

MPSI

Notes de Cours

SZ, Lycée Jean Bart

Les présentes notes constituent le fil conducteur du cours de Maths de MPSI.

Il existe quelques différences entre ces notes et le cours fait en classe.

La première est que les exemples donnés en cours peuvent être différents, plus ou moins nombreux, en fonction des questions posées, des réactions, des constatations faites sur les précédents DS etc. . .

La seconde est qu'à la fin de chaque chapitre de ce pdf, il y a un "épilogue" plus ou moins long. Cet épilogue est constitué d'exercices d'application, d'approfondissements, ou de problèmes classiques de DS en relation avec le chapitre. Ne vous étonnez pas de la longueur de certains épilogues (par exemple, celui du chapitre sur le calcul intégral est aussi long que le cours proprement dit !) : il s'agit d'une compilation d'applications données les années précédentes, et il serait hors de question de toutes les traiter en classe la même année (par manque de temps).

Bonne lecture !

SZ

Table des matières

1	Éléments de logique et de théorie des ensembles	15
1.1	Éléments de logique	15
1.1.1	Négation, conjonction et disjonction	15
1.1.2	Implication et équivalence	16
1.1.3	Règles de négation	17
1.2	Notations et vocabulaire sur les ensembles	17
1.2.1	Appartenance	17
1.2.2	Quantificateurs	18
1.2.3	Inclusion	19
1.2.4	Intersection et union	19
1.2.5	Complémentaire et différence	20
1.2.6	Produit cartésien	21
1.3	Méthodes de démonstration	22
1.3.1	Axiome, postulat, théorème	22
1.3.2	Comment prouver une implication ?	22
1.3.3	Démonstration par l'absurde	23
1.3.4	Démonstration par analyse-synthèse	23
1.3.5	Démonstration par récurrence	25
1.4	Un problème : la porte logique NOR.	29
1.5	Epilogue : à propos des Mathématiques intuitionnistes	30
2	Méthodes algébriques	35
2.1	Brefs rappels sur les suites	35
2.2	Sommes	37
2.2.1	Généralités	37
2.2.2	Sommes remarquables	40
2.3	Produits	42
2.3.1	Généralités	42
2.3.2	Factorielle d'un entier naturel	43
2.4	Coefficients binomiaux	44
2.5	Epilogue : calcul de $\sum_{k=1}^n k^4$ et de $\sum_{k=1}^n k^5$	49

3	Trigonométrie	51
3.1	Congruences modulo un nombre réel	51
3.2	Fonctions trigonométriques	52
3.2.1	Les fonctions cosinus et sinus	52
3.2.2	La fonction tangente	53
3.3	Formulaire de trigonométrie	55
3.4	Quelques équations trigonométriques	57
3.4.1	L'équation $\cos(x) = \cos(y)$	57
3.4.2	L'équation $\sin(x) = \sin(y)$	58
3.4.3	L'équation $\cos(x) = \sin(y)$	58
3.4.4	L'équation $A \cos(x) + B \sin(x) = C$	58
3.5	Epilogue : valeurs exactes de $\cos(\pi/8)$, $\cos(\pi/16)$	59
4	Nombres complexes	61
4.1	Généralités	61
4.2	Conjugaison	63
4.3	Module	65
4.4	Ensemble des nombres complexes de module 1	69
4.4.1	Généralités	69
4.4.2	Lien avec la trigonométrie	70
4.4.3	Formules d'Euler et formule de Moivre	71
4.5	Argument(s) d'un nombre complexe NON NUL	76
4.6	Exponentielle complexe	79
4.7	Equations du second degré à coefficients complexes	80
4.7.1	Racines carrées complexes	80
4.7.2	Résolution de $az^2 + bz + c = 0$	81
4.8	Racines n -ièmes	83
4.8.1	Racines n -ièmes de l'unité	83
4.8.2	Racines n -ièmes d'un complexe	90
4.9	Similitudes directes	92
4.9.1	Transformations du plan	92
4.9.2	Ecritures complexes	92
4.9.3	Propriétés des similitudes directes	93
4.9.4	Le groupe des similitudes directes	95
4.10	Epilogue	95
4.10.1	Constructibilité à la règle et au compas des pentagone et heptakaïdécagone réguliers	95
4.10.2	Un calcul de somme classique : $\sum_{k=0}^{n-1} \cos(a + kh)$	103
4.10.3	Sinus hyperbolique complexe	104
5	Fonctions usuelles	107
5.1	Généralités	107
5.1.1	Définitions	107
5.1.2	Opérations élémentaires sur les fonctions	108

5.1.3	Transformations de la courbe représentative d'une fonction	109
5.1.4	Fonctions paires et impaires	111
5.2	Fonctions et inégalités	112
5.3	Valeur absolue	113
5.4	Fonctions dérivables (rappels de T ^{ale} et un peu plus...)	115
5.5	Dérivées d'ordre supérieur	121
5.6	Fonctions numériques usuelles	124
5.6.1	Fonction valeur absolue	124
5.6.2	Fonctions exponentielle et logarithme népérien	125
5.6.3	Fonctions trigonométriques hyperboliques	126
5.6.4	Fonctions logarithmes de base a	128
5.6.5	Fonctions exponentielles de base a	129
5.6.6	Fonctions puissances	131
5.7	Epilogue	133
5.7.1	Formulaire des dérivées usuelles	133
5.7.2	Formulaire de trigonométrie hyperbolique (HP)	134
5.7.3	Utilisation des puissances 1 : étude d'un Sup	135
5.7.4	Utilisation des puissances 2 : résolution d'une inéquation	135
5.7.5	Retour sur la technique de "l'angle-moitié"	136
5.7.6	Résolution d'une inéquation avec valeurs absolues	137
6	Applications	139
6.1	Généralités	139
6.2	Applications injectives, surjectives et bijectives	141
6.3	Bijections réciproques	142
6.4	Epilogue	144
6.4.1	Fonctions indicatrices	144
6.4.2	Deux pseudo-réciproques...	146
7	Fonctions circulaires réciproques	147
7.1	Fonctions à valeurs réelles bijectives	147
7.2	Fonctions trigonométriques réciproques	149
7.2.1	La fonction arccosinus	149
7.2.2	La fonction arcsinus	150
7.2.3	La fonction arctangente	152
7.2.4	Bilan sur les fonctions circulaires réciproques	154
7.3	Epilogue	157
7.3.1	Exercice classique : $\arctan\left(\frac{1}{3}\right) + \arctan\left(\frac{1}{4}\right) = \arctan\left(\frac{7}{11}\right)$	157
7.3.2	Somme d'une série	157
7.3.3	Formule de Machin	159
7.3.4	Problème des 13 réels	160

8	Méthodes de calcul intégral	163
8.1	Fonctions de classe $\mathcal{C}^1, \mathcal{C}^n, \mathcal{C}^\infty$	163
8.2	Primitives	164
8.3	Primitives usuelles	165
8.4	Intégrale d'une fonction continue sur un segment	166
8.5	Intégration par parties	168
8.6	Changement de variable	170
8.7	Compléments	173
8.7.1	Intégrale et parité	173
8.7.2	Intégrale et périodicité	174
8.7.3	Primitives de polynômes trigonométriques	176
8.7.4	Primitives de $x \mapsto \frac{1}{x^2 + bx + c}$ (b et c réels)	177
8.8	Extension aux fonctions à valeurs complexes	178
8.9	Epilogue	179
8.9.1	$e = \lim_{n \rightarrow +\infty} \sum_{k=0}^n \frac{1}{k!}$	179
8.9.2	Intégrales de Wallis	180
8.9.3	Formule de Stirling	184
8.9.4	Lemme de Riemann-Lebesgue	189
8.9.5	Calcul de $\zeta(2)$	190
9	Equations différentielles linéaires	197
9.1	Equations différentielles linéaires d'ordre 1	197
9.1.1	Généralités	197
9.1.2	Résolution d'une EDL 1 homogène	198
9.1.3	Recherche d'une solution particulière d'une EDL1 avec second membre	201
9.1.4	Conclusion - Problème de Cauchy	203
9.2	EDL2 à coefficients constants	205
9.2.1	Généralités	205
9.2.2	Résolution d'une EDL 2 homogène	206
9.2.3	Recherche d'une solution particulière	210
9.2.4	Conclusion	215
9.3	Compléments sur les équations différentielles	217
9.3.1	EDL2 et oscillateur harmonique	217
9.3.2	Géométrie des équations différentielles	221
9.3.3	Un exemple non linéaire : équation de Bernoulli	223
9.3.4	Un problème : résolution d'une équation différentielle linéaire d'ordre 3	223
10	Nombres réels	227
10.1	Relations binaires, relations d'ordre, relations d'équivalence	227
10.2	Bornes supérieures et inférieures dans l'ensemble des nombres réels	228
10.3	Conséquences de la propriété de la borne supérieure dans $\mathbb{R}$, notion de densité	231
10.4	Partie entière d'un nombre réel	233
10.5	Application — $\mathbb{D}$ est dense dans $\mathbb{R}$	235

10.6	Epilogue	236
10.6.1	$\mathbb{C}$ peut être totalement ordonné (si, si !)	236
10.6.2	Un exemple-clef de relation d'équivalence : les congruences en arithmétique	237
10.6.3	Une partie dense de $\mathbb{R}$	238
11	Suites	241
11.1	Généralités sur les suites réelles	242
11.2	Limite d'une suite réelle	243
11.2.1	Généralités sur les limites	243
11.2.2	Opérations algébriques sur les limites	245
11.2.3	Suites réelles de limite infinie	247
11.3	Suites réelles et inégalités	248
11.3.1	Limites et inégalités	248
11.3.2	Limites et monotonie	250
11.3.3	Applications des suites adjacentes	252
11.4	Suites extraites	254
11.5	Suites complexes	259
11.6	Suites particulières	261
11.6.1	Suites arithmétiques	261
11.6.2	Suites géométriques	261
11.6.3	Suites arithmético-géométriques	262
11.6.4	Suites récurrentes linéaires d'ordre 2	263
11.6.5	Suites récurrentes non-linéaires " $u_{n+1} = f(u_n)$ "	267
11.7	Epilogue	268
11.7.1	And the winner is !... (de nouvelles croissances comparées)	268
11.7.2	Convergence en moyenne (ou au sens de Cesàro)	269
11.7.3	Développement asymptotique à un terme de la série harmonique, et constante d'Euler	271
12	Structures algébriques	273
12.1	Lois de composition internes	273
12.1.1	Généralités	273
12.1.2	Élément neutre	274
12.1.3	Éléments inversibles	275
12.2	Groupes et sous-groupes	276
12.3	Anneaux et sous-anneaux	282
12.3.1	Présentation des anneaux	282
12.3.2	Sous-anneaux	286
12.3.3	Diviseurs de zéro	286
12.3.4	Éléments nilpotents	287
12.4	Corps et sous-corps	289
12.5	Epilogue	290
12.5.1	$\mathbb{Q}$ est le plus petit sous-corps de $\mathbb{R}$	290
12.5.2	Un théorème de Lagrange sur les groupes finis	291
12.5.3	Les groupes finis de cardinal premier sont abéliens	294
12.5.4	Arithmétique dans l'anneau des entiers de Gauss	295

13 Calcul matriciel	301
13.1 Les ensembles de matrices	301
13.1.1 Généralités	301
13.1.2 Opérations sur les matrices	302
13.2 Matrices carrées	306
13.2.1 Généralités	306
13.2.2 (Quelques) sous-anneaux remarquables de $M_n(\mathbb{K})$	309
13.2.3 Commutant d'une matrice carrée	313
13.2.4 Puissances d'une matrice carrée, et matrices carrées nilpotentes	314
13.2.5 Matrices symétriques et antisymétriques	318
13.2.6 Trace d'une matrice carrée	322
13.3 Matrices inversibles	323
13.3.1 Généralités	323
13.3.2 Cas particulier des matrices inversibles dans $M_2(\mathbb{K})$	324
13.3.3 Matrices inversibles et systèmes linéaires	326
13.3.4 Synthèse (nécessaire et suffisante ?) sur les matrices inversibles	331
13.4 Matrices semblables	334
13.5 Epilogue	343
13.5.1 Un problème classique de réduction	343
13.5.2 Interprétation matricielle du corps des nombres complexes	346
14 Limites et continuité	361
14.1 Limites des fonctions à valeurs réelles	361
14.1.1 Définitions	361
14.1.2 Propriétés des limites	363
14.2 Continuité des fonctions à valeurs réelles	367
14.3 Propriétés des fonctions continues à valeurs réelles	368
14.3.1 Propriétés générales	368
14.3.2 Propriété de continuité séquentielle	369
14.3.3 Lemme de Cauchy et théorème des valeurs intermédiaires	369
14.3.4 TVI : applications et conséquences	372
14.3.5 Théorème des bornes atteintes	373
14.4 Extension aux fonctions à valeurs complexes	374
14.5 Epilogue	376
14.5.1 Quelles sont les fonctions continues telles que $f(x + y) = f(x) + f(y)$?	376
14.5.2 Ensemble des points de discontinuité d'une fonction à valeurs réelles	376
14.5.3 Etude d'une suite récurrente	377
15 Ensembles finis et groupe symétrique	381
15.1 Ensembles finis	381
15.2 Opérations sur les ensembles finis	386
15.3 Injections et permutations	389
15.4 Groupe symétrique	390
15.4.1 Généralités	390
15.4.2 Transpositions et p -cycles	391

15.4.3	Signature et groupe alterné	397
15.5	Epilogue	399
15.5.1	Description de A_6	399
15.5.2	Simplicité de A_5	401
15.5.3	Dénombrabilité de $\mathbb{Q}$	407
16	Arithmétique	411
16.1	Divisibilité dans $\mathbb{Z}$	411
16.2	Division euclidienne dans $\mathbb{Z}$ et PGCD	413
16.3	Congruences	417
16.4	Entiers premiers entre eux, Lemme de Gauss	418
16.5	L'équation diophantienne $ax + by = c$	420
16.5.1	Equation homogène associée	420
16.5.2	Equation complète	420
16.6	PPCM	421
16.7	Nombres premiers	422
16.7.1	Définition, caractérisation	422
16.7.2	Petit théorème de Fermat	423
16.7.3	Décomposition en facteurs premiers d'un entier	424
16.7.4	Valuations p -adiques	426
16.8	Epilogue	426
16.8.1	Compléments historiques	426
16.8.2	Nombres de Carmichael	427
16.8.3	Codage RSA	429
16.8.4	Théorème de Wilson	432
16.8.5	Sous-groupes de $(\mathbb{Z}, +)$	433
17	Applications de la dérivation	435
17.1	Généralités	435
17.2	Dérivation et extremums — Théorème de Rolle	437
17.3	Dérivation et sens de variation — Théorème des accroissements finis	438
17.4	Lipschitzianité	441
17.5	Extension aux fonctions à valeurs complexes	443
17.6	Epilogue - Nature des séries de Riemann	444
18	Analyse asymptotique	447
18.1	Négligeabilité	448
18.1.1	Généralités	448
18.1.2	Croissances comparées	449
18.2	Domination (ou prépondérance)	450
18.3	Equivalents	451
18.4	Développements limités	453
18.4.1	Généralités	453
18.4.2	Construction — Formule de Taylor-Young	454
18.4.3	DL et opérations algébriques	457

18.5 Applications des développements limités	459
18.5.1 Calculs de limites	459
18.5.2 Tangentes et positions relatives	459
18.5.3 Recherches d'équivalents	459
18.5.4 Asymptotes et positions relatives	460
18.6 Appendice - DL à l'ordre n de arcsin	460
19 Polynômes et fractions rationnelles	465
19.1 Généralités	465
19.2 Degré et coefficient dominant d'un polynôme	467
19.3 Arithmétique dans $\mathbb{K}[X]$	472
19.3.1 Diviseurs et multiples, division euclidienne dans $\mathbb{K}[X]$	472
19.3.2 Polynômes associés	474
19.3.3 Racines et factorisation dans $\mathbb{K}[X]$	475
19.4 Fonctions polynomiales	475
19.4.1 Racines d'un polynôme	475
19.4.2 Dérivation (formelle) des polynômes	480
19.4.3 Racines multiples	485
19.4.4 Polynômes irréductibles	487
19.4.5 Polynômes scindés, relations coefficients-racines	493
19.5 Epilogue	494
19.5.1 Exemple d'équation polynomiale	494
19.5.2 Exemple de décomposition en irréductibles, dans $\mathbb{C}[X]$ et dans $\mathbb{R}[X]$	494
19.5.3 Division euclidienne dans $\mathbb{K}[X]$, et puissances d'une matrice	495
19.5.4 Utilisation des multiplicités pour une factorisation	496
19.5.5 Somme de deux binomiales indépendantes	497
19.5.6 Polynômes de Tchebychev	498
19.5.7 Théorème de d'Alembert-Gauss	500
20 Probabilités et variables aléatoires usuelles	505
20.1 Expérience aléatoire et univers	505
20.2 Opérations sur les événements	505
20.3 Probabilités	507
20.3.1 Généralités	507
20.3.2 Formule des probabilités totales	510
20.4 Probabilités conditionnelles et événements indépendants	511
20.5 Variables aléatoires réelles	516
20.5.1 Généralités	516
20.5.2 Espérance et variance d'une VAR	517
20.5.3 Propriétés de l'espérance et la variance	518
20.6 Variables aléatoires réelles : lois usuelles	521
20.6.1 Loi uniforme	521
20.6.2 Loi de Bernoulli	522
20.6.3 Loi binomiale	522
20.7 Epilogue	526

20.7.1	Problème — Sur la loi géométrique et les fonctions génératrices	526
20.7.2	Problème — Un processus de Markov à trois positions	538
21	Espaces vectoriels et applications linéaires	545
21.1	Espaces vectoriels	545
21.2	Sous-espaces vectoriels	547
21.3	Combinaisons linéaires, familles génératrices	550
21.4	Applications linéaires	552
21.5	Noyau et image d'une application linéaire	554
21.5.1	ker et Im	554
21.5.2	Isomorphismes et automorphismes	557
21.6	Sommes de sev et sev supplémentaires	558
21.6.1	Sommes de sev	558
21.6.2	Sev supplémentaires	559
21.6.3	Projections et symétries	561
21.7	Projecteurs	562
21.8	Epilogue — Etude des symétries dans un espace vectoriel	565
22	Espaces vectoriels de dimension finie	571
22.1	Familles libres et bases d'un espace vectoriel	571
22.2	Espaces vectoriels de dimension finie	573
22.2.1	Dimension	573
22.2.2	Théorème de la base extraite	576
22.2.3	Théorème de la base incomplète	577
22.3	Dimension d'une somme de sev	579
22.4	Coordonnées d'un vecteur dans une base, matrice de passage	582
22.4.1	Coordonnées	582
22.4.2	Changement de base et matrice de passage	582
22.5	Classification des ev de dimension finie	585
22.6	Epilogue	588
22.6.1	Exemples de sev supplémentaires	588
22.6.2	Un exemple de changement de base	588
22.6.3	Interpolation polynomiale et B-splines	590
23	Intégrale sur un segment	601
23.1	Continuité uniforme	601
23.2	Subdivisions	603
23.3	Fonctions en escalier	604
23.4	Fonctions continues par morceaux	605
23.5	Intégrale d'une fonction continue par morceaux	608
23.5.1	Intégrale d'une fonction en escalier	608
23.5.2	Propriétés de l'intégrale des fonctions en escalier	609
23.5.3	Intégrale d'une fonction continue par morceaux	610
23.5.4	Propriétés de l'intégrale des fonctions continues par morceaux	611
23.6	Propriétés de l'intégrale d'une fonction continue	613

23.7	Brève extension aux fonctions à valeurs complexes	615
23.8	Primitives et intégrales	616
23.9	Formule de Taylor avec reste intégrale et inégalité de Taylor-Lagrange	617
23.10	Epilogue	619
23.10.1	Vers les séries numériques	619
23.10.2	Méthodes d'intégration numérique	620
23.10.3	Буняковски, ou l'histoire d'un mathématicien oublié	633
24	Applications linéaires en dimension finie	635
24.1	Rang d'une famille, rang d'une application linéaire	635
24.2	Le théorème du rang et ses conséquences	636
24.3	Formes linéaires et hyperplans	638
24.3.1	Formes linéaires	638
24.3.2	Correspondance entre formes linéaires et hyperplans	640
24.4	Parenthèse : inverse à gauche, inverse à droite	643
24.5	Matrice d'une application linéaire	644
24.5.1	Définition et généralités	644
24.5.2	Compatibilité avec les opérations matricielles	648
24.6	Rang d'une matrice	650
24.7	Changement de base	651
24.8	Matrices équivalentes, matrices semblables	653
24.8.1	Matrices équivalentes	653
24.8.2	Matrices semblables	654
24.8.3	Trace d'un endomorphisme	656
24.9	Calcul pratique du rang d'une matrice	657
24.10	Sous-espaces affines d'un espace vectoriel	658
24.11	Equations linéaires	661
24.12	Epilogue	663
24.12.1	Suite des noyaux itérés	663
24.12.2	Espace vectoriel dual	664
24.12.3	Un nouvel invariant de similitude : le polynôme caractéristique	665
24.12.4	Une propriété des endomorphismes nilpotents	666
24.12.5	A propos des endomorphismes cycliques	667
24.12.6	Hyperplans dans $M_n(\mathbb{K})$	672
24.12.7	Transformée de Fourier discrète	676
24.12.8	Origines de la diagonalisation	677
25	Séries numériques	683
25.1	Généralités	683
25.2	Premières séries de référence	685
25.3	Séries à termes positifs	686
25.4	Convergence absolue	689
25.5	Epilogue	690
25.5.1	Stratégie d'étude : le point sur ce qu'il faut retenir sur les séries	690
25.5.2	Séries de Riemann	695

25.5.3	Séries de Bertrand	695
25.5.4	Séries alternées	697
Index		698

Chapitre 1

Eléments de logique et de théorie des ensembles

1.1 Eléments de logique

DÉFINITION 1.1 - (assertion logique). On appelle **assertion logique** tout assemblage de mots et de symboles obéissant à une syntaxe, à laquelle on peut associer une **valeur de vérité** : Vrai (V) ou Faux (F).

Deux assertions P et Q sont **synonymes ou logiquement équivalentes** si elles ont même valeur de vérité. On le note $P \equiv Q$.

► Règles logiques

On admet les règles suivantes :

- **Règle de non contradiction** : on ne peut avoir une assertion vraie et fausse en même temps.
- **Règle du tiers exclu** : une assertion qui n'est pas vraie est fausse et une assertion qui n'est pas fausse est vraie.

1.1.1 Négation, conjonction et disjonction

DÉFINITION 1.2 - (négation) La **négation** de l'assertion P est l'assertion, notée non P , $\overline{P}$ ou $\neg P$ qui est vraie lorsque P est fausse et fausse sinon.

La valeur de vérité de $\overline{P}$ dépend de celle de P , elles satisfont à une **table de vérité**.

P	$\overline{P}$
V	F
F	V

Montrer que $\overline{\overline{P}}$ est vraie c'est trouver un **contre-exemple** à P .

Conséquence (Double négation). $\overline{\overline{P}} \equiv P$

DÉFINITION 1.3 - (Conjonction et disjonction). Soient P et Q deux assertions.

• La **conjonction** des deux assertions P et Q , notée $P \text{ et } Q$ ou $P \wedge Q$, est l'assertion qui est vraie lorsque P et Q sont toutes les deux vraies, et fausse sinon.

• La **disjonction** des deux assertions P et Q , notée $P \text{ ou } Q$ ou $P \vee Q$, est l'assertion qui est vraie lorsqu'au moins l'une des deux assertions P ou Q est vraie, et fausse sinon.

P	Q	$P \text{ et } Q$	$P \text{ ou } Q$
V	V	V	V
V	F	F	V
F	V	F	V
F	F	F	F

Observons que pour toute assertion P , l'assertion $P \wedge \overline{P}$ est toujours fausse (non-contradiction), tandis que $P \vee \overline{P}$ est toujours vraie (tiers exclu).

PROPRIÉTÉ 1.1 - Soient P , Q et R trois assertions logiques. On a :

- 1) $P \wedge P \equiv P$ et $P \vee P \equiv P$
- 2) **Commutativité** : $P \wedge Q \equiv Q \wedge P$ et $P \vee Q \equiv Q \vee P$
- 3) **Associativité** : $P \wedge (Q \wedge R) \equiv (P \wedge Q) \wedge R$ et $P \vee (Q \vee R) \equiv (P \vee Q) \vee R$
- 4) **Distributivité** : $P \wedge (Q \vee R) \equiv (P \wedge Q) \vee (P \wedge R)$ et $P \vee (Q \wedge R) \equiv (P \vee Q) \wedge (P \vee R)$

PREUVE. Les preuves de ces propriétés, sans aucune difficulté, reposent sur l'écriture de tables de vérité idoines. ☕

1.1.2 Implication et équivalence

DÉFINITION 1.4 - (implication).

Soient P et Q deux assertions. On définit $P \implies Q$ par

$$[P \implies Q] \equiv [\overline{P} \text{ ou } Q]$$

$P \implies Q$ est vraie si P est fausse ou Q est vraie et fausse sinon.

P	Q	$P \implies Q$
V	V	V
V	F	F
F	V	V
F	F	V

► Vocabulaire associé aux implications

$P \implies Q$: P est l'hypothèse et Q la conclusion.

On dit que P (est vraie) implique Q (est vraie), si P (est vraie) alors Q (est vraie).

P est une **condition suffisante** pour avoir Q : pour avoir Q , **il suffit** d'avoir P .

Q est une **condition nécessaire** pour avoir P : pour avoir P , **il faut** avoir Q .

DÉFINITION 1.5 - (équivalence). Soient P et Q deux assertions.

On définit $P \iff Q$ par

$$[P \iff Q] \equiv [P \implies Q \text{ et } Q \implies P]$$

► Vocabulaire associé aux équivalences

$P \iff Q$: P est équivalente à Q .

P (est vraie) si et seulement si Q (est vraie).

P est une **condition nécessaire et suffisante** pour avoir Q : pour avoir P , **il faut et il suffit** d'avoir Q .

1.1.3 Règles de négation

PROPRIÉTÉ 1.2 - Règles de négation (lois de Morgan). Soient P et Q deux assertions.

$$\overline{P \text{ et } Q} \equiv \overline{P} \text{ ou } \overline{Q} \quad \text{et} \quad \overline{P \text{ ou } Q} \equiv \overline{P} \text{ et } \overline{Q}$$

PREUVE. La première loi ($\text{non}(P \vee Q) \equiv \text{non } P \wedge \text{non } Q$) peut être prouvée grâce à la table de vérité ci-dessous.

P	Q	$P \vee Q$	$\text{non}(P \vee Q)$	$\text{non } P \wedge \text{non } Q$
V	V	V	F	F
V	F	V	F	F
F	V	V	F	F
F	F	F	V	V

D'après cette table, puisque les assertions $\text{non}(P \vee Q)$ et $\text{non } P \wedge \text{non } Q$ ont les mêmes valeurs de vérité, elles sont logiquement équivalentes, soit : $\text{non}(P \vee Q) \equiv \text{non } P \wedge \text{non } Q$.

La seconde loi se démontre de façon analogue.

P	Q	$P \wedge Q$	$\text{non}(P \wedge Q)$	$\text{non } P \vee \text{non } Q$
V	V	V	F	F
V	F	F	V	V
F	V	F	V	V
F	F	F	V	V

Comme plus haut, on déduit de cette table l'équivalence logique : $\text{non}(P \wedge Q) \equiv \text{non } P \vee \text{non } Q$. ☕

Conséquence. Soient P et Q deux assertions. $\overline{P \implies Q} \equiv P \text{ et } \overline{Q}$

PREUVE. C'est un corollaire immédiat des lois de Morgan. ☕

Remarque. La négation d'une implication n'est donc pas une implication!!!

1.2 Notations et vocabulaire sur les ensembles

1.2.1 Appartenance

DÉFINITION 1.6 - (ensemble). Un **ensemble** est défini soit **extension** par énumération de ses éléments, soit **en compréhension** comme les objets vérifiant une certaine propriété.

Exemple. En extension : $\{\spadesuit, \clubsuit, \heartsuit, \diamondsuit\}, \{1; 4\}, [0, 1], \mathbb{N}, \mathbb{R}, \mathbb{C}$

En compréhension : $\{M \in \mathcal{P} : MA = MB\}, \{z \in \mathbb{C} : z^n = 1\}.$

Cas particulier (intervalle d'entiers). Soient n et m deux entiers relatifs. On pose

$$\llbracket n, m \rrbracket = \{k \in \mathbb{Z} \mid n \leq k \leq m\} = [n, m] \cap \mathbb{Z}$$

Notations. Lorsque x est un élément de E , on note $x \in E$ et on lit « x appartient à E ». Si x n'est pas un élément de E on note alors $x \notin E$.

DÉFINITION 1.7 - (ensemble vide, singleton). Un ensemble qui ne contient qu'un seul élément x est appelé **singleton** et se note $\{x\}$.

L'ensemble qui ne contient aucun élément est appelé **ensemble vide** et se note $\emptyset$ ou $\{\}$.

1.2.2 Quantificateurs

On considère une assertion P dépendant d'une variable x qui appartient à un ensemble E ; on la notera $P(x)$.

DÉFINITION 1.8 - (quantificateurs)

► On définit l'assertion

$$\forall x \in E, P(x)$$

comme l'assertion qui est vraie si pour tout x élément de E , $P(x)$ est vraie.

Le symbole $\forall$ est appelé **quantificateur universel** et est lu «quel que soit», «pour tout»

► On définit l'assertion

$$\exists x \in E, P(x)$$

comme l'assertion qui est vraie s'il existe au moins un x élément de E tel que $P(x)$ soit vraie.

Le symbole $\exists$ est appelé **quantificateur existentiel** et est lu «il existe au moins un ... »

► On définit l'assertion

$$\exists! x \in E, P(x)$$

comme l'assertion qui est vraie s'il existe un unique x élément de E tel que $P(x)$ soit vraie.

Le symbole $\exists!$ est appelé **quantificateur d'unicité** et est lu «il existe un unique ... »

Remarque. Attention!!! Dans une assertion contenant plusieurs quantificateurs on peut changer l'ordre de deux quantificateurs de même nature mais on ne peut changer l'ordre de deux quantificateurs de nature différentes.

Exemple. Comparer les deux assertions suivantes :

$$\forall x \geq 0, \exists y \geq 0, y = \sqrt{x} \quad \text{et} \quad \exists y \geq 0, \forall x \geq 0, y = \sqrt{x}.$$

PROPRIÉTÉ 1.3 - Règles de négation

$$\text{non } (\forall x \in E, P(x)) \equiv \exists x \in E, \overline{P}(x)$$

$$\text{non } (\exists x \in E, P(x)) \equiv \forall x \in E, \overline{P}(x)$$

Exemple. $\text{non } (\forall x \geq 0, \exists y \geq 0, y = \sqrt{x}) \iff \dots$

Convention. L'assertion « $\exists x \in \emptyset, P(x)$ » est fausse et l'assertion « $\forall x \in \emptyset, P(x)$ » est vraie.

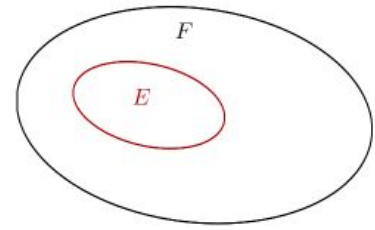
1.2.3 Inclusion

DÉFINITION 1.9 - Soient E et F deux ensembles.

On dit que E est **inclus** dans F si tout élément de E appartient à F . On le note $E \subset F$.

On dit également que F contient E , noté $F \supset E$ ou encore E est un **sous-ensemble** de F , une **partie** de F .

Dans le cas contraire, E est non inclus dans F et on note $E \not\subset F$.

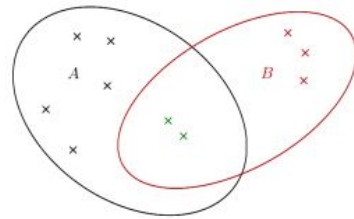


PROPRIÉTÉ 1.4 - .

➤ Pour tout ensemble E , on a $\emptyset \subset E$

➤ **Règle de double-inclusion.** $(E \subset F \text{ et } F \subset E) \iff (E = F)$

DÉFINITION 1.10 - Deux ensembles A et B sont dits **distincts** si $A \neq B$, c'est-à-dire s'ils ne contiennent pas les mêmes éléments.



DÉFINITION 1.11 - On appelle **ensemble des parties** de E , noté $\mathcal{P}(E)$, l'ensemble formé par tous les sous-ensembles de E .

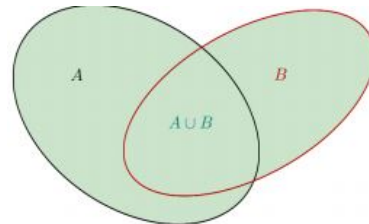
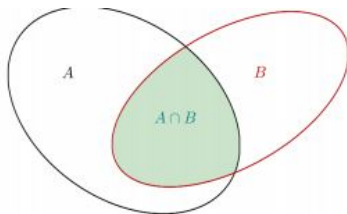
Exemple. $\mathcal{P}(\{1; 2\}) = \{\emptyset; \{1\}; \{2\}; \{1; 2\}\}$.

1.2.4 Intersection et union

DÉFINITION 1.12 - (**intersection et union**). Soient A et B deux parties d'un ensemble E .

➤ $A \cap B$ est l'**intersection de A et B** ; c'est l'ensemble des éléments qui appartiennent à A et B .

➤ $A \cup B$ est l'**union de A et B** ; c'est l'ensemble des éléments qui appartiennent à au moins l'un ensemble A et B .



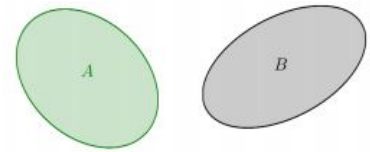
PROPRIÉTÉ 1.5 - (**commutativité et distributivité**). Soient A et B deux parties d'un ensemble E .

➤ $A \cap B = B \cap A$ et $A \cup B = B \cup A$

➤ $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ et $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$

PREUVE. C'est une traduction ensembliste de propriétés établies dans le cadre des assertions logiques.





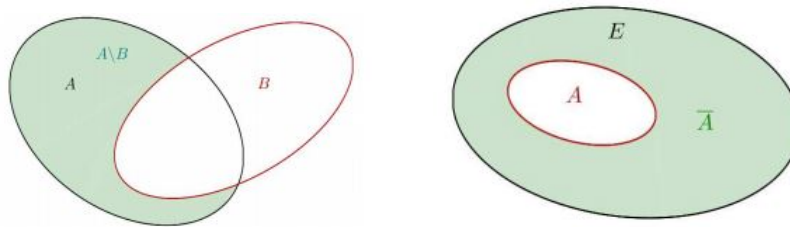
DÉFINITION 1.13 - Soient A et B deux parties d'un ensemble E .
 A et B sont dits **disjoints** si $A \cap B = \emptyset$.

1.2.5 Complémentaire et différence

DÉFINITION 1.14 - Soient A et B deux parties d'un ensemble E .

➤ $A \setminus B$ est la **différence de A par B** ; c'est l'ensemble des éléments de A qui n'appartiennent pas à B .

➤ $\bar{A}$ est le **complémentaire de A dans E** ; c'est l'ensemble des éléments de E qui n'appartiennent pas à A . Autres notations du complémentaire : $E \setminus A$ ou C_E^A .



PROPRIÉTÉ 1.6 - Soient A et B deux parties d'un même ensemble E .

➤ $\overline{(\bar{A})} = A$	➤ $\overline{A \cup B} = \bar{A} \cap \bar{B}$	➤ $A \cap \bar{A} = \emptyset$
➤ $\overline{A \cap B} = \bar{A} \cup \bar{B}$	➤ $\bar{\emptyset} = E$	➤ $A \cup \bar{A} = E$
➤ $A \setminus B = A \cap \bar{B}$		

Exemple d'application de la règle de double inclusion (et de double implication). Soient A et B deux parties d'un même ensemble E . Etablir que : $[A \subset B] \iff [A \cup B = B]$.

On raisonne par double implication pour établir l'équivalence de l'énoncé.

➤ $[A \subset B] \implies [A \cup B = B]$. Supposons donc que $A \subset B$, et prouvons l'égalité $A \cup B = B$.

Il est déjà clair que $B \subset A \cup B$. Etablissons l'inclusion réciproque : soit $x \in A \cup B$. Alors x est dans A ou B ; et puisque $A \subset B$, on en déduit que x est dans B . D'où l'inclusion $B \subset A \cup B$. Il s'ensuit (double inclusion) que : $A = B$.

Ainsi : $[A \subset B] \implies [A \cup B = B]$.

➤ $[A \subset B] \impliedby [A \cup B = B]$. Supposons donc que $A \cup B = B$. Si $x \in A$, alors a fortiori $x \in A \cup B$, donc $x \in B$ (puisque par hypothèse $A \cup B = B$). Par suite : $A \subset B$. Ce qui prouve l'implication $[A \subset B] \impliedby [A \cup B = B]$, et montre l'équivalence de l'énoncé. ☕

Exemple d'application des propriétés de l'inclusion, de l'union... Soient E un ensemble, et A et B deux parties de E . On définit la **différence symétrique de A et B** (et on note $A \Delta B$) la partie suivante :

$$A \Delta B = (A \setminus B) \cup (B \setminus A)$$

Etablir que : $A \Delta B = (A \cup B) \setminus (A \cap B)$.

Le but de l'exercice est d'établir que : $A \Delta B = (A \cup B) \setminus (A \cap B)$. On propose ci-dessous deux méthodes pour y parvenir.

► **Méthode 1 (par double inclusion)** — Commençons par prouver que $A \Delta B \subset (A \cup B) \setminus (A \cap B)$: soit donc x un élément de $A \Delta B$. Alors, par définition, deux cas sont possibles : soit $x \in A$ (et alors $x \notin B$), d'où $x \in A \cup B$ et $x \notin A \cap B$, c'est-à-dire $x \in (A \cup B) \setminus (A \cap B)$. L'autre cas est $x \in B$, qui se traite de façon analogue. On en déduit déjà que $A \Delta B \subset (A \cup B) \setminus (A \cap B)$

Montrons l'inclusion réciproque. Soit $x \in A$ (et alors $x \notin B$ sinon x serait dans $A \cap B$), d'où $x \in A \cup B$ et $x \notin A \cap B$, c'est-à-dire : $x \in (A \setminus B) \cup (B \setminus A)$. Soit $x \in B$ et un raisonnement analogue permet encore une fois de conclure. On en déduit l'inclusion réciproque, et donc finalement l'égalité : $A \Delta B = (A \setminus B) \cup (B \setminus A)$.

► **Méthode 2 (ensemblissement)** — On a :

$$A \Delta B = (A \setminus B) \cup (B \setminus A)$$

$$\iff A \Delta B = (A \cap \overline{B}) \cup (B \cap \overline{A})$$

$$\iff A \Delta B = [A \cup (B \cap \overline{A})] \cap [\overline{B} \cup (B \cap \overline{A})] \quad (\text{distributivité})$$

$$\iff A \Delta B = \left[(A \cup B) \cap \underbrace{(A \cup \overline{A})}_{=E} \right] \cap \left[\underbrace{(\overline{B} \cup B)}_{=E} \cap (\overline{B} \cup \overline{A}) \right] \quad (\text{re-distributivité})$$

$$\iff A \Delta B = [(A \cup B) \cap E] \cap [E \cap (\overline{B} \cup \overline{A})]$$

$$\iff A \Delta B = (A \cup B) \cap (\overline{A} \cup \overline{B})$$

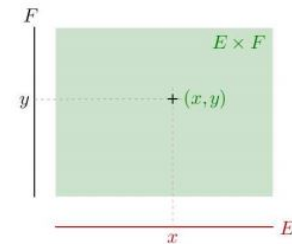
$$\iff A \Delta B = (A \cup B) \cap \overline{(A \cap B)} \quad (\text{loi de Morgan})$$

$$\iff A \Delta B = (A \cup B) \setminus (A \cap B)$$

Conclusion. $A \Delta B = (A \cup B) \setminus (A \cap B)$. ☕

1.2.6 Produit cartésien

DÉFINITION 1.15 - (produit cartésien). Soient E et F deux ensembles. On appelle **produit cartésien** de E par F , noté $E \times F$, l'ensemble des couples (x, y) avec $x \in E$ et $y \in F$.



Exemple : le rectangle $[0, 1] \times [-1, 1]$.

Cas particuliers. On note $E \times E$ par E^2 et plus généralement si $n \in \mathbb{N}^*$, $\underbrace{E \times \cdots \times E}_{n \text{ fois}}$ par E^n .

Exemple. Le plan peut donc être assimilé à $\mathbb{R}^2$ et l'espace à $\mathbb{R}^3$.

Vocabulaire. Soient $(E_i)_{i \in [1, n]}$ des ensembles. Les éléments de $E_1 \times E_2$ s'appellent des **couples**, de $E_1 \times E_2 \times E_3$ des **triplets** et plus généralement de $E_1 \times E_2 \times \cdots \times E_n$ des **n -uplets**.

1.3 Méthodes de démonstration

1.3.1 Axiome, postulat, théorème

DÉFINITION 1.16 - (axiome et postulat). On appelle **axiome** une assertion que l'on pose vraie a priori et **postulat** ou **conjecture** une assertion que l'on suppose vraie.

Exemple. L'axiome d'Euclide de la géométrie plane : «Par un point du plan on peut mener une unique parallèle à une droite donnée».

DÉFINITION 1.17 - (théorème ou tautologie). Un **théorème** de logique appelé aussi **tautologie** est une assertion vraie quelles que soient les valeurs de vérité des éléments qui la composent. Elle se déduit d'axiomes ou d'autres théorèmes.

Exemple. L'assertion $Q = [P \text{ ou } \overline{P}]$ est une tautologie.

1.3.2 Comment prouver une implication ?

Par syllogisme. Soient P, Q et R trois assertions logiques. On a :

$$[(P \implies Q) \text{ et } (Q \implies R)] \implies [P \implies R]$$

Exemple (Aristote, 350 av.JC). «Les hommes sont mortels, or tous les Grecs sont des hommes, donc tous les Grecs sont mortels».

Par disjonction de cas. Soient P, Q et R trois assertions logiques.

$$[(P \implies R) \text{ et } (Q \implies R)] \implies [(P \text{ ou } Q) \implies R]$$

Exemple. Un entier naturel et son carré ont la même parité.

Par contraposition. Soient P et Q deux assertions logiques.

$$[P \implies Q] \equiv [\overline{Q} \implies \overline{P}]$$

Cette dernière assertion est appelée **contraposée** de $P \implies Q$.

Exemple. Montrer que si $\frac{x}{1+x} < 0$ alors $x < 0$.

Conséquence. Soient P et Q deux assertions logiques.

$$P \iff Q \equiv \overline{P} \iff \overline{Q}$$

1.3.3 Démonstration par l'absurde

Principe. Pour montrer qu'une proposition P est vraie, on montre que sa négation $\overline{P}$ entraîne une proposition et son contraire. On aboutit alors à une contradiction. On peut conclure que l'hypothèse $\overline{P}$ est fausse et donc que P est vraie.

PROPRIÉTÉ 1.7 - $\sqrt{2} \in (\mathbb{R} \setminus \mathbb{Q})$ (càd : $\sqrt{2}$ est irrationnel).

PREUVE. Supposons que $\exists(p, q) \in (\mathbb{N}^*)^2$ tels que $\sqrt{2} = \frac{p}{q}$ et p et q premiers entre eux (c'est-à-dire sans diviseurs communs autres que ± 1).

On a donc : $p^2 = 2q^2$ (♠)

On en conclut que p^2 est pair, ce qui implique que p est pair¹. Par suite il existe un entier naturel k tel que : $p = 2k$.

La relation (♠) équivaut donc à : $2k^2 = q^2$.

On en déduit que q est pair.

Par suite, p et q sont divisibles par 2, ce qui est en contradiction avec l'hypothèse suivant laquelle p et q sont premiers entre eux.

Conclusion : $\sqrt{2} \in (\mathbb{R} \setminus \mathbb{Q})$. ☕

1.3.4 Démonstration par analyse-synthèse

Pour justifier l'existence et parfois l'unicité d'une solution, on peut être amené à déterminer la forme de celle-ci, forme qui n'est pas nécessairement donnée dans l'énoncé. On raisonne alors par analyse-synthèse :

➤ **Analyse** : on suppose qu'il existe au moins une solution, et on essaie d'en tirer le maximum de renseignements la concernant.

➤ **Synthèse** : on reporte dans le problème initial la ou les solution(s) trouvée(s) précédemment, ce qui permet de déterminer s'il y a bien une solution, puis une unique ou plusieurs.

➤ **Conclusion** : on énonce le résultat démontré.

THÉORÈME 1.1 - (de décomposition "paire + impaire"). Toute fonction $f : D \rightarrow \mathbb{R}$ avec $D \subset \mathbb{R}$ symétrique par rapport à 0 s'écrit de manière unique comme la somme d'une fonction paire et d'une fonction impaire.

PREUVE. • **Analyse** : supposons que $\forall x \in D, f(x) = p(x) + i(x)$ avec p paire et i impaire.

On a donc $\forall x \in D, f(-x) = p(-x) + i(-x) = p(x) - i(x)$.

En sommant on obtient $p(x) = \frac{f(x) + f(-x)}{2}$ et par différence, $i(x) = \frac{f(x) - f(-x)}{2}$.

Les fonctions p et i si elles existent sont donc uniques.

1. En vertu d'un résultat établi il y a peu !

- **Synthèse** : Posons $\forall x \in D, p(x) = \frac{f(x) + f(-x)}{2}$ et $i(x) = \frac{f(x) - f(-x)}{2}$.

On remarque que $\forall x \in D, p(x) + i(x) = f(x)$.

De plus, $p(-x) = \frac{f(-x) + f(x)}{2} = p(x)$ et $i(-x) = \frac{f(-x) - f(x)}{2} = -i(x)$ donc p est paire et i est impaire.

On a bien la décomposition voulue.

- **Conclusion** : toute fonction $f : D \rightarrow \mathbb{R}$ s'écrit de manière unique comme la somme d'une fonction paire et d'une fonction impaire. ☕

Remarque et exemple d'application. Le théorème de décomposition est un énoncé dit *constructif* dans le sens où il donne un moyen explicite d'obtenir l'écriture de f comme somme d'une fonction paire et d'une fonction impaire. Concrètement, dans le cas où la fonction f est la fonction exponentielle (définie sur $\mathbb{R}$, qui est un ensemble symétrique par rapport à zéro), on a :

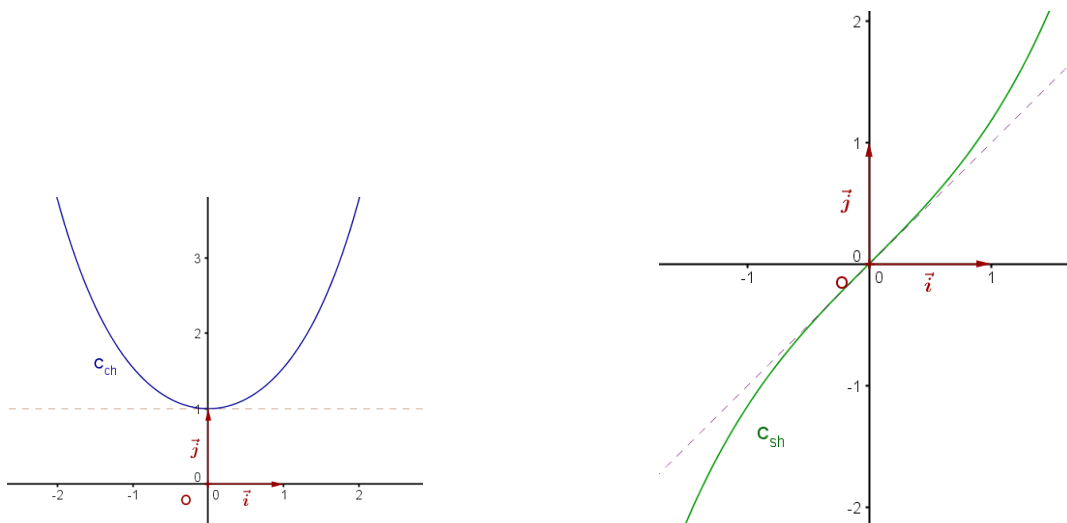
$$\forall x \in \mathbb{R}, e^x = \frac{e^x + e^{-x}}{2} + \frac{e^x - e^{-x}}{2}$$

La première des deux fonctions intervenant dans la somme ci-dessus est paire, et la seconde est impaire. Plus précisément, on appelle **cosinus hyperbolique** (*resp.* **sinus hyperbolique**) et on note **ch** (*resp.* **sh**) la première (*resp.* deuxième) fonction. On a donc :

DÉFINITIONS ET PROPRIÉTÉS.

- 1) **Définition du cosinus hyperbolique** : $\forall x \in \mathbb{R}, \text{ch}(x) = \frac{e^x + e^{-x}}{2}$
- 2) **Définition du sinus hyperbolique** : $\forall x \in \mathbb{R}, \text{sh}(x) = \frac{e^x - e^{-x}}{2}$
- 3) La fonction ch est paire, et la fonction sh est impaire.
- 4) En outre : $\exp = \text{ch} + \text{sh}$. Il revient au même d'écrire : $\forall x \in \mathbb{R}, e^x = \text{ch}(x) + \text{sh}(x)$.

Pour finir, voici à titre d'illustration les courbes représentatives des fonctions ch et sh .



1.3.5 Démonstration par récurrence

L'objectif de ce type de raisonnement est de prouver qu'une propriété $P(n)$ dépendant d'un entier naturel n est valable pour toute valeur de celui-ci. Cette méthode comporte quatre étapes :

- On donne un nom à la propriété que l'on souhaite prouver.
- **Initialisation** : on vérifie que la propriété $P(0)$ est vraie.
Remarque : dans certaines situations, la propriété $P(n)$ n'est valable que pour $n \geq 1$ (resp. pour $n \geq n_0$, avec n_0 un entier) ; dans ce cas, l'initialisation doit consister à vérifier que $P(1)$ est vraie (resp. que $P(n_0)$ est vraie). On parle alors de *récurrence décalée*.
- **Hérédité** : on suppose que $P(n)$ est vraie pour un certain entier (ou rang) n (c'est l'hypothèse de récurrence), et on prouve que la propriété $P(n+1)$ est vraie.
- **Conclusion** : on peut alors conclure que pour tout entier n la propriété $P(n)$ est vraie.

Remarque 1 : il existe plusieurs variantes du raisonnement par récurrence. La plus célèbre (la seule que vous connaissez a priori) est celle présentée plus haut, appelée *récurrence faible*.

Dans une *récurrence forte*, on suppose pour établir l'hérédité que la propriété est vraie jusqu'au rang n (et non pas simplement au rang n).

Dans une *récurrence double*, l'initialisation consiste à vérifier que $P(0)$ et $P(1)$ sont vraies ; et l'hérédité consiste à prouver que si $P(n)$ et $P(n+1)$ sont vraies, alors $P(n+2)$ l'est.

Dans une *récurrence finie*, l'initialisation consiste à vérifier que $P(0)$ est vraie ; et l'hérédité consiste à prouver que si $P(n)$ est vraie pour tout entier naturel $n \leq n_0$, alors $P(n+1)$ l'est. On prouve ainsi que pour tout entier $n \in \llbracket 0, n_0 + 1 \rrbracket$ l'assertion $P(n)$ est vraie.

Nous aurons l'occasion, en cours d'année, de rencontrer ces différentes variantes. En attendant, l'essentiel est que vous maîtrisiez le raisonnement par récurrence "classique" (la récurrence faible donc).

Remarque 2 : si, après la lecture des lignes précédentes, vous êtes très impatients de tester d'autres raisonnements par récurrence, voici ci-dessous une situation où l'on doit procéder par *récurrence double*.

Exercice. Soit (u_n) la suite réelle définie par $u_0 = 2$, $u_1 = 3$ et la relation de récurrence suivante : $\forall n \in \mathbb{N}, u_{n+2} = 3u_{n+1} - 2u_n$. Montrer que : $\forall n \in \mathbb{N}, u_n = 2^n + 1$.

Remarque 3 : si, après la lecture des lignes précédentes, vous êtes encore très impatients de tester d'autres raisonnements par récurrence, voici ci-dessous une situation où l'on doit procéder par *récurrence finie*.

Exercice. Soit N un entier naturel. On a : $\forall n \in \mathbb{N}, (x^N)^{(n)} = \begin{cases} \frac{N!}{(N-n)!} x^{N-n} & \text{si } n \in \llbracket 0, N \rrbracket \\ 0 & \text{si } n > N \end{cases}$

On donne ci-dessous quelques exemples classiques de raisonnement par récurrence.

Exemple 1 : somme des termes d'une suite arithmétique

PROPRIÉTÉ 1.8 - Soit (u_n) une suite arithmétique (de raison r avec $r \in \mathbb{C}$). On a :

$$\sum_{k=0}^n u_k = \frac{u_0 + u_n}{2} (n+1)$$

PREUVE. Démontrons la propriété par récurrence sur $n \in \mathbb{N}$.

► Notons $P(n)$ la propriété : $\sum_{k=0}^n u_k = \frac{u_0 + u_n}{2} (n+1)$.

► **Initialisation** : pour $n = 0$, on a d'une part $\sum_{k=0}^n u_k = \sum_{k=0}^0 u_k = u_0$, et d'autre part $\frac{u_0 + u_n}{2} (n+1) = \frac{u_0 + u_0}{2} (0+1) = u_0$. On en déduit que $P(0)$ est vraie.

► **Hérédité** : supposons $P(n)$ vraie pour un certain entier naturel n (hypothèse de récurrence). Alors :

$$\sum_{k=0}^{n+1} u_k = \left(\sum_{k=0}^n u_k \right) + u_{n+1} = \frac{u_0 + u_n}{2} (n+1) + u_{n+1} = \frac{(u_0 + u_n)(n+1) + 2u_{n+1}}{2} \quad (\heartsuit)$$

Puisque par ailleurs (u_n) est arithmétique de raison r , l'expression de son terme général est : $\forall n \in \mathbb{N}, u_n = u_0 + nr$. En utilisant cette expression, on peut réécrire la relation $(\heartsuit)$ comme suit :

$$\begin{aligned} \sum_{k=0}^{n+1} u_k &= \frac{(u_0 + u_0 + nr)(n+1) + 2(u_0 + (n+1)r)}{2} = \frac{(2u_0 + nr)(n+1) + 2(u_0 + (n+1)r)}{2} \\ &= \frac{2nu_0 + 2u_0 + n^2r + nr + 2u_0 + 2nr + 2r}{2} = \frac{2u_0(n+2) + r(n^2 + 3n + 2)}{2} \quad (\spadesuit) \end{aligned}$$

D'autre part, le terme de droite de l'égalité que l'on cherche à établir est :

$$\begin{aligned} \frac{u_0 + u_{n+1}}{2} (n+2) &= \frac{(u_0 + u_0 + (n+1)r)(n+2)}{2} = \frac{(2u_0 + (n+1)r)(n+2)}{2} \\ &= \frac{2u_0(n+2) + (n+1)(n+2)r}{2} = \frac{2u_0(n+2) + r(n^2 + 3n + 2)}{2} \quad (\clubsuit) \end{aligned}$$

La comparaison des expressions $(\spadesuit)$ et $(\clubsuit)$ permet alors d'affirmer que : $\sum_{k=0}^{n+1} u_k = \frac{u_0 + u_{n+1}}{2} (n+2)$; ce qui traduit le fait que la propriété $P(n+1)$ est vraie, et achève donc la preuve de l'hérédité.

► **Conclusion** : $\forall n \in \mathbb{N}, \sum_{k=0}^n u_k = \frac{u_0 + u_n}{2} (n+1)$ 

2. Ou : $\forall n \in \mathbb{N}, P(n)$ est vraie.

Exemple 2 : somme des cubes

PROPRIÉTÉ 1.9 - (somme des cubes). $\forall n \in \mathbb{N}, \sum_{k=0}^n k^3 = \frac{n^2 (n+1)^2}{4}$

PREUVE. Démontrons la propriété par récurrence sur $n \in \mathbb{N}$.

► Notons $P(n)$ la propriété : $\sum_{k=0}^n k^3 = \frac{n^2 (n+1)^2}{4}$.

► **Initialisation** : pour $n = 0$, on a d'une part $\sum_{k=0}^0 k^3 = 0$, et d'autre part $\frac{0^2 (0+1)^2}{4} = 0^3$. On en déduit que $P(0)$ est vraie.

► **Hérédité** : supposons $P(n)$ vraie pour un certain entier naturel n (hypothèse de récurrence). Alors :

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

Soit finalement : $\sum_{k=0}^{n+1} k^3 = \frac{(n+1)^2 (n+2)^2}{4}$, cette égalité signifiant que $P(n+1)$ est vraie, ce qui établit l'hérédité.

► **Conclusion** : $\forall n \in \mathbb{N}, \sum_{k=0}^n k^3 = \frac{n^2 (n+1)^2}{4}$ ☕

(Anti-)Exemple 3 : un cas où l'on peut éviter une récurrence — somme des termes d'une suite arithmétique

Le piège dans lequel il ne faut pas tomber est de penser que l'on ne peut démontrer une propriété dépendant d'un entier n que par récurrence sur n . Ce n'est évidemment pas toujours le cas, comme le prouve le raisonnement ci-dessous permettant d'établir la propriété de l'exercice 1 sur la somme des termes d'une suite arithmétique. On aurait en effet pu écrire pour tout $n \in \mathbb{N}$:

► d'une part : $\sum_{k=0}^n u_k = \sum_{k=0}^n (u_0 + kr) = \left(\sum_{k=0}^n u_0 \right) + r \left(\sum_{k=0}^n k \right) = (n+1)u_0 + r \frac{n(n+1)}{2}$

► d'autre part : $\frac{u_0 + u_n}{2} (n+1) = \frac{(u_0 + u_0 + nr)(n+1)}{2} = \frac{2u_0(n+1) + nr(n+1)}{2} = (n+1)u_0 + r \frac{n(n+1)}{2}$

Ce qui prouve la propriété directement, et nettement plus rapidement que dans le paragraphe 1. ☕

Exemple 4 (plus technique). Montrer que : $\forall n \in \mathbb{N}^*, \prod_{k=1}^n (2k+1)! \geq ((n+1)!)^{n+1}$

$$(\text{c-à-d} : 1! \times 3! \times \cdots \times (2n+1)! \geq ((n+1)!)^{n+1})$$

On raisonne par récurrence sur l'entier naturel non nul n pour établir la propriété de l'énoncé.

Notons donc $P(n)$ l'assertion : $\prod_{k=1}^n (2k+1)! \geq ((n+1)!)^{n+1}$.

► **Initialisation** : pour $n = 1$, on a d'une part $\prod_{k=1}^1 (2k+1)! = 3! = 6$, et d'autre part $((1+1)!)^{1+1} = 4$.

On en déduit que $P(1)$ est vraie.

► **Hérédité** : supposons $P(n)$ vraie pour un certain entier naturel non nul n . Alors :

$$\prod_{k=1}^{n+1} (2k+1)! = \left(\prod_{k=1}^n (2k+1)! \right) \times (2n+3)! \underset{HR}{\geq} ((n+1)!)^{n+1} (2n+3)!$$

Pour établir que $P(n+1)$ est vraie, il "n'y a plus qu'à" montrer que : $\boxed{((n+1)!)^{n+1} (2n+3)! \geq ((n+2)!)^{n+2}}$.

$$\text{En avant : } \frac{((n+1)!)^{n+1} (2n+3)!}{((n+2)!)^{n+2}} = \frac{((n+1)!)^{n+1} (2n+3)!}{((n+1)!)^{n+2} (n+2)^{n+2}} = \frac{(2n+3)!}{(n+1)! (n+2)^{n+2}}$$

$$= \frac{(n+1)! \prod_{k=2}^{n+3} (n+k)}{(n+1)! (n+2)^{n+2}} = \prod_{k=2}^{n+3} \frac{n+k}{n+2}$$

Or, pour tout entier $k \geq 2$, on a : $n+k \geq n+2$. D'où : $\forall k \in \llbracket 2, n+3 \rrbracket, \frac{n+k}{n+2} \geq 1$.

Par suite : $\prod_{k=2}^{n+3} \frac{n+k}{n+2} \geq 1$.

On en déduit que : $\frac{((n+1)!)^{n+1} (2n+3)!}{((n+2)!)^{n+2}} \geq 1$, c'est-à-dire : $\boxed{((n+1)!)^{n+1} (2n+3)! \geq ((n+2)!)^{n+2}}$.

Ce qui signifie que la propriété $P(n+1)$ est vraie, établit l'hérédité, et achève cette récurrence.

Conclusion : $\forall n \in \mathbb{N}^*, \prod_{k=1}^n (2k+1)! \geq ((n+1)!)^{n+1}$ ☕

1.4 Un problème : la porte logique NOR.

(PORTE NOR). Soient P et Q deux assertions logiques. On définit l'opérateur logique NOR par $\overline{P \vee Q}$ et on le notera par

$$P \downarrow Q = \overline{P \vee Q}$$

- 1) Dresser la table de vérité de $P \downarrow Q$.
- 2) Si A [respectivement B] est l'ensemble des éléments qui vérifient l'assertion P [respectivement l'assertion Q], quel est l'ensemble des éléments qui vérifient $P \downarrow Q$? (on pourra s'aider d'un dessin).
- 3) Montrer que $\overline{P}$ peut s'exprimer uniquement en fonction de P et du symbole $\downarrow$.
- 4) En déduire $(P \wedge Q)$ puis $(P \vee Q)$ uniquement en fonction de P , Q et du symbole $\downarrow$.
- 5) Exprimer $(P \implies Q)$ uniquement en fonction de P , Q et du symbole $\downarrow$.

Corrigé. Pour commencer, on peut observer que l'assertion $P \downarrow Q$, définie comme la proposition $\overline{P \vee Q}$, est logiquement équivalente à $\overline{P} \wedge \overline{Q}$ (c'est une des deux lois de Morgan).

- 1) Ci-contre, la table de vérité de $P \downarrow Q$.

P	Q	$\overline{P}$	$\overline{Q}$	$P \downarrow Q \equiv \overline{P} \wedge \overline{Q}$
V	V	F	F	F
V	F	F	V	F
F	V	V	F	F
F	F	V	V	V

- 2) Si A [respectivement B] est l'ensemble des éléments qui vérifient l'assertion P [respectivement l'assertion Q], alors l'ensemble des éléments qui vérifient $P \downarrow Q$ est $\overline{A \cup B}$ ou encore $\overline{A} \cap \overline{B}$.

- 3) Par définition de l'opérateur NOR, on a pour toute proposition mathématique P : $(P \downarrow P) \equiv (\overline{P} \wedge \overline{P})$ d'où : $(P \downarrow P) \equiv \overline{P}$.

- 4) Soient P et Q deux assertions mathématiques. On a :

► d'une part : $(P \downarrow Q) \downarrow (P \downarrow Q) \equiv \overline{(P \downarrow Q)} \wedge \overline{(P \downarrow Q)} \equiv \overline{(P \downarrow Q)} \equiv \overline{\overline{P} \wedge \overline{Q}} \equiv \overline{\overline{P}} \vee \overline{\overline{Q}} \equiv P \vee Q$;

► et d'autre part : $(P \downarrow P) \downarrow (Q \downarrow Q) \equiv \overline{(P \downarrow P)} \wedge \overline{(Q \downarrow Q)} \equiv \overline{(\overline{P} \wedge \overline{P})} \wedge \overline{(\overline{Q} \wedge \overline{Q})} \equiv P \wedge Q$

Conclusion : $P \wedge Q \equiv (P \downarrow P) \downarrow (Q \downarrow Q)$ et $P \vee Q \equiv (P \downarrow Q) \downarrow (P \downarrow Q)$.

- 5) Puisque : $[P \implies Q] \equiv \overline{P} \vee Q$, on déduit des questions 3 et 4 que :

$$[P \implies Q] \equiv (P \downarrow Q) \equiv [(P \downarrow P) \downarrow Q] \downarrow [(P \downarrow P) \downarrow Q]$$

1.5 Epilogue : à propos des Mathématiques intuitionnistes

Le texte reproduit ci-dessous (à titre culturel) est celui d'un article dû à Éric MORVAN, enseignant au Lycée Bertran-de-Born (Périgueux), et publié dans le Bulletin de l'UPS numéro 258 (2017).

Se moquer du tiers comme du milieu

Mathématiques constructives

Les mathématiciens font un grand usage du principe du tiers exclu, à savoir que toute proposition P est vraie ou fausse. Mais certains mathématiciens dits *intuitionnistes*, à la suite de Luitzen Egbertus Jan Brouwer en 1907, critiquent l'utilisation de ce principe, considérant qu'une proposition P ne peut être déclarée vraie que lorsqu'on en a construit une preuve directe. Par là preuve directe ils entendent un raisonnement constitué d'une succession de pas déductifs permettant de passer d'une proposition P_0 dont on sait qu'elle est vraie à la proposition P .

Les mathématiques intuitionnistes sont dites aussi mathématiques constructives car, dans ce cadre, la preuve de l'existence d'un objet mathématique passe par la possibilité de le construire, par exemple en fournissant un algorithme à cet effet.

"Constructive mathematics is distinguished from its traditional counterpart, classical mathematics, by the strict interpretation of the phrase "there exists" as "we can construct"."

Pour Brouwer, le principe du tiers exclu a le défaut de permettre d'affirmer que toute proposition P est vraie ou fausse, même si l'on ne connaît aucune démonstration de sa véracité ou de sa fausseté. On attribue à Brouwer l'adage : "there are no non-experienced truths". Sa position était en effet de refuser la possibilité qu'un théorème soit démontré par application de règles de logique sans qu'il soit nécessaire de prendre en compte son contenu sémantique.

L'opposition à l'intuitionnisme fut farouche, à commencer par celle de Hilbert : "[Accepter le programme de Brouwer] revient à démembrer et mutiler notre science ; si nous suivions de tels réformateurs, nous courrions le danger de perdre un grand nombre de nos trésors les plus précieux" (1922). Ou encore :

"Enlever le principe du tiers exclu aux mathématiciens serait la même chose, disons, que d'interdire le télescope aux astronomes ou aux boxeurs l'usage des poings." (1928)

Une proposition peut être vraie, fausse ou indécidable

On dit que Brouwer a introduit le temps en mathématiques. Au lieu de considérer qu'une proposition est vraie ou fausse de toute éternité, un mathématicien intuitionniste considère qu'elle est (peut-être temporairement) indécidable tant qu'elle n'a pas été prouvée (de manière constructive) vraie ou fausse.

Toutefois, s'ils rejettent le principe du tiers exclu, les mathématiciens intuitionnistes conservent le principe de non-contradiction : P ne peut pas être simultanément vraie ou fausse. De ce fait, il y a en logique intuitionniste trois cas possibles qui s'excluent mutuellement.

- (i) On peut construire une preuve de P ; on dira alors que P est vraie.
- (ii) À partir de l'hypothèse P , on peut construire une contradiction de P ; on dira que P est fausse.
- (iii) On ne peut ni prouver que P est vraie, ni prouver que P est fausse ; on dira que P est indécidable.

Ayant abandonné le principe du tiers exclu les mathématiciens intuitionnistes ont en échange besoin de quelques postulats consistant à déclarer que certaines propositions sont toujours décidables. Il en est ainsi de l'égalité dans $\mathbb{N}$; pour a et b entiers naturels, il n'y a que deux cas : $a = b$ ou $a \neq b$. De même le principe de récurrence est conservé.

► Un exemple classique : nous voulons démontrer qu'il existe deux nombres irrationnels a et b tels que a^b est rationnel. La preuve en logique classique repose sur le statut de vérité de la proposition P : " $\sqrt{2}^{\sqrt{2}}$ est un nombre rationnel".

1er cas : on suppose que P est vraie, c'est-à-dire que $\sqrt{2}^{\sqrt{2}}$ est rationnel. Alors on peut choisir $a = \sqrt{2}$ et $b = \sqrt{2}$.

2nd cas : on suppose que P est fausse, c'est-à-dire que $\sqrt{2}^{\sqrt{2}}$ est irrationnel. Alors :

$$\left(\sqrt{2}^{\sqrt{2}}\right)^{\sqrt{2}} = \sqrt{2}^{\sqrt{2} \times \sqrt{2}} = \sqrt{2}^2 = 2 \text{ est rationnel.}$$

On peut donc choisir $a = \sqrt{2}^{\sqrt{2}}$ et $b = \sqrt{2}$.

Il est difficile de ne pas être frustré par une telle démonstration, qui prouve l'existence d'un couple (a, b) de nombres irrationnels tels que a^b soit rationnel, mais ne permet pas de décider si le bon choix est $(a, b) = (\sqrt{2}, \sqrt{2})$ ou $(a, b) = \left(\sqrt{2}^{\sqrt{2}}, \sqrt{2}\right)$.

Cette démonstration est évidemment non constructive et n'est pas concluante dans le cadre de la logique intuitionniste puisqu'elle ne prend pas en compte le 3e cas : P peut être indécidable.⁴

Renoncer au raisonnement par l'absurde

Soit un ensemble E dans lequel, pour tout élément x , on considère une propriété $\mathcal{P}(x)$. Pour démontrer qu'il existe au moins un élément x_0 de E tel que $\mathcal{P}(x_0)$ soit vraie on peut supposer que, pour tout x de E , $\mathcal{P}(x)$ est fausse. Si ceci entraîne une contradiction, on en déduira en vertu du principe du tiers exclu qu'il existe au moins un élément x_0 de E tel que $\mathcal{P}(x_0)$ soit vraie.

Un tel raisonnement par l'absurde est inadmissible pour un mathématicien constructiviste puisqu'il ne fournit aucune méthode pour calculer x_0 . Effectivement, en logique intuitionniste, le raisonnement par l'absurde est impossible, puisque celui-ci consiste à supposer qu'une proposition P est fausse et à faire apparaître une contradiction. En logique classique, on en déduit que P n'est pas fausse et donc qu'elle est vraie. En logique intuitionniste, on en déduit seulement que le cas (ii) est impossible et donc que P est vraie (cas (i)) ou indécidable (cas (iii)).

4. On sait grâce à une démonstration due à Kuzmin que $\sqrt{2}^{\sqrt{2}}$ est irrationnel.

Il convient toutefois de distinguer le raisonnement par l'absurde, où on suppose P fausse, du raisonnement où on suppose P vraie et où on fait apparaître une contradiction (par une preuve directe). Un tel raisonnement est parfaitement accepté en mathématiques constructives puisqu'il consiste simplement à établir le cas (ii) : P est fausse.

Exemple : pour démontrer que $\sqrt{2}$ est un nombre irrationnel, on suppose que la proposition P : “ $\sqrt{2}$ est un nombre rationnel” est vraie. On sait qu'il existe une écriture $\sqrt{2} = p/q$, où p et q sont deux entiers premiers entre eux. Or on peut prouver que p et q sont tous les deux pairs (voir propriété 1.7, page 23) ce qui contredit le fait qu'ils sont premiers entre eux.

On en déduit que P est fausse : $\sqrt{2}$ n'est pas un nombre rationnel.

La négation d'une proposition

En logique classique, la proposition non- P est définie comme la proposition qui est vraie si et seulement si P est fausse. En logique intuitionniste, on convient de dire que non- P est vraie lorsque P est fausse. Pour simplifier les énoncés qui suivent, on abrègera les notations en remplaçant P vraie et P fausse par P et non- P .

En logique classique, $\text{non}(\text{non-}P) \equiv P$, d'où la possibilité de raisonner par l'absurde.

