	9	10	11	12	...
					n°0	n°1	n°2	n°3	n°4	n°5	n°6	n°7	...

Il en faut donc plus que ça pour perturber l'infini de $\mathbb{N}$: essayons de lui enlever une infinité d'éléments. Enlevons par exemple tous les nombres impairs, qui sont en quantité infinie, pour ne garder que les nombres pairs. Considérons la fonction $f : n \mapsto 2n$. Avec cette fonction, l'entier numéroté 0 sera $2 \times 0 = 0$; l'entier numéroté 1 sera $2 \times 1 = 2$; l'entier numéroté 2 sera $2 \times 2 = 4$; ... On se convainc rapidement que de cette manière, on peut numéroter les entiers naturels pairs... Il y a donc autant d'entiers naturels pairs que d'entiers tout court !

0	2	4	6	8	10	12	14	16	...
n°0	n°1	n°2	n°3	n°4	n°5	n°6	n°7	n°8	...

Mais ça ne s'arrête pas là : le même raisonnement fonctionne pour les entiers naturels impairs... Il y a donc autant d'entiers naturels pairs que impairs, et que d'entiers naturels tout court, alors que les entiers naturels pairs et impairs sont deux groupes distincts formant les entiers naturels ! Un autre exemple est celui des nombres premiers. Il est bien connu depuis Euclide qu'ils sont en nombre infini : on peut montrer qu'il y en a autant que d'entiers naturels, ce qui paraît contre intuitif au vu de leur rareté !

Bon, nous avons échoué à prendre une partie de $\mathbb{N}$ significativement plus petite tout en restant infinie... Essayons dans l'autre sens, en prenant un ensemble vraisemblablement plus grand. Dans les ensembles vus au lycée, après $\mathbb{N}$ vient $\mathbb{Z}$. La numérotation décrite par le procédé ci-dessous montre qu'il y a autant d'entiers relatifs que de naturels :

*. L'étendue de la notion de dénombrabilité fait débat au sein de la communauté mathématique. Certains définissent un ensemble dénombrable comme un ensemble dont les éléments peuvent être numérotés, sans imposer que cela se fasse avec *tous* les entiers naturels, autrement dit les ensembles finis sont aussi considérés dénombrables. D'autres excluent les ensembles finis, afin d'enrichir le mot « dénombrable » pour qu'il renvoie spécifiquement aux ensembles infinis, et donc ne pas se contraindre à préciser qu'un ensemble donné est « infini dénombrable » ou « fini dénombrable » (le deuxième cas étant redondant, puisqu'un ensemble fini est nécessairement dénombrable).

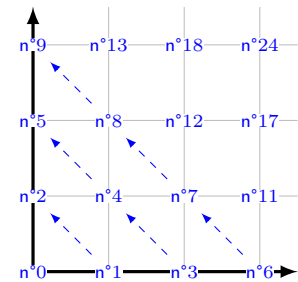
...	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	...
...	n°11	n°9	n°7	n°5	n°3	n°1	n°0	n°2	n°4	n°6	n°8	n°10	n°12	...

Encore un échec ! Essayons encore plus gros, avec $\mathbb{Q}$. Spoiler : nous allons voir qu'il est également dénombrable, mais l'argument est ici un peu plus subtil. Un nombre rationnel peut être décrit par la donnée d'un numérateur, entier relatif, et d'un entier naturel non nul. Cela peut se faire de manière unique si l'on impose la forme irréductible. Autrement dit, on peut construire une fonction f qui prend un rationnel sous forme irréductible et associe le couple (numérateur, dénominateur), par exemple $f\left(\frac{-5}{3}\right) = (-5; 3)$. Cette fonction part de $\mathbb{Q}$ et arrive dans l'ensemble des couples d'entiers sous la forme (relatif, naturel non nul), qui s'écrit avec le produit cartésien $\mathbb{Z} \times \mathbb{N}^*$. Ainsi, on s'est construit une représentation de $\mathbb{Q}$ à l'intérieur de $\mathbb{Z} \times \mathbb{N}^*$, sans répétitions (que l'on appelle alors une *injection*). Pour des raisons très compliquées *, il suffit de démontrer que $\mathbb{Z} \times \mathbb{N}^*$ est dénombrable. Comme on l'a vu plus tôt, $\mathbb{Z}$ et $\mathbb{N}^*$ ont autant d'éléments que $\mathbb{N}$ (c'est-à-dire qu'on peut trouver une correspondance un à un entre ces ensembles et $\mathbb{N}$, que nous appelons "numéroter" depuis le début), donc il suffit de montrer que $\mathbb{N} \times \mathbb{N}$ est dénombrable, en trouvant une numérotation des ses éléments.

Une manière de représenter $\mathbb{N} \times \mathbb{N}$ est de considérer ses couples d'entiers comme les coordonnées du quadrillage dans le quart de plan positif. Sur la représentation suivante, on propose une numérotation des éléments de $\mathbb{N} \times \mathbb{N}$:

Conclusion : il y a autant de paires d'entiers naturels que d'entiers naturels, *a fortiori* d'après l'explication précédente, tout autant de nombres rationnels. ‡

Dans les ensembles vus au lycée, il ne reste plus qu'un candidat pour être indénombrable... Essayons l'ensemble $\mathbb{R}$ des nombres réels.



En 1874, Georg Cantor a publié un papier autour de l'infini, présentant notamment toutes les bizarreries précédentes. C'est également lui qui donne une démonstration qui nous permettra enfin de terminer notre quête d'un infini plus grand, avec un argument qui porte désormais son nom. Cantor a montré qu'il y a infiniment plus de nombres réels entre 0 et 1 que d'entiers naturels, avec la démonstration suivante.

Par l'absurde, imaginons que l'intervalle $]0; 1[$ est dénombrable, c'est-à-dire que l'on peut numéroter ses éléments. On place les réels de $]0; 1[$ dans un tableau infini, le numéro de la ligne correspondant au numéro du réel dans notre numérotation hypothétique, et on y écrit leur développement décimal unique † :

0,	6	8	7	4	6	5	1	5	4	9	6	5	4	3	...
0,	6	7	1	6	5	7	4	3	6	5	7	6	4	3	...
0,	7	1	6	3	1	6	5	7	3	6	7	3	1	6	...
0,	6	6	1	6	5	4	1	6	5	4	1	5	4	1	...
0,	1	4	1	5	9	2	6	5	3	5	8	9	7	9	...
0,	1	8	9	7	9	3	8	3	8	6	3	6	4	3	...
0,	6	1	8	7	4	8	1	6	1	4	3	6	8	7	...
...															

À partir de cette liste, on construit un nouveau nombre réel, entre 0 et 1 : on prend le premier chiffre (après la virgule) du premier nombre, le deuxième du deuxième, le troisième du troisième, ... et on leur ajoute tous 1 (en bouclant à 0 au cas où le chiffre est 9). Par exemple, sur l'illustration, on obtiendrait 0,7877042... L'argument, dit « diagonal de Cantor », est maintenant le suivant : ce nombre n'est pas égal au premier, car son premier chiffre après la virgule diffère. Il n'est pas égal au deuxième, car son deuxième chiffre après la virgule diffère. Et ainsi de suite : on montre donc que ce nombre n'est égale à aucun des nombres de la liste. Pourtant, on était censé les avoir tous listés ! C'est donc absurde. Ainsi, il y a infiniment plus de nombres réels que de nombres entiers naturels.

*. Pour celles et ceux qui n'ont pas peur de n'absolument rien comprendre, voyez le théorème de Cantor-Bernstein.

‡. Pour les intéressé.e.s, un autre argument qui réexploite l'argument sophistiqué de l'injection utilise la fonction f qui, à un couple d'entier naturels (i, j) , associe l'entier naturel $2^i 3^j$. Par unicité de la décomposition en produit de facteurs premiers, cette fonction n'a aucune répétition, c'est-à-dire que deux couples de départ différents donnent forcément des résultats différents. C'est donc une injection, qui permet de représenter $\mathbb{N} \times \mathbb{N}$ à l'intérieur de $\mathbb{N}$, et donc conclure à sa dénombrabilité.

†. Chaque réel admet un unique développement décimal si l'on interdit celle-ci de se terminer par une répétition infinie du chiffre 9.

4 Équations différentielles

Table des matières

1	Quelques compléments	13
A	Une astuce mnémotechnique	13
B	Méthode de variation de la constante	13
C	Existence et unicité des solutions	14
2	Résolution numérique : la méthode d'Euler explicite	14
3	Les équations différentielles pour modéliser une évolution	15
A	Lien avec les suites	15
B	Désintégration radioactive	15
C	Chute d'un corps	15
D	Oscillateur harmonique	16
4	Modèles démographiques et épidémiques	16
A	Modèles de Malthus et de Verhulst	16
B	Modèle SIR	17
5	Équations différentielles linéaires d'ordre 2 à coefficients constants	19

1 – Quelques compléments

A – Une astuce mnémotechnique

Dans le cadre du cours de terminale, on considère toujours le coefficient a constant, dans l'équation différentielle. Pourtant il n'y a qu'un pas pour s'affranchir de cette contrainte. Que a est une fonction continue. Soit A une primitive de a . En guise d'exercice, démontrez que l'ensemble des solutions de $y' = a(x)y$ est $\{x \mapsto Ce^{A(x)} : C \in \mathbb{R}\}$: le principe est strictement le même que lorsque a était une constante.

Si vous avez un doute sur la formule ou sur le signe, voici un moyen mnémotechnique, que j'appelle « méthode du physicien » (attention à ne pas leur dire en face...) : on part de l'équation $y' = ay$. Imaginons qu'on puisse diviser par y : alors $\frac{y'}{y} = a$. Ce quotient ressemble furieusement à la dérivée de $\ln(y)$... Imaginons que ce soit le cas. On prend donc une primitive de part et d'autre de cette égalité. Attention à ne pas oublier la constante ! Cela donne $\ln(y) = ax + k$. Passons à l'exponentielle : $y = e^{\ln(y)} = e^{ax+k} = e^{ax}e^k$. Posons $C = e^k$: on retrouve ainsi $y = Ce^{ax}$! Bien évidemment, cette méthode ne doit jamais quitter votre brouillon... ce n'est d'ailleurs pas une méthode, c'est une astuce de notation dont chaque étape est mathématiquement douteuse.

B – Méthode de variation de la constante

Autre méthode, qui cette fois-ci n'est pas une astuce mais bel et bien une technique très répandue : la méthode dite de *variation de la constante*. Pour résoudre une équation de la forme $y' = ay + f$, il nous faut trouver une solution particulière, ce qui est très difficile en général. La méthode de variation de la constante nous permet d'en trouver une dans certains cas. On sait que les solutions de l'équation homogène sont de la forme $x \mapsto Ce^{ax}$. L'idée de cette méthode est de chercher une solution particulière de la forme $y_p : x \mapsto C(x)e^{ax}$: on ne voit plus C comme une constante, mais comme une fonction dérivable. Prenons un exemple. Considérons l'équation du dernier exercice du cours : $(E) : y' - 2y = e^{-x}$. L'équation homogène associée est $(E_0) : y' = 2y$, dont l'ensemble des solutions est

$\{x \mapsto Ce^{2x} : C \in \mathbb{R}\}$. Soit maintenant C une fonction dérivable, posons $y_p : x \mapsto C(x)e^{2x}$. Supposons que y_p est solution de (E). Alors $y_p' - 2y_p = x^2$, or :

$$e^{-x} = y_p'(x) - 2y_p(x) = C'(x)e^{2x} + 2C(x)e^{2x} - 2C(x)e^{2x} = C'(x)e^{2x}$$

donc $C'(x) = e^{-3x}$, donc on peut prendre $C(x) = \frac{-1}{3}e^{-3x}$. On vérifie ensuite en dérivant que $y_p : x \mapsto \frac{-1}{3}e^{-3x}e^{2x} = \frac{-1}{3}e^{-x}$ est bien une solution particulière de (E). Cette méthode est efficace quand on est capable de trouver une primitive de C' , ce qui peut s'avérer compliqué dans certains cas...

C – Existence et unicité des solutions

Pour terminer ces compléments, avez-vous remarqué dans les différents exercices qu'on demande souvent de trouver « l'unique solution f telle que $f(\dots) = \dots$ » ? Vous êtes-vous demandés si l'unicité venait du choix des valeurs, ou s'il y avait un résultat plus profond caché derrière ? Voici la réponse : le théorème de Cauchy-Lipschitz. Il stipule que sous certains hypothèses que l'on taira dans ce cours, il existe toujours une, et une seule solution passant par un point donné, autrement dit pour tous réels x_0 et y_0 (dans les bons domaines), il existe une et une seule solution f telle que $f(x_0) = y_0$. On parle de *condition initiale*, et cela peut s'interpréter physiquement : une équation différentielle régit l'évolution d'un système, et la condition initiale décrit son état de départ. Ce que dit ce théorème, c'est que connaissant la loi d'évolution et le point de départ, on peut décrire avec certitude le trajet exact que suivra le système.

2 – Résolution numérique : la méthode d'Euler explicite

Dans le vaste univers des équations différentielles, l'ensemble des équations que l'on est capable de résoudre explicitement est une île minuscule. Faut-il alors abandonner l'idée d'approcher une solution des ces autres équations différentielles qui nous semblent si inaccessibles ? Bien sûr que non ! Bien avant l'invention des ordinateurs, Euler a mis au point une méthode pour calculer de proche en proche une solution approximative d'une équation différentielle donnée.

Soit $y' = ay + b$ une équation différentielle (a et b sont deux réels) de condition initiale $y(t_0) = y_0$. L'approche d'Euler repose sur le constat suivant : pour tout $t \in \mathbb{R}$, $y'(t) = \lim_{h \rightarrow 0} \frac{y(t+h) - y(t)}{h}$ (par définition du nombre dérivé), donc, quand h est proche de 0 :

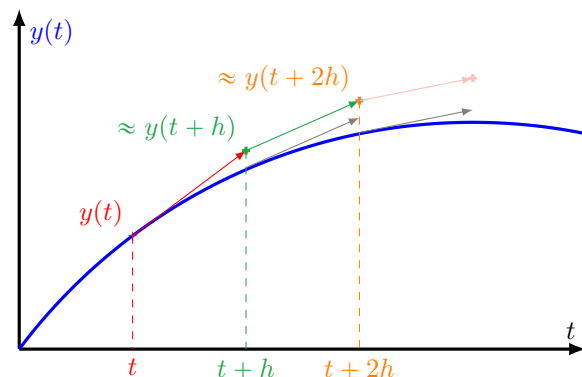
$$y(t+h) \approx y(t) + hy'(t) = y(t) + h \times (ay(t) + b)$$

Le principe de la méthode d'Euler est d'approximer de proche en proche $y(t_0 + h)$, $y(t_0 + 2h)$, $y(t_0 + 3h)$, ... en construisant une suite (y_n) par récurrence.

On fixe un $h > 0$. On considère les suites définies par :

$$\forall n \in \mathbb{N}, \quad \begin{cases} t_{n+1} = t_n + h \\ y_{n+1} = y_n + h \times (ay_n + b) \end{cases}$$

On illustre ci-contre l'idée derrière cette construction :



Cette visualisation illustre aussi le phénomène de divergence numérique : plus la simulation avance, et plus les approximations sont de moins en moins précises. En réglant la valeur de h notamment, on peut limiter cette explosion des erreurs.

Voici un exemple d'implémentation Python :

```

1 def euler_explicite(a, b, t_0, t_max, y_0, h):
2     t = [t_0]
3     y = [y_0]
4     while t[-1] < t_max:
5         t.append(t[-1] + h)
6         y.append(y[-1] + h*(a*y[-1] + b))
7     return t, y

```

Cette méthode peut être généralisée à des équations différentielles de formes plus variées.

3 – Les équations différentielles pour modéliser une évolution

A – Lien avec les suites

Voici une situation : vous avez 1000€ et chaque année vous multipliez votre capital par 1,2. Cette situation se modélise par une suite récurrente de relation $u_{n+1} = 1,2u_n$ et de terme initial $u_0 = 1000$. Les suites s'y prêtent bien car le pas de temps est constant : on regarde une année après l'autre. Les suites modélisent des événements dont le pas de temps est discret, autrement dit qu'on peut compter (le premier, le deuxième, le troisième...). Mais quand le pas de temps devient infiniment petit, que fait-on ? On tombe dans la catégorie des évolutions continues : c'est là tout l'intérêt des équations différentielles. Autrement dit, les équations différentielles sont à l'évolution continue ce que les suites sont à l'évolution discrète. Bien sûr, cette analogie a ses limites.

B – Désintégration radioactive

La loi de désintégration radioactive stipule que dans un échantillon de particules radioactives, la variation du nombre de particules qui se désintègrent est proportionnel au nombre de particules présentes. Notons N le nombre de molécules radioactives. La variation de N est représentée par la dérivée de N . La relation de proportionnalité stipule qu'il existe un réel $\lambda > 0$ tel que pour tout $t \in \mathbb{R}$, $N'(t) = -\lambda N(t)$ (le signe moins vient du fait que le nombre de particules est décroissant, donc la dérivée doit être négative sachant que l'on a pris $\lambda > 0$). Bien sûr, cela ne colle pas *exactement* à la réalité : un nombre de particule est nécessairement entier, donc N ne saurait être une fonction continue, *a fortiori* dérivable. Cependant, il s'agit là d'un exercice de *modélisation* : on tente de constituer un modèle mathématique qui représente le plus fidèlement possible le phénomène, sans tomber dans un excès de fidélité qui rendrait son étude impossible.

D'après le cours, les solutions de cette équation différentielle s'écrivent $N(t) = Ce^{-\lambda t}$, avec $C \in \mathbb{R}$ déterminée par une mesure du nombre de particules radioactives à un moment donné.

Impossible de ne pas mentionner ici la datation au carbone 14. Le carbone 14 est une particule radioactive, elle suit donc la loi de désintégration radioactive. Plutôt que de parler de nombre de particules, il est plus commode de parler de proportion : on se convainc aisément que l'équation différentielle en proportion est la même : $P'(t) = P_0 e^{-\lambda t}$, avec P_0 la proportion initiale de la particule radioactive – le carbone 14 –. On sait estimer P_0 , on mesure $P(t)$ en laboratoire, et on connaît λ : ceci nous permet d'en déduire t , qui nous donne donc l'âge de l'objet étudié.

C – Chute d'un corps

Quand on lance un ballon en l'air, il retombe : merci la gravité. Les lois de la mécanique de Newton nous permettent d'étudier plus précisément le mouvement de ce ballon : pour résumer, ces lois nous informent que l'accélération du ballon est constante, égale à l'accélération de pesanteur de la Terre (constante notée g). Notons y l'altitude du ballon. La dérivée y' de y correspond donc à sa vitesse. La dérivée y'' de y' correspond à la variation de la vitesse, ce qu'on appelle l'accélération. Ainsi, les lois de la mécanique nous permettent d'écrire que l'altitude y du ballon vérifie l'équation différentielle $y'' = -g$ (le signe moins témoigne du fait que le ballon retombe). Résolvons cette équation

différentielle : en prenant une primitive de part et d'autre de l'équation, on sait qu'il existe une constante v_0 telle que $y'(t) = v_0 - gt$ (en effet, y' est une primitive de y''). On peut renouveler l'exercice : il existe une constante p_0 telle que $y(t) = p_0 + v_0 t - g \frac{t^2}{2}$. Ainsi, nous disposons d'une équation décrivant l'altitude du ballon en fonction du temps. Les constantes p_0 et v_0 sont déterminées par la hauteur et la vitesse initiale du ballon.

D – Oscillateur harmonique

Prenez un ressort, attachez une extrémité au mur, et une masse à l'autre extrémité. Ensuite, déplacez la masse : celle-ci va se mettre à rebondir de haut en bas, tantôt tirant le ressort, tantôt tractée par celui-ci. Lorsque ce système est à l'équilibre et ne bouge pas, le ressort est fixe à une certaine longueur ℓ_0 . La masse m est soumise à son poids et la force exercée par le ressort. Elle est également repérée par la longueur du ressort : connaître la longueur $\ell(t)$ du ressort à un instant t donné permet de connaître la position de la masse, et inversement. En étudiant ce système, on trouve que les forces en présence – le poids et la traction du ressort – s'expriment respectivement mg et $-k(\ell(t) - \ell_0)$, où k est une constante qui décrit à quel point le ressort est souple ou rigide. L'accélération de la masse correspondant à la fonction ℓ'' , les lois de la mécanique nous fournissent l'équation différentielle $m\ell''(t) = mg - k(\ell(t) - \ell_0)$. En réarrangeant l'équation, on peut la mettre sous la forme $\ell'' + \omega^2 \ell = b$ avec ω et b deux réels, et $\omega > 0$. En guise d'exercice, vérifiez que les fonctions de la forme $x \mapsto \lambda \cos(\omega t) + \mu \sin(\omega t)$ sont solution de l'équation homogène.

4 – Modèles démographiques et épidémiques

A – Modèles de Malthus et de Verhulst

Thomas Malthus, dans le but d'illustrer l'épuisement des ressources vitales à la survie d'une population, propose un modèle élémentaire pour estimer la croissance de ladite population afin de prédire le moment où les besoins de cette population dépasseront les ressources limitées disponibles. Il propose un modèle de proportionnalité : la variation du nombre d'individus est proportionnelle au nombre d'individus actuellement présents. Si on note $N(t)$ le nombre d'individus vivant à l'instant t , alors la variation du nombre d'individus est mesurée par $N'(t)$. Ainsi, selon le modèle de Malthus, il existe un réel k tel que pour tout instant t , $N'(t) = kN(t)$. Le nombre $N(t)$ d'individus étant positif, pour que la population décroisse, il faut que k soit négatif, ce qui peut-être contradictoire avec la notion de proportionnalité qui ne s'énonce normalement qu'entre des grandeurs positives. Mais comme nous le verrons, ça ne pose pas de problème, ni du point de vue des formules, ni du point de vue de leur interprétation. Toutefois, dans l'esprit de Malthus, on supposera que $k > 0$, autrement dit que la population croît.

Une application élémentaire des formules vues dans ce cours permettent de montrer que pour tout instant t , $N(t) = N(0)e^{kt}$. En particulier, $\lim_{t \rightarrow +\infty} N(t) = +\infty$: la population croît sans jamais ralentir et son nombre diverge. Ce modèle n'est pas réaliste : en effet, contrainte par des ressources limitées, la croissance de la population devrait également être limitée ; plus le nombre d'individus se rapproche de la population maximale que les ressources permettent de supporter, plus la croissance doit ralentir, et le manque de ressources fait que le nombre d'individus ne peut dépasser un seuil.

Comme dit dans le paragraphe précédent, on peut raisonnablement supposer que la population est majorée par un nombre $N_{\max}$ d'individus, c'est-à-dire la population maximale que les ressources disponibles permettent de supporter. Il semble naturel de considérer que la croissance de la population dépend de l'état actuel de la population : s'il y a peu d'individus, les ressources sont abondantes et la croissance est forte ; *a contrario*, si la population est proche de la saturation, les ressources sont rares, et la croissance est presque nulle. On peut donc choisir un coefficient k qui dépend ainsi de $N(t)$: une dépendance simple est une dépendance affine, $k = aN(t) + b$. D'après ce qu'on vient de dire, Pour $N(t) = N_{\max}$, la croissance est nulle, donc $k = 0$, donc $aN_{\max} + b = 0$. D'autre part, si on note k_0 le coefficient de croissance maximal pour $N(t) = 0$ (ce qui, en un sens, est absurde, puisque si la population est vide, elle ne risque pas de croître !), alors $k_0 = a \times 0 + b$. Finalement, $k = k_0 \left(1 - \frac{N(t)}{N_{\max}}\right)$. L'équation différentielle devient alors $N'(t) = k_0 \left(1 - \frac{N(t)}{N_{\max}}\right) N(t)$. Cette équation, dit *logistique*, ne tombe dans aucune des catégories

vues en cours de terminale ; cependant, on peut ruser pour s'y ramener.

Posons $y(t) = \frac{1}{N(t)}$, en supposant la population non éteinte. Comme N est dérivable, y l'est, et :

$$y'(t) = \frac{-N'(t)}{N(t)^2} = \frac{-1}{N(t)^2} k_0 N(t) \left(1 - \frac{N(t)}{N_{\max}}\right) = -k_0 \times \frac{1}{N(t)} + \frac{k_0}{N_{\max}} = -k_0 y(t) + \frac{k_0}{N_{\max}}$$

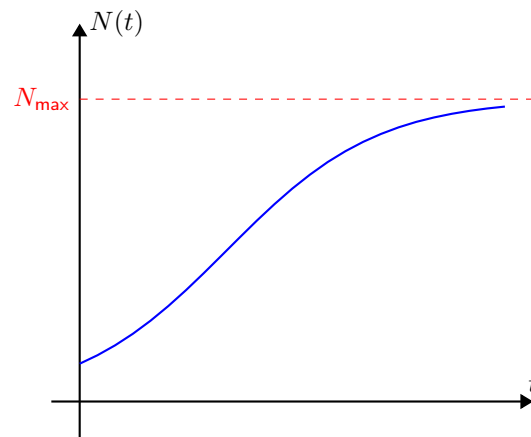
Cette fois-ci, y vérifie une équation de la forme $y' = ay + b$, que l'on sait résoudre ! En appliquant la méthode du cours, on obtient $y(t) = \frac{1}{N_{\max}} + \left(\frac{1}{N(0)} - \frac{1}{N_{\max}}\right) e^{-k_0 t}$. Finalement,

$$N(t) = \frac{1}{y(t)} = \frac{1}{\frac{1}{N_{\max}} + \left(\frac{1}{N(0)} - \frac{1}{N_{\max}}\right) e^{-k_0 t}} = \frac{N_{\max}}{1 + \left(\frac{N_{\max}}{N(0)} - 1\right) e^{-k_0 t}}$$

On peut vérifier certains aspects du modèle avec cette expression : déjà, pour $t = 0$, on retombe bien sur $N(0)$ après simplifications, et pour t qui tend vers $+\infty$, $e^{-k_0 t}$ tend vers 0 (car $k_0 > 0$), donc pas somme, produit et quotient, $N(t)$ tend vers $N_{\max}$.

Notons $\alpha = \frac{N_{\max}}{N(0)} - 1$. À supposer que la population de départ est inférieure à son maximum potentiel (ce qui n'est pas nécessairement vrai !), on a $\alpha > 0$. En dérivant une fois N , on obtient :

$$N'(t) = \frac{k_0 N_{\max} \alpha e^{-k_0 t}}{(1 + \alpha e^{-k_0 t})^2}$$



Les calculs suivants sont justifiés par le chapitre sur la convexité. En dérivant une seconde fois, on obtient – non sans peine – l'expression $N''(t) = k_0 N_{\max} \alpha e^{-k_0 t} \frac{-1 + \alpha e^{-k_0 t}}{1 + \alpha e^{-k_0 t}}$. Ainsi, pour tout instant t , $N''(t) > 0$ si, et seulement si $-1 + \alpha e^{-k_0 t} > 0$, ce qui se ramène à $t < \frac{\ln(\alpha)}{k_0}$. L'instant $t = \frac{\ln(\alpha)}{k_0}$ est donc le point d'inflexion de la courbe : c'est le moment où la croissance cesse d'accélérer et commence à ralentir.

B – Modèle SIR

(La lecture du document à l'adresse https://maths.ac-creteil.fr/IMG/pdf/5_modele_sir.pdf est conseillée pour plus de détails.)

Un des grands classiques des modèles épidémiques dits *compartimentaux* est le modèle SIR. Le S désigne les personnes saines, jamais contaminées, le I désigne les personnes infectées, et R désigne les personnes rétablies, autrement dit qui ont contracté la maladie, en ont guéri, et en sont immunisées. Le modèle SIR est un système d'équations différentielles modélisant l'évolution conjointe de ces trois grandeurs. Il formule plusieurs hypothèses fortes :

- ▶ On considère la taille de la population comme fixe, notamment la maladie ne provoque aucun décès.
- ▶ À chaque instant, les individus sains soit restent sains, soit deviennent infectés. Les individus infectés soit restent infectés, soit guérissent. Les individus rétablis ne retombent jamais malade.
- ▶ Une proportion fixe des personnes infectées guérie chaque jour.
- ▶ La probabilité de rencontre entre les individus sains et les infectés, et donc la probabilité de contamination, reste la même chaque jour.
- ▶ Au départ de l'épidémie, personne n'est immunisé et un seul patient est atteint de la maladie. Autrement dit, toute la population sauf une personne est saine.

Ces hypothèses sont contestables, mais elle vont déjà nous permettre de réfléchir un peu. Traduisons mathématiquement chaque hypothèse et leurs conséquences : à tout instant t ,

- ▶ $S(t) + I(t) + R(t) = N$ (une constante),
- ▶ S est décroissante et R est croissante, car les individus ne font que quitter (respectivement, intégrer) cette catégorie,
- ▶ La variation du nombre de rétablis est proportionnelle au nombre d'infectés : il existe $\lambda > 0$ tel que $R'(t) = \frac{1}{\lambda} I(t)$ (le choix de l'inverse s'illustrera plus tard),
- ▶ La variation du nombre de personnes saines est proportionnelle au nombre de rencontres : il existe $\beta > 0$ tel que $S'(t) = -\beta S(t)I(t)^*$,
- ▶ $S(0) \approx 1$ et $R(0) = 0$.

* : S'il y a $S(t)$ personnes saines et $I(t)$ personnes infectées, alors il y a $S(t) \times I(t)$ rencontres possibles. Si on note p la probabilité de rencontre, supposée fixe, entre une personne saine et une personne infectée, il y a donc $p \times S(t)I(t)$ rencontres effectives, à supposer que les rencontres sont équiprobables. Si l'on note q la probabilité d'infection, alors il y a $q \times pS(t)I(t)$ infections effectives. On retrouve la forme de la relation énoncée ci-dessus.

Quitte à remplacer $S(t)$, $I(t)$ et $R(t)$ par $\frac{S(t)}{N}$, $\frac{I(t)}{N}$ et $\frac{R(t)}{N}$ (ce que l'on peut faire car N est constant, ce qui n'influence pas les dérivées), on suppose que S , I et R ne représentent pas le nombre d'individus de chaque catégorie, mais leur proportion dans la population. Ainsi, ces trois quantités sont entre 0 et 1.

Plus haut, on a exprimé deux équations différentielles faisant intervenir $S'(t)$ et $R'(t)$. Qu'en est-il de $I'(t)$? Le fait que $S + I + R$ est une fonction constante permet de répondre à cette question. En effet, la dérivée d'une fonction constante est nulle. Or la dérivée d'une somme est la somme des dérivées, donc $S'(t) + I'(t) + R'(t) = 0$. Ainsi, $I'(t) = -S'(t) - R'(t) = \beta S(t)I(t) - \frac{1}{\lambda} I(t) = \left(\beta S(t) - \frac{1}{\lambda} \right) I(t)$.

L'humanité étant maintenant experte en épidémies, tout le monde est familier avec la notion de « pic de l'épidémie », qui correspond au moment où le nombre de malades cesse d'augmenter ; il est maximal et commence à redescendre. Le maximum d'une fonction implique que sa dérivée à cet endroit est nul, autrement dit avec nos notations $I'(t) = 0$.

D'après la relation trouvée précédemment, cela donne une équation produit nul : soit $S(t) = 0$, soit $\beta S(t) - \frac{1}{\lambda} = 0$, autrement dit $S(t) = \frac{1}{\beta\lambda}$. Interprétons le premier cas : s'il n'y a aucune personne saine, alors bien sûr le nombre d'infectés n'évolue pas : c'est la fin de l'épidémie. Le second cas est celui qui correspond au pic, mais son existence n'est pas garantie ! En effet, par hypothèse $S(t)$ est compris entre 0 et 1. Or $\frac{1}{\beta\lambda}$ ne l'est pas forcément : plus précisément, il l'est si, et seulement si $\beta\lambda > 1$. Cette quantité est appelée *taux de reproduction de base*, et on le note traditionnellement r_0 . Ainsi, l'épidémie présentera un pic si, et seulement si $r_0 > 1$.

On ne peut pas résoudre ces équations différentielles, mais, comme on vient de l'illustrer, ceci ne nous empêche pas de les étudier. On vient de caractériser l'existence même de l'épidémie, on peut également étudier son évolution à long terme. Grâce à une version fonctionnelle du théorème de convergence monotone,

- ▶ S est décroissante et minorée par 0, donc elle converge. On note S_∞ sa limite.
- ▶ R est croissante et majorée par 1, donc elle converge.
- ▶ $S + I + R = 1$, or S et R convergent, donc I converge.

On peut trouver une relation vérifiée par S_∞ , mais elle nécessite une grosse astuce. Remarquons que :

$$\frac{I'(t)}{S'(t)} = \frac{\beta S(t)I(t) - \frac{1}{\lambda} I(t)}{-\beta S(t)I(t)} = -1 + \frac{1}{\beta\lambda S(t)} = -1 + \frac{1}{r_0} \frac{1}{S(t)}$$

L'astuce est de considérer I comme une fonction du nombre d'individus sains, au lieu d'être une fonction du temps (ce qui est un point de vue raisonnable). L'erreur à ne pas faire cependant est celle que l'on voit le plus couramment :

cette erreur est d'écrire $I'(t) = \frac{dI}{dt}$, $S'(t) = \frac{dS}{dt}$, puis :

⚠ ABUSIF ⚠ $\frac{I'(t)}{S'(t)} = \frac{\frac{dI}{dt}}{\frac{dS}{dt}} = \frac{dI}{dt} \times \frac{dt}{dS} = \frac{dI}{dS}$ ⚠ ABUSIF ⚠

La notation $\frac{d}{dt}$ n'est pas une écriture fractionnaire, c'est une *notation*. On ne peut pas l'inverser, la simplifier, etc. comme on le souhaite. En fait, elle a un fond de vérité, elle repose sur ce qu'on appelle la *règle de la chaîne*, avec les fonctions à plusieurs variables... C'est bien au-delà du niveau terminale, donc on va l'admettre.

Ainsi, on dispose de la relation $\frac{dI}{dS} = -1 + \frac{1}{r_0} \frac{1}{S}$, donc en primitivant les deux membres comme fonctions de $S > 0$, il existe une constante C telle que $I(S) = -S + \frac{1}{r_0} \ln(S) + C$. Or, on a supposé qu'au début de l'épidémie, $S \approx 1$ et $I \approx 0$, donc C doit être égal à 1. Ainsi, $I = 1 - S + \frac{1}{r_0} \ln(S)$, ce qui se réécrit, après quelques manipulations, $S = e^{r_0(I+S-1)}$. On peut maintenant faire réapparaître la dépendance temporelle : $S(t) = e^{r_0(I(t)+S(t)-1)}$.

Or, comme dit précédemment, $S(t)$ converge vers S_∞ , et $I(t)$ converge vers 0 (les malades finissent tous par guérir un jour). Par continuité de la fonction exponentielle, le passage à la limite de la relation précédente donne $S_\infty = e^{r_0(S_\infty-1)}$. On aimerait connaître la valeur de S_∞ en fonction de r_0 , mais... on ne sait pas résoudre cette équation. On peut malgré tout l'étudier : combien de solutions a-t-elle ?

Considérons la fonction f définie sur $[0; 1]$ par $f(x) = e^{r_0(x-1)} - x$, de sorte que S_∞ vérifie l'équation $f(x) = 0$. Étudions f : elle est dérivable comme somme et composée, et $f'(x) = r_0 e^{r_0(x-1)} - 1$. Pour tout $x \in \mathbb{R}$, par stricte croissance du logarithme, et comme $r_0 > 0$,

$$f'(x) > 0 \iff r_0 e^{r_0(x-1)} > 1 \iff r_0(x-1) > \ln(1/r_0) = -\ln(r_0) \iff x > 1 - \frac{\ln(r_0)}{r_0}$$

Notons $x_0 = 1 - \frac{\ln(r_0)}{r_0}$. On peut raisonnablement restreindre notre étude à un cas épidémique, c'est-à-dire tel que $r_0 > 1$ d'après l'étude plus haut. Ainsi, $\frac{\ln(r_0)}{r_0} > 0$, donc $x_0 < 1$. De plus, par convexité du logarithme, pour tout $x > 0$, $\ln(x) \leq x - 1$, en particulier $\ln(r_0) \leq r_0 - 1 < r_0$, donc $\frac{\ln(r_0)}{r_0} < 1$, donc $x_0 > 0$. On en déduit donc le tableau de variations de f sur $[0; 1]$.

Par croissance de f sur $[x_0; 1]$, on a $f(x_0) \leq f(1) = 0$. De plus, $f(0) = e^{-r_0} > 0$. Comme f est continue et strictement décroissante sur $[0; 1]$, et que $f(0) > 0 \geq f(x_0)$, d'après le théorème de la bijection continue, l'équation $f(x) = 0$ admet une unique solution sur cet intervalle, que l'on note α . Par stricte croissance sur $[x_0; 1]$, pour tout $x < 1$, $f(x) < f(1) = 0$, donc l'équation $f(x) = 0$ n'admet pas de solution sur $[x_0; 1[$. Ainsi, l'équation $f(x) = 0$ admet exactement deux solutions : α et 1.

x	0	x_0	1
$f'(x)$		−	+
f	e^{-r_0}	$f(x_0)$	0

Autrement dit, soit la population saine tend vers une proportion de 1, soit vers une proportion α comprise entre 0 et 1 strictement. Le premier cas est contradictoire : S est décroissante, donc elle est supérieure à sa limite, c'est-à-dire $S(t) \geq 1$. Or, en tant que proportion, elle est plus petite que 1, c'est-à-dire $S(t) \leq 1$. Ainsi, $S(t)$ est à la fois plus grande et plus petite que 1, elle est donc toujours égale à 1. C'est le scénario sans épidémie, que l'on a exclu. Ainsi, $S_\infty = \alpha$: à la fin de l'épidémie, une certaine proportion de la population aura esquivé la maladie, et cette proportion vérifie une équation dépendant uniquement de r_0 .

5 – Équations différentielles linéaires d'ordre 2 à coefficients constants

Les équations différentielles linéaires d'ordre 2 sont une généralisation des équations différentielles linéaires d'ordre 1 – « coefficients constants » signifie que les quantités multipliant y et ses dérivées sont des réels et non des fonctions, ce qui est possible en toute généralité. Plus précisément, on dit que y satisfait une équation différentielle homogène linéaire d'ordre 2 à coefficients constants lorsqu'il existe b et c deux réels tels que, pour tout n entier

naturel, $y'' + by' + cy = 0$. Après avoir résolu l'équation homogène, on tient compte du second membre comme d'habitude, avec une solution particulière.

De même que pour les équations différentielles linéaires d'ordre 1 à coefficients constants, existe-t-il une formule permettant de résoudre le cas d'ordre 2 ?

La réponse est : une formule, non, mais une méthode, oui. Elle repose sur la notion d'*équation caractéristique*, $X^2 + bX + c = 0$. C'est une équation polynomiale du second degré, notons Δ son discriminant.

- ▶ Si $\Delta > 0$, alors l'équation caractéristique admet deux solutions réelles distinctes $r_1 \neq r_2$. Il existe λ et μ deux réels tels que pour tout $x \in \mathbb{R}$, $f(x) = \lambda e^{r_1 x} + \mu e^{r_2 x}$.
- ▶ Si $\Delta = 0$, alors l'équation caractéristique admet une unique solution réelle r_0 . Il existe λ et μ deux réels tels que pour tout $x \in \mathbb{R}$, $f(x) = (\lambda + \mu x)e^{r_0 x}$.
- ▶ Si $\Delta < 0$, alors l'équation caractéristique admet deux solutions complexes conjuguées $\alpha \pm i\beta$. Il existe λ et μ deux réels tels que pour tout $x \in \mathbb{R}$, $f(x) = e^{\alpha x}[\lambda \cos(\beta x) + \mu \sin(\beta x)]$.

Pour trouver λ et μ , on a besoin de deux conditions initiales au lieu d'une : l'image et la dérivée d'une abscisse donnée.

Prenons comme premier exemple la chute libre d'un corps, déjà vue plus haut, mais avec ce nouveau théorème : $y'' = -9,81$. L'équation homogène associée est $y'' + 0y' + 0y = 0$, l'équation caractéristique associée est donc $X^2 + 0X + 0 = 0$, soit $X^2 = 0$. Son discriminant est 0, elle admet comme unique solution $X = 0$. Il existe donc λ et μ deux réels tels que pour tout x réel, $y_h(x) = (\lambda x + \mu)e^{0x} = \lambda x + \mu$. Rien de surprenant, si $y_h'' = 0$, alors en primitivant une fois, y_h' est constante, et en primitivant à nouveau, y_h est affine !

Le petit bout de raisonnement précédent se répète pour trouver une solution particulière : en primitivant deux fois, on constate que $y_p(x) = -9,81 \frac{x^2}{2}$ est une solution particulière. Ainsi, les solutions de l'équation sont de la forme $y(x) = y_p(x) + y_h(x) = -9,81 \frac{x^2}{2} + \lambda x + \mu$. Supposons que $y(0) = 1$ et $y'(0) = 0$. La première condition nous donne $\mu = 1$. En remarquant que $y'(x) = -9,81x + \lambda$, on obtient $\lambda = 0$ avec la seconde condition. Finalement, $y(x) = -9,81 \frac{x^2}{2} + 1$.

Le deuxième exemple est laissé en exercice : traitez le cas de l'oscillateur harmonique vu plus haut, $y'' + \omega^2 y = 0$.

Comme troisième et dernier exemple, traitons l'oscillateur amorti : $y'' + 6y' + 13y = 0$. L'équation caractéristique est $X^2 + 6X + 13 = 0$. Son discriminant est $\Delta = -16$. Ses solutions complexes conjuguées sont donc $-3 \pm 2i$, et il existe λ et μ deux réels tels que pour tout $x \in \mathbb{R}$, $y(x) = e^{-3x}[\lambda \cos(2x) + \mu \sin(2x)]$. Exercice : trouver λ et μ tels que $y(0) = 5$ et $y'(0) = -4$.

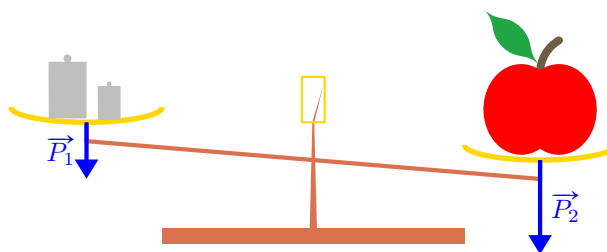
6 Géométrie et repérage dans l'espace

Table des matières

1	Introduction : balances à plateaux et romaines	21
2	Barycentre d'un système pondéré de points	22
3	Isobarycentre et médiane	24
4	Coordonnées barycentriques	25

1 – Introduction : balances à plateaux et romaines

Une balance à plateaux permet de comparer deux masses. Pour mesurer la masse d'un objet, on la compare donc à une masse de référence connue : on place l'objet à mesurer sur un plateau, et on ajuste la masse de l'autre plateau jusqu'à que ceux-ci soient au même niveau.