En logique intuitionniste, P (cas (i)) et non- P (cas (ii)) sont incompatibles. Donc, si P , supposer non- P entraîne une contradiction :

$$P \implies \text{non}(\text{non-}P)$$

Mais l'implication réciproque est fausse en général puisque non-(non- P) signifie seulement qu'il est contradictoire d'affirmer que P est fausse. Restent alors deux possibilités : P est vraie ou indécidable.

Exemple : considérons une proposition P du type “l'ensemble X est non vide”. Pour prouver non-(non- P), il suffit de démontrer que l'hypothèse non- P (càd que X est vide) conduit à une contradiction. En revanche, en mathématiques constructives, pour prouver P , il faut une démonstration permettant d'exhiber un élément de X .

De ce fait les mathématiciens intuitionnistes distinguent les ensembles habités (P est vraie) et les ensembles non vides (non-(non- P) est vraie). Tout ensemble habité est non vide, mais un ensemble non vide n'est pas forcément habité...

Les lois de Morgan

En logique classique, la première loi de Morgan est : $(\text{non-}P \text{ ou } \text{non-}Q) \equiv \text{non-}(P \text{ et } Q)$.

En logique intuitionniste, l'implication suivante reste vraie :

$$(\text{non-}P \text{ ou } \text{non-}Q) \implies \text{non-}(P \text{ et } Q).$$

En effet, si P est fausse ou Q est fausse, supposer $(P \text{ et } Q)$, c'est-à-dire P et Q vraies, contredit l'une au moins des deux hypothèses initiales. Donc $(P \text{ et } Q)$ est fausse.

Réciproquement, si l'on suppose $(P \text{ et } Q)$ fausse, il est possible par exemple que P soit vraie et Q indécidable et, dans ce cas, ni P , ni Q , ne sont fausses, donc non-($P \text{ et } Q$) n'entraîne pas (non- P ou non- Q).

On peut le comprendre ainsi : pour affirmer que $(P \text{ et } Q)$ est fausse, on a construit une preuve directe amenant à une contradiction à partir de l'hypothèse $(P \text{ et } Q)$. Mais pour prouver P fausse ou Q fausse, il faut être capable du même travail à partir d'une seule des deux hypothèses P ou Q , qui sont plus faibles.

Exemple : soit x un nombre réel. Considérons les propositions P : “ x est un nombre rationnel” et Q : “ x est un nombre irrationnel”.

La proposition $(P \text{ et } Q)$ est fausse en vertu du principe de non-contradiction. Mais prouver P fausse ou Q fausse est plus exigeant : il faut déduire une contradiction de l’hypothèse x rationnel ou de l’hypothèse x irrationnel.

En revanche, la logique intuitionniste dispose toujours de la seconde loi de Morgan :

$$(\text{non-}P \text{ et non-}Q) \implies \text{non-}(P \text{ ou } Q).$$

Retour sur le principe du tiers exclu

Malgré le rejet du principe du tiers exclu, “aucun constructiviste raisonnable ne peut prétendre qu’une instance du principe du tiers exclu est fausse”.

Soit en effet une proposition P quelconque ; la propriété du tiers exclu appliquée à P est $(P \text{ ou non-}P)$. Supposons $\text{non-}(P \text{ ou non-}P)$. D’après la seconde loi de Morgan, ceci équivaut à $(\text{non-}P \text{ et non-}(\text{non-}P))$.

Ces deux affirmations étant incompatibles, on a bien $\text{non-}(\text{non-}(P \text{ ou non-}P))$: l’affirmation que $(P \text{ ou non-}P)$ est fausse est contradictoire.

Conclusion : la propriété du tiers exclu appliquée à une proposition P est vraie ou indécidable.

Chapitre 2

Méthodes algébriques

2.1 Brefs rappels sur les suites

DÉFINITION 2.1 - Une **suite réelle** (*resp.* **complexe**) (u_n) est un procédé permettant d'associer à tout entier naturel n un nombre réel (*resp.* complexe) noté u_n .
Pour un entier naturel arbitraire n , le nombre réel (ou complexe) u_n est alors appelé **terme général** de la suite (u_n) .

Remarque. Cette définition est énoncée à titre provisoire. En effet, elle est imprécise (qu'est qu'un “procédé” ?), même si elle ressemble à s'y méprendre à celle que l'on vous a donnée pour les fonctions. Si le parallèle n'était pas trop dangeureux, on pourrait écrire qu'une suite réelle est une fonction définie sur $\mathbb{N}$. . . Mais cette comparaison est risquée, notamment car elle peut inciter à recourir à la dérivée pour étudier le sens de variation d'une suite, ce qui en général n'a pas de sens. En tout état de cause, la définition rigoureuse de suite sera donnée dans le chapitre portant sur les applications.

Par ailleurs, une suite n'est pas nécessairement définie sur $\mathbb{N}$ tout entier. Par exemple, la suite définie en posant $u_n = 1/n$ n'est définie que sur $\mathbb{N}^*$. Celle définie en posant $u_n = 1/\ln(n)$ n'est définie que sur $\mathbb{N} \setminus \{0, 1\}$. . . Néanmoins, dans la suite de ce chapitre, nous supposerons systématiquement que les suites considérées sont définies sur $\mathbb{N}$, afin d'éviter certaines lourdeurs d'écriture.

Terminologie. Certaines notions et propriétés de ce chapitre (à commencer par celles portant sur les suites arithmétiques et géométriques) seront valables aussi bien pour les suites réelles que pour les suites complexes. Pour éviter les répétitions, nous utiliserons la lettre $\mathbb{K}$ pour désigner indifféremment $\mathbb{R}$ ou $\mathbb{C}$; et nous parlerons de **scalaire** pour désigner un élément de $\mathbb{K}$, càd un nombre réel ou complexe.

Notations. On notera $\mathbb{R}^{\mathbb{N}}$ (*resp.* $\mathbb{C}^{\mathbb{N}}$, *resp.* $\mathbb{K}^{\mathbb{N}}$) l'ensemble des suites réelles (*resp.* complexes, *resp.* à valeurs dans $\mathbb{K}$).

DÉFINITION 2.2 - Une suite (réelle ou complexe) est **arithmétique** s'il existe un scalaire r tel que :

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

Lorsque tel est le cas, le scalaire r est appelé **raison** de la suite (u_n) .

Exemples. La suite des entiers naturels (de terme général $u_n = n$) est arithmétique de raison 1. La suite des entiers naturels impairs (de terme général $u_n = 2n + 1$) est arithmétique de raison 2.

Pour établir qu'une suite (u_n) est arithmétique, on prouve que pour tout entier naturel n , la différence $u_{n+1} - u_n$ est constante.

PROPRIÉTÉ 2.1 - Soit (u_n) une suite arithmétique de raison r .

$$1/ \forall n \in \mathbb{N}, u_n = u_0 + nr$$

$$2/ \forall (n, p) \in \mathbb{N}^2, u_n = u_p + (n - p)r$$

PREUVE. La première propriété provient d'un raisonnement par récurrence sans aucune difficulté ; et la seconde s'en déduit. Juste pour s'en convaincre, faisons-le !

Preuve du 1/. Soit (u_n) une suite arithmétique de raison r . Notons, pour tout entier naturel n , $P(n)$ l'assertion " $u_n = u_0 + nr$ ".

L'initialisation (le fait que $P(0)$ est vraie) est immédiate.

Passons à l'hérédité. Supposons $P(n)$ vraie pour un certain entier naturel n . On a alors :

$$u_{n+1} = u_n + r = u_0 + nr + r = u_0 + (n + 1)r$$

La première égalité ci-dessus provient du fait que (u_n) est arithmétique de raison r , et la seconde de l'hypothèse de récurrence. Au final, on a établi que : $u_{n+1} = u_0 + (n + 1)r$. Ce qui assure que $P(n + 1)$ est vraie, et prouve l'hérédité.

Conclusion. $\forall n \in \mathbb{N}, u_n = u_0 + nr$.

Preuve du 2/. Soit (u_n) une suite arithmétique de raison r , et soient n et p deux entiers naturels. D'après ce qui précède, on a : $u_n = u_0 + nr$ et $u_p = u_0 + pr$. D'où : $u_n - u_p = (n - p)r$, et donc : $u_n = u_p + (n - p)r$.

Conclusion. $\forall (n, p) \in \mathbb{N}^2, u_n = u_p + (n - p)r$. ☕

DÉFINITION 2.3 - Une suite (réelle ou complexe) est **géométrique** s'il existe un scalaire q tel que :

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

Lorsque tel est le cas, le scalaire q est appelé **raison** de la suite (u_n) .

Exemples. La suite : $1, \frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \dots$ de terme général $u_n = \frac{1}{2^n}$ est géométrique de raison $\frac{1}{2}$.

La suite : $1, i, -1, -i, 1, i, -1, -i, \dots$ de terme général $u_n = i^n$ est géométrique de raison i .

La suite de terme général $u_n = e^{ni\theta}$ est géométrique de raison $e^{i\theta}$.

PROPRIÉTÉ 2.2 - Soit (u_n) une suite géométrique de raison q .

$$1/ \forall n \in \mathbb{N}, u_n = u_0 \times q^n$$

$$2/ \forall (n, p) \in \mathbb{N}^2, u_n = u_p \times q^{n-p}$$

PREUVE. Encore une fois, la première propriété provient d'un raisonnement par récurrence sans aucune difficulté ; la seconde s'en déduit. ☕

2.2 Sommes

2.2.1 Généralités

Notation. Lorsque (u_n) est une suite, et $I = \{i_1, \dots, i_N\}$ un ensemble d'indices, on note $\sum_{i \in I} u_i$ pour désigner la somme $u_{i_1} + \dots + u_{i_N}$.

En pratique on a souvent : $I = \llbracket 0; n \rrbracket$ et on note alors $\sum_{i=0}^n u_i$ la somme $u_0 + u_1 + \dots + u_n$.

PROPRIÉTÉ 2.3 - Soient (a_n) et (b_n) deux suites, λ un scalaire et n un entier naturel.

$$1/ \text{ Linéarité : } \sum_{k=0}^n (a_k + b_k) = \sum_{k=0}^n a_k + \sum_{k=0}^n b_k \text{ et } \sum_{k=0}^n (\lambda a_k) = \lambda \sum_{k=0}^n a_k$$

$$2/ \text{ "Relation de Chasles" : } \sum_{k=n}^p a_k = \sum_{k=n}^m a_k + \sum_{k=m+1}^p a_k$$

$$3/ \text{ Sommes télescopiques : } \sum_{k=n}^p (a_{k+1} - a_k) = a_{p+1} - a_n$$

PREUVE. Exceptionnellement, on ne donne ici qu'une idée de la preuve. Chacune des propriétés ci-dessus peut-être prouvée de deux (au moins) méthodes. La première consiste à écrire les différentes sommes avec "des petits points". Par exemple, pour établir le premier énoncé, on peut écrire :

$$\sum_{k=0}^n (a_k + b_k) = (a_0 + b_0) + (a_1 + b_1) + \dots + (a_n + b_n) = (a_0 + \dots + a_n) + (b_0 + \dots + b_n) = \sum_{k=0}^n a_k + \sum_{k=0}^n b_k$$

Au sens strict du terme, cette "preuve" n'en est pas vraiment une. Une autre méthode (rigoureuse celle-ci) consiste à établir cette propriété par récurrence, ce qui constitue d'ailleurs un exercice sans difficulté (ni intérêt, il faut bien l'avouer). ☕

Exemple d'application de la relation de Chasles. Une application assez fréquente de la relation de Chasles est la décomposition :

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

Cette écriture est notamment utile lorsqu'il s'agit de prouver par récurrence une propriété faisant intervenir une somme ; voir à cet effet les formules concernant les sommes remarquables énoncées dans la section suivante.

Exemple d'application des sommes télescopiques 1. Soit n un entier naturel non nul. Calculons

$$S_n = \sum_{k=1}^n \ln \left(1 + \frac{1}{k} \right).$$

$$\text{On a : } S_n = \sum_{k=1}^n \ln \left(1 + \frac{1}{k} \right) = \sum_{k=1}^n \ln \left(\frac{k+1}{k} \right) = \sum_{k=1}^n [\ln(k+1) - \ln(k)] = \ln(n+1) - \ln(1)$$

Conclusion. $\forall n \in \mathbb{N}^*, \sum_{k=1}^n \ln \left(1 + \frac{1}{k} \right) = \ln(n+1).$

Conséquence. On peut déduire de ce qui précède que : $\lim_{n \rightarrow +\infty} \sum_{k=1}^n \ln \left(1 + \frac{1}{k} \right) = +\infty$. Plus tard dans

l'année, nous traduirons cette propriété en disant que la *série de terme général* $\ln \left(1 + \frac{1}{k} \right)$ *est divergente*.

Exemple d'application des sommes télescopiques 2. Soit n un entier naturel supérieur ou égal à 2.

$$\text{Calculons } S_n = \sum_{k=2}^n \ln \left(1 - \frac{1}{k^2} \right).$$

$$\text{On a : } S_n = \sum_{k=2}^n \ln \left(1 - \frac{1}{k^2} \right) = \sum_{k=2}^n \ln \left(\frac{k^2 - 1}{k^2} \right) = \sum_{k=2}^n \ln \left(\frac{(k-1)(k+1)}{k^2} \right)$$

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

Conclusion. $\forall n \in \mathbb{N} \setminus \{0, 1\}, \sum_{k=2}^n \ln \left(1 - \frac{1}{k^2} \right) = \ln \left(\frac{n+1}{2n} \right).$

Conséquence. On peut déduire de ce qui précède que : $\lim_{n \rightarrow +\infty} \sum_{k=1}^n \ln \left(1 - \frac{1}{k^2} \right) = \ln(1/2)$. Plus tard

dans l'année, nous traduirons cette propriété en disant que la *série de terme général* $\ln \left(1 - \frac{1}{k^2} \right)$ *est convergente, et a pour somme* $\ln(1/2)$.

Exemple d'application des sommes télescopiques 3. Soit n un entier naturel non nul. On considère la somme des inverses des entiers naturels compris entre 1 et n :

$$S_n = \sum_{k=1}^n \frac{1}{k} = 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n}$$

On souhaite étudier la limite de S_n lorsque n tend vers $+\infty$. On commence par observer que la fonction inverse est décroissante sur $\mathbb{R}_+^*$, d'où :

$$\forall k \in \mathbb{N}^*, \forall x \in \llbracket k, k+1 \rrbracket, \frac{1}{k+1} \leq \frac{1}{x} \leq \frac{1}{k}$$

En intégrant terme à terme cet encadrement sur $\llbracket k, k+1 \rrbracket$, on obtient :

$$\forall k \in \mathbb{N}^*, \forall x \in \llbracket k, k+1 \rrbracket, \frac{1}{k+1} \leq \underbrace{[\ln(x)]_k^{k+1}}_{=\ln(k+1)-\ln(k)} \leq \frac{1}{k}$$

Par suite : $\forall k \in \mathbb{N}^*, \frac{1}{k} \geq \ln(k+1) - \ln(k)$.

D'où : $\sum_{k=1}^n \frac{1}{k} \geq \sum_{k=1}^n [\ln(k+1) - \ln(k)]$. Donc : $\sum_{k=1}^n \frac{1}{k} \geq \ln(n+1)$.

En résumé, on a établi que : $\forall n \in \mathbb{N}^*, S_n \geq \ln(n+1)$. Or on sait que : $\lim_{n \rightarrow +\infty} \ln(n+1) = +\infty$. On en déduit par comparaison que : $\lim_{n \rightarrow +\infty} S_n = +\infty$.

Conclusion. $\lim_{n \rightarrow +\infty} \sum_{k=1}^n \frac{1}{k} = +\infty$.

Remarque. Plus tard dans l'année, nous traduirons cette propriété en disant que la *série harmonique* (de terme général $1/k$) est *divergente*.

Exemple d'application des sommes télescopiques 4. Soit n un entier naturel non nul. On considère la somme des inverses des carrés entiers naturels compris entre 1 et n :

$$S_n = \sum_{k=1}^n \frac{1}{k^2} = 1 + \frac{1}{4} + \frac{1}{9} + \cdots + \frac{1}{n^2}$$

On souhaite étudier la limite de S_n lorsque n tend vers $+\infty$. On peut commencer par observer que :

$$\forall k \in \mathbb{N}, k \geq 2, \frac{1}{k^2} \leq \frac{1}{k^2 - k} \quad (\spadesuit)$$

Par ailleurs, la fraction rationnelle $\frac{1}{k^2 - k}$ peut être décomposée en éléments simples :

$$\forall k \in \mathbb{N}, k \geq 2, \frac{1}{k^2 - k} = \frac{1}{k(k-1)} = \frac{1}{k-1} - \frac{1}{k} \quad (\clubsuit)$$

On déduit de $(\spadesuit)$ et de $(\clubsuit)$ que :

$$\forall k \in \mathbb{N}, k \geq 2, \frac{1}{k^2} \leq \frac{1}{k-1} - \frac{1}{k}$$

$$\text{Par suite : } \sum_{k=1}^n \frac{1}{k^2} = 1 + \sum_{k=2}^n \frac{1}{k^2} \leq 1 + \underbrace{\sum_{k=2}^n \left[\frac{1}{k-1} - \frac{1}{k} \right]}_{=1-1/n} \quad \text{Donc : } \sum_{k=1}^n \frac{1}{k^2} \leq 2 - \frac{1}{n}.$$

En résumé, on a établi que : $\forall n \in \mathbb{N}, n \geq 2, S_n \leq 2 - \frac{1}{n}$. En particulier : $\forall n \in \mathbb{N}, n \geq 2, S_n \leq 2$. Ce qui signifie que la suite (S_n) est majorée (par 2). Comme il est par ailleurs clair que la suite (S_n) est croissante, on en déduit via le théorème de la limite monotone¹ que (S_n) converge (admet une limite finie).

Conclusion. La suite de terme général $\sum_{k=1}^n \frac{1}{k^2}$ est convergente.

Remarque. Plus tard dans l'année, nous traduirons cette propriété en disant que la *série de terme général* $1/k^2$ est convergente. On peut par ailleurs montrer, mais c'est une autre histoire, que : $\lim_{n \rightarrow +\infty} \sum_{k=1}^n \frac{1}{k^2} = \frac{\pi^2}{6}$.

2.2.2 Sommes remarquables

PROPRIÉTÉ 2.4 - (somme des premiers entiers).

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

PREUVE. Etablissons la propriété par récurrence sur l'entier naturel n . Pour tout entier naturel n , notons $P(n)$ l'assertion : $\sum_{k=0}^n k = \frac{n(n+1)}{2}$.

L'initialisation (le fait que $P(0)$ est vraie) est immédiate.

Passons à l'hérédité. Supposons $P(n)$ vraie pour un certain entier naturel n . On a alors :

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

La première égalité ci-dessus provient de la relation de Chasles pour les sommes, et la seconde de l'hypothèse de récurrence. Au final, on a établi que : $\sum_{k=0}^{n+1} k = \frac{(n+1)(n+2)}{2}$. Ce qui assure que $P(n+1)$ est vraie, et prouve l'hérédité.

Conclusion. $\forall n \in \mathbb{N}, \sum_{k=0}^n k = \frac{n(n+1)}{2}$. ☕

PROPRIÉTÉ 2.5 - (somme des carrés des premiers entiers).

$$\forall n \in \mathbb{N}, \sum_{k=0}^n k^2 = \frac{n(n+1)(2n+1)}{6} \quad (\text{càd : } 0 + 1 + 4 + \dots + n^2 = \frac{n(n+1)(2n+1)}{6}).$$

1. Qui affirme que toute suite réelle croissante et majorée converge.

PREUVE. Notons $P(n)$ l'assertion : $\sum_{k=0}^n k^2 = \frac{n(n+1)(2n+1)}{6}$, et prouvons que $P(n)$ est vraie pour tout entier naturel n par récurrence sur n .

► Initialisation ($n=0$). D'une part : $\sum_{k=0}^0 k^2 = 0$. D'autre part : $\frac{0(0+1)(2 \times 0+1)}{6} = 0 \dots$ Donc $P(0)$ est vraie.

► Hérédité. Supposons $P(n)$ vraie pour un certain entier naturel n . On a alors :

$$\begin{aligned} \sum_{k=0}^{n+1} k^2 &= \sum_{k=0}^n k^2 + (n+1)^2 = \frac{n(n+1)(2n+1)}{6} + (n+1)^2 = \frac{n(n+1)(2n+1) + 6(n+1)^2}{6} \\ &= \frac{(n+1)[n(2n+1) + 6(n+1)]}{6} = \frac{(n+1)(2n^2 + 7n + 6)}{6} \quad (\spadesuit) \end{aligned}$$

Par ailleurs : $\frac{(n+1)(n+2)(2n+3)}{6} = \frac{(n+1)(2n^2 + 7n + 6)}{6} \quad (\clubsuit)$

On déduit de $(\spadesuit)$ et $(\clubsuit)$ que : $\sum_{k=0}^{n+1} k^2 = \frac{(n+1)(n+2)(2n+3)}{6}$. Ce qui signifie que $P(n+1)$ est vraie, et établit l'hérédité de la propriété.

Conclusion. $\forall n \in \mathbb{N}, \sum_{k=0}^n k^2 = \frac{n(n+1)(2n+1)}{6}$. ☕

PROPRIÉTÉ 2.6 - (somme des cubes des premiers entiers).

$$\forall n \in \mathbb{N}, \sum_{k=0}^n k^3 = \frac{n^2(n+1)^2}{4} = \left(\frac{n(n+1)}{2} \right)^2 \quad (\text{càd : } 0 + 1 + 8 + \dots + n^3 = \frac{n^2(n+1)^2}{4}).$$

PREUVE. Voir preuve de la propriété 1.9, page 27. ☕

PROPRIÉTÉ 2.7 - (somme des termes d'une suite arithmétique). Soit (u_n) une suite arithmétique. On a :

$$\sum_{k=0}^n u_k = (n+1) \frac{u_0 + u_n}{2}$$

PREUVE. Soit (u_n) une suite arithmétique. Par définition : $\exists r \in \mathbb{K}, \forall n \in \mathbb{N}, u_{n+1} = u_n + r$ (le scalaire r est la raison de (u_n)).

En outre, on a déjà établi que : $u_n = u_0 + nr$. Il s'ensuit que :

$$\begin{aligned} \sum_{k=0}^n u_k &= \sum_{k=0}^n (u_0 + kr) = \left(\sum_{k=0}^n u_0 \right) + r \left(\sum_{k=0}^n k \right) = (n+1)u_0 + \frac{n(n+1)}{2} \times r = \frac{n+1}{2} (2u_0 + nr) \\ &= \frac{n+1}{2} \left(u_0 + \underbrace{u_0 + nr}_{=u_n} \right) = \frac{n+1}{2} (u_0 + u_n) \end{aligned}$$

Conclusion. Pour (u_n) une suite arithmétique, on a : $\sum_{k=0}^n u_k = (n+1) \frac{u_0 + u_n}{2}$. ☕

PROPRIÉTÉ 2.8 - (somme des termes d'une suite géométrique). Soit (u_n) une suite géométrique de raison $q \neq 1$. On a :

$$\sum_{k=0}^n u_k = u_0 \times \frac{1 - q^{n+1}}{1 - q}$$

PREUVE. Soit (u_n) une suite géométrique de raison $q \neq 1$. On a déjà établi (voir propriété 2.2, page 36) que : $\forall k \in \mathbb{N}, u_k = u_0 \times q^k$. On doit donc établir que : $\sum_{k=0}^n u_0 \times q^k = u_0 \times \frac{1 - q^{n+1}}{1 - q}$.

Il suffit, par linéarité de la somme, de montrer que : $\sum_{k=0}^n q^k = \frac{1 - q^{n+1}}{1 - q}$. Pour ce faire, on peut procéder par récurrence sur n ; pour changer un peu, on utilise ici une méthode alternative.

Soit n un entier naturel quelconque. Posons : $S = \sum_{k=0}^n q^k$. On a :

$$(1 - q)S = (1 - q) \sum_{k=0}^n q^k = \sum_{k=0}^n (q^k - q^{k+1}) = q^0 - q^{n+1} \quad \text{d'où : } (1 - q)S = 1 - q^{n+1}$$

Puisque $q \neq 1$, on en déduit que : $S = \frac{1 - q^{n+1}}{1 - q}$, ce qu'il fallait démontrer.

Conclusion. $\forall n \in \mathbb{N}, \forall u_0 \in \mathbb{C}, \sum_{k=0}^n u_0 \times q^k = u_0 \times \frac{1 - q^{n+1}}{1 - q}$. ☕

2.3 Produits

2.3.1 Généralités

Notation. Lorsque (u_n) est une suite, et $I = \{i_1, \dots, i_N\}$ un ensemble d'indices, on note $\prod_{i \in I} u_i$ pour désigner le produit $u_{i_1} \times \dots \times u_{i_N}$.

En pratique on a souvent : $I = \llbracket 0; n \rrbracket$ et on note alors $\prod_{i=0}^n u_i$ le produit $u_0 \times u_1 \times \dots \times u_n$.

PROPRIÉTÉ 2.9 - Soient (a_n) et (b_n) deux suites, λ un scalaire et n un entier naturel.

$$1/ \prod_{k=1}^n (a_k \times b_k) = \prod_{k=1}^n a_k \times \prod_{k=1}^n b_k \text{ et } \prod_{k=1}^n (\lambda a_k) = \lambda^n \prod_{k=1}^n a_k$$

$$2/ \text{“Relation de Chasles”} : \prod_{k=n}^p a_k = \prod_{k=n}^m a_k \times \prod_{k=m+1}^p a_k$$

$$3/ \text{Produits télescopiques} : \prod_{k=n}^p \frac{a_{k+1}}{a_k} = \frac{a_{p+1}}{a_n}$$

PREUVE. Comme dans le cadre de la propriété 2.3 page 37, on ne donne ici qu'une idée de la preuve. Chacune des propriétés ci-dessus peut-être prouvée de deux (au moins) méthodes. La première consiste à écrire les différents produits avec “des petits points”. Par exemple, pour établir le premier énoncé, on peut écrire :

$$\prod_{k=1}^n (a_k \times b_k) = (a_1 \times b_1) \times \cdots \times (a_n \times b_n) = (a_1 \times \cdots \times a_n) \times (b_1 \times \cdots \times b_n) = \prod_{k=1}^n a_k \times \prod_{k=1}^n b_k$$

Les remarques faites dans la preuve de la propriété 2.3 s'appliquent évidemment encore au cas présent.



Exemple d'application de la relation de Chasles. Une application assez fréquente de la relation de Chasles est la décomposition :

$$\prod_{k=0}^{n+1} u_k = \left[\prod_{k=0}^n u_k \right] \times u_{n+1}$$

Cette écriture est notamment utile lorsqu'il s'agit de prouver par récurrence une propriété faisant intervenir un produit.

Exemple de produit télescopique 1. Soit N un entier naturel ≥ 2 . Notons : $P_N = \prod_{k=2}^N \left(1 - \frac{1}{k}\right)$.

$$\text{On a : } P_N = \prod_{k=2}^N \left(1 - \frac{1}{k}\right) = \prod_{k=2}^N \frac{k-1}{k} = \frac{1}{N}.$$

Exemple de produit télescopique 2. Soit N un entier naturel ≥ 2 . Notons : $P_N = \prod_{k=1}^N \left(1 - \frac{1}{k^2}\right)$.

$$\text{On a : } P_N = \prod_{k=2}^N \left(1 - \frac{1}{k^2}\right) = \prod_{k=2}^N \frac{k^2-1}{k^2} = \prod_{k=2}^N \frac{k-1}{k} \times \prod_{k=2}^N \frac{k+1}{k} = \frac{1}{N} \times \frac{N+1}{2} = \frac{N+1}{2N}.$$

2.3.2 Factorielle d'un entier naturel

DÉFINITION 2.4 - (Factorielle d'un entier naturel). Pour tout entier naturel n , on appelle **factorielle** de n et on note $n!$ l'entier naturel défini en posant :

$$\forall n \in \mathbb{N}^*, n! = \prod_{k=1}^n k \quad \text{et} \quad 0! = 1$$

Exemples : $0! = 1$; $1! = 1$; $2! = 2$; $3! = 6$; $4! = 24$; $5! = 120$; $6! = 720$; $7! = 5040$

PROPRIÉTÉ 2.10 - 1) $\forall n \in \mathbb{N}^*, n! = n \times (n-1)!$

2) $\forall n \in \mathbb{N}, n \geq 2, n! = n \times (n-1) \times (n-2)!$

PREUVE. Plus que des propriétés, il s'agit plutôt de remarques sur les factorielles découlant immédiatement de la définition, et de la relation de Chasles pour les produits.

2.4 Coefficients binomiaux

DÉFINITION 2.5 - (Combinaison). Soient n et p deux entiers naturels. On appelle **combinaison** de p éléments parmi n toute partie à p éléments d'un ensemble à n éléments.

Exemples : dans un ensemble à n éléments, l'ensemble vide est la seule combinaison à 0 élément. Une combinaison de 1 élément est un singleton. Dans une classe de *MPSI*, un binôme est une combinaison de deux éléments de la promo.

Notation : on note $\binom{n}{p}$ le nombre de combinaisons de p éléments parmi n (càd de parties à p éléments d'un ensemble à n éléments).

THÉORÈME 2.1 - $\forall (n, p) \in \mathbb{N}^2, \binom{n}{p} = \begin{cases} \frac{n!}{p!(n-p)!} & \text{si } p \in \llbracket 0, n \rrbracket \\ 0 & \text{sinon} \end{cases}$

PREUVE. Il est déjà clair que pour tout entier naturel n , on a : $\binom{n}{p} = 0$ pour tout entier p tel que $p < 0$ ou $p > n$.²

Montrons par récurrence que pour tout $n \in \mathbb{N}$, l'assertion $P(n) : \forall p \in \llbracket 0, n \rrbracket, \binom{n}{p} = \frac{n!}{p!(n-p)!}$ est vraie.

► Initialisation ($n = 0$). D'une part : $\frac{0!}{0!(0-0)!} = 1$. D'autre part : $\binom{0}{0} = 1$. Donc $P(0)$ est vraie.

► Hérédité. Soit n un entier naturel, et supposons $P(k)$ vraie jusqu'au rang n .³

Soit E un ensemble à $n+1$ éléments. Notons en extension : $E = \{x_1, \dots, x_n, x_{n+1}\}$.

Soit p un entier compris entre 0 et $n+1$. Distinguons deux cas.

► **Premier cas :** $p = 0$. Dans ce cas, on a : $\binom{n+1}{0} = 1$ et $\frac{(n+1)!}{0!(n+1-p)!} = 1$.

Ainsi : $\binom{n+1}{0} = \frac{(n+1)!}{0!(n+1-p)!}$ (♠).

► **Second cas :** $p > 0$. Observons que pour toute partie A de E à p éléments, soit $x_{n+1} \in A$, soit $x_{n+1} \notin A$, ces deux cas s'excluant mutuellement.

On a alors : $\binom{n+1}{p} = B_n + C_n$, en ayant noté B_n (resp. C_n) le nombre de parties à p éléments de E contenant (resp. ne contenant pas) l'élément x_{n+1} .

Or : $B_n = \binom{n}{p-1}$ puisque le choix d'une combinaison de p éléments contenant x_{n+1} est équivalent au choix de $(p-1)$ éléments parmi $x_1, \dots, x_n$.

2. Si l'on est optimiste, cela prouve la moitié du théorème!

3. Comme on l'a vu dans le premier chapitre, on utilise ici une variante du raisonnement par récurrence. Dans l'hérédité, au lieu de supposer que $P(n)$ est vraie pour un rang n particulier, on suppose que $P(0), P(1), \dots, P(n)$ sont vraies. Il s'agit d'un détail de rédaction, dont nous aurons besoin dans la présente preuve.

Par ailleurs : $C_n = \binom{n}{p}$ puisque le choix d'une combinaison de p éléments ne contenant pas x_{n+1} est équivalent au choix de p éléments parmi $x_1, \dots, x_n$.

Ainsi :

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

Donc : $\forall p \in \llbracket 1, n+1 \rrbracket, \binom{n+1}{p} = \frac{(n+1)!}{p!(n-p+1)!} \quad (\clubsuit).$

► Il résulte de ($\spadesuit$) et de ($\clubsuit$) que : $\forall p \in \llbracket 0, n+1 \rrbracket, \binom{n+1}{p} = \frac{(n+1)!}{p!(n-p+1)!}$

Ce qui signifie que $P(n+1)$ est vraie, et achève donc la preuve de l'hérédité.

Conclusion. $\forall n \in \mathbb{N}, \forall p \in \llbracket 0, n \rrbracket, \frac{n!}{p!(n-p)!}$. ☕

PROPRIÉTÉ 2.11 - (Propriétés des coefficients binomiaux).

1/ **Symétrie** : $\forall (n, p) \in \mathbb{N}^2, \binom{n}{p} = \binom{n}{n-p}$

2/ **Relation de Pascal** : $\forall (n, p) \in \mathbb{N}^2, \binom{n}{p-1} + \binom{n}{p} = \binom{n+1}{p}$

3/ **Valeurs remarquables** : pour tout entier naturel n ,

$$\binom{n}{0} = 1 = \binom{n}{n}; \quad \binom{n}{1} = n = \binom{n}{n-1}; \quad \binom{n}{2} = \binom{n}{n-2} = \frac{n(n-1)}{2}$$

PREUVE. Soit n un entier naturel arbitraire.

1/ Soit p un entier compris entre 0 et n . D'après le théorème 2.1, on a :

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

Par ailleurs, si p est un entier tel que $p > n$, alors on a : $\binom{n}{p} = 0$ car il n'existe aucune partie d'un ensemble à n éléments contenant strictement plus de n éléments. Et : $\binom{n}{n-p} = 0$ puisque $n-p < 0$.

On peut ainsi conclure que : $\forall p \in \mathbb{N}, \binom{n}{n-p} = \binom{n}{p}$.

2/ La preuve de la relation de Pascal provient essentiellement du calcul fait au cours de la preuve du théorème 2.1.

3/ ► La seule combinaison à 0 élément d'un ensemble à n éléments est l'ensemble vide, d'où : $\binom{n}{0} = 1$.

Par symétrie, on en déduit que : $\binom{n}{n} = 1$.

► Un ensemble à n éléments contient exactement n singletons (même si $n = 0$). D'où : $\binom{n}{1} = n$.

Par symétrie, on en déduit que : $\binom{n}{n-1} = n$. $\binom{n}{0} = \frac{n!}{0!}$

► Supposons $n \geq 2$. On a : $\binom{n}{2} = \frac{n!}{2!(n-2)!} = \frac{n(n-1)}{2}$.

Par symétrie, on en déduit que : $\binom{n}{n-2} = \frac{n(n-1)}{2}$.

Ces deux égalités restent valables si $n = 0$ ou $n = 1$, puisque “tout le monde est nul”.



Remarque. La relation de Pascal permet notamment de construire le fameux triangle de Pascal, donnant les coefficients binomiaux (pratiquement, pour des valeurs de n pas trop grandes...).

Triangle de Pascal

		Valeurs de p											
		0	1	2	3	4	5	6	7	8	9	10	11
Valeurs de n	0	1	0	0	0	0	0	0	0	0	0	0	0
	1	1	1	0	0	0	0	0	0	0	0	0	0
	2	1	2	1	0	0	0	0	0	0	0	0	0
	3	1	3	3	1	0	0	0	0	0	0	0	0
	4	1	4	6	4	1	0	0	0	0	0	0	0
	5	1	5	10	10	5	1	0	0	0	0	0	0
	6	1	6	15	20	15	6	1	0	0	0	0	0
	7	1	7	21	35	35	21	7	1	0	0	0	0
	8	1	8	28	56	70	56	28	8	1	0	0	0
	9	1	9	36	84	126	126	84	36	9	1	0	0
	10	1	10	45	120	210	252	210	120	45	10	1	0
	11	1	11	55	165	330	462	462	330	165	55	11	1

etc...

Ce procédé, avec le théorème qui suit, fournit de nouvelles identités remarquables...

THÉORÈME 2.2 - (Binôme de Newton). $\forall (a, b) \in \mathbb{C}^2, \forall n \in \mathbb{N}$,

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

PREUVE. Prouvons la formule par récurrence sur n . Considérons a et b deux nombres complexes arbitraires, et posons $P(n) : (a + b)^n = \sum_{k=0}^n \binom{n}{k} a^k b^{n-k}$.

► **Initialisation** : pour $n = 0$, on a $(a + b)^0 = 1$ et $\sum_{k=0}^0 \binom{0}{k} a^k b^{0-k} = \binom{0}{0} a^0 b^0 = 1$, ce qui assure que $P(0)$ est vraie.

► Hérédité : on suppose $P(n)$ vraie pour un certain entier naturel n . Alors :

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

D'où finalement : $(a+b)^{n+1} = \sum_{k=0}^{n+1} \binom{n}{k} a^k b^{n+1-k}$

La relation obtenue signifie exactement que la propriété $P(n+1)$ est vraie, ce qui établit l'hérédité de la propriété.

Conclusion : $\forall (a, b) \in \mathbb{C}^2, \forall n \in \mathbb{N}, (a+b)^n = \sum_{k=0}^n \binom{n}{k} a^k b^{n-k}$. ☕

On obtient comme conséquence du théorème 2.2 et du triangle de Pascal les quatre identités ci-dessous.

Application 1 : $\forall (a, b) \in \mathbb{C}^2, (a+b)^3 = a^3 + 3a^2b + 3ab^2 + b^3$

Application 2 : $\forall (a, b) \in \mathbb{C}^2, (a-b)^3 = a^3 - 3a^2b + 3ab^2 - b^3$

Application 3 : $\forall (a, b) \in \mathbb{C}^2, (a+b)^4 = a^4 + 4a^3b + 6a^2b^2 + 4ab^3 + b^4$

Application 4 : $\forall (a, b) \in \mathbb{C}^2, (a-b)^4 = a^4 - 4a^3b + 6a^2b^2 - 4ab^3 + b^4$

La formule ci-dessous donne l'expression de la somme des coefficients de la ligne de rang n du triangle de Pascal (c'est une application directe du théorème 2.2, avec $a = b = 1$).

Application 5 : $\forall n \in \mathbb{N}, \sum_{k=0}^n \binom{n}{k} = 2^n$

La formule ci-dessous donne l'expression de la somme alternée des coefficients de la ligne de rang n du triangle de Pascal (c'est une application directe du théorème 2.2, avec $a = 1$ et $b = -1$).

Application 6 : $\forall n \in \mathbb{N}, \sum_{k=0}^n (-1)^k \binom{n}{k} = 0$

Application 7 : soit $n \geq 2$. Calculons les sommes $S_1 = \sum_{k=0}^n k \binom{n}{k}$ et $S_2 = \sum_{k=0}^n k^2 \binom{n}{k}$.

► Soit n un entier naturel ≥ 2 . On définit sur $\mathbb{R}$ une fonction f en posant : $\forall x \in \mathbb{R}, f(x) = (1+x)^n$ (♠).

D'après le théorème 2.2, on a également : $\forall x \in \mathbb{R}, f(x) = \sum_{k=0}^n \binom{n}{k} x^k$ (♡).

En calculant $f(1)$ à l'aide des formules (♠) et (♡), on retrouve : $\sum_{k=0}^n \binom{n}{k} = 2^n$.

► La fonction f est dérivable⁴ sur $\mathbb{R}$, et on obtient deux expressions pour sa dérivée en utilisant les formules (♠) et (♡).

D'une part : $\forall x \in \mathbb{R}, f'(x) = n(1+x)^{n-1}$ (◇).

Et d'autre part : $\forall x \in \mathbb{R}, f'(x) = \sum_{k=1}^n k \binom{n}{k} x^{k-1}$ (♣)

En calculant $f'(1)$ à l'aide des formules (◇) et (♣), on obtient : $\sum_{k=1}^n k \binom{n}{k} = n2^{n-1} = \sum_{k=0}^n k \binom{n}{k}$

► A son tour, la fonction f' est dérivable⁵ sur $\mathbb{R}$, et on obtient deux expressions pour sa dérivée (la dérivée seconde de f donc) en utilisant les formules (◇) et (♣).

D'une part : $\forall x \in \mathbb{R}, f''(x) = n(n-1)(1+x)^{n-2}$ (♢)

Et d'autre part : $\forall x \in \mathbb{R}, f''(x) = \sum_{k=2}^n k(k-1) \binom{n}{k} x^{k-2}$ (♠)

En calculant $f''(1)$ à l'aide des formules (♠) et (♢), on obtient : $n(n-1)2^{n-2} = \sum_{k=2}^n k(k-1) \binom{n}{k}$ (♣).

$$\text{Or : } \sum_{k=2}^n k(k-1) \binom{n}{k} = \sum_{k=2}^n k^2 \binom{n}{k} - \sum_{k=2}^n k \binom{n}{k} = \sum_{k=1}^n k^2 \binom{n}{k} - n - \sum_{k=1}^n k \binom{n}{k} + n = \sum_{k=1}^n k^2 \binom{n}{k} - \sum_{k=1}^n k \binom{n}{k}$$

On déduit de cette égalité et de la relation (♣) que :

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

Puis, en exploitant le résultat précédent :

$$\sum_{k=1}^n k^2 \binom{n}{k} = n(n-1)2^{n-2} + n2^{n-1} = n2^{n-2}(n-1+2)$$

$$\text{D'où finalement : } \sum_{k=1}^n k^2 \binom{n}{k} = n(n+1)2^{n-2} = \sum_{k=0}^n k^2 \binom{n}{k}$$

4. Car f est polynomiale.

5. Car f' est polynomiale.

2.5 Epilogue : calcul de $\sum_{k=1}^n k^4$ et de $\sum_{k=1}^n k^5$

Soit n un entier naturel arbitraire. On pose, pour tout entier naturel i :

$$S_i = \sum_{k=0}^n k^i$$

Les valeurs de S_0 , S_1 , S_2 et S_3 sont connues (voir paragraphe sur les sommes remarquables). Le but de ce problème est de calculer S_4 et S_5 , en présentant une méthode générale permettant de déduire S_{n+1} de S_n .

1) **Calcul de S_4 .**

a) Calculer de deux façons différentes la somme $U_4 = \sum_{k=0}^n [(k+1)^5 - k^5]$.

b) Déduire de la question précédente la valeur de S_4 .

2) **Calcul de S_5 .**

a) Calculer de deux façons différentes la somme $U_5 = \sum_{k=0}^n [(k+1)^6 - k^6]$.

b) Déduire de la question précédente la valeur de S_5 .

CORRIGÉ. a) D'une part, on a : $U_4 = \sum_{k=0}^n [(k+1)^5 - k^5] = (n+1)^5 \quad (\spadesuit)$ (somme télescopique).

D'autre part, on a : $U_4 = \sum_{k=0}^n [k^5 + 5k^4 + 10k^3 + 10k^2 + 5k + 1 - k^5] = \sum_{k=0}^n [5k^4 + 10k^3 + 10k^2 + 5k + 1]$

D'où : $U_4 = 5S_4 + 10S_3 + 10S_2 + 5S_1 + S_0 \quad (\clubsuit)$

b) On déduit de $(\spadesuit)$ et $(\clubsuit)$ que :

$$5S_4 + 10S_3 + 10S_2 + 5S_1 + S_0 = (n+1)^5$$

$$\iff 5S_4 = (n+1)^5 - \frac{5n^2(n+1)^2}{2} - \frac{5n(n+1)(2n+1)}{3} - \frac{5n(n+1)}{2} - (n+1)$$

$$\iff 5S_4 = (n+1) \left[(n+1)^4 - \frac{5n^2(n+1)}{2} - \frac{5n(2n+1)}{3} - \frac{5n}{2} - 1 \right]$$

$$\iff 5S_4 = \frac{(n+1)}{6} [6(n+1)^4 - 15n^2(n+1) - 10n(2n+1) - 15n - 6]$$

$$\iff 5S_4 = \frac{(n+1)}{6} [6n^4 + 24n^3 + 36n^2 + 24n + 6 - 15n^3 - 15n^2 - 20n^2 - 10n - 15n - 6]$$

$$\iff 5S_4 = \frac{(n+1)}{6} [6n^4 + 9n^3 + n^2 - n]$$

$$\iff 5S_4 = \frac{n(n+1)}{6} [6n^3 + 9n^2 + n - 1]$$

$$\iff S_4 = \frac{n(n+1)(6n^3 + 9n^2 + n - 1)}{30}$$

3) a) D'une part, on a : $U_5 = \sum_{k=0}^n [(k+1)^6 - k^6] = (n+1)^6 \quad (\spadesuit)$ (re-somme télescopique).

D'autre part :

$$U_5 = \sum_{k=0}^n [k^6 + 6k^5 + 15k^4 + 20k^3 + 15k^2 + 6k + 1 - k^6] = \sum_{k=0}^n [6k^5 + 15k^4 + 20k^3 + 15k^2 + 6k + 1]$$

D'où : $U_5 = 6S_5 + 15S_4 + 20S_3 + 15S_2 + 6S_1 + S_0 \quad (\clubsuit)$

b) On déduit de $(\spadesuit)$ et $(\clubsuit)$ que :

$$6S_5 + 15S_4 + 20S_3 + 15S_2 + 6S_1 + S_0 = (n+1)^6$$

$$\iff 6S_5 = (n+1)^6 - \frac{n(n+1)(6n^3 + 9n^2 + n - 1)}{2} - 5n^2(n+1)^2 - \frac{5n(n+1)(2n+1)}{2} - 3n(n+1) - (n+1)$$

$$\iff 6S_5 = (n+1) \left[(n+1)^5 - \frac{n(6n^3 + 9n^2 + n - 1)}{2} - 5n^2(n+1) - \frac{5n(2n+1)}{2} - 3n - 1 \right]$$

$$\iff 6S_5 = \frac{(n+1)}{2} [2(n+1)^5 - n(6n^3 + 9n^2 + n - 1) - 10n^2(n+1) - 5n(2n+1) - 6n - 2]$$

$$\iff 6S_5 =$$

$$\frac{(n+1)}{2} [2n^5 + 10n^4 + 20n^3 + 20n^2 + 10n + 2 - 6n^4 - 9n^3 - n^2 + n - 10n^3 - 10n^2 - 10n^2 - 5n - 6n - 2]$$

$$\iff 6S_5 = \frac{(n+1)}{2} [2n^5 + 4n^4 + n^3 - n^2]$$

$$\iff S_5 = \frac{(n+1)n^2}{12} [2n^3 + 4n^2 + n - 1]$$

$$\iff S_5 = \frac{(n+1)n^2}{12} [(n+1)(2n^2 + 2n - 1)]$$

$$\iff S_5 = \frac{(n+1)^2 n^2 (2n^2 + 2n - 1)}{12}$$

Chapitre 3

Trigonométrie

3.1 Congruences modulo un nombre réel

DÉFINITION 3.1 - Soit $(x, y, A) \in \mathbb{R}^2 \times \mathbb{R}_+^*$.

Le réel y est **congru au réel x modulo A** s'il existe un entier relatif k tel que : $y = x + kA$.

Notation : lorsque tel est le cas, on note $y \equiv x [A]$ (ou simplement $y = x [A]$).

Exemples : $3 \equiv 1 \left[\frac{1}{2} \right]$; $\frac{\pi}{2} \equiv -\frac{\pi}{2} [\pi]$; $\pi \equiv 0 \left[\frac{\pi}{3} \right]$

PROPRIÉTÉ 3.1 - Soit $A \in \mathbb{R}_+^*$.

1/ **Réflexivité** : $\forall x \in \mathbb{R}, x \equiv x [A]$

2/ **Symétrie** : $\forall (x, y) \in \mathbb{R}^2, (y \equiv x [A]) \iff (x \equiv y [A])$

3/ **Transitivité** : $\forall (x, y, z) \in \mathbb{R}^3, (y \equiv x [A] \wedge z \equiv y [A]) \implies (z \equiv x [A])$

PREUVE. Soit $A \in \mathbb{R}_+^*$.

1/ Pour tout réel x , on a : $x = x + 0 \times A$, d'où : $x \equiv x [A]$.

2/ Soient x et y deux réels tels que : $y \equiv x [A]$. Alors : $\exists k \in \mathbb{Z}, y = x + kA$. D'où : $x = y + (-k)A$, soit encore : $x = y + k'A$ avec $k' = -k \in \mathbb{Z}$. D'où : $x \equiv y [A]$.

On a ainsi établi l'implication : $(y \equiv x [A]) \implies (x \equiv y [A])$. L'implication réciproque (et donc l'équivalence) provient de ce que les réels x et y jouent des rôles symétriques dans cet énoncé.

3/ Soient x, y et z trois réels tels que : $y \equiv x [A]$ et $z \equiv y [A]$. Alors : $\exists k \in \mathbb{Z}, y = x + kA$ et $\exists k' \in \mathbb{Z}, z = y + k'A$.

D'où : $z = x + (k + k')A$. Puisque $(k + k') \in \mathbb{Z}$, cette égalité signifie que : $z \equiv x [A]$.

Par suite : $\forall (x, y, z) \in \mathbb{R}^3, (y \equiv x [A] \wedge z \equiv y [A]) \implies (z \equiv x [A])$. ☕

Remarque : les trois énoncés de la propriété ci-dessus signifient que la relation “être congru à” (ou relation de congruence) est ce que l'on appelle une **relation d'équivalence**. Nous aurons de nombreuses occasions de revoir cette notion en cours d'année.

Par ailleurs, un mot sur le choix de $A \in \mathbb{R}_+^*$. On peut définir plus généralement la notion de congruence modulo un réel A quelconque.

Mais lorsque $A = 0$, il revient au même de dire que y et x sont congrus modulo 0, ou qu'ils sont égaux.

Enfin, pour un réel A quelconque, il revient au même de dire que y et x sont congrus modulo A , ou qu'ils sont congrus modulo $(-A)$.

On peut donc, sans nuire à la généralité, toujours supposer que $A > 0$.

3.2 Fonctions trigonométriques

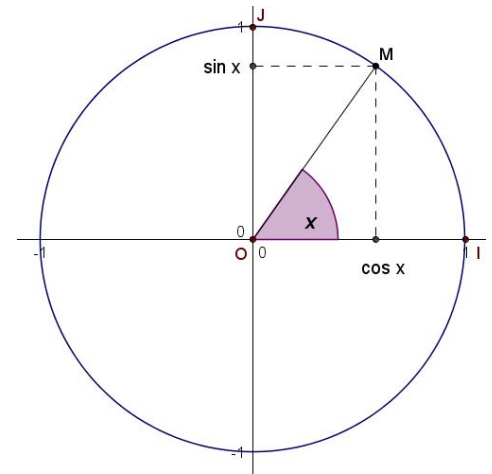
3.2.1 Les fonctions cosinus et sinus

DÉFINITION 3.2 - On se place dans le plan muni d'un repère orthonormal $(O; I, J)$.

Soit x un nombre réel.

1/ On appelle **point image** du réel x le point M du cercle trigonométrique tel qu'une mesure en radians de l'angle orienté $(\vec{OI}, \vec{OM})$ soit égale à x .

2/ On appelle **cosinus du réel x** (*resp.* **sinus du réel x**) et on note $\cos x$ (*resp.* $\sin x$) l'abscisse (*resp.* l'ordonnée) du point M .



Remarque : il résulte de la définition que les fonctions cosinus et sinus sont **définies sur $\mathbb{R}$** tout entier.

PROPRIÉTÉ 3.2 - (“Bornitude”). Les fonctions cosinus et sinus sont bornées sur $\mathbb{R}$. Plus précisément, ce sont deux fonctions définies sur $\mathbb{R}$ et à valeurs dans $[-1; 1]$:

$$\forall x \in \mathbb{R} : \begin{cases} -1 \leq \cos x \leq 1 \\ -1 \leq \sin x \leq 1 \end{cases} \quad \text{ou} \quad \forall x \in \mathbb{R} : \begin{cases} |\cos x| \leq 1 \\ |\sin x| \leq 1 \end{cases}$$

PREUVE. Tout point appartenant au cercle trigonométrique a une abscisse et une ordonnée comprise entre -1 et 1 . ☕

PROPRIÉTÉ 3.3 - (Périodicité). Les fonctions cosinus et sinus sont 2π -périodiques. En d'autres termes :

$$\forall x \in \mathbb{R} : \begin{cases} \cos(x + 2\pi) = \cos x \\ \sin(x + 2\pi) = \sin x \end{cases}$$

PREUVE. Pour tout réel x , les réels x et $(x + 2\pi)$ ont la même image sur le cercle trigonométrique. ☕

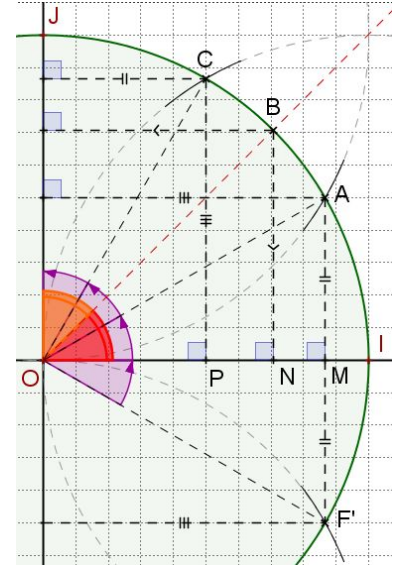
PROPRIÉTÉ 3.4 - (Parité). La fonction cosinus est paire, et la fonction sinus est impaire. En d'autres termes :

$$\forall x \in \mathbb{R}, \cos(-x) = \cos x \quad \text{et} \quad \forall x \in \mathbb{R}, \sin(-x) = -\sin x$$

PREUVE. Soit $x \in \mathbb{R}$. Les points images de x et $-x$ sont symétriques par rapport à l'axe (Ox) . ☕

PROPRIÉTÉ 3.5 - (Valeurs remarquables). La figure ci-contre représente une partie du cercle trigonométrique. Sur celui-ci, on a placé les points I, A, B, C et J , images respectives des réels $0, \frac{\pi}{6}, \frac{\pi}{4}, \frac{\pi}{3}$ et $\frac{\pi}{2}$. On appelle **valeurs remarquables** les images de ces réels par les fonctions trigonométriques.

x	0	$\frac{\pi}{6}$	$\frac{\pi}{4}$	$\frac{\pi}{3}$	$\frac{\pi}{2}$
$\cos(x)$	1	$\frac{\sqrt{3}}{2}$	$\frac{\sqrt{2}}{2}$	$\frac{1}{2}$	0
$\sin(x)$	0	$\frac{1}{2}$	$\frac{\sqrt{2}}{2}$	$\frac{\sqrt{3}}{2}$	1



PREUVE. C'est un très amusant petit exercice de géométrie élémentaire ! ☕

Remarque : on observe en particulier que $\cos x = 0$ si et seulement si $x = \pm \frac{\pi}{2} [2\pi]$, soit si et seulement si $x = \frac{\pi}{2} [\pi]$ (ou encore : $x = \frac{\pi}{2} + k\pi$ avec $k \in \mathbb{Z}$).

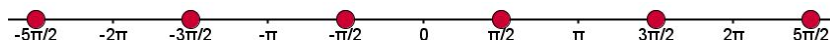
3.2.2 La fonction tangente

DÉFINITION 3.3 - Soit x un nombre réel tel que $x \neq \frac{\pi}{2} [\pi]$ (càd que x est un réel pour lequel $\cos x \neq 0$).

On appelle **tangente du réel x** et on note $\tan x$ le réel :

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

Remarque. Il découle de la définition que la fonction $\tan$ est définie sur : $\mathcal{D}_{\tan} = \left\{ x \in \mathbb{R}, x \neq \frac{\pi}{2} [\pi] \right\}$. Notons que cet ensemble est symétrique par rapport à zéro ; sur la figure ci-dessous, l'ensemble $\mathcal{D}_{\tan}$ est en effet constitué de la droite réelle "privée des points rouges".



PROPRIÉTÉ 3.6 - (Périodicité). La fonction tangente est π -périodique, c'est-à-dire :

$$\forall x \in \mathcal{D}_{\tan}, \quad \tan(x + \pi) = \tan(x)$$

PREUVE. Soit $x \in \mathcal{D}_{\tan}$. Alors $x + \pi \in \mathcal{D}_{\tan}$, et : $\tan(x + \pi) = \frac{\sin(x + \pi)}{\cos(x + \pi)} = \frac{-\sin(x)}{-\cos(x)} = \tan(x)$. ☕

PROPRIÉTÉ 3.7 - (Parité). La fonction tangente est impaire, c'est-à-dire :

$$\forall x \in \mathcal{D}_{\tan}, \quad \tan(-x) = -\tan(x)$$

PREUVE. Soit $x \in \mathcal{D}_{\tan}$. Alors $-x \in \mathcal{D}_{\tan}$, et : $\tan(-x) = \frac{\sin(-x)}{\cos(-x)} = \frac{-\sin(x)}{\cos(x)} = -\tan(x)$. ☕

PROPRIÉTÉ 3.8 - (Valeurs remarquables). Le tableau ci-dessous contient les valeurs remarquables de la fonction tangente.

x	0	$\frac{\pi}{6}$	$\frac{\pi}{4}$	$\frac{\pi}{3}$	$\frac{\pi}{2}$
$\tan(x)$	0	$\frac{\sqrt{3}}{3}$	1	$\sqrt{3}$	non-définie

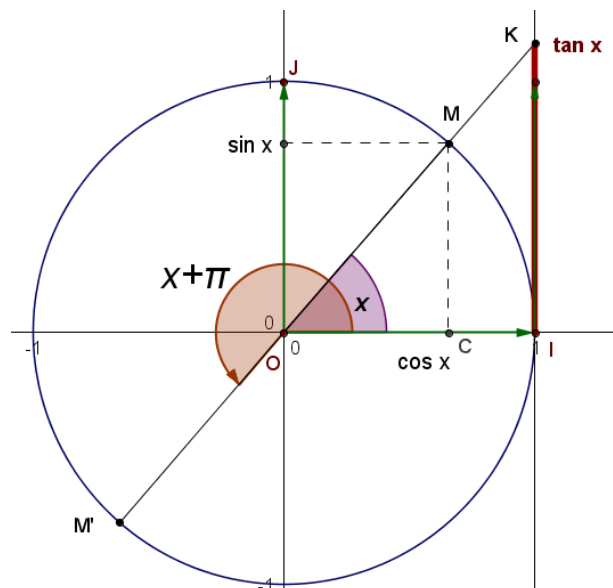
PREUVE. Conséquence immédiate de la définition de tangente, et des valeurs remarquables des fonctions cos et sin. ☕

Interprétation géométrique de la tangente.

La figure ci-contre illustre le fait que pour un réel x fixé, le réel $\tan x$ peut être lu sur l'axe orienté $(I, \overrightarrow{OJ})$. C'est finalement une conséquence du théorème de Thalès, appliquée à la configuration des points O, C, I, K et M .

La figure fait clairement apparaître la π -périodicité de la fonction tangente. En effet, "deux points du cercle diamétralement opposés vont donner la même tangente" ; soit, beaucoup plus justement, deux réels égaux modulo π ont la même tangente.

En particulier, si x est un réel (pour lequel la tangente est définie), alors les réels $x + \pi$, x et $x - \pi$ ont la même image par la fonction tangente (ce qui peut être utile notamment en Physique).



3.3 Formulaire de trigonométrie

Note : les formules suivantes sont valables pour tout nombre réel x .

Relation fondamentale : $\cos^2(x) + \sin^2(x) = 1$

Remarque : c'est une conséquence du théorème de Pythagore.

$$\cos(-x) = \cos(x)$$

$$\sin(-x) = -\sin(x)$$

Interprétation géométrique : les points images de x et $-x$ sont symétriques par rapport à l'axe des abscisses (argument déjà évoqué lors de l'étude de la parité des fonctions $\cos$ et $\sin$).

$$\cos(\pi - x) = -\cos(x)$$

$$\sin(\pi - x) = \sin(x)$$

Interprétation géométrique : les points images de x et $\pi - x$ sont symétriques par rapport à l'axe des ordonnées.

$$\cos(\pi + x) = -\cos(x)$$

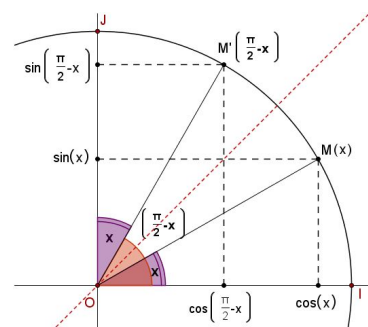
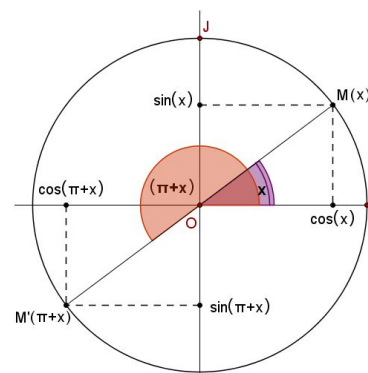
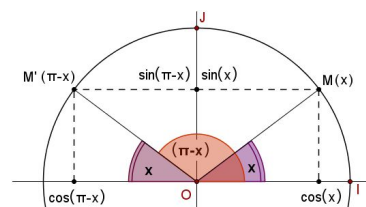
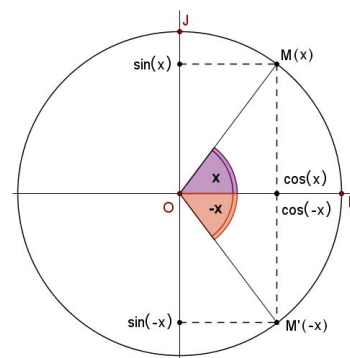
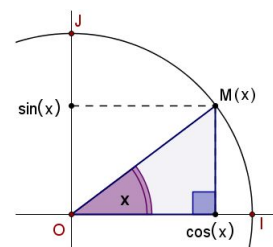
$$\sin(\pi + x) = -\sin(x)$$

Interprétation géométrique : les points images de x et $\pi + x$ sont symétriques par rapport à l'origine O du repère.

$$\cos\left(\frac{\pi}{2} - x\right) = \sin(x)$$

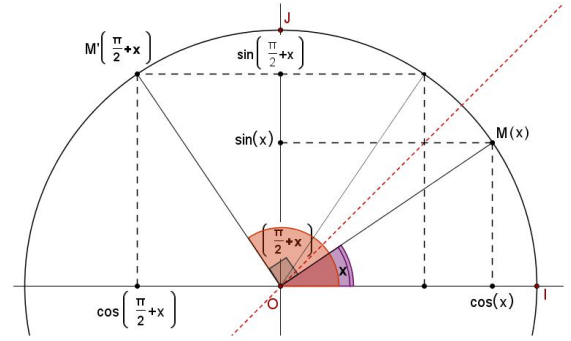
$$\sin\left(\frac{\pi}{2} - x\right) = \cos(x)$$

Interprétation géométrique : les points images de x et $\frac{\pi}{2} - x$ sont symétriques par rapport la première bissectrice des axes (la droite d'équation cartésienne $y = x$, en pointillés rouges sur la figure).



$$\begin{aligned}\cos\left(\frac{\pi}{2} + x\right) &= -\sin(x) \\ \sin\left(\frac{\pi}{2} + x\right) &= \cos(x)\end{aligned}$$

Interprétation géométrique : on passe du point image de x au point image de $\frac{\pi}{2} + x$ en faisant d'abord une symétrie par rapport la première bissectrice des axes, puis une symétrie par rapport à l'axe des ordonnées.



$$\text{Formules d'addition : } \begin{cases} \cos(a+b) = \cos(a)\cos(b) - \sin(a)\sin(b) \\ \sin(a+b) = \sin(a)\cos(b) + \sin(b)\cos(a) \\ \tan(a+b) = \frac{\tan a + \tan b}{1 - \tan a \tan b} \end{cases}$$

En utilisant la parité de la fonction cosinus, et l'imparité des fonctions sinus et tangente, on obtient déjà :

$$\text{Formules "de soustraction" : } \begin{cases} \cos(a-b) = \cos(a)\cos(b) + \sin(a)\sin(b) \\ \sin(a-b) = \sin(a)\cos(b) - \sin(b)\cos(a) \\ \tan(a-b) = \frac{\tan a - \tan b}{1 + \tan a \tan b} \end{cases}$$

On déduit de ces formules les suivantes :

$$\text{Formules de duplication : } \begin{cases} \cos(2a) = \cos^2(a) - \sin^2(a) = 2\cos^2(a) - 1 = 1 - 2\sin^2(a) \\ \sin(2a) = 2\sin(a)\cos(a) \\ \tan(2a) = \frac{2\tan a}{1 - \tan^2 a} \end{cases}$$

On peut également obtenir les formules ci-dessous à partir des formules d'addition :

$$\text{Formules de linéarisation : } \begin{cases} \cos(a)\cos(b) = \frac{\cos(a+b) + \cos(a-b)}{2} & \text{d'où : } \cos^2(a) = \frac{1 + \cos(2a)}{2} \\ \sin(a)\sin(b) = \frac{\cos(a-b) - \cos(a+b)}{2} & \text{d'où : } \sin^2(a) = \frac{1 - \cos(2a)}{2} \\ \sin(a)\cos(b) = \frac{\sin(a+b) + \sin(a-b)}{2} & \text{d'où : } \sin(a)\cos(a) = \frac{\sin(2a)}{2} \end{cases}$$

Dans le “sens contraire”, on peut transformer des sommes en produit :

$$\left\{ \begin{array}{l} \cos p + \cos q = 2 \cos \left(\frac{p+q}{2} \right) \cos \left(\frac{p-q}{2} \right) \\ \cos p - \cos q = -2 \sin \left(\frac{p+q}{2} \right) \sin \left(\frac{p-q}{2} \right) \\ \sin p + \sin q = 2 \sin \left(\frac{p+q}{2} \right) \cos \left(\frac{p-q}{2} \right) \\ \sin p - \sin q = 2 \sin \left(\frac{p-q}{2} \right) \cos \left(\frac{p+q}{2} \right) \end{array} \right.$$

Et pour finir, des relations souvent utiles en calcul intégral :

$$\text{En posant } t = \tan \left(\frac{x}{2} \right), \text{ on a : } \left\{ \begin{array}{l} \cos x = \frac{1-t^2}{1+t^2} \\ \sin x = \frac{2t}{1+t^2} \\ \tan x = \frac{2t}{1-t^2} \end{array} \right.$$

3.4 Quelques équations trigonométriques

3.4.1 L'équation $\cos(x) = \cos(y)$

PROPRIÉTÉ 3.9 - Soient x et y deux réels. On a :

$$\cos(x) = \cos(y) \quad \text{SSI} \quad y = \pm x [2\pi]$$

PREUVE. Deux points du cercle trigonométrique ont même abscisse si et seulement si ils sont confondus ou symétriques par rapport à l'axe des abscisses. ☕

Exemple. Résoudre dans $\mathbb{R}$ l'équation $\cos(2x) + \cos(6x) = 0$.

$$(E) : \quad \cos(2x) + \cos(6x) = 0 \iff \cos(6x) = -\cos(2x) \iff \cos(6x) = \cos(2x + \pi)$$

$$\iff \left\{ \begin{array}{l} 6x = 2x + \pi [2\pi] \\ 6x = -\pi - 2x [2\pi] \end{array} \right. \iff \left\{ \begin{array}{l} x = \frac{\pi}{4} [\pi/2] \\ x = -\frac{\pi}{8} [\pi/4] \end{array} \right.$$

Conclusion. L'ensemble des solutions de (E) est : $\left\{ \frac{\pi}{4} + \frac{k\pi}{2} / k \in \mathbb{Z} \right\} \cup \left\{ -\frac{\pi}{8} + \frac{k\pi}{4} / k \in \mathbb{Z} \right\}$.

3.4.2 L'équation $\sin(x) = \sin(y)$

PROPRIÉTÉ 3.10 - Soient x et y deux réels. On a :

$$\sin(x) = \sin(y) \quad \text{SSI} \quad y = x [2\pi] \text{ ou } y = \pi - x [2\pi]$$

PREUVE. Deux points du cercle trigonométrique ont même ordonnée si et seulement si ils sont confondus ou symétriques par rapport à l'axe des ordonnées. ☕

Exemple. Résoudre dans $\mathbb{R}$ l'équation $\sin(2x) + \sin(7x) = 0$.

$$(E) : \quad \sin(2x) + \sin(7x) = 0 \iff \sin(7x) = -\sin(2x) \iff \sin(7x) = \sin(2x + \pi)$$

$$\iff \begin{cases} 7x = 2x + \pi [2\pi] \\ 7x = \pi - 2x - \pi [2\pi] \end{cases} \iff \begin{cases} x = \frac{\pi}{5} [2\pi/5] \\ x = 0 [2\pi/9] \end{cases}$$

Conclusion. L'ensemble des solutions de (E) est : $\left\{ \frac{\pi}{5} + \frac{2k\pi}{5} / k \in \mathbb{Z} \right\} \cup \left\{ \frac{2k\pi}{9} / k \in \mathbb{Z} \right\}$.

3.4.3 L'équation $\cos(x) = \sin(y)$

Pour une équation du type $\cos(x) = \sin(y)$, on se ramène à l'un des deux cas précédents, par exemple en observant que $\sin(y) = \cos\left(\frac{\pi}{2} - y\right)$, ou que $\cos(x) = \sin\left(x + \frac{\pi}{2}\right)$, bref en utilisant l'une des formules permettant de transformer un "cos" en un "sin", ou un "sin" en un "cos".

3.4.4 L'équation $A \cos(x) + B \sin(x) = C$

Soient A , B et C trois nombres réels. L'objet de ce paragraphe est de résoudre dans $\mathbb{R}$ l'équation :

$$(E) : \quad A \cos(x) + B \sin(x) = C$$

On distingue deux cas.

► **Premier cas :** $(A, B) = (0, 0)$. Dans ce cas, l'ensemble des solutions de l'équation (E) est $\mathbb{R}$ si $C = 0$, et $\emptyset$ sinon.

► **Second cas :** $(A, B) \neq (0, 0)$. Dans ce cas, on observe que :

$$(E) : \quad A \cos(x) + B \sin(x) = C \iff \underbrace{\frac{A}{\sqrt{A^2 + B^2}}}_{A'} \cos(x) + \underbrace{\frac{B}{\sqrt{A^2 + B^2}}}_{B'} \sin(x) = \underbrace{\frac{C}{\sqrt{A^2 + B^2}}}_{C'}$$

Par suite, pour tout réel x :

$$x \text{ est solution de } (E) \iff A' \cos(x) + B' \sin(x) = C'$$

Il résulte de la définition des réels A' et B' que : $A'^2 + B'^2 = 1$.

Par conséquent : $\exists! \varphi \in [0; 2\pi[$, $A' = \cos(\varphi) \wedge B' = \sin(\varphi)$.

Donc : x est solution de $(E) \iff \cos(\varphi)\cos(x) + \sin(\varphi)\sin(x) = C'$

Soit : x est solution de $(E) \iff \cos(x - \varphi) = C'$

On discute alors suivant les valeurs de C' .

- Si $|C'| > 1$, càd si $|C| > \sqrt{A^2 + B^2}$, alors l'équation (E) ne possède aucune solution.
- Si $|C'| \leq 1$, càd si $|C| \leq \sqrt{A^2 + B^2}$, alors $C = \cos(\theta)$ pour un certain réel θ , et l'équation (E) devient alors équivalente à l'équation $\cos(x - \varphi) = \cos(\theta)$, que l'on sait résoudre d'après le paragraphe consacré aux équations de ce type.

3.5 Epilogue : valeurs exactes de $\cos(\pi/8)$, $\cos(\pi/16)$...

D'après la formule de duplication : $\cos\left(\frac{\pi}{4}\right) = \cos\left(2 \times \frac{\pi}{8}\right) = 2\cos^2\left(\frac{\pi}{8}\right) - 1$.

On déduit de cette égalité que $\cos\left(\frac{\pi}{8}\right)$ est solution de l'équation :

$$\frac{\sqrt{2}}{2} = 2x^2 - 1$$

$$\text{Or : } \frac{\sqrt{2}}{2} = 2x^2 - 1 \iff 2x^2 = \frac{\sqrt{2}}{2} + 1 \iff x^2 = \frac{2 + \sqrt{2}}{2} \iff x = \pm \frac{\sqrt{2 + \sqrt{2}}}{2}$$

Comme $\frac{\pi}{8}$ appartient à l'intervalle $[0, \pi/2]$, on a : $\cos\left(\frac{\pi}{8}\right) \geq 0$.

On peut donc conclure que : $\cos\left(\frac{\pi}{8}\right) = \frac{\sqrt{2 + \sqrt{2}}}{2}$.

En réitérant le processus, on peut obtenir :

$$\cos\left(\frac{\pi}{16}\right) = \frac{\sqrt{2 + \sqrt{2 + \sqrt{2}}}}{2};$$

$$\cos\left(\frac{\pi}{32}\right) = \frac{\sqrt{2 + \sqrt{2 + \sqrt{2 + \sqrt{2}}}}}{2} \dots$$

Chapitre 4

Nombres complexes

4.1 Généralités

DÉFINITION 4.1 - On appelle **nombre complexe** une écriture formelle $a + ib$, où a et b désignent deux réels, et i vérifie $i^2 = -1$. On note $\mathbb{C}$ l'ensemble des nombres complexes.

Pour un nombre complexe $z = a + ib$, le réel a (*resp.* b) est appelé **partie réelle** (*resp.* **partie imaginaire**) de z ; elle est notée $\operatorname{Re}(z)$ (*resp.* $\operatorname{Im}(z)$).

Remarque. Il résulte de la définition que $\mathbb{R} \subset \mathbb{C}$, puisque tout nombre réel x peut être vu comme le nombre complexe $x + i \times 0$.¹

DÉFINITION 4.2 - Un nombre complexe z est dit **imaginaire pur** lorsque $\operatorname{Re}(z) = 0$.

Notation. On note $i\mathbb{R}$ l'ensemble des imaginaires purs.

Opérations algébriques sur les nombres complexes. On définit la somme et le produit de deux nombres complexes, et l'inverse d'un complexe non-nul, de telle sorte que ces opérations prolongent celles connues dans $\mathbb{R}$. Explicitement, si $z = a + ib$ et $z' = a' + ib'$ sont deux complexes quelconques (avec a, b, a' et b' réels), on pose :

$$z + z' = (a + a') + i(b + b') \quad \text{et} \quad z \times z' = (aa' - bb') + i(ab' + a'b)$$

Si en outre $z \neq 0$, on pose : $\frac{1}{z} = \frac{1}{a + ib} = \frac{a}{a^2 + b^2} - i \frac{b}{a^2 + b^2}$

DÉFINITION 4.3 - Deux nombres complexes sont **égaux** si et seulement si ils ont même partie réelle et même partie imaginaire.

Cette définition est la clef de ce que l'on appelle méthode d'identification.

1. Remarque, à relire à l'issue du chapitre sur les applications : il serait plus juste d'écrire que l'application $x \mapsto x + i \times 0$ est une application injective de $\mathbb{R}$ dans $\mathbb{C}$.

PROPRIÉTÉ 4.1 - Soient z et z' deux nombres complexes. On a :

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

Plus généralement : $\forall n \in \mathbb{N}^*, \forall (z_1, \dots, z_n) \in \mathbb{C}^n, \operatorname{Re} \left(\sum_{k=1}^n z_k \right) = \sum_{k=1}^n \operatorname{Re}(z_k)$

$$\text{et } \operatorname{Im} \left(\sum_{k=1}^n z_k \right) = \sum_{k=1}^n \operatorname{Im}(z_k)$$

PREUVE. Soient $z = a + ib$ et $z' = a' + ib'$ deux complexes (avec a, b, a' et b' réels).

On a : $z + z' = (a + a') + i(b + b')$, d'où : $\operatorname{Re}(z + z') = \operatorname{Re}(z) + \operatorname{Re}(z')$ et $\operatorname{Im}(z + z') = \operatorname{Im}(z) + \operatorname{Im}(z')$.

La seconde partie de l'énoncé provient de ce résultat et d'une récurrence sans difficulté. ☕

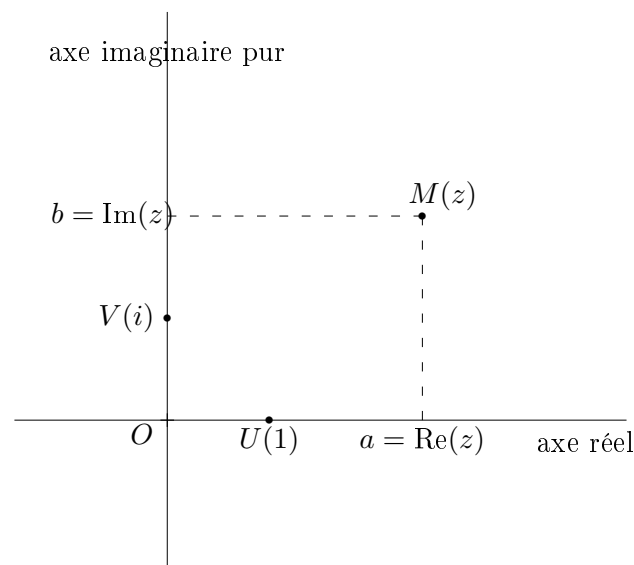
Représentation géométrique des nombres complexes.

DÉFINITION 4.4 - Soit $(O; \overrightarrow{OU}, \overrightarrow{OV})$ un repère orthonormal du plan.

- A tout nombre complexe $z = x + iy$ ($x, y \in \mathbb{R}$) on associe le point M du plan de coordonnées (x, y) . On dit que M est le **point image** de z et $\overrightarrow{OM}$ est le **vecteur image** de z .
- Tout point $M(x; y)$ du plan est le point image d'un unique complexe : $z = x + iy$. On dit que z est l'**affixe** du point M et du vecteur $\overrightarrow{OM}$.
- Ce mode de représentation des nombres complexes est appelé **plan complexe**.

En résumé, disons que pour représenter les nombres réels, il suffit de prendre une droite graduée. Pour les nombres complexes, une droite ne suffit plus, et on les représente à l'aide d'un plan, de la manière suivante.

A tout nombre complexe $z = a + ib$ on associe un point M ayant pour abscisse a et pour ordonnée b :



Les nombres réels sont les affixes des points de l'axe des abscisses ; d'où le nom d'**axe réel**. Les nombres imaginaires purs sont les affixes des points de l'axe des ordonnées ; d'où le nom d'**axe imaginaire pur**.

4.2 Conjugaison

DÉFINITION 4.5 - Soit z un nombre complexe. On appelle **conjugué** de z et on note $\bar{z}$ le complexe défini en posant :

$$\bar{z} = \operatorname{Re}(z) - i \operatorname{Im}(z)$$

PROPRIÉTÉ 4.2 - (Propriétés algébriques des conjugués).

$$1/ \forall (z, z') \in \mathbb{C}^2, \overline{z + z'} = \bar{z} + \bar{z}'$$

$$2/ \forall z \in \mathbb{C}, \overline{(\bar{z})} = z$$

$$3/ \forall z \in \mathbb{C}, z + \bar{z} = 2 \operatorname{Re}(z)$$

$$4/ \forall z \in \mathbb{C}, z - \bar{z} = 2i \operatorname{Im}(z)$$

$$5/ \forall (z, z') \in \mathbb{C}^2, \overline{zz'} = \bar{z} \bar{z}'$$

$$6/ \forall n \in \mathbb{N}, \forall z \in \mathbb{C}, \overline{z^n} = \bar{z}^n$$

$$7/ \forall z \in \mathbb{C}^*, \overline{\left(\frac{1}{z}\right)} = \frac{1}{\bar{z}}$$

PREUVE. Soient z et z' deux nombres complexes. Posons : $z = a + ib$ et $z' = a' + ib'$ avec a, a', b et b' réels.

$$1/ \text{ On a : } \overline{z + z'} = \overline{(a + ib + a' + ib')} = \overline{(a + a' + i(b + b'))} = a + a' - i(b + b') = (a - ib) + (a' - ib').$$

$$\text{On en déduit que : } \boxed{\forall (z, z') \in \mathbb{C}^2, \overline{z + z'} = \bar{z} + \bar{z}'}$$

$$2/ \text{ On a : } \overline{(\bar{z})} = \overline{(a + ib)} = \overline{(a - ib)} = a + ib. \text{ Ainsi : } \boxed{\forall z \in \mathbb{C}, \overline{(\bar{z})} = z}$$

$$3/ \text{ On a : } z + \bar{z} = (a + ib) + (a - ib) = 2a = 2 \operatorname{Re}(z). \text{ D'où : } \boxed{\forall z \in \mathbb{C}, z + \bar{z} = 2 \operatorname{Re}(z)}$$

$$4/ \text{ On a : } z - \bar{z} = (a + ib) - (a - ib) = 2ib = 2i \operatorname{Im}(z). \text{ D'où : } \boxed{\forall z \in \mathbb{C}, z - \bar{z} = 2i \operatorname{Im}(z)}$$

$$5/ \text{ D'une part : } \overline{zz'} = \overline{(a + ib)(a' + ib')} = \overline{aa' - bb' + i(a'b + ab')} = aa' - bb' - i(a'b + ab')$$

$$\text{Et d'autre part : } \bar{z} \bar{z}' = (a - ib)(a' - ib') = aa' - bb' - iab' - iab = aa' - bb' - i(a'b + ab').$$

$$\text{Ainsi : } \boxed{\forall (z, z') \in \mathbb{C}^2, \overline{zz'} = \bar{z} \bar{z}'}$$

6/ C'est une conséquence immédiate du point précédent, et d'une récurrence sur n .

7/ Supposons à présent que $z \neq 0$. Alors : $z \times \frac{1}{z} = 1$ d'où : $\overline{z \times \frac{1}{z}} = \overline{1} = 1$. On en déduit par la propriété 5

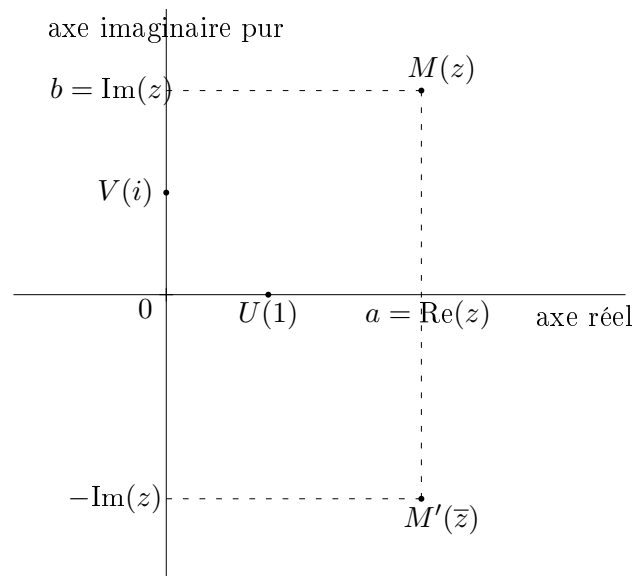
$$\text{que : } \bar{z} \times \overline{\left(\frac{1}{z}\right)} = 1 \text{ donc : } \overline{\left(\frac{1}{z}\right)} = \frac{1}{\bar{z}}. \text{ Par suite : } \boxed{\forall z \in \mathbb{C}^*, \overline{\left(\frac{1}{z}\right)} = \frac{1}{\bar{z}}} \text{ ☕}$$

Remarque. A l'instar de ce que nous avons obtenu dans la propriété 4.1 page 62, les énoncés 1 et 5 des propriétés ci-dessus peuvent être généralisés, pour obtenir à l'aide d'une récurrence (aisée) :

$$\forall n \in \mathbb{N}^*, \forall (z_1, \dots, z_n) \in \mathbb{C}^n, \overline{\sum_{k=1}^n z_k} = \sum_{k=1}^n \overline{z_k} \quad \text{et} \quad \forall n \in \mathbb{N}^*, \forall (z_1, \dots, z_n) \in \mathbb{C}^n, \overline{\prod_{k=1}^n z_k} = \prod_{k=1}^n \overline{z_k}$$

Interprétation géométrique de la conjugaison.

Soient z un complexe et M le point d'affixe z dans le plan complexe. Le point M' d'affixe $\bar{z}$ est le symétrique de M par rapport à l'axe réel.



On obtient comme corollaire de cette observation l'énoncé ci-dessous (qui peut également être prouvé algébriquement).

PROPRIÉTÉ 4.3 - Soit z un nombre complexe. On a :

$$z \in \mathbb{R} \iff \bar{z} = z \quad \text{et} \quad z \in i\mathbb{R} \iff \bar{z} = -z$$

PREUVE. Soit $z \in \mathbb{C}$.

On a : $\bar{z} = z \iff \text{Re}(z) + i \text{Im}(z) = \text{Re}(z) - i \text{Im}(z) \iff \text{Im}(z) = -\text{Im}(z) \iff \text{Im}(z) = 0 \iff z \in \mathbb{R}$

Et : $\bar{z} = -z \iff \text{Re}(z) + i \text{Im}(z) = -\text{Re}(z) + i \text{Im}(z) \iff \text{Re}(z) = -\text{Re}(z) \iff \text{Re}(z) = 0 \iff z \in i\mathbb{R}$

Ce qui prouve les deux équivalences de la propriété. ☕

COROLLAIRE 4.1 - Soit P un polynôme à coefficients réels, et soit $\alpha \in \mathbb{C}$.
Si α est racine de P , alors $\bar{\alpha}$ est racine de P .

PREUVE. Soit P un polynôme à coefficients réels. Il existe donc $(n+1)$ réels $a_0, a_1, \dots, a_n$ tels que :

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

Soit par ailleurs α un complexe, et supposons que α est racine de P .

On a alors : $P(\alpha) = \sum_{k=0}^n a_k \alpha^k = 0$. D'où : $\overline{\sum_{k=0}^n a_k \alpha^k} = 0$ ($\spadesuit$).

Or : $\overline{\sum_{k=0}^n a_k \alpha^k} = \sum_{k=0}^n \overline{a_k \alpha^k} = \sum_{k=0}^n \overline{a_k} \times \overline{\alpha^k} = \sum_{k=0}^n a_k \times \bar{\alpha}^k = P(\bar{\alpha})$ ($\clubsuit$)

On déduit de ($\spadesuit$) et ($\clubsuit$) que : $P(\bar{\alpha}) = 0$. Par suite, $\bar{\alpha}$ est racine de P .

Conclusion. $\forall P \in \mathbb{R}[X], \forall \alpha \in \mathbb{C}, [P(\alpha) = 0] \implies [P(\bar{\alpha}) = 0]$.² ☕

4.3 Module

DÉFINITION 4.6 - Soit $z \in \mathbb{C}$. On appelle **module** de z le réel positif :

$$|z| = \sqrt{z\bar{z}}$$

Remarque. En notant comme d'habitude $z = a + ib$, on a : $|z| = \sqrt{a^2 + b^2}$.

Commençons par deux propriétés dont la démonstration ne mange pas de pain, mais qu'il peut être utile de garder présentes à l'esprit lorsque l'on travaille avec les modules.

PROPRIÉTÉ 4.4 - Soit $z \in \mathbb{C}$.

$$1/ |z| \in \mathbb{R}_+$$

$$2/ |z| = 0 \iff z = 0$$

PREUVE. Soit $z \in \mathbb{C}$. Posons $z = a + ib$, avec a et b réels.

$$1/ \text{ On a : } |z| = \sqrt{a^2 + b^2}, \text{ d'où } |z| \in \mathbb{R}_+.$$

$$2/ \text{ On a : } |z| = 0 \iff \sqrt{a^2 + b^2} = 0 \iff a^2 + b^2 = 0 \iff [a = 0 \wedge b = 0] \iff z = 0. \text{ ☕}$$

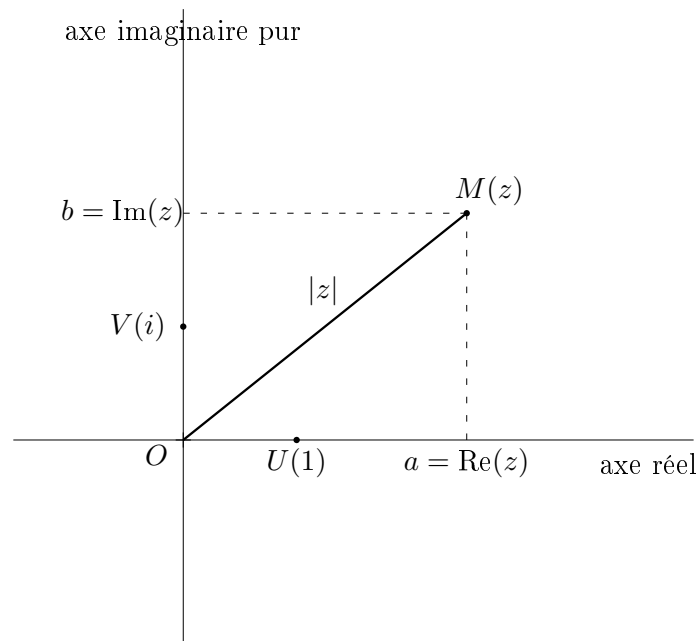
Interprétation géométrique du module.

Soit $z \in \mathbb{C}$, alors $|z|$ est la longueur $OM(z)$ (où $M(z)$ désigne le point du plan complexe d'affixe z).

Et lorsque A et B sont deux points d'affixes respectives z_A et z_B , la longueur AB est le module de $z_B - z_A$.

Il s'ensuit que le lieu des points M dont l'affixe z est telle que $|z - a| = R$ (*resp.* $|z - a| \leq R$, *resp.* $|z - a| < R$) est le cercle (*resp.* le disque fermé, *resp.* le disque ouvert) de centre le point d'affixe a et de rayon R .

2. La notation $\mathbb{R}[X]$ désigne l'ensemble des polynômes à coefficients réels.



PROPRIÉTÉ 4.5 - (Propriétés algébriques des modules).

1/ $\forall z \in \mathbb{C}, |\bar{z}| = |z|$

2/ $\forall (z, z') \in \mathbb{C}^2, |zz'| = |z| \times |z'|$

3/ $\forall z \in \mathbb{C}, \forall n \in \mathbb{N}, |z^n| = |z|^n$

4/ $\forall z \in \mathbb{C}^*, \left| \frac{1}{z} \right| = \frac{1}{|z|}$

5/ $\forall (z, z') \in \mathbb{C} \times \mathbb{C}^*, \left| \frac{z}{z'} \right| = \frac{|z|}{|z'|}$

PREUVE. Soient z et z' deux nombres complexes.

1/ On a : $|\bar{z}| = \sqrt{\bar{z} \times \overline{\bar{z}}} = \sqrt{z\bar{z}} = |z|$. D'où : $\boxed{\forall z \in \mathbb{C}, |\bar{z}| = |z|}$.

2/ Soient z et z' deux nombres complexes. On a :

$$|zz'| = \sqrt{zz' \overline{(zz')}} = \sqrt{zz' \bar{z} \bar{z}'} = \sqrt{(z\bar{z}) (z'\bar{z}')} = \sqrt{z\bar{z}} \sqrt{z'\bar{z}'} = |z| \times |z'|$$

Conclusion : $\forall (z, z') \in \mathbb{C}^2, |zz'| = |z| \times |z'|$.

3/ Soit z un nombre complexe. La propriété provient encore une fois d'une récurrence sans difficulté sur n , et juste pour s'en convaincre, faisons-la !

Notons, pour tout entier naturel n , $P(n)$ l'assertion : " $|z^n| = |z|^n$ ".

Raisonnons par récurrence sur l'entier naturel n .

► **Initialisation** ($n = 0$). D'une part : $|z^0| = |1| = 1$. D'autre part : $|z|^0 = 1$. La propriété $P(0)$ est donc vraie.

► **Hérédité**. Supposons $P(n)$ vraie pour un certain entier naturel n . On a :

$$|z^{n+1}| = |z^n \times z| = |z^n| \times |z| = |z|^n \times |z| = |z|^{n+1}$$

En résumé : $|z^{n+1}| = |z|^{n+1}$. Ce qui assure que $P(n+1)$ est vraie, et établit donc l'hérédité de la propriété.

► On peut alors affirmer que $P(n)$ est vraie pour tout entier naturel n , et puisque le nombre complexe z est arbitraire dans le raisonnement précédent, on a : $\boxed{\forall z \in \mathbb{C}, \forall n \in \mathbb{N}, |z^n| = |z|^n}$.

4/ Soit z un nombre complexe non nul. On a :

$$\left| \frac{1}{z} \right| = \sqrt{\frac{1}{z} \times \overline{\left(\frac{1}{z} \right)}} = \sqrt{\frac{1}{z} \times \frac{1}{\bar{z}}} = \sqrt{\frac{1}{z\bar{z}}} = \frac{1}{\sqrt{z\bar{z}}} = \frac{1}{|z|}$$

Conclusion : $\forall z \in \mathbb{C}^*, |1/z| = 1/|z|$.

5/ Soient z et z' deux complexes, avec $z' \neq 0$. On a :

$$\left| \frac{z}{z'} \right| = \left| z \times \frac{1}{z'} \right| = |z| \times \left| \frac{1}{z'} \right| = |z| \times \frac{1}{|z'|} = \frac{|z|}{|z'|}$$

Conclusion : $\forall (z, z') \in \mathbb{C} \times \mathbb{C}^*, \left| \frac{z}{z'} \right| = \frac{|z|}{|z'|}$



Remarque. Encore une fois, le point 2 de la propriété ci-dessus peut être généralisé (par récurrence) :

$$\forall n \in \mathbb{N}^*, \forall (z_1, \dots, z_n) \in \mathbb{C}^n, \left| \prod_{k=1}^n z_k \right| = \prod_{k=1}^n |z_k|$$

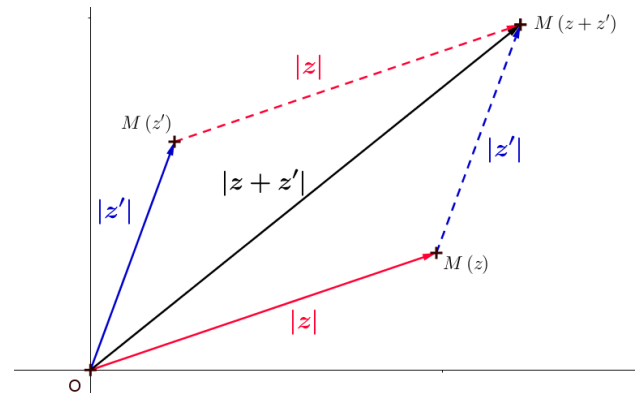
En revanche, le module de la somme de deux complexes n'est en général pas égal à la somme de leurs modules. Plus précisément :

THÉORÈME 4.1 - (Inégalité triangulaire).

$$\forall (z, z') \in \mathbb{C}^2, |z + z'| \leq |z| + |z'|$$

PREUVE. Soient z et z' deux nombres complexes. On a :

$$\begin{aligned} |z + z'|^2 &= (z + z') \overline{(z + z')} \\ \iff |z + z'|^2 &= z\bar{z} + z\bar{z}' + \bar{z}z' + z'\bar{z}' \\ \iff |z + z'|^2 &= |z|^2 + z\bar{z}' + \bar{z}z' + |z'|^2 \\ \iff |z + z'|^2 &= |z|^2 + z\bar{z}' + \overline{z\bar{z}'} + |z'|^2 \end{aligned}$$



$$\iff |z + z'|^2 = |z|^2 + 2\operatorname{Re}(z\overline{z'}) + |z'|^2$$

Or : $\operatorname{Re}(z\overline{z'}) \leq |z\overline{z'}|$.³ En observant judicieusement que $|z\overline{z'}| = |z| |z'|$, on en déduit que :

$$|z + z'|^2 \leq |z|^2 + 2|z| |z'| + |z'|^2$$

$$\iff |z + z'|^2 \leq (|z| + |z'|)^2$$

Cette inégalité ne faisant intervenir que des réels positifs, on peut conclure : $|z + z'| \leq |z| + |z'|$. ☕

Remarque. De nouveau, l'inégalité triangulaire peut être généralisée à la somme de n nombres complexes par récurrence :

$$\forall n \in \mathbb{N}^*, \forall (z_1, \dots, z_n) \in \mathbb{C}^n, \left| \sum_{k=1}^n z_k \right| \leq \sum_{k=1}^n |z_k|$$

Par ailleurs, on peut déterminer une condition nécessaire et suffisante pour avoir égalité des deux termes dans l'énoncé du théorème 4.1 :

PROPRIÉTÉ 4.6 - (inégalité triangulaire, cas d'égalité).

$$\forall (z, z') \in \mathbb{C}^2, |z + z'| \leq |z| + |z'| \quad \text{et} \quad [|z + z'| = |z| + |z'|] \iff [\exists \lambda \in \mathbb{R}_+, z' = \lambda z]$$

PREUVE. Soient z et z' deux nombres complexes. L'inégalité a été prouvée précédemment.

Montrons que : $[|z + z'| = |z| + |z'|] \iff [\exists \lambda \in \mathbb{R}_+, z' = \lambda z]$

Supposons que $|z + z'| = |z| + |z'|$. Deux cas se présentent alors : $z' = 0$ ou $z' \neq 0$.

Dans le premier cas ($z' = 0$), on a : $z' = 0 \times z$. Donc : $\exists \lambda \in \mathbb{R}_+, z' = \lambda z$.

Dans le second cas ($z' \neq 0$), on a : $\left| \frac{z}{z'} + 1 \right| = \left| \frac{z}{z'} \right| + 1$ (♠)

Posons : $Z = \frac{z}{z'}$. On cherche les complexes Z tels que : $|Z + 1| = |Z| + 1$. Posons : $Z = a + i b$ (avec a et b réels). On a :

$$|Z + 1| = |Z| + 1 \iff |Z + 1|^2 = (|Z| + 1)^2 \iff (a + 1)^2 + b^2 = a^2 + b^2 + 1 + 2\sqrt{a^2 + b^2}$$

$$\iff a = \sqrt{a^2 + b^2} \iff a = \sqrt{a^2} \iff [a \in \mathbb{R}_+ \wedge b = 0]$$

En résumé : $[|Z + 1| = |Z| + 1] \iff [\exists \lambda \in \mathbb{R}_+, Z = \lambda]$ (♣).

On déduit de (♠) et (♣) que : $[|z + z'| = |z| + |z'|] \iff [\exists \lambda \in \mathbb{R}_+, z' = \lambda z]$.

Conclusion : $\forall (z, z') \in \mathbb{C}^2, [|z + z'| = |z| + |z'|] \implies [\exists \lambda \in \mathbb{R}_+, z' = \lambda z]$.

L'implication réciproque est immédiate. ☕

3. Cette inégalité provient du fait que : $\forall z \in \mathbb{C}, |\operatorname{Re}(z)| \leq |z|$ et $|\operatorname{Im}(z)| \leq |z|$.

4.4 Ensemble des nombres complexes de module 1

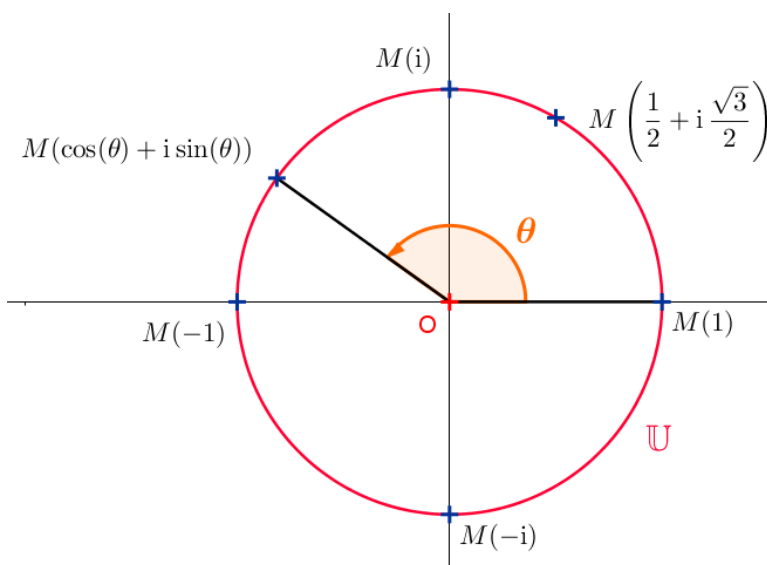
4.4.1 Généralités

DÉFINITION 4.7 - On appelle **cercle unité** et on note $\mathbb{U}$ l'ensemble des nombres complexes de module 1, càd :

$$\mathbb{U} = \{z \in \mathbb{C} / |z| = 1\}$$

Interprétation géométrique. Comme son nom le suggère, l'ensemble $\mathbb{U}$ est l'ensemble des affixes des points du plan complexe appartenant au cercle de centre l'origine du plan et de rayon 1.

Observons que $1, i, -1, -i, \frac{1}{2} + i\frac{\sqrt{3}}{2}$ et plus généralement $\cos(\theta) + i\sin(\theta)$ sont des éléments de $\mathbb{U}$.



L'énoncé ci-dessous donne une CNS pour qu'un nombre complexe appartienne au cercle unité ; explicitement un complexe appartient à $\mathbb{U}$ si et seulement si son inverse est égal à son conjugué.

LEMME 4.1 - $\forall z \in \mathbb{C}, [z \in \mathbb{U}] \iff \left[\bar{z} = \frac{1}{z} \right]$

PREUVE. Soit $z \in \mathbb{C}$. On a : $z \in \mathbb{U} \iff |z| = 1 \iff |z|^2 = 1 \iff z\bar{z} = 1 \iff \bar{z} = \frac{1}{z}$ ☕

PROPRIÉTÉ 4.7 - (Propriétés de $\mathbb{U}$).

1/ $1 \in \mathbb{U}$

2/ $\mathbb{U}$ est stable par produit : $\forall (z, z') \in \mathbb{U}^2, zz' \in \mathbb{U}$

3/ $\mathbb{U}$ est stable par conjugaison : $\forall z \in \mathbb{U}, \bar{z} \in \mathbb{U}$

4/ $\mathbb{U}$ est stable par passage à l'inverse : $\forall z \in \mathbb{U}, \frac{1}{z} \in \mathbb{U}$

PREUVE. 1/ On a : $|1| = 1$, d'où $1 \in \mathbb{U}$.

2/ Soient z et z' dans $\mathbb{U}$. Alors : $|zz'| = |z| \times |z'| = 1 \times 1 = 1$. D'où : $(z, z') \in \mathbb{U}^2 \implies zz' \in \mathbb{U}$.

3/ Soit $z \in \mathbb{U}$. Alors : $|\bar{z}| = |z| = 1$. D'où : $\bar{z} \in \mathbb{U}$. Ainsi : $z \in \mathbb{U} \implies \bar{z} \in \mathbb{U}$.

4/ Ce dernier point provient du précédent, et du lemme assurant que l'inverse et le conjugué d'un élément de $\mathbb{U}$ sont égaux. ☕

Remarque : plus tard dans l'année, nous donnerons un nom aux propriétés énoncées ci-dessus. Le point 2/, assurant que le produit de deux éléments de $\mathbb{U}$ en est encore un, signifiera que la multiplication est une **loi de composition interne** (ou *lci*), qui est **associative** (car la multiplication dans $\mathbb{C}$ l'est). Cette *lci* possède en outre un **élément neutre**, puisque d'après le point 1/, le complexe 1 (qui est l'élément neutre pour la multiplication) appartient à $\mathbb{U}$. Enfin, tout élément de $\mathbb{U}$ est **inversible** et possède un inverse dans $\mathbb{U}$ d'après le point 4/.

Dans ce contexte, nous dirons que l'ensemble $\mathbb{U}$, muni de la multiplication, est un **groupe**. Une bonne partie d'un chapitre futur sera consacrée à l'étude des groupes.

4.4.2 Lien avec la trigonométrie

Préambule. En Terminale, la fonction exponentielle a été introduite comme l'unique fonction f dérivable sur $\mathbb{R}$ telle que $f' = f$ et $f(0) = 1$. Dans ce cas, on dit aussi que la fonction exponentielle est l'unique solution de ce que l'on appelle un **problème de Cauchy**, c'est à dire la donnée d'une équation différentielle, et d'une condition initiale.

Dans le cas présent, la fonction $\exp$ est l'unique solution du problème de Cauchy : $\begin{cases} y' = y \\ y(0) = 1 \end{cases}$.

Plus généralement, pour tout réel λ , il est facile de vérifier que le problème de Cauchy : $\begin{cases} y' = \lambda y \\ y(0) = 1 \end{cases}$ admet comme unique solution la fonction $x \in \mathbb{R} \mapsto e^{\lambda x}$. Fin du premier acte.

Considérons à présent la fonction φ définie sur $\mathbb{R}$ en posant : $\forall x \in \mathbb{R}, \varphi(x) = \cos(x) + i \sin(x)$. On admet que φ est dérivable sur $\mathbb{R}$,⁴ et que : $\forall x \in \mathbb{R}, \varphi'(x) = -\sin(x) + i \cos(x)$.

On peut alors observer que : $\forall x \in \mathbb{R}, \varphi'(x) = i \varphi(x)$. En outre, $\varphi(0) = 1$.

Ces deux constatations signifient, avec la terminologie fraîchement introduite, que la fonction φ est solution du problème de Cauchy : $\begin{cases} y' = i y \\ y(0) = 1 \end{cases}$.

Par analogie avec ce qui précède, on convient de poser : $\forall x \in \mathbb{R}, e^{ix} = \varphi(x) = \cos(x) + i \sin(x)$. C'est une motivation pour introduire la notation ci-dessous.

NOTATION 4.1 - On pose : $\forall x \in \mathbb{R}, e^{ix} = \cos(x) + i \sin(x)$.

LEMME 4.2 - $\forall x \in \mathbb{R}, |e^{ix}| = 1$, soit encore : $\forall x \in \mathbb{R}, e^{ix} \in \mathbb{U}$.

PREUVE. C'est une conséquence immédiate de la relation fondamentale de la trigonométrie. ☕

4. De nouveau, nous le justifierons plus tard dans l'année, dans un futur chapitre consacré à la dérivabilité et à ses applications.

PROPRIÉTÉ 4.8 -

$$1/ \forall (x, y) \in \mathbb{R}^2, e^{i(x+y)} = e^{ix} e^{iy}$$

$$2/ \forall x \in \mathbb{R}, \overline{e^{ix}} = e^{-ix} = (e^{ix})^{-1}$$

$$3/ \forall n \in \mathbb{N}, \forall x \in \mathbb{R}, (e^{ix})^n = e^{inx}$$

PREUVE. 1/ Soient x et y deux réels. On a :

$$\begin{aligned} e^{ix} e^{iy} &= (\cos(x) + i \sin(x)) (\cos(y) + i \sin(y)) \\ &= \underbrace{(\cos(x) \cos(y) - \sin(x) \sin(y))}_{=\cos(x+y)} + i \underbrace{(\cos(x) \sin(y) + \sin(x) \cos(y))}_{=\sin(x+y)} \end{aligned}$$

Ainsi : $e^{ix} e^{iy} = \cos(x+y) + i \sin(x+y)$, soit : $e^{ix} e^{iy} = e^{i(x+y)}$, cqfd.

2/ Soit x un réel. L'égalité $\overline{e^{ix}} = (e^{ix})^{-1}$ provient des lemmes 4.1 et 4.2.

Par ailleurs : $e^{-ix} = \cos(-x) + i \sin(-x) = \cos(x) - i \sin(x) = \overline{e^{ix}}$.

Finalement : $e^{-ix} = \overline{e^{ix}} = (e^{ix})^{-1}$.

3/ Ce dernier point provient du premier par récurrence sur n . ☕

Remarque. Comme toujours, l'énoncé peut être généralisé, pour obtenir par récurrence :

$$\forall n \in \mathbb{N}^*, \forall (x_1, \dots, x_n) \in \mathbb{R}^n, \exp \left(\sum_{k=1}^n x_k \right) = \prod_{k=1}^n e^{x_k}$$

4.4.3 Formules d'Euler et formule de Moivre

PROPRIÉTÉ 4.9 - (Formules d'Euler).

$$\forall x \in \mathbb{R}, \cos(x) = \frac{e^{ix} + e^{-ix}}{2} \quad \text{et} \quad \forall x \in \mathbb{R}, \sin(x) = \frac{e^{ix} - e^{-ix}}{2i}$$

PREUVE. Soit x un réel. On a : $e^{-ix} = \overline{e^{ix}}$.

D'où : $e^{ix} + e^{-ix} = 2\operatorname{Re}(e^{ix}) = 2\cos(x)$; et $e^{ix} - e^{-ix} = 2i\operatorname{Im}(e^{ix}) = 2i\sin(x)$.

Les formules s'en déduisent immédiatement. ☕

PROPRIÉTÉ 4.10 - (Formule de Moivre).

$$\forall n \in \mathbb{N}, \forall x \in \mathbb{R}, (\cos(x) + i \sin(x))^n = \cos(nx) + i \sin(nx)$$

PREUVE. Soient x un réel et n un entier naturel.

D'après le point 3 de la propriété 4.8, on a : $(e^{ix})^n = e^{inx}$.

Par définition⁵, cette égalité s'écrit encore : $(\cos(x) + i \sin(x))^n = \cos(nx) + i \sin(nx)$. ☕

5. De la notation e^{ix} .

Application 1 — Linéarisation

Très grossièrement, linéariser une expression, c'est l'opération qui consiste à faire disparaître les puissances de celle-ci. Par exemple, nous avons déjà vu (dans le chapitre de trigonométrie) deux formules dites de linéarisation :

$$\forall a \in \mathbb{R}, \cos^2(a) = \frac{1 + \cos(2a)}{2} \quad \text{et} \quad \forall a \in \mathbb{R}, \sin^2(a) = \frac{1 - \cos(2a)}{2}$$

Dans les exemples ci-dessous, on donne d'autres exemples de linéarisation, utilisant les formules d'Euler (propriété 4.9) et la formule du binôme de Newton (théorème 2.2, page 46).

► **Exemple 1. Linéarisation de $\cos^3(\theta)$.** Soit θ un réel. On a :

$$\cos(\theta) = \frac{1}{2}(e^{i\theta} + e^{-i\theta}) \quad \text{d'où :} \quad \cos^3(\theta) = \frac{1}{8}(e^{i\theta} + e^{-i\theta})^3$$

Il s'ensuit⁶ que :

$$\cos^3(\theta) = \frac{1}{8}(e^{3i\theta} + 3e^{i\theta} + 3e^{-i\theta} + e^{-3i\theta}) = \frac{1}{8} \left(\underbrace{e^{3i\theta} + e^{-3i\theta}}_{=2\cos(3\theta)} + \underbrace{3e^{i\theta} + 3e^{-i\theta}}_{=6\cos(\theta)} \right)$$

D'où finalement : $\forall \theta \in \mathbb{R}, \cos^3(\theta) = \frac{\cos(3\theta) + 3\cos(\theta)}{4}.$

► **Exemple 2. Linéarisation de $\sin^3(\theta)$.** Soit θ un réel. On a :

$$\sin(\theta) = \frac{1}{2i}(e^{i\theta} - e^{-i\theta}) \quad \text{d'où :} \quad \sin^3(\theta) = -\frac{1}{8i}(e^{i\theta} - e^{-i\theta})^3$$

Il s'ensuit⁷ que :

$$\sin^3(\theta) = -\frac{1}{8i}(e^{3i\theta} - 3e^{i\theta} + 3e^{-i\theta} - e^{-3i\theta}) = -\frac{1}{8i} \left(\underbrace{e^{3i\theta} - e^{-3i\theta}}_{=2i\sin(3\theta)} - \underbrace{(3e^{i\theta} - 3e^{-i\theta})}_{=-6i\sin(\theta)} \right)$$

D'où finalement : $\forall \theta \in \mathbb{R}, \sin^3(\theta) = \frac{3\sin(\theta) - \sin(3\theta)}{4}.$

► **Exemple 3. Linéarisation de $\cos^4(\theta)$.** Soit θ un réel. On a :

$$\cos(\theta) = \frac{1}{2}(e^{i\theta} + e^{-i\theta}) \quad \text{d'où :} \quad \cos^4(\theta) = \frac{1}{16}(e^{i\theta} + e^{-i\theta})^4$$

Il s'ensuit⁸ que :

$$\cos^4(\theta) = \frac{1}{16}(e^{4i\theta} + 4e^{2i\theta} + 6 + 4e^{-2i\theta} + e^{-4i\theta}) = \frac{1}{16} \left(\underbrace{e^{4i\theta} + e^{-4i\theta}}_{=2\cos(4\theta)} + \underbrace{4e^{2i\theta} + 4e^{-2i\theta}}_{=8\cos(2\theta)} + 6 \right)$$

D'où finalement : $\forall \theta \in \mathbb{R}, \cos^4(\theta) = \frac{\cos(4\theta) + 4\cos(2\theta) + 3}{8}.$

6. D'après la formule du binôme de Newton : $(a+b)^3 = a^3 + 3a^2b + 3ab^2 + b^3.$

7. D'après la formule du binôme de Newton : $(a-b)^3 = a^3 - 3a^2b + 3ab^2 - b^3.$

8. D'après la formule du binôme de Newton : $(a+b)^4 = a^4 + 4a^3b + 6a^2b^2 + 4ab^3 + b^4.$

► **Exemple 4. Linéarisation de $\sin^5(\theta)$.** Soit θ un réel. On a :

$$\sin(\theta) = \frac{1}{2i} (e^{i\theta} - e^{-i\theta}) \quad \text{d'où : } \sin^5(\theta) = \frac{1}{32i} (e^{i\theta} - e^{-i\theta})^5$$

Il s'ensuit⁹ que :

$$\begin{aligned} \sin^5(\theta) &= \frac{1}{32i} (e^{5i\theta} - 5e^{3i\theta} + 10e^{i\theta} - 10e^{-i\theta} + 5e^{-3i\theta} - e^{-5i\theta}) \\ &= \frac{1}{32i} \left(\underbrace{e^{5i\theta} - e^{-5i\theta}}_{=2i \sin(5\theta)} + \underbrace{5e^{-3i\theta} - 5e^{3i\theta}}_{=-10i \sin(3\theta)} + \underbrace{10e^{i\theta} - 10e^{-i\theta}}_{=20i \sin(\theta)} \right) \end{aligned}$$

D'où finalement : $\forall \theta \in \mathbb{R}, \sin^5(\theta) = \frac{\sin(5\theta) - 5\sin(3\theta) + 10\sin(\theta)}{16}.$

Application 2 — Technique de l'angle-moitié

PROPRIÉTÉ 4.11 - $\forall \theta \in \mathbb{R}, 1 + e^{i\theta} = 2e^{i\theta/2} \cos\left(\frac{\theta}{2}\right)$

PREUVE. Soit θ un réel.

On a : $1 + e^{i\theta} = e^{i\theta/2} (e^{-i\theta/2} + e^{i\theta/2}) = 2e^{i\theta/2} \left(\frac{e^{i\theta/2} + e^{-i\theta/2}}{2} \right) = 2e^{i\theta/2} \cos\left(\frac{\theta}{2}\right)$ ☕

PROPRIÉTÉ 4.12 - $\forall \theta \in \mathbb{R}, 1 - e^{i\theta} = -2ie^{i\theta/2} \sin\left(\frac{\theta}{2}\right)$

PREUVE. Soit θ un réel.

On a : $1 - e^{i\theta} = e^{i\theta/2} (e^{-i\theta/2} - e^{i\theta/2}) = -2ie^{i\theta/2} \left(\frac{e^{i\theta/2} - e^{-i\theta/2}}{2i} \right) = -2ie^{i\theta/2} \sin\left(\frac{\theta}{2}\right)$ ☕

Au-delà de ces deux formules, ce qu'il faut retenir, c'est surtout la technique permettant de les établir, et donc voici deux exemples d'application.

Exemple 1 — Calcul de $\sum_{k=0}^n \cos(k\theta)$ (et $\sum_{k=0}^n \sin(k\theta)$) pour tout $n \in \mathbb{N}$ et pour tout $\theta \in \mathbb{R}$ (exercice classique).

Soient n et k deux entiers naturels et θ un réel. On a : $\cos(k\theta) = \operatorname{Re}(e^{ik\theta})$.¹⁰

Par conséquent : $\sum_{k=0}^n \cos(k\theta) = \sum_{k=0}^n \operatorname{Re}(e^{ik\theta}) \iff \sum_{k=0}^n \cos(k\theta) = \operatorname{Re}\left(\sum_{k=0}^n e^{ik\theta}\right).$

Et de façon analogue : $\sum_{k=0}^n \sin(k\theta) = \operatorname{Im}\left(\sum_{k=0}^n e^{ik\theta}\right)$

9. D'après la formule du binôme de Newton : $(a - b)^5 = a^5 - 5a^4b + 10a^3b^2 - 10a^2b^3 + 5ab^4 - b^5$.

10. Puisqu'en général pour tout réel $\odot$, $e^{i\odot} = \cos(\odot) + i\sin(\odot)$. Prendre $\odot = k\theta$ dans le présent cas.

Il “ne reste plus qu’à” calculer la somme entre parenthèses pour achever la question de cours. En effet, en posant :

$$S_n = \sum_{k=0}^n e^{ik\theta}, \text{ on a donc : } \boxed{\sum_{k=0}^n \cos(k\theta) = \operatorname{Re}(S_n) \quad (\spadesuit)} \text{ et } \boxed{\sum_{k=0}^n \sin(k\theta) = \operatorname{Im}(S_n) \quad (\clubsuit)}$$

$$\text{Or : } S_n = \sum_{k=0}^n e^{ik\theta} \iff S_n = \sum_{k=0}^n (e^{i\theta})^k$$

S_n est la somme des $(n+1)$ premiers termes d’une suite géométrique de raison $e^{i\theta}$. On peut donc lui appliquer la formule que vous connaissez bien¹¹, sous réserve que $e^{i\theta} \neq 1$, c’est-à-dire si $\theta \neq 0 \pmod{2\pi}$.

On suppose donc $\theta \neq 0 \pmod{2\pi}$. Alors :

$$S_n = \frac{1 - (e^{i\theta})^{n+1}}{1 - e^{i\theta}} \iff S_n = \frac{1 - e^{i(n+1)\theta}}{1 - e^{i\theta}} \iff S_n = \frac{e^{i(\frac{n+1}{2})\theta} \left(e^{-i(\frac{n+1}{2})\theta} - e^{i(\frac{n+1}{2})\theta} \right)}{e^{i\frac{\theta}{2}} \left(e^{-i\frac{\theta}{2}} - e^{i\frac{\theta}{2}} \right)}$$

$$\iff S_n = e^{i(\frac{n}{2})\theta} \frac{-2i \sin\left(\frac{n+1}{2}\theta\right)}{-2i \sin\left(\frac{\theta}{2}\right)} \text{ d'où finalement : } \boxed{S_n = e^{i(\frac{n}{2})\theta} \frac{\sin\left(\frac{n+1}{2}\theta\right)}{\sin\left(\frac{\theta}{2}\right)} \quad (\heartsuit)}$$

On déduit de $(\spadesuit)$, $(\clubsuit)$ et $(\heartsuit)$ que :

$$\forall n \in \mathbb{N}, \forall \theta \in \mathbb{R}, \theta \neq 0 \pmod{2\pi}, \left\{ \begin{array}{l} \sum_{k=0}^n \cos(k\theta) = \cos\left(\frac{n\theta}{2}\right) \frac{\sin\left(\frac{n+1}{2}\theta\right)}{\sin\left(\frac{\theta}{2}\right)} \\ \sum_{k=0}^n \sin(k\theta) = \sin\left(\frac{n\theta}{2}\right) \frac{\sin\left(\frac{n+1}{2}\theta\right)}{\sin\left(\frac{\theta}{2}\right)} \end{array} \right.$$

Dans le cas où $\theta = 0 \pmod{2\pi}$, on a $\cos \theta = 1$ et $\sin \theta = 0$ d’où

$$\boxed{\forall n \in \mathbb{N}, \forall \theta \in \mathbb{R}, \theta = 0 \pmod{2\pi}, \sum_{k=0}^n \cos(k\theta) = n+1 \text{ et } \sum_{k=0}^n \sin(k\theta) = 0.}$$

11. $\forall q \in \mathbb{C} \setminus \{1\}, \forall n \in \mathbb{N}, \sum_{k=0}^n q^k = \frac{1 - q^{n+1}}{1 - q}.$

Exemple 2 — Calcul de $\sum_{k=0}^n \binom{n}{k} \cos(k\theta)$ pour tout $n \in \mathbb{N}$ et pour tout $\theta \in \mathbb{R}$ (exercice classique).

Soient θ un réel, et n un entier naturel.

$$\begin{aligned} \text{On a : } \sum_{k=0}^n \binom{n}{k} \cos(k\theta) &= \operatorname{Re} \left(\sum_{k=0}^n \binom{n}{k} e^{ik\theta} \right) = \operatorname{Re} \left((1 + e^{i\theta})^n \right) = \operatorname{Re} \left(\left[2e^{i\theta/2} \cos\left(\frac{\theta}{2}\right) \right]^n \right) \\ &= \operatorname{Re} \left(2^n e^{in\theta/2} \cos^n\left(\frac{\theta}{2}\right) \right) \end{aligned}$$

$$\text{Conclusion. } \forall n \in \mathbb{N}, \forall \theta \in \mathbb{R}, \sum_{k=0}^n \binom{n}{k} \cos(k\theta) = 2^n \cos\left(\frac{n\theta}{2}\right) \cos^n\left(\frac{\theta}{2}\right)$$

Application 3 — Formules d'Euler généralisées

$$\text{LEMME 4.3 - } \forall (\alpha, \beta) \in \mathbb{R}^2, \begin{cases} e^{i\alpha} + e^{i\beta} = 2e^{i(\alpha+\beta)/2} \cos\left(\frac{\alpha-\beta}{2}\right) \\ e^{i\alpha} - e^{i\beta} = 2ie^{i(\alpha+\beta)/2} \sin\left(\frac{\alpha-\beta}{2}\right) \end{cases}$$

PREUVE. Soient α et β deux réels.

$$\begin{aligned} \text{On a : } e^{i\alpha} + e^{i\beta} &= e^{i(\alpha+\beta)/2} (e^{i(\alpha-\beta)/2} + e^{i(\beta-\alpha)/2}) = e^{i(\alpha+\beta)/2} (e^{i(\alpha-\beta)/2} + e^{-i(\alpha-\beta)/2}) = 2e^{i(\alpha+\beta)/2} \cos\left(\frac{\alpha-\beta}{2}\right) \\ \text{Et : } e^{i\alpha} - e^{i\beta} &= e^{i(\alpha+\beta)/2} (e^{i(\alpha-\beta)/2} - e^{i(\beta-\alpha)/2}) = e^{i(\alpha+\beta)/2} (e^{i(\alpha-\beta)/2} - e^{-i(\alpha-\beta)/2}) = 2ie^{i(\alpha+\beta)/2} \sin\left(\frac{\alpha-\beta}{2}\right) \end{aligned}$$



Application 4 — Délinéarisation

Comme son nom le suggère, c'est l'opération qui consiste à faire l'opération inverse de celle de l'application 1, et à exprimer $\cos(n\theta)$ (ou $\sin(n\theta)$) comme un polynôme en $\cos(\theta)$ et $\sin(\theta)$, par le biais suivant :

$$\cos(n\theta) = \operatorname{Re}(e^{in\theta}) = \operatorname{Re}((e^{i\theta})^n) = \operatorname{Re}((\cos(\theta) + i \sin(\theta))^n)$$

Exemple — Expression de $\cos(3a)$ en fonction de $\cos(a)$.

Soit a un nombre réel. On a :

$$\begin{aligned} \cos(3a) &= \operatorname{Re}(e^{i3a}) = \operatorname{Re}((e^{ia})^3) = \operatorname{Re}((\cos(a) + i \sin(a))^3) \\ &= \operatorname{Re}(\cos^3(a) + 3i \cos^2(a) \sin(a) - 3 \cos(a) \sin^2(a) - i \sin^3(a)) = \cos^3(a) - 3 \cos(a) \sin^2(a) \end{aligned}$$

Il ne reste plus qu'à utiliser la relation fondamentale de la trigonométrie pour obtenir :

$$\cos(3a) = \cos^3(a) - 3 \cos(a) (1 - \cos^2(a)) \text{ d'où : } \boxed{\cos(3a) = 4 \cos^3(a) - 3 \cos(a)}$$

En utilisant cette relation avec $a = \pi/9$, on obtient :

$$\cos\left(\frac{\pi}{3}\right) = 4\cos^3\left(\frac{\pi}{9}\right) - 3\cos\left(\frac{\pi}{9}\right) \text{ soit : } 4\cos^3\left(\frac{\pi}{9}\right) - 3\cos\left(\frac{\pi}{9}\right) - \frac{1}{2} = 0$$

$$\text{D'où encore : } 8\cos^3\left(\frac{\pi}{9}\right) - 6\cos\left(\frac{\pi}{9}\right) - 1 = 0.$$

Conclusion : le réel $\cos\left(\frac{\pi}{9}\right)$ est racine de l'équation $8X^3 - 6X - 1 = 0$.

Remarque. On peut déduire de cette relation que $\cos\left(\frac{\pi}{9}\right)$ est **irrationnel** (à l'issue du chapitre "Arithmétique").

4.5 Argument(s) d'un nombre complexe NON NUL

PROPRIÉTÉ 4.13 - Tout nombre complexe NON NUL z peut s'écrire sous la forme $|z|e^{i\theta}$, où θ est un réel, unique modulo 2π .

Avec des quantificateurs : $\forall z \in \mathbb{C}^*, \exists \theta \in \mathbb{R}, z = |z|e^{i\theta}$ (existence de θ)

Et : $[z = |z|e^{i\theta} = |z|e^{i\theta'}] \implies [\theta' = \theta + 2\pi]$ (unicité de θ modulo 2π)

PREUVE. Soit z un nombre complexe non nul.

$$\text{Posons : } Z = \frac{z}{|z|}. \text{ On a : } |Z| = \left| \frac{z}{|z|} \right| = \frac{|z|}{||z||} = \frac{|z|}{|z|} = 1.$$

On en déduit que $Z \in \mathbb{U}$; à ce titre, il existe un réel θ tel que : $Z = e^{i\theta}$. On en déduit que : $z = |z|e^{i\theta}$.

Supposons à présent qu'il existe deux réels θ et θ' tels que : $z = |z|e^{i\theta} = |z|e^{i\theta'}$. On en déduit, puisque $|z| \neq 0$ par hypothèse, que : $e^{i\theta} = e^{i\theta'}$. Il s'ensuit que $\theta' = \theta + 2\pi$. ☕

DÉFINITION 4.8 - Avec les notations de la propriété précédente, θ est un **argument** du nombre complexe z , et est noté $\arg(z)$.

Remarque : pour enfoncer le clou, il résulte de la définition qu'un argument d'un complexe non nul z est défini à un multiple de 2π près.

Exemples : $\arg(1) = 0 + 2\pi$ et plus généralement $\arg(x) = 0 + 2\pi$ pour tout $x \in \mathbb{R}_+^*$.

$\arg(-1) = \pi + 2\pi$ et plus généralement $\arg(x) = \pi + 2\pi$ pour tout $x \in \mathbb{R}_-^*$.

$\arg(i) = \frac{\pi}{2} + 2\pi$ et plus généralement $\arg(xi) = \frac{\pi}{2} + 2\pi$ pour tout $x \in \mathbb{R}_+^*$.

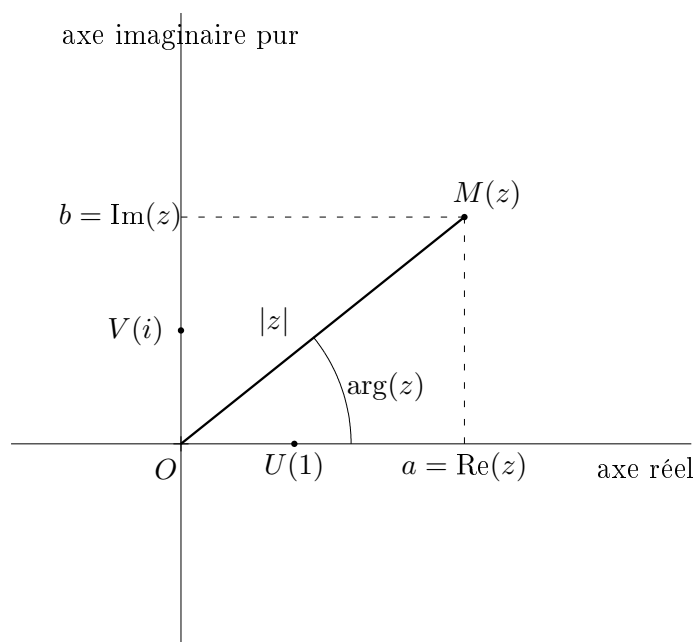
$\arg(-i) = -\frac{\pi}{2} + 2\pi$ et plus généralement $\arg(xi) = -\frac{\pi}{2} + 2\pi$ pour tout $x \in \mathbb{R}_-^*$.

$\arg(1+i) = \frac{\pi}{4} + 2\pi$ et plus généralement $\arg(x+xi) = \frac{\pi}{4} + 2\pi$ pour tout $x \in \mathbb{R}_+^*$.

$\arg(-1-i) = -\frac{3\pi}{4} + 2\pi$ et plus généralement $\arg(x+xi) = -\frac{3\pi}{4} + 2\pi$ pour tout $x \in \mathbb{R}_-^*$.

$\arg(1+i\sqrt{3}) = \frac{\pi}{3} + 2\pi$; $\arg(-1+i\sqrt{3}) = \frac{2\pi}{3} + 2\pi$; $\arg(\sqrt{3}+i) = \frac{\pi}{6} + 2\pi$; $\arg(-\sqrt{3}+i) = \frac{5\pi}{6} + 2\pi$...

Interprétation géométrique :



PROPRIÉTÉ 4.14 - (Caractérisation des réels et imaginaires purs). Soit $z \in \mathbb{C}^*$. On a :

$$[z \in \mathbb{R}^*] \iff [\arg(z) = 0 \text{ } [2\pi]] \quad \text{et} \quad [z \in i\mathbb{R}^*] \iff \left[\arg(z) = \frac{\pi}{2} \text{ } [2\pi] \right]$$

PREUVE. Soit $z \in \mathbb{C}^*$. En notant $\theta = \arg(z) \text{ } [2\pi]$, on a : $z = |z| e^{i\theta}$, d'où : $z = |z| \cos(\theta) + i |z| \sin(\theta)$.

Il s'ensuit que : $\operatorname{Re}(z) = |z| \cos(\theta)$ et $\operatorname{Im}(z) = |z| \sin(\theta)$.

D'où : $[z \in \mathbb{R}^*] \iff [\operatorname{Im}(z) = 0] \iff [|z| \sin(\theta) = 0] \underset{|z| \neq 0}{\iff} [\sin(\theta) = 0] \iff [\theta = 0 \text{ } [2\pi]]$.

De façon analogue : $[z \in i\mathbb{R}^*] \iff [\operatorname{Re}(z) = 0] \iff [|z| \cos(\theta) = 0] \underset{|z| \neq 0}{\iff} [\cos(\theta) = 0] \iff \left[\theta = \frac{\pi}{2} \text{ } [2\pi] \right]$ ☕

PROPRIÉTÉ 4.15 - (Propriétés algébriques des arguments). Soient z et z' deux nombres complexes non nuls. On a :

1/ $\arg(\bar{z}) = -\arg(z) \text{ } [2\pi]$

2/ $\arg(zz') = \arg(z) + \arg(z') \text{ } [2\pi]$

3/ $\arg(-z) = \arg(z) + \pi \text{ } [2\pi]$

4/ $\forall n \in \mathbb{N}, \arg(z^n) = n \arg(z) \text{ } [2\pi]$

5/ $\arg\left(\frac{1}{z'}\right) = -\arg(z') \text{ } [2\pi]$

6/ $\arg\left(\frac{z}{z'}\right) = \arg(z) - \arg(z') \text{ } [2\pi]$

PREUVE. Soient z et z' deux nombres complexes non nuls. Posons $\theta = \arg(z) [2\pi]$ et $\theta' = \arg(z') [2\pi]$, de telle sorte que : $z = |z| e^{i\theta}$ et $z' = |z'| e^{i\theta'}$.

1/ On a : $\bar{z} = \overline{|z| e^{i\theta}} = \overline{|z|} \times \overline{e^{i\theta}} = |z| e^{-i\theta}$. Puisqu'en outre : $|\bar{z}| = |z|$, on en déduit que : $\bar{z} = |z| e^{-i\theta}$.

Par suite : $\arg(\bar{z}) = -\arg(z) [2\pi]$.

2/ On a : $zz' = |z| e^{i\theta} |z'| e^{i\theta'} = |zz'| e^{i(\theta+\theta')}$. On en déduit que : $\arg(zz') = \arg(z) + \arg(z') [2\pi]$.

3/ D'après le point précédent : $\arg(-z) = \arg((-1)z) = \arg(-1) + \arg(z) = \pi + \arg(z) [2\pi]$.

4/ Conséquence du point 2, par une récurrence immédiate.

5/ D'après le point 2, on a : $\arg\left(z' \times \frac{1}{z'}\right) = \arg(z') + \arg\left(\frac{1}{z'}\right) [2\pi]$.

D'autre part : $\arg\left(z' \times \frac{1}{z'}\right) = \arg(1) = 0 [2\pi]$.

On déduit de ces deux égalités que : $\arg\left(\frac{1}{z'}\right) = -\arg(z') [2\pi]$.

6) Conséquence immédiate des points 2 et 5. ☕

Remarque. Une nouvelle fois, la propriété 2 peut être généralisée au produit de n nombres complexes non nuls. Explicitement :

$$\forall n \in \mathbb{N}^*, \forall (z_1, \dots, z_n) \in (\mathbb{C}^*)^n, \arg\left(\prod_{k=1}^n z_k\right) = \sum_{k=1}^n \arg(z_k) [2\pi]$$

Exemple d'application 1. Déterminer une CNS sur l'entier n pour que $(1+i)^n$ soit un réel positif.

On a : $\arg(1+i) = \frac{\pi}{4} [2\pi]$. Par suite : $\forall n \in \mathbb{N}, \arg((1+i)^n) = \frac{n\pi}{4} [2\pi]$.

Par ailleurs, pour tout $Z \in \mathbb{C}^*$, on a : $Z \in \mathbb{R}_+^* \iff \arg(Z) = 0 [2\pi]$.

On en déduit que $(1+i)^n$ soit un réel positif si et seulement si : $\frac{n\pi}{4} = 0 [2\pi] \iff n = 0 [8]$.

Conclusion. On a : $[(1+i)^n] \iff [n \text{ est multiple de } 8]$.

Exemple d'application 2. Une autre méthode pour le calcul de $\cos\left(\frac{\pi}{8}\right)$.

On pose $Z = \sqrt{2+\sqrt{2}} + i\sqrt{2-\sqrt{2}}$. On a : $|Z| = \sqrt{2+\sqrt{2}+2-\sqrt{2}} = 2$.

Par ailleurs : $Z^2 = (2+\sqrt{2}) + 2i\sqrt{2+\sqrt{2}}\sqrt{2-\sqrt{2}} - (2-\sqrt{2}) = 2\sqrt{2} + 2i\sqrt{2}$.

D'où : $Z^2 = 4e^{i\pi/4}$. Par suite, $\arg(Z^2) = \frac{\pi}{4} [2\pi]$ d'où : $2\arg(Z) = \frac{\pi}{4} [2\pi]$, donc : $\arg(Z) = \frac{\pi}{8} [\pi]$.

Puisque les parties réelle et imaginaire de Z sont positives, on peut affirmer : $\arg(Z) = \frac{\pi}{8} [2\pi]$.

$$\text{On en déduit que : } \boxed{\cos\left(\frac{\pi}{8}\right) = \frac{\operatorname{Re}(Z)}{|Z|} = \frac{\sqrt{2+\sqrt{2}}}{2}} \quad \text{et} \quad \boxed{\sin\left(\frac{\pi}{8}\right) = \frac{\operatorname{Im}(Z)}{|Z|} = \frac{\sqrt{2-\sqrt{2}}}{2}}$$

4.6 Exponentielle complexe

DÉFINITION 4.9 - Pour tout $z \in \mathbb{C}$, de forme algébrique $z = a + ib$, on pose $e^z = e^a e^{ib}$.

Exemples : $e^{\ln(2)+i\frac{\pi}{2}} = 2i$; $e^{\ln(4)+i\frac{\pi}{4}} = 2\sqrt{2} + 2i\sqrt{2}$; $e^{1+i\pi} = -e$.

PROPRIÉTÉ 4.16 - (Propriétés de l'exponentielle complexe).

$$1/ \forall z \in \mathbb{C}, |e^z| = e^{\operatorname{Re}(z)}$$

$$2/ \forall z \in \mathbb{C}, \arg(e^z) = \operatorname{Im}(z) [2\pi]$$

$$3/ \forall z \in \mathbb{C}, z \neq 0 \quad (l'exponentielle complexe ne s'annule pas sur \mathbb{C})$$

$$4/ \forall (z, z') \in \mathbb{C}^2, e^{z+z'} = e^z e^{z'}$$

$$5/ \forall n \in \mathbb{N}, \forall z \in \mathbb{C}, (e^z)^n = e^{nz}$$

$$6/ \forall z \in \mathbb{C}, e^{-z} = \frac{1}{e^z}$$

$$7/ \forall z \in \mathbb{C}, e^z = e^{z+2i\pi} \quad (2i\pi\text{-périodicité})$$

PREUVE. Soient $z = a + ib$ et $z' = a' + ib'$ (avec a, a', b et b' réels) deux complexes.

$$1/ \& 2/ \text{ Par définition, on a : } e^z = e^a e^{ib}. \text{ D'où : } |e^z| = e^a \text{ et } \arg(e^z) = b [2\pi].$$

$$\text{On en déduit que : } \forall z \in \mathbb{C}, |e^z| = e^{\operatorname{Re}(z)} \text{ et } \forall z \in \mathbb{C}, \arg(e^z) = \operatorname{Im}(z) [2\pi].$$

$$3/ \text{ D'après ce qui précède, on a : } |e^z| = e^{\operatorname{Re}(z)} \neq 0. \text{ Il s'ensuit que : } \forall z \in \mathbb{C}, e^z \neq 0.$$

$$4/ \text{ On a : } e^{z+z'} = e^{(a+a')+(b+b')} = e^{a+a'} e^{i(b+b')} = e^a e^{a'} e^{ib} e^{ib'} = e^{a+ib} e^{a'+ib'} = e^z e^{z'}.$$

$$5/ \text{ On a : } (e^z)^n = (e^a e^{ib})^n = e^{na} e^{nib} = e^{n(a+ib)} = e^{nz}.$$

$$6/ \text{ On a : } e^{-z} = e^{-a} e^{-ib} = \frac{1}{e^a} \times \frac{1}{e^{ib}} = \frac{1}{e^{a+ib}} = \frac{1}{e^z}.$$

$$7/ \text{ D'après le point 4, on a : } e^{z+2i\pi} = e^z \underbrace{e^{2i\pi}}_{=1} = e^z. \text{ ☕}$$

THÉORÈME 4.2 - (Ensemble des solutions de $e^z = \omega$, avec $\omega \in \mathbb{C}^*$).

Pour tout complexe non-nul ω , l'ensemble des solutions de l'équation $e^z = \omega$ est :

$$\{\ln|\omega| + i \arg(\omega) + 2ik\pi / k \in \mathbb{Z}\}$$

PREUVE. Soit ω un complexe non nul ; on peut écrire $\omega = |\omega| e^{i \arg(\omega)}$.

Par ailleurs soit $z = a + ib$ (avec a et b réels) un nombre complexe ;

$$\begin{aligned} & [z \text{ est solution de } e^z = \omega] \\ \iff e^a e^{ib} = |\omega| e^{i \arg(\omega)} & \iff \begin{cases} e^a = |\omega| \\ b = \arg(\omega) [2\pi] \end{cases} \iff \begin{cases} a = \ln|\omega| \\ \exists k \in \mathbb{Z}, b = \arg(\omega) + 2k\pi \end{cases} \end{aligned}$$

En résumé, un complexe z est solution de $e^z = \omega$ si et seulement si il existe un entier relatif k tel que $z = \ln |\omega| + i \arg(\omega) + 2ik\pi$.

Conclusion : pour tout nombre complexe non-nul $\arg(\omega)$, l'ensemble des nombres complexes solutions de l'équation $e^z = \arg(\omega)$ est $\{\ln |\arg(\omega)| + i \arg(\arg(\omega)) + 2ik\pi / k \in \mathbb{Z}\}$



Remarque. On savait déjà que la fonction exponentielle complexe était à valeurs dans $\mathbb{C}^*$ (puisque $e^z \neq 0$ pour tout $z \in \mathbb{C}$). Ce que précise le théorème, c'est que tout élément de $\mathbb{C}^*$ admet un antécédent (mieux, une infinité d'antécédents) par l'exponentielle complexe. Dans ce contexte, on dira que l'exponentielle complexe est *surjective* ; la notion de surjectivité (et d'autres !) sera étudiée dans le chapitre "Applications".

4.7 Equations du second degré à coefficients complexes

4.7.1 Racines carrées complexes

DÉFINITION 4.10 - Soit z un nombre complexe.

Une **racine carrée (complexe)** de z est un complexe ω tel que $\omega^2 = z$.

Exemples : 2 et -2 sont des racines carrées de 4 ; 3i et $-3i$ sont des racines carrées de -9 ; $1+i$ et $-1-i$ sont des racines carrées de $2i$; $\sqrt{2}e^{i\pi/12}$ et $-\sqrt{2}e^{i\pi/12}$ sont des racines carrées de $\sqrt{3}+i$.

THÉORÈME 4.3 - (Racines carrées dans $\mathbb{C}$).

Tout nombre complexe z non nul possède exactement deux racines carrées dans $\mathbb{C}$:

$$\pm |z|^{1/2} e^{i \arg(z)/2}$$

0 ne possède qu'une racine carrée (0 lui-même).

PREUVE. Soit z un nombre complexe non nul. On peut écrire : $z = |z| e^{i\theta}$, où $\theta = \arg(z) \in [2\pi]$.

Par définition, un nombre complexe ω est une racine carrée de z si $\omega^2 = z$. Exploitions cette relation :

$$\omega^2 = z \iff (|\omega| e^{i \arg(\omega)})^2 = |z| e^{i\theta} \iff |\omega|^2 e^{2i \arg(\omega)} = |z| e^{i\theta}$$

Cette dernière égalité entre deux formes exponentielles entraîne l'égalité des modules, et des arguments (*modulo* 2π) de chaque terme, c'est-à-dire :

$$\begin{cases} |\omega|^2 = |z| \\ 2 \arg(\omega) = \theta \pmod{2\pi} \end{cases} \iff \begin{cases} |\omega| = \sqrt{|z|} \\ \arg(\omega) = \frac{\theta}{2} \pmod{\pi} \end{cases}$$

On a ainsi établi que : $(\omega \text{ est une racine carrée de } z) \iff (|\omega| = \sqrt{|z|}) \wedge \left(\arg(\omega) = \frac{\theta}{2} \pmod{\pi}\right)$.