Sur le schéma la force exercée par la pomme sur le fléau est supérieure à celle exercée par les poids, donc le fléau tourne du côté de la pomme. Pour que le fléau soit à l'équilibre, il faut que les forces de part et d'autre du point de pivot soient égales. Enfin... pas tout à fait les forces. Si vous avez déjà ouvert une porte dans votre vie, vous devriez savoir qu'il est plus facile de la faire tourner en la poussant loin de l'axe. Ainsi, la grandeur à équilibrer n'est en réalité pas la force, mais ce qu'on appelle le *moment*, qui est le produit de la force et de la distance entre l'axe de rotation et le point d'application. Plus on applique la force loin de l'axe, plus la rotation est simple, et donc le moment est élevé ; et inversement. Et naturellement, plus la force est intense, plus la rotation du fléau est simple.

Dans le cas de la balance à plateaux, si on note ℓ la longueur des bras, on cherche donc l'égalité $\ell \vec{P}_1 = \ell \vec{P}_2$, ce qui revient à l'égalité des poids puisque les deux bras sont de longueur égale. Il manque une information tacite dans cette égalité : il faut que les deux objets soient symétriquement placés de part et d'autre du point de pivot. En effet, si on met les deux plateaux du même côté, on ne risque pas de trouver l'équilibre ! Quel objet mathématique contient à la fois une information de longueur et de direction... un vecteur ! Plutôt que d'utiliser la seule longueur ℓ , on va plutôt se servir des vecteurs reliant le point de pivot et chacun plateaux ; pour la suite, on notera A_1 le point représentant le plateau de gauche, A_2 le plateau de droite, et O le point de pivot. Ce changement de point de vue crée un nouveau problème : on ne sait pas multiplier deux vecteurs, donc la quantité $\vec{OA}_1 \times \vec{P}_1$, qui remplace $\ell \times \vec{P}_1$ n'a plus de sens. Le poids étant uniquement déterminé par la masse de l'objet, on peut remplacer $\vec{P}_1$ par m_1 qui contient autant d'informations. Pour résumer, la relation fondamentale au cœur de l'équilibre de la balance est $m_1 \vec{OA}_1 = -m_2 \vec{OA}_2$ (le signe moins traduit le fait que A_1 et A_2 sont symétriques par rapport à O), soit encore $m_1 \vec{OA}_1 + m_2 \vec{OA}_2 = \vec{0}$.

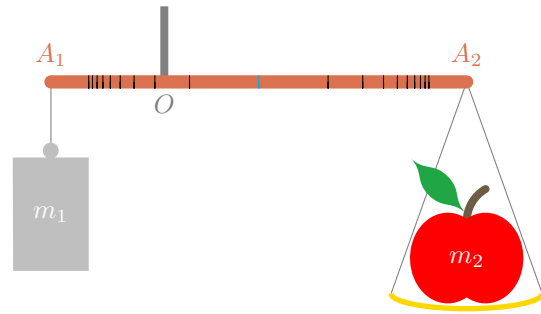
Que se passe-t-il si les deux bras ne sont pas de longueur égale ? C'est là tout le principe de la *balance romaine*. Au lieu de faire varier la masse de référence, on fait varier la position du pivot.

Le raisonnement mené précédemment est encore valable : pour atteindre l'équilibre, il faut que $m_1 \vec{OA}_1 + m_2 \vec{OA}_2 = \vec{0}$. Cependant, cette fois-ci, ce n'est plus m_1 que l'on ajuste, mais le point O .

Pour la balance à plateaux, il est assez intuitif de dire qu'il faut que la variable, m_1 , soit égale à la masse à mesurer, m_2 . Comment utilise-t-on une balance romaine ? Commençons par déterminer la position du point O pour que la balance soit à l'équilibre. La relation $m_1 \vec{OA}_1 + m_2 \vec{OA}_2 = \vec{0}$ se réécrit :

$$m_1 OA_1 \frac{\overrightarrow{OA_1}}{OA_1} = -m_2 OA_2 \frac{\overrightarrow{OA_2}}{OA_2}$$

Les vecteurs $\frac{\overrightarrow{OA_1}}{OA_1}$ et $\frac{\overrightarrow{OA_2}}{OA_2}$ sont de norme 1. Les points A_1 , O et A_2 sont alignés dans cet ordre, les deux précédents vecteurs sont opposés, et donc la relation d'équilibre est équivalente à ce que $m_1 OA_1 = m_2 OA_2$, autrement dit $m_2 = \frac{OA_1}{OA_2} m_1$.



Le rapport des longueurs des deux bras permet donc de calculer la masse de la pomme à partir de la masse de référence. En remarquant que $OA_1 + OA_2 = A_1A_2$, c'est-à-dire la longueur du fléau, on peut réécrire cette relation comme $m_2 = \frac{x}{1-x} m_1$, où $x = \frac{OA_1}{A_1A_2}$. Ainsi, si le pivot est placé à $x = \frac{1}{3}$ de la barre, ça signifie que $m_2 = 2m_1$.

Inversement, si on veut savoir où placer le pivot lorsque $m_2 = 3m_1$, il faut résoudre $\frac{x}{1-x} = 3$, ce qui donne $x = \frac{1}{4}$. Il faut donc placer le pivot au quart de la longueur. De manière générale, si on veut que $m_2 = km_1$ avec $k > 0$, il faut placer le pivot à la position $x = \frac{1}{k+1}$. Cette formule a permis de tracer les graduations sur le schéma précédent, correspondant à $m_2 = nm_1$ et $m_2 = \frac{1}{n}m_1$ pour n entier entre 2 et 10 – la graduation cyan correspondant $m_2 = m_1$, les autres étant respectivement à gauche et à droite de cette graduation –.

2 – Barycentre d'un système pondéré de points

Définition 1

Système pondéré de points, barycentre

Soient A_1 et A_2 deux points, et m_1 et m_2 deux réels. L'ensemble $\{(A_1, m_1); (A_2, m_2)\}$ est appelé un *système pondéré de points*.

S'il existe, le *barycentre* du système pondéré de points $\{(A_1, m_1); (A_2, m_2)\}$ est l'unique point G vérifiant :

$$m_1 \overrightarrow{GA_1} + m_2 \overrightarrow{GA_2} = \vec{0}$$

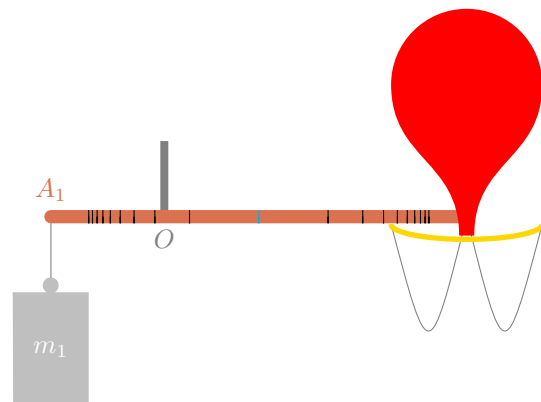
Ces définitions se généralisent naturellement à un système contenant plus de 2 points.

Remarque 2

La notion de barycentre généralise la notion de moyenne pondérée, pour des points plutôt que des nombres.

La situation de la balance romaine est donc celle du barycentre du système contenant les points aux extrémités du fléau, pondérés par les masses associées. Cependant, dans le cas de la balance romaine, les pondérations sont positives, ce qui n'est pas une contrainte de la définition : celle-ci autorise des masses négatives ! Si on voulait donner un sens physique à ce phénomène, imaginez attacher un ballon d'hélium au plateau, à la place de la pomme.

Il va sans dire que, dans la situation du schéma, le fléau n'est pas à l'équilibre : il ressent une force qui le tire vers le bas à gauche, et une autre vers le haut à droite, donc le fléau va tourner dans le sens trigonométrique.



Pourtant, on peut trouver un pivot permettant l'équilibre, autrement dit un barycentre ! En effet, celui-ci sera à gauche de la masse, et lieu d'être au milieu des deux objets. Comment peut-on caractériser les situations où l'on peut trouver le barycentre, et celles où ce n'est pas possible ? Pour nous aider dans cette quête, on introduit un nouvel outil :

Définition 3**Fonction vectorielle de Leibniz**

Soit $\{(A_1, m_1); (A_2, m_2)\}$ un système pondéré de points. La *fonction vectorielle de Leibniz* associée à ce système est la fonction f qui, à un point M du plan associe le vecteur :

$$\overrightarrow{f(M)} = m_1 \overrightarrow{MA_1} + m_2 \overrightarrow{MA_2}$$

Cette définition s'étend naturellement à des systèmes contenant plus de 2 points.

Ainsi, par définition, un système pondéré de points admet un barycentre si, et seulement si l'équation $\overrightarrow{f(M)} = \vec{0}$ admet une unique solution, où f est la fonction vectorielle de Leibniz associée au système. Jouons un peu avec cette fonction : soient M et N deux points. Alors :

$$\begin{aligned} \overrightarrow{f(M)} - \overrightarrow{f(N)} &= (m_1 \overrightarrow{MA_1} + m_2 \overrightarrow{MA_2}) - (m_1 \overrightarrow{NA_1} + m_2 \overrightarrow{NA_2}) \\ &= m_1 (\overrightarrow{MA_1} - \overrightarrow{NA_1}) + m_2 (\overrightarrow{MA_2} - \overrightarrow{NA_2}) \\ &= m_1 (\overrightarrow{MA_1} + \overrightarrow{A_1N}) + m_2 (\overrightarrow{MA_2} + \overrightarrow{A_2N}) \\ &= m_1 \overrightarrow{MN} + m_2 \overrightarrow{MN} \quad (\text{Chasles}) \\ &= (m_1 + m_2) \overrightarrow{MN} \end{aligned}$$

Ce petit calcul se généralise à des systèmes contenant plus de 2 points. Il montre que la masse totale du système, $m_1 + m_2$ joue un rôle prépondérant.

- ▶ Si $m_1 + m_2 = 0$, alors pour tous points M et N , on a $\overrightarrow{f(M)} = \overrightarrow{f(N)}$. De là, deux cas :
 - S'il existe un point M tel que $\overrightarrow{f(M)} = \vec{0}$, alors f est constante égale à $\vec{0}$, et donc l'équation $\overrightarrow{f(M)} = \vec{0}$ admet une infinité de solutions.
 - Sinon, f est constante égale à un vecteur non nul, et l'équation $\overrightarrow{f(M)} = \vec{0}$ n'admet aucune solution.

Dans tous les cas, il n'existe pas de barycentre.

- ▶ Sinon, si $m_1 + m_2 \neq 0$, alors d'après la relation précédente pour $N = O$,

$$\overrightarrow{f(M)} = \overrightarrow{f(O)} + (m_1 + m_2) \overrightarrow{MO} = m_1 \overrightarrow{OA_1} + m_2 \overrightarrow{OA_2} + (m_1 + m_2) \overrightarrow{MO}$$

Ainsi,

$$\overrightarrow{f(M)} = \vec{0} \iff m_1 \overrightarrow{OA_1} + m_2 \overrightarrow{OA_2} + (m_1 + m_2) \overrightarrow{MO} = \vec{0} \iff \overrightarrow{OM} = \frac{m_1}{m_1 + m_2} \overrightarrow{OA_1} + \frac{m_2}{m_1 + m_2} \overrightarrow{OA_2}$$

En conclusion, si la somme des pondérations n'est pas nulle, alors l'équation $\overrightarrow{f(M)} = \vec{0}$ admet une unique solution, qui donc est le centre de gravité.

Théorème 4**Existence du barycentre**

Soient $\mathcal{S}$ un système pondéré de points et f sa fonction vectorielle de Leibniz associée.

Le système $\mathcal{S}$ admet un barycentre si, et seulement si la somme de ses pondérations est non nulle.
Le cas échéant, le barycentre vérifie la relation explicitée plus haut.

Isobarycentre

Définition 5

Lorsque les pondérations d'un système de points sont toutes égales (et non nulles), le barycentre est appelé *isobarycentre* du système de points.

Propriété 6**Linéarité**

Soient $\mathcal{S} = \{(A_1, m_1); (A_2, m_2)\}$ un système pondéré de points tel que $m_1 + m_2 \neq 0$, et k un réel non nul. Le système $\{(A_1, km_1); (A_2, km_2)\}$ a le même barycentre que $\mathcal{S}$.

Corollaire 7

Lorsqu'un système pondéré de points admet un barycentre, on peut supposer que la somme de ses pondérations est 1, quitte à toutes les pondérations par leur total.

3 – Isobarycentre et médiane

Définition 8**Médianes d'un triangle et centre de gravité**

Soit ABC un triangle non aplati.

La *médiane issue de A* est le segment reliant A et le milieu de $[BC]$.

Les trois médianes issues des trois sommets de ABC sont concourantes, et leur point d'intersection est *centre de gravité* de ABC .

Théorème 9**Centre de gravité et barycentre**

Le centre de gravité d'un triangle est l'isobarycentre de ses sommets, et celui-ci est aux $\frac{2}{3}$ de la longueur de la médiane en partant du sommet.

Dans la suite, on propose un exercice guidé pour montrer cette propriété.

L'objectif est de montrer les deux propriétés suivantes :

- Le point G vérifie l'équation $\overrightarrow{GA} + \overrightarrow{GB} + \overrightarrow{GC} = \vec{0}$,
- Le point G est situé à $\frac{2}{3}$ de la longueur de chaque médiane à partir des sommets.

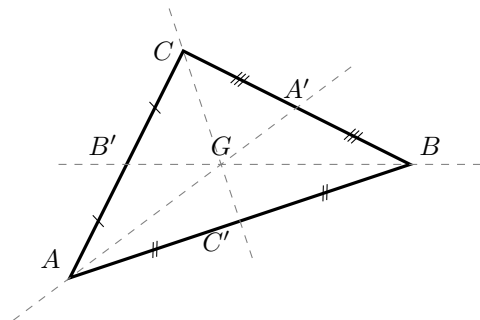
Soit M le point tel que $\overrightarrow{MA} + \overrightarrow{MB} + \overrightarrow{MC} = \vec{0}$. On admet qu'un tel point M existe et est unique. L'objectif de cette question est de montrer la première propriété : pour cela, nous allons montrer que $M = G$.

On admet que l'équation vérifiée par M le place à l'intérieur de ABC . On définit les points suivants :

- I l'intersection de (AM) et (BC) ,
- J l'intersection de (BM) et (AC) ,
- K l'intersection de (CM) et (AB) ,

On note enfin D le symétrique de A par rapport à M .

- 1► a► Faire un schéma soigné représentant la situation.
- b► À l'aide de la relation de Chasles et de la définition de D , exprimer $\overrightarrow{BD}$ uniquement en fonction de $\overrightarrow{MB}$ et $\overrightarrow{MA}$.



- c► En déduire que $\overrightarrow{BD} = \overrightarrow{MC}$.
 - d► Quelle est la nature de $BDCM$?
 - e► Justifier que G , I et D sont alignés. (*Indication : justifier que par la définition de chacun, ils appartiennent à la droite (AM) .*)
 - f► En déduire que I est le milieu de $[BC]$.
 - g► En déduire que (AM) est la médiane de ABC issue de A .
 - h► Expliquer en comment ce raisonnement se transpose pour montrer que (BM) et (CM) sont les deux autres médianes de ABC .
 - i► Conclure que $M = G$.
- 2► Montrons désormais la seconde propriété. On reprend le schéma de la question précédente avec ses notations. On rappelle que $BDCG$ est un parallélogramme de centre I .
- a► Justifier que $GD = 2GI$, et que $AG = GD$. En déduire que $GA = 2GI$.
 - b► En déduire que $GA = \frac{2}{3}GI$. Interpréter et conclure.
 - c► Expliquer comment cet raisonnement se transpose aux deux autres sommets.

4 – Coordonnées barycentriques

Dans tout ce paragraphe, A , B et C sont trois points non alignés du plan.

Lemme 10

Pour tout point M du plan, il existe un triplet de réels (a, b, c) tel que M est le barycentre du système pondéré de points $\{(A, a); (B, b); (C, c)\}$.

Démonstration. Comme A , B et C ne sont pas alignés, $(\overrightarrow{AB}; \overrightarrow{AC})$ est une base du plan. On peut donc considérer les coordonnées (b, c) du vecteur $\overrightarrow{AM}$ dans cette base, autrement dit $\overrightarrow{AM} = b\overrightarrow{AB} + c\overrightarrow{AC}$. À l'aide de la relation de Chasles, cela se réécrit, en posant $a = 1 - b - c$:

$$\overrightarrow{AM} = b(\overrightarrow{AM} + \overrightarrow{MB}) + c(\overrightarrow{AM} + \overrightarrow{MC}) \quad \text{donc} \quad a\overrightarrow{MA} + b\overrightarrow{MB} + c\overrightarrow{MC} = \vec{0}$$

Par ailleurs, $a + b + c = 1$ donc M est le barycentre du système pondéré de points $\{(A, a); (B, b); (C, c)\}$. □

Définition 11

Coordonnées barycentriques

Soit M un point du plan.

On appelle *coordonnées barycentriques* de M un triplet de réels (a, b, c) de somme non nulle, tel que M est le barycentre du système pondéré de points $\{(A, a); (B, b); (C, c)\}$.

On dit que (A, B, C) est un *repère barycentrique*. Pour ne pas confondre avec les coordonnées cartésiennes, on note $M(a : b : c)$.

Exemple 12

Notons G le centre de gravité de ABC , c'est-à-dire l'intersection de ses médianes, O le centre du cercle circonscrit à ABC , c'est-à-dire l'intersection de ses médiatrices, I le centre de son cercle inscrit, c'est-à-dire l'intersection de ses bissectrices, et, si ABC n'est pas rectangle, H son orthocentre, c'est-à-dire l'intersection de ses hauteurs.

On note a , b et c les longueurs BC , AC et AB .

- $G(1 : 1 : 1)$
- $O(\sin(\hat{A}) : \sin(\hat{B}) : \sin(\hat{C}))$
- $I(a : b : c)$
- $H(\tan(\hat{A}) : \tan(\hat{B}) : \tan(\hat{C}))$

Fun fact, pour terminer : si est M à l'intérieur de ABC , des coordonnées barycentriques de M sont les aires de ABM , ACM et BCM !

7 Limites de fonctions

Pour aller (plus) loin : branches infinies

De même que pour les suites, de même la définition d'une limite de fonction s'écrit mathématiquement de la manière suivante. Les différents cas qui ne sont pas écrits ici se rédigent de la même manière. Cette formalisation de la notion de limite a été inventée par Weierstrass, et est appelée *définition ε - δ de la limite* :

Définition 1

Définition ε - δ de la limite

- On dit f tend vers $+\infty$ en $+\infty$ lorsque :

$$\forall M > 0, \quad \exists x_0 \in \mathbb{R} : \quad \forall x \geq x_0, \quad f(x) > M$$

- On dit f tend vers $\ell \in \mathbb{R}$ en $a \in \mathbb{R}$ lorsque :

$$\forall \varepsilon > 0, \quad \exists \delta > 0 : \quad \forall x \in]a - \delta; a + \delta[, \quad |f(x) - \ell| < \varepsilon$$

Définition 2

Branche infinie

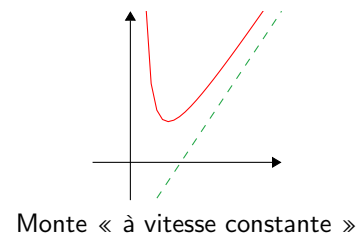
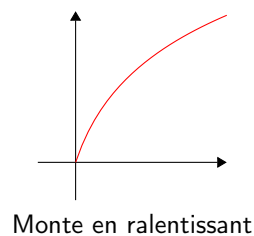
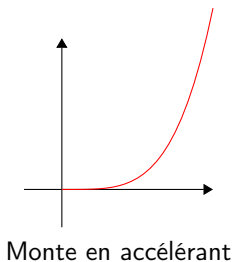
Une portion de courbe qui s'étend infiniment selon les abscisses ou selon les ordonnées possède ce que l'on appelle une *branche infinie*.

Exemple 3

Une asymptote est un cas particulier de branche infinie, qui est une droite : lorsqu'une courbe admet une asymptote horizontale ou verticale, cette asymptote est une branche infinie de la courbe, en $\pm\infty$ dans le cas d'une asymptote horizontale, et en $a \in \mathbb{R}$ dans le cas de l'asymptote verticale d'équation $x = a$.

Dans l'exemple précédent, on traite les branches infinies dans les cas de convergence en $\pm\infty$ ou de divergence en $a \in \mathbb{R}$. Dans le cas de la divergence vers l'infini en l'infini, les branches infinies sont dans certains cas appelées *branches paraboliques*, puisque leur forme rappelle celle d'une parabole. C'est l'équivalent dans l'idée d'une asymptote qui ne serait pas une droite.

Il y a trois types de divergence vers $+\infty$ en $+\infty$: soit la fonction monte de plus en plus vite, soit elle monte de plus en plus lentement, soit elle monte « à vitesse constante ».

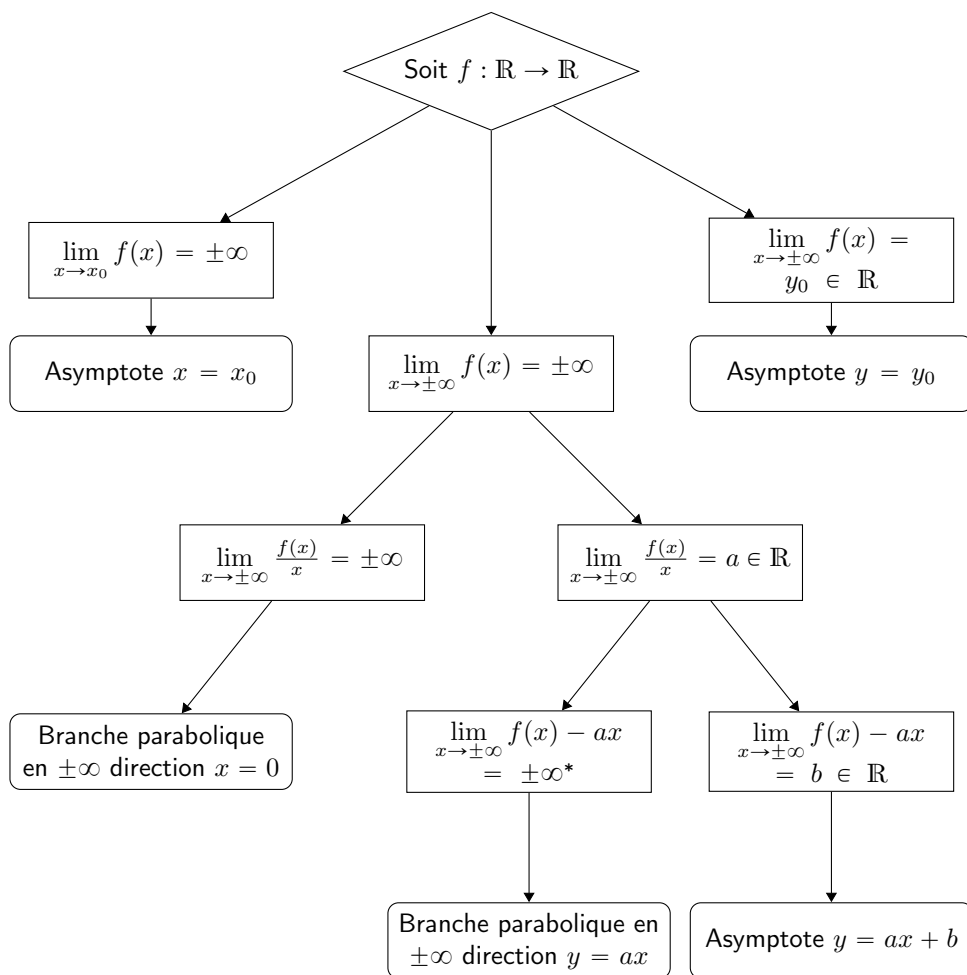


Branches paraboliques et asymptote oblique

Définition 4

- Dans le premier cas, la courbe semble devenir parallèle à l'axe des ordonnées : on dit qu'elle admet une *branche parabolique de direction l'axe des ordonnées* (abrégé Oy).
- Dans le deuxième cas, la courbe semble devenir parallèle à l'axe des abscisses : on dit qu'elle admet une *branche parabolique de direction l'axe des abscisses* (abrégé Ox).
- Dans le troisième cas, la courbe devient arbitrairement proche d'une droite oblique, que l'on appelle alors *asymptote oblique*, et donc la courbe admet une branche infinie qui est ladite asymptote.

Voici un plan d'étude complet des branches infinies, sous la forme d'un arbre de choix :



*. Note pour le graphe : si $a = 0$ alors C_f présente une branche parabolique de direction l'axe des abscisses, i.e. $y = 0$

8 Continuité

Table des matières

1	Résolution numérique d'équations	29
2	Étude d'une relation fonctionnelle	30

1 – Résolution numérique d'équations

Quelle que soit la méthode de résolution algorithmique, si l'on cherche à résoudre $f(x) = k$, on se ramène à $g(x) = 0$ pour $g : x \mapsto f(x) - k$. On fait également toujours l'hypothèse que $[a; b]$ est un encadrement valide d'une solution, déterminé au préalable grâce au théorème des valeurs intermédiaires.

Comme présenté dans la démonstration du théorème des valeurs intermédiaires, la *méthode par dichotomie* est une méthode qui permet d'encadrer de plus en plus finement une solution.

```

1 def dichotomie(f, a, b, seuil):
2     m = (a + b)/2
3     while b - a > seuil:
4         if f(m) == 0:
5             return m
6         elif f(a)*f(m) > 0:
7             a = m
8         else:
9             b = m
10        m = (a + b)/2
11    return m

```

La *méthode de la sécante*, ou *méthode de Lagrange*, fonctionne comme suit : on considère deux réels x_0, x_1 distincts (pas forcément un encadrement préalablement déterminé), et on construit (x_n) petit à petit. Une fois x_{n-1} et x_n construits, on nomme x_{n+1} l'intersection de la sécante à la courbe de f entre x_{n-1} et x_n et l'axe des abscisses.

La sécante entre x_{n+1} et x_n est la droite passant par les points de coordonnées :

$$(x_{n-1}; f(x_{n-1})) \quad \text{et} \quad (x_n; f(x_n))$$

donc elle a pour équation :

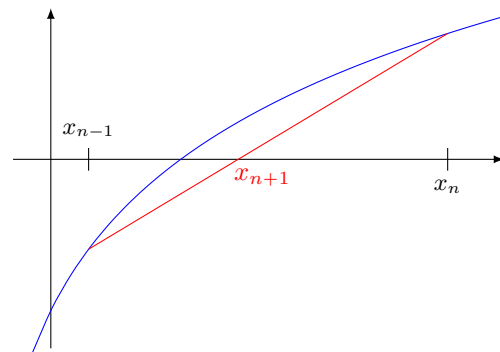
$$y = f(x_n) + \frac{f(x_n) - f(x_{n-1})}{x_n - x_{n-1}}(x - x_n)$$

La valeur de x_{n+1} est l'abscisse du point d'ordonnée 0 de cette droite, donc :

$$0 = f(x_n) + \frac{f(x_n) - f(x_{n-1})}{x_n - x_{n-1}}(x_{n+1} - x_n)$$

donc :

$$x_{n+1} = \frac{f(x_n)x_{n-1} - f(x_{n-1})x_n}{f(x_n) - f(x_{n-1})}$$



```

1 def secante(f, x0, x1, seuil, n):
2     x_old, x_new = x0, x1
3     # n représente un nombre maximal d'itérations autorisées
4     while n > 0 and x_old != x_new and abs(f(x_new)) > seuil:
5         f_old, f_new = f(x_old), f(x_new)
6         x_old, x_new = x_new, (f_new * x_old - f_old * x_new)/(f_new - f_old)
7         n = n - 1
8     return x_new

```

La *méthode de la tangente*, ou *méthode de Newton-Raphson*, nécessite quelques hypothèses supplémentaires. On suppose que f est dérivable, de dérivée continue, et que f' ne s'annule jamais sur $[a; b]$. On suppose également que $[a; b]$ n'est « pas trop grand ». Elle fonctionne comme suit : on prend $x_0 \in [a; b]$, et on construit petit à petit une suite (x_n) . Pour passer de x_n à x_{n+1} , on trace la tangente à la courbe de f en x_n , et x_{n+1} est l'abscisse de son intersection avec l'axe des abscisses.

L'équation de la tangente en x_n est :

$$y = f(x_n) + f'(x_n)(x - x_n)$$

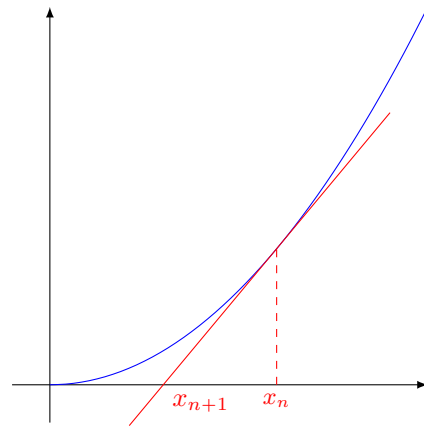
donc par définition de x_{n+1} , celui-ci vérifie :

$$0 = f(x_n) + f'(x_n)(x_{n+1} - x_n)$$

donc :

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$$

car $f'(x_n) \neq 0$ par hypothèse. Remarquez que l'on a besoin de la dérivée de f : dans le programme Python, celle-ci correspond à la fonction `df`.



```

1 def newton(f, df, x0, precision, n):
2     x = x0
3     # n représente un nombre maximal d'itérations autorisées
4     while n > 0 and abs(f(x)) > precision:
5         x = x - f(x)/df(x)
6         n = n - 1
7     return x

```

2 – Étude d'une relation fonctionnelle

Exercice 1

Déterminer les fonctions f continues qui vérifient :

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

Constatons d'abord que pour tout entier naturel n et pour tout réel x ,

$$f(nx) = f(\underbrace{x + x + \cdots + x}_n \text{ fois}) = \underbrace{f(x) + f(x) + \cdots + f(x)}_n \text{ fois} = nf(x)$$

(exercice : réécrire ce passage rigoureusement en menant avec soin une récurrence).

Ensuite, remarquons que pour tout réel x ,

$$f(0) = f(x + (-x)) = f(x) + f(-x) \quad \text{donc} \quad f(-x) = -f(x)$$

Il découle alors que pour tout entier $n \leq 0$, pour tout $x \in \mathbb{R}$, comme $-n \in \mathbb{N}$,

$$f(nx) = f(-(-nx)) = -f(-nx) = -(-nf(x)) = nf(x)$$

Soit $r \in \mathbb{Q}$. Il existe $p \in \mathbb{Z}$ et $q \in \mathbb{N}^*$ tels que $r = \frac{p}{q}$. Pour tout $x \in \mathbb{R}$,

$$qf(rx) = f(q \times rx) = f\left(q \frac{p}{q} x\right) = f(px) = pf(x)$$

et donc $f(rx) = \frac{p}{q}f(x) = rf(x)$.

Posons $a = f(1)$. Pour $x = 1$, cela montre que pour tout $r \in \mathbb{Q}$, $f(r) = f(r \times 1) = r \times f(1) = ar$.

Soit $x \in \mathbb{R}$. Soit (r_n) une suite de nombres rationnels qui converge vers x : par exemple, la suite de l'approximation décimale par défaut de x (par exemple pour π , les termes successifs seraient 3, puis 3,1, puis 3,14, puis 3,141, puis 3,1415, ...). C'est bien une suite de rationnels, car c'est une suite de décimaux. Par continuité de f sur $\mathbb{R}$, donc en x :

$$f(x) = f\left(\lim_{n \rightarrow +\infty} r_n\right) = \lim_{n \rightarrow +\infty} f(r_n) = \lim_{n \rightarrow +\infty} ar_n = a \lim_{n \rightarrow +\infty} r_n = ax$$

Ainsi, on a montré que seules les fonctions linéaires peuvent satisfaire les hypothèses initiales. On vérifie aisément que toute fonction linéaire les vérifie effectivement.

9 Loi binomiale

Table des matières

1	Problème de la surréservation	33
2	Loi de Poisson	33
3	Les cousines de la loi binomiale	34
A	Loi géométrique	34
B	Loi hypergéométrique	36
C	Loi binomiale négative	37
D	Loi multinomiale	39

1 – Problème de la surréservation

Un train possède 180 places. La compagnie qui s'occupe de la réservation des places sait qu'en moyenne, 8% des personnes ayant acheté un billet ne se présentent pas à l'embarquement. Plutôt que de vendre 180 places au risque que 8% d'entre elles soient vides et gâchées, la compagnie pratique la *surréservation* : elle vend plus de billets qu'il n'y a de places, en comptant sur le fait que la proportion attendue de passagers ne voyagera pas. Mais elle ne doit pas prendre le risque que trop de personnes se présentent réellement le jour de l'embarquement !

L'objectif est de déterminer le plus grand nombre de billets que la compagnie peut vendre en étant sûre à 95% que suffisamment de gens se désisteront et que le train ne sera pas surchargé.

On appelle n ce nombre de billets mis en vente, et on note X la variable aléatoire qui compte le nombre de passagers prenant effectivement le train. On va justifier que X suit une loi binomiale, et effectuer une recherche de seuil pour résoudre le problème.

Le fait que chaque passager prenne ou non le train est une épreuve de Bernoulli : on considère que le fait que le client honore sa réservation comme un succès, de probabilité 0,92. Les clients sont indépendants les uns des autres : cette épreuve se répète donc n fois de manière identique et indépendante. Il s'agit donc d'un schéma de Bernoulli de paramètres n et 0,92 dont X compte le nombre de succès, donc $X \sim \mathcal{B}(n; 0,92)$.

Dans la calculatrice, on rentre la fonction $x \mapsto \text{binomFrep}(x, 0,92, 180)$, et on en affiche la table : elle nous donne les valeurs de $\mathbb{P}(X \leq 180)$. On cherche le plus grand n pour lequel $\mathbb{P}(X \leq 180) > 0,95$: on trouve $n = 189$.

2 – Loi de Poisson

Soit X qui suit une loi binomiale $\mathcal{B}(n, p)$ avec n grand. Posons $\lambda = np$ (c'est l'espérance de X). En particulier, $p = \frac{\lambda}{n}$. Pour tout entier k entre 0 et n ,

$$\mathbb{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k} = \frac{n!}{k!(n-k)!} \times \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^n \left(1 - \frac{\lambda}{n}\right)^{-k}$$

Remarquons plusieurs choses :

- Comme n est grand, $\frac{n!}{(n-k)!n^k} \approx 1$. En effet,

$$\frac{n \times (n-1) \times \cdots \times (n-k+1)}{n^k} = \frac{n}{n} \times \frac{n-1}{n} \times \cdots \times \frac{n-k+1}{n} \xrightarrow{n \rightarrow +\infty} 1 \times 1 \times \cdots \times 1 = 1$$

et si un quotient tend vers 1, c'est que son numérateur et son dénominateur se ressemblent de plus en plus avec le temps.

- Comme n est grand, $\frac{\lambda}{n} \approx 0$, donc $\left(1 - \frac{\lambda}{n}\right)^{-k} \approx 1^{-k} = 1$.
- $\left(1 - \frac{\lambda}{n}\right)^n \xrightarrow{n \rightarrow +\infty} e^{-\lambda}$: c'est une limite extrêmement classique, souvent vue dans le cours sur le logarithme népérien.

En combinant toutes ces observations,

$$\mathbb{P}(X = k) = \frac{\lambda^k}{k!} \times \frac{n!}{(n-k)!n^k} \times \left(1 - \frac{\lambda}{n}\right)^n \left(1 - \frac{\lambda}{n}\right)^{-k} \approx \frac{\lambda^k}{k!} \times e^{-\lambda} \times 1 = e^{-\lambda} \frac{\lambda^k}{k!}$$

Ainsi, lorsque n devient grand, la probabilité d'obtenir k succès ne semble plus dépendre de n (qui est caché dans λ). La loi de probabilité décrite par l'expression précédente est appelée *loi de Poisson* :

Définition 1

Loi de Poisson

Une variable aléatoire X suit la *loi de Poisson* de paramètre λ , notée $\mathcal{P}(\lambda)$, lorsque :

$$\forall k \in \mathbb{N}, \quad \mathbb{P}(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}$$

Remarque 2

La loi de Poisson est plus difficile à repérer dans la nature, ses conditions d'apparition sont assez particulières. Le plus souvent, la loi de Poisson est utilisée comme approximation de la loi binomiale pour un grand nombre de répétitions – ce que nous avons fait ci-dessus.

Voici une situation plus concrète d'application de la loi de Poisson : au guichet de la Poste, les clients s'enchaînent, et le temps entre chaque client est aléatoire (selon une loi qui ressemble à la loi géométrique). Dans ces conditions, le nombre de clients qui passent au guichet sur une fenêtre d'une heure suit une loi de Poisson.

3 – Les cousines de la loi binomiale

On rappelle que la loi binomiale est celle qui modélise le nombre de succès qui surviennent lorsqu'on répète de manière identique et indépendante un certain nombre de fois une même épreuve de Bernoulli. Il y a 4 hypothèses : compter le nombre de succès, répéter de manière identique, de manière indépendante, et le fait qu'on répète des épreuves de Bernoulli. Que se passe-t-il si on se débarrasse de certaines hypothèses ? La réponse est : on obtient de nouvelles lois !

A – Loi géométrique

Dans un jeu de combat avec des créatures magiques, le joueur peut lancer une pièce pour augmenter les dégâts de ses attaques : il recommence jusqu'à tomber sur Face, et le nombre de Pile qu'il a obtenu avant cela détermine la puissance de son attaque.

On peut alors se demander : quelle est la probabilité qu'il fasse 1 Pile avant d'échouer ? 2 Piles, 3 Piles, ... ?

Un lancer de pièce est une épreuve de Bernoulli : pour des raisons qui vont s'éclaircir plus tard, on désignera comme succès le fait de tomber sur Face. Pour garder en généralité, on supposera que la pièce est déséquilibrée, et que la probabilité d'obtenir Face est p . On peut ainsi commencer à quantifier les probabilités :

- ▶ La probabilité d'obtenir Face au 1^{er} lancer est p ,
- ▶ La probabilité d'obtenir Face au 2^e lancer est $p(1 - p)$ (on obtient Pile, puis Face),
- ▶ La probabilité d'obtenir Face au 3^e lancer est $p(1 - p)^2$ (on obtient Pile, puis Pile, puis Face),
- ▶ etc.

Une logique semble se dégager. La loi qui apparaît est la *loi géométrique* :

Définition 3

Loi géométrique

Soit X la variable aléatoire qui mesure le rang du premier succès dans un schéma de Bernoulli avec un infinité potentielle d'épreuves, de probabilité de succès p .
La loi de X est la *loi géométrique de paramètre p* , notée $\mathcal{G}(p)$.
Pour tout $k \in \mathbb{N}^*$, $\mathbb{P}(X = k) = p(1 - p)^{k-1}$.

Propriété 4

Absence de mémoire

La loi géométrique est *sans mémoire*, c'est-à-dire que si X suit une loi géométrique, alors :

$$\forall n \in \mathbb{N}, \quad \forall k \in \mathbb{N}, \quad \mathbb{P}_{X > k}(X > n + k) = \mathbb{P}(X > n)$$

Remarque 5

Imaginons que l'on dispose d'une ampoule qui 1% de chance de griller chaque heure. On peut modéliser cette situation par un schéma de Bernoulli de probabilité de succès 1%, et on s'intéresse au rang du premier succès, c'est-à-dire au bout de combien de temps l'ampoule va griller.

Notons T la variable aléatoire qui donne l'instant où l'ampoule grille. Alors $T \sim \mathcal{G}(0,01)$. Comme la loi géométrique est sans mémoire, on peut écrire $\mathbb{P}_{T > 100}(T > 110) = \mathbb{P}(T > 10)$.

Qu'est-ce que cela signifie ? Le premier membre désigne la probabilité que l'ampoule dure au moins 110 heures, sachant qu'elle a déjà fonctionné 100 heures. Autrement dit, sachant qu'elle fonctionne depuis déjà 100 heures, la probabilité qu'elle dure 10 heures de plus. Le second membre est la probabilité que l'ampoule dure au moins 10 heures depuis le début. L'absence de mémoire dit que ces deux probabilités sont les mêmes : qu'elle ait déjà vécu 1 heure, 100 heures ou 1000 heures, la probabilité qu'elle résiste aux 10 prochaines heures est la même ! (Cela paraît surprenant : peut être que l'hypothèse initiale n'est pas tout à fait juste... qui sait !)

Application 6

Démontrons la propriété d'absence de mémoire de la loi géométrique. Soit $X \sim \mathcal{G}(p)$.

Préliminaires :

- 1▶ Soit $k \in \mathbb{N}^*$. Décrire avec une phrase l'événement $\{X > k\}$.
- 2▶ Déterminer $\mathbb{P}(X > k)$ en fonction de p .

Soient m et n deux entiers naturels non nuls, on note $A = \{X > m + n\}$ et $B = \{X > m\}$ deux événements.

- 3▶ Écrire la définition de $\mathbb{P}_B(A)$.
- 4▶ Interpréter l'événement $A \cap B$ avec une phrase. En déduire que $A \cap B = \{X > m + n\}$.
- 5▶ En déduire l'expression de $\mathbb{P}_B(A)$ en fonction de p .
- 6▶ En déduire que $\mathbb{P}_B(A) = \mathbb{P}(X > n)$.
- 7▶ Conclure.

Caractérisation par l'absence de mémoire

Proposition 7

Soit X une variable aléatoire qui ne prend que des valeurs entières strictement positives.
Si X vérifie la propriété d'absence de mémoire, alors X suit une loi géométrique.

Application 8

Démontrons cette caractérisation. Soit X une variable aléatoire vérifiant les hypothèses de la proposition.

Pour tout $k \in \mathbb{N}^*$, on note A_k l'événement $\{X > k\}$ et $p_k = \mathbb{P}(A_k)$.

- 1► Montrer que $\mathbb{P}_{A_k}(A_{k+1}) = q$ où $q = \mathbb{P}(X > 1)$.
(La méthodologie de l'exercice précédent peut aider à répondre à cette question.)
- 2► En déduire que la suite (p_k) est géométrique, dont on précise la raison et le premier terme.
- 3► En déduire l'expression explicite de p_k en fonction de k .
- 4► Justifier avec une phrase que $\{X > k - 1\} = \{X = k\} \cup \{X > k\}$.
- 5► En déduire que $\mathbb{P}(X = k) = p_{k-1} - p_k$.
- 6► En déduire l'expression de $\mathbb{P}(X = k)$ en fonction de p et k , puis conclure.

B – Loi hypergéométrique

Vous allez au cinéma avec vos amis, et vous achetez un seau de pop-corn qui mélange du pop-corn salé et du sucré. Vous commencez à parier sur le goût du prochain grain que vous allez manger. Vous réalisez : il s'agit d'une expérience aléatoire ! Et qui dit expérience aléatoire dit modélisation par une loi ! Quelle est donc la loi que suit le goût du prochain grain de pop-corn ?

Considérons un petit seau de pop-corn. Après quelques recherches sur Internet, on trouve que la taille standard d'un petit seau est de 1 L, et qu'un grain de pop-corn a presque la forme d'une boule de diamètre 2 cm, soit un rayon de 0,1 dm. Le volume d'un grain de pop-corn est donc $\frac{4}{3}\pi \times 0,1^3 \approx 0,004 \text{ dm}^3$. Et comme 1 L correspond à 1 dm^3 , le volume d'un grain de pop-corn est environ 0,004 L, ou plus exactement $\frac{\pi}{750}$ litres. Ainsi, le seau contient $\frac{1}{\pi/750} \approx 239$ grains. Enfin... ça serait vrai si les grains de pop-corn occupaient l'intégralité de l'espace du seau. Si vous êtes déjà passé par un étale d'oranges ou de pommes, vous aurez remarqué que, lorsqu'on empile des boules, il existe un espace entre chacune. Il doit bien y avoir une manière de calculer cet espace, non ? Johannes Kepler s'était posé la question en 1611 : des sphères empilées de la manière la plus serrée possible occupent environ 74% du volume. Ainsi, on peut estimer le nombre de grains comme $\frac{0,74 \times 1}{\pi/750} \approx 177$. Mais nous digressons. Par souci de simplification pour la suite, on estimera que le vendeur a été généreux, et qu'il vous a vendu 200 grains, avec une proportion p de grains salés, et le reste de grains sucrés, qu'on supposera parfaitement mélangés dans le seau.

Nous voilà donc au début du film, on pioche des grains de pop-corn un par un. Le goût du grain relève d'une épreuve de Bernoulli, car il n'y a que deux issues, soit salé, soit sucré. On répète cette épreuve un certain nombre n de fois. Comme le seau contient un grand nombre de grains, on peut assimiler l'expérience à un tirage avec remise : on répète donc n fois de manière identique et indépendante une épreuve de Bernoulli. Il s'agit d'un schéma de Bernoulli, et la variable aléatoire X qui compte le nombre de grains salés que l'on pioche suit une loi binomiale de paramètres n et p .

Jouons un peu avec notre modèle : on pioche une poignée de 10 grains dans le seau. On suppose que le vendeur a mis autant de grains salés que sucrés, soit $p = 0,5$. Quelle est la probabilité que cette poignée de grains ne contienne qu'un seul goût ? C'est-à-dire, soit que des grains salés, soit que des grains sucrés ? Dans le premier cas, on a $X = 10$ grains salés, il s'agit donc de calculer $P(X = 10) = \binom{10}{10} 0,5^{10}(1 - 0,5)^0 = 0,5^{10}$. Dans le second, on a 10 grains sucrés, soit $X = 0$ grains salés, il s'agit donc de calculer $P(X = 0) = \binom{10}{0} 0,5^0(1 - 0,5)^{10} = 0,5^{10}$. Ainsi, la

probabilité d'une poignée au goût unique est $P(X = 0) + P(X = 10) = 0,5^{10} + 0,5^{10} \approx 0,2\%$; très improbable, donc.

On arrive à la fin du film, il ne reste plus que 40 grains de pop-corn. Ici, il n'est plus raisonnable de considérer que le tirage se fait avec remise : 10 grains dans un seau de 200 ne représentent pas la même proportion que 10 grains parmi 40. On se retrouve donc dans une situation proche de la loi binomiale, mais on ne peut plus considérer le tirage comme étant avec remise. La loi de probabilité correspondant à un tirage sans remise s'appelle la loi hypergéométrique.

Dans cette situation : notons N_1 le nombre de grains salés, et N_2 le nombre de grains sucrés. On a donc $N_1 + N_2 = 40$ grains au total. Si l'on pioche 10 grains, la probabilité que k d'entre eux soient salés est donné par la formule :

$$\frac{\binom{N_1}{k} \times \binom{N_2}{10-k}}{\binom{40}{10}}.$$

Jouons avec ce nouveau modèle en nous posant la même question que précédemment : quelle est la probabilité que la poignée ne contienne qu'un unique goût ? On part de nouveau du principe qu'il reste autant de grains salés que sucrés, donc $N_1 = N_2 = 20$. S'il n'y a que des grains salés dans la poignée, alors $k = 10$, et en appliquant la formule, on obtient $\approx 0,02\%$. La situation est symétrique entre le salé et le sucré, donc l'autre cas a la même probabilité. Finalement, la probabilité de n'avoir qu'un goût est désormais 0,04% environ.

Loi hypergéométrique

Définition 9

On considère une urne contenant N_1 boules rouges et N_2 boules bleues, indiscernables au toucher. On pioche n boules simultanément, ou successivement, sans remise et sans compter l'ordre. Soit X la variable aléatoire comptant le nombre de boules rouges que l'on a piochées.

La variable aléatoire X suit la *loi hypergéométrique* de paramètres N_1 , N_2 et n , notée $\mathcal{HG}(N_1, N_2, n)$, et pour tout k entier entre 0 et n :

$$\mathbb{P}(X = k) = \frac{\binom{N_1}{k} \binom{N_2}{n-k}}{\binom{N_1 + N_2}{n}}$$

Démonstration. Lorsqu'on tire simultanément n boules parmi $N_1 + N_2$ boules, il ne peut pas y avoir de répétition et l'ordre ne compte pas : il n'agit d'une n -combinaison, et il y a $\binom{N_1 + N_2}{n}$ possibilités. Tous les tirages sont équiprobables, il s'agit donc de dénombrer les tirages favorables.

Pour tirer un lot de n boules contenant exactement k boules rouges, on peut tirer les k boules rouges parmi les N_1 candidates (il y a $\binom{N_1}{k}$ possibilités), et $n - k$ boules rouges parmi les N_2 candidates (il y a $\binom{N_2}{n-k}$ possibilités).

Il y a donc $\binom{N_1}{k} \binom{N_2}{n-k}$ tirages favorables. $\square$

C – Loi binomiale négative

Jules César est en guerre dans le Sud de la Gaule. Il est sollicité de toutes parts par le commandement de son armée, et n'a pas un instant pour lui. Il décide donc de se faire livrer du raisin de Pouilles, qu'il aime tant. Pompée, son meilleur ennemi alors stationné dans le Sud de l'Italie, saute sur l'occasion : il joint à la cargaison un infiltré chargé d'empoisonner le raisin de César. Pompée lui donne les consignes suivantes : « chaque jour, César mangera une grappe de raisin. Tu empoisonneras un grain de chaque grappe. » Chaque grappe porte 100 grains. Malheureusement pour l'infiltré, César n'a jamais suffisamment de temps pour se reposer et savourer pleinement son raisin, aussi il ne

mange en tout et pour tout que 20 grains sur les 100. Combien de jours l'infiltré prendra-t-il pour être sûr à 95% de réussir à empoisonner César ?

César mange 20 grains d'une grappe de 100 grains dont 1 est empoisonné. On serait tenter de dire qu'il répète 20 fois une épreuve de Bernoulli avec une probabilité de succès de 1%, de manière identique et indépendante, et donc d'utiliser la loi binomiale pour dire que la probabilité que César ne mange aucun grain empoisonné est de $P(X = 0) = \binom{20}{0} \times 0,01^0 \times 0,99^{20}$, et donc par événement contraire, $P(X \geq 1) = 1 - P(X = 0) \approx 18\%$ (ici X désigne le nombre de grains empoisonnés consommés par César).

Cependant, les événements ne sont pas identiques et indépendants, ou, pourrait-on dire, le tirage n'est pas sans remise. En effet, César mange les grains au fur et à mesure, et donc plus il mange de grains non empoisonnés, plus la probabilité que le prochain soit empoisonné augmente. On peut en fait faire le calcul à la main : calculons la probabilité que César ne mange aucun grain empoisonné. Au 1er grain, il y a 99 bons grains sur 100. Après l'avoir mangé, il reste 98 bons grains sur 99, et ainsi de suite jusqu'au vingtième. On multiplie toutes ces probabilités :

$$P(X = 0) = \underbrace{\frac{99}{100}}_{\text{1er grain}} \times \underbrace{\frac{98}{99}}_{\text{2e grain}} \times \cdots \times \underbrace{\frac{80}{81}}_{\text{20e grain}}$$

et par événement contraire, la probabilité que César mange le grain empoisonné est $1 - P(X = 0) = 20\%$.

On constate que les résultats sont proches. La loi binomiale approche la situation lorsque le nombre de grains mangés est suffisamment petit face au nombre total de grains. La loi binomiale a l'avantage de fournir une formule simple à adapter si l'on veut changer la situation, en changeant le nombre de grains par grappe, le nombre de grains mangés ou le nombre de grains empoisonnés. La contrepartie est une erreur de quelques pourcents sur la probabilité finale.

Pour éviter d'être suspecté d'empoisonnement, l'infiltré utilise un poison lent et indétectable. En contrepartie, il faut avaler au total 3 doses pour que le poison soit mortel. C'est un schéma de Bernoulli, avec une répétition identique et indépendante d'une même épreuve, mais il ne s'agit pas de compter le nombre de succès dans un nombre de répétitions fixé. Au contraire, il s'agit de trouver le nombre de répétitions nécessaires pour réaliser un nombre de succès fixé. À ce titre, ce modèle porte le nom de loi binomiale négative. On note Y le nombre de jours nécessaires pour atteindre le 3e succès, c'est-à-dire le jour où César avale la troisième dose. Soit $k \geq 3$. Pour calculer la probabilité que $Y = k$, on peut raisonner de la manière suivante : si le troisième succès arrive le k -ième jour, c'est qu'il y a eu exactement 2 succès durant les $k - 1$ jours précédents. Si on note X la variable aléatoire comptant le nombre de succès durant les $k - 1$ premiers jours, alors X suit la loi binomiale de paramètres $k - 1$ et 20%, et donc $P(X = 2) = \binom{k-1}{2} \times 0,2^2 \times 0,8^{k-1-2}$. La probabilité qu'un nouveau succès arrive le k -ième jour est de 20%, par indépendance (on est dans un schéma de Bernoulli) on multiplie donc des deux probabilités pour trouver que $P(Y = k) = P(X = k - 1) \times 0,2 = \binom{k-1}{2} \times 0,2^3 \times 0,8^{k-3}$.

Pour répondre à la question, on s'intéresse à $P(Y \leq n) = P(Y = 1) + P(Y = 2) + \cdots + P(Y = n)$, que l'on veut supérieure à 95%. On peut calculer le tableau de valeurs à l'aide d'une calculatrice, et constater qu'à partir de $n = 27$ jours, la probabilité d'assassiner César dépasse les 95%.

Définition 10

Loi binomiale négative

On considère un schéma de Bernoulli de probabilité de succès p et dont le nombre de répétition est indéterminé et arbitrairement grand. Soit $n \geq 0$, soit X la variable aléatoire donnant le rang du n -ième succès.

La loi de X est la *loi binomiale négative de paramètres n et p* , notée $\mathcal{BN}(n, p)$.

Pour tout $k \geq n$,

$$P(X = k) = \binom{k-1}{n-1} p^n (1-p)^{k-n}$$

Démonstration. Avec les notations de la définition : si le n -ième succès survient à la k -ième épreuve, c'est que la k -ième épreuve est un succès (probabilité p) et que les $(k - 1)$ -ièmes précédentes épreuves contiennent exactement

$n - 1$ succès. On est dans un schéma de Bernoulli, et si l'on se restreint aux $k - 1$ premières épreuves, alors ses paramètres sont $k - 1$ et p . Or la variable aléatoire Y qui compte le nombre de succès dans ce schéma de Bernoulli suit une loi binomiale, de paramètres $k - 1$ et p . Ainsi, la probabilité de rencontrer exactement $n - 1$ succès est $\mathbb{P}(Y = n - 1) = \binom{k-1}{n-1} p^{n-1} (1-p)^{(k-1)-(n-1)}$, et donc, après simplification et produit par p , on a bien :

$$\mathbb{P}(X = k) = \binom{k-1}{n-1} p^n (1-p)^{k-n}$$

□

D – Loi multinomiale

Alice et Bob s'en vont pêcher. Ils ont pour objectif de repartir avec 10 poissons chacun. Dans l'étendue d'eau où ils ont choisi de lancer leur canne, il y a des saumons, des thons, et des morues. On estime que la population de chaque espèce est suffisamment vaste pour que la pêche de 20 individus n'impacte pas significativement leurs effectifs. Dans l'étendue d'eau, il y a 60% de morues, 35% de saumons et 5% de thons. Alice repart avec 8 morues et 2 thons, alors que Bob repart avec 6 morues et 3 saumons et 1 thon. Qui a eu la pêche la plus chanceuse ?

Alice aurait pu pêcher d'abord 8 morues puis 2 thons, ou alors 5 morues, un premier thon, 3 morues puis le second thon, etc. On sent bien qu'il va nous falloir compter le nombre d'ordre possibles pour la pêche d'Alice. En remarquant que l'ordre de pêche d'Alice est déterminé par le rang des deux thons (le reste étant de la morue), on s'en sort un coefficient binomial : il y a 2 positions à choisir parmi les 8, donc $\binom{8}{2} = 28$.

Pour Bob, c'est plus compliqué : en effet, déterminer le rang du thon ne suffit pas, car il faut encore savoir où vont les saumons et les morues dans les places restantes. Il faut donc une autre approche : considérons que les poissons ont tous un petit nom, et que pêcher Nemo avant Dory n'est pas la même chose que de pêcher Dory avant Nemo. Ainsi, choisir l'ordre des poissons revient à choisir une permutation de 10 éléments : il y a $10!$ possibilités. Bon, trêve de fiction, tous les saumons sont les mêmes... On a donc compté des permutations en trop. Comme il y a 6 morues, il y a $6!$ de les permuter dans les emplacement précédemment choisis. On a donc compté $6!$ fois trop d'ordres. De même pour les saumons, on a compté $3!$ fois trop, et pour le thon, il n'y a pas d'ordres compté en trop puisqu'il n'y en a qu'un, mais pour des raisons qui apparaîtront plus loin, on dira qu'on compte $1!$ fois trop. Ainsi, il y a en réalité $\frac{10!}{6!3!1!} = 840$ ordres de pêche pour Bob.

On peut raisonnablement supposer la pêche d'un poisson indépendante des précédentes. Il y a donc une probabilité $0,6^8$ de pêcher 8 morues, et une probabilité $0,05^2$ de pêcher 2 thons. Ainsi, la pêche d'Alice avait une probabilité $28 \times 0,6^8 \times 0,05^2 \approx 0,0012$. De même, pour Bob, la probabilité est $840 \times 0,6^6 \times 0,35^3 \times 0,05 \approx 0,0840$. C'est donc Alice qui a été la plus chanceuse !

Loi multinomiale

Définition 11

Dans une urne, il y a une proportion p_1 de boules avec le chiffre 1, p_2 avec le chiffre 2, et p_3 avec le chiffre 3. On tire n boules successivement et avec remise ; on note N_i la variable aléatoire qui compte le nombre de boules avec le chiffre i . Pour tous k_1, k_2 et k_3 entiers naturels tels que $k_1 + k_2 + k_3 = n$, on a :

$$\mathbb{P}(\{N_1 = k_1\} \cap \{N_2 = k_2\} \cap \{N_3 = k_3\}) = \frac{n!}{k_1!k_2!k_3!} p_1^{k_1} p_2^{k_2} p_3^{k_3}$$

Démonstration. La probabilité d'obtenir 1 boule numérotée 1 est p_1 , donc la probabilité d'en obtenir k_1 est $p_1^{k_1}$ par indépendance. De même pour les boules 2 et 3. De même que la loi binomiale, il s'agit de compter le nombre de tirages favorables.

On ordonne artificiellement les tirages : il y a alors $n!$ façons de les permuter. Mais cet ordre est artificielle : on peut interchanger les boules 1 comme on veut, et il y a $k_1!$ manières de les permuter. On a donc compté $k_1!$ permutations

en trop, il faut donc diviser $n!$ par $k_1!$. De même pour les boules 2 et 3, il faut donc diviser par $k_2!$ et $k_3!$. Ainsi, il y a $\frac{n!}{k_1!k_2!k_3!}$ tirages favorables, et donc la probabilité d'avoir k_1 boules 1, k_2 boules 2 et k_3 boules 3 est bien $\frac{n!}{k_1!k_2!k_3!}p_1^{k_1}p_2^{k_2}p_3^{k_3}$. $\square$

Remarque 12

La formule s'étend naturellement à une situation avec plus de trois issues.

Remarque 13

Pour une situation à 2 issues, la loi multinomiale correspond à la loi binomiale : en effet, $k_1 + k_2 = n$ donc $k_2 = n - k_1$. De même, $p_1 + p_2 = 1$ donc $p_2 = 1 - p_1$. La formule devient alors :

$$\frac{n!}{k_1!k_2!}p_1^{k_1}p_2^{k_2} = \frac{n!}{k_1!(n-k_1)!}p_1^{k_1}(1-p_1)^{n-k_1} = \binom{n}{k_1}p_1^{k_1}(1-p_1)^{n-k_1}$$

10 Logarithme

Table des matières

1	La vraie définition de la notation puissance	41
2	Une célèbre limite	42
3	Échelles logarithmiques	42
4	Histoire du logarithme	42

1 – La vraie définition de la notation puissance

En quatrième, on définit la notation puissance comme suit : pour tout $x \in \mathbb{R}$, pour tout $n \in \mathbb{N}$,

$$x^n = \underbrace{x \times x \times \cdots \times x}_{n \text{ fois le facteur } x}$$

et si $x \neq 0$, $x^{-n} = \frac{1}{x^n}$. Mais jamais l'on ne définit une puissance non entière : que signifie $2^{1,3}$? Ça n'a pas de sens de dire « 1,3 fois le facteur 2 »...

Regardons du côté du logarithme : on sait que pour tout $x > 0$ et pour tout $n \in \mathbb{N}$, $\ln(x^n) = n \ln(x)$. En appliquant l'exponentielle de part et d'autre de cette égalité, on obtient :

$$x^n = e^{\ln(x^n)} = e^{n \ln(x)}$$

Dans l'expression $e^{n \ln(x)}$, le fait que n soit entier n'a aucune importance, car l'exponentielle est définie pour tout nombre réel. Partons de cette idée : on *définit* (et c'est bien là une *définition*) pour tout $x > 0$ et tout $\alpha \in \mathbb{R}$,

$$x^\alpha = e^{\alpha \ln(x)}$$

On peut vérifier, grâce aux relations fonctionnelles du logarithme et de l'exponentielle, que cette définition prolonge les règles de calcul sur les puissances :

$$\begin{array}{llll} \blacktriangleright x^{\alpha+\beta} = x^\alpha x^\beta & \blacktriangleright x^{\alpha-\beta} = \frac{x^\alpha}{x^\beta} & \blacktriangleright (x \times y)^\alpha = x^\alpha y^\alpha & \blacktriangleright (x^\alpha)^\beta = x^{\alpha\beta} \end{array}$$

On peut donc manipuler les puissances quelconques comme des puissances entières. Les puissances non entières les plus fréquemment rencontrées sont les racines : on a vu plus haut que $\ln(\sqrt{x}) = \frac{1}{2} \ln(x)$, donc $\sqrt{x} = e^{\frac{1}{2} \ln(x)} = x^{\frac{1}{2}}$. De la même manière, $\sqrt[3]{x} = x^{\frac{1}{3}}$ ou, plus généralement, pour tout $n \geq 2$, $\sqrt[n]{x} = x^{\frac{1}{n}}$.

Voici un exemple concret : la troisième loi de Kepler stipule que le carré de la période orbitale est proportionnel au cube du rayon de l'orbite (si celle-ci est circulaire), ce qu'on écrit souvent $\frac{T^2}{r^3} = k$ avec k une constante dépendant

de la situation. Dans cette situation, si l'on connaît T , alors on peut retrouver r en calculant $\left(\frac{T^2}{k}\right)^{1/3}$.

2 – Une célèbre limite

Parmi les formes indéterminées de limites, il y a un cas qui n'est pas répertorié dans les cours de terminale : la situation du type « $1^{+\infty}$ » (on rappelle que c'est un abus de notation...) Par exemple, soit $x \in \mathbb{R}$, pour $n \geq 1$ posons $u_n = \left(1 + \frac{x}{n}\right)^n$. Il est immédiat que, lorsque n tend vers $+\infty$, le terme à l'intérieur des parenthèses tend vers 1, et l'exposant tend vers $+\infty$. On serait tenté de penser qu'une telle limite vaut forcément 1, qu'il n'y a pas d'indétermination... prouvons le contraire en calculant la limite de (u_n) .

Supposons x assez grand, de sorte que $1 + \frac{x}{n} > 0$. Le cas échéant, on peut écrire $\ln(u_n) = n \ln\left(1 + \frac{x}{n}\right)$. On ne peut pas calculer la limite de $(\ln(u_n))$ immédiatement, car c'est une forme indéterminée.

Faisons un petit crochet pour démontrer la limite suivante :

$$\lim_{x \rightarrow 0} \frac{\ln(1+x)}{x} = 1$$

Pour cela, il suffit de remarquer que $f : x \mapsto \ln(1+x)$ est dérivable en 0, de nombre dérivé $f'(0) = \frac{1}{1+0} = 1$ (dérivation du type $\ln(u)$). Or par définition,

$$f'(0) = \lim_{x \rightarrow 0} \frac{f(x) - f(0)}{x - 0} = \lim_{x \rightarrow 0} \frac{\ln(1+x) - \ln(1+0)}{x} = \lim_{x \rightarrow 0} \frac{\ln(1+x)}{x}$$

d'où la valeur de la limite.

Revenons à notre problème initial. Comme $\lim_{n \rightarrow +\infty} \frac{x}{n} = 0$, on pose $y = \frac{x}{n}$ et on se ramène à la limite du crochet ci-dessus en effectuant le changement de variable suivant :

$$\lim_{n \rightarrow +\infty} n \ln\left(1 + \frac{x}{n}\right) = \lim_{n \rightarrow +\infty} n \times \frac{x}{n} \times \frac{\ln\left(1 + \frac{x}{n}\right)}{\frac{x}{n}} = \lim_{y \rightarrow 0} x \times \frac{\ln(1+y)}{y} = x \times 1 = x$$

Ainsi, $\lim_{n \rightarrow +\infty} \ln(u_n) = x$. Comme la fonction exponentielle est continue en x ,

$$\lim_{n \rightarrow +\infty} \left(1 + \frac{x}{n}\right)^n = \lim_{n \rightarrow +\infty} u_n = \lim_{n \rightarrow +\infty} e^{\ln(u_n)} = e^x$$

3 – Échelles logarithmiques

Parler de Benford

4 – Histoire du logarithme

Sources : [The Invention That Saved Science](#) (Ben Syversen) et [The Most Useful Curve in Mathematics \[Logarithms\]](#) (Welch Labs).

11 Orthogonalité dans l'espace

Table des matières

1	Produit vectoriel et déterminant	43
A	L'idée du déterminant dans le plan, transposée à l'espace	43
B	Produit vectoriel de deux vecteurs	44
2	Autour des sphères	46
A	Rappel : cercles du plan	46
B	Équations de sphères	47
C	Sections de sphères	47
3	Pour la culture : mais à quoi ça sert ? !	49
4	Bonus : démonstration des propriétés du produit scalaire plan	51

1 – Produit vectoriel et déterminant

Dans cette section, on se place dans une base orthonormée du plan ou de l'espace, que l'on notera $\mathcal{B}$ dans tous les cas.

A – L'idée du déterminant dans le plan, transposée à l'espace

Deux vecteurs sont colinéaires si, et seulement si leurs coordonnées sont proportionnelles. Deux jeux de données sont proportionnels si, et seulement si toutes les paires possibles de produits en croix sont égales. Soient $\vec{u}_1 \begin{pmatrix} x_1 \\ y_1 \end{pmatrix}$ et $\vec{u}_2 \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}$ deux vecteurs du plan. Il n'y a qu'une seule paire de produits en croix possible : ainsi, les coordonnées de $\vec{u}_1$ et $\vec{u}_2$ sont proportionnelles si, et seulement si $x_1 y_2 = x_2 y_1$, ce qu'on peut réécrire $x_1 y_2 - x_2 y_1 = 0$. L'intérêt de cette réécriture est de faire apparaître un critère générique, une valeur de référence indépendante du choix des vecteurs : si on note $\det_{\mathcal{B}}(\vec{u}_1, \vec{u}_2) = x_1 y_2 - x_2 y_1$, ce qu'on appelle le *déterminant du couple de vecteurs* $(\vec{u}_1, \vec{u}_2)$, alors la démonstration que l'on vient de mener permet de dire que deux vecteurs sont colinéaires si, et seulement si leur déterminant est nul. Les mathématiciens ont inventé une notation pour désigner le calcul du déterminant uniquement avec les coordonnées :

$$\det_{\mathcal{B}}(\vec{u}_1, \vec{u}_2) = \begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix} = x_1 y_2 - x_2 y_1$$

Ainsi, de même qu'un produit scalaire nul caractérise l'orthogonalité, un déterminant nul caractérise la colinéarité.

Les vecteurs de l'espace ont trois coordonnées. Pour comparer deux produits en croix, il faut choisir 2 coordonnées parmi les 3, il y a donc $\binom{3}{2} = 3$ paires de produits en croix à contrôler, ce qui, comme expliqué précédemment, correspond à l'annulation de trois déterminants :

$$\begin{vmatrix} y_1 & y_2 \\ z_1 & z_2 \end{vmatrix} = y_1 z_2 - y_2 z_1 \quad \begin{vmatrix} x_1 & x_2 \\ z_1 & z_2 \end{vmatrix} = x_1 z_2 - x_2 z_1 \quad \begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix} = x_1 y_2 - x_2 y_1$$

où ces coordonnées sont celles de $\vec{u}_1 \begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix}$ et $\vec{u}_2 \begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix}$. Définissons le vecteur $\vec{v} \begin{pmatrix} y_1 z_2 - y_2 z_1 \\ x_1 z_2 - x_2 z_1 \\ x_1 y_2 - x_2 y_1 \end{pmatrix}$ à partir des trois déterminants. Ainsi, $\vec{u}_1$ et $\vec{u}_2$ sont colinéaires si, et seulement si les coordonnées de $\vec{v}$ sont toutes nulles, autrement

dit si, et seulement si $\vec{v} = \vec{0}$. Un mot sur l'ordre des coordonnées de $\vec{v}$: on aurait pu choisir l'une des $3! = 6$ permutations possibles, mais celle-ci a une petite propriété : la coordonnée x dépend des coordonnées y et z de $\vec{u}_1$ et $\vec{u}_2$, de même la coordonnée y dépend des coordonnées x et z , et la coordonnée z dépend des coordonnées x et y . Cette symétrie est, de prime abord, simplement satisfaisante pour l'esprit, mais comme on va le voir tout de suite, elle débloquent une propriété fondamentale de l'objet que l'on est en train de construire.

Armé d'un nouveau vecteur, il est légitime de se demander ce qu'il peut avoir de spécial, notamment par rapport à ses parents $\vec{u}_1$ et $\vec{u}_2$. Par construction, il caractérise la colinéarité de ces deux derniers. Mais que peut-on dire de $\vec{v}$ par rapport à $\vec{u}_1$? Dans cette situation très générale, on dispose de peu d'outils, mais on en a au moins un : le produit scalaire. Regardons l'orthogonalité éventuelle de ces vecteurs ; on colore les différents termes similaires pour s'y retrouver dans cet océan de variables :

$$\begin{aligned}\vec{u}_1 \cdot \vec{v} &= x_1(y_1z_2 - y_2z_1) + y_1(x_1z_2 - x_2z_1) + z_1(x_1y_2 - x_2y_1) \\ &= \textcolor{red}{x_1y_1z_2} - \textcolor{green}{x_1y_2z_1} + \textcolor{red}{x_1y_1z_2} - \textcolor{blue}{x_2y_1z_1} + \textcolor{green}{x_1y_2z_1} - \textcolor{blue}{x_2y_1z_1}\end{aligned}$$

On remarque que les termes vont par paires, mais... à part les termes en vers, rien ne se simplifie. Les deux termes rouges sont affublés d'un $+$ et les deux bleus d'un $-$. Ce qui peut s'arranger très simplement ! En effet, remarquons que le deuxième groupe, issu du produit des ordonnées des deux vecteurs, contient exactement un terme rouge et

un bleu. Il suffit donc d'en renverser le signe pour que tout se compense ! Définissons le vecteur $\vec{u}_3 = \begin{pmatrix} y_1z_2 - y_2z_1 \\ -x_1z_2 + x_2z_1 \\ x_1y_2 - x_2y_1 \end{pmatrix}$.

On a simplement opposé l'ordonnée de $\vec{v}$: cela caractérise toujours la colinéarité de $\vec{u}_1$ et $\vec{u}_2$, car l'annulation d'une quantité est équivalente à l'annulation de son opposé. De là, une propriété remarquable surgit :

$$\begin{aligned}\vec{u}_1 \cdot \vec{u}_3 &= x_1(y_1z_2 - y_2z_1) + y_1(-x_1z_2 + x_2z_1) + z_1(x_1y_2 - x_2y_1) \\ &= \textcolor{red}{x_1y_1z_2} - \textcolor{green}{x_1y_2z_1} - \textcolor{red}{x_1y_1z_2} + \textcolor{blue}{x_2y_1z_1} + \textcolor{green}{x_1y_2z_1} - \textcolor{blue}{x_2y_1z_1} \\ &= 0\end{aligned}$$

On vérifie aisément qu'il en va de même pour $\vec{u}_2 \cdot \vec{u}_3$. Quelles que soient les coordonnées de $\vec{u}_1$ et $\vec{u}_2$, le vecteur $\vec{u}_3$ est orthogonal aux deux premiers !

B – Produit vectoriel de deux vecteurs

Soient $\vec{u}_1 = \begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix}$ et $\vec{u}_2 = \begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix}$ deux vecteurs de l'espace.

Définition 1

Produit vectoriel

Le *produit vectoriel* de $\vec{u}_1$ et $\vec{u}_2$ est le vecteur, noté $\vec{u}_1 \wedge \vec{u}_2$ (ou, dans la littérature anglophone, $\vec{u}_1 \times \vec{u}_2$),

de coordonnées $\begin{pmatrix} y_1z_2 - y_2z_1 \\ -x_1z_2 + x_2z_1 \\ x_1y_2 - x_2y_1 \end{pmatrix}$.

Proposition 2

Colinéarité et orthogonalité

- ▶ $\vec{u}_1 \wedge \vec{u}_2 = \vec{0}$ si, et seulement si $\vec{u}_1$ et $\vec{u}_2$ sont colinéaires.
- ▶ $\vec{u}_1 \wedge \vec{u}_2$ est orthogonal à $\vec{u}_1$ et $\vec{u}_2$

Corollaire 3

Si $\vec{u}_1$ et $\vec{u}_2$ ne sont pas colinéaires, alors $\vec{u}_1 \wedge \vec{u}_2$ est un vecteur normal au plan dirigé par $\vec{u}_1$ et $\vec{u}_2$.

Corollaire 4

Soit $\vec{w}$ un vecteur de l'espace. Les vecteurs $\vec{u}_1$, $\vec{u}_2$ et $\vec{w}$ sont coplanaires si, et seulement si $\vec{w} \cdot (\vec{u}_1 \wedge \vec{u}_2) = 0$.

Théorème 5**Norme du produit vectoriel**

$$\|\vec{u}_1 \wedge \vec{u}_2\| = \|\vec{u}_1\| \|\vec{u}_2\| |\sin(\vec{u}_1, \vec{u}_2)|$$

Démonstration. L'astuce* est de d'abord montrer l'identité suivante :

$$\|\vec{u}_1 \wedge \vec{u}_2\|^2 + (\vec{u}_1 \cdot \vec{u}_2)^2 = \|\vec{u}_1\|^2 \|\vec{u}_2\|^2$$

La démonstration est calculatoire et pénible, mais sans piège : il suffit de tout écrire.

$$\|\vec{u}_1 \wedge \vec{u}_2\|^2 = (y_1 z_2 - y_2 z_1)^2 + (-x_1 z_2 + x_2 z_1)^2 + (x_1 y_2 - x_2 y_1)^2$$

$$(\vec{u}_1 \cdot \vec{u}_2)^2 = (x_1 y_1 + y_1 y_2 + z_1 z_2)^2$$

$$\|\vec{u}_1\|^2 \|\vec{u}_2\|^2 = (x_1^2 + y_1^2 + z_1^2)(x_2^2 + y_2^2 + z_2^2)$$

On développe puis simplifie tout et tout concorde. De là en écrivant $\vec{u}_1 \cdot \vec{u}_2 = \|\vec{u}_1\| \|\vec{u}_2\| \cos(\vec{u}_1, \vec{u}_2)$:

$$\begin{aligned} \|\vec{u}_1 \wedge \vec{u}_2\|^2 &= \|\vec{u}_1\|^2 \|\vec{u}_2\|^2 - (\vec{u}_1 \cdot \vec{u}_2)^2 = \|\vec{u}_1\|^2 \|\vec{u}_2\|^2 - (\|\vec{u}_1\| \|\vec{u}_2\| \cos(\vec{u}_1, \vec{u}_2))^2 \\ &= \|\vec{u}_1\|^2 \|\vec{u}_2\|^2 (1 - \cos^2(\vec{u}_1, \vec{u}_2)) = \|\vec{u}_1\|^2 \|\vec{u}_2\|^2 \sin^2(\vec{u}_1, \vec{u}_2) \end{aligned}$$

et donc, des normes étant des quantités positives :

$$\|\vec{u}_1 \wedge \vec{u}_2\| = \sqrt{\|\vec{u}_1\|^2 \|\vec{u}_2\|^2 \sin^2(\vec{u}_1, \vec{u}_2)} = \|\vec{u}_1\| \|\vec{u}_2\| \sin(\vec{u}_1, \vec{u}_2) = \|\vec{u}_1\| \|\vec{u}_2\| |\sin(\vec{u}_1, \vec{u}_2)|$$

□

La coplanarité de trois vecteurs dans l'espace est l'analogue de la colinéarité de deux vecteurs dans le plan. Pour plus de détails sur cette analogie, voir le passage sur les familles linéairement indépendantes dans le cours complet.

Le déterminant permet de caractériser les couples de vecteurs non colinéaires, *a fortiori* susceptibles de former une base du plan. On définit de même une notion de déterminant dans l'espace, qui caractérise les triplets de vecteurs susceptibles de former une base de l'espace.

Définition 6**Déterminant dans l'espace**

Soient $\vec{u}_1 \begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix}$ et $\vec{u}_2 \begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix}$ trois vecteurs de l'espace. Le *déterminant* du triplet $(\vec{u}_1, \vec{u}_2, \vec{u}_3)$ est le nombre, noté de deux manières, et défini comme :

$$\det_{\mathcal{B}}(\vec{u}_1, \vec{u}_2, \vec{u}_3) = \begin{vmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \end{vmatrix} = \vec{u}_1 \cdot (\vec{u}_2 \wedge \vec{u}_3)$$

*. Elle se devine aisément : la formule proposée nous rappelle étrangement celle du produit scalaire qui a un cosinus à la place. Par ailleurs, on sait que $\cos^2(\theta) + \sin^2(\theta) = 1$, d'où le lien.

Remarque 7 (Le déterminant dépend de la base)

Depuis le début, on note $\det_{\mathcal{B}}$ pour le déterminant des familles de vecteurs, au lieu de simplement $\det$. En fait, la valeur du déterminant dépend du choix de la base $\mathcal{B}$: si l'on choisit une autre base, alors la valeur du déterminant change. Ce qui ne change pas, c'est s'il est nul ou non : et c'est la seule raison d'être du déterminant.

Remarque 8 (Règle de Sarrus)

Pour plus de détails sur la notation en tableau du déterminant, renseignez vous sur la « méthode de Sarrus ».

Propriété 9**Déterminant et coplanarité**

Les vecteurs $\vec{u}_1$, $\vec{u}_2$ et $\vec{u}_3$ sont coplanaires si, et seulement si leur déterminant est nul.

2 – Autour des sphères

A – Rappel : cercles du plan

On se place dans un plan $\mathcal{P}$ muni d'un repère orthonormé.

Définition 10**Cercle**

Soient A un point du plan et $r > 0$.

Le *cercle* de centre A et de *rayon* r est l'ensemble $\mathcal{C}$ des points du plan à distance r de A , autrement dit :

$$\mathcal{C} = \{M \in \mathcal{P} \mid AM = r\}$$

Dans la suite, A désigne un point du plan, r un réel strictement positif, et $\mathcal{C}$ le cercle de centre A et de rayon r .

Remarque 11 (Vocabulaire)

- ▶ Un *rayon* de $\mathcal{C}$ est un segment reliant A à un point de $\mathcal{C}$. À ne pas confondre avec le *rayon* de $\mathcal{C}$ qui est un nombre, plus précisément, r . Le *rayon* de $\mathcal{C}$ est la longueur commune de tous les rayons de $\mathcal{C}$.
- ▶ De même, un *diamètre* de $\mathcal{C}$ est un segment reliant deux points de $\mathcal{C}$ tels que A est le milieu de ce segment. Le *diamètre* de $\mathcal{C}$ est la longueur commune de tous ses diamètres.
- ▶ Le *disque* $\mathcal{D}$ de centre A et de rayon r est l'ensemble des points à distance *au plus* r de A , autrement dit $\mathcal{D} = \{M \in \mathcal{P} \mid AM \leq r\}$. Géométriquement, $\mathcal{D}$ est l'intérieur de $\mathcal{C}$. Un disque a une aire, un cercle a une longueur (aussi appelée *circonférence*). On peut parler du périmètre d'un disque ou d'un cercle, bien qu'on peut défendre le fait que la notion de périmètre ne s'applique qu'aux surfaces, ce qui n'est pas le cas d'un cercle.

Théorème 12**Équation de cercle**

Notons $A(x_A, y_A)$ les coordonnées de A dans le repère orthonormé choisi en amont.

$$M(x, y) \in \mathcal{C} \iff (x - x_A)^2 + (y - y_A)^2 = r^2$$

Cette équation est appelée *équation du cercle* $\mathcal{C}$, et on note $\mathcal{C} : (x - x_A)^2 + (y - y_A)^2 = r^2$.

Démonstration. La quantité $(x - x_A)^2 + (y - y_A)^2$ est la distance AM au carré. Une distance étant positive, $M \in \mathcal{C} \iff AM = r \iff AM^2 = r^2$, ce qui donne bien l'équation proposée. $\square$

B – Équations de sphères

On se place dans l'espace euclidien $\mathcal{E}$ muni d'un repère orthonormé.

Sphère

Définition 13

Soient A un point de l'espace et $r > 0$.

La *sphère* de centre A et de rayon r est l'ensemble $\mathcal{S}$ des points de l'espace à distance r de A , autrement dit :

$$\mathcal{S} = \{M \in \mathcal{E} \mid AM = r\}$$

Dans toute la suite, A désigne un point de l'espace, r un réel strictement positif, et $\mathcal{S}$ la sphère de centre A et de rayon r .

Remarque 14 (Vocabulaire)

- Un rayon de $\mathcal{S}$ est un segment reliant A à un point de $\mathcal{S}$. À ne pas confondre avec le rayon de $\mathcal{S}$ qui est un nombre, plus précisément, r . Le rayon de $\mathcal{S}$ est la longueur commune de tous les rayons de $\mathcal{S}$.
- De même, un diamètre de $\mathcal{S}$ est un segment reliant deux points de $\mathcal{S}$ tels que A est le milieu de ce segment. Le diamètre de $\mathcal{S}$ est la longueur commune de tous ses diamètres.
- La *boule* $\mathcal{B}$ de centre A et de rayon r est l'ensemble des points à distance *au plus* r de A , autrement dit $\mathcal{B} = \{M \in \mathcal{E} \mid AM \leq r\}$. Géométriquement, $\mathcal{B}$ est l'intérieur de $\mathcal{S}$. Une boule a un volume, une sphère a une aire. On peut parler de l'aire de la boule.

Équation de sphère

Théorème 15

Notons $A(x_A, y_A, z_A)$ les coordonnées de A dans le repère orthonormé choisi en amont.

$$M(x, y, z) \in \mathcal{S} \iff (x - x_A)^2 + (y - y_A)^2 + (z - z_A)^2 = r^2$$

Cette équation est appelée *équation de la sphère* $\mathcal{S}$, et on note $\mathcal{S} : (x - x_A)^2 + (y - y_A)^2 + (z - z_A)^2 = r^2$.

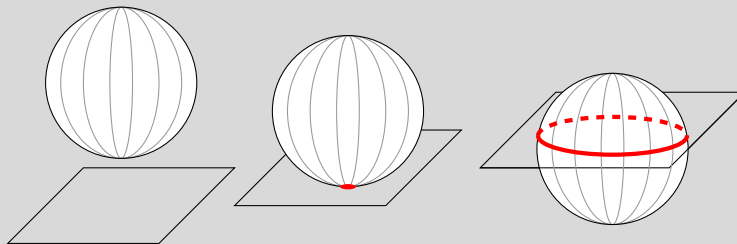
Démonstration. La démonstration est la même que pour les cercles. □

C – Sections de sphères

Section plane de sphères

Théorème 16

L'intersection d'un plan avec une sphère est soit vide, soit un point (on dit alors que le plan est *tangent* à la sphère), soit un cercle.

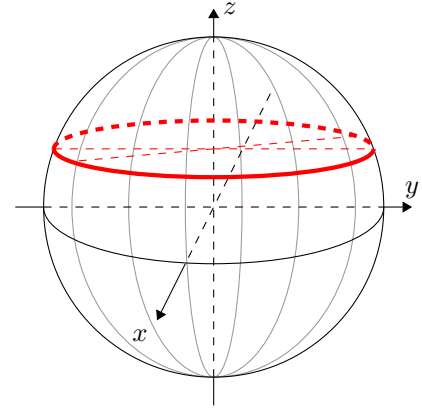


Démonstration.

Choisissons un repère de l'espace dont l'origine est le centre de la sphère, et tel que le plan $\mathcal{P}$ qui intersecte (ou pas!) la sphère a pour équation $z = k$, avec k un réel (on peut choisir un tel repère sans perte de généralité, car ça revient à tourner et translater la sphère et le plan, et les rotations et translations préservent les figures).

Pour des raisons de symétrie, supposons k positif. Commençons par remarquer la chose suivante : si on note $H(0, 0, k)$ le projeté orthogonal de O sur $\mathcal{P}$, alors pour tout $M(0, 0, k) \in \mathcal{P}$, $OM \geq OH = k$ (car le projeté orthogonal minimise la distance. Ainsi, lorsque $k > r$, pour tout $M \in \mathcal{P}$, $OM \geq k > r$, donc aucun point de $\mathcal{P}$ ne peut être à distance exactement r de O , autrement dit aucun point de $\mathcal{P}$ n'appartient à la sphère, et donc l'intersection est vide. Lorsque $k = r$, on se sert du fait que le projeté orthogonal est l'unique point à minimiser la distance, donc pour tout $M \in \mathcal{P}$ distinct de H , on a $OM > OH = k = r$, donc n'appartient pas à la sphère. Ainsi, dans ce cas, le seul point d'intersection est H .