Reste à préciser l'argument de ω :

$$\left(\arg(\omega) = \frac{\theta}{2} \pmod{\pi}\right) \iff \left(\exists k \in \mathbb{Z}, \arg(\omega) = \frac{\theta}{2} + k\pi\right)$$

D'où : $\omega = |z|^{1/2} e^{i(\frac{\theta}{2} + k\pi)} = |z|^{1/2} e^{i\frac{\theta}{2}} e^{ik\pi} = (-1)^k |z|^{1/2} e^{i\frac{\theta}{2}}$

Ainsi $\omega = |z|^{1/2} e^{i\frac{\theta}{2}}$ (si k est pair) ou $\omega = -|z|^{1/2} e^{i\frac{\theta}{2}}$ (si k est impair).

En résumé : z possède exactement deux racines carrées dans $\mathbb{C}$, qui sont $\pm \sqrt{|z|} e^{i\frac{\theta}{2}}$ (ou $\pm |z|^{1/2} e^{i\frac{\theta}{2}}$). ☕

► 2 méthodes pour déterminer les racines carrées d'un nombre complexe non nul.

Soit z un complexe non nul. On décrit ci-dessous deux procédés permettant de déterminer les deux racines carrées de z .

1ère méthode — Lorsque l'on connaît la forme exponentielle de z .

Dans ce cas, on écrit $z = |z| e^{i\theta}$ où θ désigne un argument de z , et on utilise la formule obtenue précédemment pour obtenir les racines carrées de z : $\pm |z|^{1/2} e^{i\frac{\theta}{2}}$.

Exemple. On pose $z = 4 + 4i$. On a : $|z| = \sqrt{32} e^{i\pi/4} = 2^{5/2} e^{i\pi/4}$. Il s'ensuit que les deux racines carrées complexes de z sont : $\pm 2^{5/4} e^{i\pi/8}$.

2ème méthode — Lorsque l'on ne connaît pas la forme exponentielle de z .

Dans ce cas, on recherche une racine carrée ω de z sous forme algébrique : $\omega = a + ib$ (avec a et b réels). Alors :

$$[\omega^2 = z] \iff [a^2 - b^2 + 2iab = z] \iff \begin{cases} a^2 - b^2 = \operatorname{Re}(z) & (\text{égalité des parties réelles}) \\ 2ab = \operatorname{Im}(z) & (\text{égalité des parties imaginaires}) \\ a^2 + b^2 = |z| & (\text{égalité des modules}) \end{cases}$$

Exemple. On pose $z = 3 + 4i$. En notant θ un argument de z , on a : $\cos(\theta) = 3/5$ et $\sin(\theta) = 4/5$... ce qui ne nous conduit pas vraiment à des valeurs remarquables !

On pose donc $\omega = a + ib$ (avec a et b réels), et on applique le raisonnement précédent.

$$[\omega^2 = z] \iff [a^2 - b^2 + 2iab = 3 + 4i] \iff \begin{cases} a^2 - b^2 = 3 & (\text{égalité des parties réelles}) \\ 2ab = 4 & (\text{égalité des parties imaginaires}) \\ a^2 + b^2 = 5 & (\text{égalité des modules}) \end{cases}$$

En exploitant les première et dernière ligne de ce système, on déduit facilement que : $a^2 = 4$ et $b^2 = 1$. D'où $a = \pm 1$, et $b = \pm 1$; ce qui nous donne virtuellement 4 couples (a, b) solutions, et donc 4 racines carrées pour $3 + 4i$. Heureusement, la seconde équation permet de rétablir l'équilibre : le produit ab étant positif, les réels a et b sont de même signe. Le système possède donc exactement deux couples solutions : $(3, 1)$ et $(-3, -1)$.

On peut donc conclure que les deux racines carrées du complexe $3 + 4i$ sont $1 + i$ et $-1 - i$.

4.7.2 Résolution de $az^2 + bz + c = 0$

THÉORÈME 4.4 - (Equations du second degré à coefficients complexes).

Soit $(a, b, c) \in \mathbb{C}^3$, avec $a \neq 0$. On note (E) l'équation $az^2 + bz + c = 0$, et $\Delta = b^2 - 4ac$ son discriminant.

► Si $\Delta = 0$, alors (E) possède une unique solution : $-\frac{b}{2a}$.

► Si $\Delta \neq 0$, alors (E) possède exactement deux solutions : $-\frac{b \pm \delta}{2a}$, où δ est une racine carrée dans $\mathbb{C}$ de Δ .

PREUVE. Soient a, b et c trois complexes, avec $a \neq 0$. On a :

$$(E) : az^2 + bz + c = 0 \iff a \left(z^2 + \frac{b}{a}z + \frac{c}{a} \right) = 0 \iff z^2 + \frac{b}{a}z + \frac{c}{a} = 0 \iff \left(z + \frac{b}{2a} \right)^2 - \frac{b^2}{4a^2} + \frac{c}{a} = 0$$

$$\iff \left(z + \frac{b}{2a} \right)^2 = \frac{b^2 - 4ac}{4a^2} \quad (\spadesuit)$$

Posons $\Delta = b^2 - 4ac$, et distinguons deux cas :

- Si $\Delta = 0$: alors on déduit de $(\spadesuit)$ que $\left(z + \frac{b}{2a} \right)^2 = 0$, donc que $z + \frac{b}{2a} = 0$, d'où $z = -\frac{b}{2a}$.
- Si $\Delta \neq 0$: alors Δ possède exactement deux racines carrées dans $\mathbb{C}$, que l'on note $\pm\delta$. On peut alors écrire :

$$\left(z + \frac{b}{2a} \right)^2 = \frac{b^2 - 4ac}{4a^2} \iff \left(z + \frac{b}{2a} \right)^2 = \left(\frac{\delta}{2a} \right)^2 \iff z + \frac{b}{2a} = \pm \frac{\delta}{2a} \iff z = \frac{-b \pm \delta}{2a}$$

Ce qui prouve le théorème. ☕

Exemple. Soit θ un réel quelconque. On considère l'équation (E) :

$$(E) : z^4 - 2i \sin(\theta) z^2 - 1 = 0$$

On pose $Z = z^2$. L'équation (E) se réécrit alors : $(E') \quad Z^2 - 2i \sin(\theta) Z - 1 = 0$. Calculons le discriminant de (E') . On a :

$$\Delta = (-2i \sin(\theta))^2 - 4 \times (-1) = -4 \sin^2(\theta) + 4 = 4(1 - \sin^2(\theta)) = 4 \cos^2(\theta)$$

D'où : $\Delta = \pm (2 \cos(\theta))^2$. On distingue alors deux cas :

- **Cas 1.** Si $\Delta = 0$, c'est-à-dire si $\theta = \frac{\pi}{2} [\pi]$. Alors l'équation (E') possède une unique solution : $Z_0 = i \sin(\theta)$. Or $\sin(\theta) = 1$ ou -1 suivant que $\theta = \pi/2 [2\pi]$ ou $\theta = -\pi/2 [2\pi]$; ce qui conduit à distinguer deux sous cas.
- **Sous-cas 1.1.** Si $\theta = \frac{\pi}{2} [2\pi]$. Alors $\sin(\theta) = 1$, et l'équation (E') possède comme unique solution : $Z_0 = i$. En revenant à la variable initiale, on en déduit que : $z^2 = i$, d'où : $z = \pm e^{i\pi/4}$. Dans ce cas, l'équation (E) possède donc exactement deux solutions : $\pm e^{i\pi/4}$.
- **Sous-cas 1.2.** Si $\theta = -\frac{\pi}{2} [2\pi]$. Alors $\sin(\theta) = -1$, et l'équation (E') possède comme unique solution : $Z_0 = -i$. En revenant à la variable initiale, on en déduit que : $z^2 = -i$, d'où : $z = \pm e^{-i\pi/4}$. Dans ce cas, l'équation (E) possède donc exactement deux solutions : $\pm e^{-i\pi/4}$.
- **Cas 2.** Si $\Delta \neq 0$, c'est-à-dire si $\theta \neq \frac{\pi}{2} [\pi]$. Alors l'équation (E') possède exactement deux solutions : $Z_+ = i \sin(\theta) + \cos(\theta)$ et $Z_- = i \sin(\theta) - \cos(\theta)$, soit encore : $Z_+ = e^{i\theta}$ et $Z_- = -e^{-i\theta} = e^{i(\pi-\theta)}$.

En revenant à la variable initiale, on en déduit que :

$$(E) \iff [z^2 = e^{i\theta} \vee z^2 = e^{i(\pi-\theta)}] \iff [z = \pm e^{i\theta/2} \vee z = \pm e^{i(\pi-\theta)/2}]$$

Dans ce cas, l'équation (E) possède donc exactement quatre solutions : $\pm e^{i\theta/2}$ et $\pm e^{i(\pi-\theta)/2}$.

► **Conclusion.** L'ensemble des solutions de l'équation $z^4 - 2i \sin(\theta) z^2 - 1 = 0$ est :

$$\begin{cases} \{\pm e^{i\pi/4}\} & \text{si } \theta = \pi/2 \ [2\pi] \\ \{\pm e^{-i\pi/4}\} & \text{si } \theta = -\pi/2 \ [2\pi] \\ \{\pm e^{i\theta/2}; \pm e^{i(\pi-\theta)/2}\} & \text{si } \theta \neq \pi/2 \ [\pi] \end{cases}$$

Relations coefficients-racines : avec les notations du théorème, lorsque (E) possède deux racines z_1 et z_2 , on a :

$$z_1 + z_2 = -\frac{b}{a} \quad \text{et} \quad z_1 z_2 = \frac{c}{a}$$

“Dans l'autre sens”, si z_1 et z_2 sont deux complexes, ils sont racines de l'équation $X^2 - SX + P = 0$ (où l'on a noté P et S les produit et somme de z_1 et z_2).

Exemple. Considérons l'équation (E) : $z^2 + 2017z - 2018 = 0$. Le réel 1 est racine évidente de (E). Comme il s'agit d'une équation du second degré, on en déduit que la deuxième racine est -2018 (le produit des racines étant égal à “ c/a ”).

Conclusion. (E) a exactement deux racines : 1 et -2018 .

4.8 Racines n -ièmes

4.8.1 Racines n -ièmes de l'unité

DÉFINITION 4.11 - Soit $n \in \mathbb{N}$, $n \geq 2$.

On appelle **racine n -ième de l'unité** un nombre complexe ω tel que $\omega^n = 1$.

Notation. On note $\mathbb{U}_n$ l'ensemble des racines n -ièmes de l'unité.

Remarque. Le “ $\mathbb{U}$ ” de la notation est justifié par l'observation que toute racine de l'unité est de module 1. Dans le langage ensembliste : $\mathbb{U}_n \subset \mathbb{U}$.

Exemples. $\mathbb{U}_2 = \{\pm 1\}$ et $\mathbb{U}_4 = \{\pm 1; \pm i\}$.

L'ensemble $\mathbb{U}_3$ est appelé l'ensemble des racines cubiques (plutôt de 3-ièmes) de l'unité. On a : $\mathbb{U}_3 = \{1; e^{2i\pi/3}; e^{4i\pi/3}\}$.

Notation. On note $j = e^{2i\pi/3}$. Avec cette notation, on a : $e^{4i\pi/3} = j^2$, ou $e^{4i\pi/3} = \bar{j}$. Ainsi :

$$\mathbb{U}_3 = \{1, j, j^2\} \quad \text{ou encore} \quad \mathbb{U}_3 = \{1, j, \bar{j}\}$$

PROPRIÉTÉ 4.17 - (Propriétés des racines de l'unité). Soit n un entier naturel supérieur ou égal à 2.

1/ Toute racine n -ième de l'unité est de module 1. Ensemblistement : $\mathbb{U}_n \subset \mathbb{U}$.

2/ $1 \in \mathbb{U}_n$.

3/ Soient z et z' deux éléments de $\mathbb{U}_n$. Alors $zz' \in \mathbb{U}_n$.

4/ Soit $z \in \mathbb{U}_n$. Alors $z^{-1} = \bar{z} \in \mathbb{U}_n$.

PREUVE. Soit n un entier naturel supérieur ou égal à 2.

1/ Soit $z \in \mathbb{U}_n$. Alors : $z^n = 1$, donc $|z|^n = 1$, donc $|z| = 1$. Ce qui signifie que $z \in \mathbb{U}$.

2/ Trivial.

3/ Soient z et z' deux éléments de $\mathbb{U}_n$; on a alors $z^n = 1$ et $z'^n = 1$. Sous cette hypothèse : $(zz')^n = z^n z'^n = 1$. Ce qui signifie que $zz' \in \mathbb{U}_n$.

4/ Soit $z \in \mathbb{U}_n$. Alors z est de module 1 (d'après le point 1/), d'où : $z^{-1} = \bar{z}$. En outre : $z^n = 1$ donc $\bar{z}^n = 1$ donc $\bar{z} \in \mathbb{U}_n$. ☕

Remarque : pour anticiper sur la notion de *groupe*, les propriétés ci-dessus permettent d'affirmer que $\mathbb{U}_n$ est muni d'une loi de composition interne (la multiplication, d'après le point 3/) qui est associative; en outre, cette loi possède un élément neutre dans $\mathbb{U}_n$ (le complexe 1, d'après le point 2/), et tout élément de $\mathbb{U}_n$ est inversible pour cette loi dans $\mathbb{U}_n$ (d'après le point 4/).

Dans ce contexte, on dit que $(\mathbb{U}_n, \times)$ est un **groupe**.¹²

Cette remarque faite, passons maintenant à l'énoncé suivant, qui assure qu'il existe exactement n racines n -ièmes de l'unité, et qui les donne explicitement.

THÉORÈME 4.5 - (Racines n -ièmes de l'unité).

Soit n un entier naturel supérieur ou égal à 2. On a

$$\mathbb{U}_n = \left\{ e^{\frac{2ik\pi}{n}} / k \in \llbracket 0; n-1 \rrbracket \right\} \quad \text{ou} \quad \mathbb{U}_n = \{ \omega^k / k \in \llbracket 0; n-1 \rrbracket \}$$

en ayant pris la précaution de noter $\omega = e^{\frac{2i\pi}{n}}$.

PREUVE. Soit n un entier naturel supérieur ou égal à 2, et soit z une racine n -ième de l'unité.

Par définition, on a alors : $z^n = 1$. Cette condition impliquant la non-nullité de z , on peut écrire z sous forme exponentielle : $z = |z| e^{i\theta}$ (avec $\theta = \arg(z) [2\pi]$).

$$\begin{aligned} \text{Ceci posé : } [z^n = 1] &\iff |z|^n e^{in\theta} = 1 \iff \begin{cases} |z|^n = 1 \\ n\theta = 0 [2\pi] \end{cases} \iff \begin{cases} |z| = 1 \\ \theta = 0 [2\pi/n] \end{cases} \\ &\iff \begin{cases} |z| = 1 \\ \exists k \in \mathbb{Z}, \theta = 2k\pi/n \end{cases} \end{aligned}$$

En résumé : $[z^n = 1] \iff [\exists k \in \mathbb{Z}, z = e^{2ik\pi/n}]$. On a ainsi établi que : $\mathbb{U}_n = \left\{ e^{\frac{2ik\pi}{n}} / k \in \mathbb{Z} \right\}$ (♠).

12. Pour utiliser des gros mots, le groupe multiplicatif des racines n -ièmes de l'unité.

Reste à justifier que l'ensemble $\mathbb{U}_n$ est fini, et contient n éléments. Pour cela, on **admet** provisoirement¹³ le théorème de la division euclidienne dans $\mathbb{Z}$:

Théorème (division euclidienne dans $\mathbb{Z}$). $\forall (a, b) \in \mathbb{Z} \times \mathbb{N}^*, \exists! (q, r) \in \mathbb{Z}^2, \begin{cases} a = bq + r \\ r \in \llbracket 0, b - 1 \rrbracket \end{cases}$

Posons : $E = \left\{ e^{\frac{2ik\pi}{n}} / k \in \llbracket 0; n - 1 \rrbracket \right\}$ ($\heartsuit$). Etablir le théorème, c'est établir l'égalité : $E = \mathbb{U}_n$. Pour ce faire, on peut procéder par double inclusion.

► Montrons $E \subset \mathbb{U}_n$. Il résulte clairement des descriptions respectives de E ($\heartsuit$) et de $\mathbb{U}_n$ ($\spadesuit$) que tout élément de E appartient à $\mathbb{U}_n$,¹⁴ donc $E \subset \mathbb{U}_n$.

► Montrons $\mathbb{U}_n \subset E$. Soit $z \in \mathbb{U}_n$. D'après ($\spadesuit$), il existe un entier relatif k tel que : $z = e^{\frac{2ik\pi}{n}}$. On effectue alors la division euclidienne de k par n . D'après le théorème du même nom :

$$\exists! (q, r) \in \mathbb{Z}^2, \begin{cases} k = nq + r \\ r \in \llbracket 0, n - 1 \rrbracket \end{cases}$$

Ainsi : $z = e^{\frac{2ik\pi}{n}} = e^{\frac{2i(nq+r)\pi}{n}} = e^{\frac{2inq\pi}{n}} \times e^{\frac{2ir\pi}{n}} = \underbrace{\left(e^{\frac{2in\pi}{n}} \right)^q}_{=1} \times e^{\frac{2ir\pi}{n}}$ d'où : $z = e^{\frac{2ir\pi}{n}}$. Puisque $r \in \llbracket 0, n - 1 \rrbracket$,

cette dernière écriture signifie que $z \in E$. Ce qui prouve que $\mathbb{U}_n \subset E$.

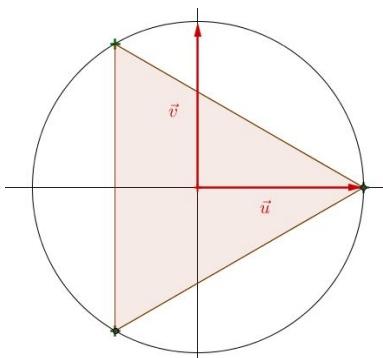
D'après la règle de double inclusion, on a : $E = \mathbb{U}_n$. Par suite : $\mathbb{U}_n = \left\{ e^{\frac{2ik\pi}{n}} / k \in \llbracket 0; n - 1 \rrbracket \right\}$ ☕

Interprétation géométrique des racines de l'unité

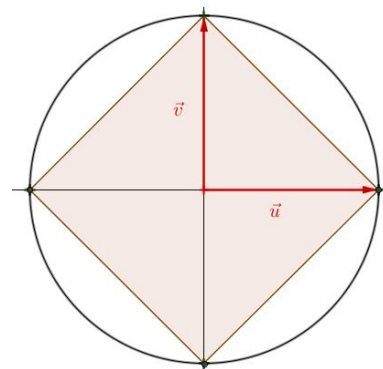
Il résulte de la description de l'ensemble des racines n -ièmes de l'unité que les éléments de $\mathbb{U}_n$ sont les affixes des sommets d'un **polygone régulier à n côtés** inscrit dans le cercle unité.

On illustre ci-dessous cette affirmation par quelques exemples.

► Les éléments de $\mathbb{U}_3$ ($1, j$ et $j^2 = \bar{j}$) sont les affixes des sommets d'un triangle équilatéral.



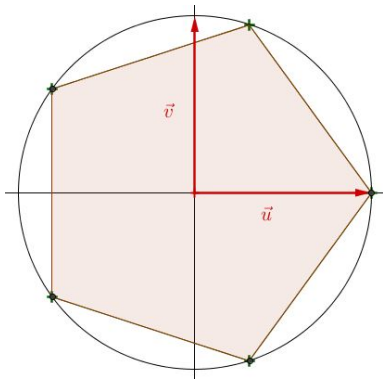
► Les éléments de $\mathbb{U}_4$ ($1, i, i^2 = -1$ et $i^3 = -i$) sont les affixes des sommets d'un carré (de plus $\mathbb{U}_4$ contient $\mathbb{U}_2$).



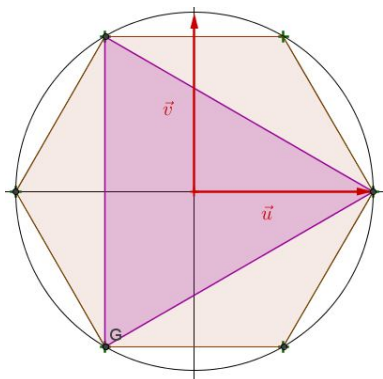
13. En attendant le chapitre d'arithmétique.

14. Ceci repose sur l'observation puissante que tout entier compris entre 0 et $n - 1$ est en particulier un entier relatif.

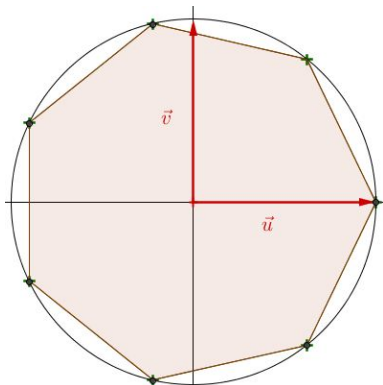
► Les éléments de $\mathbb{U}_5$ sont les affixes des sommets d'un pentagone régulier.



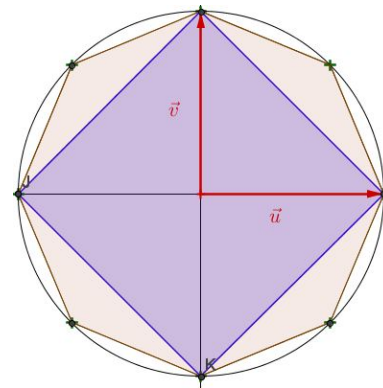
► Les éléments de $\mathbb{U}_6$ sont les affixes des sommets d'un hexagone régulier (de plus $\mathbb{U}_6$ contient $\mathbb{U}_3$ et $\mathbb{U}_2$).



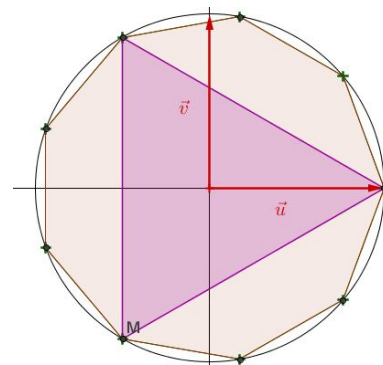
► Les éléments de $\mathbb{U}_7$ sont les affixes des sommets d'un heptagone régulier.



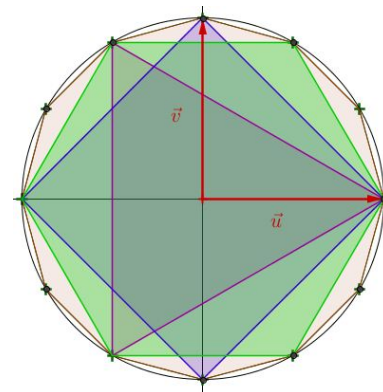
► Les éléments de $\mathbb{U}_8$ sont les affixes des sommets d'un octogone régulier (de plus $\mathbb{U}_8$ contient $\mathbb{U}_4$ et $\mathbb{U}_2$).



► Les éléments de $\mathbb{U}_9$ sont les affixes des sommets d'un enneagone régulier (et $\mathbb{U}_9$ contient $\mathbb{U}_3$).



► Les éléments de $\mathbb{U}_{12}$ sont les affixes des sommets d'un dodécagone régulier (et $\mathbb{U}_{12}$ contient $\mathbb{U}_2$, $\mathbb{U}_3$, $\mathbb{U}_4$ et $\mathbb{U}_6$).



Pour insister sur un point déjà démontré, on voit clairement sur ces illustrations que l'ensemble $\mathbb{U}_n$ est stable par conjugaison.

Par ailleurs, il est encore très “visible” que pour deux entiers naturels n et p , on a :

$$[\mathbb{U}_n \subset \mathbb{U}_p] \iff [n \text{ divise } p] \quad (\text{petit exercice d'arithmétique !})$$

PROPRIÉTÉ 4.18 - (Somme des racines de l'unité). Soit n un entier naturel supérieur ou égal à 2. On a :

$$\sum_{\omega \in \mathbb{U}_n} \omega = 0$$

Explicitement, la somme des racines n -ièmes de l'unité est nulle.

PREUVE. Soit n un entier naturel supérieur ou égal à 2. On a :

$$\sum_{\omega \in \mathbb{U}_n} \omega = \sum_{k=0}^{n-1} e^{2ik\pi/n} = \sum_{k=0}^{n-1} (e^{2i\pi/n})^k = \frac{1 - (e^{2i\pi/n})^n}{1 - e^{2i\pi/n}} = \frac{1 - 1}{1 - e^{2i\pi/n}} = 0. \text{ Ainsi : } \sum_{\omega \in \mathbb{U}_n} \omega = 0.$$

Remarque : il est légitime d'appliquer dans les calculs précédents la formule donnant la somme des termes d'une suite géométrique, puisque la raison (égale à $e^{2i\pi/n}$ dans le cas présent) est distincte de 1 (par hypothèse $n \geq 2$). ☕

Application 1 — Valeur exacte de $\cos\left(\frac{2\pi}{5}\right)$.

D'après la propriété précédente, on a : $\sum_{\omega \in \mathbb{U}_5} \omega = 0$. Explicitement : $\sum_{k=0}^4 e^{2ik\pi/5} = 0$. En particulier :

$$\operatorname{Re}\left(\sum_{k=0}^4 e^{2ik\pi/5}\right) = 0. \text{ D'où : } \sum_{k=0}^4 \cos\left(\frac{2k\pi}{5}\right) = 0.$$

$$\text{D'où : } 1 + \cos\left(\frac{2\pi}{5}\right) + \cos\left(\frac{4\pi}{5}\right) + \cos\left(\frac{6\pi}{5}\right) + \cos\left(\frac{8\pi}{5}\right) = 0$$

On peut alors observer que : $\cos\left(\frac{8\pi}{5}\right) = \cos\left(\frac{2\pi}{5}\right)$ et $\cos\left(\frac{6\pi}{5}\right) = \cos\left(\frac{4\pi}{5}\right)$ (par une double application de la formule : $\cos(2\pi - \theta) = \cos(\theta)$).

$$\text{Ainsi : } 1 + 2\cos\left(\frac{2\pi}{5}\right) + 2\cos\left(\frac{4\pi}{5}\right) = 0.$$

$$\text{Or, d'après la formule de duplication, on a : } \cos\left(\frac{4\pi}{5}\right) = 2\cos^2\left(\frac{2\pi}{5}\right) - 1.$$

$$\text{On a donc : } 4\cos^2\left(\frac{2\pi}{5}\right) + 2\cos\left(\frac{2\pi}{5}\right) - 1 = 0.$$

On en déduit que $\cos\left(\frac{2\pi}{5}\right)$ est solution de l'équation du second degré $4X^2 + 2X - 1 = 0$, qui possède exactement deux solutions réelles : $X_{+,-} = \frac{-1 \pm \sqrt{5}}{4}$.

Il reste à observer que $\cos\left(\frac{2\pi}{5}\right)$ est positif (puisque $2\pi/5$ est compris entre 0 et $\pi/2$) pour conclure que :

$$\cos\left(\frac{2\pi}{5}\right) = \frac{-1 + \sqrt{5}}{4}$$

Application 2 — Produit des racines de l'unité.

Soit n un entier naturel supérieur ou égal à 2. On a :

$$\prod_{\omega \in \mathbb{U}_n} \omega = \prod_{k=0}^{n-1} e^{2ik\pi/n} = e^{\sum_{k=0}^{n-1} (2ik\pi/n)} = e^{(2i\pi/n) \sum_{k=0}^{n-1} k} = e^{(2i\pi/n) \times (n(n-1)/2)} = e^{i\pi(n-1)} = (e^{i\pi})^{(n-1)} = (-1)^{(n-1)}.$$

Conclusion. $\prod_{\omega \in \mathbb{U}_n} \omega = (-1)^{n-1}.$

Application 3 — Résolution d'équation, et valeur exacte de $\tan\left(\frac{\pi}{5}\right)$.

On se propose de résoudre dans $\mathbb{C}$ l'équation (E) : $(1 + iz)^5 = (1 - iz)^5$

Première méthode : on développe, et on simplifie.

$$\begin{aligned} (E) &\iff 1 + 5iz - 10z^2 - 10iz^3 + 5z^4 + iz^5 = 1 - 5iz - 10z^2 + 10iz^3 + 5z^4 - iz^5 \\ &\iff 10iz - 20iz^3 + 2iz^5 = 0 \\ &\iff 5z - 10z^3 + z^5 = 0 \\ &\iff z(5 - 10z^2 + z^4) = 0 \\ &\iff [z = 0] \vee [z^4 - 10z^2 + 5 = 0] \end{aligned}$$

Reste donc à résoudre l'équation (E') : $z^4 - 10z^2 + 5 = 0$. Pour ce faire, on pose $Z = z^2$. L'équation devient alors $Z^2 - 10Z + 5 = 0$. Celle-ci possède deux solutions réelles et positives $5 \pm 2\sqrt{5}$. On en déduit que l'équation (E') possède quatre solutions réelles : $\pm\sqrt{5 \pm 2\sqrt{5}}$.

Conclusion : $\mathcal{S} = \left\{ 0; \sqrt{5 + 2\sqrt{5}}; \sqrt{5 - 2\sqrt{5}}; -\sqrt{5 + 2\sqrt{5}}; -\sqrt{5 - 2\sqrt{5}} \right\}.$

Deuxième méthode : on utilise les racines cinquièmes de l'unité.

$$\begin{aligned} (E) &\iff (1 + iz)^5 = (1 - iz)^5 \\ &\iff \left(\frac{1 + iz}{1 - iz} \right)^5 = 1 \\ &\iff \exists k \in \llbracket 0; 4 \rrbracket, \frac{1 + iz}{1 - iz} = e^{2ik\pi/5} \\ &\iff \exists k \in \llbracket 0; 4 \rrbracket, 1 + iz = e^{2ik\pi/5} (1 - iz) \\ &\iff \exists k \in \llbracket 0; 4 \rrbracket, i(1 + e^{2ik\pi/5})z = e^{2ik\pi/5} - 1 \\ &\iff \exists k \in \llbracket 0; 4 \rrbracket, iz = \frac{e^{2ik\pi/5} - 1}{e^{2ik\pi/5} + 1} \\ &\iff \exists k \in \llbracket 0; 4 \rrbracket, iz = \frac{e^{ik\pi/5} (e^{ik\pi/5} - e^{-ik\pi/5})}{e^{ik\pi/5} (e^{ik\pi/5} + e^{-ik\pi/5})} \\ &\iff \exists k \in \llbracket 0; 4 \rrbracket, iz = \frac{2i \sin(k\pi/5)}{2 \cos(k\pi/5)} \\ &\iff \exists k \in \llbracket 0; 4 \rrbracket, z = \tan(k\pi/5) \end{aligned}$$

Conclusion : $\mathcal{S} = \left\{ \tan\left(\frac{k\pi}{5}\right) / k \in \llbracket 0; 4 \rrbracket \right\}$

soit encore : $\mathcal{S} = \left\{ 0; \tan\left(\frac{\pi}{5}\right); \tan\left(\frac{2\pi}{5}\right); \tan\left(\frac{3\pi}{5}\right); \tan\left(\frac{4\pi}{5}\right) \right\}.$

Synthèse :

► En identifiant les ensembles de solutions trouvés par les deux méthodes, on obtient la valeur exacte de $\tan\left(\frac{\pi}{5}\right)$; on obtient également la valeur exacte de $\tan 0$, mais c'est un résultat moins spectaculaire.

Explicitement : $\tan\left(\frac{\pi}{5}\right) = \sqrt{5 - 2\sqrt{5}}$ (la plus petite des deux racines strictement positives de (E)) ; on

a également $\tan\left(\frac{2\pi}{5}\right) = \sqrt{5 + 2\sqrt{5}}$ (la plus grande des deux racines strictement positives de (E)).

Enfin : $\tan\left(\frac{4\pi}{5}\right) = -\sqrt{5 - 2\sqrt{5}}$ et $\tan\left(\frac{3\pi}{5}\right) = -\sqrt{5 + 2\sqrt{5}}$, ces deux valeurs se déduisant des précédentes par la relation $\tan(\pi - x) = -\tan(x)$.

Application 4 — Résolution d'équation (bis).

On considère l'équation (E) : $\bar{z} = z^n$, où n désigne un entier naturel supérieur ou égal à 2.

► Observons déjà que $z = 0$ est solution de (E) .

► Par la suite, on ne cherche plus que les solutions non nulles de (E) (en clair, on suppose $z \neq 0$).

On a alors : $(E) \iff z\bar{z} = z^{n+1} \iff |z|^2 = z^{n+1}$.

En prenant les modules de chaque terme de l'équation, on obtient : $|z|^2 = |z|^{n+1}$ et puisque l'on a supposé $n \geq 2$, cette condition implique : $|z| = 1$.

Par suite : $(E) \iff z^{n+1} = 1 \iff z \in \mathbb{U}_{n+1}$.

Conclusion : $\mathcal{S} = \{0\} \cup \mathbb{U}_{n+1}$ ou encore : $\mathcal{S} = \{0\} \cup \{e^{2ik\pi/(n+1)}, k \in \llbracket 0; n \rrbracket\}$

Application 5 — Factorisation de polynômes.

LEMME 4.4 - $\forall n \in \mathbb{N}^*, \forall (a, b) \in \mathbb{C}^2, a^{n+1} - b^{n+1} = (a - b) \sum_{k=0}^n a^k b^{n-k}$

PREUVE. Soit n un entier naturel supérieur non nul, et soient a et b deux complexes. On a :

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

COROLLAIRE 4.2 - Soit n un entier naturel supérieur ou égal à 2. On a :

$$X^n - 1 = (X - 1) \left(\sum_{k=0}^{n-1} X^k \right)$$

Autrement écrit : $X^n - 1 = (X - 1)(1 + X + \dots + X^{n-1})$.

PREUVE. Formellement, c'est une conséquence immédiate du lemme précédent. ☕

PROPRIÉTÉ 4.19 - Soit n un entier naturel supérieur ou égal à 2. On a :

$$1/ \quad X^n - 1 = \prod_{\omega \in \mathbb{U}_n} (X - \omega) = \prod_{k=0}^{n-1} (X - e^{2i k \pi / n})$$

$$2/ \quad \sum_{k=0}^{n-1} X^k = \prod_{\omega \in \mathbb{U}_n \setminus \{1\}} (X - \omega) = \prod_{k=1}^{n-1} (X - e^{2i k \pi / n})$$

PREUVE.¹⁵ Les racines n -ièmes de l'unité (*resp.* les éléments de $\mathbb{U}_n \setminus \{1\}$) sont racines des polynômes $X^n - 1$ et $\prod_{\omega \in \mathbb{U}_n} (X - \omega)$ (*resp.* $\sum_{k=0}^{n-1} X^k$ et $\prod_{k=1}^{n-1} (X - e^{2i k \pi / n})$).

Ces polynômes étant de degré n (*resp.* $n - 1$) et unitaires, ils sont égaux (principe du prolongement algébrique). ☕

Exemple 1 : le polynôme $1 + X + X^2$ a pour racines j et $\bar{j}$. On a donc : $X^2 + X + 1 = (X - j)(X - \bar{j})$.

Exemple 2 : le polynôme $1 + X + X^2 + X^3$ a pour racines -1 et $\pm i$. On a donc :

$$X^3 + X^2 + X + 1 = (X + 1)(X - i)(X + i).$$

4.8.2 Racines n -ièmes d'un complexe

DÉFINITION 4.12 - Soit $n \in \mathbb{N}$, $n \geq 2$, et soit z un complexe quelconque.

On appelle **racine n -ième de z** un nombre complexe ω tel que $\omega^n = z$.

Cette définition posée, il est tentant de faire le lien avec le paragraphe précédent.

THÉORÈME 4.6 - Tout nombre complexe z non nul possède exactement n racines n -ièmes, les nombres

$$|z|^{1/n} e^{i(\frac{\arg(z)+2k\pi}{n})} \quad \text{soit} \quad |z|^{1/n} e^{i\frac{\arg(z)}{n}} e^{i\frac{2k\pi}{n}}$$

avec k entier compris entre 0 et $n - 1$.

PREUVE. Soit $z \in \mathbb{C}^*$. On a : $z = \left(|z|^{1/n} e^{i\theta/n}\right)^n$, où l'on a posé : $\theta = \arg(z) [2\pi]$.

Soit Z un complexe arbitraire. On a :

$$Z^n = z \iff Z^n = \left(|z|^{1/n} e^{i\theta/n}\right)^n \iff \frac{Z}{|z|^{1/n} e^{i\theta/n}} \in \mathbb{U}_n \iff \exists k \in \llbracket 0, n-1 \rrbracket, \frac{Z}{|z|^{1/n} e^{i\theta/n}} = e^{2i k \pi / n}$$

Conclusion : Z est une racine n -ième de z SSI $\exists k \in \llbracket 0, n-1 \rrbracket, Z = |z|^{1/n} e^{i\theta/n} e^{2i k \pi / n}$.



Méthode (recherche des racines n -ièmes d'un complexe z non nul) : sauf cas particulier (où vous seriez guidés de toute façon), les racines n -ièmes d'un complexe non nul z ne peuvent être déterminées (pour $n > 2$) qu'en utilisant la forme exponentielle de z . En outre, pour les déterminer toutes, il suffit d'en trouver une que l'on multiplie ensuite par toutes les racines n -ièmes de l'unité.

¹⁵. A relire après le chapitre portant sur les polynômes.

Exemple : 2 est une racine quatrième de 16 dans $\mathbb{C}$. Les 4 racines quatrièmes de 16 dans $\mathbb{C}$ sont donc : $2, 2i, -2$ et $-2i$.

Il revient au même de dire que l'ensemble des racines quatrièmes de 16 est : $\{2i^k / k \in \llbracket 0, 3 \rrbracket\}$.

Application à la résolution de (E) : $(1+z)^4 = 16(1-z)^4$.

Soit z un nombre complexe. En observant que 1 n'est pas solution de l'équation (E), on peut écrire que :

$$[z \text{ est solution de } (E)]$$

$$\iff \frac{(1+z)^4}{(1-z)^4} = 16$$

$$\iff \left(\frac{1+z}{1-z}\right)^4 = 16$$

$$\iff \exists k \in \llbracket 0, 3 \rrbracket, \frac{1+z}{1-z} = 2i^k$$

$$\iff \exists k \in \llbracket 0, 3 \rrbracket, 1+z = 2i^k(1-z)$$

$$\iff \exists k \in \llbracket 0, 3 \rrbracket, z(2i^k + 1) = 2i^k - 1$$

$$\iff \exists k \in \llbracket 0, 3 \rrbracket, z = \frac{2i^k - 1}{2i^k + 1}$$

$$\iff \exists k \in \llbracket 0, 3 \rrbracket, z = \frac{4(-1)^k - 4i^k + 1}{4(-1)^k - 1}$$

Notons $Z_k = \frac{4(-1)^k - 4i^k + 1}{4(-1)^k - 1}$. On a : $Z_0 = \frac{1}{3}$; $Z_1 = \frac{3+4i}{5}$; $Z_2 = 3$ et $Z_3 = \frac{3-4i}{5}$.

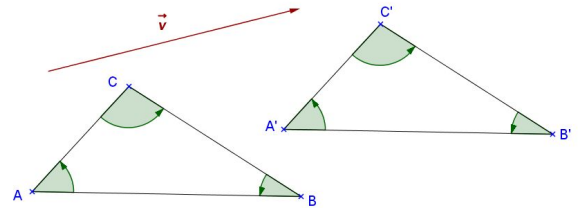
On déduit de ces calculs et de la chaîne d'équivalences précédente que :

$$[z \text{ est solution de } (E)] \iff \left[z \in \left\{ \frac{1}{3}; 3; \frac{3+4i}{5}; \frac{3-4i}{5} \right\} \right]$$

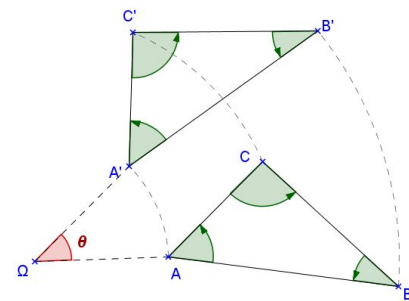
4.9 Similitudes directes

4.9.1 Transformations du plan

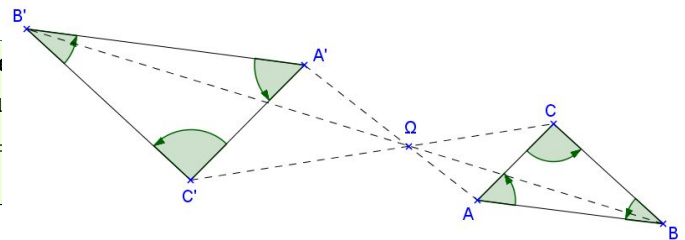
DÉFINITION 4.13 - On appelle **translation** de vecteur $\vec{v}$ la transformation du plan qui à tout point M associe l'unique point M' tel que $\overrightarrow{MM'} = \vec{v}$.



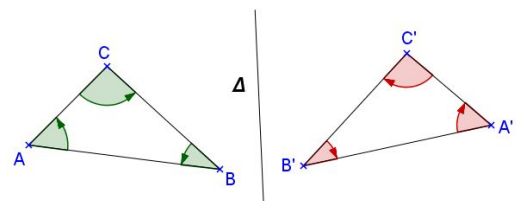
DÉFINITION 4.14 - On appelle **rotation de centre** Ω et d'angle θ la transformation du plan qui à tout point M associe l'unique point M' tel que $\Omega M' = \Omega M$ et $(\overrightarrow{\Omega M}, \overrightarrow{\Omega M'}) = \theta$.



DÉFINITION 4.15 - On appelle **homothétie de centre** Ω et de rapport $k \in \mathbb{R}^*$ la transformation du plan qui à tout point M associe l'unique point M' tel que $\overrightarrow{\Omega M'} = k\overrightarrow{\Omega M}$.



DÉFINITION 4.16 - On appelle **réflexion (ou symétrie axiale)** d'axe Δ la transformation du plan qui à tout point M associe l'unique point M' tel que Δ soit la médiatrice de $[MM']$ si $M \notin \Delta$, et M si $M \in \Delta$.



4.9.2 Ecritures complexes

- $z \mapsto z + b$ est une écriture complexe de la translation de vecteur $\vec{v}(b)$ (avec $b \in \mathbb{C}$);
- $z \mapsto \bar{z}$ est une écriture complexe de la réflexion d'axe réel;
- $z \mapsto k(z - \omega) + \omega$ est une écriture complexe de l'homothétie de centre $\Omega(\omega)$, et de rapport k (avec $k \in \mathbb{R}^*$ et $\omega \in \mathbb{C}$);
- $z \mapsto e^{i\theta}(z - \omega) + \omega$ est une écriture complexe de la rotation de centre $\Omega(\omega)$, et d'angle θ (avec $\theta \in \mathbb{R}$ et $\omega \in \mathbb{C}$);
- $z \mapsto az + b$ (avec $a \in \mathbb{C}^*$ et $b \in \mathbb{C}$) est une écriture complexe de ce que l'on appelle une **similitude directe**.

Remarque : les translations, rotations et homothéties sont en particulier des similitudes directes.

4.9.3 Propriétés des similitudes directes

PROPRIÉTÉ 4.20 - La composée de deux similitudes directes est une similitude directe.

PREUVE. Soient S et S' deux similitudes directes d'écritures complexes respectives $z \mapsto az + b$ et $z \mapsto a'z + b'$. Leur composée $S \circ S'$ a pour écriture complexe $z \mapsto aa'z + ab' + b$. ☕

PROPRIÉTÉ 4.21 - L'identité du plan est une similitude directe, qui est l'élément neutre pour la composition.

PREUVE. Triviale. ☕

PROPRIÉTÉ 4.22 - Pour toute similitude directe S , il existe une unique similitude directe S' telle que $S \circ S' = S' \circ S = \text{id}$.

PREUVE. Soient S une similitude directe d'écriture complexe $z \mapsto az + b$. Un calcul sans difficulté assure que l'unique similitude directe S' de l'énoncé est celle d'écriture complexe $z \mapsto a^{-1}z - a^{-1}b$. ☕

PROPRIÉTÉ 4.23 - Les similitudes directes conservent les angles orientés et les rapports de longueurs.

PREUVE. Explicitement, il s'agit de prouver ceci : soit S une similitude directe, et soient M, N, P et Q quatre points du plan tels que $M \neq N$ et $P \neq Q$. On note M', N', P' et Q' les images respectives de ces points par S . Alors :

$$\left(\overrightarrow{M'N'}, \overrightarrow{P'Q'} \right) = \left(\overrightarrow{MN}, \overrightarrow{PQ} \right) \quad \text{et} \quad \frac{P'Q'}{M'N'} = \frac{PQ}{MN}$$

On se place donc dans les hypothèses des propriétés, et on convient de noter $m, m', n, n', \dots$ les affixes respectives des points $M, M', N, N', \dots$. On note par ailleurs $z' = az + b$ (avec $(a, b) \in \mathbb{C}^* \times \mathbb{C}$) une écriture complexe de la similitude S .

Avec ces notations, on a :

$$\begin{aligned} \left(\overrightarrow{M'N'}, \overrightarrow{P'Q'} \right) &= \arg \left(\frac{q' - p'}{n' - m'} \right) = \arg \left(\frac{aq + b - (ap + b)}{an + b - (am + b)} \right) = \arg \left(\frac{a(q - p)}{a(n - m)} \right) \\ &= \arg \left(\frac{q - p}{n - m} \right) = \left(\overrightarrow{MN}, \overrightarrow{PQ} \right) \end{aligned} \quad 16$$

Par ailleurs :

$$\frac{P'Q'}{M'N'} = \left| \frac{q' - p'}{n' - m'} \right| = \left| \frac{aq + b - (ap + b)}{an + b - (am + b)} \right| = \left| \frac{a(q - p)}{a(n - m)} \right| = \left| \frac{q - p}{n - m} \right| = \frac{PQ}{MN}$$

Conclusion : $\left(\overrightarrow{M'N'}, \overrightarrow{P'Q'} \right) = \left(\overrightarrow{MN}, \overrightarrow{PQ} \right)$ et $\frac{P'Q'}{M'N'} = \frac{PQ}{MN}$. Ce qui prouve que toute similitude directe conserve les angles orientés et les rapports de longueurs. ☕

16. Toutes ces égalités étant à lire *modulo* 2π .

Plus précisément, il ressort de la preuve précédente qu'une similitude directe "multiplie les longueurs par $|a|$ ". D'où le :

COROLLAIRE 4.3 - la similitude directe d'écriture complexe $z \mapsto az + b$ est une isométrie si et seulement si $|a| = 1$.

PREUVE. Conséquence immédiate de la preuve de la propriété précédente. ☕

THÉORÈME 4.7 - (Classification des similitudes directes). Soit S une similitude directe, d'écriture complexe $z \mapsto az + b$ (avec a et b complexes, $a \neq 0$).

- Si $a = 1$, alors S est une translation (de vecteur d'affixe b).
- Si $a \neq 1$, alors S admet un unique point invariant, noté Ω ($\omega = b/(1-a)$). Une écriture complexe de S est alors $z \mapsto a(z - \omega) + \omega$; en écrivant $a = |a|e^{i\theta}$, S est la composée d'une homothétie h de rapport $|a|$ et d'une rotation R d'angle θ , toutes deux de centre Ω (et dans ce cas particulier, $h \circ R = R \circ h$).

PREUVE. Soit S une similitude directe, d'écriture complexe $z \mapsto az + b$ (avec a et b complexes, $a \neq 0$).

➤ Si $a = 1$. Alors S a pour écriture complexe $z \mapsto z + b$, qui est évidemment l'écriture d'une translation.

➤ Si $a \neq 1$. **Commençons par établir que S admet un unique point fixe (ou invariant).** Le point $\Omega(\omega)$ est un point fixe de S si $S(\Omega) = \Omega$, c'est-à-dire si : $a\omega + b = \omega$. Or :

$$a\omega + b = \omega \iff (1-a)\omega = b \iff \omega = \frac{b}{1-a}$$

(la division par $1-a$ étant légitime puisque $a \neq 1$).

Ainsi : S admet un unique point fixe d'affixe $\omega = b/(1-a)$.

Prouvons que S admet pour écriture complexe : $z \mapsto a(z - \omega) + \omega$ (avec $\omega = b/(1-a)$). Soit $z \in \mathbb{C}$ arbitraire. On a :

$$a(z - \omega) + \omega = a\left(z - \frac{b}{1-a}\right) + \frac{b}{1-a} = az - \frac{ab}{1-a} + \frac{b}{1-a} = az + \frac{(1-a)b}{1-a} = az + b$$

On a établi que : $\forall z \in \mathbb{C}, a(z - \omega) + \omega = az + b$.

Ce qui prouve que S admet pour écriture complexe $z \mapsto a(z - \omega) + \omega$ (toujours avec $\omega = b/(1-a)$).

Pour conclure, on note $a = |a|e^{i\theta}$. On définit h et R comme les transformations d'écritures complexes respectives : $z \mapsto |a|(z - \omega) + \omega$ et $z \mapsto e^{i\theta}(z - \omega) + \omega$ (h est une homothétie et R une rotation). Muni de ces définitions, on vérifie par un calcul aisé (ci-dessous) que $(h \circ R)(z) = S(z)$ (et l'égalité $(R \circ h)(z) = S(z)$ pourra être admise), ce qui achève la preuve.

$$\forall z \in \mathbb{C}, (h \circ R)(z) = h(R(z)) = h(e^{i\theta}(z - \omega) + \omega) = |a|[(e^{i\theta}(z - \omega) + \omega) - \omega] + \omega = \underbrace{|a|e^{i\theta}}_{=a}(z - \omega) + \omega$$

$\omega = S(z)$. ☕

Remarque pratique : le théorème fournit une méthode pratique pour déterminer la nature d'une similitude S d'écriture $z \mapsto az + b$. Il suffit en effet de regarder a "dans les yeux" pour y parvenir. Si vous avez beaucoup de chance, $a = 1$ et S est une translation. Si $a \neq 1$, mais que $a \in \mathbb{R}$, S est alors une homothétie de rapport a . Si $a \in \mathbb{U}$, alors S est une rotation d'angle $\arg(a)$. Enfin, dans le cas plus général ($a \notin \mathbb{U}$), S est la composée d'une rotation et d'une homothétie, comme indiqué dans le théorème.

Pour conclure, l'affixe ω du centre de S (c'est-à-dire de l'unique point invariant ou fixe, qui existe dès lors que S n'est pas une translation) est solution de l'équation $\omega = a\omega + b$.

COROLLAIRE 4.4 - Deux similitudes directes de même centre commutent.

PREUVE. Soient S et S' deux similitudes directes, qui ne sont pas des translations. On suppose que S et S' ont même centre. Il existe donc a, a' et ω trois complexes, avec a et a' non nuls, tels que S et S' admettent pour écritures complexes respectives : $z' = a(z - \omega) + \omega$ et $z' = a'(z - \omega) + \omega$.

Alors : $\forall z \in \mathbb{C}, (S' \circ S)(z) = S'(S(z)) = S'(a(z - \omega) + \omega) = a'a(z - \omega) + \omega$

Et : $\forall z \in \mathbb{C}, (S \circ S')(z) = S(S'(z)) = S(a'(z - \omega) + \omega) = aa'(z - \omega) + \omega$

D'où : $\forall z \in \mathbb{C}, (S' \circ S)(z) = (S \circ S')(z)$. Il revient au même d'écrire que : $S' \circ S = S \circ S'$.

Conclusion. S et S' deux similitudes directes de même centre commutent. ☕

4.9.4 Le groupe des similitudes directes

Bilan : notons $\text{Sim}^+(\mathbb{C})$ l'ensemble des similitudes directes : $\text{Sim}^+(\mathbb{C})$ possède les propriétés suivantes.

Prop 1 : $\forall (T, T') \in (\text{Sim}^+(\mathbb{C}))^2, (T' \circ T) \in \text{Sim}^+(\mathbb{C})$. La composée de deux similitudes directes est une similitude directe. On dit que la composition est une *loi de composition interne (lci)* pour $\text{Sim}^+(\mathbb{C})$.

Prop 2 : $\forall T \in \text{Sim}^+(\mathbb{C}), (T \circ \text{id}_{\mathbb{C}}) = T = (\text{id}_{\mathbb{C}} \circ T)$. L'identité du plan complexe $\text{id}_{\mathbb{C}}$ est l'*élément neutre* pour la composition des applications.

Prop 3 : $\forall (T, T', T'') \in (\text{Sim}^+(\mathbb{C}))^3, T'' \circ (T' \circ T) = (T'' \circ T') \circ T$. La composition dans $\text{Sim}^+(\mathbb{C})$ est *associative*.

Prop 4 : $\forall T \in \text{Sim}^+(\mathbb{C}), \exists ! T' \in \text{Sim}^+(\mathbb{C}), (T' \circ T) = \text{id}_{\mathbb{C}} = (T \circ T')$. Tout élément T de $\text{Sim}^+(\mathbb{C})$ admet un *inverse* (unique) pour la composition.

Dans ce contexte, on dit que $(\text{Sim}^+(\mathbb{C}), \circ)$ est un **groupe**. En outre, ce groupe est dit **non commutatif** (ou **non abélien**) puisque nous avons observé que la composition des similitudes n'est pas commutative.

4.10 Epilogue

4.10.1 Constructibilité à la règle et au compas des pentagone et heptakaïdécagone réguliers

L'objectif de ce problème est de déterminer les valeurs exactes de $\cos\left(\frac{\pi}{5}\right)$ de $\cos\left(\frac{\pi}{17}\right)$.

Plus précisément, il s'agit de montrer que ces cosinus peuvent s'exprimer **par radicaux carrés**, c'est-à-dire (grossièrement) que l'on peut en trouver des expressions ne faisant intervenir que des combinaisons finies de quotients et de racines carrées.

A titre d'exemples, $\sin\left(\frac{\pi}{3}\right) = \frac{\sqrt{3}}{2}$, $\cos\left(\frac{\pi}{8}\right) = \frac{\sqrt{2+\sqrt{2}}}{2}$ et $e^{-i\pi/4} = \frac{\sqrt{2}}{2} - i\frac{\sqrt{2}}{2}$ sont des expressions par radicaux carrés.

Première partie — Calcul de $\cos\left(\frac{\pi}{5}\right)$

Le calcul de la valeur exacte de $\cos\left(\frac{\pi}{5}\right)$ a été récemment fait en cours, comme application de la propriété affirmant que la somme des racines cinquièmes de l'unité est nulle. On propose dans cette partie une méthode alternative pour déterminer cette valeur exacte. A cette fin, on considère l'équation

$$(E) : \quad z^4 + z^3 + z^2 + z + 1 = 0$$

D'après le cours, les racines de cette équation sont toutes les racines cinquièmes de l'unité sauf 1, c'est à dire explicitement les nombres complexes $e^{2i\pi/5}$, $e^{4i\pi/5}$, $e^{6i\pi/5}$ et $e^{8i\pi/5}$.

On pose : $\forall z \in \mathbb{C}$, $Q(z) = z^4 + z^3 + z^2 + z + 1$, de telle sorte que l'équation (E) s'écrit : $Q(z) = 0$.

1) Déterminer trois réels a , b et c tels que :

$$\forall z \in \mathbb{C}^*, \quad \frac{Q(z)}{z^2} = a \left(z + \frac{1}{z}\right)^2 + b \left(z + \frac{1}{z}\right) + c$$

2) Résoudre dans $\mathbb{C}$ l'équation : $aZ^2 + bZ + c = 0$ (on tâchera d'exprimer les solutions par radicaux carrés).

3) Résoudre dans $\mathbb{C}$ l'équation : $Q(z) = 0$.

4) Dédire des questions précédentes les valeurs exactes de : $\cos\left(\frac{2\pi}{5}\right)$, $\cos\left(\frac{4\pi}{5}\right)$ et $\cos\left(\frac{\pi}{5}\right)$.

(au regard de l'objectif du problème, on tâchera d'exprimer ces valeurs par radicaux carrés).

Deuxième partie — Calcul de $\cos\left(\frac{\pi}{17}\right)$

5) Jusqu'à la fin du problème, on pose $\theta = \frac{\pi}{17}$.

En outre, on pose :

$$x_1 = \cos(3\theta) + \cos(5\theta) + \cos(7\theta) + \cos(11\theta)$$

$$x_2 = \cos(\theta) + \cos(9\theta) + \cos(13\theta) + \cos(15\theta)$$

a) Montrer que $x_1 > 0$.

b) Calculer $x_1 + x_2$ en s'aidant de la propriété énoncée dans l'exercice de la section suivante (*indication* : le résultat est un rationnel "simple").

c) Calculer le produit $x_1 x_2$ (*indication*¹⁷ : on établira que $x_1 x_2 = -2(x_1 + x_2)$).

d) Dédire des questions précédentes les valeurs exactes de x_1 et x_2 (que l'on tâchera d'exprimer par radicaux carrés).

17. Plus que jamais, l'indication donnée ici a pour objectif de vous permettre d'aller plus loin dans le sujet (pour peu que vous ayez traité la question précédente). En clair, il n'y a pas de méthode expéditive pour obtenir la relation " $x_1 x_2 = -2(x_1 + x_2)$ "; il faut de la patience et de l'astuce pour parvenir à l'établir.

6) On pose à présent :

$$y_1 = \cos(3\theta) + \cos(5\theta)$$

$$y_2 = \cos(7\theta) + \cos(11\theta)$$

$$y_3 = \cos(\theta) + \cos(13\theta)$$

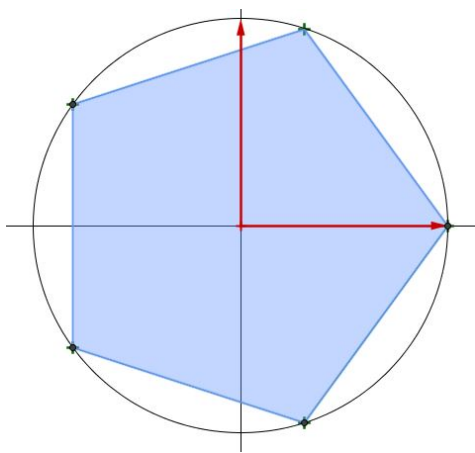
$$y_4 = \cos(9\theta) + \cos(15\theta)$$

- a) En s'inspirant de la question précédente, calculer les produits $y_1 y_2$ et $y_3 y_4$.
- b) En déduire les valeurs exactes de y_1, y_2, y_3 et y_4 (que l'on tâchera d'exprimer par radicaux carrés).
- c) Calculer le produit $P = \cos(\theta) \cos(13\theta)$ et en déduire une méthode qui permette d'obtenir l'expression de $\cos(\theta)$ par radicaux carrés.

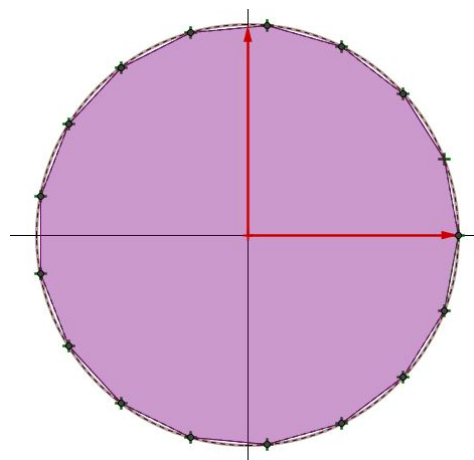
Remarque : à l'issue de la dernière question du problème, on obtient une équation dont la résolution (que je ne vous demande pas de faire ici¹⁸) donne

$$\cos\left(\frac{\pi}{17}\right) = \frac{1 - \sqrt{17} + \sqrt{34 - 2\sqrt{17}} + \sqrt{68 + 12\sqrt{17} + 2\sqrt{680 + 152\sqrt{17}}}}{16}$$

On peut déduire de ce résultat la valeur exacte de $\cos\left(\frac{2\pi}{17}\right)$. Il “suffit” en effet d'utiliser la formule de duplication pour le cosinus... Puis, à l'aide de cette valeur exacte, on peut construire très précisément (à l'aide d'une règle non-graduée et d'un compas) un polygone régulier à dix-sept cotés. Pour des raisons algébriques assez profondes, une telle construction n'est pas possible pour un polygone régulier à n côtés avec n arbitraire (c'est impossible par exemple pour $n = 13$ ou $n = 14$ ou $n = 19$).



PENTAGONE RÉGULIER



HEPTAKAÏDÉCAGONE RÉGULIER

18. C'est gentil, n'est-ce pas ?

CORRIGÉ.

Première partie — Calcul de $\cos\left(\frac{\pi}{5}\right)$

1) D'après l'énoncé : $\forall z \in \mathbb{C}^*, \frac{Q(z)}{z^2} = \frac{1}{z^2} + \frac{1}{z} + 1 + z + z^2 \quad (\spadesuit).$

D'autre part : $\forall z \in \mathbb{C}^*, a\left(z + \frac{1}{z}\right)^2 + b\left(z + \frac{1}{z}\right) + c = az^2 + 2a + \frac{a}{z^2} + bz + \frac{b}{z} + c$

$$\iff \forall z \in \mathbb{C}^*, a\left(z + \frac{1}{z}\right)^2 + b\left(z + \frac{1}{z}\right) + c = \frac{a}{z^2} + \frac{b}{z} + 2a + c + bz + az^2 \quad (\clubsuit)$$

En identifiant les deux expressions obtenues dans $(\spadesuit)$ et $(\clubsuit)$, on a : $\begin{cases} a = 1 \\ b = 1 \\ 2a + c = 1 \end{cases}$ d'où : $\begin{cases} a = 1 \\ b = 1 \\ c = -1 \end{cases}$

Conclusion : $\forall z \in \mathbb{C}^*, \frac{Q(z)}{z^2} = \left(z + \frac{1}{z}\right)^2 + \left(z + \frac{1}{z}\right) - 1$

2) L'équation $Z^2 + Z - 1 = 0$ a pour discriminant $\Delta = 5$. Elle a donc deux racines réelles : $\frac{-1 \pm \sqrt{5}}{2}$.

3) En opérant le changement de variable $Z = z + \frac{1}{z}$, l'équation $Q(z) = 0$ se réécrit : $Z^2 + Z - 1 = 0$. En d'autres termes :

$$(z \text{ est solution de } Q(z) = 0) \iff (Z = z + \frac{1}{z} \text{ est solution de } Z^2 + Z - 1 = 0)$$

On déduit de la question précédente que : z est solution de $Q(z) = 0$ si et seulement si $z + \frac{1}{z} = Z_1$ ou $z + \frac{1}{z} = Z_2$ (en ayant noté $Z_1 = \frac{-1 + \sqrt{5}}{2}$ et $Z_2 = \frac{-1 - \sqrt{5}}{2}$). Reste à résoudre ces deux équations.

► Réolvons (E1) : $z + \frac{1}{z} = Z_1$.

$$z + \frac{1}{z} = Z_1 \iff z^2 + 1 = Z_1 z \iff z^2 - Z_1 z + 1 = 0.$$

Le discriminant est $\Delta = \left(\frac{-1 + \sqrt{5}}{2}\right)^2 - 4 = \frac{6 - 2\sqrt{5}}{4} - 4 = \frac{-10 - 2\sqrt{5}}{4} = -\frac{5 + \sqrt{5}}{2}$. Comme Δ est ici un réel strictement négatif (et que l'équation est à coefficients réels), on en déduit directement que (E1) a deux solutions complexes conjuguées :

$$\frac{\frac{-1 + \sqrt{5}}{2} \pm i\sqrt{\frac{5 + \sqrt{5}}{2}}}{2} \text{ c'est-à-dire : } \frac{-1 + \sqrt{5}}{4} \pm \frac{i}{2}\sqrt{\frac{5 + \sqrt{5}}{2}}$$

► Réolvons (E2) : $z + \frac{1}{z} = Z_2$.

$$z + \frac{1}{z} = Z_2 \iff z^2 + 1 = Z_2 z \iff z^2 - Z_2 z + 1 = 0.$$

Le discriminant est $\Delta = \left(\frac{-1 - \sqrt{5}}{2}\right)^2 - 4 = \frac{6 + 2\sqrt{5}}{4} - 4 = \frac{-10 + 2\sqrt{5}}{4} = -\frac{5 - \sqrt{5}}{2}$. Comme précédemment, on en déduit que (E2) a deux solutions complexes conjuguées :

$$\frac{-\frac{1 + \sqrt{5}}{2} \pm i\sqrt{\frac{5 - \sqrt{5}}{2}}}{2} \text{ c'est-à-dire : } -\frac{1 + \sqrt{5}}{4} \pm \frac{i}{2}\sqrt{\frac{5 - \sqrt{5}}{2}}$$

Conclusion : l'équation $Q(z) = 0$ a exactement quatre solutions dans $\mathbb{C}$:

$$\frac{-1 + \sqrt{5}}{4} \pm \frac{i}{2}\sqrt{\frac{5 + \sqrt{5}}{2}} \text{ et } -\frac{1 + \sqrt{5}}{4} \pm \frac{i}{2}\sqrt{\frac{5 - \sqrt{5}}{2}}$$

4) On sait (d'après l'énoncé... et le cours) que les racines du polynôme Q sont donc les racines cinquièmes de l'unité distinctes de 1. Explicitement les racines de Q sont : $e^{2i\pi/5}$, $e^{4i\pi/5}$, $e^{6i\pi/5}$ et $e^{8i\pi/5}$. En observant que les complexes $e^{2i\pi/5}$ et $e^{8i\pi/5}$ (*resp.* $e^{4i\pi/5}$ et $e^{6i\pi/5}$) sont conjugués, on peut affirmer qu'ils ont la même partie réelle, et que cette partie réelle est positive puisque $2\pi/5 \in [0; \pi/2]$ (*resp.* cette partie réelle est négative puisque $4\pi/5 \in [\pi/2; \pi]$). Par suite, en comparant avec les formes algébriques des racines de Q obtenues à la question précédente, on a :

$$\cos\left(\frac{2\pi}{5}\right) = \frac{-1 + \sqrt{5}}{4} \text{ et } \cos\left(\frac{4\pi}{5}\right) = -\frac{1 + \sqrt{5}}{4}$$

Il reste à noter que $\cos\left(\frac{\pi}{5}\right) = -\cos\left(\frac{4\pi}{5}\right)$ ¹⁹ pour obtenir : $\cos\left(\frac{\pi}{5}\right) = \frac{1 + \sqrt{5}}{4}$

Deuxième partie — Calcul de $\cos\left(\frac{\pi}{17}\right)$

5) Jusqu'à la fin du problème, on pose $\theta = \frac{\pi}{17}$.

a) Pour commencer, indiquons que les réels $\cos(3\theta)$, $\cos(5\theta)$ et $\cos(7\theta)$ sont positifs (puisque $3\pi/17$, $5\pi/17$ et $7\pi/17$ appartiennent à $[0; \pi/2]$).

D'autre part : $\cos(11\theta) = \cos(11\pi/17) = -\cos(6\pi/17)$. Or $\cos(6\pi/17) < \cos(5\pi/17)$ (puisque la fonction $\cos$ est strictement décroissante sur $[0; \pi/2]$).

Donc : $\cos(5\pi/17) - \cos(6\pi/17) > 0$ d'où $\cos(5\pi/17) + \cos(11\pi/17) > 0$.

A fortiori : $\cos(3\pi/17) + \cos(7\pi/17) + \cos(5\pi/17) + \cos(11\pi/17) > 0$, c'est-à-dire : $x_1 > 0$.

b) On a : $x_1 + x_2 = \cos(3\theta) + \cos(5\theta) + \cos(7\theta) + \cos(11\theta) + \cos(\theta) + \cos(9\theta) + \cos(13\theta) + \cos(15\theta)$
 $= \cos(\theta) + \cos(3\theta) + \cos(5\theta) + \cos(7\theta) + \cos(9\theta) + \cos(11\theta) + \cos(13\theta) + \cos(15\theta) = \sum_{k=0}^7 \cos(\theta + 2k\theta)$

19. Via $\cos(\pi - \theta) = -\cos(\theta)$.

Ainsi, avec les notations de l'exercice de la section suivante, on a : $x_1 + x_2 = C(\theta, 2\theta)$ avec $n = 8$. D'après cette même question 4, on a donc :

$$x_1 + x_2 = \frac{\sin(8\theta) \cos(8\theta)}{\sin(\theta)} = \frac{1}{2} \frac{\sin(16\theta)}{\sin(\theta)} = \frac{1}{2}^{20} \quad \boxed{\text{Conclusion : } x_1 + x_2 = \frac{1}{2}}.$$

c) Accrochons nos ceintures :

$$\begin{aligned} x_1 x_2 &= [\cos(3\theta) + \cos(5\theta) + \cos(7\theta) + \cos(11\theta)] \times [\cos(\theta) + \cos(9\theta) + \cos(13\theta) + \cos(15\theta)] \\ \Leftrightarrow x_1 x_2 &= \cos(3\theta) \cos(\theta) + \cos(3\theta) \cos(9\theta) + \cos(3\theta) \cos(13\theta) + \cos(3\theta) \cos(15\theta) + \cos(5\theta) \cos(\theta) + \\ &\cos(5\theta) \cos(9\theta) + \cos(5\theta) \cos(13\theta) + \cos(5\theta) \cos(15\theta) + \cos(7\theta) \cos(\theta) + \cos(7\theta) \cos(9\theta) + \cos(7\theta) \cos(13\theta) + \\ &\cos(7\theta) \cos(15\theta) + \cos(11\theta) \cos(\theta) + \cos(11\theta) \cos(9\theta) + \cos(11\theta) \cos(13\theta) + \cos(11\theta) \cos(15\theta) \end{aligned}$$

Or cette expression peut être linéarisée²¹ en appliquant 16 fois la formule :

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

En avant :

$$\begin{aligned} x_1 x_2 &= \frac{1}{2} (\cos(2\theta) + \cos(4\theta) + \cos(6\theta) + \cos(12\theta) + \cos(10\theta) + \cos(16\theta) + \cos(12\theta) + \cos(18\theta) + \cos(4\theta) + \\ &\cos(6\theta) + \cos(4\theta) + \cos(14\theta) + \cos(8\theta) + \cos(18\theta) + \cos(10\theta) + \cos(20\theta) + \cos(6\theta) + \cos(8\theta) + \cos(2\theta) + \\ &\cos(16\theta) + \cos(6\theta) + \cos(20\theta) + \cos(8\theta) + \cos(22\theta) + \cos(10\theta) + \cos(12\theta) + \cos(20\theta) + \cos(2\theta) + \cos(24\theta) + \\ &\cos(2\theta) + \cos(4\theta) + \cos(26\theta)) \end{aligned}$$

$$\begin{aligned} \Leftrightarrow x_1 x_2 &= \frac{1}{2} (4 \cos(2\theta) + 4 \cos(4\theta) + 4 \cos(6\theta) + 3 \cos(8\theta) + 3 \cos(10\theta) + 3 \cos(12\theta) + \cos(14\theta) + 2 \cos(16\theta) \\ &+ 2 \cos(18\theta) + 3 \cos(20\theta) + \cos(22\theta) + \cos(24\theta) + \cos(26\theta)) \end{aligned}$$

Maintenant que le plus dur est fait²², on peut utiliser la relation $\cos(2\pi - x) = \cos(x)$ pour avancer que : $\cos(18\theta) = \cos(16\theta)$; $\cos(20\theta) = \cos(14\theta)$; $\cos(22\theta) = \cos(12\theta)$; $\cos(24\theta) = \cos(10\theta)$ et $\cos(26\theta) = \cos(8\theta)$.