Supposons désormais $0 \leq k < r$. Soit M dans l'intersection de $\mathcal{P}$ et $\mathcal{S}$. Notons $H(0, 0, k)$ le projeté orthogonal de O sur $\mathcal{P}$. Comme OHM est rectangle en H (par projection), d'après le théorème de Pythagore, $OH^2 + HM^2 = OM^2$. Or $M \in \mathcal{S}$ et O est le centre de $\mathcal{S}$, donc $OM = r$, donc la relation de Pythagore se réécrit $k^2 + HM^2 = r^2$, soit $HM^2 = r^2 - k^2$. Or $0 \leq k < r$, donc $k^2 < r^2$ (par stricte croissance de la fonction carré sur $\mathbb{R}^+$), donc on peut poser $\rho = \sqrt{r^2 - k^2} > 0$, de sorte que $HM^2 = \rho^2$, donc $HM = \rho$ (car ce sont des grandeurs positives). Comme ρ ne dépend pas de x ni de y , on peut affirmer que M dans l'intersection est un point du plan à distance de ρ de H ; finalement l'intersection de $\mathcal{S}$ et $\mathcal{P}$ est le cercle de $\mathcal{P}$ de centre H et de rayon ρ . $\square$



Théorème 17

Intersection de deux sphères

L'intersection de deux sphères est soit vide, soit un point (on dit alors que les sphères sont *tangentes*), soit un cercle.

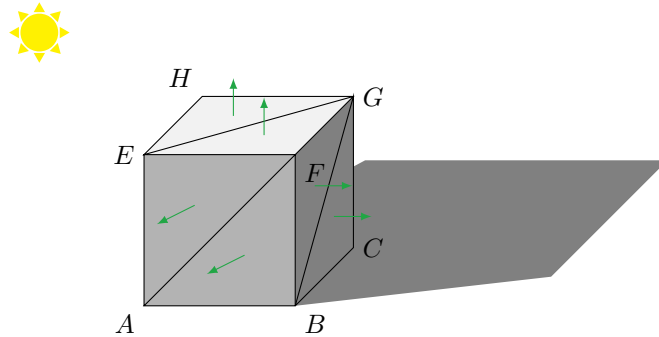
Démonstration. Soient $\mathcal{S}$ et $\mathcal{S}'$ deux sphères, de centres O et O' , et de rayons r et r' . Notons $k = OO'$, et choisissons un repère orthonormé de l'espace tel que $O(0, 0, 0)$ et $O'(0, 0, k)$. Dans ce repère, $\mathcal{S} : x^2 + y^2 + z^2 = r^2$ et $\mathcal{S}' : x^2 + y^2 + (z - k)^2 = r'^2$.

Soit $M(x, y, z)$ dans l'éventuelle intersection de $\mathcal{S}$ et $\mathcal{S}'$. Ses coordonnées vérifient simultanément les équations des deux sphères. En soustrayant la relation donnée par $\mathcal{S}$ à celle donnée par $\mathcal{S}'$, on obtient l'équation $(z - k)^2 - z^2 = r'^2 - r^2$, soit $z = -(r'^2 - r^2 - k^2)/(2k)$. Cette expression ne dépend que de k , r et r' , autrement dit pas de M . Cela montre que tout point de l'éventuelle intersection de $\mathcal{S}$ et $\mathcal{S}'$ appartient au plan $\mathcal{P}$ d'équation $z = -(r'^2 - r^2 - k^2)/(2k)$. Or on a vu précédemment que l'intersection d'un plan et d'une sphère est vide, ponctuelle ou circulaire, et le raisonnement que l'on vient de mener montre que tout point de l'intersection de $\mathcal{S}$ et $\mathcal{S}'$ est aussi dans l'intersection de $\mathcal{S}$ et $\mathcal{P}$, ce qui montre la proposition. $\square$

3 – Pour la culture : mais à quoi ça sert ? !

Les produits scalaires, projetés orthogonaux et autres vecteurs normaux sont des outils mathématiques très présents quand des problèmes de géométrie dite « linéaire » se présentent ; et pas seulement dans des exercices théoriques !

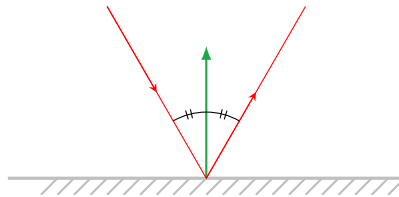
Un exemple d'application dans notre monde moderne est la gestion des espaces virtuels tridimensionnels. La majorité du temps, toute surface est subdivisée en polygones élémentaires – la plupart du temps des triangles – pour créer ce qu'on appelle un *maillage polygonal* (*polygonal mesh*, en anglais). Les mondes virtuels obéissent (outre les escapades fantaisistes) aux règles de la physique réelles ; notamment en ce qui concerne l'éclairage.



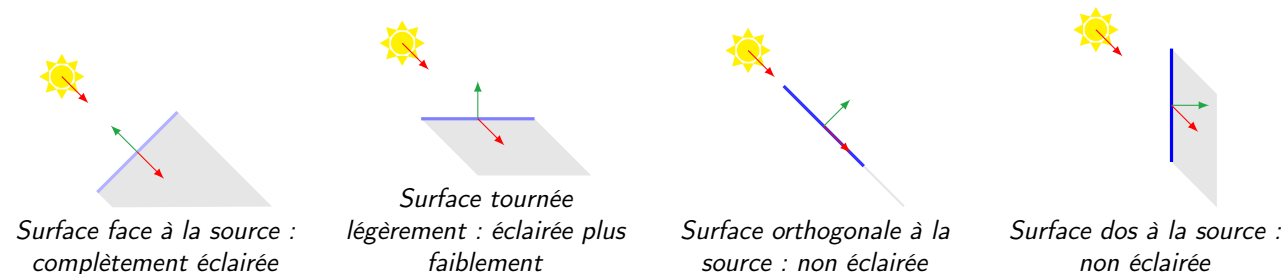
Maillage polygonal d'un cube, vecteur normaux du maillage et ombres

Il y a deux types de surfaces : les *surfaces réfléchissantes*, comme les miroirs, qui *conservent*, en un sens, les rayons, et les *surfaces diffuses*, qui dispersent les rayons dans toutes les directions : par exemple, on ne voit pas le reflet des rayons du Soleil dans un rocher, ceux-ci sont dispersés dans tous les sens.

Les surfaces réfléchissantes obéissent à une règle simple : tout rayon incident se réfléchit symétriquement par rapport à la normale à la surface.

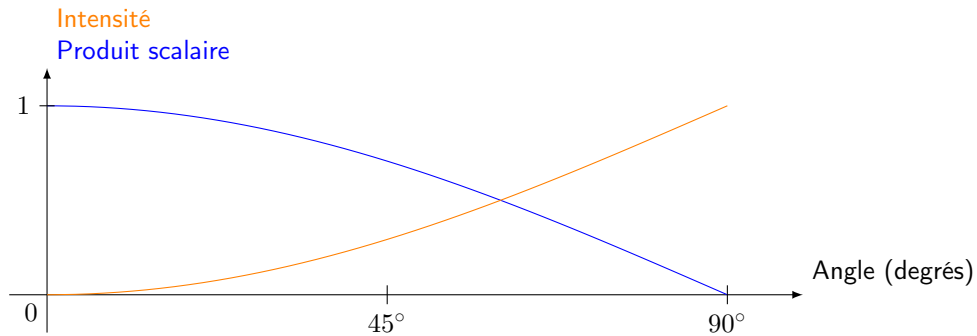


Concernant les surfaces diffuses, l'idée est de quantifier à quel point une surface est en face ou dos à une source lumineuse. Les vecteurs normaux représentent justement la direction à laquelle les surfaces font face, autrement dit ils *pointent* dans la direction qui est *devant* la surface. (Remarque : si $\vec{n}$ est un vecteur normal, alors $-\vec{n}$ en est aussi un. Dans ce genre de contexte, on calcule les vecteurs normaux d'une manière particulière, afin d'orienter les surfaces, et décider d'un *avant* et d'un *arrière*). À mesure qu'on tourne la surface, son intensité lumineuse apparente faiblit, jusqu'à ce que la surface soit orthogonale à la source, où l'intensité est nulle, et elle reste nulle tant que la surface est dos à la source.

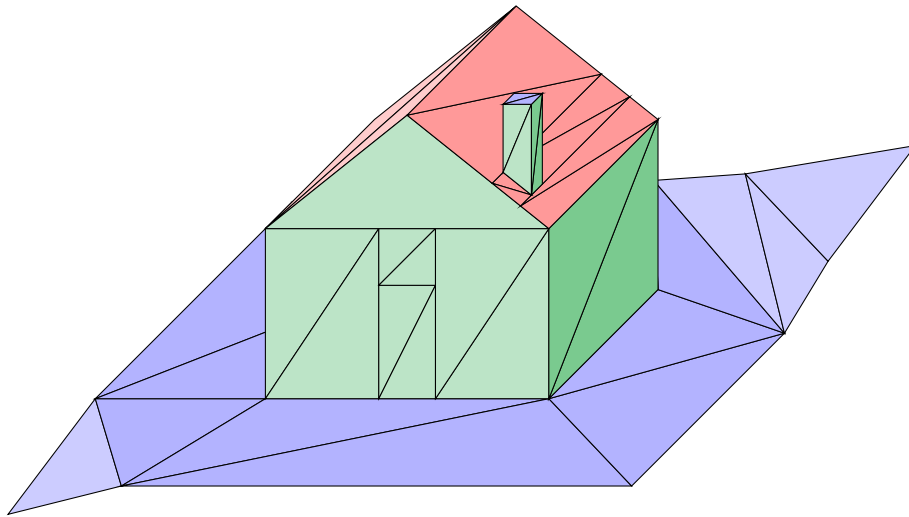


La source lumineuse émet des rayons dans une certaine direction, et selon un certain sens : ils sont donc représentés par des vecteurs. L'orientation d'une surface est représentée par son vecteur normal. Et pour déterminer l'inclinaison relative de deux directions, on peut utiliser le produit scalaire entre leurs vecteurs directeurs, ici le vecteur source et

le vecteur normal ! Dans ce genre de contexte, on considère toujours des vecteurs *normalisés**, c'est-à-dire de norme 1, afin que les produits scalaires, qui sont donc égaux aux cosinus des angles, soient eux aussi normalisés, entre -1 et 1 . Notez la relation inverse entre le cosinus de l'angle et l'intensité de l'éclairage. Ici, on considère le vecteur source de la source vers l'objet. C'est un choix, mais selon lequel lorsque la surface est parfaitement en face de la source, les vecteurs source et normal sont colinéaires et de sens opposés, formant donc un angle de 180° . Ainsi, la source est plus ou moins éclairée pour un angle entre 90° et 180° , avec un éclairage maximal à 180° et nul à 90° , et un éclairage nul pour un angle entre 0° et 90° , correspondant au cas où la surface est dos à la source lumineuse. (Remarque : bien sûr, une surface dos à une source lumineuse est un minimum éclairée, grâce aux réflexions secondaires sur les surfaces environnantes.)



Connaître l'angle que fait le vecteur normal par rapport à la direction du sol permet également de classer les surfaces selon si elles sont des sols sur lesquels le personnage peut marcher, des plafonds ou des murs qui ne peuvent pas être traversés, voire des pentes sur lesquelles le personnage doit glisser. Et c'est encore une fois en faisant un produit scalaire que l'on obtient cette information, puisque ce dernier est lié à l'angle formé entre les deux (directions des) vecteurs. (Remarque : là encore, un choix d'orientation est fait entre $\vec{n}$ et $-\vec{n}$ pour pouvoir discriminer, par exemple, les sols des plafonds, qui seront parallèles à la direction de référence, mais qui n'ont pas le même statut.)

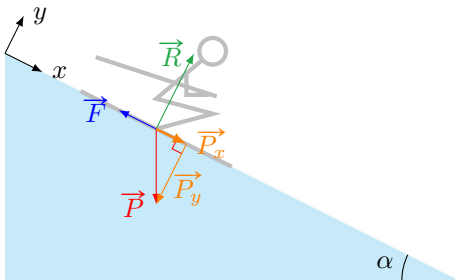


*. Normaliser un vecteur, c'est le diviser par sa norme, autrement dit la racine carrée de la somme des carrés de ses coordonnées. Mathématiquement, c'est simple, mais pour un ordinateur, diviser et calculer une racine carrée, c'est compliqué... Alors quand on doit le faire pour des millions de polygones de surface et de rayons, c'est très coûteux en temps de calcul, tant et si bien qu'il est impossible de les calculer en temps réel dans un jeu ! Enfin... jusqu'à ce que les développeurs de *Quake* trouvent un algorithme extrêmement efficace, maintenant connu sous le nom de *Fast Inverse Square Root Algorithm*.



Type de surface en fonction de l'orientation du vecteur normal

En physique, l'aspect géométrique et vectoriel intervient naturellement lorsque l'on travaille avec des forces. Le cas d'école est celui d'un skieur dévalant ou remontant une pente. Il est soumis à son poids, à la résistance du sol et à des forces frottement. La réaction du sol n'influence pas son mouvement, car elle est orthogonale à sa direction, puisque dirigée selon la normale au sol. Les forces de frottement le ralentissent, dans le sens opposé à son mouvement. Pour le poids, il y a une subtilité : ce n'est pas *tout* le poids qui influence le mouvement. En effet, on peut décomposer le poids selon deux composantes, l'une colinéaire au mouvement, et l'autre orthogonale. De même que pour la réaction du sol, la composante orthogonale du poids n'agit pas sur le mouvement. C'est purement la composante colinéaire qui a une incidence, autrement dit la projection orthogonale du poids (en faisant un abus de langage, en parlant de projection orthogonale d'un vecteur sur une droite) sur la direction du mouvement qui nous intéresse.



Lorsqu'il descend, le skieur n'a pas besoin d'aide pour avancer, la gravité est le moteur de son mouvement. En revanche, lorsqu'il doit remonter, il a besoin d'une force motrice pour contrer la gravité qui le tire vers le bas. C'est pourquoi les remontées mécaniques, ou *tire-fesses* existent. Le tire-fesses exerce sur le skieur une force dans la direction de son mouvement de remontée, mais pas nécessairement colinéaire pour des raisons matérielles. Intuitivement, on imagine que le tire-fesses est d'autant plus efficace que la force qu'il exerce est colinéaire à la direction du mouvement. Inversement, il est d'autant moins efficace qu'il exerce une force orthogonale : autrement dit ici si le tire-fesses vous tire vers le haut, vous avancerez plus difficilement que s'il vous tirait tout droit.

Cette idée est quantifiée par la notion de *travail* d'une force ; on a intuité que le travail d'une force est d'autant plus grand que l'angle est proche de 0° , et d'autant plus proche de 0 lorsque l'angle devient droit : on retrouve le comportement du produit scalaire entre le vecteur dirigeant le mouvement et celui dirigeant la force, tant et si bien que le travail d'une force se calcule justement grâce à un produit scalaire.

4 – Bonus : démonstration des propriétés du produit scalaire plan

Avant de commencer les démonstrations, quelques éléments de culture : au lycée, le produit scalaire est présenté à l'envers. La définition avec le cosinus n'est pas une définition, mais une proposition. * Par ailleurs, on appelle « produit scalaire » n'importe quelle opération qui à deux vecteurs associe un nombre réel, qui est symétrique (échanger ses deux composantes ne change pas le résultat), bilinéaire (voir le cours complet), et qui donne un réel positif quand elle est appliquée à deux fois le même vecteur. En particulier, la formule avec les coordonnées est un produit scalaire car elle vérifie ces propriétés, que nous allons montrer.

Avant toute chose, nous devons régler un problème : la définition avec les coordonnées... dépend des coordonnées ! Et donc, jusqu'à preuve du contraire, la valeur d'un produit scalaire dépend du choix de la base (car les coordonnées elles-mêmes sont relatives à une base).

La démonstration de certaines choses dans l'espace nécessite des outils et un formalisme bien au-delà du niveau lycée, aussi pour ces démonstrations nous nous placerons dans le plan.

*. Posez vous cette question : comment définit-on (la mesure de) l'angle entre deux vecteurs ? Réponse : avec le produit scalaire... On ne peut pas définir un objet à partir d'une formule qui utilise cet objet, c'est circulaire !

Lemme 18 (Indépendance de la base)

Soient $\mathcal{B} = (\vec{i}, \vec{j})$ et $\mathcal{B}' = (\vec{i}', \vec{j}')$ deux bases orthonormées du plan. Soient $\vec{u}$ et $\vec{v}$ deux vecteurs du plan.

On note $\begin{pmatrix} x_u \\ y_u \end{pmatrix}$ et $\begin{pmatrix} x_v \\ y_v \end{pmatrix}$ dans $\mathcal{B}$ et de coordonnées $\begin{pmatrix} x'_u \\ y'_u \end{pmatrix}$ et $\begin{pmatrix} x'_v \\ y'_v \end{pmatrix}$ dans $\mathcal{B}'$. Alors :

$$x_u x_v + y_u y_v = x'_u x'_v + y'_u y'_v$$

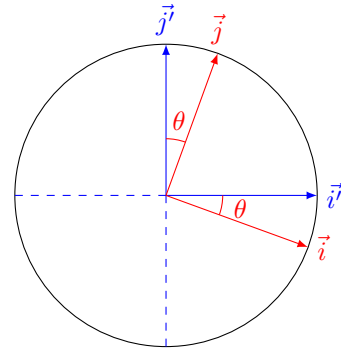
Démonstration. On va supposer que $\mathcal{B}$ et $\mathcal{B}'$ sont « dans le même sens », c'est-à-dire que $\mathcal{B}'$ s'obtient à partir de $\mathcal{B}$ à partir d'une rotation. L'autre cas, où ce n'est pas possible, c'est-à-dire que $\vec{i}'$ et $\vec{j}'$ sont inversés, est analogue. On note θ une mesure de l'angle de cette rotation.

Comme les bases sont orthonormées, on peut représenter ces bases dans un cercle trigonométrique :

On lit sur ce dessin que $\vec{i} = \cos(-\theta)\vec{i}' + \sin(-\theta)\vec{j}' = \cos(\theta)\vec{i}' - \sin(\theta)\vec{j}'$
et que $\vec{j} = \cos(\pi/2 - \theta)\vec{i}' + \sin(\pi/2 - \theta)\vec{j}' = \sin(\theta)\vec{i}' + \cos(\theta)\vec{j}'$.

Si $\vec{a}$ a pour coordonnées $\begin{pmatrix} x \\ y \end{pmatrix}$ dans $\mathcal{B}$, c'est-à-dire si $\vec{a} = x\vec{i} + y\vec{j}$, alors :

$$\begin{aligned} \vec{a} &= x\vec{i} + y\vec{j} \\ &= x(\cos(\theta)\vec{i}' - \sin(\theta)\vec{j}') + y(\sin(\theta)\vec{i}' + \cos(\theta)\vec{j}') \\ &= [x \cos(\theta) + y \sin(\theta)]\vec{i}' + [-x \sin(\theta) + y \cos(\theta)]\vec{j}' \end{aligned}$$



Donc les coordonnées de $\vec{a}$ dans $\mathcal{B}'$ sont $\begin{pmatrix} x \cos(\theta) + y \sin(\theta) \\ -x \sin(\theta) + y \cos(\theta) \end{pmatrix}$. En appliquant cette formule à $\vec{u}$ et $\vec{v}$:

$$\begin{aligned} &x'_u x'_v + y'_u y'_v \\ &= [x_u \cos(\theta) + y_u \sin(\theta)][x_v \cos(\theta) + y_v \sin(\theta)] \\ &\quad + [-x_u \sin(\theta) + y_u \cos(\theta)][-x_v \sin(\theta) + y_v \cos(\theta)] \\ &= x_u x_v \cos^2(\theta) + x_u y_v \cos(\theta) \sin(\theta) + x_v y_u \cos(\theta) \sin(\theta) \\ &\quad + y_u y_v \sin^2(\theta) + x_u x_v \sin^2(\theta) - x_u y_v \cos(\theta) \sin(\theta) \\ &\quad - x_v y_u \cos(\theta) \sin(\theta) + y_u y_v \cos^2(\theta) \\ &= x_u x_v [\cos^2(\theta) + \sin^2(\theta)] + y_u y_v [\cos^2(\theta) + \sin^2(\theta)] \\ &\quad + [x_u y_v + x_v y_u - x_u y_v - x_v y_u] \cos(\theta) \sin(\theta) \\ &= x_u x_v + y_u y_v \end{aligned}$$

□

Maintenant que l'on a montré que cette formule indépendante de la base, on peut proprement définir le produit scalaire :

Définition 19**Produit scalaire**

Soient $\vec{u}$ et $\vec{v}$ deux vecteurs du plan, notons $\begin{pmatrix} x \\ y \end{pmatrix}$ et $\begin{pmatrix} x' \\ y' \end{pmatrix}$ leurs coordonnées dans une base orthonormée quelconque.

Le *produit scalaire* de $\vec{u}$ et $\vec{v}$ est le réel, indépendant de la base orthonormée :

$$\vec{u} \cdot \vec{v} = xx' + yy'$$

Propriété 20

Le produit scalaire est bilinéaire et symétrique.

Démonstration. Soient $\vec{u}$, $\vec{v}$ et $\vec{w}$ trois vecteurs du plan, notons $\begin{pmatrix} x_u \\ y_u \end{pmatrix}$, $\begin{pmatrix} x_v \\ y_v \end{pmatrix}$ et $\begin{pmatrix} x_w \\ y_w \end{pmatrix}$ leurs coordonnées dans une base orthonormée quelconque. Soit k un nombre réel.

Pour la symétrie : $\vec{u} \cdot \vec{v} = x_u x_v + y_u y_v = x_v x_u + y_v y_u = \vec{v} \cdot \vec{u}$

Pour la bilinéarité : d'une part, $(k\vec{u}) \cdot \vec{v} = kx_u x_v + ky_u y_v = k(x_u x_v + y_u y_v) = k(\vec{u} \cdot \vec{v})$, d'autre part,

$$\begin{aligned} \vec{u} \cdot (\vec{v} + \vec{w}) &= x_u(x_v + x_w) + y_u(y_v + y_w) = x_u x_v + x_u x_w + y_u y_v + y_u y_w = (x_u x_v + y_u y_v) + (x_u x_w + y_u y_w) \\ &= \vec{u} \cdot \vec{v} + \vec{u} \cdot \vec{w} \end{aligned}$$

□

Remarque 21

En réalité, la norme n'est pas définie comme la longueur d'un vecteur... c'est l'inverse ! La réelle définition de la norme de $\vec{u}$ est $\|\vec{u}\| = \sqrt{\vec{u} \cdot \vec{u}}$. Celle-ci est bien définie car $\vec{u} \cdot \vec{u} = x^2 + y^2 \geq 0$ avec les notations évidentes. On remarque aussi avec l'égalité $\|\vec{u}\|^2 = x^2 + y^2$ que la norme d'un vecteur est nulle si, et seulement si le vecteur est nul.

Propriété 22

Définition géométrique du produit scalaire

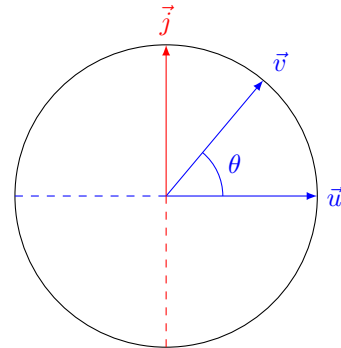
Soient $\vec{u}$ et $\vec{v}$ deux vecteurs.

$$\vec{u} \cdot \vec{v} = \|\vec{u}\| \times \|\vec{v}\| \times \cos(\vec{u}, \vec{v})$$

Démonstration. Si $\vec{u}$ ou $\vec{v}$ est nul, alors l'égalité est immédiate.

Supposons d'abord que $\|\vec{u}\| = \|\vec{v}\| = 1$. On note $\theta = (\vec{u}, \vec{v})$. Soit $\vec{j}$ un vecteur unitaire orthogonal à $\vec{u}$, de sorte que $\mathcal{B} = (\vec{u}, \vec{j})$ est une base orthonormée. Dans cette base, $\vec{u}$ a pour coordonnées $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$, et $\vec{v}$ a pour coordonnées $\begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix}$. Ainsi,

$$\vec{u} \cdot \vec{v} = 1 \times \cos(\theta) + 0 \times \sin(\theta) = \cos(\theta)$$



Traisons le cas général. Comme $\vec{u}$ et $\vec{v}$ ne sont pas nuls, $\|\vec{u}\|$ et $\|\vec{v}\|$ ne le sont pas non plus. Remarquons alors que $\frac{\vec{u}}{\|\vec{u}\|}$ et $\frac{\vec{v}}{\|\vec{v}\|}$ sont unitaires, donc d'après ce qui précède, en remarquant que multiplier $\vec{u}$ et $\vec{v}$ par un nombre ne change pas l'angle entre eux deux :

$$\begin{aligned} \vec{u} \cdot \vec{v} &= \left(\|\vec{u}\| \frac{\vec{u}}{\|\vec{u}\|} \right) \cdot \left(\|\vec{v}\| \frac{\vec{v}}{\|\vec{v}\|} \right) = \|\vec{u}\| \times \|\vec{v}\| \times \left(\frac{\vec{u}}{\|\vec{u}\|} \cdot \frac{\vec{v}}{\|\vec{v}\|} \right) = \|\vec{u}\| \times \|\vec{v}\| \times \cos \left(\frac{\vec{u}}{\|\vec{u}\|}; \frac{\vec{v}}{\|\vec{v}\|} \right) \\ &= \|\vec{u}\| \times \|\vec{v}\| \times \cos(\vec{u}, \vec{v}) \end{aligned}$$

□

12 Intégration

Table des matières

1	Série harmonique	55
2	Intégration numérique : méthode de Monte-Carlo	57
3	Intégrales, longueurs et volumes	59
A	Intro-digression : l'approximation de π par Archimède	59
B	Intégrales curvilignes	60
C	Intégrales et volumes	61
4	Changement de variable	62

1 – Série harmonique

Pour $n \in \mathbb{N}$, notons $H_n = \sum_{k=1}^n \frac{1}{k} = 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n}$. On l'appelle la *série harmonique*. Comme cette somme se comporte-t-elle quand n tend vers $+\infty$?

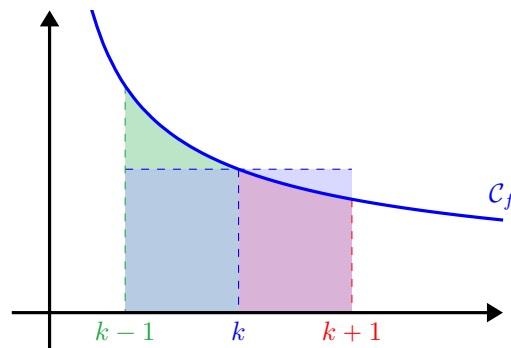
Notons $f : x \mapsto \frac{1}{x}$ la fonction inverse sur $]0; +\infty[$, qui est continue. Soit $k > 1$ un entier. Par décroissance de f :

- Pour tout $x \in [k, k+1]$, $f(x) \leq f(k)$, donc $\int_k^{k+1} f(x) dx \leq \int_k^{k+1} f(k) dx = f(k) \int_k^{k+1} 1 dx = f(k)$.
- Pour tout $x \in [k-1, k]$, $f(k) \leq f(x)$, donc $f(k) = \int_{k-1}^k f(k) dx \leq \int_{k-1}^k f(x) dx$.

Ainsi,

$$\int_k^{k+1} f(x) dx \leq f(k) \quad \text{et} \quad f(k) \leq \int_{k-1}^k f(x) dx$$

Cet encadrement peut également se lire sur un schéma inspiré de la méthode des rectangles : on considère les domaines sous la courbe de f entre $k-1$ et k , en vert sur le schéma, et entre k et $k+1$, en rouge sur le schéma. Leurs aires sont données par les deux intégrales qui forment l'encadrement. L'aire verte est plus grande que l'aire du rectangle bleu, de largeur 1 et de hauteur $f(k)$, et l'aire rouge est plus petite que l'aire du même rectangle bleu.



On somme ces inégalités : à gauche, pour k allant de 1 à n , à droite pour k allant de 2 à n :

$$\sum_{k=1}^n \int_k^{k+1} f(x) dx \leq \sum_{k=1}^n f(k) \quad \text{et} \quad \sum_{k=2}^n f(k) \leq \sum_{k=2}^n \int_{k-1}^k f(x) dx$$

Remarquons que $\sum_{k=2}^n f(k) = f(2) + f(3) + \dots + f(n)$, ce que l'on réécrit $(f(1) + f(2) + \dots + f(n)) - f(1) = \sum_{k=1}^n f(k) - f(1)$. Concernant les intégrales, la relation de Chasles permet de simplifier les expressions : en effet, regardons l'expression de la borne supérieure. On ajoute les intégrales de f entre 1 et 2 (pour $k = 2$), puis entre 2 et 3 (pour $k = 3$), etc. jusqu'entre $n - 1$ et n (pour $k = n$). La relation de Chasles nous permet de recoller tous ces morceaux en un seul : l'intégrale de f entre 1 et n . La borne inférieure se simplifie de la même manière, en une simple intégrale de 1 à $n + 1$. Ainsi :

$$\int_1^{n+1} f(x) \, dx \leq \sum_{k=1}^n f(k) \quad \text{et} \quad \sum_{k=1}^n f(k) - f(1) \leq \int_1^n f(x) \, dx$$

En remplaçant f par son expression :

$$\int_1^{n+1} \frac{1}{x} \, dx \leq \sum_{k=1}^n \frac{1}{k} \quad \text{et} \quad \sum_{k=1}^n \frac{1}{k} - 1 \leq \int_1^n \frac{1}{x} \, dx$$

La somme est égale par définition à H_n . Les intégrales se calculent à l'aide du logarithme, qui est une primitive de f :

$$\ln(n+1) - \ln(1) \leq H_n \quad \text{et} \quad H_n - 1 \leq \ln(n) - \ln(1)$$

Ce qui se réécrit :

$$\ln(n+1) \leq H_n \leq \ln(n) + 1$$

En particulier le théorème de comparaison justifie à ce point que $\lim_{n \rightarrow +\infty} H_n = +\infty$.

Si vous avez survécu jusque là, félicitations ! Le cadre de l'approfondissement proposé par le programme s'arrête ici. Mais si votre curiosité n'est pas encore rassasiée, on peut pousser le raisonnement précédent un peu plus loin pour faire apparaître une nouvelle constante mathématique.

Comme $\ln(n) > 0$, on peut diviser ces inégalités pour écrire $\frac{\ln(n+1)}{\ln(n)} \leq \frac{H_n}{\ln(n)} \leq 1 + \frac{1}{\ln(n)}$. Comme $\lim_{n \rightarrow +\infty} \ln(n) = +\infty$, la borne supérieure tend vers 1. Pour la borne inférieure, cette forme indéterminée requiert une petite astuce de calcul :

$$\frac{\ln(n+1)}{\ln(n)} = \frac{\ln(n+1) - \ln(n)}{\ln(n)} + 1 = \frac{\ln\left(1 + \frac{1}{n}\right)}{\ln(n)} + 1$$

Par continuité du logarithme en 1, le numérateur tend vers 0, donc la borne supérieure tend vers 1. Ainsi, d'après le théorème des gendarmes, $\lim_{n \rightarrow +\infty} \frac{H_n}{\ln(n)} = 1$. Autrement dit, quand n est grand, la suite (H_n) se comporte à peu près comme le logarithme !

Cette proximité est d'ailleurs encore plus profonde : l'écart entre la série harmonique et le logarithme tend vers une constante γ , appelée *constante d'Euler-Mascheroni*. Démontrons-le. Pour tout $n \in \mathbb{N}^*$, posons $u_n = H_n - \ln(n)$ et $v_n = u_n - \frac{1}{n}$.

- Premièrement, pour tout $n \in \mathbb{N}^*$, $u_n - v_n = \frac{1}{n} \geq 0$ donc $\lim_{n \rightarrow +\infty} u_n - v_n = 0$.
- Ensuite, pour tout $n \in \mathbb{N}^*$,

$$\begin{aligned} u_{n+1} - u_n &= [H_{n+1} - \ln(n+1)] - [H_n - \ln(n)] \\ &= [H_{n+1} - H_n] + [\ln(n+1) - \ln(n)] \\ &= \frac{1}{n+1} + \ln(n+1) - \ln(n) \\ &= \frac{1}{n+1} - \ln\left(1 - \frac{1}{n+1}\right) \end{aligned}$$

(Le lecteur est invité à détailler les étapes de calcul précédentes, ainsi que l'inégalité suivante.) Or par concavité du logarithme, pour tout $x > -1$, $\ln(1+x) \leq x$, donc $u_{n+1} - u_n \leq 0$, donc (u_n) est décroissante.

► Enfin, pour tout $n \in \mathbb{N}^*$,

$$\begin{aligned} v_{n+1} - v_n &= \left[u_{n+1} - \frac{1}{n+1} \right] - \left[u_n - \frac{1}{n} \right] \\ &= [u_{n+1} - u_n] + \left[\frac{1}{n} - \frac{1}{n+1} \right] \\ &= \frac{1}{n+1} + \ln(n+1) - \ln(n) + \frac{1}{n} - \frac{1}{n+1} \\ &= \frac{1}{n} - \ln\left(1 + \frac{1}{n}\right) \geq 0 \end{aligned}$$

par le même argument de concavité que précédemment. Ainsi, $v_{n+1} - v_n \geq 0$, donc (v_n) est croissante.

On a donc montré que (u_n) et (v_n) sont adjacentes, donc d'après le théorème des suites adjacentes*, elles convergent toutes les deux vers une limite commune que l'on note γ , la constante d'Euler-Mascheroni.

$$\lim_{n \rightarrow +\infty} H_n - \ln(n) = \gamma \approx 0,5772156649$$

2 – Intégration numérique : méthode de Monte-Carlo

On a vu en introduction que la méthode des rectangles ou des trapèzes permet de créer une suite approchant la valeur d'une intégrale recherchée, et que cette suite est facile à implémenter. Il existe une autre méthode, probabiliste : la méthode de Monte-Carlo. Pour expliquer son principe, voici une analogie : imaginez que vous êtes si mauvais aux fléchettes que lorsque vous lancez, vous tirez aléatoirement dans le cadre carré sur lequel vous avez accroché votre cible. La méthode de Monte-Carlo part alors du principe que en proportion, le nombre de fléchettes qui ont touché la cible par rapport au total est proche rapport de l'aire de la cible par rapport au cadre. Pour l'intégration numérique c'est la même chose : on enferme la courbe dans un rectangle, d'aire connue, on tire un grand nombre de points dans ce rectangle, et on mesure la proportion qui tombe sous la courbe.

On rappelle que la courbe d'une fonction f a pour équation $y = f(x)$. Autrement dit, les points de la courbe sont ceux pour lesquels l'ordonnée y est exactement égale à $f(x)$. Si un point est en-dessous de la courbe, c'est donc que son ordonnée y est inférieure à $f(x)$: cela donne un critère facile à vérifier pour savoir si un point est tombé au bon endroit ou non.

Les codes ci-dessous peuvent paraître compliqués, mais cela vient de la volonté de leur faire afficher un joli résultat, illustré plus bas. La logique est pourtant simple :

- On délimite le rectangle de tirage (notons, par exemple, les abscisses entre a et b , et les ordonnées entre c et d),
- On tire un nombre aléatoire entre a et b , un autre en c et d , pour donner les coordonnées d'un point aléatoire dans le rectangle,
- On vérifie si le point tiré est dans la région désirée ou non (un disque dans le second exemple, sous la courbe de la fonction dans le premier).
 - Si oui, on stocke les coordonnées dans la liste des bons points (colorés en vert),
 - Si non, on stocke les coordonnées dans la liste des mauvais points (colorés en rouge),
- On calcule le nombre de bons points par rapport au total, ce qui permet de retrouver l'aire recherchée par proportionnalité.

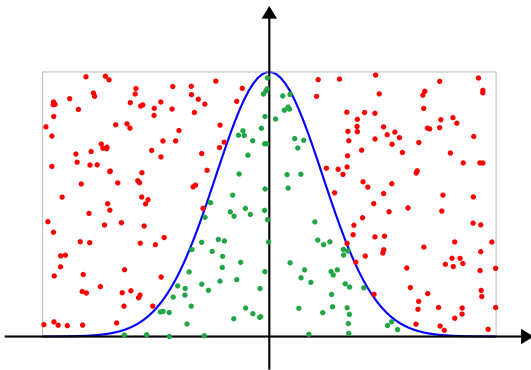
*. Voir la démonstration du théorème des valeurs intermédiaires dans le chapitre 8 pour la démonstration de ce théorème.

AIRE SOUS UNE COURBE

```

1 import numpy as np
2 from matplotlib import pyplot as plt
3
4 def alea_bornes(a, b):
5     """ Réel aléatoire de [a,b]. """
6     return a + (b-a)*np.random.random()
7
8 def f(x):
9     """ Fonction à intégrer. """
10    return np.exp(-x**2)
11
12 N = 10000 # Nombre d'essais
13
14 Xok, Yok = [], [] # Points dedans
15 Xrt, Yrt = [], [] # Points dehors
16 for _ in range(N):
17     x = alea_bornes(-5,5)
18     y = alea_bornes(0,1)
19     if y < f(x):
20         Xok.append(x)
21         Yok.append(y)
22     else:
23         Xrt.append(x)
24         Yrt.append(y)
25
26 _ = plt.scatter(Xok,Yok,s=2.,c="green")
27 _ = plt.scatter(Xrt,Yrt,s=.8,c="red")

```

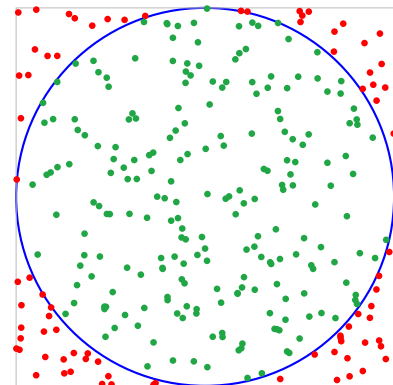


AIRE D'UN DISQUE

```

1 import numpy as np
2 from matplotlib import pyplot as plt
3
4 def alea_bornes(a, b):
5     """ Réel aléatoire de [a,b]. """
6     return a + (b-a)*np.random.random()
7
8
9
10
11
12 N = 10000 # Nombre d'essais
13 r = 1. # Rayon du cercle
14 Xok, Yok = [], [] # Points dedans
15 Xrt, Yrt = [], [] # Points dehors
16 for _ in range(N):
17     x = alea_bornes(-r,r)
18     y = alea_bornes(-r,r)
19     if x**2 + y**2 <= r**2:
20         Xok.append(x)
21         Yok.append(y)
22     else:
23         Xrt.append(x)
24         Yrt.append(y)
25
26 _ = plt.scatter(Xok,Yok,s=2.,c="green")
27 _ = plt.scatter(Xrt,Yrt,s=.8,c="red")

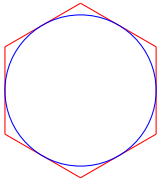
```



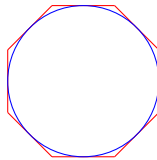
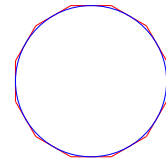
3 – Intégrales, longueurs et volumes

A – Intro-digression : l'approximation de π par Archimède

Le nombre π est défini comme le coefficient de proportionnalité entre le diamètre et la circonférence d'un cercle. Cette définition ne permet pas de donner une valeur de π : on a besoin d'un pont entre le monde géométrique dans lequel est défini π , et le monde numérique. L'idée d'Archimède était la suivante : on sait calculer le périmètre d'un polygone, comme somme des longueurs des segments qui constituent ses côtés. On sait également que la circonférence d'un cercle de rayon 1 est 2π . Ainsi, pour estimer π , il suffit de trouver un polygone qui ressemble le plus possible à un cercle !

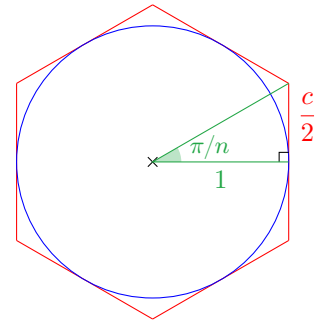


Avec un hexagone...

Avec un octogone...
un peu mieux !Avec un dodécagone...
encore mieux !

Ces trois exemples nous font comprendre qu'il y a pas de « meilleur » polygone pour approcher un cercle... On devine que plus le polygone a de côtés, meilleure sera l'approximation... Autrement dit, plus le nombre de côtés du polygone devient grand, plus son périmètre se rapproche du périmètre... c'est une limite ! Ainsi, la limite du périmètre du polygone tend vers 2π lorsque le nombre de côtés tend vers $+\infty$. Si l'on pouvait calculer le périmètre du polygone... on aurait une suite qui permettrait de calculer des valeurs approchées de π !

Considérons un polygone régulier à n côtés tel que le cercle unité est inscrit en son intérieur. On note c la longueur de ses côtés. Le polygone admet n sommets, donc divise le tour complet en n secteurs. Ainsi, l'angle au centre formé entre deux sommets consécutifs est $\frac{2\pi}{n}$. Le cercle étant inscrit dans le polygone, les côtés de ce dernier sont tangents au cercle. Quelques arguments géométriques élémentaires montrent que le point d'intersection entre un côté et le cercle est le milieu du côté.



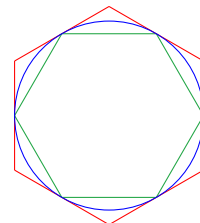
Par tangence, celui-ci est perpendiculaire au rayon du cercle correspondant, donc dans le triangle rectangle représenté ci-contre, $\tan(\pi/n) = \frac{c/2}{1} = \frac{c}{2}$. Ainsi, le périmètre du polygone est $n \times 2 \times c = 2n \tan(\pi/n)$. La limite de cette expression est indéterminée : en effet, $\frac{\pi}{n}$ tend vers 0, or la fonction tangente est continue en 0, donc $\lim_{n \rightarrow +\infty} \tan(\pi/n) = \tan(0) = 0$. L'astuce est d'écrire :

$$n \tan(\pi/n) = \pi \frac{\tan(\pi/n) - 0}{\frac{\pi}{n} - 0} = \pi \frac{\tan(x) - \tan(0)}{x - 0}$$

en écrivant $x = \frac{\pi}{n}$. Lorsque n tend vers $+\infty$, par quotient x tend vers 0 : on reconnaît alors la limite du taux d'accroissement de $\tan$ en 0, qui est $\tan'(0) = 1$ puisque $\tan$ est dérivable en 0. Ainsi, $\lim_{n \rightarrow +\infty} n \tan(\pi/n) = \pi$, ce qui prouve que le périmètre du polygone tend bien vers le périmètre du cercle.