On peut donc réécrire la relation précédente :

$$\begin{aligned} x_1 x_2 &= \frac{1}{2} (4 \cos(2\theta) + 4 \cos(4\theta) + 4 \cos(6\theta) + 4 \cos(8\theta) + 4 \cos(10\theta) + 4 \cos(12\theta) + 4 \cos(14\theta) + 4 \cos(16\theta)) \\ \text{soit encore : } x_1 x_2 &= 2 (\cos(2\theta) + \cos(4\theta) + \cos(6\theta) + \cos(8\theta) + \cos(10\theta) + \cos(12\theta) + \cos(14\theta) + \cos(16\theta)) \end{aligned}$$

Une octuple application de la formule $\cos(\pi - x) = -\cos(x)$ nous autorise à réécrire la relation précédente :

$$x_1 x_2 = -2 (\cos(15\theta) + \cos(13\theta) + \cos(11\theta) + \cos(9\theta) + \cos(7\theta) + \cos(5\theta) + \cos(3\theta) + \cos(\theta))$$

C'est-à-dire : $x_1 x_2 = -2(x_1 + x_2)$. Ainsi, d'après la question précédente : $\boxed{x_1 x_2 = -1}$.

d) Rappelons que deux complexes (a fortiori deux réels) sont racines de l'équation $X^2 - SX + P = 0$, où S et P désignent les somme et produit de ces deux nombres. Puisque $x_1 + x_2 = \frac{1}{2}$ et $x_1 x_2 = -1$, les réels

x_1 et x_2 sont racines de l'équation (E3) : $X^2 - \frac{X}{2} - 1 = 0$.

20. La dernière simplification venant de $\sin(16\pi/17) = \sin(\pi/17)$.

21. Dans l'allégresse.

22. Dans cette question en tout cas.

La résolution (aisée) de cette équation fournit deux solutions : $\frac{1 \pm \sqrt{17}}{4}$. Ces deux racines sont de signes contraires. Et comme on a établi que $x_1 > 0$ (question 5-a), on peut en déduire que :

$$x_1 = \frac{1 + \sqrt{17}}{4} \text{ et } x_2 = \frac{1 - \sqrt{17}}{4}$$

$$\textbf{Conclusion : } \cos\left(\frac{3\pi}{17}\right) + \cos\left(\frac{5\pi}{17}\right) + \cos\left(\frac{7\pi}{17}\right) + \cos\left(\frac{11\pi}{17}\right) = \frac{1 + \sqrt{17}}{4}$$

$$\text{et } \cos\left(\frac{\pi}{17}\right) + \cos\left(\frac{9\pi}{17}\right) + \cos\left(\frac{13\pi}{17}\right) + \cos\left(\frac{15\pi}{17}\right) = \frac{1 - \sqrt{17}}{4}$$

$$6) \text{ a) On a : } y_1 y_2 = (\cos(3\theta) + \cos(5\theta))(\cos(7\theta) + \cos(11\theta))$$

$$\iff y_1 y_2 = \cos(3\theta)\cos(7\theta) + \cos(3\theta)\cos(11\theta) + \cos(5\theta)\cos(7\theta) + \cos(5\theta)\cos(11\theta)$$

$$\iff 2y_1 y_2 = \cos(10\theta) + \cos(4\theta) + \cos(14\theta) + \cos(8\theta) + \cos(2\theta) + \cos(12\theta) + \cos(16\theta) + \cos(6\theta)$$

$$\iff 2y_1 y_2 = -(\cos(7\theta) + \cos(13\theta) + \cos(3\theta) + \cos(9\theta) + \cos(15\theta) + \cos(5\theta) + \cos(\theta) + \cos(11\theta))$$

$$\iff 2y_1 y_2 = -(x_1 + x_2) \iff 2y_1 y_2 = -\frac{1}{2} \iff \boxed{y_1 y_2 = -\frac{1}{4}}$$

$$\text{On a : } y_1 y_2 = (\cos(3\theta) + \cos(5\theta))(\cos(7\theta) + \cos(11\theta))$$

D'autre part :

$$y_3 y_4 = (\cos(\theta) + \cos(13\theta))(\cos(9\theta) + \cos(15\theta))$$

$$\iff y_3 y_4 = \cos(\theta)\cos(9\theta) + \cos(\theta)\cos(15\theta) + \cos(13\theta)\cos(9\theta) + \cos(13\theta)\cos(15\theta)$$

$$\iff 2y_3 y_4 = \cos(10\theta) + \cos(8\theta) + \cos(16\theta) + \cos(14\theta) + \cos(22\theta) + \cos(4\theta) + \cos(28\theta) + \cos(2\theta)$$

Or $\cos(22\theta) = \cos(12\theta)$ et $\cos(28\theta) = \cos(6\theta)$ (car $\cos(\pi + x) = \cos(\pi - x)$) d'où :

$$\iff 2y_3 y_4 = \cos(10\theta) + \cos(8\theta) + \cos(16\theta) + \cos(14\theta) + \cos(12\theta) + \cos(4\theta) + \cos(6\theta) + \cos(2\theta)$$

$$\iff 2y_3 y_4 = -(\cos(7\theta) + \cos(9\theta) + \cos(\theta) + \cos(3\theta) + \cos(5\theta) + \cos(13\theta) + \cos(11\theta) + \cos(15\theta))$$

$$\iff 2y_3 y_4 = -(x_1 + x_2) \iff 2y_3 y_4 = -\frac{1}{2} \iff \boxed{y_3 y_4 = -\frac{1}{4}}$$

b) On a prouvé plus haut que : $y_1 y_2 = -1/4$. Par ailleurs, il résulte de la définition de y_1 et y_2 que : $y_1 + y_2 = x_1$. Par conséquent y_1 et y_2 sont les racines de l'équation²³ :

$$(E4) \quad X^2 - x_1 X - \frac{1}{4} = 0 \text{ c'est-à-dire : } (E4) \quad X^2 - \left(\frac{1 + \sqrt{17}}{4}\right) X - \frac{1}{4} = 0$$

Réolvons (E4) : $X^2 - \left(\frac{1 + \sqrt{17}}{4}\right) X - \frac{1}{4} = 0 \iff 4X^2 - (1 + \sqrt{17})X - 1 = 0$. Le discriminant, dont on sait à l'avance qu'il est strictement positif puisque les racines (y_1 et y_2) sont réelles et distinctes, est égal à : $\Delta = (1 + \sqrt{17})^2 + 16 = 34 + 2\sqrt{17}$.

23. C'est une nouvelle application de la propriété affirmant que deux nombres complexes z_1 et z_2 sont racines de $X^2 - SX + P = 0$, avec $S = z_1 + z_2$ et $P = z_1 z_2$.

Il s'ensuit que les racines de (E4) sont :

$$\frac{(1 + \sqrt{17}) \pm \sqrt{34 + 2\sqrt{17}}}{8}$$

Ces deux racines sont de signes opposés (leur produit étant égal à $-1/4$) ; on remarque alors que y_1 est positif (puisque $y_1 = \cos(3\theta) + \cos(5\theta)$, et $\cos(3\theta)$ et $\cos(5\theta)$ sont tous deux positifs) pour conclure.

Conclusion : $y_1 = \frac{(1 + \sqrt{17}) + \sqrt{34 + 2\sqrt{17}}}{8}$ et $y_2 = \frac{(1 + \sqrt{17}) - \sqrt{34 + 2\sqrt{17}}}{8}$

On conduit un raisonnement analogue pour obtenir y_3 et y_4 . Mutatis mutandis, on obtient :

Conclusion : $y_3 = \frac{(1 - \sqrt{17}) + \sqrt{34 - 2\sqrt{17}}}{8}$ et $y_4 = \frac{(1 - \sqrt{17}) - \sqrt{34 - 2\sqrt{17}}}{8}$

c) En piste pour le final ! Une n -ième application de la formule $\cos(a)\cos(b) = \frac{\cos(a+b) + \cos(a-b)}{2}$ donne :

$$2\cos(\theta)\cos(13\theta) = \cos(12\theta) + \cos(14\theta) = -(\cos(5\theta) + \cos(3\theta)) = -y_1$$

Donc $\cos(\theta)\cos(13\theta) = -y_1/2$ soit :

$\cos(\theta)\cos(13\theta) = -\frac{(1 + \sqrt{17}) + \sqrt{34 + 2\sqrt{17}}}{16}$

Par ailleurs, et par définition du réel y_3 , on a : $y_3 = \cos(\theta) + \cos(13\theta)$ soit :

$\cos(\theta) + \cos(13\theta) = \frac{(1 - \sqrt{17}) + \sqrt{34 - 2\sqrt{17}}}{8}$

Il s'ensuit que les réels $\cos(\theta)$ et $\cos(13\theta)$ sont racines de l'équation (E5) $X^2 - y_3X - y_1/2 = 0$. Cette équation possède deux racines de signes opposés (de nouveau en observant que le produit de ces racines, égal à $-y_1/2$, est négatif), ce qui tombe drôlement bien car $\cos(\theta)$ est positif tandis que $\cos(13\theta)$ est négatif (13θ appartenant à l'intervalle $[\pi; 2\pi]$).

Il "suffit" donc à présent de résoudre (E5), et de choisir la racine positive pour avoir la valeur exacte de $\cos(\theta)$.

Remarque : le sujet demandait de s'arrêter ici, et de ne pas effectuer les calculs ci-dessous.

Réolvons (E5) : $X^2 - y_3X - y_1/2 = 0$

$$\iff X^2 - \left(\frac{(1 - \sqrt{17}) + \sqrt{34 - 2\sqrt{17}}}{8} \right) X - \frac{(1 + \sqrt{17}) + \sqrt{34 + 2\sqrt{17}}}{16} = 0$$

$$\iff 16X^2 - 2 \left((1 - \sqrt{17}) + \sqrt{34 - 2\sqrt{17}} \right) X - \left((1 + \sqrt{17}) + \sqrt{34 + 2\sqrt{17}} \right) = 0$$

Calculons son discriminant :

$$\Delta = 4 \left((1 - \sqrt{17}) + \sqrt{34 - 2\sqrt{17}} \right)^2 + 64 \left((1 + \sqrt{17}) + \sqrt{34 + 2\sqrt{17}} \right)$$

$$\iff \Delta = 4 \left[\left((1 - \sqrt{17}) + \sqrt{34 - 2\sqrt{17}} \right)^2 + 16 \left((1 + \sqrt{17}) + \sqrt{34 + 2\sqrt{17}} \right) \right]$$

$$\iff \Delta = 4 \left(18 - 2\sqrt{17} + 34 - 2\sqrt{17} + 2(1 - \sqrt{17})\sqrt{34 - 2\sqrt{17}} + 16 + 16\sqrt{17} + 16\sqrt{34 + 2\sqrt{17}} \right)$$

$$\iff \Delta = 4 \left(68 + 12\sqrt{17} + 2(1 - \sqrt{17})\sqrt{34 - 2\sqrt{17}} + 16\sqrt{34 + 2\sqrt{17}} \right)$$

On en déduit que les racines de (E5) sont :

$$\frac{2 \left((1 - \sqrt{17}) + \sqrt{34 - 2\sqrt{17}} \right) \pm 2\sqrt{68 + 12\sqrt{17} + 2(1 - \sqrt{17})\sqrt{34 - 2\sqrt{17}} + 16\sqrt{34 + 2\sqrt{17}}}}{32}$$

soit encore :

$$\frac{1 - \sqrt{17} + \sqrt{34 - 2\sqrt{17}} \pm \sqrt{68 + 12\sqrt{17} + 2(1 - \sqrt{17})\sqrt{34 - 2\sqrt{17}} + 16\sqrt{34 + 2\sqrt{17}}}}{16}$$

En choisissant la racine positive, on obtient enfin :

$$\boxed{\cos\left(\frac{\pi}{17}\right) = \frac{1 - \sqrt{17} + \sqrt{34 - 2\sqrt{17}} + \sqrt{68 + 12\sqrt{17} + 2(1 - \sqrt{17})\sqrt{34 - 2\sqrt{17}} + 16\sqrt{34 + 2\sqrt{17}}}}{16}}$$

Pour complément d'information, une valeur approchée de $\cos\left(\frac{\pi}{17}\right)$ à 10^{-6} près est : 0,982973.

4.10.2 Un calcul de somme classique : $\sum_{k=0}^{n-1} \cos(a + kh)$

Soient a et h deux réels, et n un entier naturel non nul. On pose :

$$C(a, h) = \sum_{k=0}^{n-1} \cos(a + kh)$$

On suppose que $h \neq 0 \pmod{2\pi}$. Montrer que

$$C(a, h) = \frac{\sin\left(\frac{nh}{2}\right) \cos\left(a + (n-1)\frac{h}{2}\right)}{\sin\left(\frac{h}{2}\right)}$$

CORRIGÉ. Soient a et h deux réels avec $h \neq 0 \pmod{2\pi}$, et n un entier naturel non nul. Calculons :

$$C(a, h) = \sum_{k=0}^{n-1} \cos(a + kh)$$

$$\iff C(a, h) = \operatorname{Re} \left(\sum_{k=0}^{n-1} e^{i(a+kh)} \right)$$

$$\iff C(a, h) = \operatorname{Re} \left(e^{ia} \sum_{k=0}^{n-1} e^{ikh} \right)$$

$$\iff C(a, h) = \operatorname{Re} \left(e^{ia} \sum_{k=0}^{n-1} (e^{ih})^k \right)$$

$$\iff C(a, h) = \operatorname{Re} \left(e^{ia} \frac{1 - e^{inh}}{1 - e^{ih}} \right)$$

$$\iff C(a, h) = \operatorname{Re} \left(e^{ia} \frac{e^{inh/2} (e^{-inh/2} - e^{inh/2})}{e^{ih/2} (e^{-ih/2} - e^{ih/2})} \right)$$

$$\iff C(a, h) = \operatorname{Re} \left(e^{i(a+(n-1)h/2)} \frac{-2i \sin(nh/2)}{-2i \sin(h/2)} \right)$$

$$\iff C(a, h) = \operatorname{Re} \left(e^{i(a+(n-1)h/2)} \frac{\sin(nh/2)}{\sin(h/2)} \right)$$

Par suite :
$$C(a, h) = \frac{\sin\left(\frac{nh}{2}\right) \cos\left(a + (n-1)\frac{h}{2}\right)}{\sin\left(\frac{h}{2}\right)} \quad 24$$

4.10.3 Sinus hyperbolique complexe

ENONCÉ. On rappelle que la fonction exponentielle complexe est celle qui à un complexe $z = a + ib$ (a et b réels) associe $e^z = e^a e^{ib}$. La fonction **sinus hyperbolique complexe** (notée SH) est alors définie en posant :

$$\forall z \in \mathbb{C}, \quad \operatorname{SH}(z) = \frac{e^z - e^{-z}}{2}$$

1) Montrer que la fonction SH est impaire ; puis montrer qu'elle est périodique, de période $2i\pi$.

2) Résoudre dans $\mathbb{C}$ l'équation (E1) : $\operatorname{SH}(z) = 0$.

3) Soit λ un nombre réel. Résoudre dans $\mathbb{C}$ l'équation (E2) : $\operatorname{SH}(z) = \lambda$.

CORRIGÉ.

1) Commençons par observer que l'ensemble de définition de SH, qui est $\mathbb{C}$, est symétrique par rapport à zéro ($\forall z \in \mathbb{C}, -z \in \mathbb{C}$). Cette observation faite, on note que : $\forall z \in \mathbb{C}, \operatorname{SH}(-z) = \frac{e^{-z} - e^z}{2} = -\frac{e^z - e^{-z}}{2} = -\operatorname{SH}(z)$. La fonction SH est donc impaire.

► Par ailleurs : $\forall z \in \mathbb{C}, \operatorname{SH}(z + 2i\pi) = \frac{e^{z+2i\pi} - e^{-z-2i\pi}}{2}$. Puisque la fonction exponentielle complexe est $2i\pi$ -périodique, on en déduit que : $\forall z \in \mathbb{C}, \operatorname{SH}(z + 2i\pi) = \frac{e^z - e^{-z}}{2}$, soit : $\forall z \in \mathbb{C}, \operatorname{SH}(z + 2i\pi) = \operatorname{SH}(z)$.

La fonction SH est $2i\pi$ -périodique.

c) $[\operatorname{SH}(z) = 0] \iff [e^z - e^{-z} = 0] \iff [e^z = e^{-z}] \iff [e^{2z} = 1]$. Une application directe du cours permet alors d'affirmer que z est solution de cette équation si et seulement si il existe un entier relatif k tel que : $2z = 2ik\pi$.

Subséquentement : [$\operatorname{SH}(z) = 0$] $\iff$ [$\exists k \in \mathbb{Z}, z = ki\pi$].

24. Et
$$S(a, h) = \sum_{k=0}^{n-1} \sin(a + kh) = \frac{\sin\left(\frac{nh}{2}\right) \sin\left(a + (n-1)\frac{h}{2}\right)}{\sin\left(\frac{h}{2}\right)}.$$

$$\begin{aligned} \text{d) Soit } \lambda \text{ un réel. Alors : } [\text{SH}(z) = \lambda] &\iff \left[\frac{e^z - e^{-z}}{2} = \lambda \right] \iff [e^z - e^{-z} = 2\lambda] \iff [e^{2z} - 1 = 2\lambda e^z] \\ &\iff [e^{2z} - 2\lambda e^z - 1 = 0] \end{aligned}$$

On pose alors : $X = e^z$. L'équation précédente se réécrit : $X^2 - 2\lambda X - 1 = 0$. Cette équation du second degré a pour discriminant $\Delta = 4(\lambda^2 + 1)$, qui est un réel strictement positif. L'équation (en X) possède donc deux solutions :

$$X_1 = \frac{2\lambda + \sqrt{4(\lambda^2 + 1)}}{2} \text{ et } X_2 = \frac{2\lambda - \sqrt{4(\lambda^2 + 1)}}{2} \quad \text{c'est-à-dire : } X_1 = \lambda + \sqrt{\lambda^2 + 1} \text{ et } X_2 = \lambda - \sqrt{\lambda^2 + 1}$$

Faisons le point. Nous venons d'établir que z est solution de l'équation $\text{SH}(z) = \lambda$ si et seulement si $e^z = X_1$ ou $e^z = X_2$. Pour conclure, il reste donc à résoudre ces deux équations (en déterminant les modules et arguments de X_1 et X_2).

- ➡ Aucun problème pour X_1 : puisque c'est un réel strictement positif, $|X_1| = X_1$ et $\arg(X_1) = 0 \pmod{2\pi}$.
- ➡ Attention au petit piège pour X_2 , réel strictement négatif (puisque $\sqrt{\lambda^2 + 1} > \lambda$)²⁵. D'où $|X_2| = -X_2$ et $\arg(X_2) = \pi \pmod{2\pi}$.

Conclusion : pour tout réel λ , l'ensemble des solutions dans $\mathbb{C}$ de l'équation $\text{SH}(z) = \lambda$ est

$$\{\ln(\lambda + \sqrt{\lambda^2 + 1}) + 2ik\pi \mid k \in \mathbb{Z}\} \cup \{\ln(\sqrt{\lambda^2 + 1} - \lambda) + i\pi + 2ik\pi \mid k \in \mathbb{Z}\}$$

25. Tiens, pourquoi au fait ?

Chapitre 5

Fonctions usuelles

5.1 Généralités

5.1.1 Définitions

On commence par rappeler la notion de fonction que vous connaissez déjà, et que nous affinerons en cours d'année.

DÉFINITION 5.1 - Une **fonction numérique** f est un procédé permettant d'associer à tout réel x appartenant à une partie E de $\mathbb{R}$ un unique réel noté $f(x)$.

Dans ce contexte, l'ensemble E est appelé **ensemble de définition** de la fonction f ; le réel $f(x)$ est appelé **image** du réel x par f ; et x est un **antécédent** de $f(x)$ par f .

Remarque : si l'image d'un réel par une fonction est unique, un réel peut avoir un ou plusieurs ou aucun antécédent. Par exemple, le réel 4 a pour antécédents 2 et -2 par la fonction carrée, 0 a un unique antécédent par la même fonction et -1 n'en a aucun.

DÉFINITION 5.2 - L'**image** de f , notée $\text{Im}(f)$ ou $f(E)$ est l'ensemble :

$$f(E) = \{y \in \mathbb{R} / \exists x \in E, f(x) = y\} \quad \text{ou} \quad f(E) = \{f(x) / x \in E\}$$

C'est l'ensemble des valeurs de la fonctions, ou encore l'ensemble des réels admettant au moins un antécédent par f .

Exemples. On a $\exp(\mathbb{R}) = \mathbb{R}_+^*$; $\cos(\mathbb{R}) = [-1, 1]$; $\sin([0, \pi]) = [0, 1]$...

5.1.2 Opérations élémentaires sur les fonctions

DÉFINITION 5.3 - Soient f et g deux fonctions numériques définies sur le même ensemble E . On appelle **somme**, **produit** et **quotient** de f et de g les trois fonctions respectivement définies en posant pour tout réel x dans E :

$$(f + g)(x) = f(x) + g(x) \quad (fg)(x) = f(x) \times g(x) \quad \left(\frac{f}{g}\right)(x) = \frac{f(x)}{g(x)}$$

Ces opérations algébriques sur les fonctions sont associatives et commutatives.

Notation : lorsque E et F sont deux parties de $\mathbb{R}$, on notera F^E l'ensemble des fonctions définies sur E et à valeurs dans F . Par exemple, la fonction exponentielle est un élément de $]0; +\infty[^{\mathbb{R}}$ (et de $\mathbb{R}^{\mathbb{R}}$, qui peut le plus peut le moins...); la fonction sinus est un élément de $[-1; 1]^{\mathbb{R}}$ (et de $\mathbb{R}^{\mathbb{R}}$).

DÉFINITION 5.4 - Etant données deux fonctions $f \in F^E$ et $g \in G^F$, la **composée** de g et f , est la fonction notée $g \circ f$ de G^E définie en posant :

$$\forall x \in E, (g \circ f)(x) = g(f(x))$$

Exemple. On considère la fonction $f : x \in \mathbb{R} \mapsto x^2$ et la fonction $g : x \in \mathbb{R} \mapsto x + 1$. Dans cette situation, on a :

$$\forall x \in \mathbb{R}, (g \circ f)(x) = x^2 + 1 \quad \text{tandis que :} \quad \forall x \in \mathbb{R}, (f \circ g)(x) = (x + 1)^2.$$

En particulier, **la composition n'est pas commutative!** En revanche :

PROPRIÉTÉ 5.1 - La composition des fonctions est une opération associative.

Explicitement, cela signifie que pour tout triplet de fonctions (f, g, h) telles que toutes les composées écrites ci-dessous soient définies, on a :

$$f \circ (g \circ h) = (f \circ g) \circ h$$

PREUVE. Sous réserve que tout le monde soit défini, on a par définition de composée :

$$\forall x \in E, [f \circ (g \circ h)](x) = f((g \circ h)(x)) = f(g(h(x)))$$

et :

$$\forall x \in E, [(f \circ g) \circ h](x) = (f \circ g)(h(x)) = f(g(h(x))) \quad \text{☕}$$

DÉFINITION 5.5 - Soit E une partie de $\mathbb{R}$. L'**identité** de E est la fonction notée id_E définie en posant :

$$\forall x \in E, \text{id}_E(x) = x$$

PROPRIÉTÉ 5.2 - L'identité est l'élément neutre pour la composition, dans le sens plus explicite suivant :

$$\forall f \in F^E, f \circ \text{id}_E = f \quad \text{et} \quad \forall f \in F^E, \text{id}_F \circ f = f$$

PREUVE. Triviale.



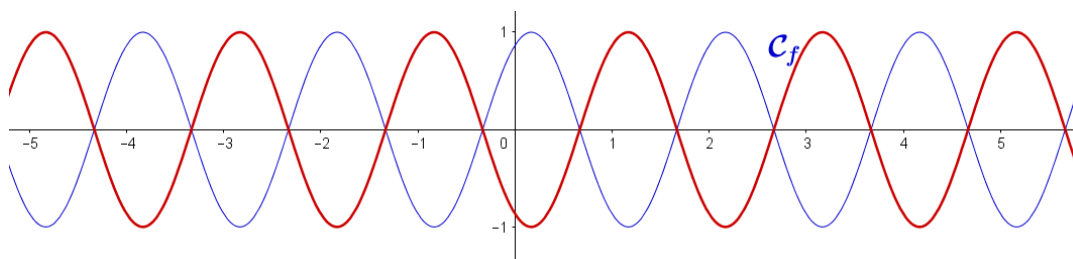
5.1.3 Transformations de la courbe représentative d'une fonction

Soit f une fonction définie sur $\mathbb{R}$, dont on note $\mathcal{C}_f$ la courbe représentative dans un repère orthonormal du plan.

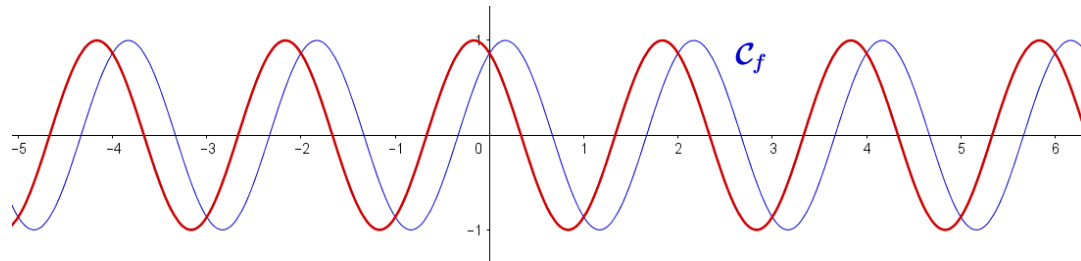
- la courbe représentative de la fonction $x \in \mathbb{R} \mapsto -f(x)$ est l'image de $\mathcal{C}_f$ par une symétrie axiale par rapport à l'axe des abscisses ;
- la courbe représentative de la fonction $x \in \mathbb{R} \mapsto f(-x)$ est l'image de $\mathcal{C}_f$ par une symétrie axiale par rapport à l'axe des ordonnées ;
- la courbe représentative de la fonction $x \in \mathbb{R} \mapsto f(x) + a$ est l'image de $\mathcal{C}_f$ par la translation de vecteur $\vec{v}(0, a)$ (où a est un réel arbitraire) ;
- la courbe représentative de la fonction $x \in \mathbb{R} \mapsto f(x + a)$ est l'image de $\mathcal{C}_f$ par la translation de vecteur $\vec{v}(-a, 0)$ (où a est un réel arbitraire) ;
- la courbe représentative de la fonction $x \in \mathbb{R} \mapsto af(x)$ est l'image de $\mathcal{C}_f$ par une affinité d'axe celui des abscisses, de directrice l'axe des ordonnées, et de rapport a (où a est un réel arbitraire) ;
- la courbe représentative de la fonction $x \in \mathbb{R} \mapsto f(ax)$ est l'image de $\mathcal{C}_f$ par une affinité d'axe celui des ordonnées, de directrice l'axe des abscisses, et de rapport $1/a$ (où a est un réel non nul arbitraire).

On illustre ci-dessous ces différentes transformations, où l'on a choisi $a = 2$ dans toutes les situations (où intervient le réel $a \dots$).

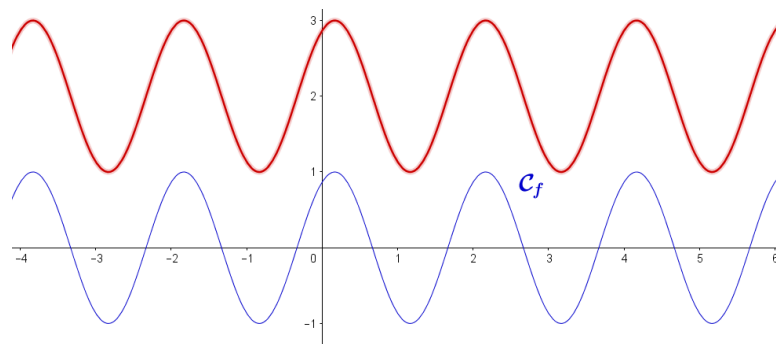
- Courbes représentatives de f et de $x \in \mathbb{R} \mapsto -f(x)$



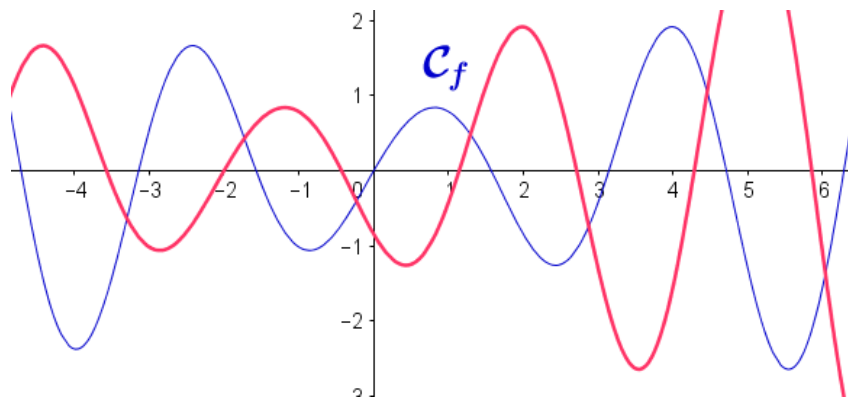
→ Courbes représentatives de f et de $x \in \mathbb{R} \mapsto f(-x)$



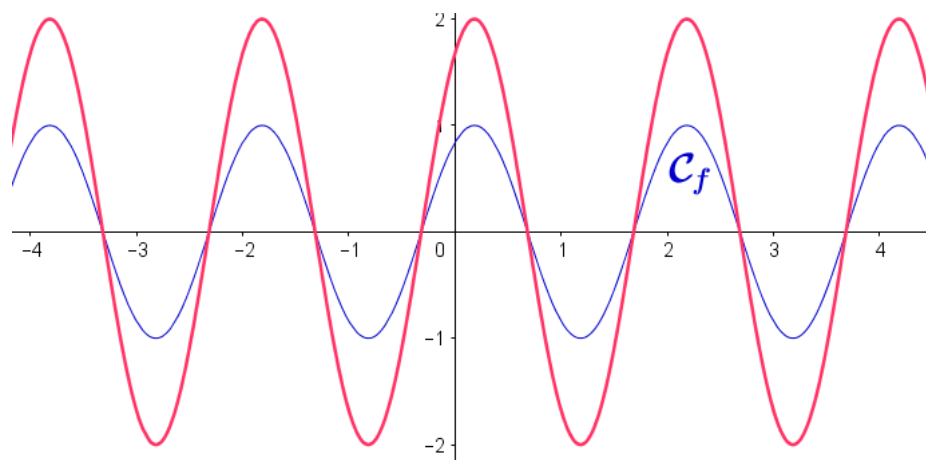
→ Courbes représentatives de f et de $x \in \mathbb{R} \mapsto f(x) + a$



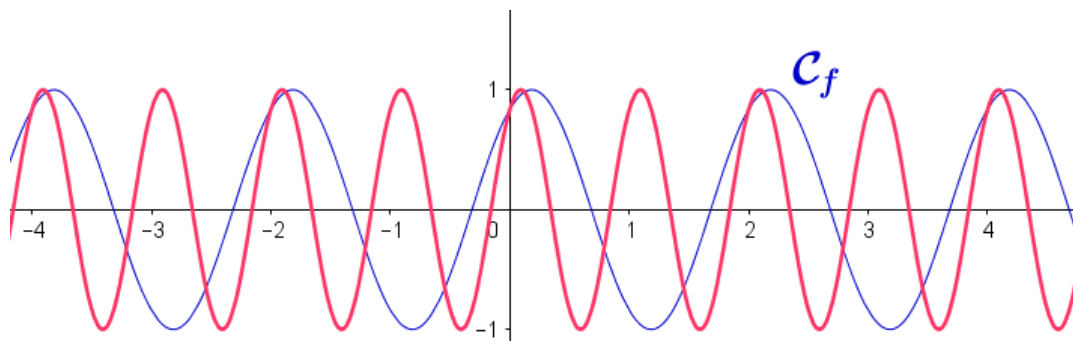
→ Courbes représentatives de f et de $x \in \mathbb{R} \mapsto f(x + a)$



→ Courbes représentatives de f et de $x \in \mathbb{R} \mapsto af(x)$



→ Courbes représentatives de f et de $x \in \mathbb{R} \mapsto f(ax)$



5.1.4 Fonctions paires et impaires

DÉFINITION 5.6 - Soit f une fonction définie sur une partie E de $\mathbb{R}$ et à valeurs dans $\mathbb{R}$. La fonction f est **paire** (*resp.* **impaire**) si :

1/ E est **symétrique par rapport à zéro** : $\forall x \in E, (-x) \in E$

2/ $\forall x \in E, f(-x) = f(x)$ (*resp.* $f(-x) = -f(x)$)

Exemples. 1) La fonction $x \in \mathbb{R} \mapsto x^2$ est paire. La fonction $x \in \mathbb{R}_+ \mapsto x^2$ n'est ni paire ni impaire ($\mathbb{R}_+$ n'est pas symétrique par rapport à 0).

2) La fonction $x \in \mathbb{R} \mapsto x^{2n}$ est paire (pour tout entier naturel n).

3) La fonction $x \in \mathbb{R} \mapsto x^{2n+1}$ est impaire (pour tout entier naturel n).

4) La fonction $x \in \mathbb{R}^* \mapsto 1/x$ est impaire.

5) La fonction $x \in \mathbb{R}^* \mapsto 1/x^2$ est paire.

6) La fonction cosinus est paire, et la fonction sinus est impaire.

7) Les fonctions exponentielle, racine carrée et logarithme népérien sur leurs ensembles de définition usuels ne sont ni paires ni impaires.

Remarque méthodologique. Pour établir qu'une fonction f est impaire, on peut (après avoir vérifié que son ensemble de définition est symétrique par rapport à zéro) vérifier l'identité :

$$\forall x \in D_f, \quad f(x) + f(-x) = 0$$

Cette remarque banale peut se révéler bien pratique dans certains cas, notamment pour établir que la fonction $f : x \in \mathbb{R} \mapsto \ln(x + \sqrt{x^2 + 1})$ est impaire (exercice).

PROPRIÉTÉ 5.3 - Dans un repère orthonormal du plan, la courbe représentative d'une fonction paire (*resp.* impaire) est symétrique par rapport à l'axe des ordonnées (*resp.* symétrique par rapport à l'origine du repère).

PREUVE. Dans un repère orthonormal du plan, les points de coordonnées $(x, f(x))$ et $(-x, f(x))$ sont symétriques par rapport à l'axe des ordonnées ; et les points de coordonnées $(x, f(x))$ et $(-x, -f(x))$ sont symétriques par rapport à l'origine du repère. ☕

L'énoncé ci-dessous permet en pratique de n'étudier les variations d'une fonction paire ou impaire uniquement "d'un côté de zéro" pour connaître ses variations sur tout son ensemble de définition.

PROPRIÉTÉ 5.4 - Soit f une fonction définie sur une partie E de $\mathbb{R}$ symétrique par rapport à zéro. On note :

$$E_+ = \{x \in E / x \geq 0\} \quad \text{et} \quad E_- = \{x \in E / x \leq 0\}$$

- 1) Si f est paire, et si f est croissante sur E_+ , alors f est décroissante sur E_- ;
- 2) Si f est impaire, et si f est croissante sur E_+ , alors f est croissante sur E_- .

PREUVE. Triviale. ☕

Remarque. On peut bien entendu permuter les rôles de "croissante" et "décroissante" dans l'énoncé précédent.

On rappelle pour conclure sur ce thème un énoncé établi au cours du premier chapitre.

THÉORÈME 5.1 - (de décomposition "paire + impaire"). Toute fonction $f : D \rightarrow \mathbb{R}$ avec $D \subset \mathbb{R}$ symétrique par rapport à 0 s'écrit de manière unique comme la somme d'une fonction paire et d'une fonction impaire.

PREUVE. Voir page 23, preuve du théorème 1.1. ☕

5.2 Fonctions et inégalités

DÉFINITION 5.7 - Soit f une fonction définie sur une partie E de $\mathbb{R}$, et à valeurs réelles. La fonction f est :

- 1/ **minorée sur E** si : $\exists m \in \mathbb{R}, \forall x \in E, f(x) \geq m$
- 2/ **majorée sur E** si : $\exists M \in \mathbb{R}, \forall x \in E, f(x) \leq M$
- 3/ **bornée sur E** si elle est minorée et majorée sur E .

Terminologie : avec les notations de la définition ci-dessus, le réel m est appelé un **minorant** de f , et le réel M un **majorant** de f . Dès qu'une fonction f admet un minorant m (*resp.* un majorant M), alors tout réel inférieur ou égal à m (*resp.* supérieur ou égal à M) est un minorant (*resp.* un majorant) de f .

Exemples. La fonction exponentielle est minorée par 0 sur $\mathbb{R}$, mais n'est pas majorée. La fonction logarithme népérien n'est ni minorée, ni majorée sur $\mathbb{R}_+^*$. Les fonctions cosinus et sinus sont minorées par -1 et majorées par 1.

DÉFINITION 5.8 - Soit f une fonction définie sur une partie E de $\mathbb{R}$, et à valeurs réelles. La fonction f est :

- 1/ **croissante sur E** si : $\forall (x, y) \in E^2, [x \leq y] \implies [f(x) \leq f(y)]$
- 2/ **décroissante sur E** si : $\forall (x, y) \in E^2, [x \leq y] \implies [f(x) \geq f(y)]$
- 3/ **monotone sur E** si elle est croissante ou décroissante sur E .

Remarque. Comme vous le savez, on dispose également de versions strictes de ces définitions.

Exemples. La fonction carrée est croissante sur $\mathbb{R}_+$, décroissante sur $\mathbb{R}_-$. La fonction cube, la fonction exponentielle, la fonction logarithme népérien sont monotones sur leurs ensembles de définition respectifs.

DÉFINITION 5.9 - Soit f une fonction définie sur un ensemble E , et soit a un réel appartenant à E . On dit que :

- 1/ f admet un **maximum** en a , ou que $f(a)$ est le **maximum de f sur E** si :

$$\forall x \in E, \quad f(x) \leq f(a)$$

- 2/ f admet un **minimum** en a , ou que $f(a)$ est le **minimum de f sur E** si :

$$\forall x \in E, \quad f(x) \geq f(a)$$

- 3/ f admet un **extremum** en a , ou que $f(a)$ est un **extremum de f sur E** si $f(a)$ est le minimum ou le maximum de f sur E .

5.3 Valeur absolue

DÉFINITION 5.10 - Pour tout réel x , la **valeur absolue** de x est le réel noté $|x|$ et défini en posant :

$$|x| = \max(-x, x)$$

PROPRIÉTÉ 5.5 - $\forall x \in \mathbb{R}, |x| = \sqrt{x^2}$

PREUVE. Le réel positif dont le carré est égal à x^2 est x lorsque x est positif, et $-x$ sinon. ☕

Conséquence : pour un nombre réel, la **valeur absolue coïncide avec le module**. Ceci implique que la valeur absolue possède les propriétés ci-dessous (établies pour les modules dans $\mathbb{C}$) :

PROPRIÉTÉ 5.6 - (Propriétés de la valeur absolue).

1/ $\forall (x, y) \in \mathbb{R}^2, |x \times y| = |x| \times |y|$

2/ En particulier : $\forall x \in \mathbb{R}, |-x| = |x|$

3/ $\forall (x, n) \in \mathbb{R} \times \mathbb{N}, |x^n| = |x|^n$

4/ $\forall y \in \mathbb{R}^*, \left| \frac{1}{y} \right| = \frac{1}{|y|}$

5/ $\forall (x, y) \in \mathbb{R} \times \mathbb{R}^*, \left| \frac{x}{y} \right| = \frac{|x|}{|y|}$

6/ Inégalité triangulaire : $\forall (x, y) \in \mathbb{R}^2, |x + y| \leq |x| + |y|$

PREUVE. Voir chapitre sur les nombres complexes. ☕**PROPRIÉTÉ 5.7 - (Inégalité triangulaire “renversée”).**

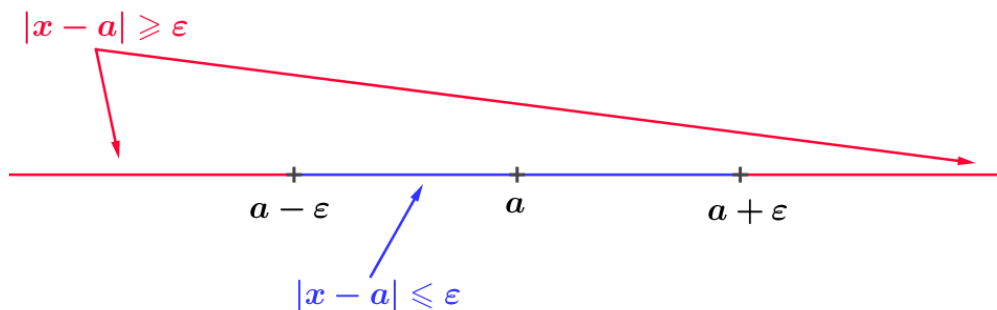
$$\forall (x, y) \in \mathbb{R}^2, ||x| - |y|| \leq |x - y|$$

PREUVE. Soient x et y deux nombres réels. D’après l’inégalité triangulaire, on a :

$$|x| \leq |y| + |x - y| \quad \text{d'où : } |x| - |y| \leq |x - y|$$

D'où : $||x| - |y|| \leq |x - y|$. ☕**Interprétation géométrique de la valeur absolue (très important en pratique dans la suite de cours de cette année) :**Pour tout réel x , on a : $|x| = d(x, 0)$ (distance entre les points d’abscisse 0 et x).Pour tout couple de réels (x, a) , on a : $d(x, a) = |x - a|$. Il s’ensuit que : $d(x, a) = 0 \iff x = a$.Soit en outre ε un réel strictement positif. Alors :

- $|x - a| = \varepsilon$ SSI $x = a \pm \varepsilon$
- $|x - a| \leq \varepsilon$ SSI $x \in [a - \varepsilon; a + \varepsilon]$;
- $|x - a| \geq \varepsilon$ SSI $x \in]-\infty; a - \varepsilon] \cup [a + \varepsilon; +\infty[$.



On finit ce paragraphe par une nouvelle propriété pratiquement importante pour la suite de cours de cette année.

LEMME 5.1 - Soit f une fonction à valeurs réelles.

$$[f \text{ est bornée}] \iff [|f| \text{ est majorée}]$$

PREUVE. Supposons f bornée. Il existe deux réels m et M tels que : $m \leq f \leq M$.

➤ Premier cas. Si $0 \leq m \leq M$. Alors il résulte de l'hypothèse que : $0 \leq f \leq M$, d'où en particulier : $|f| \leq M$.

➤ Second cas. Si $m \leq M \leq 0$. Alors il résulte de l'hypothèse que : $m \leq f \leq 0$, d'où en particulier : $|f| \leq -m$.

➤ Troisième cas. Si $m \leq 0 \leq M$. Alors il résulte de l'hypothèse que : $m \leq f \leq 0$ et $0 \leq f \leq M$. D'où en particulier : $|f| \leq \max(-m, M)$.

Dans tous les cas, si f est bornée, alors $|f|$ est majorée.

L'implication réciproque est immédiate. ☕

5.4 Fonctions dérivables (rappels de Tale et un peu plus...)

Tout au long de ce paragraphe, f désignera une fonction à valeurs réelles définie sur un intervalle ouvert de $\mathbb{R}$, et a un élément de I .

DÉFINITION 5.11 - La fonction f est **dérivable en** a si la limite du taux d'accroissement de f entre a et $a + h$ lorsque h tend vers 0 existe et est finie.

Lorsque tel est le cas, on appelle **nombre dérivé de f en a** et on note $f'(a)$ cette limite. Ainsi, sous réserve d'existence et de finitude :

$$f'(a) = \lim_{h \rightarrow 0} \frac{f(a+h) - f(a)}{h} \quad \text{ou} \quad f'(a) = \lim_{x \rightarrow a} \frac{f(x) - f(a)}{x - a}$$

Exemples : en tout réel a , la fonction $f : x \mapsto x^n$ (avec $n \in \mathbb{N}^*$) est dérivable et $f'(a) = na^{n-1}$. La fonction $g : x \mapsto \sqrt{x}$ est dérivable en tout réel non nul a , et $g'(a) = \frac{1}{2\sqrt{a}}$; elle n'est en revanche pas dérivable (à droite) en 0, puisque la limite (à droite) du taux d'accroissement de g entre 0 et h quand h tend vers 0 est égale à $+\infty$. La fonction $h : x \mapsto |x|$ n'est pas dérivable en 0 puisque les limites du taux d'accroissement de f à gauche et à droite de 0 existent, sont finies (respectivement égales à -1 et 1), mais sont distinctes.

DÉFINITION 5.12 - Lorsque la fonction f est dérivable en tout $a \in I$, on dit que la fonction f est **dérivable sur** I et on note f' la dérivée de f (c'est la fonction définie sur I , et qui à tout réel de a de I associe le nombre dérivé $f'(a)$).

THÉORÈME 5.2 - f est dérivable en a SSI il existe un réel ℓ et une fonction ε définie sur I telle que :

$$\forall h \in \mathbb{R}, (a+h) \in I, f(a+h) = f(a) + \ell h + h\varepsilon(h) \text{ avec } \lim_{h \rightarrow 0} \varepsilon(h) = 0$$

PREUVE. On raisonne par double implication.

► **Sens direct** : supposons que f soit dérivable en a .

On écrit, pour tout réel h (tel que $(a+h) \in I$) :

$$f(a+h) = f(a) + hf'(a) + f(a+h) - f(a) - hf'(a) \quad ^1$$

D'où : $f(a+h) = f(a) + hf'(a) + h \left(\frac{f(a+h) - f(a)}{h} - f'(a) \right)$ d'où : $f(a+h) = f(a) + hf'(a) + h\varepsilon(h)$ (♠)

en ayant posé : $\varepsilon(h) = \frac{f(a+h) - f(a)}{h} - f'(a)$.

Or, f étant dérivable en a , on a : $\lim_{h \rightarrow 0} \frac{f(a+h) - f(a)}{h} = f'(a)$ et donc $\lim_{h \rightarrow 0} \varepsilon(h) = 0$.

En résumé, on a établi l'implication :

$$[f \text{ est dérivable en } a] \implies [\forall h \in \mathbb{R}, (a+h) \in I, f(a+h) = f(a) + \ell h + h\varepsilon(h) \text{ avec } \lim_{h \rightarrow 0} \varepsilon(h) = 0] \quad (\heartsuit)$$

► **Réciproquement** : supposons qu'il existe un réel ℓ tel que f vérifie : $f(a+h) = f(a) + \ell h + h\varepsilon(h)$ avec $\lim_{h \rightarrow 0} \varepsilon(h) = 0$.

Alors pour tout réel h non nul tel que $(a+h) \in I$, on a : $\frac{f(a+h) - f(a)}{h} = \ell + \varepsilon(h)$.

D'où : $\lim_{h \rightarrow 0} \frac{f(a+h) - f(a)}{h} = \ell$. Ce qui signifie que la fonction f est dérivable en a (et que $f'(a) = \ell$). Ce qui assure que :

$$[\forall h \in \mathbb{R}, (a+h) \in I, f(a+h) = f(a) + \ell h + h\varepsilon(h) \text{ avec } \lim_{h \rightarrow 0} \varepsilon(h) = 0] \implies [f \text{ est dérivable en } a] \quad (\clubsuit)$$

Conclusion. f est dérivable en a SSI il existe un réel ℓ et une fonction ε définie au voisinage de zéro tels que :

$$\forall h \in \mathbb{R}, (a+h) \in I, f(a+h) = f(a) + \ell h + h\varepsilon(h) \text{ avec } \lim_{h \rightarrow 0} \varepsilon(h) = 0.$$



1. Diabolique, n'est-ce pas ?

Lorsque la propriété du théorème précédent est satisfaite est le cas, on a : $\ell = f'(a)$. L'égalité du théorème se réécrit alors :

$$\forall h \in \mathbb{R}, (a+h) \in I, f(a+h) = f(a) + hf'(a) + h\varepsilon(h) \text{ avec } \lim_{h \rightarrow 0} \varepsilon(h) = 0$$

Le terme de droite de l'égalité ci-dessus est alors appelé **développement limité à l'ordre 1 de f en a** ; l'expression $f(a) + hf'(a)$ est l'**approximation affine** de f au voisinage de a .

On déduit alors du théorème les DL 1 usuels en zéro, que voici :

$$1/ \quad e^h = 1 + h + h\varepsilon(h)$$

$$2/ \quad \sin(h) = h + h\varepsilon(h)$$

$$3/ \quad \tan(h) = h + h\varepsilon(h)$$

$$4/ \quad \operatorname{sh}(h) = h + h\varepsilon(h)$$

$$5/ \quad \operatorname{th}(h) = h + h\varepsilon(h)$$

$$6/ \quad \ln(1+h) = h + h\varepsilon(h)$$

$$7/ \quad \cos(h) = 1 + h\varepsilon(h)$$

$$8/ \quad \operatorname{ch}(h) = 1 + h\varepsilon(h)$$

$$9/ \quad \text{Pour tout } n \in \mathbb{N},$$

$$(1+h)^n = 1 + nh + h\varepsilon(h)$$

$$10/ \quad \text{Et même, pour tout } \alpha \in \mathbb{R},$$

$$(1+h)^\alpha = 1 + \alpha h + h\varepsilon(h)$$

$$11/ \quad \text{En particulier, pour } \alpha = 1/2,$$

$$\sqrt{1+h} = 1 + \frac{h}{2} + h\varepsilon(h)$$

$$12/ \quad \text{En particulier (bis), pour } \alpha = -1,$$

$$\frac{1}{1+h} = 1 - h + h\varepsilon(h)$$

$$13/ \quad \frac{1}{1-h} = 1 + h + h\varepsilon(h)$$

$$14/ \quad \arccos(h) = \frac{\pi}{2} - h + h\varepsilon(h)$$

$$15/ \quad \arcsin(h) = h + h\varepsilon(h)$$

$$16/ \quad \arctan(h) = h + h\varepsilon(h)$$

► Conséquences

$$1/ \quad e^{1/n} = 1 + \frac{1}{n} + \frac{1}{n} \varepsilon\left(\frac{1}{n}\right)$$

$$(\text{avec } \lim_{n \rightarrow +\infty} \varepsilon\left(\frac{1}{n}\right) = 0)$$

$$2/ \quad \sin\left(\frac{1}{n}\right) = \frac{1}{n} + \frac{1}{n} \varepsilon\left(\frac{1}{n}\right)$$

$$3/ \quad \tan\left(\frac{1}{n}\right) = \frac{1}{n} + \frac{1}{n} \varepsilon\left(\frac{1}{n}\right)$$

$$4/ \quad \ln\left(1 + \frac{1}{n}\right) = \frac{1}{n} + \frac{1}{n} \varepsilon\left(\frac{1}{n}\right)$$

$$5/ \quad \text{Etc.} \dots$$

Exemple d'application - Exercice classique. Calcul de $\lim_{n \rightarrow +\infty} \left(1 + \frac{1}{n}\right)^n$.

Pour tout $n \in \mathbb{N}^*$, on a : $\left(1 + \frac{1}{n}\right)^n = e^{n \ln(1 + \frac{1}{n})}$ ($\spadesuit$).

Par ailleurs : $\forall h > -1$, $\ln(1 + h) = h + h\varepsilon(h)$ avec $\lim_{h \rightarrow 0} \varepsilon(h) = 0$.

En posant $h = 1/n$ (avec $n \in \mathbb{N}^*$), le réel h tend vers 0 lorsque n tend vers $+\infty$, et il s'ensuit que :

$$\forall n \in \mathbb{N}^*, \ln\left(1 + \frac{1}{n}\right) = \frac{1}{n} + \frac{1}{n} \varepsilon\left(\frac{1}{n}\right) \text{ avec } \lim_{n \rightarrow +\infty} \varepsilon\left(\frac{1}{n}\right) = 0$$

D'où pour tout entier naturel n non nul : $n \ln\left(1 + \frac{1}{n}\right) = 1 + \varepsilon\left(\frac{1}{n}\right)$ avec $\lim_{n \rightarrow +\infty} \varepsilon\left(\frac{1}{n}\right) = 0$.

Par suite : $\lim_{n \rightarrow +\infty} n \ln\left(1 + \frac{1}{n}\right) = 1$ ($\heartsuit$). On déduit de ($\spadesuit$) et ($\heartsuit$) que : $\lim_{n \rightarrow +\infty} \left(1 + \frac{1}{n}\right)^n = e$.

Lien avec le cours de Physique : “pour des petites valeurs de θ , $\sin(\theta) = \theta$ ”. Cette phrase, mathématiquement fausse, doit être comprise en Physique comme $\sin(\theta) = \theta + \theta\varepsilon(\theta)$ avec $\lim_{\theta \rightarrow 0} \varepsilon(\theta) = 0$. Ce qui signifie grossièrement que lorsque θ est petit (“petit = proche de zéro”), on ne commet pas une grande erreur en remplaçant $\sin(\theta)$ par son approximation affine au voisinage de 0.

PROPRIÉTÉ 5.8 - (Propriétés algébriques des dérivées). Soient f et g deux fonctions définies sur un voisinage ouvert d'un réel a . On suppose que f et g sont dérivables en a . Alors :

1/ $f + g$ est dérivable en a et : $(f + g)'(a) = f'(a) + g'(a)$

2/ $f \times g$ est dérivable en a et : $(f \times g)'(a) = f'(a) \times g(a) + f(a) \times g'(a)$

3/ Si $g(a) \neq 0$, $\frac{1}{g}$ est dérivable en a et $\left(\frac{1}{g}\right)'(a) = -\frac{g'(a)}{g^2(a)}$

4/ Si $g(a) \neq 0$, $\frac{f}{g}$ est dérivable en a et $\left(\frac{f}{g}\right)'(a) = \frac{f'(a) \times g(a) - f(a) \times g'(a)}{g^2(a)}$

PREUVE. Sous les hypothèses de l'énoncé, il existe deux fonctions ε_1 et ε_2 que l'on peut supposer SNALG définies sur un même voisinage V de a telles que :

$$\forall x \in V, f(x) = f(a) + (x - a)f'(a) + (x - a)\varepsilon_1(x) \quad \text{et} \quad \forall x \in V, g(x) = g(a) + (x - a)g'(a) + (x - a)\varepsilon_2(x)$$

avec : $\lim_a \varepsilon_1 = \lim_a \varepsilon_2 = 0$.

1/ On a : $\forall x \in V, (f + g)(x) = (f + g)(a) + (x - a)(f'(a) + g'(a)) + (x - a)[\varepsilon_1(x) + \varepsilon_2(x)]$.

En posant : $\varepsilon_3(x) = \varepsilon_1(x) + \varepsilon_2(x)$, on a : $\lim_a \varepsilon_3 = 0$.

Il vient alors : $\forall x \in V, (f + g)(x) = (f + g)(a) + (x - a)(f'(a) + g'(a)) + (x - a)\varepsilon_3(x)$. On déduit de cette écriture et du théorème 5.2 que $(f + g)$ est dérivable en a , et que : $(f + g)'(a) = f'(a) + g'(a)$.

2/ Soit $x \in V$. On a :

$$(f \times g)(x) = [f(a) + (x-a)f'(a) + (x-a)\varepsilon_1(x)] [g(a) + (x-a)g'(a) + (x-a)\varepsilon_2(x)].$$

$$\iff (f \times g)(x) = (fg)(a) + (x-a)(f'(a)g(a) + f(a)g'(a))$$

$$+ (x-a)[f(a)\varepsilon_2(x) + g(a)\varepsilon_1(x) + (x-a)f'(a)g'(a) + (x-a)f'(a)\varepsilon_2(x) + (x-a)g'(a)\varepsilon_1(x) + (x-a)\varepsilon_1(x)\varepsilon_2(x)]$$

On pose alors :

$$\varepsilon_3(x) = f(a)\varepsilon_2(x) + g(a)\varepsilon_1(x) + (x-a)f'(a)g'(a) + (x-a)f'(a)\varepsilon_2(x) + (x-a)g'(a)\varepsilon_1(x) + (x-a)\varepsilon_1(x)\varepsilon_2(x).$$

On observe avec soulagement que : $\lim_a \varepsilon_3 = 0$.

Par suite :

$$\forall x \in V, (f \times g)(x) = (f \times g)(a) + (x-a)(f'(a)g(a) + f(a)g'(a)) + (x-a)\varepsilon_3(x)$$

On déduit de cette écriture et du théorème 5.2 que $(f \times g)$ est dérivable en a , et que :

$$(f \times g)'(a) = f'(a)g(a) + f(a)g'(a).$$

3/ Supposons donc en outre que g ne s'annule pas en a . Alors g ne s'annule pas sur un voisinage de a , et pour tout réel x appartenant à ce voisinage, on a :

$$\left(\frac{1}{g}\right)(x) = \frac{1}{g(x)} = \frac{1}{g(a) + (x-a)g'(a) + (x-a)\varepsilon_2(x)}$$

D'où :

$$\left(\frac{1}{g}\right)(x) = \frac{1}{g(a)} \times \frac{1}{1 + (x-a)\frac{g'(a)}{g(a)} + (x-a)\varepsilon_3(x)} \text{ en posant : } \varepsilon_3(x) = \frac{\varepsilon_2(x)}{g(a)}.$$

Par conséquent :

$$\left(\frac{1}{g}\right)(x) = \frac{1}{g(a)} \times \left[1 - (x-a)\frac{g'(a)}{g(a)} + (x-a)\varepsilon_4(x)\right] \text{ avec } \lim_a \varepsilon_4 = 0.$$

Finalement :

$$\left(\frac{1}{g}\right)(x) = \frac{1}{g(a)} - (x-a)\frac{g'(a)}{g^2(a)} + (x-a)\varepsilon_4(x) \text{ avec } \lim_a \varepsilon_4 = 0.$$

On déduit de cette écriture et du théorème 5.2 que $(1/g)$ est dérivable en a , et que :

$$\left(\frac{1}{g}\right)'(a) = -\frac{g'(a)}{g^2(a)}.$$

4/ Conséquence immédiate des propriétés 2 et 3. ☕

PROPRIÉTÉ 5.9 - (Linéarité de la dérivation). Soient f et g deux fonctions définies sur un voisinage ouvert d'un réel a . On suppose que f et g sont dérivables en a .

Alors pour tout couple (λ, μ) de réels, la fonction $(\lambda f + \mu g)$ est dérivable en a et :

$$(\lambda f + \mu g)'(a) = \lambda f'(a) + \mu g'(a)$$

PREUVE. Triviale. ☕

THÉORÈME 5.3 - (Dérivée d'une composée). Soient I et J deux intervalles ouverts de $\mathbb{R}$, et $a \in I$.

Soient encore $f \in \mathbb{R}^I$ et $g \in \mathbb{R}^J$ deux fonctions à valeurs réelles, avec $f(I) \subset J$.

Si f est dérivable en a et si g est dérivable en $f(a)$, alors $g \circ f$ est dérivable en a et :

$$(g \circ f)'(a) = g'(f(a)) \times f'(a)$$

PREUVE. Soient a, I, J, f et g comme indiqués dans l'énoncé du théorème.²

Puisque f est dérivable en a , on a pour tout réel h "raisonnable"³ : $f(a+h) = f(a) + hf'(a) + h\varepsilon(h)$ avec $\lim_{h \rightarrow 0} \varepsilon(h) = 0$.

De même, g étant dérivable en $f(a)$, on a pour tout réel H "raisonnable"⁴ :

$$g(f(a) + H) = g(f(a)) + Hg'(f(a)) + H\widehat{\varepsilon}(H) \text{ avec } \lim_{H \rightarrow 0} \widehat{\varepsilon}(H) = 0$$

On en déduit que :

$$\begin{aligned} (g \circ f)(a+h) &= g(f(a+h)) = g\left(f(a) + \underbrace{hf'(a) + h\varepsilon(h)}_{= "H''"}\right) \\ &= g(f(a)) + (hf'(a) + h\varepsilon(h))g'(f(a)) + (hf'(a) + h\varepsilon(h))\widehat{\varepsilon}(H) \end{aligned}$$

D'où finalement :

$$(g \circ f)(a+h) = g(f(a)) + hf'(a)g'(f(a)) + h[\varepsilon(h)g'(f(a)) + (f'(a) + \varepsilon(h))\widehat{\varepsilon}(H)]$$

Soit encore :

$$(g \circ f)(a+h) = g(f(a)) + hf'(a)g'(f(a)) + h\widetilde{\varepsilon}(h) \text{ en ayant noté : } \widetilde{\varepsilon}(h) = \varepsilon(h)g'(f(a)) + (f'(a) + \varepsilon(h))\widehat{\varepsilon}(H).$$

Puisqu'il est assez immédiat que $\lim_{h \rightarrow 0} \widetilde{\varepsilon}(h) = 0$, on déduit de cette écriture (et du théorème 5.2) que :

$$g \circ f \text{ est dérivable en } a \text{ et : } (g \circ f)'(a) = g'(f(a)) \times f'(a). \quad \text{☕}$$

2. Non mais !

3. C'est-à-dire tel que $a+h$ appartienne à I , l'intervalle de définition de la fonction f .

4. C'est-à-dire tel que $f(a)+H$ appartienne à J , l'intervalle de définition de la fonction g .

Applications : colonne de droite du formulaire des dérivées (voir un peu plus loin dans ce chapitre).

Remarque - Attention ! : on ne peut en revanche pas conclure à la non-dérivabilité de $g \circ f$ si l'une des deux fonctions f ou g n'est pas dérivable. Explicitement, notons g la fonction racine carrée (non dérivable en 0), f_1 la fonction $x \mapsto x^2$, et f_2 la fonction $x \mapsto x^4$.

Dans ce cas $g \circ f_1$ (*resp.* $g \circ f_2$) est la fonction valeur absolue (*resp.* la fonction carrée). Ainsi $g \circ f_1$ n'est pas dérivable en 0, tandis que $g \circ f_2$ est dérivable en 0.

5.5 Dérivées d'ordre supérieur

DÉFINITION 5.13 - Soit f une fonction définie sur un intervalle ouvert I de $\mathbb{R}$. On appelle **dérivée n -ième** de f la fonction définie (sous réserve d'existence) récursivement en posant :

$$f^{(0)} = f \quad \text{et} \quad \forall N \in \mathbb{N}^*, f^{(N)} = (f^{(N-1)})'$$

Exemples — Dérivées n -ièmes usuelles

Les formules ci-dessous sont valables pour tout entier naturel n , et pour tout réel x raisonnable.

- $[e^x]^{(n)} = [e^x]$; et plus généralement $[e^{ax+b}]^{(n)} = [a^n e^{ax+b}]$
- $[\cos(x)]^{(n)} = \left[\cos\left(x + n\frac{\pi}{2}\right) \right]$
- $[\sin(x)]^{(n)} = \left[\sin\left(x + n\frac{\pi}{2}\right) \right]$
- $[\operatorname{ch}(x)]^{(n)} = \frac{e^x + (-1)^n e^{-x}}{2} = \begin{cases} \operatorname{ch}(x) & \text{si } n \text{ est pair} \\ \operatorname{sh}(x) & \text{si } n \text{ est impair} \end{cases}$
- $\left[\frac{1}{x}\right]^{(n)} = \left[(-1)^n \frac{n!}{x^{n+1}}\right]$
- $[\ln(x)]^{(n)} = \left[(-1)^{n+1} \frac{(n-1)!}{x^n}\right] \quad (n \in \mathbb{N}^*)$
- $[x^N]^{(n)} = \begin{cases} \frac{N!}{(N-n)!} x^{N-n} & \text{si } n \leq N \\ 0 & \text{si } n > N \end{cases}$
- $[P(x)]^{(n)} = 0$ si P polynôme, et $n > \deg(P)$
- $[f(ax+b)]^{(n)} = a^n [f^{(n)}(ax+b)]$ pour toute f (n fois dérivable)

PREUVE. Récurrence, récurrence, récurrence, récurrence...



PROPRIÉTÉ 5.10 - (“Linéarité de la dérivée n -ième”). Si f et g sont n fois dérivables sur I , alors pour tout couple (λ, μ) de réels, $(\lambda f + \mu g)$ l'est également et :

$$(\lambda f + \mu g)^{(n)} = \lambda f^{(n)} + \mu g^{(n)}$$

PREUVE. Récurrence immédiate à partir de la propriété de linéarité de la dérivation. ☕

THÉORÈME 5.4 - (Formule de Leibnitz). Si f et g sont n fois dérivables sur I , alors (fg) l'est également et :

$$(fg)^{(n)} = \sum_{k=0}^n \binom{n}{k} f^{(k)} g^{(n-k)} \quad \text{ou} \quad (fg)^{(n)} = \sum_{k=0}^n \binom{n}{k} f^{(n-k)} g^{(k)}$$

PREUVE. Prouvons le théorème par récurrence sur n .

Posons $P(n) : (fg)^{(n)} = \sum_{k=0}^n \binom{n}{k} f^{(k)} g^{(n-k)}$ (où f et g désignent deux fonctions n fois dérivables arbitraires).

► **Initialisation** : pour $n = 0$, on a $(fg)^{(0)} = fg$ et $\sum_{k=0}^0 \binom{0}{k} f^{(k)} g^{(0-k)} = \binom{0}{0} f^{(0)} g^{(0)} = fg$. Ainsi $P(0)$ est vraie.

► **Hérédité** : on suppose $P(n)$ vraie pour un certain entier naturel n .⁵ Soient f et g deux fonctions $(n+1)$ fois dérivables. Alors :

$$\begin{aligned} (fg)^{(n+1)} &= ((fg)^{(n)})' = \sum_{k=0}^n \binom{n}{k} (f^{(k)} g^{(n-k)})' = \sum_{k=0}^n \binom{n}{k} f^{(k+1)} g^{(n-k)} + \sum_{k=0}^n \binom{n}{k} f^{(k)} g^{(n+1-k)} \\ &= \sum_{k=1}^{n+1} \binom{n}{k-1} f^{(k)} g^{(n+1-k)} + \sum_{k=0}^n \binom{n}{k} f^{(k)} g^{(n+1-k)} = f^{(n+1)} + \left[\sum_{k=1}^n \left(\binom{n}{k-1} + \binom{n}{k} \right) f^{(k)} g^{(n+1-k)} \right] + g^{(n+1)} \end{aligned}$$

D'où, en appliquant la relation de Pascal⁶ : $(fg)^{(n+1)} = \sum_{k=0}^{n+1} \binom{n+1}{k} f^{(k)} g^{(n+1-k)}$

Cette relation assure que la propriété $P(n+1)$ est vraie, ce qui établit l'hérédité.

Conclusion. Pour tout entier naturel n , et pour tout couple de fonctions (f, g) n fois dérivables sur I , fg est n fois dérivable sur I et :

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

5. Hypothèse de récurrence.

6. $\forall n \in \mathbb{N}^*, \forall k \in \llbracket 1, n \rrbracket, \binom{n}{k-1} + \binom{n}{k} = \binom{n+1}{k}$

Application — Exercice classique. Calcul de $\sum_{k=0}^n \binom{n}{k}^2$.

Soit n un entier naturel arbitraire. Pour tout réel x , on pose $\varphi(x) = x^{2n}$. La fonction φ est n fois dérivable sur $\mathbb{R}$.⁷ On se propose de calculer $\varphi^{(n)}$ de deux façons différentes.

→ Calcul direct de $\varphi^{(n)}$. $\forall x \in \mathbb{R}, \varphi^{(n)}(x) = \frac{(2n)!}{n!} x^n \quad (\spadesuit)$.⁸

→ Calcul de $\varphi^{(n)}$ via la formule de Leibniz. On écrit judicieusement : $\varphi(x) = \underbrace{x^n}_{f(x)} \times \underbrace{x^n}_{g(x)}$.

Les fonctions f et g étant dérivables n fois sur $\mathbb{R}$, la formule de Leibniz permet alors d'affirmer que pour tout réel x :

$$\begin{aligned} \varphi^{(n)}(x) &= \sum_{k=0}^n \binom{n}{k} f^{(k)}(x) g^{(n-k)}(x) = \sum_{k=0}^n \binom{n}{k} \frac{n!}{(n-k)!} x^{n-k} \frac{n!}{k!} x^{n-(n-k)} = \sum_{k=0}^n \binom{n}{k} \frac{n!n!}{(n-k)!k!} x^n \\ &= n! \sum_{k=0}^n \binom{n}{k} \underbrace{\frac{n!}{(n-k)!k!}}_{=\binom{n}{k}} x^n = n! \sum_{k=0}^n \binom{n}{k}^2 x^n \end{aligned}$$

En résumé : $\forall x \in \mathbb{R}, \varphi^{(n)}(x) = n! \sum_{k=0}^n \binom{n}{k}^2 x^n \quad (\clubsuit)$

→ Epilogue. D'après $(\spadesuit)$, on a : $\varphi^{(n)}(1) = \frac{(2n)!}{n!}$. Et d'après $(\clubsuit)$, on a : $\varphi^{(n)}(1) = n! \sum_{k=0}^n \binom{n}{k}^2$. On en déduit que :

$$n! \sum_{k=0}^n \binom{n}{k}^2 = \frac{(2n)!}{n!} \quad \text{d'où :} \quad \sum_{k=0}^n \binom{n}{k}^2 = \frac{(2n)!}{n!n!} \quad \text{soit finalement :} \quad \sum_{k=0}^n \binom{n}{k}^2 = \binom{2n}{n}.$$

7. Puisque c'est une fonction polynomiale, elle est dérivable sur $\mathbb{R}$ "autant de fois qu'on le souhaite". Une telle fonction sera dite **de classe** $\mathcal{C}^\infty$ sur $\mathbb{R}$.

8. On applique la formule mentionnée dans le programme de colle, et dont vous devez pouvoir retrouver l'expression assez rapidement : $[x^N]^{(n)} = \frac{N!}{(N-n)!} x^{N-n}$ si $n \in \llbracket 0, N \rrbracket$.

5.6 Fonctions numériques usuelles

5.6.1 Fonction valeur absolue

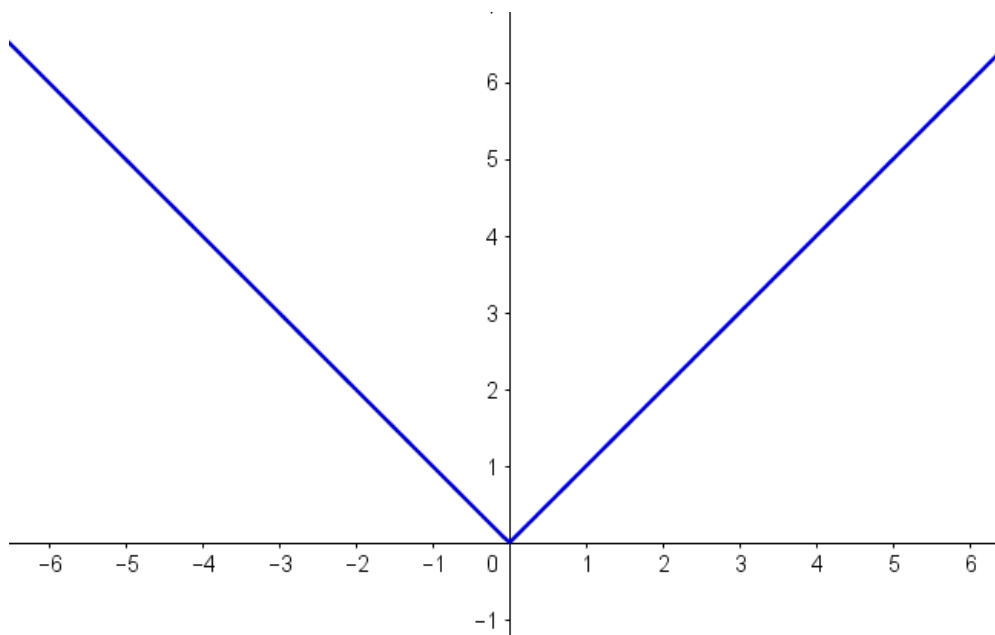
DÉFINITION 5.14 - La fonction **valeur absolue** est la fonction qui à tout réel x associe le réel $|x|$.