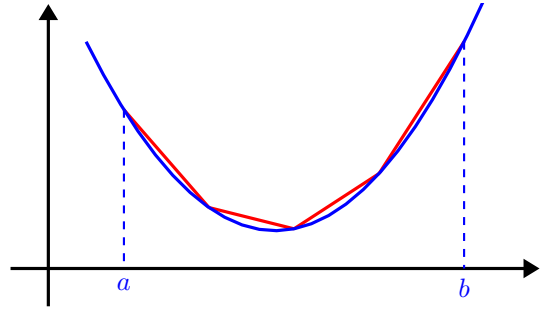
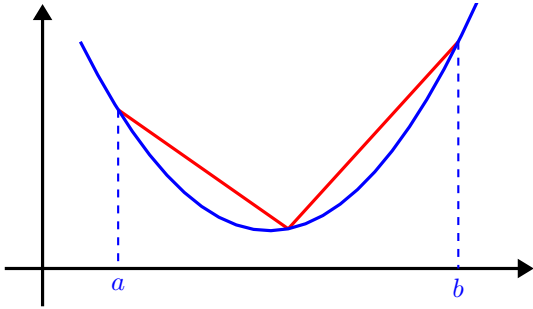
En réalité, l'approche d'Archimède consiste à encadrer le périmètre du cercle par deux polygones, l'un inscrit dans le cercle et l'autre inscrivant le cercle.

Exercice : mener la même étude que précédemment pour le polygone inscrit ! Cela donnait à Archimède un encadrement de π , ce qui permet de quantifier la qualité de son approximation.



B – Intégrales curvilignes

L'idée de Riemann pour déterminer l'aire sous une courbe était d'approcher celle-ci par des rectangles toujours plus fins. L'idée d'Archimède était d'approcher le cercle par des polygones avec toujours plus de côtés. Peut-on fusionner ces deux idées pour calculer la longueur de la courbe d'une fonction entre deux abscisses ? Essayons !



Suivant les pas d'Archimède et de Riemann, il est raisonnable de penser que cette méthode permet bien d'approcher la longueur de la courbe, et plus précisément que cette longueur est la limite de la longueur de la ligne brisée que l'on construit avec une subdivision de plus en plus fine.

La question maintenant est : peut-on calculer la longueur de cette ligne brisée, et sa limite ? Zoomons un peu :

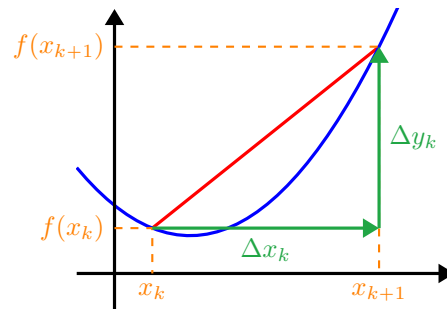
Si on note x_k et x_{k+1} les abscisses des deux extrémités d'un segment de la ligne brisée, alors, d'après le théorème de Pythagore, sa longueur est $\sqrt{(\Delta x_k)^2 + (\Delta y_k)^2}$, où $\Delta x_k = x_{k+1} - x_k$ et $\Delta y_k = f(x_{k+1}) - f(x_k)$. Remarquons que :

$$\sqrt{(\Delta x_k)^2 + (\Delta y_k)^2} = \Delta x_k \sqrt{1 + \left(\frac{\Delta y_k}{\Delta x_k}\right)^2}$$

(note : $\Delta x_k > 0$ car on considère les x_k dans l'ordre croissant).

Or, si f est dérivable, alors :

$$\lim_{\Delta x_k \rightarrow 0} \frac{\Delta y_k}{\Delta x_k} = \lim_{\Delta x_k \rightarrow 0} \frac{f(x_{k+1}) - f(x_k)}{\Delta x_k} = \lim_{\Delta x_k \rightarrow 0} \frac{f(x_k + \Delta x_k) - f(x_k)}{\Delta x_k} = f'(x_k)$$



Supposons l'espacement de la subdivision régulier, c'est-à-dire $\Delta x_k = h$ pour tout k . Pour résumer : la longueur totale de la ligne brisée est (pour plus de détails, revoir l'introduction du chapitre) :

$$\sum_{k=1}^n \sqrt{1 + \left(\frac{\Delta y_k}{\Delta x_k}\right)^2} \times \Delta x_k \quad \text{c'est-à-dire} \quad \sum_{k=1}^n g_k(x_k) \Delta x_k$$

où $g_k : x \mapsto \sqrt{1 + \left(\frac{f(x+h) - f(x)}{h}\right)^2}$. Dans la méthode des rectangles de Riemann, on avait écrit $\sum_{k=1}^n f(x_k) \Delta x_k$.

Dans ce cas, la fonction est la même, indépendamment de la subdivision. Ici, g_k varie à chaque k . On admet que la méthode de Riemann fonctionne de la même manière avec la limite de $g_k(x)$, pour k tendant vers $+\infty$; que l'on a établi comme étant $\sqrt{1 + f'(x)^2}$. Finalement, la longueur de la portion de courbe est donnée par :

$$\int_a^b \sqrt{1 + f'(x)^2} \, dx$$

Exemple 1

Soit $f : x \mapsto \sqrt{1-x^2}$. Elle est dérivable sur $\left[0; \frac{\sqrt{2}}{2}\right]$ (comme somme et composée), et $f'(x) = \frac{-2x}{2\sqrt{1-x^2}} = \frac{-x}{\sqrt{1-x^2}}$. D'après le formule précédente, la longueur de la courbe de f entre 0 et $\frac{\sqrt{2}}{2}$ est :

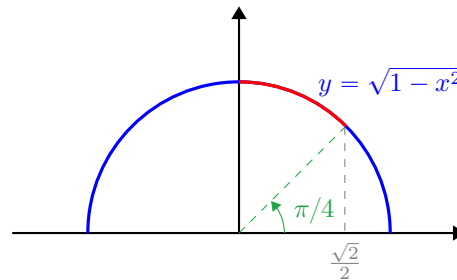
$$\int_0^{\sqrt{2}/2} \sqrt{1 + \left(\frac{-x}{\sqrt{1-x^2}}\right)^2} dx = \int_0^{\sqrt{2}/2} \sqrt{1 + \frac{x^2}{1-x^2}} dx = \int_0^{\sqrt{2}/2} \frac{1}{\sqrt{1-x^2}} dx$$

En posant le changement de variable $x = \sin(y)$ (donc $y = \arcsin(x)$) comme développé dans la section suivante, on obtient :

$$\int_0^{\sqrt{2}/2} \frac{1}{\sqrt{1-x^2}} dx = \int_{\pi/4}^{\pi/2} \frac{1}{\sqrt{1-\sin(y)^2}} \cos(y) dy = \int_{\pi/4}^{\pi/2} \frac{1}{\cos(y)} \cos(y) dy = 1 \left(\frac{\pi}{2} - \frac{\pi}{4} \right) = \frac{\pi}{4}$$

Prenons un peu de recul... Soit $y = f(x) = \sqrt{1-x^2}$. Alors $y^2 = 1-x^2$, donc $x^2 + y^2 = 1$: c'est l'équation du cercle unité !

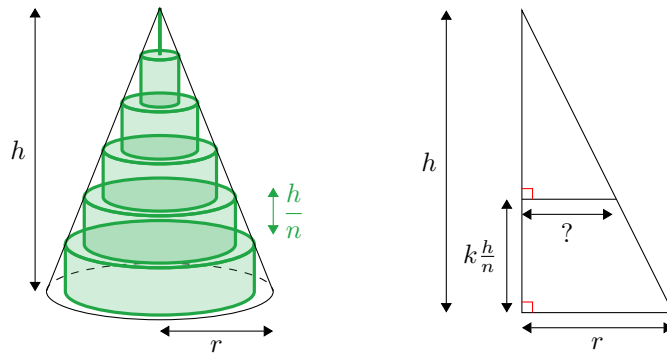
La longueur que l'on a calculée est celle du huitième de cercle entre $\frac{\pi}{4}$ et $\frac{\pi}{2}$: de notre calcul, on en déduit que la circonférence du cercle est $8 \times \frac{\pi}{4} = 2\pi$... sans surprise !



C – Intégrales et volumes

On a calculé des longueurs, des aires... qu'en est-il des volumes ? Spoiler : une intégrale bien pensée fera l'affaire !

Toujours dans la même idée que Riemann, on découpe notre volume en des formes élémentaires, d'épaisseur devenant de plus en plus fine. Pour s'exercer, considérons un cône de rayon r et de hauteur h . On y insère n tranches cylindriques, d'épaisseur $\frac{h}{n}$, et de rayon ? = $r \left(1 - \frac{k}{n}\right)$ d'après le théorème de Thalès appliqué dans une tranche verticale du cône.



Le volume de la k -ième tranche est donc $\frac{h}{n} \times \pi r^2 \left(1 - \frac{k}{n}\right)^2$, et le volume des n tranches est la somme du volume de toutes les tranches :

$$\sum_{k=1}^n \frac{h}{n} \times \pi r^2 \left(1 - \frac{k}{n}\right)^2 = h\pi r^2 \sum_{k=1}^n \left(1 - \frac{k}{n}\right)^2 \times \frac{1}{n}$$

Soit $f : x \mapsto (1-x)^2$. Pour k entre 1 et n , notons $x_k = \frac{k}{n}$. On a $\Delta x_k = x_{k+1} - x_k = \frac{1}{n}$. Ainsi, on retrouve dans

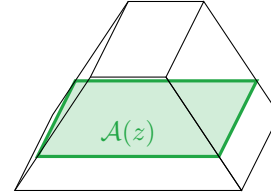
l'expression précédente la méthode des rectangles de Riemann de f entre 0 et 1 :

$$\sum_{k=1}^n \left(1 - \frac{k}{n}\right)^2 \times \frac{1}{n} = \sum_{k=1}^n f(x_k) \times \Delta x_k \xrightarrow[n \rightarrow +\infty]{\text{Riemann}} \int_0^1 f(x) dx = \int_0^1 (1-x)^2 dx = \left[-\frac{(1-x)^3}{3}\right]_0^1 = \frac{1}{3}$$

En conclusion, comme l'aire de la base est $\mathcal{A}_{\text{base}} = \pi r^2$, on retrouve la formule du volume d'un cône : $\mathcal{V} = \frac{1}{3} \mathcal{A}_{\text{base}} h$.

En pratique, on peut simplifier le calcul : on peut montrer que si un solide s'étend de $z = a$ à $z = b$, et si l'on note $\mathcal{A}(z)$ l'aire de la section du solide à la cote z , alors le volume du solide est donné par :

$$\mathcal{V} = \int_a^b \mathcal{A}_z dz$$



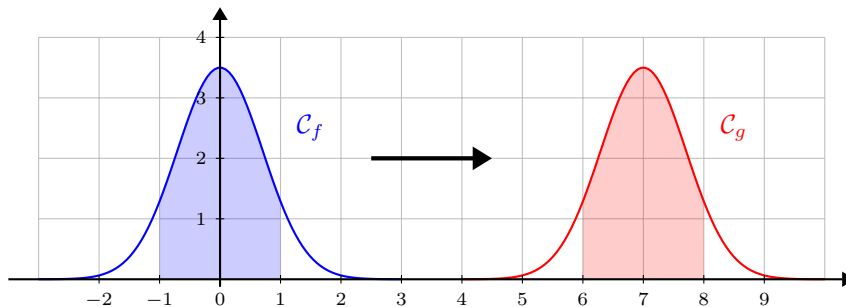
Application 2 (Volume d'une boule)

On considère une boule de rayon r . Pour le reste des calculs, on considérera la demi-boule formée par son hémisphère nord.

- 1► À quel intervalle la cote des points de ce solide appartient-elle ?
- 2► Tracer une tranche verticale de cette demi-boule, et faire figurer sur ce schéma la section de cette demi-boule par un plan horizontale de cote z .
- 3► Montrer que le carré du rayon de cette section est $r^2 - z^2$.
- 4► Quelle est la nature de cette section ? En déduire son aire.
- 5► En déduire que le volume de l'hémisphère nord est $\frac{2}{3}\pi r^3$.
- 6► En déduire le volume de la boule.

4 – Changement de variable

Soit f une fonction positive. L'intégrale de f entre -1 et 1 correspond alors à l'aire sous la courbe de f entre les abscisses -1 et 1 . Soit $g : x \mapsto f(x-7)$. La courbe de g est identique à celle de f , à la différence près que la courbe de g est translatée de 7 unités vers la droite par rapport à celle de f . Sur le schéma suivant, l'aire en rouge est égale à l'aire en bleu, car les translations préservent les aires. Cependant l'aire en rouge n'est pas intégrale de g entre -1 et 1 : c'est l'intégrale de g entre 6 et 8. Autrement dit, les bornes d'intégration ont elle aussi été translatées.



Ainsi, on peut écrire :

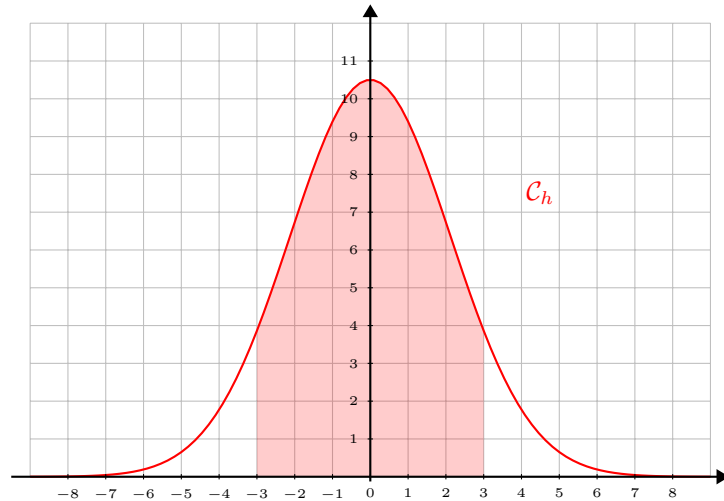
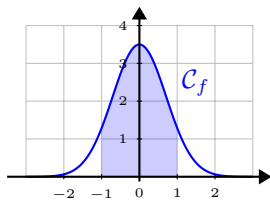
$$\int_{-1}^1 f(x) dx = \int_6^8 g(x) dx = \int_6^8 f(x-7) dx$$

De manière générale, pour tous réels a , b et k tels que $b > a$,

$$\int_a^b f(x-k) dx = \int_{a+k}^{b+k} f(y) dy$$

On dit qu'on a effectué le *changement de variable* $y = x - 7$, que l'on peut aussi écrire $x = y + 7$.

Intéressons-nous à une autre situation. Considérons $h : x \mapsto 3h(x/3)$. Sa courbe représentative est identique à celle de f , mais agrandie d'un facteur 3. On sait que, lors d'un agrandissement de facteur 3, les aires sont multipliées par $3^2 = 9$; sur le schéma ci-dessous, l'aire en rouge est donc 9 fois plus grande que la bleue.



L'aire bleue est mesurée par $\int_{-1}^1 f(x) \, dx$. L'agrandissement de facteur 3 triple les bornes de l'intégrale mesurant l'aire rouge : celle-ci est donc $\int_{-3}^3 h(x) \, dx$. Ainsi, par linéarité de l'intégrale :

$$\int_{-3}^3 3f(x/3) \, dx = \int_{-3}^3 h(x) \, dx = 9 \int_{-1}^1 f(x) \, dx = \int_{-1}^1 9f(x) \, dx$$

En divisant les membres les plus à droite et à gauche par 3, en réarrangeant un peu l'écriture, on obtient :

$$\int_{-3}^3 f(x/3) \, dx = \int_{-1}^1 f(x) \times 3 \, dx$$

Par un même raisonnement, on montre que pour tous réels a et b tels que $b > a$, pour tout $\lambda > 0$,

$$\int_{\lambda a}^{\lambda b} f(x/\lambda) \, dx = \int_a^b f(y) \times \lambda \, dy$$

On dit qu'on a effectué le changement de variable $y = \frac{x}{\lambda}$, que l'on peut aussi écrire $x = \lambda y$.

Pourquoi dans le second cas, y a-t-il une multiplication par 3 en plus ? Quelle différence fondamentale y a-t-il entre les deux situations ? Si, dans le premier cas, on écrit $1 \, dy$ à la place de dy tout seul, on commence à deviner qu'il faut trouver un procédé qui lie $y + 7$ à 1, et $3y$ à 3... c'est la dérivée ! De manière plus générale que ces deux exemples, si l'on dérive $y + k$ comme une fonction de y , on obtient 1. Si l'on dérive λy comme une fonction de y , on obtient λ . On retrouve dans chaque cas la dérivée de l'expression de x en fonction de la nouvelle variable. Cela se généralise :

Théorème 3

Changement de variable

Soient f une fonction continue, φ une fonction dérivable telle que g' est continue, et a et b deux réels.

$$\int_a^b f(\varphi(x)) \times \varphi'(x) \, dx = \int_{\varphi(a)}^{\varphi(b)} f(y) \, dy$$

On dit qu'on a effectué le changement de variable $y = \varphi(x)$.

Démonstration. Comme f est continue, elle admet au moins une primitive F d'après le théorème fondamental de l'intégration. Soit $g = F \circ \varphi$. Cette fonction est dérivable comme composée, et de dérivée continue. Par ailleurs, pour tout $x \in [a, b]$, $g'(x) = F'(\varphi(x)) \times \varphi'(x) = f(\varphi(x)) \times \varphi'(x)$. Or g est une primitive de g' donc :

$$\int_a^b f(\varphi(x)) \times \varphi'(x) \, dx = \int_a^b g'(x) \, dx = [g(x)]_a^b = g(b) - g(a) = F(\varphi(b)) - F(\varphi(a))$$

D'autre part,

$$\int_{\varphi(a)}^{\varphi(b)} f(y) \, dy = [F(y)]_{\varphi(a)}^{\varphi(b)} = F(\varphi(b)) - F(\varphi(a))$$

ce qui montre le théorème. □

Exemple 4 (Présentation-type)

Calculons $\int_1^4 \frac{1 - \sqrt{x}}{\sqrt{x}} \, dx$ avec le changement de variable $y = \sqrt{x}$.

Pour x entre 1 et 4, $y = \sqrt{x} \iff x = y^2$ (autrement dit $\varphi(x) = \sqrt{x}$). On peut appliquer la formule apprise par cœur et avec un peu d'astuce :

$$\int_1^4 \frac{1 - \sqrt{x}}{\sqrt{x}} \, dx = \int_1^4 2(1 - \sqrt{x}) \times \frac{1}{2\sqrt{x}} \, dx = \int_{\sqrt{1}}^{\sqrt{4}} 2(1 - y) \, dy = [-(1 - y)^2]_1^2 = -1$$

On peut aussi utiliser une présentation un peu standard regroupant des astuces mnémotechniques :

$$\begin{array}{c|c} x & y = \sqrt{x} \\ \hline 4 & 2 \\ 1 & 1 \end{array}$$

$$\frac{dx}{dy} = \frac{d}{dy}[y^2] = 2y \quad \blacktriangle \text{ donc } \blacktriangle \, dx = 2y dy.$$

(**⚠** Attention, ce passage est abusif, car la notation $\frac{dx}{dy}$ n'est pas une écriture fractionnaire, c'est une notation !

Toutefois, cette astuce fonctionne dans le cadre d'un changement de variable.)

Ainsi :

$$\int_1^4 \frac{1 - \sqrt{x}}{\sqrt{x}} \, dx = \int_1^2 \frac{1 - \sqrt{y^2}}{\sqrt{y^2}} \times 2y \, dy = \int_1^2 2(1 - y) \, dy = [-(1 - y)^2]_1^2 = -1$$

13 Variables aléatoires

Table des matières

1	Covariance et corrélation	65
2	Intervalle de confiance	68

1 – Covariance et corrélation

Nous avons vu que la variance ne se comporte bien qu'avec les sommes que sous l'hypothèse d'indépendance. Démontrons la relation du cours pour s'en convaincre.

Avant cela, nous devons démontrer une relation centrale de l'indépendance de variables aléatoires :

Lemme 1

Si X et Y sont deux variables aléatoires indépendantes, alors $\mathbb{E}[XY] \stackrel{II}{=} \mathbb{E}[X] \mathbb{E}[Y]$.

Démonstration. Notons $x_1, x_2, \dots, x_n$ les valeurs prises par X , et $y_1, y_2, \dots, y_m$ les valeurs prises par Y . On rappelle que l'espérance d'une variable aléatoire est la somme des produits des valeurs par la probabilité de valoir cette valeur, donc :

$$\mathbb{E}[X] = \sum_{i=1}^n x_i \mathbb{P}(X = x_i) \quad \text{et} \quad \mathbb{E}[Y] = \sum_{j=1}^m y_j \mathbb{P}(Y = y_j)$$

La question que l'on se pose maintenant est : comment calculer $\mathbb{E}[XY]$? Analysons la variable aléatoire XY pour mieux le comprendre. Prenons un exemple : imaginons que X et Y ont pour lois :

x_i	2	3	6
$\mathbb{P}(X = x_i)$	0,1	0,5	0,4

y_j	4	5	6
$\mathbb{P}(Y = y_j)$	0,2	0,3	0,5

Pour connaître la loi de XY , on peut commencer par déterminer ses valeurs en traçant la table de multiplication de X et Y :

	2	3	6
4	8	12	24
5	10	15	30
6	12	18	36

Pour connaître les probabilités, on peut par exemple constater que 8 n'apparaît que dans un cas, donc $\mathbb{P}(XY = 8) = \mathbb{P}(\{X = 2\} \cap \{Y = 4\})$, intersection que l'on note $\mathbb{P}(X = 2, Y = 4)$ dans un souci de lisibilité. Comme 12 apparaît deux fois, on peut écrire :

$$\mathbb{P}(XY = 12) = \mathbb{P}(X = 2, Y = 6) + \mathbb{P}(X = 3, Y = 4)$$

Remarquons cependant que nous n'aurons pas besoin de nous préoccuper de ces cas de recouvrement : en effet, nous cherchons à calculer l'espérance, nous allons donc calculer :

$$\begin{aligned} 12\mathbb{P}(XY = 12) &= 12(\mathbb{P}(X = 2, Y = 6) + \mathbb{P}(X = 3, Y = 4)) \\ &= 12\mathbb{P}(X = 2, Y = 6) + 12\mathbb{P}(X = 3, Y = 4) \end{aligned}$$

Ainsi, on se convainc que :

$$\begin{aligned}\mathbb{E}[XY] &= x_1 y_1 \mathbb{P}(X = x_1, Y = y_1) + x_1 y_2 \mathbb{P}(X = x_1, Y = y_2) + \cdots + x_1 y_m \mathbb{P}(X = x_1, Y = y_m) \\ &\quad + x_2 y_1 \mathbb{P}(X = x_2, Y = y_1) + x_2 y_2 \mathbb{P}(X = x_2, Y = y_2) + \cdots + x_2 y_m \mathbb{P}(X = x_2, Y = y_m) \\ &\quad + \cdots \\ &\quad + x_n y_1 \mathbb{P}(X = x_n, Y = y_1) + x_n y_2 \mathbb{P}(X = x_n, Y = y_2) + \cdots + x_n y_m \mathbb{P}(X = x_n, Y = y_m)\end{aligned}$$

En ce qui concerne $\mathbb{E}[X] \mathbb{E}[Y]$, il s'agit d'une banale distributivité, prenez votre temps pour la comprendre (on distribue le premier terme de X sur chacun de Y , puis le deuxième, etc.) :

$$\begin{aligned}\mathbb{E}[XY] &= x_1 y_1 \mathbb{P}(X = x_1) \mathbb{P}(Y = y_1) + x_1 y_2 \mathbb{P}(X = x_1) \mathbb{P}(Y = y_2) + \cdots + x_1 y_m \mathbb{P}(X = x_1) \mathbb{P}(Y = y_m) \\ &\quad + x_2 y_1 \mathbb{P}(X = x_2) \mathbb{P}(Y = y_1) + x_2 y_2 \mathbb{P}(X = x_2) \mathbb{P}(Y = y_2) + \cdots + x_2 y_m \mathbb{P}(X = x_2) \mathbb{P}(Y = y_m) \\ &\quad + \cdots \\ &\quad + x_n y_1 \mathbb{P}(X = x_n) \mathbb{P}(Y = y_1) + x_n y_2 \mathbb{P}(X = x_n) \mathbb{P}(Y = y_2) + \cdots + x_n y_m \mathbb{P}(X = x_n) \mathbb{P}(Y = y_m)\end{aligned}$$

La correspondance est alors frappante : pour conclure, il suffit de dire que :

$$\mathbb{P}(X = x_i, Y = y_j) = \mathbb{P}(X = x_i) \mathbb{P}(Y = y_j)$$

pour chaque i et j . Est-ce que ça ne vous rappelle rien... ? N'oubliez pas que $\mathbb{P}(X = x_i, Y = y_j)$ est une notation qui désigne $\mathbb{P}(\{X = x_i\} \cap \{Y = y_j\})$. Ainsi, on voudrait que la probabilité de l'intersection soit égale au produit des probabilités... c'est la définition de l'indépendance, qui par le plus grand des hasards n'a pas été utilisée jusque là !

Ainsi l'indépendance se présente comme la clef de voûte de cette relation. □

Forts de cette nouvelle relation, nous pouvons désormais étudier $\mathbb{V}[X + Y]$. Afin de gagner en lisibilité, on note $\tilde{X} = X - \mathbb{E}[X]$ et $\tilde{Y} = Y - \mathbb{E}[Y]$. Remarquons qu'alors, par définition de la variance d'une variable aléatoire, $\mathbb{V}[X] = \mathbb{E}[\tilde{X}^2]$ et $\mathbb{V}[Y] = \mathbb{E}[\tilde{Y}^2]$.

$$\begin{aligned}\mathbb{V}[X + Y] &= \mathbb{E}[(X + Y - \mathbb{E}[X + Y])^2] \\ &= \mathbb{E}[(\tilde{X} + \tilde{Y})^2] \\ &= \mathbb{E}[\tilde{X}^2 + \tilde{Y}^2 + 2\tilde{X}\tilde{Y}] \\ &= \mathbb{E}[\tilde{X}^2] + \mathbb{E}[\tilde{Y}^2] + 2\mathbb{E}[\tilde{X}\tilde{Y}] \quad (\text{par linéarité de l'espérance}) \\ &= \mathbb{V}[X] + \mathbb{V}[Y] + 2\mathbb{E}[\tilde{X}\tilde{Y}]\end{aligned}$$

On remarque donc que la variance de la somme est égale à la somme des variances, à un terme parasite près...

On se convainc que si X et Y sont indépendantes, alors décaler l'une et l'autre d'une constante ne change en rien cette indépendance, et donc $\tilde{X}$ et $\tilde{Y}$ sont également indépendantes. D'après le lemme, on a donc :

$$\mathbb{E}[\tilde{X}\tilde{Y}] \stackrel{\text{II}}{=} \mathbb{E}[\tilde{X}] \mathbb{E}[\tilde{Y}] = \mathbb{E}[X - \mathbb{E}[X]] \mathbb{E}[Y - \mathbb{E}[Y]]$$

Or par linéarité de l'espérance, $\mathbb{E}[X - \mathbb{E}[X]] = \mathbb{E}[X] - \mathbb{E}[\mathbb{E}[X]] = \mathbb{E}[X] - \mathbb{E}[X] = 0$ (on rappelle que l'espérance d'une variable aléatoire est un nombre, et l'espérance d'un nombre est ce nombre lui-même). Ainsi, $\mathbb{E}[\tilde{X}\tilde{Y}] = 0$, d'où la formule.

Cette espérance parasite porte un nom :

Définition 2**Covariance**

La *covariance* de X et Y et le nombre :

$$\text{Cov}(X, Y) = \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])]$$

On s'en sert notamment pour mesurer la corrélation de deux variables aléatoires :

Définition 3**Corrélation**

Le *coefficient de corrélation linéaire* des variables aléatoires X et Y de variance non nulle est le nombre :

$$r(X, Y) = \frac{\text{Cov}(X, Y)}{\sigma(X)\sigma(Y)}$$

Propriété 4

$$r(X, Y) \in [-1; 1]$$

Le coefficient de corrélation linéaire de deux variables aléatoires quantifie la distinction suivante :

- Plus il est proche des extrêmes (1 ou -1), plus on est en droit d'écrire une relation affine $Y = aX + b$ (le signe de r donnant le sens de variation de cette fonction affine),
- Plus il est proche de 0, moins une telle relation est plausible.

On a vu au fil des démonstrations que deux variables indépendantes ont une covariance nulle, *a fortiori* un coefficient de corrélation linéaire nul. On dit alors que deux variables indépendantes sont *non corrélées*. La réciproque est cependant fautive ! En effet, deux variables non corrélées ne sont pas forcément indépendantes. Un contre-exemple abordable est le suivant : on considère U et V indépendantes suivant une loi de Bernoulli de paramètre 0,5, et on pose $X = U + V$ et $Y = U - V$. Remarquons rapidement avant de détailler le contre-exemple que par linéarité de l'espérance, $\text{Cov}(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$. La loi de Bernoulli a ceci de particulier que U^2 et V^2 suivant également la loi $\mathcal{B}(p)$, puisque prenant les mêmes valeurs avec les mêmes probabilités. Par conséquent, $\mathbb{E}[U^2] = \mathbb{E}[U] = \mathbb{E}[V^2] = \mathbb{E}[V] = 0,5$. Ainsi, par linéarité de l'espérance,

$$\begin{aligned} \text{Cov}(X, Y) &= \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] \\ &= \mathbb{E}[(U + V)(U - V)] - \mathbb{E}[U + V]\mathbb{E}[U - V] \\ &= \mathbb{E}[U^2] - \mathbb{E}[V^2] - \mathbb{E}[U]^2 + \mathbb{E}[V]^2 \\ &= 0,5 - 0,5 - 0,5^2 + 0,5^2 \\ &= 0 \end{aligned}$$

Et dont X et Y ne sont pas corrélées. En revanche, elles ne sont pas indépendantes. En effet, montrons que $\mathbb{P}(X = 0, Y = 0) \neq \mathbb{P}(X = 0)\mathbb{P}(Y = 0)$.

Commençons par le premier membre. Si $X = Y = 0$, c'est-à-dire $U + V = 0$ et $U - V = 0$, alors $(U + V) + (U - V) = 0$, c'est-à-dire $2U = 0$, donc $U = 0$, par conséquent $V = 0$ également. Par indépendance de U et V ,

$$\mathbb{P}(X = 0, Y = 0) = \mathbb{P}(U = 0, V = 0) = \mathbb{P}(U = 0)\mathbb{P}(V = 0) = 0,5 \times 0,5 = 0,25$$

Concernant le second membre, le seul moyen pour que $(X = 0) \wedge (Y = 0)$ est que $U = 0$ et $V = 0$, toujours par indépendance on a donc $\mathbb{P}(U + V = 0) = \mathbb{P}(U = 0)\mathbb{P}(V = 0) = 0,25$. En revanche, il y a deux possibilités pour

que $(Y = 1)U - V = 0$: soit $U = V = 0$, soit $U = V = 1$. Par un calcul similaire, on montre que $\mathbb{P}(U - V = 0) = 0,5$. Ainsi, $\mathbb{P}(X = 0)\mathbb{P}(Y = 0) = 0,25 \times 0,5 = 0,125$ ce qui est bien différent. Ainsi X et Y ne sont pas indépendantes.

2 – Intervalles de confiance

Dans le lointain pays d'Algérie, deux candidates s'affrontent pour l'élection de la nouvelle cheffe du gouvernement : Ada Noether et Emmy Lovelace. L'institut Germain a réalisé une enquête d'intention de vote, auprès de la population : sur 1 000 sondés, 52% déclarent vouloir voter pour Noether. Malgré ce résultat favorable, Noether ne proclame pas victoire, et redouble d'efforts. Pourquoi se méfier, alors que d'après le sondage, plus de la moitié de la population semble vouloir voter pour elle ?

Soit $(X_1, \dots, X_n)$ un échantillon d'électeurs sondés. On suppose la population d'Algérie suffisamment grande pour assimiler le sondage à un tirage avec remise, si bien que les réponses des sondés sont indépendantes et de même loi. Si on note 1 lorsqu'un sondé annonce voter pour Noether, et 0 sinon, alors chaque X_i suit une loi de Bernoulli de paramètre p , la proportion réelle d'électeurs et électrices qui désirent voter pour Noether.

Notons S_n la variable aléatoire qui compte le nombre d'intentions de votes pour Noether dans notre échantillon : d'après les hypothèses susmentionnées, cela revient à compter le nombre de succès dans un schéma de Bernoulli de paramètres n et p , donc $S_n \sim \mathcal{B}(n, p)$. En particulier, $\mathbb{E}[S_n] = np$ et $\mathbb{V}[S_n] = np(1-p)$.

Notons $M_n = \frac{S_n}{n}$: c'est la proportion de l'échantillon qui déclare souhaiter voter pour Noether. Dans notre situation : $M_{1000} = 0,52$. Par linéarité de l'espérance, $\mathbb{E}[M_n] = \frac{S_n}{n} = p$. Pour la variance : $\mathbb{V}[M_n] = \frac{1}{n^2}\mathbb{V}[S_n] = \frac{p(1-p)}{n}$. D'après l'inégalité de Bienaymé-Tchebychev, pour tout $a > 0$,

$$\mathbb{P}(|M_n - \mathbb{E}[M_n]| \geq a) \leq \frac{\mathbb{V}[M_n]}{a^2} \quad \text{c'est-à-dire} \quad \mathbb{P}(|M_n - p| \geq a) \leq \frac{p(1-p)}{na^2}$$

Remarquons que $p(1-p) \leq \frac{1}{4}$: en effet, la fonction $f : x \mapsto x(1-x) = -x^2 + x$ est une fonction polynomiale de degré 2, de racines 0 et 1 et de coefficient dominant $-1 < 0$. En particulier, son sommet est un maximum, atteint en $\frac{0+1}{2} = \frac{1}{2}$, valant $f\left(\frac{1}{2}\right) = \frac{1}{4}$. Ainsi, $f(p) = p(1-p) = \frac{1}{4}$.

Ainsi,

$$\mathbb{P}(|M_n - p| \geq a) \leq \frac{p(1-p)}{na^2} \leq \frac{1}{4na^2} \quad \text{donc} \quad \mathbb{P}(|M_n - p| < a) \geq 1 - \frac{1}{4na^2} \quad \text{c'est-à-dire} \quad \mathbb{P}(p \in]M_n - a, M_n + a[) \geq 1 - \frac{1}{4na^2}$$

La minoration de la probabilité dépend encore de n . Pour des raisons qui s'éclairciront après, faisons en sorte que cette dépendance disparaisse. Notre seule variable d'ajustement est a , qui est au carré, donc si on écrit a sous la forme $a = \sqrt{\frac{\delta}{n}}$, on obtient :

$$\mathbb{P}\left(p \in \left]M_n - \sqrt{\frac{\delta}{n}}, M_n + \sqrt{\frac{\delta}{n}}\right[\right) \geq 1 - \frac{1}{4n(\sqrt{\delta/n})^2} = 1 - \frac{1}{4\delta}$$

Définition 5

Intervalle de confiance

Un intervalle de confiance au niveau de confiance α est un intervalle I tel que $\mathbb{P}(p \in I) \geq \alpha$.

Soit $\alpha \in]0; 1[$. Pour régler le niveau de confiance, il s'agit de choisir un bon δ . Le niveau de confiance ne devant pas dépendre de n , c'est ce qui a motivé la manipulation précédente.

$$1 - \frac{1}{4\delta} = \alpha \iff 4\delta = \frac{1}{1-\alpha} \iff \delta = \frac{1}{4(1-\alpha)} \quad \text{et} \quad \sqrt{\frac{\delta}{n}} = \frac{1}{2\sqrt{n}\sqrt{1-\alpha}}$$

En injectant cette expression dans l'intervalle de confiance trouvé ci-dessus :

$$\mathbb{P} \left(p \in \left[M_n - \frac{1}{2\sqrt{n}\sqrt{1-\alpha}}, M_n + \frac{1}{2\sqrt{n}\sqrt{1-\alpha}} \right] \right) \geq \alpha$$

Il est courant de prendre $\alpha = 0,95$, c'est-à-dire que l'on s'assure à 95% que la valeur recherchée est dans l'intervalle de confiance. Cela donne :

$$\mathbb{P} \left(p \in \left[M_n - \sqrt{\frac{5}{n}}, M_n + \sqrt{\frac{5}{n}} \right] \right) \geq 0,95$$

Application numérique : Dans la situation exposée en introduction, $n = 1\,000$ et $M_{1\,000} = 0,52$. L'intervalle de confiance au niveau de confiance $\alpha = 0,95$ donne :

$$\mathbb{P} \left(p \in \left[0,52 - \sqrt{\frac{5}{1\,000}}; 0,52 + \sqrt{\frac{5}{1\,000}} \right] \right) \geq 0,95 \quad \text{soit} \quad \mathbb{P}(p \in]0,44; 0,60]) \geq 0,95$$

Remarque importante : $0,52 - \sqrt{\frac{5}{1\,000}} \approx 0,449$ et $0,52 + \sqrt{\frac{5}{1\,000}} \approx 0,591$. On arrondit la borne inférieure par défaut et la supérieure par excès : en effet, cela agrandit l'intervalle, et donc augmente la probabilité que p y appartienne, qui reste nécessairement supérieure au niveau de confiance choisi. *A contrario*, si on arrondissait au centième le plus proche, cela rétrécirait l'intervalle, ce qui réduirait la probabilité y appartienne : auquel cas, rien ne nous garantit que cette probabilité réduite est toujours supérieure au niveau de confiance choisi.

Conclusion : Noether a toutes les raisons de se méfier, la majorité absolue n'est pas garantie à 95% !

Justement : en supposant que la moyenne d'échantillon donne toujours 52% en faveur de Noether, combien faudrait-il sonder d'électeurs et électrices pour être sûr à 95% que plus de la moitié de l'électorat souhaite voter pour Noether ? Avec les termes mathématique établis précédemment, cela revient à chercher à partir de quel n la borne inférieure de l'intervalle de confiance devient supérieure à 0,5 :

$$0,52 - \sqrt{\frac{5}{n}} > 0,5 \iff \sqrt{\frac{5}{n}} < 0,02 \iff \sqrt{n} > \frac{\sqrt{5}}{0,02} \iff n > \left(\frac{\sqrt{5}}{0,02} \right)^2 = 12\,500$$

Ainsi, pour $n = 12\,500$ sondés, l'intervalle de confiance donne :

$$\mathbb{P}(p \in]0,5; 0,54]) \geq 0,95$$

14 Convexité

Table des matières

1	Dérivée de la dérivée de la dérivée de la...	71
2	Inégalité arithmético-géométrique	73
3	Cercle osculateur et étude qualitative	73

1 – Dérivée de la dérivée de la dérivée de la...

Ce chapitre a été l'occasion de dériver la fonction dérivée pour obtenir la dérivée seconde... mais pourquoi s'arrêter en si bon chemin !

Définition 1

Dérivée n -ième

Soient f une fonction définie sur un intervalle et n un entier naturel non nul.
On dit que f est n fois dérivable sur I si f est dérivable sur I , que sa dérivée est aussi dérivable sur I , que la dérivée de la dérivée est aussi dérivable sur I , etc. jusqu'à avoir dérivé n fois les résultats successifs.
On note $f^{(n)}$ la n -ième dérivée de f . Par convention on note $f = f^{(0)}$.

Attention, cette notation se place en exposant, mais ce n'est pas une puissance ! C'est une notation, rien de plus. Cette définition peut vous sembler un peu « avec les mains », et à raison ; on peut donner une définition plus rigoureuse en « mode récurrence » :

Définition 2

Dérivée n -ième, alternative

Soient f une fonction définie sur un intervalle et n un entier supérieur ou égal à 2.
On dit que f est n fois dérivable si f est dérivable et si f' est $(n-1)$ -fois dérivable.

Exercice 1 (Formule de Leibniz)

Soient f et g deux fonctions n fois dérivables sur un intervalle I .
Montrer par récurrence que pour tout entier $n \geq 1$,

$$(f \times g)^{(n)} = \sum_{k=0}^n \binom{n}{k} f^{(k)} g^{(n-k)}$$

Exercice 2 (Degré d'un polynôme)

Soit n un entier naturel non nul.
On dit qu'une fonction f est *polynomiale de degré n* si elle s'écrit sous la forme :

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0$$

où $a_0, a_1, \dots, a_n$ sont des réels et $a_n \neq 0$ (autrement dit la plus grande puissance de x qui apparaît est x^n).
Soit f une fonction polynomiale. On admet qu'une fonction est dérivable autant de fois que l'on veut sur $\mathbb{R}$.
Montrer que f est de degré n si, et seulement si $f^{(n+1)} = 0$ et $f^{(n)} \neq 0$.

Exercice 3 (Formule intégrale de Taylor)

Soit n un entier naturel. Soit f une fonction définie sur un intervalle I ouvert, soit $a \in I$. L'objectif de cet exercice est de montrer que si f est $(n+1)$ fois dérivable et que $f^{(n+1)}$ est continue sur I , alors on dispose de la magnifique formule suivante, qui est pour ainsi dire le porte-drapeau de l'analyse du supérieur :

$$\forall x \in I, \quad f(x) = \sum_{k=0}^n \frac{f^{(k)}(a)}{k!} (x-a)^k + \int_a^x \frac{f^{(n+1)}(t)}{n!} (x-t)^n dt$$

Pour cela nous allons procéder par récurrence. On note $\mathcal{P}_n$ la propriété au rang n , c'est-à-dire l'énoncé ci-dessus.

1► Commençons par l'initialisation.

a► On rappelle que le symbole $\sum_{k=0}^n u_k$ signifie que l'on ajoute les valeurs de u_k pour chaque indice k entre 0 et n (comme une boucle `for`), c'est-à-dire que l'on calcule $u_0 + u_1 + u_2 + \dots + u_{n-1} + u_n$.

Ici, n vaut 0, donc on le symbole s'écrit $\sum_{k=0}^0$. Quelle(s) est/sont la/les valeur(s) prises par k ? En déduire la valeur de la somme dans la formule.

b► Citer une primitive de f' sur I .

c► En déduire la valeur de l'intégrale dans la formule.

d► En déduire que $\mathcal{P}_0$ est vraie.

2► Faire proprement la rédaction entre l'initialisation et l'hérédité. On réécrira toute la formule dans l'hypothèse de récurrence, et on explicitera l'objectif de l'hérédité sans raccourci (« montrons que $\mathcal{P}_{n+1}$ est vraie, c'est-à-dire [...] »).

3► Occupons-nous de l'hérédité. On part de l'hypothèse de récurrence.

a► Concentrons-nous sur le terme intégral de la formule. Réaliser une intégration par parties en prenant

$$u(t) = \frac{f^{(n+1)}(t)}{n!} \text{ et } v'(t) = (x-t)^n.$$

b► Faisons un petit aparté sur le symbole somme. Expliquer pourquoi :

$$\sum_{k=0}^{n+1} u_k = \left(\sum_{k=0}^n u_k \right) + u_{n+1}$$

c► Conclure l'hérédité.

4► Rédiger la conclusion de la récurrence.