PROPRIÉTÉ 5.11 - (**Propriétés de la fonction valeur absolue**). La fonction valeur absolue est :

- 1/ définie sur $\mathbb{R}$;
- 2/ paire :
- 3/ croissante sur $\mathbb{R}_+$, décroissante sur $\mathbb{R}_-$;
- 4/ continue sur $\mathbb{R}$;
- 5/ dérivable sur $\mathbb{R}^*$, et non dérivable en 0 ;
- 6/ telle que : $\lim_{x \rightarrow +\infty} |x| = +\infty$ et $\lim_{x \rightarrow -\infty} |x| = +\infty$.

PREUVE. Triviale. ☕

Courbe représentative de la fonction valeur absolue.



5.6.2 Fonctions exponentielle et logarithme népérien

La fonction logarithme népérien

► **Dérivabilité** : la fonction $x \mapsto \ln x$ est définie et dérivable sur $\mathbb{R}_+^*$. Sa dérivée est donnée par :

$$\forall x \in \mathbb{R}_+^*, \quad \ln' x = \frac{1}{x}$$

► **Sens de variation** : strictement croissante sur $\mathbb{R}_+^*$.

► **Limites aux bornes** :

$$\lim_{x \rightarrow +\infty} \ln x = +\infty \quad \text{et} \quad \lim_{x \xrightarrow{x>0} 0} \ln x = -\infty$$

(asymptote verticale d'équation $x = 0$).

► **Tableau de variation** :

x	0	$+\infty$
$\ln x$	$-\infty$	$+\infty$

► **Bijection** : la fonction $\ln$ induit une bijection de $\mathbb{R}_+^*$ vers $\mathbb{R}$ et de plus :

$$\forall (x, y) \in (\mathbb{R}_+^*)^2, \quad \ln(x \times y) = \ln(x) + \ln(y)$$

Enfin vous savez que ces deux fonctions sont réciproques l'une de l'autre, ce qui se traduit par les identités suivantes (*et on prendra bien garde aux domaines de validité respectifs de ces identités*) :

$$\forall x \in \mathbb{R}_+^*, \quad e^{\ln x} = x \quad \text{et} \quad \forall x \in \mathbb{R}, \quad \ln(e^x) = x$$

Sur la figure ci-contre sont tracées les courbes représentatives des fonctions logarithme népérien et exponentielle, dans un repère orthonormé du plan ; ces courbes sont symétriques par rapport à la droite d'équation $y = x$.

La fonction exponentielle

► **Dérivabilité** : la fonction $x \mapsto e^x$ est définie et dérivable sur $\mathbb{R}$. Sa dérivée est donnée par :

$$\forall x \in \mathbb{R}, \quad e'(x) = e^x$$

► **Sens de variation** : strictement croissante sur $\mathbb{R}$.

► **Limites aux bornes** :

$$\lim_{x \rightarrow +\infty} e^x = +\infty \quad \text{et} \quad \lim_{x \rightarrow -\infty} e^x = 0$$

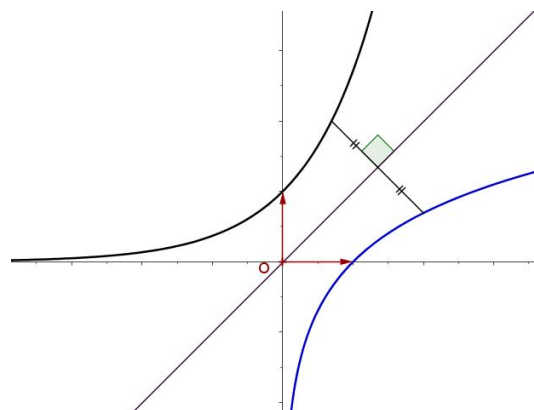
(asymptote horizontale d'équation $y = 0$).

► **Tableau de variation** :

x	$-\infty$	$+\infty$
e^x	0	$+\infty$

► **Bijection** : la fonction $\exp$ induit une bijection de $\mathbb{R}$ vers $\mathbb{R}_+^*$ et de plus :

$$\forall (x, y) \in \mathbb{R}^2, \quad e^{x+y} = e^x \times e^y$$



5.6.3 Fonctions trigonométriques hyperboliques

► La fonction cosinus hyperbolique (ch).

La fonction ch est **définie** sur $\mathbb{R}$ par : $\text{ch} : \mathbb{R} \longrightarrow \mathbb{R}$. ch est **paire**.

$$x \longmapsto \frac{e^x + e^{-x}}{2}$$

La fonction ch est **dérivable** (et même de classe $\mathcal{C}^1$) sur $\mathbb{R}$, et $\text{ch}' = \text{sh}$. Il s'ensuit que c'est une fonction strictement décroissante sur $\mathbb{R}_-$, et strictement croissante sur $\mathbb{R}^+$. Elle admet donc un minimum en 0, égal à 1. Pour finir, on a : $\lim_{x \rightarrow +\infty} \text{ch}(x) = +\infty$, et (par parité) : $\lim_{x \rightarrow -\infty} \text{ch}(x) = +\infty$.

La tangente à la courbe représentative $\mathcal{C}_{\text{ch}}$ au point d'abscisse 0 a pour équation $y = 1$ (tangente horizontale), et $\mathcal{C}_{\text{ch}}$ est située au-dessus de cette tangente (puisque ch est minorée par 1).

► La fonction sinus hyperbolique (sh).

La fonction sh est **définie** sur $\mathbb{R}$ par : $\text{sh} : \mathbb{R} \longrightarrow \mathbb{R}$. La fonction sh est **impaire**.

$$x \longmapsto \frac{e^x - e^{-x}}{2}$$

La fonction sh est **dérivable** (et même de classe $\mathcal{C}^1$) sur $\mathbb{R}$, et $\text{sh}' = \text{ch}$. Il s'ensuit que c'est une fonction strictement croissante sur $\mathbb{R}$. Enfin : $\lim_{x \rightarrow +\infty} \text{sh}(x) = +\infty$, et (comme sh est impaire) : $\lim_{x \rightarrow -\infty} \text{sh}(x) = -\infty$.

La tangente à la courbe représentative $\mathcal{C}_{\text{sh}}$ au point d'abscisse 0 a pour équation $y = x$ (première bissectrice des axes), et $\mathcal{C}_{\text{sh}}$ est située au-dessus (*resp.* en-dessous) de cette tangente sur $\mathbb{R}_+$ (*resp.* sur $\mathbb{R}_-$).⁹

► La fonction tangente hyperbolique (th) — (HP).

La fonction th est **définie** sur $\mathbb{R}$ par : $\text{th} : \mathbb{R} \longrightarrow \mathbb{R}$. La fonction th est **impaire**.

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

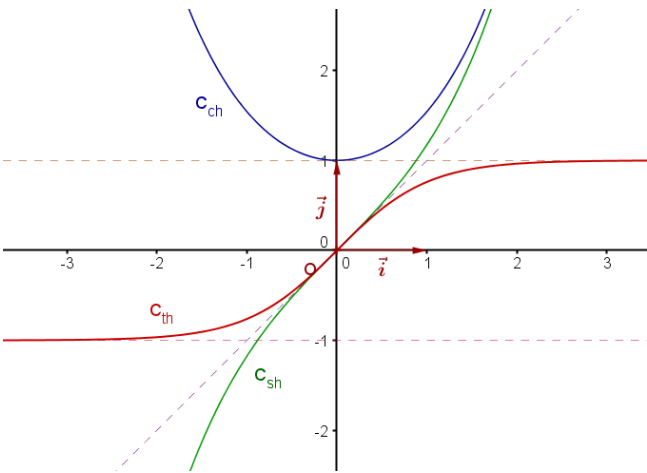
La fonction th est **dérivable** (et même de classe $\mathcal{C}^1$) sur $\mathbb{R}$, et $\text{th}' = \frac{1}{\text{ch}^2} = 1 - \text{th}^2$. Il s'ensuit que c'est une fonction strictement croissante sur $\mathbb{R}$. Enfin : $\lim_{x \rightarrow +\infty} \text{th}(x) = 1$, et (comme th est impaire) : $\lim_{x \rightarrow -\infty} \text{th}(x) = -1$.

La tangente à la courbe représentative $\mathcal{C}_{\text{th}}$ au point d'abscisse 0 a pour équation $y = x$ (première bissectrice des axes), et $\mathcal{C}_{\text{th}}$ est située en-dessous (*resp.* au-dessus) de cette tangente sur $\mathbb{R}_+$ (*resp.* sur $\mathbb{R}_-$).¹⁰

x	$-\infty$	0	$+\infty$
$\text{ch}'(x)$	-	0	+
$\text{ch}(x)$	$+\infty$	1	$+\infty$

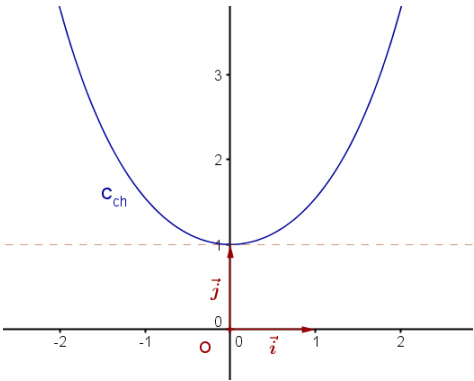
x	$-\infty$	0	$+\infty$
$\text{sh}'(x)$		+	
$\text{sh}(x)$	$-\infty$	0	$+\infty$

x	$-\infty$	0	$+\infty$
$\text{th}'(x)$		+	
$\text{th}(x)$	-1	0	1



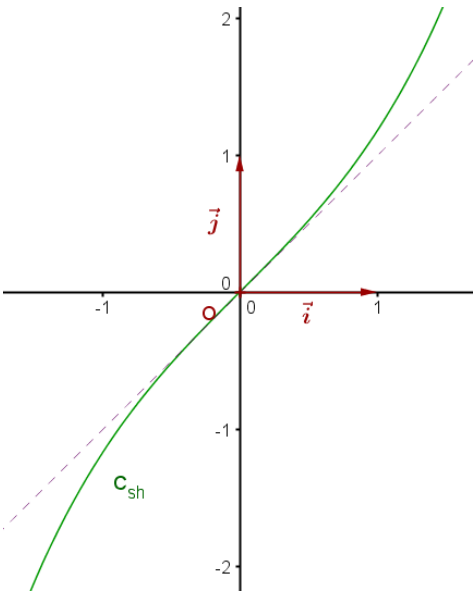
COSINUS HYPERBOLIQUE (CH OU COSH)

x	$-\infty$	0	$+\infty$
$ch'(x)$	$-$	0	$+$
$ch(x)$	$+\infty$	1	$+\infty$

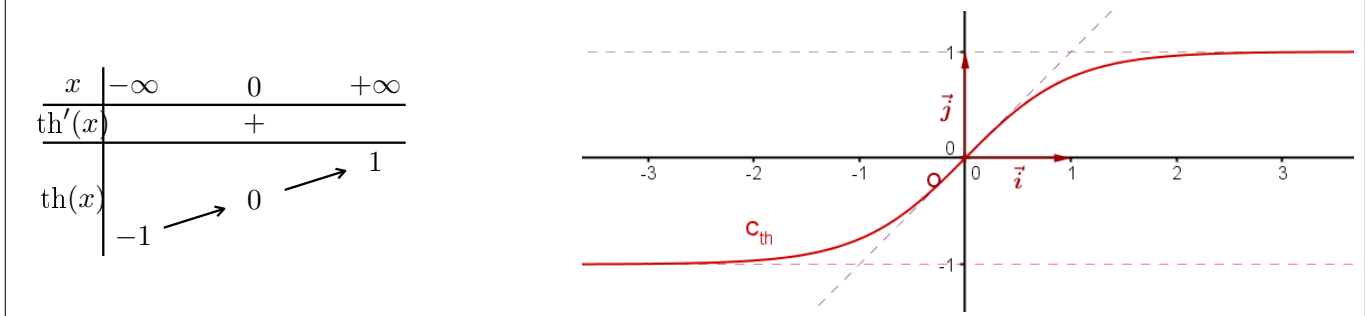


SINUS HYPERBOLIQUE (SH OU SINH)

x	$-\infty$	0	$+\infty$
$sh'(x)$	$+$	$+$	$+$
$sh(x)$	$-\infty$	0	$+\infty$



TANGENTE HYPERBOLIQUE (TH OU TANH)



5.6.4 Fonctions logarithmes de base a

DÉFINITION 5.15 - Soit a un réel strictement positif et différent de 1. On appelle fonction **logarithme de base a** et on note $\log_a$ la fonction définie sur $\mathbb{R}_+^*$ par :

$$\forall x \in \mathbb{R}_+^*, \quad \log_a(x) = \frac{\ln x}{\ln a}$$

Si l'enseignement de la Chimie n'a pas trop changé depuis mon époque, il me semble que vous devez y utiliser à tour de bras la fonction logarithme en base 10 (notée donc $\log_{10}$) pour calculer des pH par exemple.

On peut par ailleurs observer que la fonction $\log_a$ est obtenue en multipliant $\ln$ par une constante : $1/\ln a$. Or cette constante peut être strictement positive (lorsque $a > 1$) ou strictement négative (lorsque $a < 1$). Deux cas peuvent donc se présenter :

☞ Si $a > 1$: la fonction $\log_a$ est strictement croissante sur $\mathbb{R}_+^*$ et on a

$$\lim_{x \rightarrow 0} \log_a x = -\infty \quad \text{et} \quad \lim_{x \rightarrow +\infty} \log_a x = +\infty$$

☞ Si $0 < a < 1$: la fonction $\log_a$ est strictement **décroissante** sur $\mathbb{R}_+^*$ et on a

$$\lim_{x \rightarrow 0} \log_a x = +\infty \quad \text{et} \quad \lim_{x \rightarrow +\infty} \log_a x = -\infty$$

En revanche dans tous les cas, la fonction $\log_a$ réalise une bijection qui “transforme les produits en sommes”. Plus précisément :

PROPRIÉTÉ 5.12 - Soit a un réel strictement positif, tel que $a \neq 1$.

La fonction $\log_a : x \mapsto \log_a x$ réalise une bijection $\mathbb{R}_+^*$ vers $\mathbb{R}$, et en outre :

$$\forall (x, y) \in (\mathbb{R}_+^*)^2, \quad \log_a(x \times y) = \log_a(x) + \log_a(y)$$

PREUVE. Triviale. ☕

5.6.5 Fonctions exponentielles de base a

DÉFINITION 5.16 - Soit a un réel strictement positif et différent de 1. On appelle fonction **exponentielle de base a** et on note $\exp_a$ fonction définie sur $\mathbb{R}$ en posant :

$$\forall x \in \mathbb{R}, \quad a^x = e^{x \ln a}$$

Remarque. En particulier : $\forall x \in \mathbb{R}, \quad a^x > 0$.

DÉFINITION ALTERNATIVE. Soit a un réel strictement positif et différent de 1. On appelle fonction **exponentielle de base a** et on note $\exp_a$ la bijection réciproque de la fonction $\log_a$. Autrement dit, c'est la fonction qui associe à tout réel x l'unique solution de l'équation (d'inconnue X) : $\log_a X = x$. Avec des quantificateurs :

$$\forall x \in \mathbb{R}, \quad y = \exp_a x \iff x = \log_a y$$

En particulier, $\exp_a$ réalise une bijection de $\mathbb{R}$ sur $\mathbb{R}_+^*$.

PROPRIÉTÉ 5.13 - Soit a un réel strictement positif, et $a \neq 1$.

La fonction $\exp_a : x \mapsto a^x$ est dérivable sur $\mathbb{R}$ et :

$$\forall x \in \mathbb{R}, \quad (\exp_a)'(x) = (\ln a) \times a^x$$

PREUVE. Triviale. ☕

Cette propriété précédente permet d'obtenir aisément le sens de variation des fonctions exponentielles de base a (voir page suivante).

Etude de $x \mapsto a^x$ ($0 < a < 1$)

► **Sens de variation** : strictement décroissante sur $\mathbb{R}$.

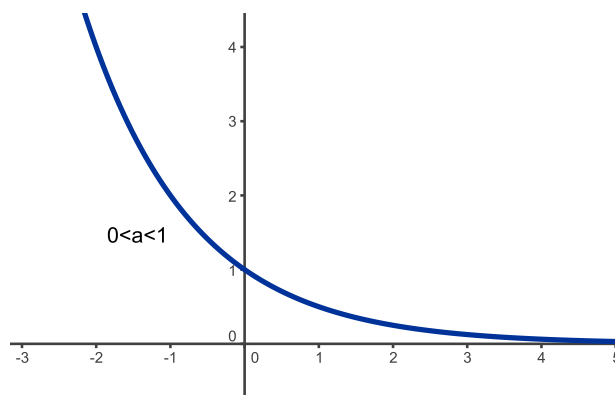
► **Limites aux bornes** :

$$\lim_{x \rightarrow -\infty} a^x = +\infty \quad \text{et} \quad \lim_{x \rightarrow +\infty} a^x = 0$$

► **Tableau de variation** :

x	$-\infty$	$+\infty$
$(\exp_a)'$		
a^x	$+\infty$	0

► **Exemple de courbe représentative** :



Etude de $x \mapsto a^x$ ($a > 1$)

► **Sens de variation** : strictement croissante sur $\mathbb{R}$.

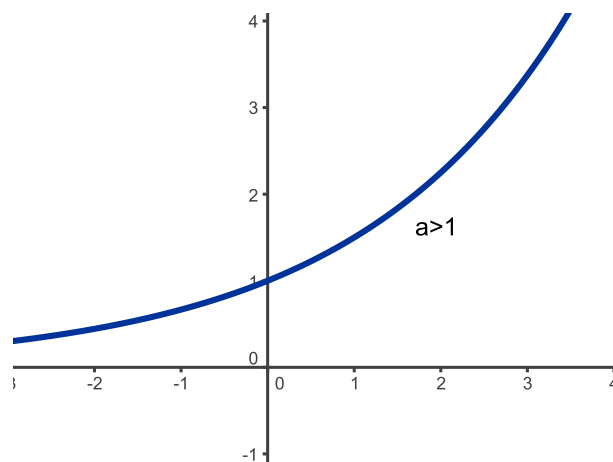
► **Limites aux bornes** :

$$\lim_{x \rightarrow -\infty} a^x = 0 \quad \text{et} \quad \lim_{x \rightarrow +\infty} a^x = +\infty$$

► **Tableau de variation** :

x	$-\infty$	$+\infty$
$(\exp_a)'$		
a^x	0	$+\infty$

► **Exemple de courbe représentative** :



Pour conclure, on notera encore que pour tout réel a (strictement positif et $\neq 1$), la fonction exponentielle de base a réalise une bijection de $\mathbb{R}$ vers $\mathbb{R}_+^*$, et “transforme les sommes en produits”.

5.6.6 Fonctions puissances

DÉFINITION 5.17 - Soit a un réel (quelconque). On appelle fonction **puissance d'exposant a** et nous noterons P_a la fonction définie sur $\mathbb{R}_+^*$ définie en posant :

$$\forall x \in \mathbb{R}_+^*, P_a(x) = x^a = e^{a \ln x}$$

Etude de $x \mapsto x^a$ ($a > 0$)

► **Sens de variation** : strictement croissante sur $\mathbb{R}_+^*$.

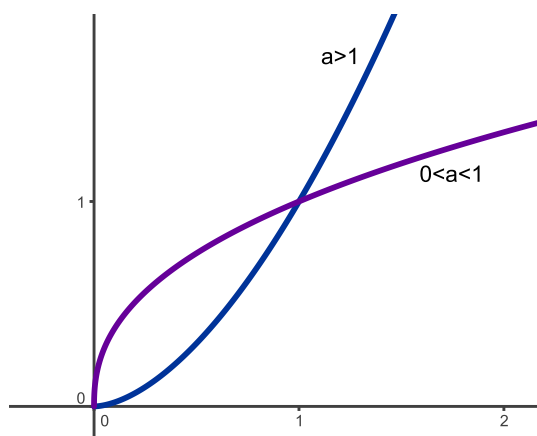
► **Limites aux bornes** :

$$\lim_{x \xrightarrow{>} 0} x^a = 0 \quad \text{et} \quad \lim_{x \rightarrow +\infty} x^a = +\infty$$

► **Tableau de variation** :

x	0	$+\infty$
$P'_a(x)$		+
x^a	0	$+\infty$

► **Exemples de courbes représentatives** :



Etude de $x \mapsto x^a$ ($a < 0$)

► **Sens de variation** : strictement décroissante sur $\mathbb{R}_+^*$.

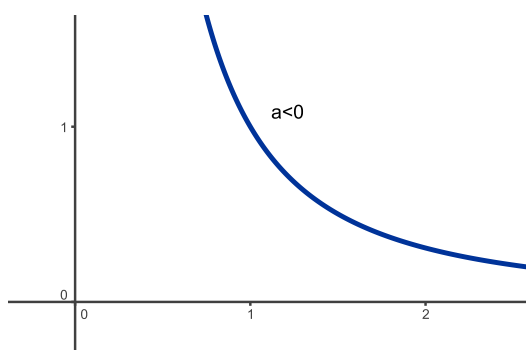
► **Limites aux bornes** :

$$\lim_{x \xrightarrow{>} 0} x^a = +\infty \quad \text{et} \quad \lim_{x \rightarrow +\infty} x^a = 0$$

► **Tableau de variation** :

x	0	$+\infty$
$P'_a(x)$		-
x^a	$+\infty$	0

► **Exemple de courbe représentative** :



Remarques. Soit n un entier naturel strictement positif, et x un réel strictement positif. On appelle racine $n^{\text{ième}}$ de x , et on note $\sqrt[n]{x}$ le nombre $x^{\frac{1}{n}}$. C'est un cas particulier des fonctions x^a étudiées ci-dessus.

D'autre part, le graphique suggère que dans le cas $a > 0$, la fonction $P_a : x \mapsto x^a$ peut être *prolongée par continuité en 0* en posant $P_a(0) = 0$. Ce choix est rendu légitime par le fait que si $a > 0$, la limite en 0 (à droite) de la fonction P_a est égale à 0.

Remarque : concernant les deux paragraphes précédents, le point-clef (qui permet de retrouver les propriétés des fonctions puissances et exponentielles) est

$$\forall a \in \mathbb{R}_+^*, \forall b \in \mathbb{R}, a^b = e^{b \ln(a)}$$

5.7 Epilogue

5.7.1 Formulaire des dérivées usuelles

$f(x)$	$f'(x)$
x^α	$\alpha x^{\alpha-1}$
$\frac{1}{x}$	$-\frac{1}{x^2}$
$\sqrt{x}$	$\frac{1}{2\sqrt{x}}$
$\ln(x)$	$\frac{1}{x}$
e^x	e^x
$\cos(x)$	$-\sin(x)$
$\sin(x)$	$\cos(x)$
$\tan(x)$	$1 + \tan^2(x) = \frac{1}{\cos^2(x)}$
$\operatorname{ch}(x)$	$\operatorname{sh}(x)$
$\operatorname{sh}(x)$	$\operatorname{ch}(x)$
$\arccos(x)$	$-\frac{1}{\sqrt{1-x^2}}$
$\arcsin(x)$	$\frac{1}{\sqrt{1-x^2}}$
$\arctan(x)$	$\frac{1}{1+x^2}$

$f(x)$	$f'(x)$
$(v \circ u)(x)$	$v'(u(x)) \times u'(x)$
$u(ax+b)$	$au'(ax+b)$
$\frac{1}{v(x)}$	$-\frac{v'(x)}{v^2(x)}$
$u^\alpha(x) \ (\alpha \neq 0)$	$\alpha u^{\alpha-1}(x)u'(x)$
$u^2(x)$	$2u(x)u'(x)$
$\sqrt{u(x)}$	$\frac{u'(x)}{2\sqrt{u(x)}}$
$\ln(u(x))$	$\frac{u'(x)}{u(x)}$
$e^{u(x)}$	$u'(x)e^{u(x)}$
$\cos(u(x))$	$-u'(x)\sin(u(x))$
$\sin(u(x))$	$u'(x)\cos(u(x))$
$\tan(u(x))$	$\frac{u'(x)}{\cos^2(u(x))}$
$\operatorname{ch}(u(x))$	$u'(x)\operatorname{sh}(u(x))$
$\operatorname{sh}(u(x))$	$u'(x)\operatorname{ch}(u(x))$

5.7.2 Formulaire de trigonométrie hyperbolique (HP)

Note : les formules suivantes sont valables pour tout nombre réel x

$$\textbf{Relation fondamentale : } \quad \text{ch}^2(x) - \text{sh}^2(x) = 1$$

$$\text{ch}(-x) = \text{ch}(x)$$

$$\text{sh}(-x) = -\text{sh}(x)$$

Les formules suivantes sont hors-programme.

$$\textbf{Formules d'addition : } \left\{ \begin{array}{l} \text{ch}(a+b) = \text{ch}(a)\text{ch}(b) + \text{sh}(a)\text{sh}(b) \\ \text{sh}(a+b) = \text{sh}(a)\text{ch}(b) + \text{sh}(b)\text{ch}(a) \\ \text{th}(a+b) = \frac{\text{tha} + \text{th}b}{1 + \text{th}a\text{th}b} \end{array} \right.$$

En utilisant la parité de la fonction ch , et l'imparité des fonctions sh et th , on obtient déjà :

$$\textbf{Formules "de soustraction" : } \left\{ \begin{array}{l} \text{ch}(a-b) = \text{ch}(a)\text{ch}(b) - \text{sh}(a)\text{sh}(b) \\ \text{sh}(a-b) = \text{sh}(a)\text{ch}(b) - \text{sh}(b)\text{ch}(a) \\ \text{th}(a-b) = \frac{\text{tha} - \text{th}b}{1 - \text{th}a\text{th}b} \end{array} \right.$$

On déduit de ces formules les suivantes :

$$\textbf{Formules de duplication : } \left\{ \begin{array}{l} \text{ch}(2a) = \text{ch}^2(a) + \text{sh}^2(a) = 2\text{ch}^2(a) + 1 = 1 + 2\text{sh}^2(a) \\ \text{sh}(2a) = 2\text{sh}(a)\text{ch}(a) \\ \text{th}(2a) = \frac{2\text{tha}}{1 + \text{th}^2a} \end{array} \right.$$

On peut également obtenir les formules ci-dessous à partir des formules d'addition :

$$\textbf{Formules de linéarisation : } \left\{ \begin{array}{l} \text{ch}(a)\text{ch}(b) = \frac{\text{ch}(a+b) + \text{ch}(a-b)}{2} \quad \text{d'où : } \text{ch}^2(a) = \frac{1 + \text{ch}(2a)}{2} \\ \text{sh}(a)\text{sh}(b) = \frac{\text{ch}(a+b) - \text{ch}(a-b)}{2} \quad \text{d'où : } \text{sh}^2(a) = \frac{\text{ch}(2a) - 1}{2} \\ \text{sh}(a)\text{ch}(b) = \frac{\text{sh}(a+b) + \text{sh}(a-b)}{2} \quad \text{d'où : } \text{sh}(a)\text{ch}(a) = \frac{\text{sh}(2a)}{2} \end{array} \right.$$

5.7.3 Utilisation des puissances 1 : étude d'un Sup

Exercice : Déterminer $\sup_{n \in \mathbb{N}^*} \sqrt[n]{n}$.

Notons que pour tout entier naturel non nul : $\sqrt[n]{n} = n^{1/n} = e^{\ln(n)/n}$. Définissons donc la fonction f sur $\mathbb{R}_+^*$ en posant : $\forall x > 0, f(x) = e^{\ln(x)/x}$.

La fonction f est dérivable sur $\mathbb{R}_+^*$, et : $\forall x > 0, f'(x) = \left(\frac{1 - \ln(x)}{x^2} \right) e^{\ln(x)/x}$. Par suite le réel $f'(x)$ est du signe de $1 - \ln(x)$; celui-ci est positif $]0; e]$, négatif sur $[e; +\infty[$.

La fonction f est donc croissante sur $]0; e]$ et décroissante sur $[e; +\infty[$. Elle admet donc un maximum en e (qui vaut $e^{1/e}$).

En se restreignant aux valeurs $f(n)$, on a donc : $\sup_{n \in \mathbb{N}^*} \sqrt[n]{n} = f(2) = \sqrt{2}$ ou $\sup_{n \in \mathbb{N}^*} \sqrt[n]{n} = f(3) = \sqrt[3]{3}$.

On note que : $(\sqrt{2})^6 = 8$ tandis que $(\sqrt[3]{3})^6 = 9$. Puisque la fonction $x \mapsto x^6$ est croissante sur $\mathbb{R}_+$ (et même strictement croissante sur $\mathbb{R}_+^*$), on en déduit que : $\sqrt{2} < \sqrt[3]{3}$.

Conclusion : $\sup_{n \in \mathbb{N}^*} \sqrt[n]{n} = \sqrt[3]{3}$

5.7.4 Utilisation des puissances 2 : résolution d'une inéquation

Exercice : Soit $a > 0$. Résoudre en fonction de a l'inéquation (I) $a^{(x^2)} < (\sqrt{a})^{7x-3}$.

Notons que : $(I) \iff e^{x^2 \ln(a)} < e^{(7x-3) \ln(\sqrt{a})} \iff e^{x^2 \ln(a)} < e^{\frac{(7x-3) \ln(a)}{2}} \iff x^2 \ln(a) < \frac{(7x-3) \ln(a)}{2}$

$$\iff \ln(a) (2x^2 - 7x + 3) < 0$$

Etudions le signe de l'expression $P(x) = 2x^2 - 7x + 3$. Ce polynôme de degré 2 possède deux racines (3 et 1/2), et on déduit donc son signe sur $\mathbb{R}$:

x	$-\infty$	$1/2$	3	$+\infty$
Signe de $P(x)$	+	-	+	+

On déduit de ce tableau de signes et du signe de la fonction $\ln$ sur $\mathbb{R}_+^*$ que :

- si $a > 1$, l'inéquation a pour solution $] -\infty; 1/2[\cup] 3; +\infty[$
- si $a = 1$, l'inéquation n'a aucune solution.
- si $0 < a < 1$, l'inéquation a pour solution $] 1/2; 3[$.

5.7.5 Retour sur la technique de “l’angle-moitié”

► Calcul de $\sum_{k=0}^n \operatorname{ch}(kx)$ (et $\sum_{k=0}^n \operatorname{sh}(kx)$) pour tout $n \in \mathbb{N}$ et pour tout $x \in \mathbb{R}$

Soient n et k deux entiers naturels et x un réel. On a : $\operatorname{ch}(kx) = \frac{e^{kx} + e^{-kx}}{2}$.

Fixons x un réel arbitraire, et notons : $C(x) = \sum_{k=0}^n \operatorname{ch}(kx)$. En vertu de la remarque précédente :

$$C(x) = \sum_{k=0}^n \frac{e^{kx} + e^{-kx}}{2} = \frac{1}{2} (S_1 + S_2) \quad (\heartsuit), \text{ en ayant posé } S_1 = \sum_{k=0}^n e^{kx} \text{ et } S_2 = \sum_{k=0}^n e^{-kx} \quad 11$$

Calculons S_1 : si $e^x \neq 1$ (càd si $x \neq 0$), on a : $S_1 = \sum_{k=0}^n (e^x)^k = \frac{1 - e^{(n+1)x}}{1 - e^x}$. En appliquant la technique de “l’angle-moitié”, on obtient :

$$S_1 = \frac{e^{(n+1)x/2}}{e^{x/2}} \times \frac{e^{-(n+1)x/2} - e^{(n+1)x/2}}{e^{-x/2} - e^{x/2}} = e^{nx/2} \frac{-2\operatorname{sh}\left(\frac{(n+1)x}{2}\right)}{-2\operatorname{sh}\left(\frac{x}{2}\right)} \text{ soit : } S_1 = e^{nx/2} \frac{\operatorname{sh}\left(\frac{(n+1)x}{2}\right)}{\operatorname{sh}\left(\frac{x}{2}\right)} \quad (\spadesuit)$$

Il résulte du calcul précédent ¹² que pour tout réel x non nul :

$$S_2 = \sum_{k=0}^n e^{-kx} = e^{-nx/2} \frac{\operatorname{sh}\left(\frac{(n+1)x}{2}\right)}{\operatorname{sh}\left(\frac{x}{2}\right)} \quad (\clubsuit)$$

D’après $(\heartsuit)$, $(\spadesuit)$ et $(\clubsuit)$ on a pour tout réel x non nul :

$$\sum_{k=0}^n \operatorname{ch}(kx) = \frac{1}{2} (e^{nx/2} + e^{-nx/2}) \frac{\operatorname{sh}\left(\frac{(n+1)x}{2}\right)}{\operatorname{sh}\left(\frac{x}{2}\right)} \iff \forall x \in \mathbb{R}^*, \quad \sum_{k=0}^n \operatorname{ch}(kx) = \operatorname{ch}\left(\frac{nx}{2}\right) \frac{\operatorname{sh}\left(\frac{(n+1)x}{2}\right)}{\operatorname{sh}\left(\frac{x}{2}\right)}$$

Par ailleurs, pour $x = 0$: $C(0) = \sum_{k=0}^n \operatorname{ch}(k0) = \sum_{k=0}^n 1 = n + 1$.

$$\textbf{Conclusion.} \quad \forall n \in \mathbb{N}, \quad \sum_{k=0}^n \operatorname{ch}(kx) = \begin{cases} \operatorname{ch}\left(\frac{nx}{2}\right) \frac{\operatorname{sh}\left(\frac{(n+1)x}{2}\right)}{\operatorname{sh}\left(\frac{x}{2}\right)} & \text{si } x \in \mathbb{R}^* \\ n + 1 & \text{si } x = 0 \end{cases}$$

Les calculs précédents et quelques changements de signes aux bonnes places permettent d’obtenir :

$$\forall n \in \mathbb{N}, \quad \sum_{k=0}^n \operatorname{sh}(kx) = \begin{cases} \operatorname{sh}\left(\frac{nx}{2}\right) \frac{\operatorname{sh}\left(\frac{(n+1)x}{2}\right)}{\operatorname{sh}\left(\frac{x}{2}\right)} & \text{si } x \in \mathbb{R}^* \\ 0 & \text{si } x = 0 \end{cases}$$

11. Avec les mêmes notations, on a clairement : $\sum_{k=0}^n \operatorname{sh}(kx) = \frac{1}{2} (S_1 - S_2)$.

12. Et du fait que la fonction sh est impaire.

5.7.6 Résolution d'une inéquation avec valeurs absolues

Résoudre l'équation : (E) $\ln(x+2) + \ln(4-x) = \ln|8x-8|$.

1) Ensemble de résolution de l'équation : en premier lieu, on détermine pour quelles valeurs de x les expressions intervenant dans l'équation sont définies simultanément.

- $\ln(x+2)$ est défini SSI $x+2 > 0$ c'est-à-dire pour : $x > -2$.
- $\ln(4-x)$ est défini SSI $4-x > 0$ c'est-à-dire pour : $x < 4$.
- $\ln|8x-8|$ est défini SSI $|8x-8|$ c'est-à-dire pour : $x \neq 1$.

L'ensemble de résolution de l'équation est donc : $E =]-2, 1[\cup]1, 4[$

2) Résolution de (E)

a) Résolution sur $I_1 =]1, 4[$: pour tout $x \in I_1$, on a : $\ln|8x-8| = \ln(8x-8)$. Par suite :

$$(E) \iff \ln(x+2) + \ln(4-x) = \ln(8x-8) \iff \ln(-x^2 + 2x + 8) = \ln(8x-8) \iff x^2 + 6x - 16 = 0$$

2 est racine évidente de cette équation ; la seconde est donc -8 (le produit des racines étant égal à " c/a "). Dans l'intervalle I_1 , l'équation (E) possède une unique solution qui est 2.

b) Résolution sur $I_2 =]-2, 1[$: pour tout $x \in I_2$, on a : $\ln|8x-8| = \ln(8-8x)$. Par suite :

$$(E) \iff \ln(x+2) + \ln(4-x) = \ln(8-8x) \iff \ln(-x^2 + 2x + 8) = \ln(8-8x) \iff x^2 - 10x = 0$$

Cette équation possède deux racines : 0 et 10. Dans l'intervalle I_2 , l'équation (E) possède une unique solution qui est 0.

Conclusion : l'équation $\ln(x+2) + \ln(4-x) = \ln|8x-8|$ possède exactement deux solutions : 0 et 2.

Chapitre 6

Applications

6.1 Généralités

DÉFINITION 6.1 - (“naïve”). Soient E et F deux ensembles. Une **application** de E dans F est un procédé qui permet d’associer à tout élément x de l’ensemble E un unique élément noté $f(x)$ de l’ensemble F , appelé **image** de x par f .

En fait il est plus précis de définir une application par identification avec son graphe, explicitement :

DÉFINITION 6.2 - (“la vraie”). Soient E et F deux ensembles. Une **application** de E dans F est une partie Γ_f de $E \times F$ telle que :

$$\forall x \in E, \exists ! y \in F, (x, y) \in \Gamma_f$$

Ainsi, pour tout $x \in E$, l’unique élément y de F tel que le couple (x, y) appartienne à Γ_f est noté $f(x)$ et est appelé **image** de x par f .

Exemples : les fonctions usuelles sont des applications (en général d’un intervalle I de $\mathbb{R}$, vers $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$) ; le fait d’associer à un polynôme $P \in \mathbb{R}_n[X]$ son polynôme dérivé $P' \in \mathbb{R}_{n-1}[X]$ est une application ; la conjugaison complexe ($z \mapsto \bar{z}$) est une application (de $\mathbb{C}$ dans $\mathbb{C}$) ; *etc...*

Remarque : le point-clef est qu’une application n’est pas “qu’une formule”. C’est la donnée d’un nom (f souvent), d’un ensemble de définition (ou domaine), d’un ensemble d’arrivée (ou codomaine), et de la description du procédé qui à un élément de l’ensemble de définition associe son image. Une des notations ci-dessous doit donc être respectée lorsque l’on se réfère à une application :

$$f : E \longrightarrow F \quad \text{ou} \quad f : x \in E \longmapsto f(x) \in F$$
$$x \longmapsto f(x)$$

Notation. On note F^E l’ensemble des applications de E dans F .

DÉFINITION 6.3 - . Soit E un ensemble. L’**identité** de E est l’application qui à tout élément x de E associe x .

Elle est notée id_E .

DÉFINITION 6.4 - Soit $f \in F^E$.

➤ L'**image de f** , notée $\text{Im}(f)$ ou $f(E)$ est la partie de F suivante :

$$f(E) = \{f(x), x \in E\} \text{ ou } f(E) = \{y \in F, \exists x \in E, y = f(x)\}$$

➤ Plus généralement, lorsque A est une partie de E , l'**image (directe) de A par f** , notée $f(A)$ est la partie de F suivante :

$$f(A) = \{f(x), x \in A\} \text{ ou } f(A) = \{y \in F, \exists x \in A, y = f(x)\}$$

➤ Lorsque B est une partie de F , l'**image réciproque de B par f** , notée $f^{-1}(B)$ est la partie de E suivante :

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

Exemples. Cf cours.

DÉFINITION 6.5 - Soit $f \in F^E$. Pour toute partie A de E , on appelle **restriction de f à A** et on note $f|_A$ l'application de F^A définie en posant : $\forall x \in A, f|_A(x) = f(x)$. Dans cette situation, on dit que f est un **prolongement** de $f|_A$ à E .

DÉFINITION 6.6 - Soient E, F et G trois ensembles ; $f : E \longrightarrow F$ et $g : F \longrightarrow G$ deux applications. On définit la **composée de f et g** et on note $g \circ f$ l'application de E dans G qui à tout élément x de E associe $g(f(x))$, ce que l'on écrit :

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

Exemple : on note $\exp : x \in \mathbb{R} \mapsto e^x \in \mathbb{R}_+^*$ et $\ln : x \in \mathbb{R}_+^* \mapsto \ln(x) \in \mathbb{R}$. Alors $\exp \circ \ln = \text{id}_{\mathbb{R}_+^*}$; tandis que $\ln \circ \exp = \text{id}_{\mathbb{R}}$. Les composées $\exp \circ \ln$ et $\ln \circ \exp$ sont donc distinctes. Voir plus généralement la remarque ci-dessous.

Remarque : on a déjà observé (dans le cadre des similitudes) que **la composition des applications n'est pas commutative**. En revanche, la composition est associative :

PROPRIÉTÉ 6.1 - Soient E, F, G et H quatre ensembles, et $f \in F^E, g \in G^F, h \in H^G$ trois applications. On a : $h \circ (g \circ f) = (h \circ g) \circ f$.

PREUVE. Triviale. ☕

PROPRIÉTÉ 6.2 - Pour toute application $f : E \longrightarrow F$, on a $f \circ \text{id}_E = f$ et $\text{id}_F \circ f = f$ (l'identité est l'élément neutre pour la composition des applications).

PREUVE. Triviale. ☕

6.2 Applications injectives, surjectives et bijectives

DÉFINITION 6.7 - Soient E et F deux ensembles, f une application de E dans F . On dit que f est **injective** (*resp.* **surjective**, *resp.* **bijective**) si tout élément de F admet au plus un (*resp.* au moins un, *resp.* exactement un) antécédent par f dans E .

Traduction à l'aide de quantificateurs : soit $f : E \longrightarrow F$ une application.

$$\triangleright [f \text{ injective}] \iff [\forall (x, x') \in E^2, f(x) = f(x') \implies x = x']$$

$$\triangleright [f \text{ surjective}] \iff [\forall y \in F, \exists x \in E, y = f(x)]$$

$$\triangleright [f \text{ bijective}] \iff [\forall y \in F, \exists! x \in E, y = f(x)]$$

Exemples : l'application $f : x \in \mathbb{R} \mapsto x^2 \in \mathbb{R}$ n'est ni injective ($f(2) = f(-2)$ par exemple), ni surjective (tout réel $y < 0$ n'a aucun antécédent par f). L'application $g : x \in \mathbb{R}^+ \mapsto x^2 \in \mathbb{R}$ est injective et non surjective. L'application $h : x \in \mathbb{R}^+ \mapsto x^2 \in \mathbb{R}^+$ est bijective.

La propriété ci-dessous signifie que les notions d'injectivité et de surjectivité sont stables par composition.

PROPRIÉTÉ 6.3 - Soient E, F et G trois ensembles ; $f : E \longrightarrow F$ et $g : F \longrightarrow G$ deux applications. Si f et g sont injectives (*resp.* surjectives), alors la composée $g \circ f$ est injective (*resp.* surjective).

PREUVE. Soient $f : E \longrightarrow F$ et $g : F \longrightarrow G$ deux applications.

► Supposons f et g injectives.

Soient x et x' deux éléments de E . Alors :

$$[(g \circ f)(x) = (g \circ f)(x')] \iff [g(f(x)) = g(f(x'))] \underset{g \text{ injective}}{\implies} [f(x) = f(x')] \underset{f \text{ injective}}{\implies} [x = x']$$

Ce qui prouve l'injectivité de $g \circ f$. **Conclusion** : si f et g sont injectives, alors $g \circ f$ est injective.

► Supposons à présent f et g surjectives.

Soit $z \in G$.

Alors, l'application g étant surjective : $\exists y \in F, g(y) = z$.

Et puisque f est surjective : $\exists x \in E, f(x) = y$.

En exploitant ces deux relations, on a : $g(f(x)) = z$.

Puisque z est un élément arbitraire de G , on vient d'établir que : $\forall z \in G, \exists x \in E, (g \circ f)(x) = z$.

Ce qui prouve la surjectivité de $g \circ f$. **Conclusion** : si f et g sont surjectives, alors $g \circ f$ est surjective. ☕

Par suite, la notion de bijectivité est elle aussi stable par composition.

COROLLAIRE 6.1 - Soient E, F et G trois ensembles ; $f : E \longrightarrow F$ et $g : F \longrightarrow G$ deux applications. Si f et g sont bijectives, alors la composée $g \circ f$ est bijective.

PREUVE. Conséquence immédiate de la propriété précédente. ☕

DÉFINITION 6.8 - Deux ensembles E et F sont **équipotents** lorsqu'il existe une bijection entre E et F . On le note $E \sim F$.

Exemples : $\mathbb{R}$ et $\mathbb{R}_+^*$ sont équipotents puisque la fonction exponentielle induit une bijection du premier ensemble vers le second.

DÉFINITION 6.9 - Un ensemble E est **dénombrable** lorsqu'il existe une bijection entre E et une partie de $\mathbb{N}$.

Exemples : l'ensemble des étudiants de MPSI est dénombrable, puisqu'il est en bijection avec $\llbracket 1, 45 \rrbracket$, qui est une partie de $\mathbb{N}$. $\mathbb{N}^*$ est dénombrable, puisque l'application $n \mapsto n - 1$ induit une bijection de $\mathbb{N}^*$ vers $\mathbb{N}$. Plus rigolo, $\mathbb{Z}$ est dénombrable, puisqu'il est en bijection avec $\mathbb{N}$ (ce qui peut sembler contre-intuitif au premier abord). Encore plus fort : $\mathbb{Q}$ est dénombrable (cf dernier exo de la feuille d'exos sur les applications + théorème de Cantor-Bernstein). Pour l'avenir (en Spé) : $\mathbb{R}$ n'est pas dénombrable.

Preuve de la dénombrabilité de $\mathbb{Z}$: on établit l'existence d'une bijection entre $\mathbb{Z}$ et $\mathbb{N}$, en considérant l'application :

$$f : \mathbb{N} \longrightarrow \mathbb{Z}$$

$$n \longmapsto \frac{-n}{2} \text{ si } n \text{ est pair ; } \frac{n+1}{2} \text{ sinon}$$

On définit également

$$g : \mathbb{Z} \longrightarrow \mathbb{N}$$

$$n \longmapsto -2n \text{ si } n \text{ est négatif ou nul ; } 2n - 1 \text{ sinon}$$

On vérifie ensuite aisément que $g \circ f = \text{id}_{\mathbb{N}}$ et $f \circ g = \text{id}_{\mathbb{Z}}$, d'où la conclusion.

6.3 Bijections réciproques

THÉORÈME 6.1 - Soient E et F deux ensembles, et $f : E \longrightarrow F$ une application.

f est bijective SSI il existe une application $g : F \longrightarrow E$ telle que $f \circ g = \text{id}_F$ et $g \circ f = \text{id}_E$.

PREUVE. ► Sens direct : supposons f bijective.

Puisque f est bijective, tout élément y de F admet un unique antécédent dans E par f , que nous noterons x_y .

On définit alors une application $g : F \longrightarrow E$ en posant (pour tout y de F) $g(y) = x_y$ l'unique antécédent de y par f .

Vérifions que $f \circ g = \text{id}_F$ et $g \circ f = \text{id}_E$.

Pour la première égalité : soit $y \in F$. Alors $(f \circ g)(y) = f(g(y)) = f(x_y) = y$ (puisque x_y est l'unique antécédent de y par f). Ainsi : $f \circ g = \text{id}_F$.

Passons à la seconde égalité : soit $x \in E$. Alors $(g \circ f)(x) = g(f(x)) = x$ (puisque, f étant bijective, x est l'unique antécédent de $f(x)$ par f). Ainsi : $g \circ f = \text{id}_E$.

Conclusion intermédiaire : si f est bijective, alors il existe une application $g : F \longrightarrow E$ telle que $f \circ g = \text{id}_F$ et $g \circ f = \text{id}_E$.

► **Réciproque** : supposons qu'il existe une application $g : F \longrightarrow E$ telle que $f \circ g = \text{id}_F$ et $g \circ f = \text{id}_E$.

Soient x et x' deux éléments de E tels que : $f(x) = f(x')$. On a alors : $g(f(x)) = g(f(x'))$ et puisque $g \circ f = \text{id}_E$, on en déduit $x = x'$. Ce qui prouve que f est injective.

Soit y un élément de E . On a : $y = \text{id}_E(y) = (f \circ g)(y) = f(g(y))$. Donc y admet un antécédent par f (qui est $g(y)$). Puisque y est un élément arbitraire de E dans ce petit raisonnement, on en déduit que f est surjective.

L'application f étant injective et surjective, elle est bijective.

Conclusion intermédiaire bis : s'il existe une application $g : F \longrightarrow E$ telle que $f \circ g = \text{id}_F$ et $g \circ f = \text{id}_E$, alors f est bijective.

Conclusion (des conclusions !) : f est bijective SSI il existe une application $g : F \longrightarrow E$ telle que

$$f \circ g = \text{id}_F \text{ et } g \circ f = \text{id}_E. \quad \text{☕}$$

PROPRIÉTÉ 6.4 - (Propriété, définition et notation). Lorsque f est bijective, l'application g du théorème ci-dessus est **unique**, appelée **bijection réciproque** de f , et notée f^{-1} .

PREUVE. Seule l'unicité est à établir. Soit donc $f \in F^E$ une bijection, et supposons qu'il existe deux applications g et h dans E^F telles que $f \circ g = \text{id}_F$ et $g \circ f = \text{id}_E$, et $f \circ h = \text{id}_F$ et $h \circ f = \text{id}_E$.

On a : $g = g \circ \text{id}_F = g \circ (f \circ h) = (g \circ f) \circ h = \text{id}_E \circ h = h. \quad \text{☕}$

PROPRIÉTÉ 6.5 - (Propriétés des bijections réciproques). Soient $f \in F^E$ et $g \in G^F$ deux applications.

1/ Si f est bijective, alors f^{-1} est bijective et $(f^{-1})^{-1} = f$.

2/ Si f et g sont bijectives, alors $(g \circ f)$ est bijective et $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$.

PREUVE. Repose sur des simples vérifications. ☕

Exemples.

- ➡ La fonction racine carrée (de $\mathbb{R}_+$ dans $\mathbb{R}_+$) est la bijection réciproque de la fonction carrée (de $\mathbb{R}^+$ dans $\mathbb{R}^+$).
- ➡ La fonction logarithme népérien (de $\mathbb{R}_+^*$ dans $\mathbb{R}$) est la bijection réciproque de la fonction exponentielle (de $\mathbb{R}$ dans $\mathbb{R}_+^*$).
- ➡ C'est l'occasion de rappeler que les fonctions carrée (de $\mathbb{C}$ dans $\mathbb{C}$) et exponentielle (de $\mathbb{C}$ dans $\mathbb{C}^*$) n'étant pas bijectives, elles n'admettent pas de réciproque ; dans ce contexte, il serait donc hautement déconseillé d'écrire un nombre complexe sous une " $\sqrt{}$ " ou derrière un " $\ln$ " !

- La fonction inverse (de $\mathbb{R}_+^*$ dans $\mathbb{R}_+^*$) est sa propre bijection réciproque (pour tout réel x non nul, l'inverse de l'inverse de x est x lui-même). Une bijection qui est sa propre réciproque est appelée une **involution**. D'autres exemples d'involution sont la conjugaison complexe, ou une symétrie centrale. Plus précisément :

DÉFINITION 6.10 - Soit E un ensemble. Une **involution** de E est une application f de E dans E telle que : $f \circ f = \text{id}_E$.

PROPRIÉTÉ 6.6 - Soit E un ensemble. Toute involution de E est bijective, et est sa propre bijection réciproque.

PREUVE. Corollaire immédiat de la définition d'involution et du théorème 6.1. ☕

6.4 Epilogue

6.4.1 Fonctions indicatrices

DÉFINITION 6.11 - Soient E un ensemble, et A une partie de E . La **fonction indicatrice de A** est l'application de $\{0, 1\}^E$ notée $\mathbb{1}_A$ définie en posant :

$$\forall x \in E, \mathbb{1}_A(x) = 1 \text{ si } x \in A \text{ et } \mathbb{1}_A(x) = 0 \text{ si } x \notin A$$

PROPRIÉTÉ 6.7 - Soient E un ensemble, et A et B deux parties de E . On a :

$$1) \mathbb{1}_{\bar{A}} = 1 - \mathbb{1}_A; \quad 2) \mathbb{1}_{A \cap B} = \mathbb{1}_A \times \mathbb{1}_B; \quad 3) \mathbb{1}_{A \cup B} = \mathbb{1}_A + \mathbb{1}_B - \mathbb{1}_{A \cap B}$$

PREUVE. Soient E un ensemble, et A et B deux parties de E .

1) Soit x un élément de E .

Si $x \in A$, alors $1 - \mathbb{1}_A(x) = 0$ et $\mathbb{1}_{\bar{A}}(x) = 0$ (puisque alors $x \notin \bar{A}$).

Sinon, on a $x \notin A$ d'où $1 - \mathbb{1}_A(x) = 1$ et $\mathbb{1}_{\bar{A}}(x) = 1$.

Conclusion. $\forall x \in E, \mathbb{1}_{\bar{A}}(x) = 1 - \mathbb{1}_A(x)$, soit : $\mathbb{1}_{\bar{A}} = 1 - \mathbb{1}_A$

2) Soit x un élément de E .

Si $x \in A \cap B$, alors $\mathbb{1}_{A \cap B}(x) = 1$ et $\mathbb{1}_A(x) \times \mathbb{1}_B(x) = 1$.

Sinon, on a $x \notin A$ ou $x \notin B$ d'où $\mathbb{1}_{A \cap B}(x) = 0$ et $\mathbb{1}_A(x) \times \mathbb{1}_B(x) = 0$ (puisque alors on a $\mathbb{1}_A(x) = 0$ ou $\mathbb{1}_B(x) = 0$).

Conclusion. $\forall x \in E, \mathbb{1}_{A \cap B}(x) = \mathbb{1}_A(x) \times \mathbb{1}_B(x)$, soit : $\mathbb{1}_{A \cap B} = \mathbb{1}_A \times \mathbb{1}_B$

3) Soit x un élément de E .

➤ Si $x \in A \cup B$, alors $\mathbb{1}_{A \cup B}(x) = 1$. Puis on distingue deux cas : $x \in A \cap B$ et $x \in (A \cup B) \setminus (A \cap B)$.

➤ Dans le premier cas ($x \in A \cap B$) : alors $\mathbb{1}_A(x) = 1, \mathbb{1}_B(x) = 1$ et $\mathbb{1}_{A \cap B}(x) = 1$. D'où : $\mathbb{1}_A(x) + \mathbb{1}_B(x) - \mathbb{1}_{A \cap B}(x) = 1$.

➤ Dans le second cas $(x \in (A \cup B) \setminus (A \cap B))$: on distingue deux sous-cas. Soit $(x \in A) \wedge (x \notin B)$, soit $(x \in B) \wedge (x \notin A)$. Dans le premier de ces sous-cas on a : $\mathbb{1}_A(x) = 1$, $\mathbb{1}_B(x) = 0$ et $\mathbb{1}_{A \cap B}(x) = 0$. D'où : $\mathbb{1}_A(x) + \mathbb{1}_B(x) - \mathbb{1}_{A \cap B}(x) = 1$. Le second se traite de manière analogue (ou en observant que A et B jouent des rôles symétriques).

En résumé : $\boxed{\forall x \in A \cup B, \mathbb{1}_{A \cup B}(x) = \mathbb{1}_A(x) + \mathbb{1}_B(x) - \mathbb{1}_{A \cap B}(x)}$

➤ Si $x \notin A \cup B$, alors $x \notin A$ et $x \notin B$. D'où : $\mathbb{1}_{A \cup B}(x) = 0$, $\mathbb{1}_A(x) = 0$, $\mathbb{1}_B(x) = 0$ et $\mathbb{1}_{A \cap B}(x) = 0$.

Il s'ensuit que : $\boxed{\forall x \in E \setminus A \cup B, \mathbb{1}_{A \cup B}(x) = \mathbb{1}_A(x) + \mathbb{1}_B(x) - \mathbb{1}_{A \cap B}(x)}$

Conclusion. $\forall x \in E, \mathbb{1}_{A \cup B}(x) = \mathbb{1}_A(x) + \mathbb{1}_B(x) - \mathbb{1}_{A \cap B}(x)$, soit : $\mathbb{1}_{A \cup B} = \mathbb{1}_A + \mathbb{1}_B - \mathbb{1}_{A \cap B}$



THÉORÈME 6.2 - L'application

$$\begin{aligned} \varphi : \mathcal{P}(E) &\longrightarrow \{0, 1\}^E \text{ est bijective.} \\ A &\longmapsto \mathbb{1}_A \end{aligned}$$

PREUVE. ➤ **Injectivité.** Pour deux parties A et B de E , on a : $[\mathbb{1}_A = \mathbb{1}_B] \iff [A = B]$ (lemme). Ce qui établit l'injectivité de φ_A .

Lemme. Soient A et B deux parties de E . On a : $[\mathbb{1}_A \leq \mathbb{1}_B] \iff [A \subset B]$.

Preuve du lemme : montrons l'implication $[\mathbb{1}_A \leq \mathbb{1}_B] \implies [A \subset B]$. On fait donc l'hypothèse que : $\mathbb{1}_A \leq \mathbb{1}_B$. Soit x un élément de A . Alors $\mathbb{1}_A(x) = 1$, et d'après notre hypothèse : $\mathbb{1}_B(x) = 1$, ce qui signifie que $x \in B$. D'où l'inclusion $A \subset B$, ce qui établit la première implication.

Réciproquement, supposons que $A \subset B$. Soit $x \in E$. Alors $x \in A$ ou $x \notin A$. Dans le premier cas : $x \in A$ donc $\mathbb{1}_A(x) = 1$; et comme $A \subset B$, on a aussi $x \in B$ donc $\mathbb{1}_B(x) = 1$. En particulier : $\mathbb{1}_A(x) \leq \mathbb{1}_B(x)$.

Dans le second cas : $x \notin A$ donc $\mathbb{1}_A(x) = 0$, donc $\mathbb{1}_A(x) \leq \mathbb{1}_B(x)$ (puisque une fonction indicatrice prend comme valeur 0 ou 1).

Par suite $\forall x \in E, \mathbb{1}_A(x) \leq \mathbb{1}_B(x)$. Ce qui établit la réciproque et achève la preuve du lemme.

➤ **Surjectivité.** Soit $f \in \{0, 1\}^E$. Posons $A = f^{-1}(\{1\})$, et montrons que $f = \mathbb{1}_A$.

Soit x un élément de E . Alors x est un antécédent de 1 ou de 0 par f .

— Si x est un antécédent de 1 par f , alors $x \in A$ donc $f(x) = 1$ et $\mathbb{1}_A(x) = 1$.

— Sinon, x est un antécédent de 0 par f , donc $x \notin A$ donc $f(x) = 0$ et $\mathbb{1}_A(x) = 0$.

Par conséquent : $\forall x \in E, f(x) = \mathbb{1}_A(x)$. Donc : $f = \mathbb{1}_A$. Ce qui prouve que φ est surjective.

Conclusion : l'application $\varphi : \mathcal{P}(E) \longrightarrow \{0, 1\}^E$ est bijective.

$$A \longmapsto \mathbb{1}_A$$



6.4.2 Deux pseudo-réciproques...

Les propriétés ci-dessous, qui constituent des exercices classiques, sont à rapprocher des propriétés affirmant que les notions d'injectivité et de surjectivité sont stables par composition.

PROPRIÉTÉ 6.8 - Soient E , F et G trois ensembles, et soient $f \in F^E$ et $g \in G^F$.

On a :

$$1/ [g \circ f \text{ injective}] \implies [f \text{ injective}]$$

$$2/ [g \circ f \text{ surjective}] \implies [g \text{ surjective}]$$

PREUVE. 1/ Soient x et x' deux éléments de E . Supposons que $f(x) = f(x')$. Alors : $(g \circ f)(x) = (g \circ f)(x')$. Puisque $g \circ f$ est injective par hypothèse, on en déduit que $x = x'$.

Finalement : $\forall (x, x') \in E^2, [f(x) = f(x')] \implies [x = x']$. Ce qui signifie que f est injective.

Ainsi : $[g \circ f \text{ injective}] \implies [f \text{ injective}]$.

2/ Soit $z \in G$. Puisque $g \circ f$ est surjective par hypothèse, il existe un élément x de E tel que : $g(f(x)) = z$. D'où $z \in g(F)$, ce qui signifie que g est surjective.

Ainsi : $[g \circ f \text{ surjective}] \implies [g \text{ surjective}]$. ☕

Chapitre 7

Fonctions circulaires réciproques

7.1 Fonctions à valeurs réelles bijectives

PROPRIÉTÉ 7.1 - Soit $f : I \longrightarrow \mathbb{R}$ une fonction définie sur un intervalle non vide de $\mathbb{R}$. Si f est strictement monotone sur I , alors f réalise une bijection de I vers $f(I)$.

PREUVE. La surjectivité provient de ce que l'on choisit $f(I)$ comme codomaine, et l'injectivité de la stricte monotonie. ☕

Exemple : la fonction $\exp$ réalise une bijection de $\mathbb{R}$ vers $\mathbb{R}_+^*$.

PROPRIÉTÉ 7.2 - La bijection réciproque d'une fonction strictement croissante (*resp.* décroissante) est strictement croissante (*resp.* décroissante).

PREUVE. On suppose que $f : I \longrightarrow f(I)$ est une fonction définie sur un intervalle I et à valeurs réelles, bijective et strictement croissante. Montrons que f^{-1} est strictement croissante.

Soient a et b deux réels de $f(I)$ tels que $a < b$.

Supposons $f^{-1}(a) \geq f^{-1}(b)$ (♠).

Alors, f étant croissante, on en déduit que : $f(f^{-1}(a)) \geq f(f^{-1}(b))$, soit $a \geq b$, ce qui est en contradiction avec l'hypothèse initiale ($a < b$). Ceci implique que l'assertion (♠) est fausse.

Par suite $f^{-1}(a) < f^{-1}(b)$. En résumé, on a établi l'implication : $\forall (a, b) \in (f(I))^2, (a < b) \implies (f^{-1}(a) < f^{-1}(b))$.

Conclusion : sous les hypothèses de l'énoncé, si f est strictement croissante, alors f^{-1} est strictement croissante.

La propriété "si f est bijective et strictement décroissante, alors f^{-1} est strictement décroissante" se déduit du raisonnement précédent, en modifiant un seul signe... ☕

PROPRIÉTÉ 7.3 - La bijection réciproque d'une fonction impaire est impaire.

Remarque : arrêtez-vous deux secondes pour vous demander pourquoi l'énoncé ci-dessus n'a aucune chance de tenir si l'on remplace "impaire" par "paire" !

PREUVE. ► Montrons que si f est impaire, alors f^{-1} l'est.

On suppose que $f : I \longrightarrow f(I)$ est une bijection impaire (en particulier, I est symétrique par rapport à zéro).

► Pour établir l'imparité de f^{-1} , on commence par établir que $f(I)$ est symétrique par rapport à zéro.

Si y est dans $f(I)$, alors il existe un élément x de I tel que : $y = f(x)$. D'où $-y = -f(x) = f(-x)$ (f étant impaire). Or $(-x) \in I$, puisque $x \in I$ et I est supposé symétrique par rapport à zéro. Donc $-y \in f(I)$.

En résumé, on a établi l'implication : $(y \in f(I)) \implies (-y \in f(I))$.

D'où $f(I)$ est symétrique par rapport à zéro.

► Ceci fait, il ne reste plus qu'à vérifier que : $\forall y \in f(I), f^{-1}(-y) = -f^{-1}(y)$.

Soit $y \in f(I)$. D'une part : $f(f^{-1}(-y)) = -y$ ($\spadesuit$) puisque $f \circ f^{-1} = \text{id}_{f(I)}$.

D'autre part : $f(-f^{-1}(y)) = -f(f^{-1}(y))$ (f étant impaire). D'où : $f(-f^{-1}(y)) = -y$ ($\clubsuit$).

D'après ($\spadesuit$) et ($\clubsuit$) : $f(-f^{-1}(y)) = f(f^{-1}(-y))$. Puisque f est injective (car bijective), on en déduit que : $-f^{-1}(y) = f^{-1}(-y)$.

Conclusion. Sous les hypothèses de l'énoncé, si f est impaire, alors f^{-1} est impaire. ☕

PROPRIÉTÉ 7.4 - Soit $f : I \longrightarrow f(I)$ une fonction bijective définie sur un intervalle non vide de $\mathbb{R}$, et soit $a \in I$. Si f est dérivable en a et $f'(a) \neq 0$, alors f^{-1} est dérivable en $f(a)$ et :

$$(f^{-1})'(f(a)) = \frac{1}{f'(a)}$$

PREUVE. Prouvons la propriété. Soient f, I et a comme dans l'énoncé. La fonction $f^{-1} \circ f$ est dérivable sur I , puisque pour tout réel $x \in I$ on a : $(f^{-1} \circ f)(x) = x$. Il s'ensuit en particulier que : $(f^{-1} \circ f)'(a) = 1$.

Par ailleurs, en admettant la dérivabilité de f^{-1} en $f(a)$, on a : $(f^{-1} \circ f)'(a) = (f^{-1})'(f(a)) \times f'(a)$, en vertu de la propriété donnant la dérivée d'une composée de fonctions dérivables.

On déduit des deux identités précédentes que : $(f^{-1})'(f(a)) \times f'(a) = 1$.

D'où : $(f^{-1})'(f(a)) = \frac{1}{f'(a)}$ (car $f'(a) \neq 0$). ☕

Les 4 propriétés ci-dessus sont des outils fort pratiques pour définir et étudier les réciproques des fonctions trigonométriques ; c'est l'objet du paragraphe suivant.

7.2 Fonctions trigonométriques réciproques

7.2.1 La fonction arccosinus

La fonction cosinus réalise une bijection de $[0; \pi]$ dans $[-1; 1]$; sa bijection réciproque est appelée fonction **arccosinus** et est notée **arccos** (plutôt que $\cos^{-1}$, notation que vous avez sans doute déjà vue sur vos calculatrices).

La fonction arccosinus est “explicitement” donnée par la construction de “ g ” dans le théorème 6.1 page 142 du chapitre précédent. Précisément :

$$\begin{array}{ccc} \text{arccos} : [-1; 1] & \longrightarrow & [0; \pi] \\ y & \longmapsto & \text{unique solution dans } [0; \pi] \\ & & \text{de l'équation } \cos(x) = y \end{array}$$

Il résulte de la définition que (attention aux intervalles !) :

$$[\forall x \in [-1; 1], \cos(\arccos(x)) = x] \wedge [\forall x \in [0; \pi], \arccos(\cos(x)) = x]$$

Quelques valeurs : $\arccos(0) = \frac{\pi}{2}$; $\arccos(1) = 0$; $\arccos(-1) = \pi$; $\arccos\left(\cos\left(\frac{\pi}{12}\right)\right) = \frac{\pi}{12}$; mais

attention : $\arccos\left(\cos\left(\frac{3\pi}{2}\right)\right) = \frac{\pi}{2}$.

PROPRIÉTÉ 7.5 - La fonction arccos est strictement décroissante sur $[-1, 1]$.

PREUVE. La fonction arccos est la bijection réciproque de la fonction cos restreinte à $[0, \pi]$. Puisque cette dernière est strictement décroissante sur $[0, \pi]$, arccos est strictement décroissante sur $[-1, 1]$ en vertu de la propriété 7.2 (qui affirme qu’une bijection strictement monotone et sa bijection réciproque ont le même sens de variation. ☕)

PROPRIÉTÉ 7.6 - La fonction arccos est dérivable sur $] -1, 1[$, et : $\forall x \in] -1, 1[$, $\arccos'(x) = \frac{-1}{\sqrt{1-x^2}}$.

PREUVE. Soit x un réel de $] -1, 1[$. Il existe un unique réel $a \in]0, \pi[$ tel que : $x = \cos(a)$. Puisque la fonction cos est dérivable en a , et que sa dérivée (-sin) ne s’annule pas en a ,¹ il résulte de la propriété 7.4 que la fonction arccos est dérivable en $\cos(a)$ et que :

$$\arccos'(\cos(a)) = \frac{-1}{\sin(a)}$$

Or d’après la relation fondamentale de la trigonométrie : $\sin^2(a) = 1 - \cos^2(a)$, d’où : $|\sin(a)| = \sqrt{1 - \cos^2(a)}$. En observant que $\sin(a)$ est positif (puisque $a \in]0, \pi[$), on en déduit que : $\sin(a) = \sqrt{1 - \cos^2(a)}$. Ainsi :

$$\arccos'(\cos(a)) = \frac{-1}{\sqrt{1 - \cos^2(a)}}$$

En revenant à notre variable initiale, on a donc : $\arccos'(x) = \frac{-1}{\sqrt{1-x^2}}$.

1. $-\sin(a) \neq 0$ puisque $0 < a < \pi$.

Le réel x étant un réel arbitraire de $] -1, 1[$ dans le raisonnement précédent, on a établi que la fonction arccos est dérivable sur $] -1, 1[$, et que : $\forall x \in] -1, 1[$, $\arccos'(x) = \frac{-1}{\sqrt{1-x^2}}$. ☕

Remarque. La propriété précédente ne dit rien sur la dérivabilité (ou non) de la fonction arccos en ± 1 . De fait, arccos n'est dérivable ni en -1 , ni en 1 . Une preuve de cette affirmation repose sur une propriété que nous verrons ultérieurement (le théorème des accroissements finis). Un argument (qui n'est pas une preuve rigoureuse) graphique consiste à observer que la courbe représentative de la fonction arccos possède aux points d'abscisses -1 et 1 des demi-tangentes verticales.

COROLLAIRE 7.1 - (DL à l'ordre 1 en 0 de arccos).

$$\forall h \in] -1, 1[, \arccos(h) = \frac{\pi}{2} - h + h\varepsilon(h) \quad \text{et} \quad \forall n \in \mathbb{N} \setminus \{0, 1\}, \arccos\left(\frac{1}{n}\right) = \frac{\pi}{2} - \frac{1}{n} + \frac{1}{n}\varepsilon\left(\frac{1}{n}\right)$$

avec $\lim_{h \rightarrow 0} \varepsilon(h) = 0$

PREUVE. Conséquence immédiate du théorème 5.2 page 116 et de la propriété précédente. ☕

7.2.2 La fonction arcsinus

La fonction sinus réalise une bijection de $[-\pi/2; \pi/2]$ dans $[-1; 1]$; sa bijection réciproque est appelée fonction **arcsinus** et est notée **arcsin**. La fonction arcsinus est donnée par :

$$\begin{array}{ccc} \arcsin : [-1; 1] & \longrightarrow & [-\pi/2; \pi/2] \\ y & \longmapsto & \text{unique solution dans } [-\pi/2; \pi/2] \\ & & \text{de l'équation } \sin(x) = y \end{array}$$

Il résulte de la définition que (re-attention aux intervalles !) :

$$[\forall x \in [-1; 1], \sin(\arcsin(x)) = x] \wedge [\forall x \in [-\pi/2; \pi/2], \arcsin(\sin(x)) = x]$$

Quelques valeurs : $\arcsin(0) = 0$; $\arcsin(1) = \frac{\pi}{2}$; $\arcsin(-1) = -\frac{\pi}{2}$; $\arcsin\left(\sin\left(\frac{\pi}{12}\right)\right) = \frac{\pi}{12}$; mais attention : $\arcsin\left(\sin\left(\frac{2\pi}{3}\right)\right) = \frac{\pi}{3}$.

PROPRIÉTÉ 7.7 - La fonction arcsin est strictement croissante sur $[-1, 1]$.

PREUVE. La fonction arcsin est la bijection réciproque de la fonction sin restreinte à $[-\pi/2; \pi/2]$. Puisque cette dernière est strictement croissante sur $[-\pi/2; \pi/2]$, arcsin est strictement croissante sur $[-1, 1]$ en vertu de la propriété 7.2. ☕

PROPRIÉTÉ 7.8 - La fonction arcsin est impaire.

PREUVE. La fonction arcsin est la bijection réciproque de la fonction sin restreinte à $[-\pi/2; \pi/2]$. Puisque cette dernière est impaire, arcsin est impaire en vertu de la propriété 7.3. ☕

PROPRIÉTÉ 7.9 - La fonction arcsin est dérivable sur $] -1, 1[$, et : $\forall x \in] -1, 1[$, $\arcsin'(x) = \frac{1}{\sqrt{1-x^2}}$.

PREUVE. Soit x un réel de $] -1, 1[$. Il existe un unique réel $a \in] -\frac{\pi}{2}, \frac{\pi}{2}[$ tel que : $x = \sin(a)$. Puisque la fonction sin est dérivable en a , et que sa dérivée (cos) ne s'annule pas en a , il résulte de la propriété 7.4 que la fonction arcsin est dérivable en $\sin(a)$ et que :

$$\arcsin'(\sin(a)) = \frac{1}{\cos(a)}$$

Or d'après la relation fondamentale de la trigonométrie : $\cos^2(a) = 1 - \sin^2(a)$, d'où : $|\cos(a)| = \sqrt{1 - \sin^2(a)}$. En observant que $\cos(a)$ est positif (puisque $a \in] -\frac{\pi}{2}, \frac{\pi}{2}[$), on en déduit que : $\cos(a) = \sqrt{1 - \sin^2(a)}$. Ainsi :

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

En revenant à notre variable initiale, on a donc : $\arcsin'(x) = \frac{1}{\sqrt{1-x^2}}$.

Le réel x étant un réel arbitraire de $] -1, 1[$ dans le raisonnement précédent, on a établi que la fonction arcsin est dérivable sur $] -1, 1[$, et que : $\forall x \in] -1, 1[$, $\arcsin'(x) = \frac{1}{\sqrt{1-x^2}}$. ☕

Remarque. La propriété précédente ne dit rien sur la dérivabilité (ou non) de la fonction arcsin en ± 1 . De nouveau, nous pourrions prouver que la fonction arcsin n'est dérivable ni en -1 , ni en 1 .

COROLLAIRE 7.2 - (DL à l'ordre 1 en 0 de arcsin).

$$\forall h \in] -1, 1[, \arcsin(h) = h + h\varepsilon(h) \quad \text{et} \quad \forall n \in \mathbb{N} \setminus \{0, 1\}, \arcsin\left(\frac{1}{n}\right) = \frac{1}{n} + \frac{1}{n}\varepsilon\left(\frac{1}{n}\right)$$

avec $\lim_{h \rightarrow 0} \varepsilon(h) = 0$

PREUVE. Conséquence immédiate du théorème 5.2 page 116 et de la propriété précédente. ☕

Remarque. On peut observer (ce serait même un crime de ne pas le faire !) que les fonctions arccos et arcsin ont la même dérivée, au signe près. Une application de cette observation est la question standard suivante.

EXERCICE CLASSIQUE. Montrer que : $\forall x \in [-1, 1], \arccos(x) + \arcsin(x) = \frac{\pi}{2}$

PREUVE. Notons f la fonction définie en posant : $\forall x \in [-1, 1], f(x) = \arccos(x) + \arcsin(x)$. La fonction f est dérivable sur $] - 1, 1[$, et : $\forall x \in] - 1, 1[, f'(x) = \frac{-1}{\sqrt{1-x^2}} + \frac{1}{\sqrt{1-x^2}} = 0$. On en déduit que f est constante sur $] - 1, 1[$. Pour déterminer la valeur de cette constante, il suffit de calculer (par exemple) $f(0)$.² Par définition : $f(0) = \arccos(0) + \arcsin(0) = \frac{\pi}{2}$.

On a donc établi que : $\forall x \in] - 1, 1[, f(x) = \frac{\pi}{2}$, càd : $\forall x \in] - 1, 1[, \arccos(x) + \arcsin(x) = \frac{\pi}{2}$.

Pour “fermer les crochets”, on peut observer que :

$$f(1) = \arccos(1) + \arcsin(1) = 0 + \frac{\pi}{2} = \frac{\pi}{2} \quad \text{et} \quad f(-1) = \arccos(-1) + \arcsin(-1) = \pi + \left(-\frac{\pi}{2}\right) = \frac{\pi}{2}$$

Ainsi : $\forall x \in [-1, 1], \arccos(x) + \arcsin(x) = \frac{\pi}{2}$ ☕

7.2.3 La fonction arctangente

La fonction tangente réalise une bijection de $] - \frac{\pi}{2}; \frac{\pi}{2}[$ dans $\mathbb{R}$; sa bijection réciproque est appelée **arctangente** et est notée **arctan** (plutôt que $\tan^{-1}$).

La fonction arctangente est donnée par :

$$\begin{aligned} \arctan : \mathbb{R} &\longrightarrow \left] - \frac{\pi}{2}; \frac{\pi}{2}[\right. \\ y &\longmapsto \text{unique solution dans } \left] - \frac{\pi}{2}; \frac{\pi}{2}[\right. \\ &\quad \text{de l'équation } \tan(x) = y \end{aligned}$$

Il résulte de la définition que (re-re-attention aux intervalles !) :

$$[\forall x \in \mathbb{R}, \tan(\arctan(x)) = x] \wedge \left[\forall x \in \left] - \frac{\pi}{2}; \frac{\pi}{2}[, \arctan(\tan(x)) = x \right]$$

Quelques valeurs : $\arctan(0) = 0$; $\arctan(1) = \frac{\pi}{4}$; $\arctan(-1) = -\frac{\pi}{4}$; $\arctan(\sqrt{3}) = \frac{\pi}{3}$;

$$\arctan\left(\tan\left(\frac{\pi}{12}\right)\right) = \frac{\pi}{12}; \text{ mais, attention (ce n'est plus une surprise) : } \arctan\left(\tan\left(\frac{5\pi}{4}\right)\right) = \frac{\pi}{4}.$$

PROPRIÉTÉ 7.10 - La fonction arctan est strictement croissante sur $\mathbb{R}$.

PREUVE. La fonction arctan est la bijection réciproque de la fonction $\tan$ restreinte à $\left] - \frac{\pi}{2}; \frac{\pi}{2}[$. Puisque cette dernière est strictement croissante sur $\left] - \frac{\pi}{2}; \frac{\pi}{2}[$, arctan est strictement croissante sur $\mathbb{R}$ en vertu de la propriété 7.2. ☕

². Ou l'image par f de n'importe quel réel appartenant à $] - 1, 1[$.

PROPRIÉTÉ 7.11 - La fonction arctan est impaire.

PREUVE. La fonction arctan est la bijection réciproque de la fonction tan restreinte à $\left] -\frac{\pi}{2}, \frac{\pi}{2} \right[$. Puisque cette dernière est impaire, arctan est impaire en vertu de la propriété 7.3. ☕

PROPRIÉTÉ 7.12 - La fonction arctan est dérivable sur $\mathbb{R}$, et : $\forall x \in \mathbb{R}, \arctan'(x) = \frac{1}{1+x^2}$.

PREUVE. Soit x un réel. Il existe un unique réel $a \in \left] -\frac{\pi}{2}, \frac{\pi}{2} \right[$ tel que : $x = \tan(a)$. Puisque la fonction tan est dérivable en a , et que sa dérivée $(1+\tan^2)$ ne s'annule pas en a , il résulte de la propriété 7.4 que la fonction arctan est dérivable en $\tan(a)$ et que :

$$\arctan'(\tan(a)) = \frac{1}{1+\tan^2(a)}$$

En revenant à notre variable initiale, on a donc : $\arctan'(x) = \frac{1}{1+x^2}$.

Le réel x étant arbitraire dans le raisonnement précédent, on a établi que la fonction arctan est dérivable sur $\mathbb{R}$, et que : $\forall x \in \mathbb{R}, \arctan'(x) = \frac{1}{1+x^2}$. ☕

COROLLAIRE 7.3 - (DL à l'ordre 1 en 0 de arctan).

$$\forall h \in \mathbb{R}, \arctan(h) = h + h\varepsilon(h) \quad \text{et} \quad \forall n \in \mathbb{N}^*, \arctan\left(\frac{1}{n}\right) = \frac{1}{n} + \frac{1}{n}\varepsilon\left(\frac{1}{n}\right)$$

avec $\lim_{h \rightarrow 0} \varepsilon(h) = 0$

PREUVE. Conséquence immédiate du théorème 5.2 page 116 et de la propriété précédente. ☕

On achève ce paragraphe consacré à la fonction arctangente avec une propriété remarquable de cette dernière.

EXERCICE CLASSIQUE. Montrer que : $\forall x \in \mathbb{R}^*, \arctan(x) + \arctan\left(\frac{1}{x}\right) = \frac{\pi}{2} \operatorname{sgn}(x)$

où $\operatorname{sgn}(x)$ est égal à 1 si $x > 0$, et égal à -1 si $x < 0$.

PREUVE. Notons f la fonction définie en posant : $\forall x \in \mathbb{R}^*, f(x) = \arctan(x) + \arctan\left(\frac{1}{x}\right)$. La fonction f est dérivable sur $\mathbb{R}^*$ (théorèmes généraux), et : $\forall x \in \mathbb{R}^*, f'(x) = \frac{1}{1+x^2} + \frac{-1/x^2}{1+\frac{1}{x^2}} = 0$. On en déduit que f est constante sur chacun des intervalles $\mathbb{R}_+^*$ et $\mathbb{R}_-^*$.

Pour déterminer la valeur de ces constante, il suffit de calculer (par exemple) $f(1)$ et $f(-1)$.

Par définition : $f(1) = \arctan(1) + \arctan(1) = \frac{\pi}{4} + \frac{\pi}{4} = \frac{\pi}{2}$. Et par “imparité” : $f(-1) = -\frac{\pi}{2}$.

Ainsi : $\forall x > 0, \arctan(x) + \arctan\left(\frac{1}{x}\right) = \frac{\pi}{2}$ et $\forall x < 0, \arctan(x) + \arctan\left(\frac{1}{x}\right) = -\frac{\pi}{2}$ ☕

7.2.4 Bilan sur les fonctions circulaires réciproques

► La fonction arccosinus (arccos).

arccos est **définie** sur $[-1; 1]$ par : $\text{arccos} : [-1; 1] \longrightarrow [0; \pi]$

$$x \longmapsto y$$

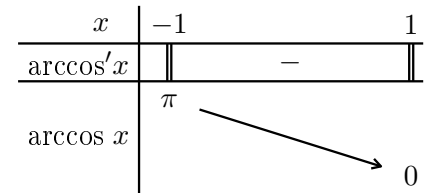
(où y : unique solution dans $[0; \pi]$ de $\cos(y) = x$).

La fonction arccos est **dérivable** (et même de classe $\mathcal{C}^1$) sur $] -1; 1[$, et **n'est pas dérivable** en 1, ni en -1 .

De plus : $\forall x \in] -1; 1[, \text{arccos}'x = -\frac{1}{\sqrt{1-x^2}}$.

La fonction arccos est strictement décroissante sur $[-1; 1]$.

La tangente à la courbe représentative $\mathcal{C}_{\text{arccos}}$ au point d'abscisse 0 a pour équation $y = \frac{\pi}{2} - x$, et $\mathcal{C}_{\text{arccos}}$ est située au-dessus (*resp.* en-dessous) de cette tangente sur $[-1; 0]$ (*resp.* sur $[0; 1]$).



► La fonction arcsinus (arcsin).

arcsin est **définie** sur $[-1; 1]$ par : $\text{arcsin} : [-1; 1] \longrightarrow \left[-\frac{\pi}{2}; \frac{\pi}{2}\right]$

$$x \longmapsto y$$

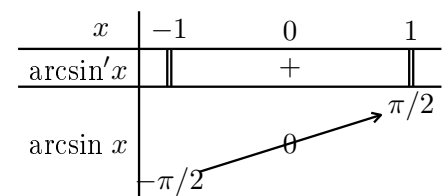
(où y : unique solution dans $\left[-\frac{\pi}{2}; \frac{\pi}{2}\right]$ de $\sin(y) = x$).

La fonction arcsin est **impaire**, **dérivable** (et même de classe $\mathcal{C}^1$) sur $] -1; 1[$, et **n'est pas dérivable** en 1, ni en -1 .

De plus : $\forall x \in] -1; 1[, \text{arcsin}'x = \frac{1}{\sqrt{1-x^2}}$.

La fonction arcsin est strictement croissante sur $[-1; 1]$.

La tangente à la courbe représentative $\mathcal{C}_{\text{arcsin}}$ au point d'abscisse 0 a pour équation $y = x$, et $\mathcal{C}_{\text{arcsin}}$ est située au-dessus (*resp.* en-dessous) de cette tangente sur $[0; 1]$ (*resp.* sur $[-1; 0]$).



► La fonction arctangente (arctan).

arctan est **définie** sur $\mathbb{R}$ par : $\text{arctan} : \mathbb{R} \longrightarrow]-\frac{\pi}{2}; \frac{\pi}{2}[$

$$x \longmapsto y$$

(où y : unique solution dans $]-\frac{\pi}{2}; \frac{\pi}{2}[$ de $\tan(y) = x$).

La fonction arctan est **impaire**, **dérivable** (et même de classe $\mathcal{C}^1$) sur $\mathbb{R}$.

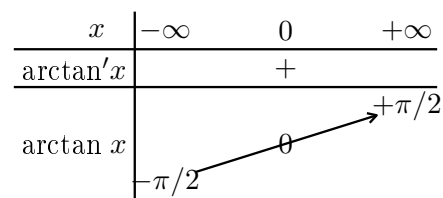
De plus : $\forall x \in \mathbb{R}, \text{arctan}'x = \frac{1}{1+x^2}$.

La fonction arctan est strictement croissante sur $\mathbb{R}$.

De plus elle est **bornée** (entre $-\pi/2$ et $\pi/2$).

La tangente à la courbe représentative $\mathcal{C}_{\text{arctan}}$ au point d'abscisse 0 a pour équation $y = x$, et $\mathcal{C}_{\text{arctan}}$ est située au-dessus (*resp.* en-dessous) de cette tangente sur $\mathbb{R}_-$ (*resp.* sur $\mathbb{R}_+$).

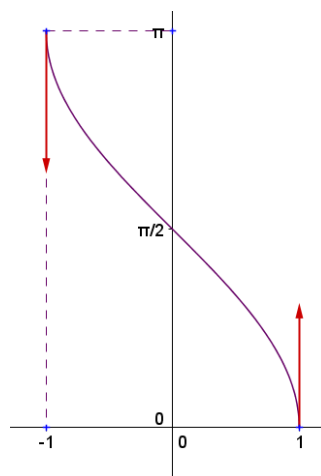
La courbe $\mathcal{C}_{\text{arctan}}$ admet au voisinage de $+\infty$ (*resp.* de $-\infty$) une asymptote horizontale d'équation $y = \pi/2$ (*resp.* $y = -\pi/2$).



► On achève ce paragraphe par les courbes représentatives des fonctions circulaires réciproques.

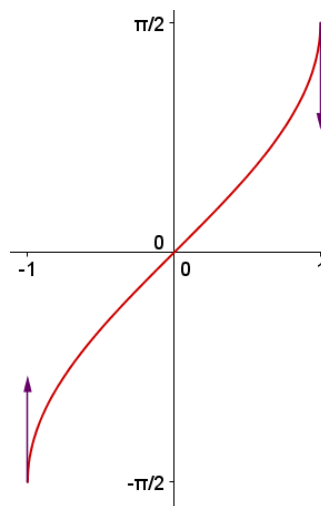
ARCCOSINUS (ARCCOS)

x	-1	1
$\arccos' x$	$\parallel$	$-$
$\arccos x$	π	0



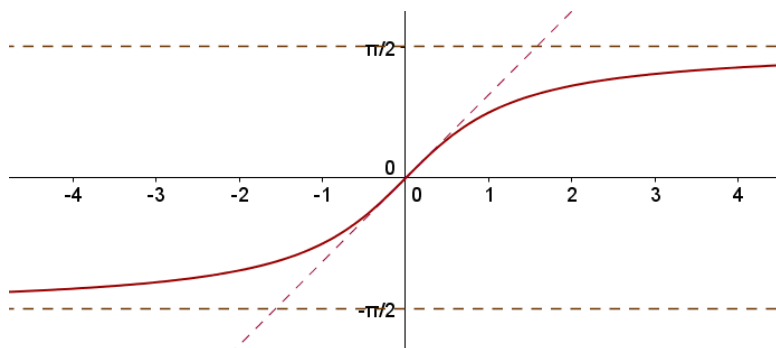
ARCSINUS (ARCSIN)

x	-1	0	1
$\arcsin' x$	$\parallel$	$+$	$\parallel$
$\arcsin x$	$-\pi/2$	0	$\pi/2$



ARCTANGENTE (ARCTAN)

x	$-\infty$	0	$+\infty$
$\arctan' x$		$+$	
$\arctan x$	$-\pi/2$	0	$+\pi/2$



7.3 Epilogue

7.3.1 Exercice classique : $\arctan\left(\frac{1}{3}\right) + \arctan\left(\frac{1}{4}\right) = \arctan\left(\frac{7}{11}\right)$

Montrons que : $\arctan\left(\frac{1}{3}\right) + \arctan\left(\frac{1}{4}\right) = \arctan\left(\frac{7}{11}\right)$.

En utilisant la formule d'addition pour la fonction tangente (et le fait que $\tan \circ \arctan = \text{id}_{\mathbb{R}}$), on obtient :

$$\tan\left(\arctan\left(\frac{1}{3}\right) + \arctan\left(\frac{1}{4}\right)\right) = \frac{\frac{1}{3} + \frac{1}{4}}{1 - \frac{1}{3} \times \frac{1}{4}} = \frac{7/12}{11/12}.$$

$$\text{D'où : } \tan\left(\arctan\left(\frac{1}{3}\right) + \arctan\left(\frac{1}{4}\right)\right) = \frac{7}{11}.$$

$$\text{On a donc : } \tan\left(\arctan\left(\frac{1}{3}\right) + \arctan\left(\frac{1}{4}\right)\right) = \tan\left(\arctan\left(\frac{7}{11}\right)\right).$$

$$\text{Par suite : } \arctan\left(\frac{1}{3}\right) + \arctan\left(\frac{1}{4}\right) = \arctan\left(\frac{7}{11}\right) \quad [\pi].$$

On peut conclure en notant³ que $\arctan(7/11) \in [0, \pi/2[$, et que $\arctan(1/3)$ et $\arctan(1/4)$ sont dans $[0, \pi/4[$.

$$\text{Par suite : } \boxed{\arctan\left(\frac{1}{3}\right) + \arctan\left(\frac{1}{4}\right) = \arctan\left(\frac{7}{11}\right)}$$

7.3.2 Somme d'une série

Le but de l'exercice ci-dessous est de calculer :

$$\lim_{n \rightarrow +\infty} \left[\sum_{p=0}^n \arctan\left(\frac{1}{p^2 + p + 1}\right) \right]$$

1/ Soit p un entier naturel. Calculer : $\delta_p = \arctan(p+1) - \arctan(p)$

2/ Pour tout entier naturel n , on pose : $S_n = \sum_{p=0}^n \arctan\left(\frac{1}{p^2 + p + 1}\right)$

Calculer S_n , puis la limite de S_n lorsque n tend vers $+\infty$.

3. Comme dans la récente question de cours.

CORRIGÉ.

1/ Soit p un entier naturel. On a : $\tan(\arctan(p+1) - \arctan(p)) = \frac{(p+1) - p}{1 + (p+1)p}$

$$\text{soit : } \tan(\arctan(p+1) - \arctan(p)) = \frac{1}{p^2 + p + 1}$$

L'idée est maintenant d'appliquer la fonction arctangente aux deux membres de cette égalité. Il suffit pour cela d'observer que $\arctan(p+1) - \arctan(p)$ appartient à $[0, \pi/2[$ (essentiellement en lisant le tableau de variation de la fonction arctangente). De même, le réel $\arctan\left(\frac{1}{p^2 + p + 1}\right)$ est dans $[0, \pi/2[$.

On en déduit que : $\forall p \in \mathbb{N}, \arctan(p+1) - \arctan(p) = \arctan\left(\frac{1}{p^2 + p + 1}\right)$

2/ D'après la question précédente :

$$\begin{aligned} \forall n \in \mathbb{N}, \sum_{p=0}^n \arctan\left(\frac{1}{p^2 + p + 1}\right) &= \sum_{p=0}^n \arctan(p+1) - \arctan(p) \\ \iff \forall n \in \mathbb{N}, \sum_{p=0}^n \arctan\left(\frac{1}{p^2 + p + 1}\right) &= \arctan(n+1) - \arctan(0) \end{aligned}$$

Conclusion : $\forall n \in \mathbb{N}, \sum_{p=0}^n \arctan\left(\frac{1}{p^2 + p + 1}\right) = \arctan(n+1)$

Par suite, comme la fonction arctangente tend vers $\frac{\pi}{2}$ en $+\infty$: $\lim_{n \rightarrow +\infty} \left[\sum_{p=0}^n \arctan\left(\frac{1}{p^2 + p + 1}\right) \right] = \frac{\pi}{2}$

Remarque : un peu plus tard dans l'année, nous traduirons ce dernier résultat en disant que la **série** de terme général $\arctan\left(\frac{1}{p^2 + p + 1}\right)$ est **convergente** et a pour somme $\pi/2$, et nous le noterons

$$\sum_{p=0}^{+\infty} \arctan\left(\frac{1}{p^2 + p + 1}\right) = \frac{\pi}{2}$$

7.3.3 Formule de Machin

Exercice — (Formule de Machin).⁴

- 1) Soient x et y deux réels tels que $0 < x < y$. Montrer que : $\arctan\left(\frac{x}{y}\right) + \arctan\left(\frac{y-x}{y+x}\right) = \frac{\pi}{4}$.
- 2) Calculer $A = 4\arctan\left(\frac{1}{5}\right)$.
- 3) Montrer que : $4\arctan\left(\frac{1}{5}\right) - \arctan\left(\frac{1}{239}\right) = \frac{\pi}{4}$.

CORRIGÉ.

- 1) Soient x et y deux réels tels que $0 < x < y$.

D'après la formule d'addition pour la tangente, on a :

$$\tan\left(\arctan\left(\frac{x}{y}\right) + \arctan\left(\frac{y-x}{y+x}\right)\right) = \frac{\frac{x}{y} + \frac{y-x}{y+x}}{1 - \frac{x}{y} \times \frac{y-x}{y+x}} = \frac{\left(\frac{x(y+x) + (y-x)y}{y(y+x)}\right)}{\left(\frac{y(y+x) - x(y-x)}{y(y+x)}\right)} = \frac{\left(\frac{x^2 + y^2}{y(y+x)}\right)}{\left(\frac{x^2 + y^2}{y(y+x)}\right)} = 1.$$

D'autre part : $\tan(\pi/4) = 1$.

On en déduit que : $\arctan\left(\frac{x}{y}\right) + \arctan\left(\frac{y-x}{y+x}\right) = \frac{\pi}{4} \quad [\pi]$.

Les hypothèses faites sur les réels x et y permettent alors d'affirmer que les réels $\frac{x}{y}$ et $\frac{y-x}{y+x}$ sont strictement compris entre 0 et 1. Donc les images par la fonction arctangente de ces deux derniers réels appartiennent à l'intervalle $]0, \pi/4[$; par suite leur somme est dans l'intervalle $]0, \pi/2[$, et à plus forte raison : $\arctan\left(\frac{x}{y}\right) + \arctan\left(\frac{y-x}{y+x}\right) \in]0, \pi[$.

Puisque $\arctan\left(\frac{x}{y}\right) + \arctan\left(\frac{y-x}{y+x}\right) = \frac{\pi}{4} \quad [\pi]$ et que les deux réels intervenant dans cette congruence appartiennent au même intervalle $]0, \pi[$, on peut conclure qu'ils sont égaux.

Conclusion : $\forall (x, y) \in \mathbb{R}^2, [0 < x < y] \implies \left[\arctan\left(\frac{x}{y}\right) + \arctan\left(\frac{y-x}{y+x}\right) = \frac{\pi}{4}\right]$.

$$2) \text{ On a : } \tan\left(2\arctan\left(\frac{1}{5}\right)\right) = \frac{2/5}{1 - (1/5)^2} = \frac{2/5}{24/25} = \frac{5}{12}.$$

$$\text{D'où : } \tan\left(4\arctan\left(\frac{1}{5}\right)\right) = \tan\left(2 \times 2\arctan\left(\frac{1}{5}\right)\right) = \frac{5/6}{1 - (5/12)^2} = \frac{5/6}{119/144} = \frac{120}{119}.$$

⁴ John Machin fut un mathématicien anglais du 18^{ème} siècle (1680-1751), notamment connu pour avoir démontré en 1706 la formule que l'on vous demande d'établir dans cet exercice, qui lui a permis d'obtenir une remarquable (pour l'époque) approximation du nombre π (100 décimales).

On en déduit que : $4\arctan\left(\frac{1}{5}\right) = \arctan\left(\frac{120}{119}\right) \pmod{\pi}$.

Puisque $1/5$ est strictement compris entre 0 et 1, son image par la fonction arctangente appartient à l'intervalle $]0, \pi/4[$; par suite $4\arctan\left(\frac{1}{5}\right) \in]0, \pi[$.

Puisque $4\arctan\left(\frac{1}{5}\right) = \arctan\left(\frac{120}{119}\right) \pmod{\pi}$, et que les deux réels intervenant dans cette congruence appartiennent au même intervalle $]0, \pi[$, on peut conclure qu'ils sont égaux.

Conclusion : $4\arctan\left(\frac{1}{5}\right) = \arctan\left(\frac{120}{119}\right)$.

3) D'après la relation de la question 1, appliquée à $x = 119$ et $y = 120$ on a :

$$\arctan\left(\frac{119}{120}\right) + \arctan\left(\frac{1}{239}\right) = \frac{\pi}{4}$$

Par ailleurs : $\arctan\left(\frac{119}{120}\right) = \frac{\pi}{2} - \arctan\left(\frac{120}{119}\right)$.⁵ D'où : $\arctan\left(\frac{119}{120}\right) = \frac{\pi}{2} - 4\arctan\left(\frac{1}{5}\right)$.

On déduit de ce qui précède que : $\frac{\pi}{2} - 4\arctan\left(\frac{1}{5}\right) + \arctan\left(\frac{1}{239}\right) = \frac{\pi}{4}$.

Ainsi :

$4\arctan\left(\frac{1}{5}\right) - \arctan\left(\frac{1}{239}\right) = \frac{\pi}{4}$.

7.3.4 Problème des 13 réels

Exercice.

1/ Calculer la valeur exacte de $\tan(\pi/12)$.

2/ Soient $x_1, \dots, x_{13}$ treize réels arbitraires. Etablir que :

$$\exists (i, j) \in \llbracket 1, 13 \rrbracket, i \neq j, 0 \leq \frac{x_j - x_i}{1 + x_i x_j} \leq 2 - \sqrt{3}$$

CORRIGÉ.

1/ Pour déterminer la valeur exacte de $\tan(\pi/12)$, on peut observer que $\frac{\pi}{12} = \frac{\pi}{3} - \frac{\pi}{4}$, et utiliser la formule de soustraction pour la tangente :

$$\tan\left(\frac{\pi}{12}\right) = \frac{\tan\left(\frac{\pi}{3}\right) - \tan\left(\frac{\pi}{4}\right)}{1 + \tan\left(\frac{\pi}{3}\right)\tan\left(\frac{\pi}{4}\right)} = \frac{\sqrt{3} - 1}{1 + \sqrt{3}} = \frac{(\sqrt{3} - 1)^2}{2} = \frac{4 - 2\sqrt{3}}{2} = 2 - \sqrt{3}$$

Ainsi :

$\tan\left(\frac{\pi}{12}\right) = 2 - \sqrt{3}$.

⁶

5. Puisque : $\forall x \in \mathbb{R}^*, \arctan(x) + \arctan\left(\frac{1}{x}\right) = \frac{\pi}{2} \operatorname{sgn}(x)$.

6. Quelle merveilleuse coïncidence!...

2/ Il s'agit entre autres d'exploiter la remarquable coïncidence observée à la question précédente...

Chapitre 8

Méthodes de calcul intégral

8.1 Fonctions de classe $\mathcal{C}^1$, $\mathcal{C}^n$, $\mathcal{C}^\infty$

DÉFINITION 8.1 - Soit f une fonction définie sur un intervalle I et à valeurs réelles.

- f est **de classe $\mathcal{C}^1$ sur I** si f est dérivable sur I et f' est continue sur I .
- Plus généralement pour $n \in \mathbb{N}$, f est **de classe $\mathcal{C}^n$ sur I** lorsque $f, f', \dots, f^{(n-1)}$ sont dérivables sur I , et $f^{(n)}$ est continue sur I .
- Enfin f est **de classe $\mathcal{C}^\infty$ sur I** lorsque f admet une dérivée à tout ordre sur I (c'est-à-dire si $f^{(n)}$ existe pour tout entier naturel n).

Remarques : une fonction de classe $\mathcal{C}^0$ sur I est une fonction continue sur I . Une fonction de classe $\mathcal{C}^1$ (*resp.* $\mathcal{C}^n$, *resp.* $\mathcal{C}^\infty$) est dite **continûment dérivable** (*resp.* **n fois continûment dérivable**, *resp.* **indéfiniment dérivable**) sur I .

Notation. On note $\mathcal{C}^1(I, \mathbb{R})$ (*resp.* $\mathcal{C}^n(I, \mathbb{R})$, *resp.* $\mathcal{C}^\infty(I, \mathbb{R})$) l'ensemble des fonctions de classe $\mathcal{C}^1$ (*resp.* $\mathcal{C}^n$, *resp.* $\mathcal{C}^\infty$) sur I .

Il résulte des définitions que l'on a une chaîne d'inclusions :

$$\mathcal{C}^\infty(I, \mathbb{R}) \subset \dots \subset \mathcal{C}^{n+1}(I, \mathbb{R}) \subset \mathcal{C}^n(I, \mathbb{R}) \subset \mathcal{C}^{n-1}(I, \mathbb{R}) \subset \dots \subset \mathcal{C}^1(I, \mathbb{R}) \subset \mathcal{C}^0(I, \mathbb{R})$$

On dispose par ailleurs de théorèmes généraux sur les fonctions de classe $\mathcal{C}^N$, que l'on peut énoncer un peu informellement comme suit :

THÉORÈME 8.1 - (Théorème général 1 sur les fonctions de classe $\mathcal{C}^N$). Soit $N \in \mathbb{N} \cup \{\infty\}$. Les fonctions usuelles sont de classe $\mathcal{C}^N$ sur leurs ensembles de dérivabilité respectifs.

En particulier, les fonctions $\exp$, $\sin$, $\cos$, th , arctan , polynomiales sont de classe $\mathcal{C}^\infty$ sur $\mathbb{R}$; les fonction $\ln$ et " $\sqrt{}$ " sont de classe $\mathcal{C}^\infty$ sur $\mathbb{R}_+^*$; les fonctions arcsin et arccos sont de classe $\mathcal{C}^\infty$ sur $] -1, 1[$; la fonction valeur absolue est de classe $\mathcal{C}^\infty$ sur $\mathbb{R}^*$.

THÉORÈME 8.2 - (Théorème général 2 sur les fonctions de classe $\mathcal{C}^N$). Soit $N \in \mathbb{N} \cup \{\infty\}$. La somme, le produit, le quotient (partout où il est défini), la composée (sous réserve qu'elle soit définie) de fonctions de classe $\mathcal{C}^N$ est une fonction de classe $\mathcal{C}^N$.

Ce second énoncé provient par récurrence des propriétés analogues sur les fonctions dérivables (une fois!).

Exemples : la fonction $x \mapsto x^2 e^{1/x}$ est de classe $\mathcal{C}^\infty$ sur $\mathbb{R}^*$. La fonction $x \mapsto \sin(x-1) \ln(1+x^2)$ est de classe $\mathcal{C}^\infty$ sur $\mathbb{R}$...

8.2 Primitives

DÉFINITION 8.2 - Soit f une fonction définie sur un intervalle I et à valeurs réelles. Une primitive de f sur I est une fonction F dérivable sur I telle que $F' = f$.

Exemples : la fonction $\ln$ est une primitive de la fonction inverse sur $\mathbb{R}_+^*$. La fonction $(-\cos)$ est une primitive de $\sin$ sur $\mathbb{R}$.

PROPRIÉTÉ 8.1 - Soit f une fonction définie sur un intervalle I et à valeurs réelles. Deux primitives de f sur I diffèrent d'une constante.

Explicitement, si F et G sont deux primitives de f sur I , alors : $\exists K \in \mathbb{R}, \forall x \in I, F(x) - G(x) = K$.

PREUVE. Sous les hypothèses de l'énoncé : $F' = G'$ d'où $(F - G)' = 0$ donc $(F - G)$ est constante. ☕

PROPRIÉTÉ 8.2 - Soient f et g deux fonctions définies sur un intervalle I et à valeurs réelles.

Si F et G désignent des primitives respectives de f et g sur I , alors $F + G$ est une primitive de $f + g$ sur I ; λF est une primitive de λf (pour tout réel λ) sur I . Il revient au même de dire que pour tout couple de réels (λ, μ) , la fonction $\lambda F + \mu G$ est une primitive de $\lambda f + \mu g$ sur I .

PREUVE. Vérifications immédiates. ☕

THÉORÈME 8.3 - (Théorème fondamental) Toute fonction continue sur un intervalle I admet une primitive sur I (donc une infinité de primitives sur I).

Explicitement, la fonction $x \in I \mapsto \int_a^x f(t) dt$ est la primitive de f sur I s'annulant en a (avec $a \in I$).

Remarque : ce théorème, dont je vous ai donné une idée de la preuve en cours, ne peut être démontré rigoureusement qu'une fois la notion d'intégrale elle-même définie rigoureusement, ce qui sera fait au second semestre dans le cadre des compléments sur l'intégration que nous verrons alors.

8.3 Primitives usuelles

Fonction f	Primitives F
$x^\alpha \ (\alpha \neq -1)$	$\frac{1}{\alpha+1} x^{\alpha+1} + c$
$\frac{1}{x}$	$\ln x + c$
$\frac{1}{x^2}$	$-\frac{1}{x} + c$
$\frac{1}{x^n} \ (n \neq 1)$	$\frac{-1}{(n-1)x^{n-1}} + c$
$\frac{1}{\sqrt{x}}$	$2\sqrt{x} + c$
e^x	$e^x + c$
e^{ax+b}	$\frac{1}{a} e^{ax+b} + c$
$\sin(x)$	$-\cos(x) + c$
$\cos(x)$	$\sin(x) + c$
$\sin(ax+b)$	$-\frac{1}{a} \cos(ax+b) + c$
$\cos(ax+b)$	$\frac{1}{a} \sin(ax+b) + c$
$\frac{1}{\cos^2(x)}$	$\tan(x) + c$
$\tan(x)$	$-\ln(\cos(x))$
$\operatorname{ch}(x)$	$\operatorname{sh}(x)$
$\operatorname{sh}(x)$	$\operatorname{ch}(x)$

Fonction f	Primitives F
$\frac{1}{\sqrt{1-x^2}}$	$\arcsin(x) + c$
$-\frac{1}{\sqrt{1-x^2}}$	$\arccos(x) + c$
$\frac{1}{1+x^2}$	$\arctan(x) + c$
$u'(x) u(x)$	$\frac{1}{2} u^2(x) + c$
$\frac{u'(x)}{u(x)}$	$\ln u(x) + c$
$u'(x) u^\alpha(x)$	$\frac{1}{\alpha+1} u^{\alpha+1}(x) + c$
$u'(x) e^{u(x)}$	$e^{u(x)} + c$
$u'(x) \sin(u(x))$	$-\cos(u(x)) + c$
$u'(x) \cos(u(x))$	$\sin(u(x)) + c$
$\frac{u'(x)}{\cos^2(u(x))}$	$\tan(u(x)) + c$
$u'(x) \operatorname{ch}(u(x))$	$\operatorname{sh}(u(x)) + c$
$u'(x) \operatorname{sh}(u(x))$	$\operatorname{ch}(u(x)) + c$
$\frac{u'(x)}{\sqrt{1-u^2(x)}}$	$\arcsin(u(x)) + c$
$-\frac{u'(x)}{\sqrt{1-u^2(x)}}$	$\arccos(u(x)) + c$
$\frac{u'(x)}{1+u^2(x)}$	$\arctan(u(x)) + c$

Cas particuliers – Primitives et composées

- a) Une primitive de $\frac{u'}{u}$ est $\ln |u|$;
- b) Une primitive de $u'e^u$ est e^u ;
- c) Une primitive de $u'u^\alpha$ est $\frac{u^{\alpha+1}}{\alpha+1}$ (avec $\alpha \in \mathbb{R}, \alpha \neq -1$).

8.4 Intégrale d'une fonction continue sur un segment

DÉFINITION 8.3 - Soit f une fonction continue sur $[a; b]$.

L'**intégrale de f sur $[a, b]$** est le réel : $\int_a^b f(t)dt = F(b) - F(a)$, où F désigne une primitive de f sur $[a, b]$.

En notant $[F(t)]_a^b = F(b) - F(a)$, la relation précédente peut encore s'écrire : $\int_a^b f(t)dt = [F(t)]_a^b$.

Remarque. Il convient de remarquer que la valeur de l'intégrale $\int_a^b f(t)dt$ est (heureusement !) indépendante de la primitive F de f choisie pour la calculer. En effet, si F et G désignent deux primitives de f , il existe une constante K telle que $F = G + K$. Dans ces conditions, on a clairement : $[F(t)]_a^b = [G(t)]_a^b$.

Exemples.

1/ On a : $\int_0^1 \frac{1}{1+t} dt = [\ln(1+t)]_0^1 = \ln(2)$

2/ On a : $\int_{-3}^{-1} \frac{1}{t} dt = [\ln |t|]_{-3}^{-1} = -\ln(3)$

3/ On a : $\int_0^{\pi/4} \frac{1}{\cos^2 x} dx = [\tan(x)]_0^{\pi/4} = \ln(2)$

4/ On a : $\int_0^{\pi/2} \cos^2 x dx = \int_0^{\pi/2} \frac{1 + \cos(2x)}{2} dx = \left[\frac{x}{2} + \frac{\sin(2x)}{4} \right]_0^{\pi/2} = \frac{\pi}{4}$

PROPRIÉTÉ 8.3 - (Propriétés de l'intégrale). Soient f et g deux fonctions continues sur $[a; b]$.

1/ $\forall (\lambda, \mu) \in \mathbb{R}^2, \int_a^b (\lambda f + \mu g)(t) dt = \lambda \int_a^b f(t)dt + \mu \int_a^b g(t)dt$ (**linéarité**)

2/ $\int_a^b f(t) dt = \int_a^c f(t)dt + \int_c^b f(t)dt$ (**relation de Chasles**)

3/ $[f \text{ positive sur } [a, b]] \implies \left[\int_a^b f(t) dt \geq 0 \right]$ (**positivité**)

4/ $\left| \int_a^b f(t) dt \right| \leq \int_a^b |f(t)| dt$ (**majoration de la valeur absolue**)

PREUVE. Soient f et g deux fonctions continues sur $[a; b]$.

1/ Notons F et G deux primitives de f et g respectivement, et considérons λ et μ deux scalaires arbitraires. D'après la propriété 8.2 page 164, la fonction $\lambda F + \mu G$ est une primitive de $\lambda f + \mu g$, d'où :

$$\begin{aligned} \int_a^b (\lambda f + \mu g)(t) dt &= [\lambda F(t) + \mu G(t)]_a^b = \lambda F(b) - \lambda F(a) + \mu G(b) - \mu G(a) = \lambda [F(t)]_a^b + \mu [G(t)]_a^b \\ &= \lambda \int_a^b f(t) dt + \mu \int_a^b g(t) dt \end{aligned}$$

2/ Soit $c \in [a, b]$. On a : $\int_a^b f(t) dt = F(b) - F(a) = F(b) - F(c) + F(c) - F(a) = \int_c^b f(t) dt + \int_a^c f(t) dt$.

3/ Supposons f positive sur $[a, b]$. Alors F est croissante sur $[a, b]$, d'où $F(b) - F(a) \geq 0$, ce qui équivaut à $\int_a^b f(t) dt \geq 0$.

4/ Sur $[a, b]$, la fonction $|f| - f$ est positive. D'après le point précédent, on a donc : $\int_a^b |f(t)| - f(t) dt \geq 0$.

Par linéarité de l'intégrale, on en déduit que : $\int_a^b |f(t)| dt \geq \int_a^b f(t) dt$.

Si $\int_a^b f(t) dt \geq 0$, on a : $\int_a^b f(t) dt = \left| \int_a^b f(t) dt \right|$; et on déduit alors de ce qui précède que :

$$\left| \int_a^b f(t) dt \right| \leq \int_a^b |f(t)| dt$$

Dans le cas où $\int_a^b f(t) dt \leq 0$, on peut reprendre le raisonnement précédent en l'appliquant à $(-f)$, pour obtenir de nouveau : $\left| \int_a^b f(t) dt \right| \leq \int_a^b |f(t)| dt$. ☕

Exemple 1. Soit n un entier naturel. Posons : $I_n = \int_0^1 \cos(t) t^n dt$.

Puisque la fonction $t \mapsto \cos(t) t^n$ est positive sur $[0, 1]$, on a : $I_n \geq 0$ (par positivité de l'intégrale).

Par ailleurs, pour tout réel t dans $[0, 1]$, on a : $\cos(t) t^n \leq t^n$.

On en déduit que pour tout $n \in \mathbb{N}$, on a : $0 \leq I_n \leq \int_0^1 t^n dt$.

D'où : $\forall n \in \mathbb{N}, \quad 0 \leq I_n \leq \frac{1}{n+1}$.

En particulier, d'après le théorème des gendarmes : $\lim_{n \rightarrow +\infty} I_n = 0$

Exemple 2. On se propose de généraliser dans cet exemple la situation précédente. Soit n un entier naturel, et soit f une fonction continue sur $[0, 1]$. Posons : $I_n = \int_0^1 f(t) t^n dt$. Le but est de prouver que I_n tend vers 0 lorsque n tend vers $+\infty$.

Puisque f est continue sur $[0, 1]$, f est bornée (et atteint ses bornes¹). En particulier :

$$\exists M \in \mathbb{R}_+, \forall t \in [0, 1], |f(t)| \leq M$$

D'après le point 4 des propriétés précédentes, on a pour tout entier naturel n :

$$|I_n| \leq \int_0^1 |f(t) t^n| dt \quad (\spadesuit)$$

$$\text{Or : } \int_0^1 |f(t) t^n| dt = \int_0^1 |f(t)| t^n dt \leq M \int_0^1 t^n dt \quad (\clubsuit)$$

On déduit de $(\spadesuit)$ et de $(\clubsuit)$ que pour tout entier naturel n :

$$|I_n| \leq M \int_0^1 t^n dt \quad \text{d'où : } |I_n| \leq \frac{M}{n+1}.$$

Une nouvelle application du théorème des gendarmes permet de conclure que : $\lim_{n \rightarrow +\infty} |I_n| = 0$.

Il s'ensuit que : $\lim_{n \rightarrow +\infty} I_n = 0$.

On a donc établi que pour toute fonction f continue sur $[0, 1]$, on a : $\lim_{n \rightarrow +\infty} \int_0^1 f(t) t^n dt = 0$.

On achève ce paragraphe par une propriété qui n'est finalement qu'une retraduction de celle de positivité de l'intégrale ; mais son importance pratique est suffisamment grande pour ne pas la passer sous silence.

COROLLAIRE 8.1 - (Croissance de l'intégrale). Soient f et g deux fonctions continues sur $[a, b]$. On a :

$$[f \leq g \text{ sur } [a, b]] \implies \left[\int_a^b f(t) dt \leq \int_a^b g(t) dt \right]$$

Remarque. Il est essentiel dans cette propriété de faire l'hypothèse que les bornes sont “dans le bon sens”, càd que $a \leq b$.

PREUVE. Si $f \leq g$, alors $g - f \geq 0$. Par positivité de l'intégrale (propriété 8.3), on en déduit que : $\int_a^b g(t) - f(t) dt \geq 0$, d'où la conclusion, par linéarité de l'intégrale. ☕

8.5 Intégration par parties

THÉORÈME 8.4 - (Formule d'intégration par parties) Soient u et v deux fonctions de classe $\mathcal{C}^1$ sur un intervalle I , et soient a et b deux réels dans I . On a :

$$\int_a^b u'v = [uv]_a^b - \int_a^b u'v$$

1. C'est ce que l'on appellera le **théorème des bornes atteintes**, valide pour une fonction continue sur un segment ; voir chapitre “Limites et Continuité”.

PREUVE. Soient u et v deux fonctions de classe $\mathcal{C}^1$ sur I . Alors uv est dérivable sur I et :

$\forall x \in I, (uv)'(x) = u'(x)v(x) + u(x)v'(x)$. Par intégration (légitime car u, v, u' et v' sont continues sur I), on en déduit que : $\int_a^b (uv)'(x) dx = \int_a^b u'(x)v(x) dx + \int_a^b u(x)v'(x) dx$ d'où :

$$\forall (u, v) \in \mathcal{C}^1(I, \mathbb{R}), \int_a^b u'(x)v(x) dx = [u(x)v(x)]_a^b - \int_a^b u(x)v'(x) dx$$



Exemple. Posons : $I = \int_0^1 te^{-t} dt$. Posons pour tout réel t : $u(t) = t$ et $v(t) = -e^{-t}$.

D'après les théorèmes généraux, les fonctions u et v sont de classe $\mathcal{C}^1$ sur $\mathbb{R}$, et pour tout réel t on a : $u'(t) = 1$ et $v'(t) = e^{-t}$.

On peut légitimement utiliser une intégration par parties pour obtenir :

$$I = [-te^{-t}]_0^1 - \int_0^1 -e^{-t} dt = -e^{-1} - [e^{-t}]_0^1 = -e^{-1} - e^{-1} + 1 \quad \text{soit : } \boxed{I = 1 - 2e^{-1}}.$$

COROLLAIRE 8.2 - (IPP pour le calcul d'une primitive) Soient u et v deux fonctions de classe $\mathcal{C}_1$ sur un intervalle I .

Alors : une primitive de $u'v$ sur I est $uv - \int u'v$.

PREUVE. Soient u et v deux fonctions de classe $\mathcal{C}_1$ sur un intervalle I . Alors uv est dérivable sur I et :

$\forall x \in I, (uv)'(x) = u'(x)v(x) + u(x)v'(x)$. Par intégration (légitime car u, v, u' et v' sont continues), on en déduit que : $\int (uv)'(x) dx = \int u'(x)v(x) dx + \int u(x)v'(x) dx$ d'où :

$$\forall (u, v) \in \mathcal{C}^1(I, \mathbb{R}), \int u'(x)v(x) dx = u(x)v(x) - \int u(x)v'(x) dx (+cste)$$



➤ Application 1.

$$\int \ln(x) dx = \int 1 \times \ln(x) dx = x \ln(x) - \int x \times \frac{1}{x} dx = x \ln(x) - x$$

La fonction $x \in \mathbb{R}_+^* \mapsto x \ln(x) - x$ est une primitive de la fonction $\ln$ sur $\mathbb{R}_+^*$.

Remarque : on aura bien entendu observé que les fonctions $x \mapsto x$ et $x \mapsto \ln(x)$ sont de classe $\mathcal{C}^1$ sur $\mathbb{R}_+^*$, ce qui autorise l'utilisation d'une intégration par parties.

➤ Application 2.

$$\int \arctan(x) dx = \int 1 \times \arctan(x) dx = x \arctan(x) - \int \frac{x}{1+x^2} dx = x \arctan(x) - \ln(\sqrt{1+x^2})$$

La fonction $x \in \mathbb{R} \mapsto x \arctan(x) - \ln(\sqrt{1+x^2})$ est une primitive de la fonction Arctan sur $\mathbb{R}$.

► **Application 3.**

$$\int \arccos(x) dx = \int 1 \times \arccos(x) dx = x \arccos(x) + \int \frac{x}{\sqrt{1-x^2}} dx = x \arccos(x) - \sqrt{1-x^2}$$

La fonction $x \in]-1, 1[\mapsto x \arccos(x) - \sqrt{1-x^2}$ est une primitive de la fonction Arccos sur $] - 1, 1 [$.

► **Application 4.**

$$\int \arcsin(x) dx = \int 1 \times \arcsin(x) dx = x \arcsin(x) - \int \frac{x}{\sqrt{1-x^2}} dx = x \arcsin(x) + \sqrt{1-x^2}$$

La fonction $x \in]-1, 1[\mapsto x \arcsin(x) + \sqrt{1-x^2}$ est une primitive de la fonction Arcsin sur $] - 1, 1 [$.

Remarque. Comme autres exemples d'application (parmi les innombrables conséquences de la formule d'IPP), voir aussi en fin de chapitre les compléments consacrés aux intégrales de Wallis, au lemme de Riemann-Lebesgue, et à l'écriture de e comme limite d'une somme.

8.6 Changement de variable

THÉORÈME 8.5 - (Formule du changement de variable pour les intégrales) Soit f une fonction continue sur un intervalle I , et φ une fonction de classe $\mathcal{C}^1$ sur $[a, b]$ et à valeurs dans I . On a :

$$\int_a^b f(\varphi(t)) \varphi'(t) dt = \int_{\varphi(a)}^{\varphi(b)} f(x) dx$$

PREUVE. Soient f une fonction continue sur un intervalle I , et φ une fonction de classe $\mathcal{C}^1$ sur $[a, b]$ et à valeurs dans I . Puisque f est continue sur I , elle admet des primitives sur I : notons F l'une d'entre elles. La fonction $F \circ \varphi$ est de classe $\mathcal{C}^1$ sur $[a, b]$ et pour tout réel $t \in [a, b]$, on a :

$$(F \circ \varphi)'(t) = F'(\varphi(t)) \varphi'(t) = f(\varphi(t)) \varphi'(t)$$

Par intégration sur l'intervalle $[a, b]$, on en déduit que : $\int_a^b (F \circ \varphi)'(t) dt = \int_a^b f(\varphi(t)) \varphi'(t) dt$ (♠).

Or : $\int_a^b (F \circ \varphi)'(t) dt = [(F \circ \varphi)(t)]_a^b = F(\varphi(b)) - F(\varphi(a)) = \int_{\varphi(a)}^{\varphi(b)} f(t) dt$ (♣).

On déduit de (♠) et de (♣) que : $\int_a^b f(\varphi(t)) \varphi'(t) dt = \int_{\varphi(a)}^{\varphi(b)} f(x) dx$. ☕

Exemple 1. Calcul de $I = \int_1^e \frac{1}{x\sqrt{1+\ln(x)}} dx$.

La fonction $x \mapsto 1/x\sqrt{1+\ln(x)}$ est continue sur $[1, e]$; l'intégrale I est donc bien définie.

Posons : $u = \ln(x)$. Lorsque x varie de 1 à e , u varie de 0 à 1. Par suite :

$$I = \int_0^1 \frac{1}{e^u \sqrt{1+u}} e^u du = \int_0^1 \frac{1}{\sqrt{1+u}} du = 2 \left[(1+u)^{1/2} \right]_0^1 = 2(2^{1/2} - 1)$$

Conclusion : $I = \int_1^e \frac{1}{x \sqrt{1+\ln(x)}} dx = 2(2^{1/2} - 1).$

Exemple 2. Pour tout entier naturel n , on pose $I_n = \int_0^{\pi/2} \cos^n(t) dt$ et $J_n = \int_0^{\pi/2} \sin^n(t) dt$.

Dans l'intégrale I_n , procédons au changement de variable : $u = \frac{\pi}{2} - t$ (de la sorte : $t = \frac{\pi}{2} - u$). On a alors :

$$I_n = \int_{\pi/2}^0 \cos^n\left(\frac{\pi}{2} - u\right) (-du) = \int_0^{\pi/2} \sin^n(t) dt = J_n$$

Conclusion : pour tout n entier naturel, on a : $\int_0^{\pi/2} \cos^n(t) dt = \int_0^{\pi/2} \sin^n(t) dt$

Exemple 3. Posons $I = \int_0^{\pi/3} \frac{1}{\cos(x)} dx$.

Un mot sur la définition de I : la fonction $1/\cos$ est continue sur tout intervalle dans lequel la fonction $\cos$ ne s'annule pas, c'est à dire tout intervalle ne contenant aucun réel congru à $\pi/2$ modulo π .

Calculons I . On utilise le changement de variable $u = \tan(x/2)$, la motivation étant d'utiliser la célèbre formule de trigonométrie : $\cos(x) = \frac{1 - \tan^2(x/2)}{1 + \tan^2(x/2)}$

Avec ce changement de variable, on a $x = 2\arctan(u)$ d'où $dx = \frac{2du}{1+u^2}$. Par ailleurs, lorsque x varie de 0 à $\pi/3$, $u = \tan(x/2)$ varie de 0 à $\sqrt{3}/3$, d'où :

$$I = \int_0^{\sqrt{3}/3} \frac{1+u^2}{1-u^2} \frac{2du}{1+u^2} = \int_0^{\sqrt{3}/3} \frac{2du}{1-u^2}$$

On effectue la décomposition en éléments simples (aisée) de $2/(1-u^2)$ pour obtenir : $\frac{2}{1-u^2} = \frac{1}{1+u} + \frac{1}{1-u}$

D'où : $I = \int_0^{\sqrt{3}/3} \frac{1}{1+u} + \frac{1}{1-u} du = \left[\ln \left| \frac{1+u}{1-u} \right| \right]_0^{\sqrt{3}/3}$ Soit : $\int_0^{\pi/3} \frac{1}{\cos(x)} dx = \ln \left| \frac{3+\sqrt{3}}{3-\sqrt{3}} \right|$.

THÉORÈME 8.6 - (Formule du changement de variable pour les primitives) Soit f une fonction continue et φ une fonction de classe $\mathcal{C}^1$, toutes deux définies sur des intervalles adéquats. Alors :

$$\int f(\varphi(t)) \varphi'(t) dt = \int f(x) dx$$

PREUVE. Conséquence directe du théorème précédent. ☕

Exemple 1. Une primitive de $\frac{1}{\cosh}$ sur $\mathbb{R}$ est $x \mapsto 2\arctan(e^x)$.

Calculons : $I = \int \frac{1}{\operatorname{ch}(x)} dx = \int \frac{2}{e^x + e^{-x}} dx.$

Posons $u = e^x$; alors $x = \ln(u)$ et $dx = \frac{1}{u} du.$

Il s'ensuit que : $I = 2 \int \frac{1}{u + \frac{1}{u}} \times \frac{1}{u} du = 2 \int \frac{1}{u^2 + 1} du = 2\arctan(u).$

En n'oubliant pas de revenir à la variable initiale, on en déduit que : $\boxed{\int \frac{1}{\operatorname{ch}(x)} dx = 2\arctan(e^x)}.$

Exemple 2. Une primitive de $\frac{1}{\operatorname{sh}}$ sur $\mathbb{R}_+^*$ est $x \mapsto \ln\left(\frac{e^x - 1}{e^x + 1}\right).$

Calculons : $I = \int \frac{1}{\operatorname{sh}(x)} dx = \int \frac{2}{e^x - e^{-x}} dx.$

Posons $u = e^x$; alors $x = \ln(u)$ et $dx = \frac{1}{u} du.$

Il s'ensuit que : $I = 2 \int \frac{1}{u - \frac{1}{u}} \times \frac{1}{u} du = 2 \int \frac{1}{u^2 - 1} du = \int \frac{1}{u - 1} du - \int \frac{1}{u + 1} du = \ln\left|\frac{u - 1}{u + 1}\right|.$

D'où, en revenant à la variable initiale : $\boxed{\int \frac{1}{\operatorname{sh}(x)} dx = \ln\left(\frac{e^x - 1}{e^x + 1}\right)}.$

Exemple 3. Une primitive sur $] -\sqrt{a}; \sqrt{a}[$ de $x \mapsto \frac{1}{x^2 - a}$ est

$$x \mapsto \frac{1}{\sqrt{a}} \ln \sqrt{\left| \frac{x - \sqrt{a}}{x + \sqrt{a}} \right|}$$

Soit a un réel strictement positif. Calculons : $I = \int \frac{1}{x^2 - a} dx = \int \frac{1}{(x - \sqrt{a})(x + \sqrt{a})} dx.$

Il existe deux réels α et β tels que : $\forall x \in] -\sqrt{a}; \sqrt{a}[, \quad \frac{1}{(x - \sqrt{a})(x + \sqrt{a})} = \frac{\alpha}{x - \sqrt{a}} + \frac{\beta}{x + \sqrt{a}}$

Pour déterminer α et β , on peut procéder par identification ou utiliser l'astuce que vous a expliquée Mr Deloecker. En multipliant par $(x - \sqrt{a})$ la relation précédente, on obtient :

$$\frac{1}{x + \sqrt{a}} = \alpha + \frac{\beta(x - \sqrt{a})}{x + \sqrt{a}}$$

d'où, en donnant à x la valeur $\sqrt{a}$:² $\alpha = \frac{1}{2\sqrt{a}}.$

On obtient de manière analogue : $\beta = -\frac{1}{2\sqrt{a}}.$

2. Ou, si vous préférez, en prenant la limite de chaque terme de l'égalité lorsque x tend vers $\sqrt{a}$.

Par suite, pour tout réel x dans $] -\sqrt{a}; \sqrt{a}[$, on a : $\frac{1}{x^2 - a} = \frac{1}{2\sqrt{a}} \left(\frac{1}{x - \sqrt{a}} - \frac{1}{x + \sqrt{a}} \right)$

$$\text{D'où : } \int \frac{1}{x^2 - a} dx = \frac{1}{2\sqrt{a}} \left(\int \frac{dx}{x - \sqrt{a}} - \int \frac{dx}{x + \sqrt{a}} \right) = \frac{1}{2\sqrt{a}} (\ln(x - \sqrt{a}) - \ln(x + \sqrt{a}))$$

$$\text{soit finalement : } \boxed{\int \frac{1}{x^2 - a} dx = \frac{1}{\sqrt{a}} \ln \sqrt{\left| \frac{x - \sqrt{a}}{x + \sqrt{a}} \right|}}$$

Exemple 4. Une primitive sur $\mathbb{R}$ de $x \mapsto \frac{1}{x^2 + a}$ est $x \mapsto \frac{1}{\sqrt{a}} \arctan\left(\frac{x}{\sqrt{a}}\right)$.

Soit a un réel strictement positif. Calculons :

$$I = \int \frac{1}{x^2 + a} dx = \int \frac{1}{a \left(\frac{x^2}{a} + 1 \right)} dx = \frac{1}{a} \int \frac{1}{\left(\frac{x}{\sqrt{a}} \right)^2 + 1} dx.$$

On pose $u = \frac{x}{\sqrt{a}}$. Alors $x = \sqrt{a} u$ d'où $dx = \sqrt{a} du$. Ainsi :

$$I = \frac{1}{a} \int \frac{1}{u^2 + 1} \sqrt{a} du = \frac{1}{\sqrt{a}} \arctan(u).$$

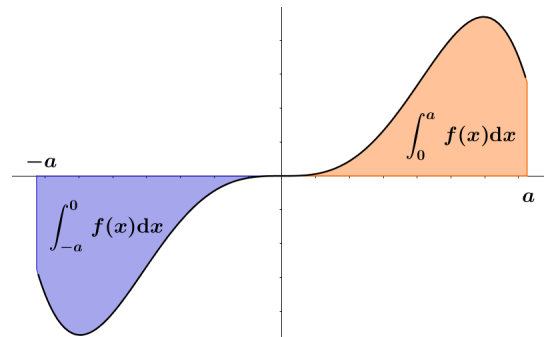
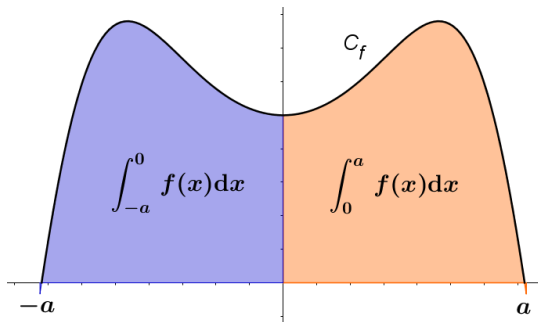
$$\text{D'où, en revenant à la variable initiale : } \boxed{\int \frac{1}{x^2 + a} dx = \frac{1}{\sqrt{a}} \arctan\left(\frac{x}{\sqrt{a}}\right)}$$

8.7 Compléments

8.7.1 Intégrale et parité

PROPRIÉTÉ 8.4 - (Intégrale et parité). Soit $f \in \mathcal{C}^0([-a, a], \mathbb{R})$ (avec $a \in \mathbb{R}_+$).

Si f est paire, alors : $\int_{-a}^a f(x) dx = 2 \int_0^a f(x) dx$ Si f est impaire, alors : $\int_{-a}^a f(x) dx = 0$



PREUVE. Soit $f \in \mathcal{C}^0([-a, a], \mathbb{R})$ (avec $a \in \mathbb{R}_+$).

D'après la relation de Chasles pour les intégrales, on a : $\int_{-a}^a f(x) dx = \int_{-a}^0 f(x) dx + \int_0^a f(x) dx$ (♠)

Le changement de variable $u = -x$ donne : $\int_{-a}^0 f(x) dx = \int_a^0 f(-u) (-du)$

$$\text{Soit : } \int_{-a}^0 f(x) dx = \int_0^a f(-u) du \quad (\clubsuit)$$

► Si f est paire : $\int_0^a f(-u) du = \int_0^a f(u) du$. On en déduit, avec (♣) et (♠) que :

$$\int_{-a}^a f(x) dx = \int_0^a f(x) dx + \int_0^a f(u) du. \text{ Par conséquent : } [f \text{ paire}] \implies \left[\int_{-a}^a f(x) dx = 2 \int_0^a f(x) dx \right]$$

► Si f est impaire : $\int_0^a f(-u) du = \int_0^a -f(u) du = - \int_0^a f(u) du$. On en déduit, avec (♣) et (♠) que :

$$\int_{-a}^a f(x) dx = \int_0^a f(x) dx - \int_0^a f(u) du. \text{ Par conséquent : } [f \text{ impaire}] \implies \left[\int_{-a}^a f(x) dx = 0 \right] \quad \text{☕}$$

Exemple 1. Pour tout entier naturel k , on a : $\int_{-\pi/2}^{\pi/2} \sin^{2k+1}(x) dx = 0$.

Exemple 2. Déterminer une condition nécessaire et suffisante sur les réels a , b et c pour que la fonction polynomiale $x \mapsto ax^2 + bx + c$ soit d'intégrale nulle sur $[-1, 1]$.

CORRIGÉ. Soient a , b et c trois réels. On a :

$$\int_{-1}^1 ax^2 + bx + c dx = \int_{-1}^1 ax^2 + c dx + b \underbrace{\int_{-1}^1 x dx}_{=0} = 2 \int_0^1 ax^2 + c dx = 2 \left(\frac{a}{3} + c \right)$$

Conclusion. $\left[\int_{-1}^1 ax^2 + bx + c dx = 0 \right] \iff [a = -3c].$

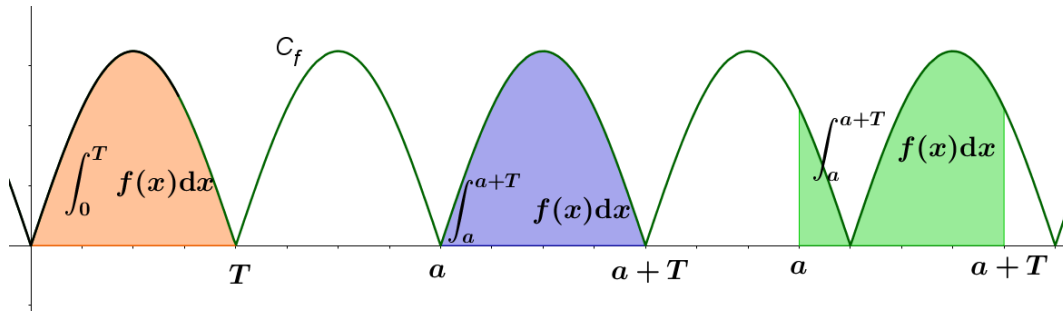
8.7.2 Intégrale et périodicité

PROPRIÉTÉ 8.5 - (Intégrale et périodicité). Soit $f \in \mathcal{C}^0(\mathbb{R}, \mathbb{R})$, T -périodique (avec $T \in \mathbb{R}_+^*$). On a :

$$\forall a \in \mathbb{R}, \int_a^{a+T} f(x) dx = \int_0^T f(x) dx$$

L'intégrale d'une fonction T périodique sur un segment de longueur T est indépendante de l'origine de ce segment.

PREUVE. Soit $f \in \mathcal{C}^0(\mathbb{R}, \mathbb{R})$, T -périodique (avec $T \in \mathbb{R}_+^*$), et soit a un réel. On distingue deux cas suivant que a est un multiple entier de la période T ou non.



► **Premier cas :** $a = 0 \quad [T]$. Dans ce cas : $\exists! k \in \mathbb{Z}$, $a = kT$. On pose alors : $u = x - kT$ (ce qui implique : $du = dx$).

D'après la formule du changement de variable, on a : $\int_a^{a+T} f(x) dx = \int_0^T \underbrace{f(u+kT)}_{=f(u) \text{ car } f \text{ } T\text{-pér.}} du = \int_0^T f(u) du = \int_0^T f(x) dx$

► **Second cas :** $a \neq 0 \quad [T]$. Dans ce cas : $\exists! k \in \mathbb{Z}$, $(k-1)T < a < kT$. On utilise alors la relation de Chasles pour écrire : $\int_a^{a+T} f(x) dx = \int_a^{kT} f(x) dx + \int_{kT}^{a+T} f(x) dx$

Dans la première (*resp.* seconde) intégrale intervenant au second membre de cette égalité, on procède au changement de variable $u = x - (k-1)T$ (*resp.* $u = x - kT$). On obtient alors :

$$\int_a^{a+T} f(x) dx = \int_{a-(k-1)T}^T \underbrace{f(u+(k-1)T)}_{=f(u)} du + \int_0^{a-(k-1)T} \underbrace{f(u+kT)}_{=f(u)} du$$

$$\text{D'où : } \int_a^{a+T} f(x) dx = \int_{a-(k-1)T}^T f(u) du + \int_0^{a-(k-1)T} f(u) du = \int_0^T f(u) du = \int_0^T f(x) dx$$

Conclusion. Pour f continue et T -périodique, on a : $\forall a \in \mathbb{R}$, $\int_a^{a+T} f(x) dx = \int_0^T f(x) dx$



8.7.3 Primitives de polynômes trigonométriques

DÉFINITION 8.4 - Un **polynôme trigonométrique** est une fonction f définie sur $\mathbb{R}$ par :

$$\forall x \in \mathbb{R}, f(x) = \sum_{i=0}^n \sum_{j=0}^p a_{ij} \cos^i(x) \sin^j(x)$$

Par linéarité, le calcul d'une primitive d'un polynôme trigonométrique se ramène au calcul de :

$$F_{i,j} = \int \cos^i(x) \sin^j(x) dx$$

➤ Notons que si $i = 0$ (ou $j = 0$), on peut penser à linéariser. Exemples : les calculs des primitives de $\cos^3$, $\sin^5$, $\cos^6 \dots$ peuvent se faire à l'aide de la formule d'Euler.

➤ Si i est impair (alors $\exists l \in \mathbb{Z}, i = 2l + 1$) : on pose $u = \sin(x)$, pour se ramener au calcul d'une primitive d'une fonction polynomiale $\int (1 - u^2)^l u^j du$. Explicitement :

$$\forall x \in \mathbb{R}, F_{i,j}(x) = \sum_{k=0}^l (-1)^k \binom{l}{k} \frac{\sin^{2k+j+1}(x)}{2k + j + 1}$$

➤ Si j est impair (alors $\exists l \in \mathbb{Z}, j = 2l + 1$) : on pose $u = \cos(x)$, pour se ramener au calcul d'une primitive d'une fonction polynomiale $-\int u^i (1 - u^2)^l du$. Explicitement :

$$\forall x \in \mathbb{R}, F_{i,j}(x) = \sum_{k=0}^l (-1)^{k+1} \binom{l}{k} \frac{\cos^{2k+i+1}(x)}{2k + i + 1}$$

➤ Si i et j sont tous deux pairs, on peut tenter de linéariser et croiser les doigts bien fort pour que les calculs ne soient pas trop longs...

PROPRIÉTÉ 8.6 - Une primitive sur $\mathbb{R}$ de $x \mapsto \cos^{2l+1}(x) \sin^j(x)$ est :

$$x \mapsto \sum_{k=0}^l \binom{l}{k} (-1)^k \frac{\sin^{2k+j+1}(x)}{2k + j + 1}$$

PREUVE. Soient j et l deux entiers naturels. La fonction $f_{ij} : x \in \mathbb{R} \mapsto \cos^{2l+1}(x) \sin^j(x)$ est continue sur $\mathbb{R}$;³ à ce titre, elle admet des primitives sur $\mathbb{R}$. Notons F_{ij} l'une d'entre elles. On a pour tout réel x :

$$F_{ij}(x) = \int \cos^{2l+1}(x) \sin^j(x) dx = \int \cos^{2l}(x) \sin^j(x) \cos(x) dx = \int (1 - \sin^2(x))^l \sin^j(x) \cos(x) dx$$

3. Et même de classe $\mathcal{C}^\infty$ sur $\mathbb{R}$.

On pose : $u = \sin(x)$. On a alors : $du = \cos(x) dx$, et :

$$\begin{aligned} \int (1 - \sin^2(x))^l \sin^j(x) \cos(x) dx &= \int (1 - u^2)^l u^j du = \int \sum_{k=0}^l \binom{l}{k} (-1)^k u^{2k} u^j du \\ &= \sum_{k=0}^l \binom{l}{k} (-1)^k \int u^{2k+j} du = \sum_{k=0}^l \binom{l}{k} (-1)^k \frac{u^{2k+j+1}}{2k+j+1} \end{aligned}$$

Une primitive sur $\mathbb{R}$ de $x \mapsto \cos^{2l+1}(x) \sin^j(x)$ est : $x \mapsto \sum_{k=0}^l \binom{l}{k} (-1)^k \frac{\sin^{2k+j+1}(x)}{2k+j+1}$ ☕

Remarque. Pour achever ce paragraphe, observons que les mêmes raisonnements peuvent être adaptés au calcul d'une primitive de $x \mapsto \operatorname{ch}^i(x) \operatorname{sh}^j(x)$ sur $\mathbb{R}$.

8.7.4 Primitives de $x \mapsto \frac{1}{x^2 + bx + c}$ (b et c réels)

Dans ce paragraphe, il est important de ne pas retenir les formules, mais d'avoir parfaitement compris les techniques permettant de conclure.

Soient b et c deux réels. Pour déterminer une primitive de $f : x \mapsto \frac{1}{x^2 + bx + c}$, on distingue 3 cas, suivant que le discriminant Δ du polynôme $P = X^2 + bX + c$ est strictement positif, nul ou strictement négatif.