Exercice 4 (Développement de Taylor)

On admet que sous certaines hypothèses très restrictives, la formule de l'exercice précédent permet de montrer que :

$$\forall x \in I, \quad f(x) = \lim_{n \rightarrow +\infty} \sum_{k=0}^n \frac{f^{(k)}(a)}{k!} (x-a)^k$$

On admet que la fonction exponentielle vérifie ces hypothèses. Utiliser ce résultat pour montrer que pour tout $x \in \mathbb{R}$,

$$\exp(x) = \lim_{n \rightarrow +\infty} \sum_{k=0}^n \frac{x^k}{k!}$$

2 – Inégalité arithmético-géométrique

Exercice 5

Soient un entier $n \geq 2$ et $a_1, a_2, \dots, a_n$ des réels strictement positifs. L'objectif de cet exercice est de montrer l'inégalité suivante :

$$(a_1 a_2 \cdots a_n)^{\frac{1}{n}} \leq \frac{a_1 + a_2 + \cdots + a_n}{n}$$

- 1► a► Par un argument de convexité sur la fonction exponentielle, montrer que pour tout t réel, $e^t \geq 1 + t$.
b► En déduire que pour tout x réel, $x \leq e^{x-1}$.
- 2► Résultat intermédiaire : montrer que si $0 < a \leq b$ et $0 < c \leq d$, alors $ac \leq bd$. (Indication : comparer des quotients à 1...)

- 3► On note $M = \frac{a_1 + a_2 + \cdots + a_n}{n}$. Montrer que :

$$\frac{a_1}{M} \times \frac{a_2}{M} \times \cdots \times \frac{a_n}{M} \leq \exp\left(\frac{a_1}{M} - 1\right) \times \exp\left(\frac{a_2}{M} - 1\right) \times \cdots \times \exp\left(\frac{a_n}{M} - 1\right)$$

- 4► En déduire que :

$$\frac{a_1 a_2 \cdots a_n}{M^n} \leq \exp\left(\left[\frac{a_1}{M} + \frac{a_2}{M} + \cdots + \frac{a_n}{M}\right] - n\right)$$

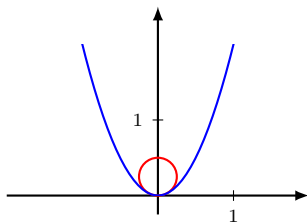
- 5► Montrer que le membre de droite est égal à 1.
- 6► Conclure.

3 – Cercle osculateur et étude qualitative

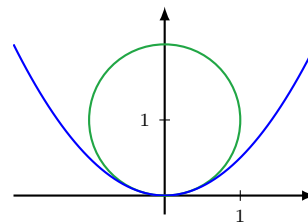
Vous l'avez peut-être oublié à force de la réduire à un simple exercice calculatoire, mais l'essence de la dérivée d'une fonction est de répondre à la question « quelle est la droite qui fournit la meilleure approximation de ma fonction à cet endroit ? »

Les droites, c'est bien gentil, mais c'est très... droit. En général les courbes sont... courbes, et quel objet géométrique courbe et simple rencontre-t-on en primaire ? Le cercle ! La question « quel est le cercle qui fournit la meilleure approximation de ma fonction à cet endroit ? » donne naissance à un nouvel objet : le *cercle osculateur*.

Comme vu en première, la dérivée permet de trouver l'équation réduite de la tangente, ou plus simplement sa pente, qui le plus souvent est ce qui nous intéresse. Elle nous indique à quelle point une fonction « monte » ou « descend » vite. Pour le cercle osculateur, on va utiliser la dérivée seconde. Son interprétation géométrique est la suivante : la courbe longe le cercle, donc plus le cercle est petit, plus la courbe « tourne vite », et plus le cercle est grand, plus le virage est doux.



Petit cercle osculateur en 0



Grand cercle osculateur en 0

Et qu'est-ce qui dicte la taille d'un cercle ? En effet, c'est son rayon ! Vous en avez sûrement entendu parler maintes et maintes fois dans votre vie, mais vous avez maintenant à peu près les outils nécessaires pour le comprendre :

Définition 3**Rayon de courbure**

Le *rayon de courbure* (ou la *courbure* tout court) d'une courbe en un point est rayon de son cercle osculateur en ce point.

Malheureusement, son calcul n'est pas aussi simple, loin de là ; on admettra volontiers que sa formule est la suivante :

Proposition 4

Soit f une fonction deux fois dérivable sur un intervalle I , soit $a \in I$. Si $f''(a) \neq 0$, alors la rayon de courbure de f en a vaut :

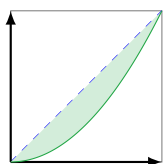
$$\rho_f(a) = \left| \frac{(1 + f'(a)^2)^{3/2}}{f''(a)} \right|$$

Par convention, on dit que le rayon de courbure est infini lorsque $f''(a) = 0$.

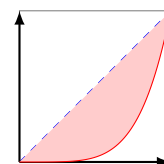
Lorsque la tangente est proche de l'horizontale, la dérivée est presque nulle, ce qui permet de donner l'approximation $\rho_f(a) \approx \frac{1}{|f''(a)|}$. Ainsi,

- Plus la dérivée seconde est grande (en valeur absolue), plus le rayon de courbure est petit, et plus la courbe « tourne fort », autrement dit elle est très creusée ou pointue,
- Plus la dérivée seconde est proche de 0, plus le rayon de courbure est grand, et plus la courbe est rectiligne.

Pour celles et ceux qui suivant la spécialité SES, la question de savoir si une courbe est plus ou moins creusée est intervenue avec la courbe de Lorenz et le coefficient de Gini qui lui est associé. C'est une courbe convexe, et le coefficient de Gini est lié à son intégrale (plus précisément à l'aire entre la courbe et la diagonale). Avec ces contraintes, l'aire sous la courbe est fortement liée à sa convexité : plus l'aire est proche de 0, plus la fonction est « tirée vers la diagonale », *a fortiori* elle a un grand rayon de courbure, autrement dit une dérivée seconde proche de 0 ; inversement, plus le coefficient est proche de son maximum 1, plus la courbe est « poussée vers le coin », ce qui la creuse : elle a donc un faible rayon de courbure, autrement dit une grande dérivée seconde.

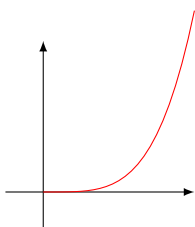


Faible coefficient

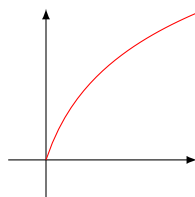


Grand coefficient

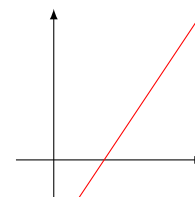
Le rayon de courbure permet de « quantifier la convexité » d'une fonction, mais c'est un outil bien puissant pour nous qui ne cherchons que de l'étude qualitative de courbes. En effet, ce qui nous intéresse, c'est surtout la vitesse de croissance ou de décroissance d'une courbe. Prenons le cas d'une fonction croissante. Il y a plusieurs manières de croître : avec un rythme constant, qui s'accélère, ou qui ralentit.



Croît en accélérant



Croît en ralentissant



Croissance à vitesse constante

Avec une dérivée seconde positive, dérivée est croissante, mais la dérivée régit la vitesse de croissance, donc notre fonction croît de plus en plus vite. Ainsi, plus la dérivée seconde est positive, plus la croissance accélère rapidement. Inversement, plus la dérivée seconde est négative, plus la croissance ralentit rapidement.

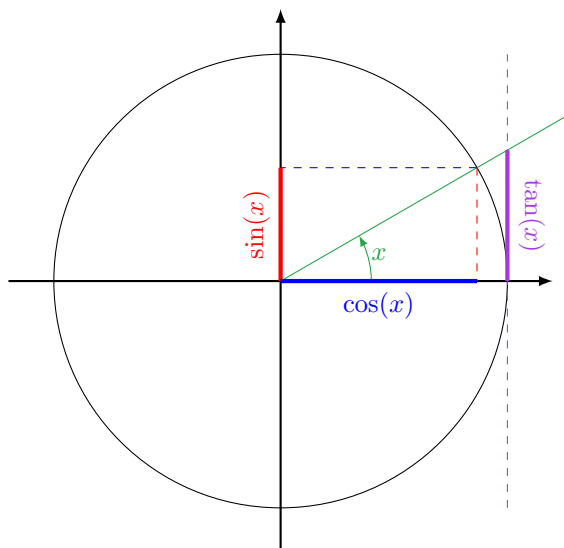
15 Trigonométrie

Table des matières

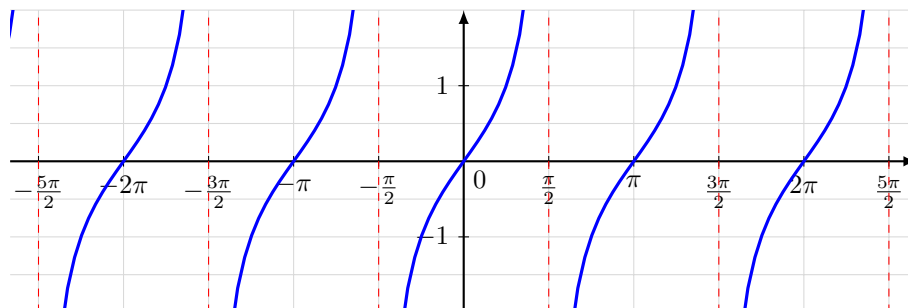
1	Pour aller (plus) loin : la fonction tangente	75
2	Pour la culture : mais à quoi ça sert ? !	76
3	La trigonométrie post-bac	77
A	Fonctions inverses	77
B	Relations trigonométriques	78
C	Écriture trigonométrique des nombres complexes – fonctions hyperboliques	79

1 – Pour aller (plus) loin : la fonction tangente

Dans le vaste monde des fonctions périodiques, trois fonctions règnent en maîtres. Le cosinus et le sinus sont deux l'une des autres, et le troisième membre de ce triumvirat, vu au collège lors des relations trigonométriques du triangle rectangle, est la *tangente*. D'abord définie comme le rapport des longueurs des côtés opposé et adjacent, sa définition fonctionnelle est $\tan = \frac{\sin}{\cos}$. Comme ses consœurs, la tangente peut se lire sur le cercle trigonométrique – lecture qui lui vaut son nom de *tangente* :



Son graphe est le suivant :



On constate sur ce graphique :

- Elle est π -périodique,
- Elle possède une asymptote verticale à chaque mesure d'un angle droit : en effet, dans ces conditions, le cosinus au dénominateur s'annule.

La fonction tangente est dérivable partout où elle est définie, et sa dérivée peut se calculer alors la formule de dérivation d'un quotient :

$$\tan' = \frac{\sin' \cos - \sin \cos'}{\cos^2} = \frac{\cos \cos - \sin(-\sin)}{\cos^2} = \frac{\cos^2 + \sin^2}{\cos^2}$$

Cette dernière expression peut se simplifier de deux façons :

- ▶ Soit on utilise l'identité trigonométrique $\cos^2 + \sin^2 = 1$, ce qui donne $\tan' = \frac{1}{\cos^2}$,
- ▶ Soit on scinde la fraction en deux et on la simplifie :

$$\tan' = \frac{\cos^2}{\cos^2} + \frac{\sin^2}{\cos^2} = 1 + \left(\frac{\sin}{\cos}\right)^2 = 1 + \tan^2$$

2 – Pour la culture : mais à quoi ça sert ? !

Ouvrez Geogebra et tracez une fonction $f : x \mapsto A \cos(wx + p)$ où A , w et p sont trois paramètres que vous créerez à l'aide de trois curseurs. Faites varier ces curseurs : que constatez-vous ? Le paramètre A influence l'*amplitude* des vagues, le paramètre w influence leur fréquence, leur vitesse d'oscillation – on parle de *fréquence* ou de *pulsation* selon le contexte –, et le paramètre p glisse la courbe horizontalement – on parle de *phase* de la fonction –. Cette simulation doit vous convaincre que n'importe quelle fonction qui ressemble à un sinus ou un cosinus déformé peut s'obtenir en ajustant ces trois paramètres.

Essayez maintenant d'ajouter quelques telles fonctions, avec des paramètres A , w et p qui leur sont propres : vous obtenez une courbe biscornue, mais périodique, autrement dit que se répète. Faites varier les différents paramètres : cela change la forme de la fonction somme, mais encore une fois celle-ci reste périodique, aussi complexe soit-elle. On peut alors se demander : toute fonction périodique peut-elle être obtenue en ajoutant des fonctions de la forme de f ? La réponse est : oui, c'est très exactement le fondement de ce qu'on appelle l'*analyse de Fourier*. Cette propriété semble de prime abord n'être qu'une curiosité mathématique, pourtant l'analyse de Fourier est aujourd'hui omniprésente, elle est un des outils centraux dans le traitement de l'information.

Commençons par un exemple concret qui permet de comprendre l'utilité d'une *série de Fourier*, c'est-à-dire le fait d'écrire une fonction périodique comme une somme de fonctions trigonométriques. Voici une question : qu'est-ce qui fait la différence entre la La d'un piano et celui d'un violon ? La même note est jouée par les deux instruments, pourtant ils produisent un son bien différent... Le son peut être modélisé par une fonction périodique, et donc peut être décomposée en série de Fourier. Et c'est là que se trouve la différence : ce sont les différents coefficients A , w et p qui caractérisent un son. Autrement dit, si on connaît tous les A , w et p , alors on peut dire avec certitude « ce son correspond au La d'un piano ». Mais cela marche aussi dans l'autre sens : connaissant les valeurs des coefficients A , w et p , un ordinateur peut calculer la somme de fonctions trigonométriques correspondante, et donc recréer la fonction qui correspond au son demandé : on crée ainsi un synthétiseur. Cette méthode permet de vraiment simuler le son d'un instrument, ce qui serait impossible en enregistrant séparément chaque note, car le synthétiseur peut ajuster en temps réel les différents paramètres pour moduler le son, ce qui serait impossible avec des sons fixes préenregistrés.

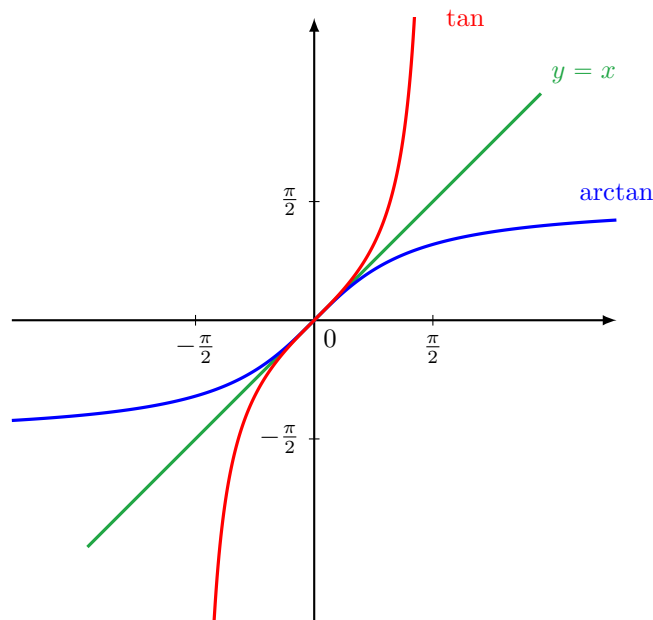
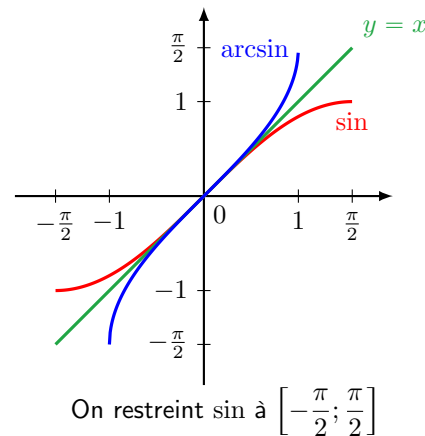
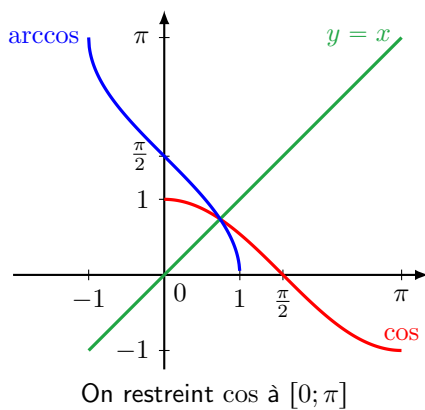
Ainsi on comprend que l'analyse de Fourier est un moyen d'étudier les fondements d'un signal. De la compression fichiers JPEG ou MP3 à l'analyse de données sismologiques pendant la guerre froide pour déterminer les pays travaillant sur l'arme nucléaire, l'analyse de Fourier trouve une application dès lors qu'il s'agit de traiter de l'information.

3 – La trigonométrie post-bac

A – Fonctions inverses

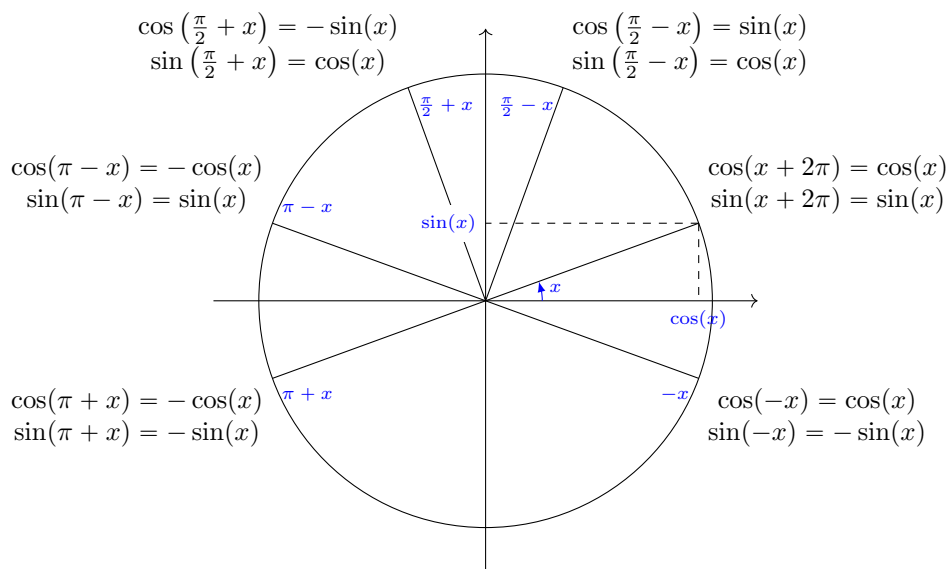
En terminale, on utilise la touche `arccos` pour calculer des angles dans le contexte des produits scalaires, sa vraiment savoir ce qu'est la fonction arccosinus.

Si on restreint les fonctions cosinus et sinus à des intervalles particuliers, on obtient des fonctions strictement monotones. Souvenez-vous de la fonction logarithme népérien : son graphe est symétrique à celui de la fonction exponentielle par rapport à la droite d'équation $y = x$. On a également vu que cette propriété était liée à l'égalité $e^{\ln(x)} = \ln(e^x) = x$. Le principe est le même ici : la fonction arccosinus vérifie $\arccos(\cos(x)) = \cos(\arccos(x)) = x$ pour les valeurs de x où la fonction est bien définie, sa courbe représentative est symétrique à la courbe de la fonction cosinus, restreinte au bon intervalle. Le principe est le même pour le sinus, et même la tangente.



B – Relations trigonométriques

Les fonctions trigonométriques sont des fonctions très rigides, riches en diverses relations. Les formules de transformation résumées si le cercle trigonométrique ci-dessous sont très utiles pour ramener certains calculs de sinus ou de cosinus à des angles plus simples, dans le premier quadrant. Les tableaux plus bas listent les autres formules que l'on peut rencontrer.



FORMULES DE DUPLICATION

$$\cos(2a) = \cos^2(a) - \sin^2(a)$$

$$= 1 - 2 \sin^2(a)$$

$$= 2 \cos^2(a) - 1$$

$$\sin(2a) = 2 \cos(a) \sin(a)$$

FORMULES DE DÉVELOPPEMENT

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

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

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

FORMULES D'ADDITION

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

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

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

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

$$\tan(x+y) = \frac{\tan(x) + \tan(y)}{1 - \tan(x) \tan(y)}$$

$$\tan(x-y) = \frac{\tan(x) - \tan(y)}{1 + \tan(x) \tan(y)}$$

FORMULES DE FACTORISATION

$$\cos(a) + \cos(b) = 2 \cos\left(\frac{a+b}{2}\right) \cos\left(\frac{a-b}{2}\right)$$

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

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

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

ARC MOITIÉ ($t = \tan(x/2)$)	
$\cos(x) = \frac{1-t^2}{1+t^2}$	$\sin(x) = \frac{2t}{1+t^2}$
$\tan(x) = \frac{2t}{1-t^2}$	

C – Écriture trigonométrique des nombres complexes – fonctions hyperboliques

Si vous suivez le cours de mathématiques expertes, vous avez sûrement vu la notation $e^{i\theta}$ définie par $e^{i\theta} = \cos(\theta) + i \sin(\theta)$. À l'aide de la conjugaison complexe, on peut renverser cette formule pour définir le cosinus et le sinus d'un nombre en fonction de cette exponentielle complexe : c'est ce que l'on appelle les formules d'Euler.

FORMULES D'EULER	
$\cos(x) = \frac{e^{ix} + e^{-ix}}{2}$	$\sin(x) = \frac{e^{ix} - e^{-ix}}{2i}$

Cette notation exponentielle est, en vérité, plus qu'une notation. On peut en fait la manipuler presque comme une exponentielle classique, et ceci permet de démontrer les diverses formules de la section précédente.

Et si on oubliait l'exponentielle complexe, pour se restreindre à une exponentielle réelle ? En mimant les formules d'Euler, cette idée définit les fonctions *cosinus hyperbolique*, notée ch , et *sinus hyperbolique*, notée sh :

FONCTIONS HYPERBOLIQUES	
$\text{ch}(x) = \frac{e^x + e^{-x}}{2}$	$\text{sh}(x) = \frac{e^x - e^{-x}}{2}$

On définit de même la *tangente hyperbolique* par $\text{th} = \frac{\text{sh}}{\text{ch}}$. Fait amusant : pour obtenir le formulaire de trigonométrie hyperbolique, il suffit de changer le signe de chaque produit de deux sinus !

Contrairement à leurs homologues trigonométriques, les fonctions trigonométriques ne sont pas périodiques. Elles trouvent cependant d'autres applications. La plus notable d'entre elles est sûrement l'équation de la chaînette : lorsque vous laissez pendre une corde attachée en ses deux extrémités, la courbe qu'elle décrit est celle d'un cosinus hyperbolique !

A Rappels de probabilités

1 – Rappels : vocabulaire et formulaire de probabilités

Vocabulaire

Définition 1

- ▶ Une *expérience aléatoire* est une expérience pour laquelle on ne peut pas prévoir avec certitude le résultat.
- ▶ On appelle *issue* d'une expérience aléatoire tout résultat possible de cette expérience.
- ▶ L'*univers* associé à une expérience aléatoire est l'ensemble des résultats possibles, ou issues, de cette expérience aléatoire. On le note souvent Ω (que l'on lit « omega »).
- ▶ Un *événement* est une partie ou un sous ensemble de l'univers.
- ▶ Un *événement élémentaire* est un événement qui ne contient qu'un seul élément, qu'une seule issue.

Remarque 2 (Événement vs. issue)

Pour bien comprendre la différence entre issue et événement, faisons une analogie. Quand vous faites des courses, vous achetez certains articles. Chaque article (l'objet physique, en lui-même) est une issue différente. Un événement, c'est ce que vous mettez dans votre panier. Votre panier (événement) peut contenir plusieurs articles (issues), ou un seul (événement élémentaire).

Lorsque vous arrivez en caisse, vous payez pour le contenu de votre panier, autrement dit le prix de votre panier est la somme du prix de chaque article. De la même manière, la probabilité d'un événement est la somme des probabilités des issues qu'il contient.

Probabilité, union, intersection, complémentaire

Définition 3

Soient A et B deux événements d'un même univers Ω fini.

- ▶ La *probabilité* d'un événement A est une mesure de la fréquence de réalisation de l'événement A .
- ▶ Lorsque tous les événements élémentaires ont tous la même probabilité, on dit qu'ils sont *équiprobables*.
- ▶ L'événement $A \cap B$ est l'événement « A et B », constitué de toutes les issues appartenant à la fois à A et à B .
- ▶ L'événement $A \cup B$ est l'événement « A ou B », formé des issues appartenant à au moins l'un des deux événements A et B .
- ▶ Deux événements A et B sont dits *incompatibles* lorsque leur intersection $A \cap B$ est vide. Cela équivaut à dire qu'aucune issue ne peut réaliser à la fois A et B , que la réalisation d'un des deux événements A ou B exclut celle de l'autre.
- ▶ L'*événement contraire* de A , noté $\bar{A}$, est l'ensemble de toutes les issues de l'univers Ω n'appartenant pas à A .

Événements extrêmes

Propriété 4

- ▶ Pour tout événement A , on a $0 \leq \mathbb{P}(A) \leq 1$.
- ▶ $\mathbb{P}(\emptyset) = 0$ (on dit que l'événement $\emptyset$ est *impossible*)
- ▶ $\mathbb{P}(\Omega) = 1$ (on dit que l'événement Ω est *certain*)

Équiprobabilité**Proposition 5**

Dans une situation d'équiprobabilité :

$$\mathbb{P}(A) = \frac{\text{nombre d'issues de } A}{\text{nombre d'issues de } \Omega}$$

S'il y a n issues élémentaires, la probabilité d'avoir une de ces issues est $\frac{1}{n}$.

Proposition 6**Formule du crible – Événement contraire**

Pour tous événements A et B :

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B) \quad \text{et} \quad \mathbb{P}(\overline{A}) = 1 - \mathbb{P}(A)$$

Définition 7**Probabilité conditionnelle**

Soient A et B deux événements tel que $\mathbb{P}(B) \neq 0$. La *probabilité de A sachant que B* , notée $\mathbb{P}_B(A)$, est la probabilité que l'événement A se réalise, en ayant l'information que l'événement B s'est déjà réalisé. Elle vaut :

$$\mathbb{P}_B(A) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$$

Remarque 8

La probabilité de A sachant B ne correspond pas à la probabilité que A et B se réalisent : cet événement est $A \cap B$. Quand on dit « sachant B », on entend que B s'est déjà réalisé. Quand on attend $A \cap B$, l'événement B ne s'est pas encore réalisé.

Remarque 9

Il existe une autre notation, privilégiée dans le monde anglophone : $\mathbb{P}_B(A) = \mathbb{P}(A|B)$.

Proposition 10**Formule des probabilités composées**

$$\mathbb{P}(A \cap B) = \mathbb{P}(A) \mathbb{P}_A(B)$$

Proposition 11**Formule des probabilités totales**

$$\mathbb{P}(A) = \mathbb{P}(A \cap B) + \mathbb{P}(A \cap \overline{B})$$

Définition 12**Indépendance**

On dit que deux événements A et B sont *indépendants* lorsque :

$$\mathbb{P}(A \cap B) \stackrel{\text{id}}{=} \mathbb{P}(A) \times \mathbb{P}(B)$$

Corollaire 13

Les événements A et B sont indépendants si, et seulement si $\mathbb{P}_B(A) \stackrel{\text{id}}{=} \mathbb{P}(A)$.

Indépendance et contraire**Proposition 14**

A et B sont indépendants si, et seulement A et $\overline{B}$ le sont.

Application 15

On lance deux fois une pièce équilibrée, de manière indépendante.

- 1► Représenter la situation à l'aide d'un arbre pondéré.
- 2► Déterminer l'univers des possibles.
- 3► Déterminer les probabilités d'obtenir 0 fois Pile, 1 fois Pile, 2 fois Pile.

Exercices conseillés : 1, 2

2 – Succession d'épreuves indépendantes ou non**Proposition 16**

Dans un arbre de probabilités, la somme des probabilités des branches issues d'un même nœud est égale à 1.

Proposition 17**Formule des probabilités composées**

Dans un arbre de probabilités, la probabilité d'un chemin est le produit des probabilités rencontrées en chemin.

Proposition 18**Formule des probabilités totales**

Dans un arbre de probabilités, la probabilité d'un événement est la somme des probabilités des chemins qui mènent à cet événement.

Application 19

Une urne contient 5 boules blanches. On lance une pièce équilibrée : si elle tombe sur Pile, on rajoute 3 boules noires dans l'urne, et si elle tombe sur Face on rajoute 10 boules noires dans l'urne. On tire ensuite une boule dans l'urne.

- 1► Représenter la situation à l'aide d'un arbre de probabilités.
- 2► Déterminer la probabilité de tirer Pile puis une boule blanche.
- 3► Déterminer la probabilité de tirer une boule blanche.

Proposition 20

L'ensemble des issues d'une expérience aléatoire à plusieurs épreuves est le produit cartésien des ensembles des issues de chaque épreuve.

Exemple 21

On lance une pièce puis un dé à 4 faces, numérotées de 1 à 4. L'ensemble des issues de la première épreuve est $\{P, F\}$, l'ensemble des issues de la seconde épreuve est $\{1; 2; 3; 4\}$. L'ensemble des issues de l'expérience aléatoire est donc $\{(P, 1); (P, 2); (P, 3); (P, 4); (F, 1); (F, 2); (F, 3); (F, 4)\}$.

Application 22

On lance un jeton dont les deux faces sont numérotées 1 et 2, puis un dé à 4 faces, puis un dé à 6 faces. On note L_1 la variable aléatoire qui mesure le résultat du lancer du jeton, L_2 la variable aléatoire qui mesure le résultat du lancer du dé à 4 faces, et L_3 la variable aléatoire qui mesure le résultat du lancer du dé à 6 faces.

- 1► Déterminer l'univers des possibles, c'est-à-dire l'ensemble des possibilités pour le triplet (L_1, L_2, L_3) .
- 2► Déterminer la probabilité de l'issue $(2; 4; 6)$.
- 3► Déterminer la probabilité de l'issue $(4; 2; 6)$.

Application 23

On dispose de trois cartes notées A , B et C . On tire une carte parmi ces trois, on la remet en jeu et on tire à nouveau, de sorte que les tirages sont indépendants.

- 1► Représenter la situation à l'aide d'un arbre de probabilités.
- 2► Déterminer la probabilité de tirer deux fois la même carte.

Exercices conseillés : 3, 4, 5, 6, 7

Exercices**Exercice 1**

Dans un supermarché, 22% des paiements se font à la caisse automatique, dont 74% correspondent à des courses de moins de dix articles.

Pour les paiements en caisse non automatique, 11% correspondent à des courses de moins de dix articles.

Pour un paiement effectué dans ce magasin, on considère les événements :

- A : «Le paiement a été fait en caisse automatique»
- M : «Le paiement correspond à des courses de moins de dix articles»

- 1► Réaliser un arbre des possibles.
- 2► Calculer $\mathbb{P}(A \cap M)$ puis $\mathbb{P}(M)$. En déduire $\mathbb{P}_M(A)$.

Exercice 2

Un jeu de grattage coûtant 1€ rapporte :

- 100€ avec une probabilité de 0,005
- 20€ avec une probabilité de 0,01
- 5€ avec une probabilité de 0,02
- 1€ avec une probabilité de 0,1
- 0€ avec une probabilité de 0,865.

On considère la variable aléatoire G donnant le gain algébrique à ce jeu (tenant compte du prix du ticket).

- 1► Donner la loi de probabilité de G sous forme de tableau.
- 2► Calculer l'espérance et l'écart type de G . Ce jeu est-il équitable ?
- 3► Calculer $\mathbb{P}(G \geq 15)$.

Exercice 3

On considère l'expérience aléatoire consistant à lancer un dé équilibré à 4 faces numérotées de 1 à 4 puis une pièce équilibrée à Pile ou Face.

- 1► Représenter cette succession de deux épreuves indépendantes par un arbre.
- 2► Déterminer la probabilité que le résultat du dé soit inférieur ou égal à 3 et que la pièce tombe sur Pile.

Exercice 4

Le samedi soir, deux amis vont au cinéma, au concert et au théâtre sans jamais faire la même activité deux samedis de suite. Le choix entre deux activités possibles pour le samedi suivant est équiprobable.

- 1► On considère l'expérience aléatoire dont le résultat est l'activité faite par ces deux amis un samedi soir. Expliquer pourquoi la succession de cette expérience aléatoire les trois semaines à venir n'est pas une succession d'épreuves indépendantes.
- 2► Représenter cette succession de trois épreuves par un arbre sachant que samedi dernier, les amis sont allés au théâtre.
- 3► Déterminer la probabilité qu'ils aillent deux fois au cinéma au cours des 3 prochains samedis.

Exercice 5

On tire avec remise six personnes dans une population contenant 87% de droitiers (D) et 12% de gauchers (G) et 1% d'ambidextres (A), et on regarde si elles sont droitères, gauchères ou ambidextres.

Donner l'univers associé à cette succession de six épreuves indépendantes puis calculer la probabilité de l'issue (D, D, G, D, D, A) .

Exercice 6

Une urne contient 100 boules de 5 couleurs différentes : 23 bleues, 12 rouges, 9 vertes, 45 oranges et 11 violettes. On tire trois boules avec remise dans cette urne et on note à chaque fois la couleur obtenue.

- 1► Est-ce une succession de trois épreuves indépendantes ?
- 2► Donner l'univers associé à cette succession de 3 épreuves.
- 3► Déterminer la probabilité d'obtenir une boule bleue puis une boule rouge puis une boule violette.

Exercice 7

Dans un jeu télévisé, on tire au sort des boules sans remise parmi 20. Il y a 17 boules blanches numérotées qui donnent droit de deviner un mot, et 3 boules noires qui font passer la main à l'équipe adverse.

On considère la succession des trois premiers tirages de boules dans ce jeu selon qu'elles sont noires ou non.

- 1► Est-ce une succession de trois épreuves indépendantes ?
- 2► Représenter la situation par un arbre.
- 3► Quelle est la probabilité de devoir passer la main exactement une fois sur les trois premiers tirages ?

B Second degré : algèbre et géométrie

Table des matières

1	Définition et forme développée	87
2	Forme canonique, aspects graphiques et variations	88
A	Représentation graphique : parabole	88
B	Forme canonique, conséquences graphiques	90
C	Variations	91
3	(In)équations de degré 2 exploitant la symétrie	92
4	Pour aller (plus) loin	93

1 – Définition et forme développée

Définition 1

Fonction polynôme de degré 2

Une *fonction polynôme de degré 2* (ou *polynôme du second degré*) est une fonction f définie sur $\mathbb{R}$ par une expression de la forme :

$$f(x) = ax^2 + bx + c$$

où a , b et c sont des réels, et $a \neq 0$.

Les nombres a , b et c sont appelés *coefficients* de f ; le nombre c est appelé *coefficient constant*, et a est appelé *coefficient dominant*.

Remarque 2

En mathématiques, il y a plusieurs types d'objets : les nombres, les « formes géométriques », les fonctions... les polynômes en sont un autre. La subtilité est qu'un *polynôme* et une *fonction polynôme* ne sont pas le même objet, puisqu'ils ne sont pas du même type (une analogie est, par exemple, avec le langage Python, `12` et `12.0` désignent la même *valeur*, mais ne sont pas le même objet, le premier étant de type `int`, le second de type `float`).

Cette distinction, bien au-delà des attendus du lycée, peut se résumer ainsi : tant que l'on ne fait que manipuler l'expression algébrique, en développant, factorisant, en évaluant ou en dérivant, cela peut relever du domaine des polynômes. Dès lors que l'on étudie le signe et les variations, cela relève nécessairement du domaine des fonctions.

Exemple 3

- ▶ $f(x) = 3x^2 - 5x + 6$ est une fonction polynôme de degré 2 avec $a = 3$, $b = -5$ et $c = 6$.
- ▶ $f(x) = x^2 - 8$ est une fonction polynôme de degré 2 avec $a = 1$, $b = 0$ et $c = -8$.
- ▶ $f(x) = 7x + 9$ n'est *pas* une fonction polynôme de degré 2.

Théorème 4

Unicité des coefficients

Les trois coefficients d'une fonction polynôme de degré 2 sont *uniques*, autrement dit il n'y a qu'une manière d'écrire une fonction polynôme de degré 2 sous la forme $ax^2 + bx + c$.

Démonstration. Soit f une fonction polynôme de degré 2. Supposons qu'il existe des coefficients tels que pour tout $x \in \mathbb{R}$, $f(x) = a_1x^2 + b_1x + c_1$ et $f(x) = a_2x^2 + b_2x + c_2$.

Alors pour tout $x \in \mathbb{R}$,

$$0 = f(x) - f(x) = (a_1 - a_2)x^2 + (b_1 - b_2)x + (c_1 - c_2)$$

Posons $a = a_1 - a_2$, $b = b_1 - b_2$ et $c = c_1 - c_2$. Ainsi, pour tout $x \in \mathbb{R}$, $ax^2 + bx + c = 0$.

En évaluant pour $x = 0$, on en déduit que $c = 0$. Ainsi, pour tout $x \in \mathbb{R}$, $ax^2 + bx = x(ax + b) = 0$. En évaluant pour $x = 1$, on en déduit que $a + b = 0$. Pour $x = -1$, on trouve que $-a + b = 0$. En ajoutant ces deux égalités, on trouve $2b = 0$, ainsi $b = 0$.

Ainsi, pour tout $x \in \mathbb{R}$, $ax^2 = 0$, et substituer 1 à x montre que $a = 0$.

Finalement, $a = a_1 - a_2 = 0$, donc $a_1 = a_2$. De même, $b_1 = b_2$, et $c_1 = c_2$: on a montré que si l'on souhaite écrire une fonction polynôme de degré 2 de deux manières, alors les coefficients sont égaux deux à deux, donc il n'existe *de facto* qu'une façon d'écrire une fonction polynôme de degré 2 d'une telle manière. $\square$

Forme développée

Définition 5

L'écriture $f(x) = ax^2 + bx + c$ est appelée *forme développée* de f .

2 – Forme canonique, aspects graphiques et variations

A – Représentation graphique : parabole

Parabole, orientation

Définition 6

La courbe représentative d'une fonction polynôme de degré 2 est appelée *parabole*. Lorsque le coefficient dominant est positif (respectivement, négatif), on dit qu'elle est *orientée vers le haut* (respectivement, *vers le bas*).

Axe de symétrie

Propriété 7

Une parabole admet un axe de symétrie vertical.

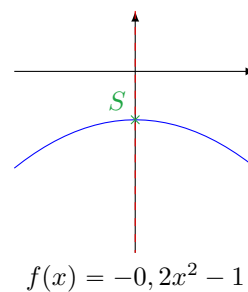
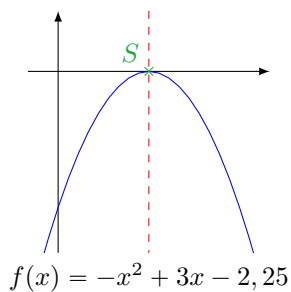
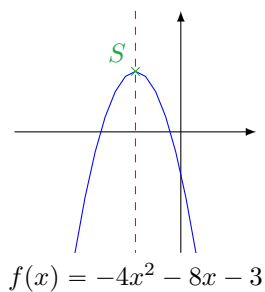
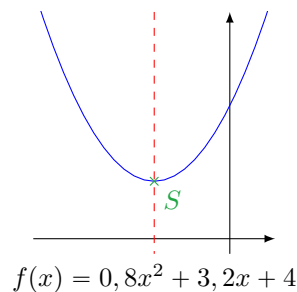
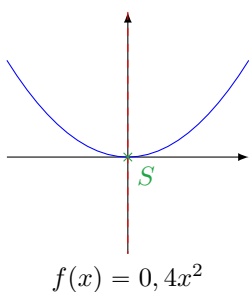
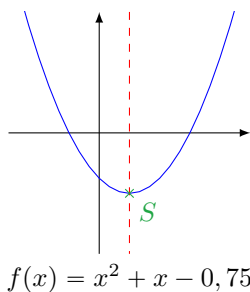
Démonstration. Voir la Proposition 14. $\square$

Sommet d'une parabole

Définition 8

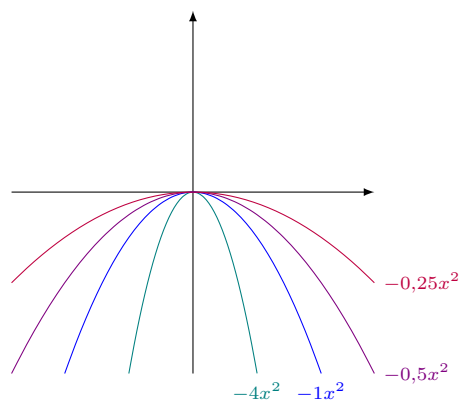
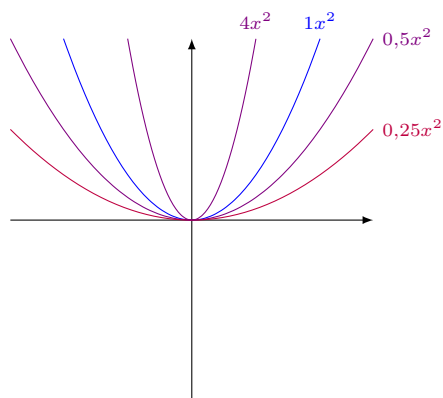
- Le *sommet* d'une parabole est le point d'intersection de celle-ci avec son axe de symétrie.
- Les deux parties de la courbe de part et d'autre de l'axe de symétrie sont appelées *branches* de la parabole.

Exemple 9

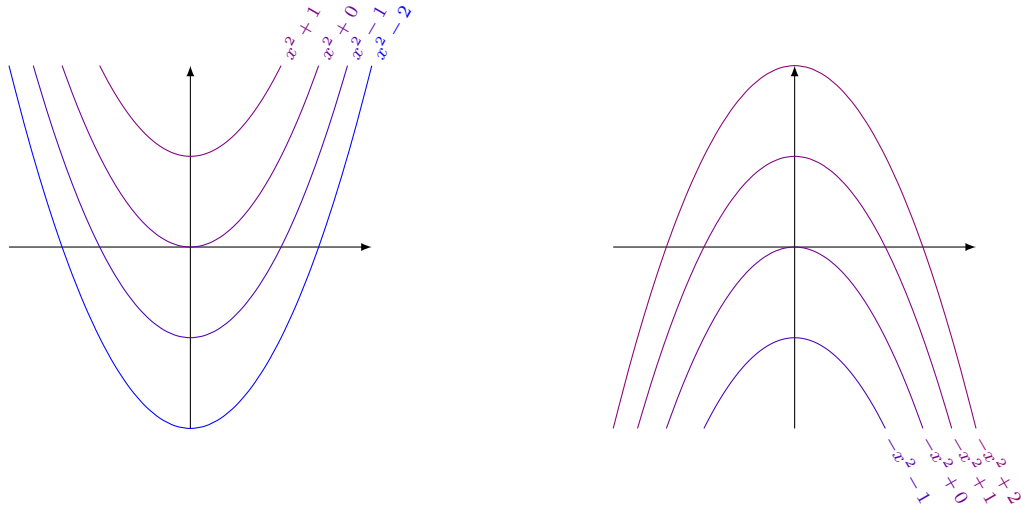


Exemple 10 (Cas particuliers remarquables)

- Les paraboles des fonctions du type $f(x) = ax^2$: leur axe de symétrie est l'axe des ordonnées, et leur sommet est l'origine du repère.



- Les paraboles des fonctions du type $f(x) = ax^2 + c$: elles correspondent à une translation de la parabole de $f(x) = ax^2$ le long de l'axe des ordonnées. Leur axe de symétrie est donc l'axe des ordonnées, et leur sommet est le point de coordonnée $(0, c)$.



B – Forme canonique, conséquences graphiques

Lemme 11 (Existence et unicité de la forme canonique)

Soit f une fonction polynôme de degré 2 de forme développée $f(x) = ax^2 + bx + c$.
Il existe deux réels α et β , uniques, tels que pour tout $x \in \mathbb{R}$,

$$f(x) = a(x - \alpha)^2 + \beta$$

Démonstration. Supposons que de tels réels existent. Alors pour tout $x \in \mathbb{R}$,

$$a(x - \alpha)^2 + \beta = a(x^2 - 2\alpha x + \alpha^2) + \beta = ax^2 + (-2a\alpha)x + (a\alpha^2 + \beta)$$

Cette écriture est une forme développée de f , donc par unicité des coefficients, $a = a$ (sans surprise), $b = -2a\alpha$ et $c = a\alpha^2 + \beta$. Comme $a \neq 0$, $\alpha = -\frac{b}{2a}$ et $\beta = c - a\alpha^2 = c - a\frac{b^2}{4a^2} = \frac{-b^2 + 4ac}{4a}$: ainsi, si de tels coefficients α et β existent, alors on peut les exprimer en fonction de a , b et c . Or ces derniers étant uniques, cela montre que α et β le sont également.

L'existence de α et β découle des expressions précédentes : on vérifie en développant et en réduisant l'expression que $a(x - \alpha)^2 + \beta = ax^2 + bx + c$. $\square$

Définition 12

Forme canonique

L'(unique) écriture du lemme est appelée *forme canonique* de f .

Remarque 13

Le coefficient a dans la forme développée et la forme canonique est le même !

Propriété 14

Soit f une fonction polynôme de degré 2 de forme canonique $f(x) = a(x - \alpha)^2 + \beta$. On note $\mathcal{C}_f$ sa courbe représentative (qui est une parabole).

► $\alpha = -\frac{b}{2a}$ et $\beta = f(\alpha)$.

► L'axe de symétrie de $\mathcal{C}_f$ a pour équation $x = \alpha$.

► Le sommet de $\mathcal{C}_f$ a pour coordonnées $S(\alpha, \beta)$.

Démonstration. L'expression de α a été donnée dans la démonstration du lemme. Le fait que $\beta = f(\alpha)$ découle de l'expression de f sous forme canonique.

Pour montrer que la droite d'équation $x = \alpha$ est un axe de symétrie de $\mathcal{C}_f$, il faut et il suffit de montrer que pour tout $h > 0$, $f(\alpha - h) = f(\alpha + h)$: soit $h > 0$.

$$f(\alpha - h) = a((\alpha - h) - \alpha)^2 + \beta = a(-h)^2 + \beta = ah^2 + \beta = a((\alpha + h) - \alpha)^2 + \beta = f(\alpha + h)$$

Le sommet S étant l'intersection de $\mathcal{C}_f$ et de son axe de symétrie, S appartient à l'axe de symétrie, donc son abscisse est α d'après ce qui précède. Comme $S \in \mathcal{C}_f$, son ordonnée est $f(\alpha) = \beta$. Ainsi, $S(\alpha, \beta)$. $\square$

C – Variations

Lemme 15 (Variations de la fonction carré)

La fonction carré est strictement décroissante sur $] -\infty; 0]$ et strictement croissante sur $[0; +\infty[$.

Démonstration. Soient x et y deux réels.

- Supposons que $x < y \leq 0$. D'une part, $x^2 > xy$ (en multipliant l'inégalité $x < y$ de l'hypothèse par x qui est strictement négatif). D'autre part, $xy \geq y^2$ (en multipliant $x < y$ par y qui est négatif). Par conséquent, $x^2 > xy \geq y^2$, en particulier $x^2 > y^2$, et donc la fonction carrée est strictement décroissante sur $[0; +\infty[$.
- Supposons que $x > y \geq 0$. D'une part, $x^2 > xy$ (en multipliant l'inégalité $x > y$ de l'hypothèse par x , qui est positif). D'autre part, $xy \geq y^2$ (en multipliant $x < y$ par y , qui est positif). Par conséquent, $x^2 > xy \geq y^2$, en particulier $x^2 > y^2$, et donc la fonction carrée est strictement croissante sur $[0; +\infty[$.

$\square$

Propriété 16**Variations**

Soit f une fonction polynôme de degré 2 de forme canonique $f(x) = a(x - \alpha)^2 + \beta$.

Cas $a > 0$

x	$-\infty$	α	$+\infty$
f			

Cas $a < 0$

x	$-\infty$	α	$+\infty$
f			

Démonstration. (un peu informelle)

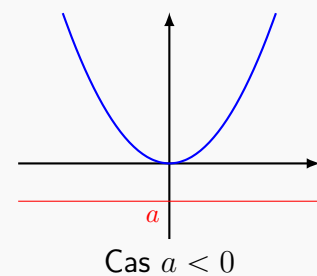
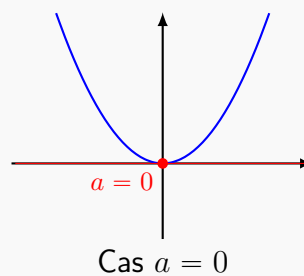
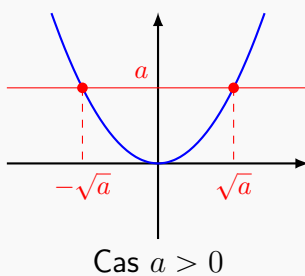
Cela découle des variations de la fonction carré : la composante $x - \alpha$ revient à traduire les abscisses de α , ou d'un point de vue réciproque, traduire la parabole de α . De même, ajouter β revient à traduire la parabole de β . La multiplication par a revient à dilater ou contracter les abscisses, éventuellement en inversant les ordonnées si a est négatif. Une translation ou une dilatation ne changent pas l'allure des variations. $\square$

3 – (In)équations de degré 2 exploitant la symétrie

Proposition 17

Soit $a \in \mathbb{R}$. On considère l'équation $x^2 = a$.

- Si $a > 0$, alors elle possède deux solutions distinctes, $\sqrt{a}$ et $-\sqrt{a}$.
- Si $a = 0$, alors elle possède 0 comme unique solution.
- Si $a < 0$, alors elle n'a pas de solution réelle*.



Exemple 18

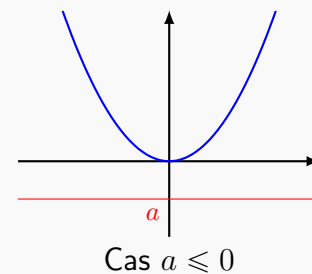
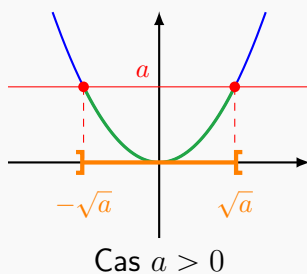
L'ensemble des solutions sur $\mathbb{R}$ de l'équation...

- ... $x^2 = 2$ est $S = \{\sqrt{2}, -\sqrt{2}\}$.
- ... $x^2 = -3$ est $S = \emptyset$ (il n'y a pas de solution réelle).

Proposition 19

Soit $a \in \mathbb{R}$. On considère l'inéquation $x^2 < a$.

- Si $a > 0$, alors l'ensemble des solutions est l'intervalle $S =]-\sqrt{a}, \sqrt{a}[$.
- Sinon, elle n'a pas de solution (autrement dit l'ensemble des solutions est $S = \emptyset$).



Exemple 20

L'ensemble des solutions réelles de l'inéquation...

- ... $x^2 < 2$ est $S =]-\sqrt{2}, \sqrt{2}[$.

*. Les experts ici savent qu'il existe deux solutions complexes conjuguées !

- ... $x^2 \geq 5$ est $S =]-\infty, -\sqrt{5}] \cup [\sqrt{5}, +\infty[$ (en effet, l'inéquation $x^2 \geq 5$ est la négation de l'inéquation $x^2 < 5$, dont l'ensemble des solutions est $] -\sqrt{5}; \sqrt{5}[$; ainsi on doit en considérer le complémentaire).

4 – Pour aller (plus) loin

Théorème 21

Définition géométrique d'une parabole

Soient $\mathcal{D}$ une droite du plan et F un point n'appartenant pas $\mathcal{D}$. L'ensemble des points du plan équidistants de F et $\mathcal{D}$:

$$\{M \in \mathcal{P} \mid MF = d(M, \mathcal{D})\}$$

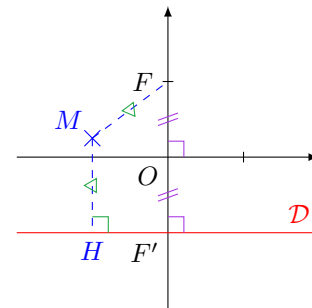
est une parabole. On appelle F le *foyer* de celle-ci, et $\mathcal{D}$ sa *directrice*.

Démonstration. Soit F' le projeté orthogonal de F sur $\mathcal{D}$. Comme $F \notin \mathcal{D}$, F' est distinct de F , et on peut considérer la droite (FF') , orthogonale à $\mathcal{D}$ par définition du projeté orthogonal. Soit O le milieu $[FF']$, on fixe le repère orthonormé du plan d'origine O , d'axe des ordonnées (FF') et d'unité OF .

Comme $\mathcal{D}$ et l'axe des abscisses sont toutes deux perpendiculaires à (FF') , elles sont parallèles entre elles. Comme OF est l'unité du repère orthonormé, F a pour coordonnées $(0; 1)$. Comme $OF = OF'$, les coordonnées de $F' \in \mathcal{D}$ sont $(0; -1)$. Ainsi, $\mathcal{D}$ a pour équation $y = -1$ dans le repère.

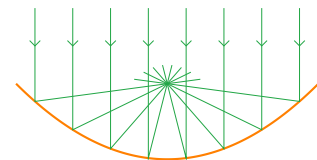
Soit $M(x, y)$ un point du plan équidistant de F et de $\mathcal{D}$. Notons H le projeté orthogonal de M sur $\mathcal{D}$. En particulier, H a pour coordonnées $(x; -1)$. La distance d'un point à une droite étant celle dudit point à son projeté, $MF = MH$, et donc $MF^2 = MH^2$, autrement dit :

$$\begin{aligned} (x-0)^2 + (y-1)^2 &= (x-x)^2 + (y-(-1))^2 \\ \text{c'est-à-dire} \quad x^2 + y^2 - 2y + 1 &= y^2 + 2y + 1 \\ \text{c'est-à-dire} \quad y &= \frac{1}{4}x^2 \end{aligned}$$



Comme MF et MH sont des distances, elles sont positives, et donc tout le paragraphe précédent est un raisonnement par équivalences. Ainsi, M est équidistant de F et $\mathcal{D}$ si, et seulement si les coordonnées (x, y) de M vérifient $y = \frac{1}{4}x^2$, autrement dit si, et seulement si M appartient à la courbe représentative de la fonction polynôme de degré 2 définie par $f(x) = \frac{1}{4}x^2$. Or la courbe d'une fonction polynôme de degré 2 est une parabole, donc l'ensemble des points équidistants de F et $\mathcal{D}$ décrit bien une parabole. $\square$

De cette définition découle une propriété notable des paraboles, qui justifie que l'on utilise des antennes *paraboliques* : orientée face à la source de rayons, une parabole réfléchit ces derniers vers son foyer. Ainsi, les rayons sont concentrés en un unique point. Cette propriété de concentration est également utilisée par les télescopes, pour focaliser les rayons si distants qu'ils en deviennent parallèles. Mais les rayons peuvent se propager dans l'autre sens ! Par exemple, les phares des voitures sont des miroirs paraboliques, dont on a placé sur le foyer une ampoule : ainsi, le faisceau qui en ressort est confiné dans une seule direction, pour éclairer une zone relativement précise.



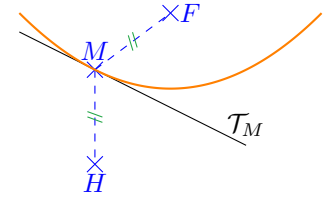
Démonstration. On reprend les notations de la preuve précédente. Nous y avons vu qu'en choisissant bien le repère, on peut assimiler la parabole à la courbe de $f : x \mapsto \frac{1}{4}x^2$. Notons $\mathcal{T}_M$ la tangente à la parabole en M .

Dans le repère, $\mathcal{T}_M$ a pour pente $f'(x_M) = \frac{1}{2}x_M$, et n'étant pas verticale (par dérivabilité), le vecteur $\vec{u}$ de coordonnée $\begin{pmatrix} 1 \\ x_M/2 \end{pmatrix}$ dirige $\mathcal{T}_M$.

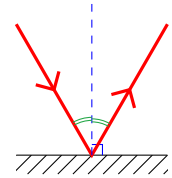
Le vecteur $\overrightarrow{FH}$ dirige la droite (FH) , ses coordonnées dans le repère sont $\begin{pmatrix} x_M - 0 \\ -1 - 1 \end{pmatrix} = \begin{pmatrix} x_M \\ -2 \end{pmatrix}$. Remarquons alors que :

$$\vec{u} \cdot \overrightarrow{FH} = 1 \times x_M + \frac{x_M}{2} \times (-2) = 0$$

Autrement dit, $\vec{u}$ et $\overrightarrow{FH}$ sont orthogonaux, il en va donc de même pour $\mathcal{T}_M$ et $[FH]$. Or $M \in \mathcal{T}_M$, donc $\mathcal{T}_M$ est la hauteur issue de M dans le triangle MFH . Or par construction de la parabole, $MF = MH$, autrement dit MFH est isocèle en M , et donc la hauteur issue de son sommet, M , en est aussi une bissectrice !



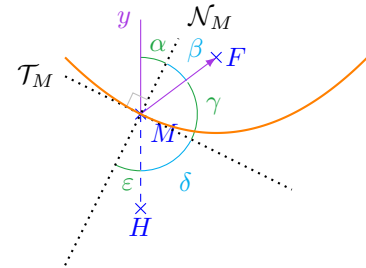
Un rayon réfléchi par une surface plane l'est de sorte qu'il soit symétrique par rapport à la normale à la surface au point de contact. On admet que lorsque la surface réfléchissante est courbe, on considère la réflexion par rapport à la tangente à la courbe en ce point.



Dans la situation que l'on étudie, les rayons arrivent parallèlement à l'axe des ordonnées (il est relativement naturel de vouloir orienter son antenne parabolique face à la source !). Un rayon qui touche la parabole en M se prolonge alors jusqu'en H , car les rayons sont perpendiculaires à la directrice de la parabole.

Notons $[My)$ la demi-droite correspondant au rayon source touchant la parabole en M . On note également $\mathcal{N}_M$ la droite perpendiculaire à $\mathcal{T}_M$ passant par M . On considère les angles α , β , γ , δ et ε tels que reportés sur la figure.

La droite $\mathcal{T}_M$ étant une bissectrice de $\widehat{HMF}$, on a $\delta = \gamma$. Comme $\mathcal{T}_M$ et $\mathcal{N}_M$ sont perpendiculaires, on a $\beta + \gamma = \delta + \varepsilon = 90^\circ$. Enfin, $\beta + \gamma + \delta + \varepsilon = 180^\circ$ car ils forment un angle plat le long de $\mathcal{N}_M$. De toutes ces relations, on déduit que $\beta = \varepsilon$. Or α et ε sont opposés par le sommet, ils ont donc même mesure : d'après ce qui précède $\alpha = \varepsilon = \beta$. Autrement dit $\mathcal{N}_M$ est une bissectrice de l'angle entre $[My)$ et $[MF)$, lesquelles sont donc symétriques, ce qui correspond au critère de réflexion d'un rayon.



□

C Second degré : analyse

Table des matières

1	Racines d'une fonction polynôme	95
2	Discriminant d'une fonction polynôme du second degré	96
3	Forme factorisée et racines	97
4	Racines et représentation graphique	99
5	Étude du signe	100

1 – Racines d'une fonction polynôme

Définition 1

Racine

Soit f une fonction polynôme de degré 2.

Une *racine* de f est un réel r tel que $f(r) = 0$.

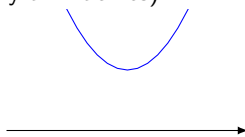
Exemple 2

Soit f définie pour tout $x \in \mathbb{R}$ par $f(x) = 2x^2 - 10x + 12$.

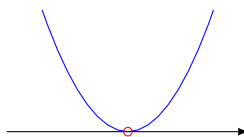
- ▶ $x = 2$ est une racine de f car $f(2) = 2 \times 2^2 - 10 \times 2 + 12 = 0$
- ▶ $x = 3$ est une racine de f car $f(3) = 2 \times 3^2 - 10 \times 3 + 12 = 0$
- ▶ $x = 4$ n'est pas une racine de f car $f(4) = 2 \times 4^2 - 10 \times 4 + 12 = 4 \neq 0$

Remarque 3

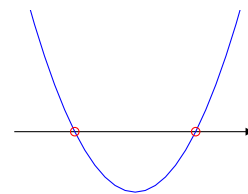
Les racines d'une fonction polynôme de degré 2 correspondent aux abscisses auxquelles la parabole coupe l'axe des abscisses. Graphiquement, on voit qu'il existe 3 cas : soit la parabole ne coupe pas l'axe des abscisses (auquel cas il n'y a pas de racine), soit la parabole touche juste l'axe des abscisses et reste toujours d'un côté ou de l'autre de ce dernier (auquel cas il y a 1 racine), soit la parabole traverse deux fois l'axe des abscisses (auquel cas il y a 2 racines).



Pas de racines



1 racine



2 racines

2 – Discriminant d'une fonction polynôme du second degré

Définition 4

Soit f une fonction polynôme de degré 2 définie sous forme développée par $f(x) = ax^2 + bx + c$.
Le *discriminant* de f est le réel :

$$\Delta = b^2 - 4ac$$

Théorème 5

Discriminant et racines

- ▶ Si $\Delta > 0$, alors f a deux racines réelles distinctes,
- ▶ Si $\Delta = 0$, alors f a une racine réelle,
- ▶ Si $\Delta < 0$, alors f n'a pas de racine réelle.

Démonstration. On adopte les notations introduites précédemment. Pour tout $x \in \mathbb{R}$, en partant de la forme canonique de f :

$$\begin{aligned} f(x) &= a(x - \alpha)^2 + \beta \\ &= a \left(x + \frac{b}{2a} \right)^2 + \frac{4ac - b^2}{4a} \\ &= a \left[\left(x + \frac{b}{2a} \right)^2 - \frac{b^2 - 4ac}{(2a)^2} \right] \\ &= a \left[\left(x + \frac{b}{2a} \right)^2 - \frac{\Delta}{(2a)^2} \right] \end{aligned}$$

- ▶ Supposons que $\Delta > 0$. Alors $\sqrt{\Delta}$ existe, et on peut écrire $\frac{\Delta}{(2a)^2} = \left(\frac{\sqrt{\Delta}}{2a} \right)^2$. Ainsi, pour tout $x \in \mathbb{R}$,

$$\begin{aligned} f(x) &= a \left[\left(x + \frac{b}{2a} \right)^2 - \left(\frac{\sqrt{\Delta}}{2a} \right)^2 \right] = a \left(x + \frac{b}{2a} + \frac{\sqrt{\Delta}}{2a} \right) \left(x + \frac{b}{2a} - \frac{\sqrt{\Delta}}{2a} \right) \\ &= a \left(x - \frac{-b - \sqrt{\Delta}}{2a} \right) \left(x - \frac{-b + \sqrt{\Delta}}{2a} \right) \end{aligned}$$

En posant $r_1 = \frac{-b + \sqrt{\Delta}}{2a}$ et $r_2 = \frac{-b - \sqrt{\Delta}}{2a}$, on a donc pour tout $x \in \mathbb{R}$, $f(x) = a(x - r_1)(x - r_2)$.

- ▶ Supposons que $\Delta = 0$. Alors pour tout $x \in \mathbb{R}$, $f(x) = a \left(x + \frac{b}{2a} \right)^2$. En posant $r_0 = -\frac{b}{2a}$, on a bien pour tout $x \in \mathbb{R}$, $f(x) = a(x - r_0)^2$.
- ▶ Supposons que $\Delta < 0$. Alors $-\Delta > 0$, et $-\frac{\Delta}{(2a)^2} > 0$. Comme pour tout $x \in \mathbb{R}$, $\left(x + \frac{b}{2a} \right)^2 \geq 0$, on a pour tout $x \in \mathbb{R}$,

$$\left(x + \frac{b}{2a} \right)^2 - \frac{\Delta}{(2a)^2} > 0$$

En particulier, cette quantité ne s'annule jamais ; la multiplier par $a \neq 0$ ne change rien à ce fait. D'après le calcul préliminaire, cela montre que $f(x)$ n'est jamais nul, autrement dit que f n'a aucune racine (réelle).

□

3 – Forme factorisée et racines

Lemme 6 (Nombre de racines)

Une fonction polynôme de degré 2 a au plus 2 racines, c'est-à-dire soit 0, soit 1, soit 2 racines.

Démonstration. Soient a , b et c trois réels avec $a \neq 0$, soit $f : x \mapsto ax^2 + bx + c$ une fonction polynôme de degré 2, supposons par l'absurde qu'il existe r_1 , r_2 et r_3 trois réels distincts tels que $f(r_1) = f(r_2) = f(r_3) = 0$. On traduit ceci dans un système :

$$\begin{aligned} \begin{cases} ar_1^2 + br_1 + c = 0 & (f(r_1) = 0) \\ ar_2^2 + br_2 + c = 0 & (f(r_2) = 0) \\ ar_3^2 + br_3 + c = 0 & (f(r_3) = 0) \end{cases} & \iff \begin{cases} ar_1^2 + br_1 + c = 0 & (L_1) \\ a(r_2^2 - r_1^2) + b(r_2 - r_1) = 0 & (L_2 \leftarrow L_2 - L_1) \\ a(r_3^2 - r_1^2) + b(r_3 - r_1) = 0 & (L_3 \leftarrow L_3 - L_1) \end{cases} \\ \iff (*) & \iff \begin{cases} ar_1^2 + br_1 + c = 0 & (L_1) \\ a(r_2 + r_1) + b = 0 & (\div (r_2 - r_1) \neq 0) \\ a(r_3 + r_1) + b = 0 & (\div (r_3 - r_1) \neq 0) \end{cases} \iff \begin{cases} ar_1^2 + br_1 + c = 0 & (L_1) \\ a(r_2 + r_1) + b = 0 & (L_2) \\ a(r_3 - r_2) = 0 & (L_3 \leftarrow L_3 - L_2) \end{cases} \end{aligned}$$

(*) : il s'agit de l'identité remarquable $r_i^2 - r_1^2 = (r_i + r_1)(r_i - r_1)$

Or $r_3 - r_2 \neq 0$ (car r_1 , r_2 et r_3 sont distincts), donc (L_3) implique que $a = 0$, contredisant le fait que f est une fonction polynôme de degré 2. $\square$

Remarque 7

À la fin de la preuve précédente, en injectant ceci dans (L_2) , on obtient immédiatement que $b = 0$. En injectant ces deux valeurs dans (L_1) , on trouve finalement que $c = 0$. Ainsi, pour tout $x \in \mathbb{R}$, $f(x) = 0x^2 + 0x + 0 = 0$, autrement dit, sous ces conditions, f est nécessairement la fonction nulle.

On peut montrer que si une fonction de la forme $g(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0$ admet au moins $n + 1$ racines, alors tous ses coefficients sont nuls, autrement dit g est la fonction nulle.

Théorème 8

Factorisation et racines

Soit $f(x) = ax^2 + bx + c$ une fonction polynôme de degré 2.

- Si f admet deux racines distinctes r_1 et r_2 , alors f admet la forme factorisée $f(x) = a(x - r_1)(x - r_2)$.
- Si f admet une unique racine r_0 , alors f admet la forme factorisée $f(x) = a(x - r_0)^2$.
- Si f n'admet pas de racines, alors elle n'admet pas de forme factorisée.

(Note importante : le a dans la forme factorisée est le même a que dans la forme développée)

Démonstration. Supposons que f admet deux racines distinctes que l'on note r_1 et r_2 . Soit $g : x \mapsto a(x - r_1)(x - r_2)$. Pour tout $x \in \mathbb{R}$,

$$g(x) = a(x^2 - r_1x - r_2x + r_1r_2) = ax^2 - a(r_1 + r_2)x + ar_1r_2$$

et donc pour tout $x \in \mathbb{R}$,

$$f(x) - g(x) = ax^2 + bx + c - ax^2 + a(r_1 + r_2)x - ar_1r_2 = [b + a(r_1 + r_2)]x + (c - ar_1r_2)$$

Ainsi $f - g$ est une fonction affine. Or $f(r_1) = f(r_2) = g(r_1) = g(r_2) = 0$, donc $f(r_1) - g(r_1) = 0$ et $f(r_2) - g(r_2) = 0$: en particulier, les points de coordonnées $(r_1, 0)$ et $(r_2, 0)$ appartiennent à la courbe de $f - g$. Or étant affine, sa courbe est une droite, dont on a trouvé deux points distincts : nécessairement, la courbe de $f - g$ est confondue avec l'axe des abscisses, et donc $f - g = 0$, autrement dit $f = g$. Cela démontre que f admet la forme factorisée :

$$f(x) = a(x - r_1)(x - r_2)$$

Supposons que f admet une unique racine que l'on note r_0 . Par souci de simplicité, supposons que $a > 0$ (le cas $a < 0$ est similaire). On note $f(x) = a(x - \alpha)^2 + \beta$ la forme canonique de f . Par l'absurde, si $\beta > 0$, alors pour tout $x \in \mathbb{R}$, $f(x) = \underbrace{a(x - \alpha)^2}_{\geq 0} + \beta \geq \beta > 0$, ce qui contredit le fait que $f(r_0) = 0$. Ainsi, $\beta \leq 0$: notons $\delta = -\frac{\beta}{a} \geq 0$.

Ainsi, pour tout $x \in \mathbb{R}$,

$$f(x) = a(x - \alpha)^2 + \beta = a[(x - \alpha)^2 - \delta] = a[(x - \alpha)^2 - (\sqrt{\delta})^2] = a(x - \alpha + \sqrt{\delta})(x - \alpha - \sqrt{\delta})$$

En particulier, $f(\alpha - \sqrt{\delta}) = f(\alpha + \sqrt{\delta}) = 0$, autrement dit $\alpha - \sqrt{\delta}$ et $\alpha + \sqrt{\delta}$ sont deux racines de f : or par hypothèse r_0 est l'unique racine de f , donc $r_0 = \alpha - \sqrt{\delta} = \alpha + \sqrt{\delta}$. En soustrayant ces deux quantités égales, on trouve $0 = 2\sqrt{\delta}$, on déduit que $\delta = 0$, *a fortiori* $\beta = 0$. De cela découle que $\alpha = r_0$. Ainsi, pour tout $x \in \mathbb{R}$, $f(x) = a(x - \alpha)^2 + \beta = a(x - r_0)^2$, ce qui est la forme factorisée recherchée.

Le dernier cas est immédiat par contraposée : si f admet une forme factorisée, alors f admet au moins une racine (il s'agit d'une équation produit). $\square$

Relations coefficients-racines, ou formules de Viète

Proposition 9

Si f admet une forme factorisée, alors celle-ci est unique. Plus précisément, si $f(x) = ax^2 + bx + c$ est la forme développée de f , et r_1 et r_2 sont deux racines de f (éventuellement égales), alors :

$$\triangleright r_1 + r_2 = -\frac{b}{a}$$

$$\triangleright r_1 r_2 = \frac{c}{a}$$

Démonstration. D'après le théorème précédent, si f admet une forme factorisée, alors celle-ci est de la forme $a(x - r_1)(x - r_2)$ (avec éventuellement $r_1 = r_2$, correspondant alors au r_0 du théorème). En développant :

$$f(x) = ax^2 - a(r_1 + r_2)x + ar_1 r_2$$

Ceci est une forme développée de f : par unicité des coefficients, on a, sans surprise, $a = a$, mais aussi $b = -a(r_1 + r_2)$ et $c = ar_1 r_2$, ce qui donne immédiatement les relations demandées. $\square$

Forme factorisée

Définition 10

La *forme factorisée* d'une fonction polynôme de degré 2, si elle existe, est une réécriture de f comme un produit de deux facteurs dits *linéaires*, plus précisément :

$$f(x) = a(x - r_1)(x - r_2)$$

où a est le coefficient dominant de f , et r_1 et r_2 sont des réels (éventuellement égaux).

Exemple 11

Vérifions que $x \mapsto 5(x - 6)(x + 2)$ est la forme factorisée de $x \mapsto 5x^2 - 20x - 60$.

$$\begin{aligned} 5(x - 6)(x + 2) &= 5[x \times x + x \times 2 - 6 \times x - 6 \times 2] \\ &= 5[x^2 - 4x - 12] \\ &= 5x^2 - 20x - 60 \end{aligned}$$

Méthode 12 (Factorisation en connaissant une racine)

Pour déterminer la forme factorisée de f lorsqu'on connaît une racine r , on nomme l'autre racine inconnue v (éventuellement $v = r$), on écrit $f(x) = a(x-r)(x-v)$, on développe l'expression, et on identifie les coefficients b et c pour déterminer v (ce qui revient à exploiter les relations coefficients-racines).

Exemple 13

- ▶ Déterminons une forme factorisée de $f(x) = 3x^2 - 15x + 18$ sachant que $x = 2$ en est une racine. Soit v la deuxième racine (éventuellement $v = 2$). On a alors $f(x) = 3(x-2)(x-v)$, c'est-à-dire :

$$\begin{aligned} 3x^2 - 15x + 18 &= 3(x-2)(x-v) = 3[x^2 - 2x - vx + 2v] \\ &= 3x^2 + 3(-2-v)x + 6v \end{aligned}$$

On en déduit que $-15 = 3(-2-v)$ et $18 = 6v$, et donc que $v = 3$.
Ainsi, $3x^2 - 15x + 18 = 3(x-2)(x-3)$.

- ▶ Déterminons une forme factorisée de $f(x) = 4x^2 - 24x + 36$ sachant que $x = 3$ en est une racine. Soit v la deuxième racine (éventuellement $v = 3$). On a alors $f(x) = 4(x-3)(x-v)$, c'est-à-dire :

$$\begin{aligned} 4x^2 - 24x + 36 &= 4(x-3)(x-v) = 4[x^2 - 3x - vx + 3v] \\ &= 4x^2 + 4(-3-v)x + 12v \end{aligned}$$

On en déduit que $-24 = 4(-3-v)$ et $36 = 12v$, et donc que $v = 3$.
Ainsi, $4x^2 - 24x + 36 = 4(x-3)(x-3) = 4(x-3)^2$.

4 – Racines et représentation graphique

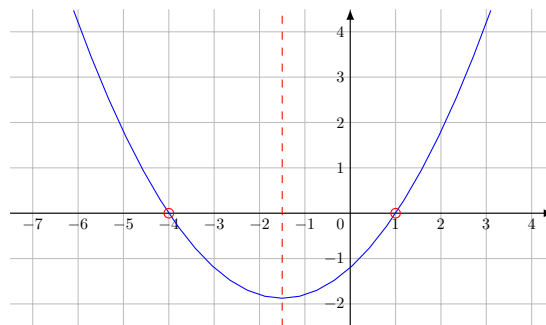
Proposition 14

L'axe de symétrie d'une parabole est toujours au milieu de ses racines, si elles existent.
Autrement dit : l'axe de symétrie de $f(x) = a(x-r_1)(x-r_2)$ (avec éventuellement $r_1 = r_2$) est la droite verticale d'équation $x = \frac{r_1 + r_2}{2}$.

Exemple 15

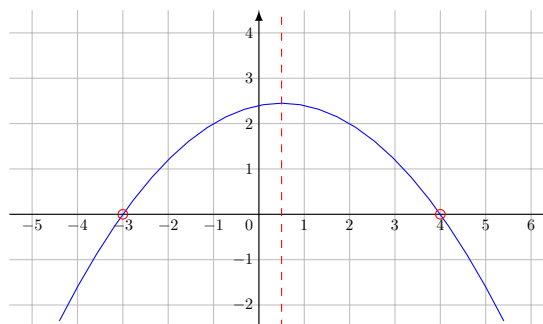
On lit sur le graphique que les racines de cette parabole sont -4 et 1 : l'axe de symétrie a donc pour équation :

$$x = \frac{-4 + 1}{2} = -1,5$$



On lit sur le graphique que les racines de cette parabole sont -3 et 4 : l'axe de symétrie a donc pour équation :

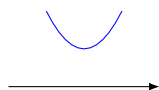
$$x = \frac{-3 + 4}{2} = 0,5$$



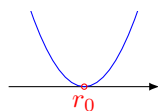
5 – Étude du signe

Méthode 16 (Tableau de signes)

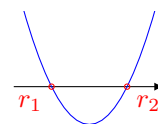
Une fonction polynôme de degré 2 est « du signe de a » en dehors de ses racines. Il est conseillé de réaliser un petit schéma au brouillon ou dans sa tête, avec la parabole orientée vers le haut ou vers le bas selon le signe de a , en y faisant apparaître l'axe des abscisses avec les racines.



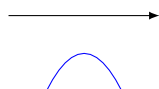
x	$-\infty$	$+\infty$
$f(x)$	+	



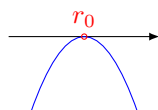
x	$-\infty$	r_0	$+\infty$
$f(x)$	+	○	+



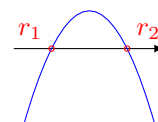
x	$-\infty$	r_1	r_2	$+\infty$	
$f(x)$	+	○	-	○	+



x	$-\infty$	$+\infty$
$f(x)$	-	



x	$-\infty$	r_0	$+\infty$
$f(x)$	-	○	-



x	$-\infty$	r_1	r_2	$+\infty$	
$f(x)$	$-$	$\circ$	$+$	$\circ$	$-$

D Logique mathématique

Table des matières

1	Pour la culture : la méthode axiomatique	102
A	Un « peu » d'histoire : l'exemple des axiomes d'Euclide	102
B	À quoi ça sert, la logique ?	103
2	Logique propositionnelle	104
A	Propositions mathématiques	104
B	Connecteurs logiques et quantificateurs	105
i	Les opérateurs logiques	105
ii	Les quantificateurs	106
iii	Les connecteurs logiques	106
3	Les différents types de raisonnement	108
A	Les contre-exemples	108
B	Le raisonnement par déduction	109
C	Le raisonnement par contraposée	109
D	Le raisonnement par l'absurde	111
E	Le raisonnement par disjonction de cas	111

Programme

Vocabulaire ensembliste et logique

L'apprentissage des notations mathématiques et de la logique est transversal à tous les chapitres du programme. Aussi, il importe d'y travailler d'abord dans des contextes où ils se présentent naturellement, puis de prévoir des temps où les concepts et types de raisonnement sont étudiés, après avoir été rencontrés plusieurs fois en situation.

Les élèves apprennent en situation à :

- ▶ reconnaître ce qu'est une proposition mathématique, à utiliser des variables pour écrire des propositions mathématiques ;
- ▶ lire et écrire des propositions contenant les connecteurs « et », « ou » ;
- ▶ formuler la négation de propositions simples, pouvant contenir un ou deux quantificateurs ;
- ▶ mobiliser un contre-exemple pour montrer qu'une proposition est fausse ;
- ▶ formuler une implication, une équivalence logique, et à les mobiliser dans un raisonnement simple ;
- ▶ formuler la réciproque d'une implication, ou sa contraposée ;
- ▶ lire et écrire des propositions contenant une quantification universelle ou existentielle (les symboles $\forall$ et $\exists$ ne sont pas exigibles) ;
- ▶ raisonner par disjonctions des cas, par l'absurde, par contraposée ;
- ▶ raisonner par équivalence, utiliser une propriété caractéristique ;
- ▶ distinguer condition nécessaire et condition suffisante ;
- ▶ démontrer une propriété par récurrence.

1 – Pour la culture : la méthode axiomatique

À quoi ça sert, la logique ? Pour répondre à cette question, il faut d'abord comprendre ce qu'est un axiome.

A – Un « peu » d'histoire : l'exemple des axiomes d'Euclide

Un théorème, c'est quelque chose qui part de faits supposés vrais (les hypothèses) pour produire un nouveau fait vrai (la conclusion). Les hypothèses peuvent elles-mêmes provenir d'autres théorèmes, et ainsi de suite... Mais alors, c'est une boucle sans fin ! En effet, pour produire un théorème, il faut donc disposer d'au moins un fait vrai qui sert de base, qui ne vient pas lui-même d'un théorème. Ces fondements vrais sont ce que l'on appelle des *axiomes*. Tout le monde est d'accord pour dire qu'ils sont vrais, et l'on se fonde là-dessus pour énoncer les théorèmes. C'est un peu comme des briques de Lego : les axiomes sont les briques élémentaires, on peut les assembler pour former de nouvelles structures – des théorèmes –, et on peut assembler ces structures entre elles pour faire des structures encore plus grandes, etc.

La *géométrie euclidienne* est la branche de la géométrie qui part des cinq axiomes suivants*, formulés par Euclide :

- 1► Par deux points passe une droite.
- 2► Tout segment peut être prolongé suivant sa direction en une droite.
- 3► Étant donné un segment de droite quelconque, on peut toujours tracer un cercle dont ce segment est un rayon.
- 4► Tous les angles droits sont égaux entre eux.
- 5► Si une droite tombant sur deux droites fait les angles intérieurs et du même côté plus petits que deux droits, les deux droites indéfiniment prolongées, se rencontreront du côté où sont les angles plus petits que deux droits.

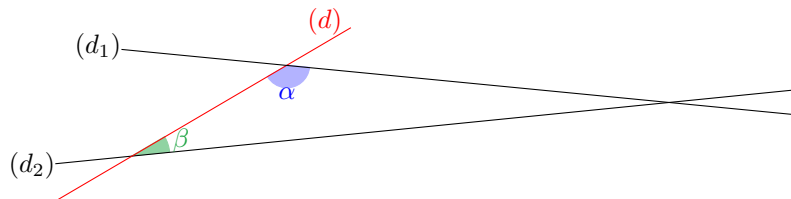
Notez ici une faille dans la rigueur : rappelons que nous partons du néant, aucun objet ni concept n'existe pour le moment. Ainsi, lorsque Euclide parle de points, de droites, de segments, ... dans ses axiomes, à quels objets fait-il référence ? Aucune définition n'est donnée actuellement !

Essayez de réfléchir à la définition d'une droite... Grâce au calcul vectoriel, on peut définir la droite passant par A et dirigée par $\vec{u}$ comme l'ensemble des points M du plan tels que $\overrightarrow{AM}$ et $\vec{u}$ sont colinéaires. Mais Euclide n'a certainement pas attendu 2000 ans pour que le calcul vectoriel soit inventé ! Sa définition – la première connue de l'Histoire – est la suivante : « Une ligne est une longueur sans largeur. La ligne droite est une ligne dont l'extension entre deux quelconques de ses points est égale à la distance entre ces points. ». Une formulation intriquée qui, si on y réfléchit bien, est aussi vide que circulaire.

Cette difficulté à définir rigoureusement un objet aussi simple qu'une droite illustre tout le défi que représente le fait de donner un cadre mathématique rigoureux à ce qui nous semble intuitif !

Ainsi, dans tout ce qui suit, on ne définira pas les objets les plus rudimentaires.

Le cinquième axiome est également étrange. Les quatre premiers semblent tous relativement naturels, et tout aussi facile à comprendre qu'à accepter, et le cinquième est le seul dont la formulation est obscure. Si l'on dessine la situation, et que l'on adopte des notations et un langage plus moderne, cela revient à dire que dans le schéma suivant, si $\alpha + \beta < 180^\circ$, alors (d_1) et (d_2) sont sécantes.



De nos jours, on lui trouve cette formulation équivalente† :

*. Certains les appellent « Postulat », et nomment « Axiomes » d'autres énoncés d'Euclide. Pour simplifier cette introduction déjà technique, nous esquivons cette nuance.

†. « Équivalente » signifie, dans ce contexte mathématique, que la véracité de l'une entraîne la véracité de l'autre. Autrement dit, l'une est conséquence de l'autre et inversement. Ainsi, deux propositions équivalentes sont interchangeables.

5► *Il existe une et une seule parallèle à une droite passant par un point donné.*

Cette singularité a laissé un goût amer dans la bouche des mathématiciennes et mathématiciens qui ont tenté de démontrer ce cinquième axiome à partir des quatre premiers, en vrai. Jusqu'à ce que János Bolyai se fasse une réflexion : et si cet axiome était faux ? Que se passe-t-il si l'on accepte l'existence d'une infinité de parallèles passant par un même point, ou, à l'inverse, à l'existence de points par lesquels aucune parallèle ne passe ? Ce faisant, Bolyai a inventé ce que l'on appelle aujourd'hui la *géométrie hyperbolique*. Pour en apprendre davantage, visionnez les vidéos [Les géométries non-euclidiennes | Voyage au pays des maths](#) de ARTE, ou [How One Line in the Oldest Math Text Hinted at Hidden Universes](#) de Veritasium sur YouTube.

B – À quoi ça sert, la logique ?

Plusieurs analogies pour comprendre :

- Analogie gourmande : la logique est le pain du sandwich. Si on change la garniture, on change le goût du sandwich ; mais pas le fait que c'est un sandwich. On peut imaginer que certains ingrédients vont bien ensemble, ou qu'au contraire leur combinaison créerait une atrocité ! Mais pour le vérifier, il faut assembler le sandwich, et le goûter. Traduction : on peut choisir différents axiomes, mais certaines combinaisons ne sont pas compatibles. Lorsqu'on essaye d'associer des axiomes incompatibles, on arrive à une absurdité, car l'association permet de démontrer que quelque chose est à la fois vrai et faux.
- AFOL (*Analogie Fan Of Lego*) : la logique, c'est la contrainte physique des briques. Si on essaye d'emboîter deux briques, on doit aligner les boutons de l'une avec les tubes de l'autre. Si on ne les aligne pas, soit ça ne s'emboîte pas, soit on casse les briques. En respectant cette règle universelle, on peut construire n'importe quoi... à condition de choisir les bonnes briques pour le faire ! Traduction : si on choisit des axiomes différents, on construit des mathématiques différentes, mais les règles pour « assembler » les axiomes sont les mêmes dans tous les cas.

Ce qu'il faut en retenir : la logique est fondamentale. Les règles des mathématiques sont celles dictées par la logique, mais, pour explorer différentes mathématiques, celle-ci ne se suffit pas à elle-même. En changeant les axiomes, on change l'univers mathématique.

2 – Logique propositionnelle

A – Propositions mathématiques

Définition 1

Proposition mathématique

Une *proposition mathématique* est une phrase (sujet + verbe + complément, écrite en français ou en mathématiques) qui est soit vraie, soit fausse.

Une proposition $\mathcal{P}$ peut dépendre d'une variable x : auquel cas on note $\mathcal{P}(x)$ ou $\mathcal{P}_x$ pour marquer la dépendance de $\mathcal{P}$ en x .

Exemple 2

Voici quelques exemples de propositions mathématiques dépendant ou non d'une variable, vraies ou fausses :

- $\mathcal{P}$: «Bordeaux est la capitale de la France.»
- $\mathcal{P}$: «Tous les triangles ont 3 côtés.»
- $\mathcal{P}(x)$: « $x > 3, 14$.»
- $\mathcal{P}_n$: «Le nombre $n^3 - n$ est divisible par 3.»

Définition 3

Négation

La *négation* d'une proposition $\mathcal{P}$ est la proposition $\mathcal{Q}$ construite avec les règles de la Proposition 15 (on la note parfois NON $\mathcal{P}$ ou $\neg \mathcal{P}$) qui est fausse quand $\mathcal{P}$ est vraie, et inversement.

$\mathcal{P}$	NON $\mathcal{P}$
V	F
F	V

Exemple 4

La négation de $\mathcal{P}_x$: « $x > 3, 14$ » est NON $\mathcal{P}_x$: « $x \leq 3, 14$ ».

Définition 5

Implication

Soient $\mathcal{P}$ et $\mathcal{Q}$ deux propositions.

On dit que $\mathcal{P}$ *implique* $\mathcal{Q}$ lorsque le fait que $\mathcal{P}$ soit vraie entraîne que $\mathcal{Q}$ est aussi vraie. On note alors $\mathcal{P} \implies \mathcal{Q}$. En français, cela correspond à la proposition «Si $\mathcal{P}$ est vraie, alors $\mathcal{Q}$ est aussi vraie».

La relation $\implies$ s'appelle une *implication*. Dans le cas où $\mathcal{P}$ implique $\mathcal{Q}$, on dit que :

- $\mathcal{P}$ est une *condition suffisante* à $\mathcal{Q}$ (il suffit que $\mathcal{P}$ soit vraie pour que $\mathcal{Q}$ le soit aussi),
- $\mathcal{Q}$ est une *condition nécessaire* à $\mathcal{P}$ (si $\mathcal{P}$ est vraie, alors nécessairement $\mathcal{Q}$ est aussi vraie).

Exemple 6

- «Si il fait beau, alors je vais à la plage.»
- «Si un polygone a 3 côtés, alors c'est un triangle.»
- $x \in \mathbb{R} \implies x^2 \geq 0$

Définition 7

Contraposée et réciproque

Soient $\mathcal{P}$ et $\mathcal{Q}$ deux propositions.

- La *contraposée* de $\mathcal{P} \implies \mathcal{Q}$ est l'implication NON $\mathcal{Q} \implies$ NON $\mathcal{P}$.
- La *réciproque* de $\mathcal{P} \implies \mathcal{Q}$ est l'implication $\mathcal{Q} \implies \mathcal{P}$.

Exemple 8

- ▶ Considérons l'implication «Si j'habite à Aubusson, alors j'habite en France».
 - ▶ Contraposée : «Si je n'habite pas en France, alors je n'habite pas à Aubusson.» (Vrai)
 - ▶ Réciproque : «Si j'habite en France, alors j'habite à Aubusson.» (Faux)
- ▶ Considérons l'implication «Si ABC est rectangle en A , alors $BC^2 = BA^2 + AC^2$ ».
 - ▶ Contraposée : «Si $BC^2 \neq BA^2 + AC^2$, alors ABC n'est pas rectangle en A .» (Vrai)
 - ▶ Réciproque : «Si $BC^2 = BA^2 + AC^2$, alors ABC est rectangle en A .» (Vrai)

Propriété 9

La contraposée d'une implication est toujours vraie. La réciproque peut-être soit vraie, soit fausse.

Définition 10

Soient $\mathcal{P}$ et $\mathcal{Q}$ deux propositions.

On dit que $\mathcal{P}$ et $\mathcal{Q}$ sont *équivalentes* lorsque $\mathcal{P} \implies \mathcal{Q}$ et $\mathcal{Q} \implies \mathcal{P}$. On note alors $\mathcal{P} \iff \mathcal{Q}$. En français, cela correspond à la proposition « $\mathcal{P}$ est vraie si, et seulement si $\mathcal{Q}$ est vraie».

La relation $\iff$ s'appelle une *équivalence*. Parfois, pour dire que $\mathcal{P} \iff \mathcal{Q}$, on dit que $\mathcal{P}$ est une *condition nécessaire et suffisante (CNS)* à $\mathcal{Q}$.

Exemple 11

Exemples d'équivalence :

- ▶ « ABC est rectangle en A si, et seulement si $BC^2 = BA^2 + AC^2$.»
- ▶ « n est pair $\iff n^2$ est pair.»

Exemples de non équivalence :

- ▶ «J'habite en France si, et seulement si j'habite à Aubusson» : c'est faux, car on peut habiter à Bordeaux, qui est aussi en France.
- ▶ « $x \geq 0 \iff x^2 \geq 0$ » : c'est faux, car $-3 < 0$ et $(-3)^2 = 9 \geq 0$.

B – Connecteurs logiques et quantificateurs

i – Les opérateurs logiques

Conjonction et disjonction

Définition 12

Soient $\mathcal{P}$ et $\mathcal{Q}$ deux propositions.

- ▶ La proposition $\mathcal{P}$ OU $\mathcal{Q}$ (parfois notée $\mathcal{P} \vee \mathcal{Q}$) est la proposition qui est vraie quand au moins une des deux propositions parmi $\mathcal{P}$ et $\mathcal{Q}$ est vraie.
- ▶ La proposition $\mathcal{P}$ ET $\mathcal{Q}$ (parfois notée $\mathcal{P} \wedge \mathcal{Q}$) est la proposition qui est vraie quand $\mathcal{P}$ et $\mathcal{Q}$ sont simultanément vraies.

$\mathcal{P}$	$\mathcal{Q}$	$\mathcal{P}$ OU $\mathcal{Q}$
V	V	V
V	F	V
F	V	V
F	F	F

$\mathcal{P}$	$\mathcal{Q}$	$\mathcal{P}$ ET $\mathcal{Q}$
V	V	V
V	F	F
F	V	F
F	F	F

ii – Les quantificateurs

Quantificateurs

Définition 13

- ▶ «(Pour) tout» est ce qu'on appelle le *quantificateur universel*. On le note $\forall$.
- ▶ «Il existe» est ce qu'on appelle le *quantificateur existentiel*. On le note $\exists$.

Exemple 14

Propositions utilisant le quantificateur universel (vraies ou fausses) :

- ▶ «Pour tout $x \in \mathbb{R}$, $x^2 \geq 0$.»
- ▶ «Tous les triangles ont 3 côtés.»
- ▶ « $\forall n \in \mathbb{N}$, n est pair.»

Propositions utilisant le quantificateur existentiel (vraies ou fausses) :

- ▶ «Il existe un éléphant rose sur Terre.»
- ▶ «Il existe $x \in \mathbb{R}$ tel que $x^2 < 0$.»
- ▶ « $\exists n \in \mathbb{N} : 4 = 2 \times n$.»

Propositions utilisant les deux quantificateurs (vraies ou fausses) :

- ▶ «Dans toutes les salles du lycée, il existe une fenêtre ouverte.»
- ▶ «Il existe une salle du lycée telle que toutes ses fenêtres sont ouvertes.»

Proposition 15

Négations des connecteurs et quantificateurs

- ▶ La négation d'un OU est un ET, et inversement.
- ▶ La négation d'un «Pour tout» est un «Il existe», et inversement.

Méthode 16 (Négation d'une proposition complexe)

Pour trouver la négation d'une proposition, on identifie chaque élément, et on formule leur négation sans changer l'ordre :

Pour tout $n \in \mathbb{N}$,	n^2 est pair	OU	$n(n+1)$ est pair
↓	↓	↓	↓
Il existe $n \in \mathbb{N}$ tel que	n^2 est impair	ET	$n(n+1)$ est impair

À retenir

« Quand il y a un "Pour tout", on écrit "Soit" » :

Quand on cherche à montrer que quelque chose est vrai, par exemple, *pour tout* $x \in \mathbb{R}$, la démonstration doit commencer par « Soit $x \in \mathbb{R}$ » (ou « Considérons $x \in \mathbb{R}$ », ou « Notons x un nombre réel quelconque »).

iii – Les connecteurs logiques

Implication et équivalence

Définition 17

Soient $\mathcal{P}$ et $\mathcal{Q}$ deux propositions.

- ▶ La proposition $\mathcal{P} \implies \mathcal{Q}$ (lue « $\mathcal{P}$ implique $\mathcal{Q}$ ») est une autre proposition, qui stipule que SI $\mathcal{P}$ est vraie, ALORS $\mathcal{Q}$ est aussi vraie.

- La proposition $P \iff Q$ (lue « P est équivalente à Q ») est une autre proposition, qui stipule que P et Q ont toujours la même valeur de vérité.

Remarque 18

On peut dresser les tables de vérité de l'implication et de l'équivalence, bien que pour la première, la situation où P est fausse est un peu contre-intuitive, et se traduit mal dans le langage courant.

P	Q	$P \implies Q$
V	V	V
V	F	F
F	V	V
F	F	V

P	Q	$P \iff Q$
V	V	V
V	F	F
F	V	F
F	F	V

Proposition 19

Double implication

$$P \iff Q \text{ a même valeur de vérité que } \begin{cases} P \implies Q \\ \text{ET} \\ Q \implies P \end{cases}$$

Application 20

En dressant des tables de vérité,

- 1► Démontrer la proposition précédente.
- 2► Démontrer que $\text{NON } (P \implies Q)$ a même valeur de vérité que $P \text{ ET } (\text{NON } Q)$.
- 3► Démontrer que $\text{NON } (P \text{ ET } Q)$ a même valeur de vérité que $(\text{NON } P) \text{ OU } (\text{NON } Q)$.
- 4► Démontrer que $P \text{ ET } (Q \text{ OU } R)$ a même valeur de vérité que $(P \text{ ET } Q) \text{ OU } (P \text{ ET } R)$ (on parle de distributivité!)

Remarque 21 (Dédution et implications)

Dans une démonstration par déduction, on utilise le mot «donc», et jamais le symbole $\implies$. En effet, ils ont des sens très différents :

- $P \implies Q$ signifie que «Si P est vraie, alors Q l'est aussi».
- « P donc Q » signifie que «Je sais que P est vraie, j'en déduis que Q l'est aussi».

Dans le cas d'une implication, on ne sait pas que P est vraie : on dit que dans le cas hypothétique où P serait vraie, alors Q le serait aussi.

Exemple 22

Illustrons cette remarque au travers d'un exemple élémentaire.

Question : On cherche à démontrer que pour tout réel x , si $x = 2$ alors $x^2 = 4$. Laquelle des deux rédactions ci-dessous est correcte ?

Soit x un nombre réel. Supposons que $x = 2$.

$$x = 2 \implies x^2 = 2^2 = 4$$

Finalement $x^2 = 4$.

Soit x un nombre réel. Supposons que $x = 2$.

$$x = 2 \text{ donc } x^2 = 2^2 = 4$$

Finalement $x^2 = 4$.

Réponse : c'est la rédaction de droite, avec le mot « donc ». Traduisons les deux démonstration en cheminement de pensée :

- ▶ On prend un nombre réel que l'on appelle x .
- ▶ On suppose que x est égal 2. C'est un fait désormais établi, à partir de maintenant x vaut 2 et rien d'autre.
- ▶ **SI** jamais x était égal à 2, **ALORS** dans ce cas le calcul de x^2 donnerait 4.
- ▶ On conclut que l'on a bien démontré ce que l'on attendait.

- ▶ On prend un nombre réel que l'on appelle x .
- ▶ On suppose que x est égal 2. C'est un fait désormais établi, à partir de maintenant x vaut 2 et rien d'autre.
- ▶ Comme on a fixé que x était égal à 2, par voie de conséquence, le calcul de x^2 donne 4.
- ▶ On conclut que l'on a bien démontré ce que l'on attendait.

Voyez-vous la différence ? Dans la rédaction erronée de gauche, au moment du calcul, on se replace sous l'hypothèse de **SI** x est égal à 2, alors qu'à l'étape précédente on avait fixé ce fait ! *A contrario*, dans la rédaction juste de droite, le mot « donc » signifie « ok, on m'a dit que x était égal à 2, quelle conséquences puis-je en tirer ? ».

À retenir : Bien que légèrement abusive, en Terminale on se contentera de retenir la règle suivante :

*Dans une démonstration, on n'utilise JAMAIS le symbole $\implies$.
(On utilise « DONC » pour exprimer la conséquence.)*

Application 23

Une réunion de mathématiciens et mathématiciennes du monde entier a lieu à Paris. Les mathématicien(ne)s américain(e)s portent tous et toutes une chemise rouge.

- 1▶ Un convive vêtu d'une chemise blanche franchit la porte. Est-il américain ?
- 2▶ Une convive vêtue d'une chemise rouge franchit la porte. Est-elle américaine ?
- 3▶ On annonce l'arrivée d'un convive allemand. Peut-il porter une chemise rouge ?
- 4▶ Une convive américaine arrive portant un manteau ? Peut-elle porter une chemise rouge ?

3 – Les différents types de raisonnement

A – Les contre-exemples

Méthode 24 (Utiliser un contre-exemple)

Pour montrer qu'une proposition est fausse, on peut chercher un *contre-exemple*, c'est-à-dire une situation qui met la proposition en défaut.

Exemple 25

À l'aide d'un contre-exemple, montrons que l'affirmation « Pour tout $x \in \mathbb{R}$, $x^2 - x - 1 > 0$ » est fausse :
On remarque que pour $x = 0$, on a $x^2 - x - 1 = -1 \leq 0$, ce qui est le contraire de l'affirmation, donc celle-ci est fausse.

Application 26 (Contre-exemples)

Montrer à l'aide de contre-exemples que les affirmations suivantes sont fausses :

- 1▶ « Si $x^2 = 9$, alors $x = 3$. »
- 2▶ « Si un nombre se termine par le chiffre 3, alors il est divisible par 3. »
- 3▶ « Si la somme des chiffres d'un nombre est un multiple de 7, alors ce nombre est multiple de 7. »

- 4► « Pour tous a, b, c et d entiers avec b, d et $b + d$ non nuls, $\frac{a}{b} + \frac{c}{d} = \frac{a+c}{b+d}$. »
- 5► « Pour tous réels positifs a et b , on a $\sqrt{a+b} = \sqrt{a} + \sqrt{b}$ »
- 6► « Si un quadrilatère a 4 côtés de même longueur, alors c'est un carré. »
- 7► « Pour tout $n \in \mathbb{N}$ tel que $n \geq 1$, le nombre $2n^2 + 29$ est premier. »

B – Le raisonnement par déduction

Raisonnement déductif

Définition 27

Le *raisonnement déductif* prend la forme suivante : partant d'un point de départ $\mathcal{A}$ (une hypothèse), on montre une conséquence de $\mathcal{A}$, puis une conséquence de cette conséquence, etc. en cascade jusqu'à arriver au point d'arrivée $\mathcal{B}$.

Un raisonnement par déduction repose essentiellement sur les connecteurs logiques «donc», «or» et «car» et leurs synonymes.

Exemple 28 (Rédaction-type)

Montrons que pour tout $n \in \mathbb{N}$, $n^3 - n$ est divisible par 3.

Soit $n \in \mathbb{N}$. On remarque que $n^3 - n = n(n^2 - 1) = n(n-1)(n+1)$, donc $n^3 - n$ peut s'écrire comme le produit de trois nombres entiers consécutifs. Or dans trois entiers consécutifs, au moins l'un est divisible par 3, donc leur produit est divisible par 3 (car au moins d'un des facteur l'est), donc $n^3 - n$ est divisible par 3.

Application 29

En conduisant un raisonnement déductif, démontrer les propositions suivantes :

- 1► « La somme de deux multiples de 7 est aussi un multiple de 7. »
- 2► « Les diagonales d'un carré sont perpendiculaires. » (Indication : utiliser la réciproque du théorème de Pythagore.)
- 3► « Une suite arithmétique de raison positive est croissante. »

C – Le raisonnement par contraposée

Contraposée d'une implication

Définition 30

Soient $\mathcal{P}$ et $\mathcal{Q}$ deux propositions.

La *contraposée* de l'implication $\mathcal{P} \implies \mathcal{Q}$ est l'implication $(\text{NON } \mathcal{P}) \implies (\text{NON } \mathcal{Q})$.

On rappelle que sans les notations mathématiques, l'implication s'écrit « **SI** $\mathcal{P}$ est vraie **ALORS** $\mathcal{Q}$ aussi », et sa contraposée s'écrit alors « **SI** $\mathcal{Q}$ est fausse **ALORS** $\mathcal{P}$ aussi. ».

Exemple 31

Vous avez vu en classe de quatrième que le théorème de Pythagore stipule que :

« **SI** ABC est rectangle en A **ALORS** $BC^2 = AB^2 + AC^2$. »

Sa contraposée est donc :

« **SI** $BC^2 \neq AB^2 + AC^2$ **ALORS** ABC n'est pas rectangle en A . »

À ne pas confondre avec la réciproque !

Proposition 32

La contraposée d'une implication a même valeur de vérité que celle-ci.

Application 33

Démontrer cette proposition à l'aide d'une table de vérité.

Méthode 34 (Raisonnement par contraposée)

Pour démontrer une implication, il est parfois plus simple de démontrer sa contraposée.

Exemple 35 (Rédaction-type)

Montrons par contraposée que pour tout $n \in \mathbb{N}$, si n^2 est pair, alors n est pair.
Supposons n impair. Il existe donc $k \in \mathbb{N}$ tel que $n = 2k + 1$. Par conséquent,

$$n^2 = (2k + 1)^2 = 4k^2 + 2k + 1 = 2(2k^2 + k) + 1$$

Comme $2k^2 + k \in \mathbb{Z}$, on en déduit que n^2 est impair.
Par contraposée, cela justifie que si n^2 est pair, alors n est pair.

Application 36

Formuler la contraposée des implications suivantes (on se fiche de savoir si elles sont vraies ou contre-intuitives!) :

- 1► « S'il pleut, alors le sol est mouillé. »
- 2► « S'il fait beau, alors je vais à la plage. »
- 3► « Si je travaille suffisamment, alors j'ai de bonnes notes. »
- 4► « Si toutes les fenêtres sont fermées, alors il n'y a pas de courant d'air. »
- 5► « Si au moins une porte est ouverte, alors je peux rentrer. »

Application 37

Démontrer les propositions suivantes grâce à leur contraposée :

- 1► « Pour tous entiers m et n , si $m \times n$ est impair, alors m et n sont impairs. »
- 2► « Pour tous m et n entiers naturels, si $m \times n = 1$, alors $m = 1$ et $n = 1$. »
- 3► « Pour tous x et y réels, si $xy = 0$ alors $x = 0$ ou $y = 0$. »

Application 38 (Nombres premiers de Mersenne)

Soit n un entier naturel non nul.

On suppose n composé, c'est-à-dire qu'il existe m et k deux entiers supérieurs ou égaux à 2 tels que $n = mk$.

- 1► Que vaut $1 + 2^k + 2^{2k} + \dots + 2^{(m-1)k}$?
- 2► En déduire une factorisation de $2^n - 1$, et que cet entier est composé.
- 3► On suppose $2^n - 1$ premier. Que peut-on en déduire sur n ?

D – Le raisonnement par l'absurde

Définition 39

Raisonnement par l'absurde

Raisonnement par l'absurde signifie que l'on suppose que la proposition que l'on souhaite montrer est fausse (autrement dit, on suppose vraie sa négation), et on montre que cette supposition *contredit* les hypothèses, autrement dit entraîne une *absurdité*.

Exemple 40 (Rédaction-type)

Montrons par l'absurde que $\sqrt{2}$ est irrationnel.

Supposons le contraire : supposons que $\sqrt{2}$ est rationnel. Il existe donc p et q deux nombres entiers strictement positifs et premiers entre eux tels que $\sqrt{2} = \frac{p}{q}$. On en déduit que $p = q\sqrt{2}$, puis que $p^2 = 2q^2$, donc p^2 est pair, donc p est pair (voir l'Exemple 35). On en déduit qu'il existe $k \in \mathbb{N}$ tel que $p = 2k$, donc $2q^2 = p^2 = (2k)^2 = 4k^2$, donc $q^2 = 2k^2$. Par conséquent, q^2 est pair, donc q est aussi pair : c'est absurde, car p et q sont supposés premiers entre eux.

Ainsi, $\sqrt{2}$ est irrationnel.

Application 41

Imaginons que vous devez démontrer les implications suivantes par l'absurde. Écrire la supposition initiale dans chaque cas.

- 1► « S'il pleut, alors le sol est mouillé. »
- 2► « S'il fait beau, alors je vais à la plage. »
- 3► « Si je travaille suffisamment, alors j'ai de bonnes notes. »
- 4► « Si toutes les fenêtres sont fermées, alors il n'y a pas de courant d'air. »
- 5► « Si au moins une porte est ouverte, alors je peux rentrer. »

Application 42

On rappelle que l'inverse d'un nombre x est l'unique nombre y tel que $x \times y = 1$, que l'on note alors $y = \frac{1}{x}$.

- 1► Démontrer par l'absurde que 0 n'a pas d'inverse.
- 2► Démontrer par l'absurde que 1 est le seul entier naturel dont l'inverse est entier.

E – Le raisonnement par disjonction de cas

Définition 43

Disjonction de cas

Dans une démonstration d'une proposition dépendant d'une variable, il est parfois utile de séparer plusieurs sous-cas complémentaires : par exemple, selon le signe de x , selon la parité de n , ...

Exemple 44 (Rédaction-type)

Montrons que le produit de trois entiers consécutifs est toujours divisible par 3.

Soit P le produit de trois entiers consécutifs, il existe $n \in \mathbb{Z}$ tel que $P = n(n+1)(n+2)$. Soient q et r les quotient et reste de la division euclidienne de n par 3, autrement dit $n = 3q + r$ et $0 \leq r < 3$. Il y a trois possibilités : soit $r = 0$, soit $r = 1$, soit $r = 2$.

- 1^{er} cas : $r = 0$

Dans ce cas, $n = 3q$ et $P = 3q \times (3q + 1) \times (3q + 2) = 3 \times \underbrace{[q(3q + 1)(3q + 2)]}_{\in \mathbb{Z}}$ est divisible par 3,

► **2^e cas** : $r = 1$

Dans ce cas, $n = 3q + 1$ et $P = (3q + 1) \times (3q + 2) \times (3q + 3) = 3 \times \underbrace{[(3q + 1)(3q + 2)(q + 1)]}_{\in \mathbb{Z}}$ est divisible par 3,

► **3^e cas** : $r = 2$

Dans ce cas, $n = 3q + 2$ et $P = (3q + 2) \times (3q + 3) \times (3q + 4) = 3 \times \underbrace{[(3q + 2)(q + 1)(3q + 4)]}_{\in \mathbb{Z}}$ est divisible par 3.

Dans tous les cas, P est divisible par 3.

Application 45

Démontrer par disjonction de cas les propositions suivantes :

- 1► « Pour tout $x \in \mathbb{R}$, le réel x^2 est positif. »
- 2► « Pour tout $x \in \mathbb{R}$, on a $\sqrt{x^2} = |x|$. »
- 3► Pour tout $n \in \mathbb{N}$, le nombre $\frac{n(n+1)}{2}$ est entier. »
- 4► « Pour tous x et y réels, $|x + y| \leq |x| + |y|$. » (*Indication : séparer les cas $x + y > 0$ et $x + y < 0$.*)
- 5► « Pour tous x et y réels, $\max\{x, y\} = \frac{1}{2}(x + y + |x - y|)$. »
- 6► « Le reste de la division euclidienne de n'importe quel carré par 3 n'est jamais 2. »
- 7► « Pour tout x réel, $|x - 1| \leq x^2 - x + 1$. »

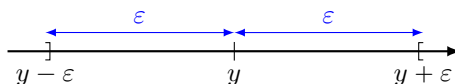
E

Compléments sur le logarithme

1 – Propriétés analytiques du logarithme : démonstration

A – Continuité

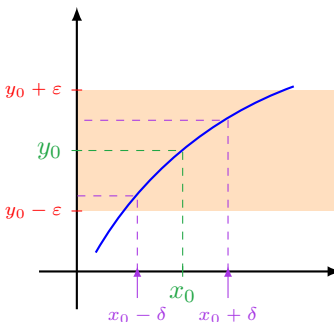
Démontrer la continuité d'une fonction est difficile en général. On rappelle qu'une fonction f est continue en un réel a lorsque $\lim_{x \rightarrow a} f(x) = f(a)$. En terminale, on ne donne qu'une définition « avec les mains » de ce que signifie une limite. On dit en général que, dans le cas présent, plus x se rapproche de a , plus $f(x)$ se rapproche de $f(a)$. Mais « se rapprocher » n'est pas une notion mathématique, et « plus quelque chose fait ça » l'est encore moins, du moins pas énoncés ainsi. Pour être plus rigoureux, on dira que quelque soit le niveau de rapprochement que l'on souhaite entre $f(x)$ et $f(a)$, aussi fin soit-il, on peut trouver un certain écart en-dessous duquel toutes les abscisses x proches de a avec cette précision font que $f(x)$ est proche de $f(a)$ avec le niveau de rapprochement que l'on désire. Pour ce qui est de la traduction mathématique du verbe « se rapprocher », il faut le penser en termes de distance, ce qui chez les réels se traduit grâce à la valeur absolue : on dira que x et y sont « proches au niveau de précision » ε lorsque $|x - y| < \varepsilon$, autrement dit lorsque $x \in]y - \varepsilon; y + \varepsilon[$.



Accrochez-vous, car il est temps d'écrire la formulation mathématique de la continuité du logarithme ! Pour montrer que $\ln$ est continue en $y_0 > 0$:

$$\forall \varepsilon > 0, \quad \exists \delta > 0 : \quad \forall y > 0, \quad |y - y_0| < \delta \implies |\ln(y) - \ln(y_0)| < \varepsilon$$

Re-traduisons cette ligne en français : quel que soit le niveau de rapprochement ε souhaité, on peut trouver un niveau de précision δ tel que si y est proche de y_0 avec une précision de δ au maximum, alors $\ln(y)$ est proche de $\ln(y_0)$ avec une précision de ε au maximum.



Remarquez le fait que les antécédents sont appelés y au lieu de x comme à l'accoutumée : en effet, les fonctions $\ln$ et $\exp$ étant réciproques l'une de l'autre, les antécédents de l'une sont les images de l'autre, et inversement. Et bien sûr, la continuité du logarithme va être une conséquence de celle de l'exponentielle.

Il est temps de démarrer la démonstration !

Par commodité, on notera $x_0 = \ln(y_0)$. On sait, par définition, qu'alors $y_0 = e^{x_0}$.

L'assertion que nous devons montrer comment par un « pour tout » : on écrit donc « soit ». Soit $\varepsilon > 0$. Le quantificateur suivant est un « il existe » : nous devons donc trouver un certain $\delta > 0$ qui satisfera nos désirs.

Étant pour le moment en manque de pistes à creuser, analysons notre point d'arrivée. Nous voulons :

$$\begin{aligned}
 & |\ln(y) - \ln(y_0)| < \varepsilon \\
 \text{C'est-à-dire} & \quad |\ln(y) - x_0| < \varepsilon \\
 \text{C'est-à-dire} & \quad -\varepsilon < \ln(y) - x_0 < \varepsilon \\
 \text{C'est-à-dire} & \quad -\varepsilon < \ln(y) - x_0 < \varepsilon \\
 \text{Autrement dit} & \quad x_0 - \varepsilon < \ln(y) < x_0 + \varepsilon \\
 \text{Ce que l'on aurait si} & \quad e^{x_0 - \varepsilon} < y < e^{x_0 + \varepsilon}
 \end{aligned}$$

par croissance de la fonction $\ln$. Prenons un peu de recul et regardons ce que l'on a écrit : pour arriver à la conclusion souhaitée, il nous suffirait de confiner y entre deux valeurs données... Cela ne vous semble-t-il pas familier ? C'est précisément la partie « $|y - y_0| < \delta$ », qui pour rappel se réécrit « $y_0 - \delta < y < y_0 + \delta$ », de ce que l'on cherche à montrer ! Nous devons donc trouver un δ qui pourrait lier $y_0 - \delta$ avec $e^{x_0 - \varepsilon}$, et $y_0 + \delta$ avec $e^{x_0 + \varepsilon}$.

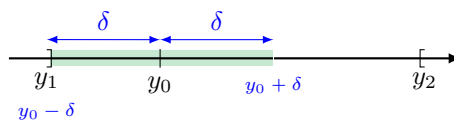
En général, la symétrie dans l'écriture d'une démonstration est un indicateur de bonne direction, essayons donc de la conserver. Dans le raisonnement ci-dessus, dans la dernière ligne, on se retrouve avec de l'exponentielle de part et d'autre de y qui – pauvre de lui – se retrouve à nu. Plutôt que de revenir en arrière en écrivant que $y = e^{\ln(y)}$ (qui fait réapparaître le $\ln(y)$ précédemment supprimé), nous allons plutôt nous débarrasser des exponentielles, ce qui aura de plus le mérite de simplifier les écritures. Posons $y_1 = e^{x_0 - \varepsilon}$ et $y_2 = e^{x_0 + \varepsilon}$, ce qui donnerait $y_1 < y < y_2$ pour la dernière ligne.

Remarquez aussi le brin de symétrie dans les définitions de y_1 et y_2 ... Ce $x_0 - \varepsilon$ et $x_0 + \varepsilon$ ne demandent rien de plus que notre attention ! Par ailleurs, on travaille avec y qui est le seul élément encore inconnu à ce stade, et nous n'avons pas exploité y_0 qui est notre point de départ. La connexion se fait ici toute seule : y_0 et x_0 sont liés par définition-même ! Afin de parachever cette parfaite symétrie, on meurt donc d'envie d'écrire :

$$\begin{aligned}
 & x_0 - \varepsilon < x_0 < x_0 + \varepsilon \\
 \text{Mais alors} & \quad e^{x_0 - \varepsilon} < e^{x_0} < e^{x_0 + \varepsilon} \\
 \text{C'est-à-dire} & \quad y_1 < y_0 < y_2
 \end{aligned}$$

par stricte croissance de la fonction exponentielle.

Ne perdez pas de vue que ce qui précède n'est rien de plus qu'un brouillon de réflexion visant à nous donner des pistes de recherche pour notre véritable objectif : trouver un $\delta > 0$ qui fonctionne ! Et bien heureusement, nos pérégrinations nous ont donné toutes les clés nécessaires. Celles-ci nous ont mené à écrire $y_1 < y_0 < y_2$. Représentons la situation sur un dessin :



Cette double inégalité indique que le segment $[y_1; y_2]$ est coupé en deux morceaux, de part et d'autre de y_0 . Prenons $\delta = \min\{y_0 - y_1; y_2 - y_0\}$, ce qui est une manière pompeuse de dire que δ est la distance et y_0 et le bord le plus proche du segment. De cette manière, si on prend l'intervalle $]y_0 - \delta; y_0 + \delta[$, par construction celui-ci est contenu dans $]y_1; y_2[$, autrement dit :

$$y_1 \leq y_0 - \delta < y_0 + \delta \leq y_2.$$

Il ne nous reste plus qu'à assembler tous les blocs : pour tout y tel que $|y - y_0| < \delta$, c'est-à-dire $y \in]y_0 - \delta; y_0 + \delta[$, autrement dit $y_0 - \delta < y < y_0 + \delta$, on a :

$$y_1 \leq y_0 - \delta < y < y_0 + \delta \leq y_2$$

et en particulier $y_1 < y < y_2$. Il ne nous reste plus qu'à redérouler la réflexion du début : sachant cette double inégalité, la stricte croissance de $\ln$ permet d'écrire $\ln(y_1) < \ln(y) < \ln(y_2)$, ce qui par définition de y_1 et y_2

donne $x_0 - \varepsilon < \ln(y) < x_0 + \varepsilon$, ce qui après quelques manipulations devient $|\ln(y) - x_0| < \varepsilon$, autrement dit $|\ln(y) - \ln(y_0)| < \varepsilon$. Nous voilà à destination, en bonne et due forme !

B – Dérivabilité et dérivée

Face au travail titanesque que la continuité a demandé, la dérivabilité semblera enfantine. Rappelons que $\ln$ est dérivable en $a > 0$ si $\lim_{x \rightarrow a} \frac{\ln(x) - \ln(a)}{x - a}$ existe et est finie, sa valeur donnant alors $\ln'(a)$. Comme à chaque fois, les propriétés du logarithme découle de celles de l'exponentielle, l'astuce est donc de la faire apparaître :

$$\lim_{x \rightarrow a} \frac{\ln(x) - \ln(a)}{x - a} = \lim_{x \rightarrow a} \frac{\ln(x) - \ln(a)}{e^{\ln(x)} - e^{\ln(a)}} = \lim_{y \rightarrow b} \frac{y - b}{e^y - e^b}$$

où $b = \ln(a)$ et avec le changement de variable $y = \ln(x)$. Attention, la validité d'un changement de variable n'est pas automatique : il faut vérifier que $y \rightarrow b$ lorsque $x \rightarrow a$. Et en effet, cela signifie que $\ln(x) \rightarrow \ln(a)$ lorsque $x \rightarrow a$, ce qui est vrai car c'est la définition de la continuité de $\ln$ en a , ce que nous avons démontré précédemment.

On reconnaît après le changement de variable l'inverse du taux d'accroissement de $\exp$ en b : on sait que $\exp$ est dérivable en b , donc :

$$\lim_{y \rightarrow b} \frac{e^y - e^b}{y - b} = \exp'(b) = e^b \neq 0 \quad \text{donc} \quad \lim_{y \rightarrow b} \frac{y - b}{e^y - e^b} = \frac{1}{e^b} \quad \text{donc} \quad \lim_{x \rightarrow a} \frac{\ln(x) - \ln(a)}{x - a} = \frac{1}{e^b} = \frac{1}{e^{\ln(a)}} = \frac{1}{a}$$

par continuité de la fonction inverse en $e^b \neq 0$ (exercice : expliquez cette justification!).

Ainsi, la limite du taux d'accroissement en a existe et est finie, donc $\ln$ est dérivable en a , de nombre dérivé $\ln'(a) = \frac{1}{a}$.

Comme a était quelconque, $\ln$ est dérivable sur $]0; +\infty[$, et pour tout $x > 0$, $\ln'(x) = \frac{1}{x}$.

2 – Définition alternative du logarithme et de l'exponentielle

L'une des nombreuses caractérisations du logarithme, la plus souvent admise comme définition, est celle stipulant que $\ln$ est l'unique primitive de $x \mapsto \frac{1}{x}$ sur $\mathbb{R}^{+*}$ qui s'annule en 1. Le théorème fondamental de l'analyse nous assure de la validité de cette définition car la fonction inverse est continue. Elle nous donne immédiatement la dérivabilité du logarithme, et par conséquent sa continuité que nous avons si difficilement démontrée.

Ce qui manque avec cette définition, c'est la relation de réciprocity avec l'exponentielle, et la relation fonctionnelle. Démontrons que cette définition donne bien accès à ces relations fondamentales.

Soit $y > 0$, soit φ la fonction définie sur $\mathbb{R}^{+*}$ par $\varphi(x) = \ln(xy) - \ln(x) - \ln(y)$. Elle est dérivable comme somme et composée de fonctions dérivables, et pour tout $x > 0$, $\varphi'(x) = y \ln'(xy) - \ln'(x) - 0 = y \times \frac{1}{xy} - \frac{1}{x} = 0$ (notez bien que y est fixé, et donc traité comme une constante, donc le premier terme est la fonction logarithme composée par la multiplication par y). Ainsi φ est de dérivée nulle sur l'intervalle $\mathbb{R}^{+*}$, il existe donc une constante $c \in \mathbb{R}$ telle que pour tout $x > 0$, $\varphi(x) = c$. Pour $x = 1$, cela donne $c = \varphi(1) = \ln(x) - \ln(x) - \ln(1) = 0$. Finalement, $\varphi(x) = 0$, et donc :

$$\forall x > 0, \quad \forall y > 0, \quad \ln(xy) = \ln(x) + \ln(y)$$

On définit alors l'exponentielle avec le théorème de la bijection continue à partir du logarithme, dans la logique symétrique de l'activité d'introduction. En effet $\ln$ est continue car dérivable par construction, et strictement croissante car de dérivée strictement positive. Ce crée définit la fonction réciproque du logarithme népérien, donc on montre la continuité de l'exponentielle de la même manière que dans la section précédente, en inversant simplement les $\ln$ et les $\exp$.

F Des lois de probabilité continues

1 – Notion de loi continue

Définition 1

Variable aléatoire continue

Une variable aléatoire (réelle) X est dite *continue* si elle peut prendre toutes les valeurs d'un intervalle de $\mathbb{R}$.

Remarque 2

Contrairement aux variables aléatoires discrètes, pour les variables aléatoires continues on ne calcule pas $\mathbb{P}(X = k)$, on calcule plutôt $\mathbb{P}(X \in [a, b]) = \mathbb{P}(a \leq X \leq b)$. Cela s'illustrera d'autant plus avec la prochaine définition.

Définition 3

Variables aléatoire à densité

On dit qu'une variable aléatoire est à *densité* s'il existe une fonction f positive telle que pour tout intervalle $[a, b]$,

$$\mathbb{P}(a \leq X \leq b) = \int_a^b f(t) dt$$

On dit que f est une *densité de probabilité*.

Dans toute la suite, on considérera une variable aléatoire X à densité, dont f est la densité de probabilité est continue.

Propriété 4

Pour tout $a \in \mathbb{R}$, $\mathbb{P}(X = a) = 0$.

Démonstration. En remarquant que $\mathbb{P}(X = a) = \mathbb{P}(a \leq X \leq a)$, par définition, $\mathbb{P}(X = a) = \int_a^a f(t) dt = 0$. $\square$

Remarque 5

D'après la probabilité précédente, la probabilité que X soit égale à une borne de l'intervalle est nulle. Ainsi :

$$\mathbb{P}(X \in [a, b]) = \mathbb{P}(X \in]a, b]) = \mathbb{P}(X \in [a, b[) = \mathbb{P}(X \in]a, b[)$$

Autrement dit :

$$\mathbb{P}(a \leq X \leq b) = \mathbb{P}(a < X \leq b) = \mathbb{P}(a \leq X < b) = \mathbb{P}(a < X < b)$$

Définition 6

Intégrale impropre

Pour tout $x \in \mathbb{R}$, sous réserve d'existence de la limite,

$$\int_x^{+\infty} f(t) dt = \lim_{y \rightarrow +\infty} \int_x^y f(t) dt \quad \int_{-\infty}^x f(t) dt = \lim_{y \rightarrow -\infty} \int_y^x f(t) dt \quad \int_{-\infty}^{+\infty} f(t) dt = \lim_{x \rightarrow +\infty} \int_{-\infty}^x f(t) dt$$

Variable à densité, extension de la définition

Définition 7

Pour $x \in \mathbb{R}$, on définit :

$$\mathbb{P}(X \leq x) = \mathbb{P}(X \in]-\infty, x]) = \int_{-\infty}^x f(t) \, dt \quad \text{et} \quad \mathbb{P}(X \geq x) = \mathbb{P}(X \in [x; +\infty[) = \int_x^{+\infty} f(t) \, dt$$

Également :

Propriété 8

$$\int_{-\infty}^{+\infty} f(t) \, dt = 1$$

Démonstration. Par définition, $\int_{-\infty}^{+\infty} f(t) \, dt = \mathbb{P}(X \in]-\infty, \infty[) = \mathbb{P}(X \in \mathbb{R})$

□

2 – Loi uniforme

3 – Loi exponentielle

APPLI : nombre de piqures de moustique en restant dehors un certain temps

On considère un temps d'attente exponentiel entre 2 piqures (POURQUOI ??)

T_i = temps d'attente à la i -ième piqure

S_n = somme T_k

N_t = nombre de piqures en t minutes

$(S_n > t) = (N_t < t)$

ADMIS : $f_{S_n}(x)$ = densité gamma n , lambda

IPP : $P(S_n > t) = \text{poisson}(\text{lambda} - 1) + P(S_n - 1 > t)$

On retourne avec N_t

Et pour N_t suit $P(\text{lambda}t)$

4 – Lois de puissance

Parler de "Benford" (voir vidéo)

5 – Loi normale

G La géométrie des courbes de fonctions

croissance, parité, translation horizontale, verticale, dilatation

Mathématiques

Terminale spécialité – Extras

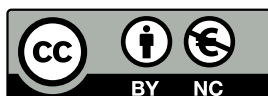
La série « Mathématiques – Terminale spécialité » est composée de quatre ouvrages :

- *Cours complet* est une version enrichie du cours, avec des remarques supplémentaires et toutes les démonstrations, qu'elles soient exigibles par le programme ou bien au-delà de sa portée, tout en restant abordables pour quiconque peut faire preuve de suffisamment de motivation et de patience. Cette version a été écrite avec plusieurs objectifs en tête. Permettre aux absents et absentes de rattraper le cours sans trop de pertes. Donner à celles et ceux qui souhaitent s'avancer, notamment pour le Grand Oral, suffisamment de matériel pour le faire en autonomie. Satisfaire l'appétit des esprits les plus gourmands, en culture comme en technique. Donner un aperçu de ce que sont les mathématiques dans le supérieur – et à ce titre rester une ressource pertinente plusieurs années après la terminale –.
- *Squelette de cours* est, comme son nom le suggère, une extraction de la version complète, qui ne contient que les essentiels du cours parfois à compléter, et les exercices. Il n'y a pas les démonstrations ni les détails supplémentaires. Cet élagage a pour objectif de créer un support de cours à la fois synthétique et fourni, afin d'optimiser la prise de notes en classe.
- *Extras* est un complément au *Cours complet*, avec des approfondissements, des points culture supplémentaires, et des annexes contenant des rappels « éclairés » (des notions vues les années passées, agrémentées de démonstrations permettant une compréhension plus profonde) et des outils qui ne sont pas ou plus au programme, mais qui peuvent rester utiles, notamment pour le Grand Oral.
- *Corrections* rassemble tous les exercices présents dans les feuilles de fin de chapitre, ainsi que leur correction plus ou moins rédigée. Il s'agit d'un support d'aide pour le travail en autonomie, aussi certains exercices ne bénéficient que d'éléments de correction. Voir la préface du manuel en question pour plus d'explications.

Cours complet (et *Squelette de cours* par conséquent) regroupe tous les cours et exercices dispensés lors de l'année de terminale, en spécialité mathématique. Les chapitres sont ordonnés selon une progression arbitraire, qui n'est pas la seule envisageable.

Cours & illustrations réalisés par E. Suong

Les exercices sont issus de sources multiples. Pour tout ajout de référence, adaptation ou retrait, merci de me contacter.



Version du 30 août 2026