1/ Lorsque $\Delta > 0$: le polynôme P possède exactement deux racines réelles α et β . On utilise alors une **décomposition en éléments simples**, càd que l'on détermine deux réels k_1 et k_2 tels que :

$$\forall x \in \mathbb{R} \setminus \{\alpha, \beta\}, \quad \frac{1}{x^2 + bx + c} = \frac{k_1}{x - \alpha} + \frac{k_2}{x - \beta}$$

D'où : $\int \frac{1}{x^2 + bx + c} dx = k_1 \ln |x - \alpha| + k_2 \ln |x - \beta|$

2/ Lorsque $\Delta = 0$: le polynôme P possède une unique racine réelle α . Alors :

$$\forall x \in \mathbb{R} \setminus \{\alpha\}, \quad \frac{1}{x^2 + bx + c} = \frac{1}{(x - \alpha)^2}$$

D'où : $\int \frac{1}{x^2 + bx + c} dx = \frac{-1}{x - \alpha}$

3/ Lorsque $\Delta < 0$: le polynôme P ne possède aucune racine réelle. Moyennant une transformation détaillée dans le cours, on peut alors se ramener au calcul d'une primitive sur $\mathbb{R}$ de $u \mapsto \frac{1}{u^2 + d^2}$ (avec $d \in \mathbb{R}_+^*$) :

On obtient alors : $\int \frac{1}{u^2 + d^2} du = \frac{1}{d} \arctan\left(\frac{u}{d}\right)$ (poser $v = \frac{u}{d}$)

Exemple. $\int \frac{dx}{1 + x + x^2} = \frac{2}{\sqrt{3}} \arctan\left(\frac{2}{\sqrt{3}} \left(x + \frac{1}{2}\right)\right)$

8.8 Extension aux fonctions à valeurs complexes

DÉFINITION 8.5 - Soit f une fonction définie sur un intervalle I et à valeurs dans $\mathbb{C}$. On appelle **partie réelle** (*resp.* **imaginaire**) de f la fonction notée $\operatorname{Re}(f)$ (*resp.* $\operatorname{Im}(f)$) définie sur I en posant :

$$\forall x \in I, [\operatorname{Re}(f)](x) = \operatorname{Re}(f(x)) \quad \text{resp.} \quad \forall x \in I, [\operatorname{Im}(f)](x) = \operatorname{Im}(f(x))$$

Remarque. Les fonctions $\operatorname{Re} f$ et $\operatorname{Im} f$ sont à valeurs réelles.

DÉFINITION 8.6 - Une fonction $f \in \mathbb{C}^I$ est **continue** (*resp.* **dérivable**, *resp.* **de classe** $\mathcal{C}^1, \dots$) sur I si les fonctions $\operatorname{Re} f$ et $\operatorname{Im} f$ sont continues (*resp.* dérivables, *resp.* de classe $\mathcal{C}^1, \dots$) sur I .

En outre, lorsque f est dérivable sur I , on pose : $f' = (\operatorname{Re} f)' + i(\operatorname{Im} f)'$.

Dans “l'autre sens”, une **primitive de f sur I** est une fonction $F \in \mathbb{C}^I$ dérivable sur I telle que $F' = f$.

Remarque. Dans le courant de cette année, nous justifierons la formule : $f' = (\operatorname{Re} f)' + i(\operatorname{Im} f)'$.

PROPRIÉTÉ 8.7 - La fonction $x \in \mathbb{R} \mapsto \frac{1}{\alpha} e^{\alpha x}$ est une primitive sur $\mathbb{R}$ de $x \in \mathbb{R} \mapsto e^{\alpha x}$ (avec $\alpha \in \mathbb{C}^*$).

PREUVE. Soit $\alpha \in \mathbb{C}^*$. Notons $\alpha = a + ib$ (avec a et b réels), et $F : x \in \mathbb{R} \mapsto \frac{1}{\alpha} e^{\alpha x}$.

La fonction F est à valeurs complexes, et on a donc : $F' = (\operatorname{Re} F)' + i(\operatorname{Im} F)'$. Or pour tout réel x on a :

$$\begin{aligned} F(x) &= \frac{1}{a + ib} e^{(a+ib)x} = \frac{a - ib}{a^2 + b^2} e^{ax} (\cos(bx) + i \sin(bx)) \\ \iff F(x) &= \underbrace{\frac{1}{a^2 + b^2} (ae^{ax} \cos(bx) + be^{ax} \sin(bx))}_{=(\operatorname{Re} F)(x)} + i \underbrace{\left(\frac{1}{a^2 + b^2} (ae^{ax} \sin(bx) - be^{ax} \cos(bx)) \right)}_{=(\operatorname{Im} F)(x)} \end{aligned}$$

Ainsi pour tout réel x :

$$\begin{aligned} F'(x) &= \frac{1}{a^2 + b^2} (a^2 e^{ax} \cos(bx) - abe^{ax} \sin(bx) + abe^{ax} \sin(bx) + b^2 e^{ax} \cos(bx)) \\ &\quad + i \left(\frac{1}{a^2 + b^2} (a^2 e^{ax} \sin(bx) + abe^{ax} \cos(bx) - abe^{ax} \cos(bx) + b^2 e^{ax} \sin(bx)) \right) \end{aligned}$$

D'où pour tout réel x :

$$F'(x) = \frac{1}{a^2 + b^2} (a^2 + b^2) e^{ax} \cos(bx) + i \frac{1}{a^2 + b^2} (a^2 + b^2) e^{ax} \sin(bx) = e^{ax} (\cos(bx) + i \sin(bx)) = e^{(a+ib)x}$$

Conclusion. $\forall \alpha \in \mathbb{C}^*, \int e^{\alpha x} dx = \frac{1}{\alpha} e^{\alpha x}$



COROLLAIRE 8.3 - Soient k et μ deux réels non simultanément nuls.

➤ Une primitive sur $\mathbb{R}$ de $x \in \mathbb{R} \mapsto e^{kx} \cos(\mu x)$ est $x \in \mathbb{R} \mapsto \frac{e^{kx}}{k^2 + \mu^2} (k \cos(\mu x) + \mu \sin(\mu x))$.

➤ Une primitive sur $\mathbb{R}$ de $x \in \mathbb{R} \mapsto e^{kx} \sin(\mu x)$ est $x \in \mathbb{R} \mapsto \frac{e^{kx}}{k^2 + \mu^2} (k \sin(\mu x) - \mu \cos(\mu x))$.

PREUVE. Soient k et μ deux réels non simultanément nuls.

Commençons par observer que : $\forall x \in \mathbb{R}, e^{kx} \cos(\mu x) = \operatorname{Re}(e^{kx} e^{i\mu x})$ et $\forall x \in \mathbb{R}, e^{kx} \sin(\mu x) = \operatorname{Im}(e^{kx} e^{i\mu x})$.

Il s'ensuit que :

$$\int e^{kx} \cos(\mu x) dx = \operatorname{Re} \left(\int e^{kx} e^{i\mu x} dx \right) \text{ et } \int e^{kx} \sin(\mu x) dx = \operatorname{Im} \left(\int e^{kx} e^{i\mu x} dx \right) \quad (\spadesuit)$$

Le calcul d'une seule primitive (d'une fonction à valeurs complexes) fournira d'après la précédente formule deux primitives (de fonctions à valeurs réelles) :

$$\int e^{kx} e^{i\mu x} dx = \int e^{(k+i\mu)x} dx \text{ d'où }^4 : \boxed{\int e^{kx} e^{i\mu x} dx = \frac{1}{k+i\mu} e^{(k+i\mu)x} dx} \quad (\clubsuit).$$

Il ne reste plus qu'à extraire la partie réelle et la partie imaginaire de cette dernière expression :

$$\forall x \in \mathbb{R}, \frac{1}{k+i\mu} e^{(k+i\mu)x} = \frac{k-i\mu}{k^2+\mu^2} e^{kx} (\cos(\mu x) + i \sin(\mu x))$$

D'où :

$$\boxed{\forall x \in \mathbb{R}, \frac{1}{k+i\mu} e^{(k+i\mu)x} = \underbrace{\frac{e^{kx}}{k^2+\mu^2} (k \cos(\mu x) + \mu \sin(\mu x))}_{\text{partie réelle}} + i \underbrace{\frac{e^{kx}}{k^2+\mu^2} (k \sin(\mu x) - \mu \cos(\mu x))}_{\text{partie imaginaire}} \quad (\heartsuit)}.$$

On déduit de $(\spadesuit)$, $(\clubsuit)$ et $(\heartsuit)$ les formules données dans l'énoncé. 

8.9 Epilogue

$$\mathbf{8.9.1} \quad e = \lim_{n \rightarrow +\infty} \sum_{k=0}^n \frac{1}{k!}$$

On pose pour tout $n \in \mathbb{N} : I_n = \int_0^1 \frac{(1-x)^n}{n!} e^x dx$.

Etape 1 — Montrons que (I_n) tend vers 0.

Soit n un entier naturel. On a : $|I_n| \leq \int_0^1 \frac{|1-x|^n}{n!} |e^x| dx$ d'où : $|I_n| \leq \int_0^1 \frac{e}{n!} dx$ soit encore : $|I_n| \leq \frac{e}{n!}$.

On en déduit que $\boxed{\lim_{n \rightarrow +\infty} I_n = 0}$.

4. Puisqu'une primitive sur $\mathbb{R}$ de la fonction $x \mapsto e^{\alpha x}$ est la fonction $x \mapsto e^{\alpha x}/\alpha$ (pour tout complexe α non nul).

Remarque : le recours aux valeurs absolues n'est ici pas indispensable, puisque tous les termes intervenant dans la fonction à intégrer sont positifs sur $[0, 1]$. On peut donc bien évidemment reprendre le raisonnement précédent en observant donc que le réel I_n est positif pour tout n , et en écrivant $0 \leq I_n \leq e/n!$.

Etape 2 — Montrons que pour tout $n \in \mathbb{N}$: $I_n = \frac{1}{(n+1)!} + I_{n+1}$.

Soit n un entier naturel. On a, grâce à une intégration par parties⁵ :

$$I_{n+1} = \int_0^1 \frac{(1-x)^{n+1}}{(n+1)!} e^x dx = \frac{1}{(n+1)!} \left([(1-x)^{n+1} e^x]_0^1 + (n+1) \int_0^1 (1-x)^n e^x dx \right)$$

$$\text{soit } I_{n+1} = -\frac{1}{(n+1)!} + \frac{1}{n!} \int_0^1 (1-x)^n e^x dx \quad \text{D'où : } \boxed{I_{n+1} = -\frac{1}{(n+1)!} + I_n}$$

Conclusion.

Soit n un entier naturel non nul. Alors :

$$\sum_{k=0}^n \frac{1}{k!} = 1 + \sum_{k=1}^n \frac{1}{k!} =_{k=K+1} 1 + \sum_{K=0}^{n-1} \frac{1}{(K+1)!} \text{ et d'après ce qui précède : } \frac{1}{(K+1)!} = I_K - I_{K+1}.$$

$$\text{D'où : } \sum_{k=0}^n \frac{1}{k!} = 1 + \sum_{K=0}^{n-1} (I_K - I_{K+1}) \text{ soit : } \sum_{k=0}^n \frac{1}{k!} = 1 + I_0 - I_n \quad (\clubsuit)$$

Il reste à observer que $I_0 = e - 1$ (calcul immédiat) et $\lim_{n \rightarrow +\infty} I_n = 0$ pour déduire de $(\clubsuit)$ la limite :

$$\boxed{e = \lim_{n \rightarrow +\infty} \sum_{k=0}^n \frac{1}{k!}}$$

Remarque : un peu plus tard dans l'année, nous pourrions traduire ce résultat en disant que la *série* de terme général $1/k!$ est *convergente* et a pour *somme* le réel e . Et nous le noterons alors : $e = \sum_{k=0}^{+\infty} \frac{1}{k!}$

8.9.2 Intégrales de Wallis

On appelle **intégrales de Wallis** les intégrales définies en posant pour tout entier naturel n :

$$I_n = \int_0^{\pi/2} \cos^n(t) dt \quad \text{et} \quad J_n = \int_0^{\pi/2} \sin^n(t) dt$$

Observons déjà que :

$$\boxed{J_0 = \int_0^{\pi/2} 1 dt = \frac{\pi}{2}} \text{ et } \boxed{J_1 = \int_0^{\pi/2} \sin(t) dt = [-\cos(t)]_0^{\pi/2} = 1}.$$

5. Légitime car les fonctions polynomiales et la fonction exponentielle sont de classe $\mathcal{C}^1$ sur $\mathbb{R}$ (donc en particulier sur $[0; 1]$).

Par ailleurs : $J_2 = \int_0^{\pi/2} \sin^2(t) dt = \frac{1}{2} \int_0^{\pi/2} 1 - \cos(2t) dt = \frac{1}{2} [t - \sin(2t)/2]_0^{\pi/2}$. Ainsi : $J_2 = \frac{\pi}{4}$.

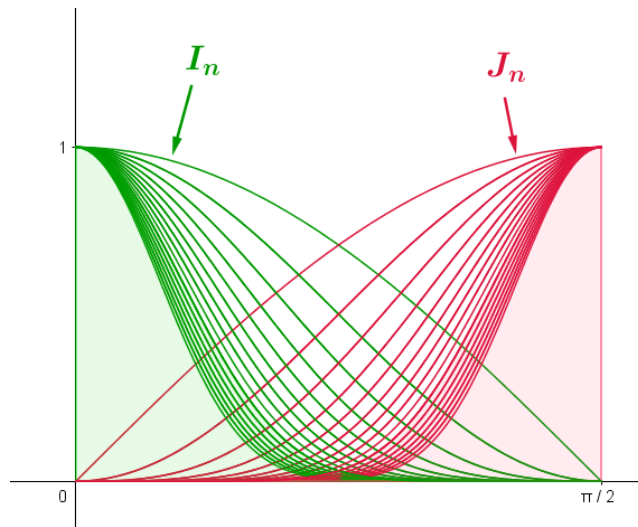
► Propriétés des intégrales de Wallis

PROPRIÉTÉ 8.8 - $\forall n \in \mathbb{N}, I_n = J_n$

PREUVE. Posons : $u = \frac{\pi}{2} - t$. Lorsque t varie de 0 à $\pi/2$, u varie de $\pi/2$ à 0. En outre : $dt = -du$. D'après la formule du changement de variable, on a donc pour tout entier naturel n :

$$I_n = \int_0^{\pi/2} \cos^n(t) dt = - \int_{\pi/2}^0 \cos^n\left(\frac{\pi}{2} - u\right) du = \int_0^{\pi/2} \sin^n(u) du$$

Conclusion. $\forall n \in \mathbb{N}, I_n = J_n$.



PROPRIÉTÉ 8.9 - $\forall n \in \mathbb{N}, I_n \geq 0$.

PREUVE. La fonction $t \mapsto \cos^n(t)$ est positive sur $[0, \pi/2]$. Par positivité de l'intégrale, on en déduit que : $\forall n \in \mathbb{N}, I_n \geq 0$ ☕

PROPRIÉTÉ 8.10 - La suite (I_n) (ou (J_n)) est décroissante.

PREUVE. Soit n un entier naturel. Par linéarité de l'intégrale, on a :

$$J_{n+1} - J_n = \int_0^{\pi/2} \sin^{n+1}(t) dt - \int_0^{\pi/2} \sin^n(t) dt = \int_0^{\pi/2} \sin^{n+1}(t) - \sin^n(t) dt = \int_0^{\pi/2} \sin^n(t) (\sin(t) - 1) dt$$

Or : $\forall t \in [0, \pi/2], \sin^n(t) \geq 0$ et $\sin(t) - 1 \leq 0$. On en déduit que : $\forall t \in [0, \pi/2], \sin^n(t) (\sin(t) - 1) \leq 0$.

Par conséquent : $\int_0^{\pi/2} \sin^n(t) (\sin(t) - 1) dt \leq 0$. D'où : $\forall n \in \mathbb{N}, J_{n+1} - J_n \leq 0$. ☕

COROLLAIRE 8.4 - La suite (I_n) (ou (J_n)) est convergente.

PREUVE. D'après ce qui précède, la suite (I_n) est décroissante et minorée (par 0); d'après le théorème de la limite monotone, elle converge donc (et sa limite est positive ou nulle). ☕

PROPRIÉTÉ 8.11 - $\forall n \in \mathbb{N}, I_{n+2} = \frac{n+1}{n+2} I_n$

PREUVE. Soit n un entier naturel. On a : $I_{n+2} = \int_0^{\pi/2} \cos^{n+2}(t) dt = \int_0^{\pi/2} \cos(t) \cos^{n+1}(t) dt$

On pose alors : $\forall t \in [0, \pi/2], \begin{cases} u(t) = \sin(t) \\ v(t) = \cos^{n+1}(t) \end{cases}$ d'où : $\forall t \in [0, \pi/2], \begin{cases} u'(t) = \cos(t) \\ v'(t) = -(n+1) \sin(t) \cos^n(t) \end{cases}$

Les fonctions u et v sont de classe $\mathcal{C}^1$ sur $[0, \pi/2]$ (théorèmes généraux).

Il est donc légitime d'utiliser une intégration par parties pour écrire :

$$I_{n+2} = \underbrace{[\sin(t) \cos^{n+1}(t)]_0^{\pi/2}}_{=0} + (n+1) \int_0^{\pi/2} \sin^2(t) \cos^n(t) dt$$

$$\text{D'où : } I_{n+2} = (n+1) \int_0^{\pi/2} (1 - \cos^2(t)) \cos^n(t) dt = (n+1) \int_0^{\pi/2} \cos^n(t) dt - (n+1) \int_0^{\pi/2} \cos^{n+2}(t) dt$$

C'est-à-dire : $I_{n+2} = (n+1) I_n - (n+1) I_{n+2}$ d'où : $\forall n \in \mathbb{N}, I_{n+2} = \frac{n+1}{n+2} I_n$ ☕

PROPRIÉTÉ 8.12 - Pour tout entier naturel p :

$$I_{2p} = \frac{(2p)!}{2^{2p} (p!)^2} \frac{\pi}{2} \quad \text{et} \quad I_{2p+1} = \frac{2^{2p} (p!)^2}{(2p+1)!}$$

PREUVE. On souhaite établir que la propriété $P(p)$: $I_{2p} = \frac{(2p)!}{2^{2p} (p!)^2} \frac{\pi}{2}$ est vraie pour tout entier naturel p .

Initialisation : pour $p = 0$, on a d'une part $I_0 = \frac{\pi}{2}$ et d'autre part $\frac{(0)!}{2^0 (0!)^2} \frac{\pi}{2} = \frac{\pi}{2}$. La propriété est initialisée.

Hérédité : supposons $P(p)$ vraie pour un certain entier naturel p , et montrons que $P(p+1)$ l'est.

$$\text{On a : } I_{2(p+1)} = I_{2p+2} = \frac{2p+1}{2p+2} I_{2p} = \frac{2p+1}{2p+2} \frac{(2p)!}{2^{2p} (p!)^2} \frac{\pi}{2} = \frac{2p+2}{2p+2} \frac{2p+1}{2p+2} \frac{(2p)!}{2^{2p} (p!)^2} \frac{\pi}{2} = \frac{(2p+2)!}{2^{2p+2} ((p+1)!)^2} \frac{\pi}{2}$$

Ce qui signifie que $P(p+1)$ est vraie, établit l'hérédité de la propriété, et achève donc cette récurrence.

Conclusion. $\forall p \in \mathbb{N}, I_{2p} = \frac{(2p)!}{2^{2p} (p!)^2} \frac{\pi}{2}$

La preuve de la formule donnant I_{2p+1} est analogue. ☕

PROPRIÉTÉ 8.13 - $\forall n \in \mathbb{N}, (n+1) I_{n+1} I_n = \frac{\pi}{2}$

PREUVE. Pour tout entier naturel n posons : $u_n = (n+1) I_{n+1} I_n$. On a : $u_{n+1} = (n+2) I_{n+2} I_{n+1}$.

Or, d'après une propriété précédente, on a : $u_{n+1} = (n+2) \frac{n+1}{n+2} I_n I_{n+1} = (n+1) I_{n+1} I_n$, soit : $u_{n+1} = u_n$.

On en déduit que (u_n) est constante.

Par suite : $\forall n \in \mathbb{N}, u_n = u_0 = I_1 \times I_0 = \frac{\pi}{2}$.

Conclusion. $\forall n \in \mathbb{N}, (n+1) I_{n+1} I_n = \frac{\pi}{2}$. ☕

PROPRIÉTÉ 8.14 - $\lim_{n \rightarrow +\infty} \frac{I_{n+1}}{I_n} = 1$

PREUVE. Soit n un entier naturel.

La fonction $\cos^n$ étant continue, positive, et ne s'annulant qu'en $\pi/2$ (sur l'intervalle $[0, \pi/2]$), on peut affirmer que : $\forall n \in \mathbb{N}, I_n > 0$.

On sait par ailleurs que la suite (I_n) est décroissante.

On en déduit que : $\forall n \in \mathbb{N}, 0 < I_{n+1} \leq I_n$. Par suite : $\forall n \in \mathbb{N}, 0 < \frac{I_{n+1}}{I_n} \leq 1$ (♠).

Par ailleurs, puisque la suite (I_n) est décroissante, on a : $0 < I_{n+2} \leq I_{n+1} \leq I_n$.

Il s'ensuit que : $\frac{I_{n+1}}{I_n} \geq \frac{I_{n+2}}{I_{n+1}}$.

D'après la propriété 8.11, on a donc : $\frac{I_{n+1}}{I_n} \geq \frac{n+1}{n+2}$.

Ainsi : $\forall n \in \mathbb{N}, \frac{I_{n+1}}{I_n} \geq \frac{n+1}{n+2}$ (♣).

D'après (♠) et (♣) : $\forall n \in \mathbb{N}, \frac{n+1}{n+2} \leq \frac{I_{n+1}}{I_n} \leq 1$. Puisque $\lim_{n \rightarrow +\infty} \frac{n+1}{n+2} = 1$, le théorème d'encadrement

permet d'affirmer que : $\lim_{n \rightarrow +\infty} \frac{I_{n+1}}{I_n} = 1$ ☕

PROPRIÉTÉ 8.15 - $\lim_{p \rightarrow +\infty} \frac{2(2p+1)}{\pi} I_{2p}^2 = 1$

PREUVE. Soit p un entier naturel arbitraire. On a :

$$\begin{aligned} \frac{2(2p+1)}{\pi} I_{2p}^2 &= \frac{2(2p+1)}{\pi} I_{2p}^2 \frac{I_{2p+1}}{I_{2p+1}} = \frac{2}{\pi} \underbrace{[(2p+1) I_{2p+1} I_{2p}]}_{= \frac{\pi}{2} \text{ (d'après 7)}} \times \frac{I_{2p}}{I_{2p+1}} \\ &= \frac{\pi}{2} \text{ (d'après 7)} \end{aligned}$$

D'où : $\frac{2(2p+1)}{\pi} I_{2p}^2 = \frac{I_{2p}}{I_{2p+1}}$. Or, d'après la propriété précédente : $\lim_{p \rightarrow +\infty} \frac{I_{2p}}{I_{2p+1}} = 1$.

Conclusion. $\lim_{p \rightarrow +\infty} \frac{2(2p+1)}{\pi} I_{2p}^2 = 1$. ☕

COROLLAIRE 8.5 - $\lim_{n \rightarrow +\infty} I_n = 0$

PREUVE. On sait déjà que la suite (I_n) est convergente vers une limite ℓ positive ou nulle. Si ℓ était strictement positive, alors on aurait $\lim_{p \rightarrow +\infty} \frac{2(2p+1)}{\pi} I_{2p}^2 = +\infty$, ce qui contredirait la propriété précédente.

On en déduit que : $\lim_{n \rightarrow +\infty} I_n = 0$. ☕

COROLLAIRE 8.6 - $\pi = \lim_{p \rightarrow +\infty} \frac{2^{4p} (p!)^4}{p ((2p)!)^2}$

PREUVE. D'après les propriétés 8.15 et 8.12, on a : $\lim_{p \rightarrow +\infty} \frac{2(2p+1)}{\pi} \left(\frac{(2p)!}{2^{2p} (p!)^2} \frac{\pi}{2} \right)^2 = 1$.

D'où : $\lim_{p \rightarrow +\infty} (2p+1) \frac{\pi}{2} \frac{((2p)!)^2}{2^{4p} (p!)^4} = 1$. Donc : $\lim_{p \rightarrow +\infty} \frac{(2p+1)}{2p} \times (2p) \times \frac{\pi}{2} \frac{((2p)!)^2}{2^{4p} (p!)^4} = 1$

Par suite : $\lim_{p \rightarrow +\infty} \pi \frac{p ((2p)!)^2}{2^{4p} (p!)^4} = 1$, et donc : $\lim_{p \rightarrow +\infty} \frac{p ((2p)!)^2}{2^{4p} (p!)^4} = \frac{1}{\pi}$.

Finalement : $\pi = \lim_{p \rightarrow +\infty} \frac{2^{4p} (p!)^4}{p ((2p)!)^2}$. ☕

8.9.3 Formule de Stirling

L'énoncé ci-dessous est extrait d'un problème dont la première partie était consacrée aux intégrales de Wallis. La partie reproduite ici a pour objectif d'établir la formule de Stirling, qui donne un équivalent de $n!$.

On définit trois suites réelles u , S et v en posant :

$$\forall n \in \mathbb{N}^*, u_n = \left(n + \frac{1}{2}\right) \ln \left(1 + \frac{1}{n}\right) - 1; \quad \forall n \in \mathbb{N}^*, S_n = \sum_{k=1}^n u_k; \quad \text{et : } \forall n \in \mathbb{N}, n \geq 2, v_n = e^{1-S_{n-1}}$$

Soit n un entier naturel non nul.

14) Etablir que : $u_n = \int_n^{n+1} \frac{2(n-t)+1}{2t} dt$; puis que $u_n = \int_0^1 \frac{1-2u}{2(n+u)} du$

Pour la seconde formule, on utilisera un changement de variable très simple.

15) On pose : $r_n = \int_0^{1/2} \frac{1-2u}{2(n+u)} du$ et $t_n = \int_{1/2}^1 \frac{2u-1}{2(n+u)} du$.

a) A l'aide du sens de variation de $u \mapsto \frac{1}{n+u}$, établir les encadrements suivants :

$$\frac{1}{4(2n+1)} \leq r_n \leq \frac{1}{8n} \quad \text{et} \quad \frac{1}{8(n+1)} \leq t_n \leq \frac{1}{4(2n+1)}$$

b) En déduire un encadrement de u_n , puis montrer que : $\forall n \in \mathbb{N}^*, S_n \leq \frac{1}{8}$.

NB : on peut observer que la suite (S_n) est croissante, et majorée (par $1/8$) ; à ce titre, elle est convergente. Il s'ensuit que la suite (v_n) est convergente. On pose, jusqu'à la fin de ce problème :

$$\ell = \lim_{n \rightarrow +\infty} v_n$$

16) a) Montrer que pour tout $n \in \mathbb{N}^*$, on a :

$$S_n = \left(n + \frac{1}{2}\right) \ln(n+1) - n - \ln(n!)$$

b) En déduire l'expression de v_n en fonction de n .

c) A partir de l'expression de v_n , montrer que : $\lim_{n \rightarrow +\infty} \frac{n!}{\ell n^n e^{-n} \sqrt{n}} = 1$

► PARTIE E - Conclusion

17) A l'aide de la question 16-c, établir que : $\lim_{n \rightarrow +\infty} \frac{I_{2n} \ell \sqrt{2n}}{\pi} = 1$

18) En déduire la valeur de ℓ .

CORRIGÉ.

14) Soit n un entier naturel non nul. On a :

$$\int_n^{n+1} \frac{2(n-t)+1}{2t} dt = \int_n^{n+1} -1 + \frac{2n+1}{2t} dt = -1 + \frac{2n+1}{2} [\ln(t)]_n^{n+1} = -1 + \left(n + \frac{1}{2}\right) (\ln(n+1) - \ln(n))$$

$$\text{D'où : } \int_n^{n+1} \frac{2(n-t)+1}{2t} dt = -1 + \left(n + \frac{1}{2}\right) \left(\ln\left(1 + \frac{1}{n}\right)\right).$$

$$\text{Par suite : } \forall n \in \mathbb{N}^*, u_n = \int_n^{n+1} \frac{2(n-t)+1}{2t} dt.$$

Dans l'intégrale précédente, on procède au changement de variable : $t = n + u$. Lorsque t varie de n à $(n+1)$, $u = t - n$ varie de 0 à 1 et $dt = du$. On en déduit grâce à la formule du changement de variable que :

$$\int_n^{n+1} \frac{2(n-t)+1}{2t} dt = \int_0^1 \frac{-2u+1}{2(n+u)} du = \int_0^1 \frac{1-2u}{2(n+u)} du$$

$$\text{Ainsi : } \forall n \in \mathbb{N}^*, u_n = \int_0^1 \frac{1-2u}{2(n+u)} du.$$

15) a) La fonction $n \mapsto \frac{1}{n+u}$ est décroissante sur $[0, 1]$. Or une fonction décroissante sur un segment $[a, b]$ atteint son maximum en a , et son minimum en b . On en déduit que :

$$\left[\begin{array}{l} \forall u \in \left[0, \frac{1}{2}\right], \quad \frac{1}{2n+1} \leq \frac{1}{2(n+u)} \leq \frac{1}{2n} \quad (\spadesuit) \\ \forall u \in \left[\frac{1}{2}, 1\right], \quad \frac{1}{2(n+1)} \leq \frac{1}{2(n+u)} \leq \frac{1}{2n+1} \quad (\clubsuit) \end{array} \right.$$

► Soit u un réel dans $\left[0, \frac{1}{2}\right]$. On a alors : $(1-2u) \geq 0$. D'où, grâce à $(\spadesuit)$:

$$\forall u \in \left[0, \frac{1}{2}\right], \quad \frac{1-2u}{2n+1} \leq \frac{1-2u}{2(n+u)} \leq \frac{1-2u}{2n}$$

Par croissance de l'intégrale, on obtient en intégrant terme à terme cet encadrement :

$$\frac{1}{2n+1} \int_0^{1/2} 1-2u \, du \leq r_n \leq \frac{1}{2n} \int_0^{1/2} 1-2u \, du$$

$$\text{Or : } \int_0^{1/2} 1-2u \, du = [u - u^2]_0^{1/2} = \frac{1}{4}. \quad \text{Donc : } \frac{1}{2n+1} \times \frac{1}{4} \leq r_n \leq \frac{1}{2n} \times \frac{1}{4}.$$

Ainsi : $\forall n \in \mathbb{N}^*, \quad \frac{1}{4(2n+1)} \leq r_n \leq \frac{1}{8n} \quad (\heartsuit)$

► Soit u un réel dans $\left[\frac{1}{2}, 1\right]$. On a alors : $(2u-1) \geq 0$. D'où, grâce à $(\clubsuit)$:

$$\forall u \in \left[\frac{1}{2}, 1\right], \quad \frac{2u-1}{2(n+1)} \leq \frac{2u-1}{2(n+u)} \leq \frac{2u-1}{2n+1}$$

Par croissance de l'intégrale, on obtient en intégrant terme à terme cet encadrement :

$$\frac{1}{2(n+1)} \int_{1/2}^1 2u-1 \, du \leq t_n \leq \frac{1}{2n+1} \int_{1/2}^1 2u-1 \, du$$

$$\text{Or : } \int_{1/2}^1 2u-1 \, du = [u^2 - u]_{1/2}^1 = \frac{1}{4}. \quad \text{Donc : } \frac{1}{2(n+1)} \times \frac{1}{4} \leq t_n \leq \frac{1}{2n+1} \times \frac{1}{4}.$$

Ainsi : $\forall n \in \mathbb{N}^*, \quad \frac{1}{8(n+1)} \leq t_n \leq \frac{1}{4(2n+1)} \quad (\diamondsuit)$

b) Soit n un entier naturel non nul. D'après la relation de Chasles pour les intégrales, on a :

$$u_n = \int_0^{1/2} \frac{1-2u}{2(n+u)} \, du + \int_{1/2}^1 \frac{1-2u}{2(n+u)} \, du = \int_0^{1/2} \frac{1-2u}{2(n+u)} \, du - \int_{1/2}^1 \frac{2u-1}{2(n+u)} \, du$$

En d'autres termes : $u_n = r_n - t_n$. On déduit de cette remarque et des encadrements $(\heartsuit)$ et $(\diamondsuit)$ de la question précédente que :

$$\forall n \in \mathbb{N}^*, \quad 0 \leq u_n \leq \frac{1}{8} \left(\frac{1}{n} - \frac{1}{n+1} \right).$$

On en déduit que : $S_n = \sum_{k=1}^n u_k \leq \sum_{k=1}^n \frac{1}{8} \left(\frac{1}{k} - \frac{1}{k+1} \right)$. D'où, "télescopiquement" : $S_n \leq \frac{1}{8} \left(1 - \frac{1}{n+1} \right)$.

En particulier : $\boxed{\forall n \in \mathbb{N}^*, \quad S_n \leq \frac{1}{8}}.$

16) a) Prouvons par récurrence sur l'entier naturel non nul n l'assertion :

$$P(n) : \quad S_n = \left(n + \frac{1}{2} \right) \ln(n+1) - n - \ln(n!)$$

Initialisation : pour $n = 1$, on a d'une part $S_1 = u_1 = \frac{3}{2} \ln(2) - 1$.

D'autre part : $\left(1 + \frac{1}{2} \right) \ln(1+1) - 1 - \ln(1!) = \frac{3}{2} \ln(2) - 1$.

Il s'ensuit que $P(1)$ est vraie.

Hérédité : supposons $P(n)$ vraie pour un certain entier naturel non nul n , et montrons que $P(n+1)$ l'est.

$$\text{On a : } S_{n+1} = S_n + u_{n+1} = \underbrace{\left(n + \frac{1}{2} \right) \ln(n+1) - n - \ln(n!)}_{=S_n \text{ (hyp. de réc.)}} + \underbrace{\left(n + \frac{3}{2} \right) \ln\left(1 + \frac{1}{n+1} \right) - 1}_{=u_{n+1} \text{ (énoncé)}}$$

$$\text{D'où : } S_{n+1} = \left(n + \frac{1}{2} \right) \ln(n+1) - n - \ln(n!) + \left(n + \frac{3}{2} \right) \ln(n+2) - \left(n + \frac{3}{2} \right) \ln(n+1) - 1$$

$$\iff S_{n+1} = \left(n + \frac{3}{2} \right) \ln(n+2) - (n+1) - \ln(n!) - \ln(n+1)$$

$$\iff S_{n+1} = \left(n + \frac{3}{2} \right) \ln(n+2) - (n+1) - \ln((n+1)!)$$

Ce qui signifie que $P(n+1)$ est vraie, et établit l'hérédité de la propriété.

$$\boxed{\text{Conclusion. } \forall n \in \mathbb{N}^*, \quad S_n = \left(n + \frac{1}{2} \right) \ln(n+1) - n - \ln(n!)}.$$

b) Soit n un entier supérieur ou égal à 2. D'après l'énoncé et la question précédente :

$$v_n = \exp(1 - S_{n-1}) = \exp\left(1 - \left(n - \frac{1}{2} \right) \ln(n) + n - 1 + \ln((n-1)!) \right) = e^n n^{-n+1/2} (n-1)!$$

$$\text{Donc : } v_n = \frac{e^n \sqrt{n} (n-1)!}{n^n} = \frac{e^n n!}{n^n \sqrt{n}} = \frac{n!}{n^n e^{-n} \sqrt{n}}.$$

$$\boxed{\text{Conclusion. } \forall n \in \mathbb{N}, \quad n \geq 2, \quad v_n = \frac{n!}{n^n e^{-n} \sqrt{n}}}.$$

c) D'après l'énoncé : $\lim_{n \rightarrow +\infty} v_n = \ell$. D'où : $\lim_{n \rightarrow +\infty} \frac{v_n}{\ell} = 1$.⁶

On en déduit avec la question précédente que : $\lim_{n \rightarrow +\infty} \frac{n!}{\ell n^n e^{-n} \sqrt{n}} = 1$.

► PARTIE E - Conclusion

17) Soit n un entier naturel supérieur ou égal à 2. On a, d'après la question 5 :

$$\frac{I_{2n} \ell \sqrt{2n}}{\pi} = \frac{(2n)!}{2^{2n} (n!)^2} \frac{\pi}{2} \times \frac{\ell \sqrt{2n}}{\pi} \quad \text{soit :} \quad \frac{I_{2n} \ell \sqrt{2n}}{\pi} = \frac{(2n)!}{(n!)^2} \frac{\ell \sqrt{2n}}{2^{2n+1}} \quad (\spadesuit)$$

Or, d'après la question précédente, on a :

$$\lim_{n \rightarrow +\infty} \frac{(2n)!}{\ell (2n)^{2n} e^{-2n} \sqrt{2n}} = 1 \quad (\heartsuit) \quad \text{et} \quad \lim_{n \rightarrow +\infty} \frac{((n!)^2)}{\ell^2 n^{2n} e^{-2n} n} = 1 \quad (\diamondsuit)$$

On écrit donc judicieusement :

$$\frac{I_{2n} \ell \sqrt{2n}}{\pi} = \frac{(2n)!}{\ell (2n)^{2n} e^{-2n} \sqrt{2n}} \times \frac{\ell^2 n^{2n} e^{-2n} n}{(n!)^2} \times \underbrace{\frac{\ell (2n)^{2n} e^{-2n} \sqrt{2n}}{\ell^2 n^{2n} e^{-2n} n}}_{=1 \text{ (formidable!)}} \times \frac{\ell \sqrt{2n}}{2^{2n+1}}$$

$$\text{Ainsi :} \quad \frac{I_{2n} \ell \sqrt{2n}}{\pi} = \frac{(2n)!}{\ell (2n)^{2n} e^{-2n} \sqrt{2n}} \times \frac{\ell^2 n^{2n} e^{-2n} n}{(n!)^2}.$$

Il résulte de cette égalité, de $(\heartsuit)$ et de $(\diamondsuit)$ que : $\lim_{n \rightarrow +\infty} \frac{I_{2n} \ell \sqrt{2n}}{\pi} = 1$.

18) On peut déduire de la question 12 que : $\lim_{n \rightarrow +\infty} \sqrt{\frac{2}{\pi}} \times \sqrt{2n} \times I_{2n} = 1$.

Donc : $\lim_{n \rightarrow +\infty} \frac{\sqrt{2n} \times I_{2n}}{\pi} = \frac{1}{\sqrt{2\pi}}$. On en déduit, grâce à la question 17, que : $\ell = \sqrt{2\pi}$.

On a ainsi démontré la **formule de Stirling** : $\lim_{n \rightarrow +\infty} \frac{n!}{\sqrt{2\pi n} n^n e^{-n} \sqrt{n}} = 1$

ou encore : $\lim_{n \rightarrow +\infty} \frac{n!}{\sqrt{2n\pi} \left(\frac{n}{e}\right)^n} = 1$

Autrement énoncée :

THÉORÈME 8.7 - (Formule de Stirling)

$$n! \sim_{+\infty} \sqrt{2n\pi} \left(\frac{n}{e}\right)^n$$

6. On pouvait admettre que $\ell \neq 0$ à ce point du devoir. En voici toutefois la justification. D'après la question 15, la suite (S_n) admet une limite finie, disons ℓ' . Donc la suite (v_n) converge vers $\ell = e^{1-\ell'}$, qui est non nulle.

8.9.4 Lemme de Riemann-Lebesgue

PROPRIÉTÉ 8.16 - (Riemann-Lebesgue). Pour toute fonction ψ de classe $\mathcal{C}^1$ sur $[0, \pi]$, on a :

$$\lim_{n \rightarrow +\infty} \int_0^\pi \psi(t) \sin(nt) dt = 0$$

PREUVE. Soient $n \in \mathbb{N}^*$ et ψ une fonction de classe $\mathcal{C}^1$ sur $[0, \pi]$.

On pose :
$$\begin{cases} u(t) = \psi(t) \\ v(t) = -\frac{\cos(nt)}{n} \end{cases} \quad \text{d'où : } \begin{cases} u'(t) = \psi'(t) \\ v'(t) = \sin(nt) \end{cases}$$

Les fonctions u et v sont de classe $\mathcal{C}^1$ sur $[0, \pi]$, et on peut donc utiliser une intégration par parties pour obtenir :

$$\int_0^\pi \psi(t) \sin(nt) dt = \left[-\frac{\cos(nt)}{n} \psi(t) \right]_0^\pi + \frac{1}{n} \int_0^\pi \cos(nt) \psi'(t) dt$$

On en déduit que :

$$\int_0^\pi \psi(t) \sin\left(\left(n + \frac{1}{2}\right)t\right) dt = \underbrace{\frac{\psi(0) + (-1)^{n+1} \psi(\pi)}{n}}_{=A_n} + \frac{1}{n} \underbrace{\int_0^\pi \cos(nt) \psi'(t) dt}_{=B_n}$$

D'après l'inégalité triangulaire, on a : $\left| \int_0^\pi \psi(t) \sin(nt) dt \right| \leq |A_n| + \frac{1}{n} |B_n|$.

► Il est clair que : $|A_n| = \frac{|\psi(0) + (-1)^{n+1} \psi(\pi)|}{n}$

► Par ailleurs : $|B_n| \leq \int_0^\pi \underbrace{|\cos(nt)|}_{\leq 1} \times \underbrace{|\psi'(t)|}_{\leq M} dt \leq \int_0^\pi M dt = M\pi$

On déduit de ces inégalités que pour toute fonction ψ de classe $\mathcal{C}^1$ sur $[0, \pi]$:

$$\exists M \in \mathbb{R}_+, \forall n \in \mathbb{N}^*, \left| \int_0^\pi \psi(t) \sin(nt) dt \right| \leq \frac{|\psi(0) + (-1)^{n+1} \psi(\pi)|}{n} + \frac{M\pi}{n}$$

Or : $\lim_{n \rightarrow +\infty} \left[\frac{|\psi(0) + (-1)^{n+1} \psi(\pi)|}{n} + \frac{M\pi}{n} \right] = 0$. On en déduit que : $\lim_{n \rightarrow +\infty} \left| \int_0^\pi \psi(t) \sin(nt) dt \right| = 0$.

D'où finalement :
$$\lim_{n \rightarrow +\infty} \int_0^\pi \psi(t) \sin(nt) dt = 0$$
 ☕

8.9.5 Calcul de $\zeta(2)$

ENONCÉ.

Nous avons établi en début d'année que lorsque l'on fait tendre n vers $+\infty$, la somme

$$S_n = 1 + \frac{1}{4} + \frac{1}{9} + \cdots + \frac{1}{n^2}$$

tend vers une limite finie. Cette limite finie est notée $\zeta(2)$, et on a donc

$$\zeta(2) = \lim_{n \rightarrow +\infty} \left(\sum_{k=1}^n \frac{1}{k^2} \right)$$

L'objectif de ce problème est de déterminer la valeur exacte de cette limite.

PARTIE 1 : POUR SE RAFRAÎCHIR LA MÉMOIRE

- 1) Etablir que pour tout entier $k \geq 2$, on a : $\frac{1}{k^2} \leq \frac{1}{k-1} - \frac{1}{k}$.
- 2) En déduire que pour tout entier naturel n non nul on a : $1 \leq \sum_{k=1}^n \frac{1}{k^2} \leq 2$.
- 3) Pour tout entier naturel n non nul, on pose : $u_n = \sum_{k=1}^n \frac{1}{k^2}$.
 - a) Etablir que la suite (u_n) est croissante.
 - b) Déduire des questions précédentes que la suite est convergente.

PARTIE 2 : ÉTUDE D'UNE FONCTION

On considère la fonction h définie sur $[0, \pi]$ par : $\forall t \in [0, \pi], h(t) = \frac{t^2}{2\pi} - t$.

Et on définit une seconde fonction φ sur $[0, \pi]$ en posant :

$$\forall t \in [0, \pi], \quad \varphi(t) = \begin{cases} \frac{h(t)}{2 \sin\left(\frac{t}{2}\right)} & \text{si } t \neq 0 \\ -1 & \text{si } t = 0 \end{cases}$$

- 4) Quelle est la limite de $\frac{\sin(t)}{t}$ lorsque t tend vers 0 ? La réponse devra être justifiée.
- 5) Déduire de la question précédente les limites de $\frac{t}{2 \sin\left(\frac{t}{2}\right)}$ puis de $\frac{h(t)}{2 \sin\left(\frac{t}{2}\right)}$ lorsque t tend vers 0.

Dans la suite du problème, on pourra supposer que φ est de classe $\mathcal{C}^1$ sur $[0, \pi]$.

PARTIE 3 : SOMMES ET INTÉGRALES

6) Montrer que pour tout entier naturel k , on a : $\cos(k\pi) = (-1)^k$.

7) Etablir que :

$$\forall k \in \mathbb{N}^*, \int_0^\pi t \cos(kt) \, dt = \frac{(-1)^k - 1}{k^2}$$

8) Etablir que :

$$\forall k \in \mathbb{N}^*, \int_0^\pi t^2 \cos(kt) \, dt = \frac{2\pi(-1)^k}{k^2}$$

9) Pour tout entier naturel k non nul, on pose :

$$I_k = \int_0^\pi h(t) \cos(kt) \, dt$$

Déduire de ce qui précède l'expression exacte de I_k en fonction de k .

10) Soit $n \in \mathbb{N}^*$. Montrer que :

$$\forall t \in]0, \pi], \sum_{k=1}^n \cos(kt) = \cos\left(\left(\frac{n+1}{2}\right)t\right) \frac{\sin\left(\frac{nt}{2}\right)}{\sin\left(\frac{t}{2}\right)}$$

11) Déduire de la question précédente l'existence d'une constante réelle λ telle que :

$$\forall t \in]0, \pi], \sum_{k=1}^n \cos(kt) = \frac{\sin\left(\left(n + \frac{1}{2}\right)t\right)}{2 \sin\left(\frac{t}{2}\right)} - \lambda$$

PARTIE 4 : ÉPILOGUE

12) Soit ψ une fonction de classe $\mathcal{C}^1$ sur $[0; \pi]$. Montrer que :

$$\forall n \in \mathbb{N}^*, \left| \int_0^\pi \psi(t) \sin\left(\left(n + \frac{1}{2}\right)t\right) dt \right| \leq \frac{|\psi(0)|}{n + \frac{1}{2}} + \left(\frac{\pi}{n + \frac{1}{2}} \times M \right)$$

où l'on a posé : $M = \max_{[0, \pi]} |\psi'(t)|$.⁷

13) Déduire de la question précédente que pour toute fonction ψ de classe $\mathcal{C}^1$ sur $[0; \pi]$, on a :

$$\lim_{n \rightarrow +\infty} \left(\int_0^\pi \psi(t) \sin\left(\left(n + \frac{1}{2}\right)t\right) dt \right) = 0$$

7. La fonction ψ' est continue sur $[0, \pi]$, puisque l'on a supposé ψ de classe $\mathcal{C}^1$. Or nous verrons prochainement (et nous admettons provisoirement) qu'une fonction continue définie sur un intervalle borné est elle-même bornée. Donc dans le cas présent, ψ' est bornée sur $[0, \pi]$, ce qui assure l'existence d'un réel M tel que : $\forall t \in [0, \pi], |\psi'(t)| \leq M$.

14) Etablir que :

$$\forall n \in \mathbb{N}^*, \quad \sum_{k=1}^n \frac{1}{k^2} = \int_0^\pi h(t) \left(\sum_{k=1}^n \cos(kt) \right) dt$$

15) Dédurre de ce qui précède la valeur exacte de :

$$\zeta(2) = \lim_{n \rightarrow +\infty} \left(\sum_{k=1}^n \frac{1}{k^2} \right)$$

CORRIGÉ

PARTIE 1 : POUR SE RAFRAÎCHIR LA MÉMOIRE

1) Soit $k \geq 2$ un entier. On a : $\boxed{\frac{1}{k-1} - \frac{1}{k} = \frac{1}{k^2 - k} \geq \frac{1}{k^2}}$.

2) Soit n un entier naturel non nul. On a :

$$\sum_{k=1}^n \frac{1}{k^2} = 1 + \sum_{k=2}^n \frac{1}{k^2} \leq 1 + \sum_{k=2}^n \frac{1}{k-1} - \frac{1}{k} \text{ d'où : } \sum_{k=1}^n \frac{1}{k^2} \leq 1 + 1 - \frac{1}{n}$$

On en déduit que : $\forall n \in \mathbb{N}^*, 1 \leq \sum_{k=1}^n \frac{1}{k^2} \leq 2 - \frac{1}{n}$ donc : $\boxed{\forall n \in \mathbb{N}^*, 1 \leq \sum_{k=1}^n \frac{1}{k^2} \leq 2}$

3) a) $\forall n \in \mathbb{N}^*, u_{n+1} - u_n = \frac{1}{(n+1)^2} \geq 0$. Donc : $\boxed{(u_n) \text{ est croissante}}$.

b) La suite (u_n) est croissante (d'après la propriété précédente) et majorée (d'après 2), donc $\boxed{\text{convergente}}$.

PARTIE 2 : ÉTUDE D'UNE FONCTION

4) On a : $\boxed{\lim_{t \rightarrow 0} \frac{\sin(t)}{t}} = \lim_{t \rightarrow 0} \frac{\sin(t) - \sin(0)}{t} = \sin'(0) = \cos(0) \boxed{= 1}$.

5) On commence par observer que d'après ce qui précède : $\lim_{t \rightarrow 0} \frac{t}{\sin(t)} = 1$.

Puis, en procédant au changement de variable $T = t/2$, on obtient : $\lim_{t \rightarrow 0} \frac{t}{2 \sin(t/2)} = \lim_{T \rightarrow 0} \frac{T}{\sin(T)} = 1$.

Par ailleurs, pour tout t non nul : $\frac{h(t)}{2 \sin(\frac{t}{2})} = \frac{1}{2\pi} \frac{t^2}{2 \sin(\frac{t}{2})} - \frac{t}{2 \sin(\frac{t}{2})}$.

On a déjà vu que : $\lim_{t \rightarrow 0} \frac{t}{2 \sin(t/2)} = 1$ d'où $\lim_{t \rightarrow 0} \frac{t^2}{2 \sin(t/2)} = 0$. Par suite : $\boxed{\lim_{t \rightarrow 0} \frac{h(t)}{2 \sin(\frac{t}{2})} = -1}$.

PARTIE 3 : SOMMES ET INTÉGRALES

$$6) \boxed{\forall k \in \mathbb{N}, \cos(k\pi) = \frac{e^{ik\pi} + e^{-ik\pi}}{2} = \frac{(e^{i\pi})^k + (e^{-i\pi})^k}{2} = (-1)^k}.$$

$$7) \text{ Soit } k \in \mathbb{N}^*. \text{ On pose : } \begin{cases} u(t) = t \\ v(t) = \frac{\sin(kt)}{k} \end{cases} \quad \text{d'où : } \begin{cases} u'(t) = 1 \\ v'(t) = \cos(kt) \end{cases}$$

Les fonctions u et v sont de classe $\mathcal{C}^1$ sur $[0, \pi]$, et on peut donc utiliser une intégration par parties pour obtenir :

$$\int_0^\pi t \cos(kt) \, dt = \underbrace{\left[\frac{t \sin(kt)}{k} \right]_0^\pi}_{=0} - \frac{1}{k} \int_0^\pi \sin(kt) \, dt = \frac{1}{k^2} [\cos(kt)]_0^\pi = \frac{\cos(k\pi) - 1}{k^2}$$

$$\text{d'où : } \boxed{\int_0^\pi t \cos(kt) \, dt = \frac{(-1)^k - 1}{k^2}}$$

$$8) \text{ Soit } k \in \mathbb{N}^*. \text{ On pose : } \begin{cases} u(t) = t^2 \\ v(t) = \frac{\sin(kt)}{k} \end{cases} \quad \text{d'où : } \begin{cases} u'(t) = 2t \\ v'(t) = \cos(kt) \end{cases}$$

Les fonctions u et v sont de classe $\mathcal{C}^1$ sur $[0, \pi]$, et on peut donc utiliser une intégration par parties pour obtenir :

$$\int_0^\pi t^2 \cos(kt) \, dt = \underbrace{\left[\frac{t^2 \sin(kt)}{k} \right]_0^\pi}_{=0} - \underbrace{\frac{2}{k} \int_0^\pi t \sin(kt) \, dt}_{=J_k} \quad (\spadesuit)$$

$$\text{Calculons } J_k. \text{ On pose : } \begin{cases} U(t) = t \\ V(t) = -\frac{\cos(kt)}{k} \end{cases} \quad \text{d'où : } \begin{cases} U'(t) = 1 \\ V'(t) = \sin(kt) \end{cases}$$

Les fonctions U et V sont de classe $\mathcal{C}^1$ sur $[0, \pi]$, et on peut donc utiliser une intégration par parties pour obtenir :

$$J_k = \int_0^\pi t \sin(kt) \, dt = \left[-\frac{t \cos(kt)}{k} \right]_0^\pi + \frac{1}{k} \int_0^\pi \cos(kt) \, dt = -\frac{\pi \cos(k\pi)}{k} + \frac{1}{k^2} \underbrace{[\sin(kt)]_0^\pi}_{=0} = -\frac{\pi(-1)^k}{k} + \frac{1}{k^2} \quad (\clubsuit)$$

$$\text{On déduit alors de } (\spadesuit) \text{ et de } (\clubsuit) \text{ que : } \boxed{\forall k \in \mathbb{N}^*, \int_0^\pi t^2 \cos(kt) \, dt = \frac{2\pi(-1)^k}{k^2}}$$

9) Soit $k \in \mathbb{N}^*$.

$$\begin{aligned} \text{On a : } I_k &= \int_0^\pi h(t) \cos(kt) \, dt = \int_0^\pi \left(\frac{t^2}{2\pi} - t \right) \cos(kt) \, dt = \frac{1}{2\pi} \int_0^\pi t^2 \cos(kt) \, dt - \int_0^\pi t \cos(kt) \, dt \\ &= \frac{(-1)^k}{k^2} - \frac{(-1)^k - 1}{k^2} = \frac{1}{k^2} \end{aligned}$$

$$\boxed{\text{Conclusion : } \forall k \in \mathbb{N}^*, I_k = \int_0^\pi h(t) \cos(kt) \, dt = \frac{1}{k^2} .}$$

10) Soient $t \in]0; \pi]$ et $n \in \mathbb{N}^*$. Observons que : $\sum_{k=1}^n \cos(kt) = \operatorname{Re} \left(\underbrace{\sum_{k=1}^n \exp^{ikt}}_S \right)$, et calculons S .

$S = \sum_{k=1}^n \exp^{ikt} = \sum_{k=1}^n (\exp^{it})^k = e^{it} \frac{1 - e^{int}}{1 - e^{it}}$ (puisque $e^{it} \neq 1$). Puis on applique la technique de l'angle moitié :

$$S = e^{it} \frac{e^{int/2} e^{-int/2} - e^{int/2}}{e^{it/2} e^{-it/2} - e^{it/2}} = e^{i \frac{n+1}{2} t} \frac{-2i \sin(nt/2)}{-2i \sin(t/2)} \quad \text{d'où : } \boxed{S = e^{i \frac{n+1}{2} t} \frac{\sin(nt/2)}{\sin(t/2)}}$$

On en déduit, en identifiant les parties réelles de l'égalité ci-dessus que :

$$\boxed{\forall t \in]0; \pi], \forall n \in \mathbb{N}^*, \sum_{k=1}^n \cos(kt) = \cos\left(\frac{n+1}{2}t\right) \frac{\sin(nt/2)}{\sin(t/2)}}.$$

11) Soient $t \in]0; \pi]$ et $n \in \mathbb{N}^*$. On a :

$$\begin{aligned} \cos\left(\frac{n+1}{2}t\right) \sin\left(\frac{nt}{2}\right) &= \frac{\sin\left(\frac{nt}{2} + \frac{n+1}{2}t\right) + \sin\left(\frac{nt}{2} - \frac{n+1}{2}t\right)}{2} = \frac{\sin\left(\frac{2n+1}{2}t\right) + \sin\left(-\frac{t}{2}\right)}{2} \\ &= \frac{\sin\left(\left(n + \frac{1}{2}\right)t\right) - \sin\left(\frac{t}{2}\right)}{2} \end{aligned}$$

De cette identité et de la question précédente, on déduit que :

$$\boxed{\forall t \in]0; \pi], \forall n \in \mathbb{N}^*, \sum_{k=1}^n \cos(kt) = \frac{\sin\left(\left(n + \frac{1}{2}\right)t\right)}{2 \sin(t/2)} - \frac{1}{2} \quad (\text{d'où } \lambda = \frac{1}{2})}.$$

PARTIE 4 : ÉPILOGUE

12) Soient $n \in \mathbb{N}^*$ et ψ une fonction de classe $\mathcal{C}^1$ sur $[0, \pi]$.

$$\text{On pose : } \begin{cases} u(t) = \psi(t) \\ v(t) = -\frac{\cos\left(\left(n + \frac{1}{2}\right)t\right)}{n + \frac{1}{2}} \end{cases} \quad \text{d'où : } \begin{cases} u'(t) = \psi'(t) \\ v'(t) = \sin\left(\left(n + \frac{1}{2}\right)t\right) \end{cases}$$

Les fonctions u et v sont de classe $\mathcal{C}^1$ sur $[0, \pi]$, et on peut donc utiliser une intégration par parties pour obtenir :

$$\int_0^\pi \psi(t) \sin\left(\left(n + \frac{1}{2}\right)t\right) dt = \left[-\frac{\cos\left(\left(n + \frac{1}{2}\right)t\right)}{n + \frac{1}{2}} \psi(t) \right]_0^\pi + \frac{1}{n + \frac{1}{2}} \int_0^\pi \cos\left(\left(n + \frac{1}{2}\right)t\right) \psi'(t) dt$$

En observant que $\cos\left(\left(n + \frac{1}{2}\right)\pi\right) = 0$, on déduit de ce qui précède :

$$\int_0^\pi \psi(t) \sin\left(\left(n + \frac{1}{2}\right)t\right) dt = \underbrace{\frac{\psi(0)}{n + \frac{1}{2}}}_{=A} + \frac{1}{n + \frac{1}{2}} \underbrace{\int_0^\pi \cos\left(\left(n + \frac{1}{2}\right)t\right) \psi'(t) dt}_{=B}$$

D'après l'inégalité triangulaire, on a : $\left| \int_0^\pi \psi(t) \sin\left(\left(n + \frac{1}{2}\right)t\right) dt \right| \leq |A| + \frac{1}{n + \frac{1}{2}} |B|$.

► Il est clair que : $|A| = \frac{|\psi(0)|}{n + \frac{1}{2}}$

► Par ailleurs : $|B| \leq \int_0^\pi \underbrace{\left| \cos\left(\left(n + \frac{1}{2}\right)t\right) \right|}_{\leq 1} \times \underbrace{|\psi'(t)|}_{\leq M} dt \leq \int_0^\pi M dt = M\pi$

On déduit de ces inégalités que pour toute fonction ψ de classe $\mathcal{C}^1$ sur $[0, \pi]$:

$$\forall n \in \mathbb{N}^*, \left| \int_0^\pi \psi(t) \sin\left(\left(n + \frac{1}{2}\right)t\right) dt \right| \leq \frac{|\psi(0)|}{n + \frac{1}{2}} + \frac{M\pi}{n + \frac{1}{2}}$$

13) D'après la question précédente, si ψ est de classe $\mathcal{C}^1$ sur $[0, \pi]$, alors :

$$\forall n \in \mathbb{N}^*, \left| \int_0^\pi \psi(t) \sin\left(\left(n + \frac{1}{2}\right)t\right) dt \right| \leq \frac{|\psi(0)|}{n + \frac{1}{2}} + \frac{M\pi}{n + \frac{1}{2}}$$

Comme : $\lim_{n \rightarrow +\infty} \frac{|\psi(0)|}{n + \frac{1}{2}} + \frac{M\pi}{n + \frac{1}{2}} = 0$, on en déduit que $\lim_{n \rightarrow +\infty} \left| \int_0^\pi \psi(t) \sin\left(\left(n + \frac{1}{2}\right)t\right) dt \right| = 0$ d'où :

$$\lim_{n \rightarrow +\infty} \int_0^\pi \psi(t) \sin\left(\left(n + \frac{1}{2}\right)t\right) dt = 0$$

14) D'après la question 9, on a : $\forall k \in \mathbb{N}^*, \int_0^\pi h(t) \cos(kt) dt = \frac{1}{k^2}$

En sommant les égalités obtenues en faisant varier k de 1 à n (avec $n \in \mathbb{N}^*$) dans la relation ci-dessus, on obtient :

$$\sum_{k=1}^n \frac{1}{k^2} = \sum_{k=1}^n \int_0^\pi h(t) \cos(kt) dt = \int_0^\pi h(t) \sum_{k=1}^n \cos(kt) dt$$

la deuxième égalité provenant de la linéarité de l'intégrale.

$$\textbf{Conclusion.} \forall n \in \mathbb{N}^*, \sum_{k=1}^n \frac{1}{k^2} = \int_0^\pi h(t) \sum_{k=1}^n \cos(kt) dt$$

15) Soit $n \in \mathbb{N}^*$. D'après la question précédente :

$$\sum_{k=1}^n \frac{1}{k^2} = \int_0^\pi h(t) \sum_{k=1}^n \cos(kt) \, dt$$

Et d'après la question 11 :
$$\sum_{k=1}^n \frac{1}{k^2} = \int_0^\pi h(t) \left(\frac{\sin\left(\left(n + \frac{1}{2}\right)t\right)}{2 \sin(t/2)} - \frac{1}{2} \right) dt$$

D'où :
$$\sum_{k=1}^n \frac{1}{k^2} = \int_0^\pi \frac{h(t)}{2 \sin(t/2)} \sin\left(\left(n + \frac{1}{2}\right)t\right) dt - \frac{1}{2} \int_0^\pi h(t) dt$$

Soit :
$$\sum_{k=1}^n \frac{1}{k^2} = \int_0^\pi \varphi(t) \sin\left(\left(n + \frac{1}{2}\right)t\right) dt - \frac{1}{2} \int_0^\pi h(t) dt \quad (*)$$

D'après l'énoncé, la fonction φ est de classe $\mathcal{C}^1$ sur $[0, \pi]$. On peut donc lui appliquer le résultat de la question 13 pour obtenir :
$$\lim_{n \rightarrow +\infty} \int_0^\pi \varphi(t) \sin\left(\left(n + \frac{1}{2}\right)t\right) dt = 0$$

On déduit de cette observation et de la relation (*) que :
$$\lim_{n \rightarrow +\infty} \sum_{k=1}^n \frac{1}{k^2} = -\frac{1}{2} \int_0^\pi h(t) dt \quad (**)$$

Il reste à calculer :
$$I = \int_0^\pi h(t) dt = \int_0^\pi \frac{t^2}{2\pi} - t dt = \frac{1}{6\pi} [t^3]_0^\pi - \frac{1}{2} [t^2]_0^\pi = -\frac{\pi^2}{3}$$

On déduit de ce calcul et de (**) que :

$$\lim_{n \rightarrow +\infty} \left(\sum_{k=1}^n \frac{1}{k^2} \right) = \frac{\pi^2}{6} \quad \text{c'est à dire} \quad \zeta(2) = \frac{\pi^2}{6}$$

Chapitre 9

Equations différentielles linéaires

Convention : dans ce chapitre, I est un intervalle non-vide de $\mathbb{R}$, et $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$.

9.1 Equations différentielles linéaires d'ordre 1

9.1.1 Généralités

DÉFINITION 9.1 - Soient a , b et c trois fonctions de $\mathcal{C}^0(I, \mathbb{K})$. Une fonction $f \in \mathcal{C}^1(I, \mathbb{K})$ est dite **solution sur I de l'équation différentielle** $a(x)y' + b(x)y = c(x)$ si :

$$\forall x \in I, \quad a(x)f'(x) + b(x)f(x) = c(x)$$

Terminologie. (E) $a(x)y' + b(x)y = c(x)$ est appelée **équation** différentielle linéaire d'ordre 1 (ou EDL1) **avec second membre**; et (H) $a(x)y' + b(x)y = 0$ est appelée **équation** (différentielle linéaire d'ordre 1) **homogène** (ou **sans second membre**) **associée à (E)** .

Par ailleurs, on note parfois simplement $ay' + by = c$ l'équation différentielle E .

Exemples. La fonction exponentielle est solution sur $\mathbb{R}$ de l'EDL1 homogène $y' - y = 0$. La fonction $\sin$ est solution sur $\mathbb{R}$ de l'EDL1 avec second membre $\cos(x)y' + \sin(x)y = 1$.

THÉORÈME 9.1 - (Structure de l'ensemble des solutions d'une EDL1). Soient a , b et c trois fonctions de $\mathcal{C}^0(I, \mathbb{K})$.

On note (E) l'EDL1 : $a(x)y' + b(x)y = c(x)$, et (H) : $a(x)y' + b(x)y = 0$ l'équation homogène associée à (E) .

On suppose qu'il existe une solution φ de (E) sur I .

Alors, pour toute fonction $f \in \mathcal{C}^1(I, \mathbb{K})$:

$$[f \text{ est solution de } (E)] \iff [(f - \varphi) \text{ est solution de } (H)]$$

PREUVE. Soit f une fonction de classe $\mathcal{C}^1$ sur I à valeurs dans $\mathbb{K}$. Avec les notations et hypothèses de l'énoncé :

f est solution de (E)

$$\iff \forall x \in I, \quad a(x)f'(x) + b(x)f(x) = c(x)$$

$$\iff \forall x \in I, \quad a(x)f'(x) + b(x)f(x) = a(x)\varphi'(x) + b(x)\varphi(x)$$

$$\iff \forall x \in I, \quad a(x)(f'(x) - \varphi'(x)) + b(x)(f(x) - \varphi(x)) = 0$$

$$\iff \forall x \in I, \quad a(x)(f - \varphi)'(x) + b(x)(f - \varphi)(x) = 0 \quad (\text{linéarité de la dérivation})$$

$$\iff (f - \varphi) \text{ est solution de } (H) \quad \text{☕}$$

En d'autres termes, et avec les notations précédemment introduites, toute solution f de (E) s'écrit $f_H + f_P$ où f_H (*resp.* f_P) désigne une solution quelconque de l'équation homogène associée à (E) (*resp.* une solution particulière de (E)).

Ce résultat fournit donc une méthode (et le plan de ce paragraphe consacré aux EDL1) de résolution, que voici :

MÉTHODE “UNIVERSELLE” DE RÉOLUTION DES EDL

1) Choix de l'intervalle de résolution (s'il n'est pas imposé dans l'énoncé).

2) Résolution de l'équation homogène (ou sans second membre) associée.

3) Recherche d'une solution particulière de l'équation complète (avec second membre).

4) Solutions de l'EDL complète obtenues en “faisant la somme des deux étapes précédentes”.

9.1.2 Résolution d'une EDL 1 homogène

THÉORÈME 9.2 - (Solution générale d'une EDL1 homogène). Soient a et b deux fonctions de $\mathcal{C}^0(I, \mathbb{K})$.

On note (H) l'EDL1 : $a(x)y' + b(x)y = 0$.

On suppose que a ne s'annule pas sur I : $\forall x \in I, a(x) \neq 0$.

Alors, pour toute fonction $f \in \mathcal{C}^1(I, \mathbb{K})$:

$$[f \text{ est solution de } (H)] \iff [\exists K \in \mathbb{K}, \forall x \in I, f(x) = Ke^{-A(x)}]$$

où A désigne une primitive sur I de $\frac{b}{a}$.

Terminologie. Le théorème ci-dessus se traduit aussi en écrivant que la **solution générale** de H est :

$$\forall K \in \mathbb{K}, \forall x \in I, f(x) = Ke^{-A(x)}$$

Entre nous, et si personne ne nous écoute, nous pourrions parfois commettre l'abus d'écrire que la solution générale de E est : $f(x) = Ke^{-\int (b(x)/a(x)) dx}$.

PREUVE. Sous les hypothèses de l'énoncé, la fonction b/a est continue sur I , en tant que quotient de fonctions continues sur I dont le dénominateur ne s'annule pas sur I . A ce titre, b/a admet des primitives sur I ; notons A l'une d'entre elles.

Soit φ une solution de (E). Introduisons une fonction auxiliaire g en posant : $\forall x \in I, g(x) = \varphi(x)e^{A(x)}$.¹

La fonction g est dérivable (et même de classe $\mathcal{C}^1$) sur I , et :

$$\begin{aligned}\forall x \in I, g'(x) &= \varphi'(x)e^{A(x)} + A'(x)\varphi(x)e^{A(x)} \\ \forall x \in I, g'(x) &= \varphi'(x)e^{A(x)} + \frac{b(x)}{a(x)}\varphi(x)e^{A(x)} \\ \iff \forall x \in I, g'(x) &= e^{A(x)} \left(\varphi'(x) + \frac{b(x)}{a(x)}\varphi(x) \right)\end{aligned}$$

La fonction a ne s'annulant pas sur I , cette dernière assertion est encore équivalente à la suivante :

$$\forall x \in I, a(x)g'(x) = e^{A(x)} \underbrace{(a(x)\varphi'(x) + b(x)\varphi(x))}_{=0}$$

la nullité de la parenthèse du terme de droite provenant de ce que φ est solution de (E).

Par suite : $\forall x \in I, a(x)g'(x) = 0$ d'où : $\forall x \in I, g'(x) = 0$ (puisque a ne...).

Il s'ensuit que g est constante sur I , explicitement :

$$\exists K \in \mathbb{K}, g(x) = K \iff \exists K \in \mathbb{K}, \varphi(x)e^{A(x)} = K \iff \exists K \in \mathbb{K}, \varphi(x) = Ke^{-A(x)}$$

On vient donc d'établir que toute solution de (E) est "de la forme $Ke^{-A(x)}$ ".

La réciproque (toute fonction "de la forme $ke^{-A(x)}$ " est solution de (E)) est une vérification aisée (et laissée en exercice) qui permet de conclure :

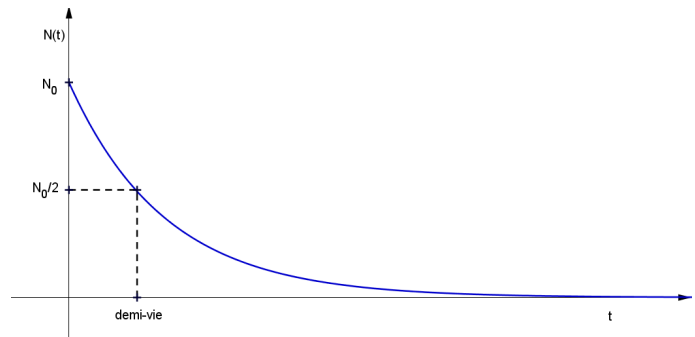
La solution générale de l'EDL $a(x)y' + b(x)y = 0$ est : $\forall x \in I, f_K(x) = Ke^{-A(x)}$ ($K \in \mathbb{K}$), où A désigne une primitive sur I de la fonction $\frac{b}{a}$. ☕

Exemple 1 (Désintégration radioactive). Dans un tissu radioactif, les lois de la Physique permettent d'affirmer que la vitesse de désintégration des noyaux radioactifs (à l'instant t) est proportionnelle au nombre $N(t)$ de noyaux radioactifs présents dans le tissu à l'instant t .²

Ce phénomène, appelé désintégration radioactive, peut être modélisé par le *problème de Cauchy*³ suivant :

$$\begin{cases} \forall t \in [0; +\infty[, & \frac{dN(t)}{dt} + \lambda N(t) = 0 \\ N(0) = N_0 \end{cases}$$

où N_0 est le nombre de noyaux radioactifs présents à l'instant $t = 0$, et λ désigne une constante strictement positive.



1. Dans un certain sens, la fonction g "mesure l'écart" séparant une solution de (H) de la fonction $x \in I \mapsto e^{-A(x)}$. Au regard de l'énoncé, l'objectif est de prouver que g est constante sur I .

2. Pour plus de précisions à ce sujet, consulter M Roveillo!

3. On appelle problème de Cauchy (d'ordre 1) la donnée d'une EDL1 et d'une condition initiale. Ce type de problème est d'usage fréquent en Physique et en Chimie.

Déterminons l'expression de $N(t)$ pour tout réel $t \geq 0$. Avec les notations introduites depuis le début de ce chapitre, ceci revient à résoudre sur $I = [0; +\infty[$ l'EDL1 homogène : $(H) \ y' + \lambda y = 0$.

Dans la présente situation, on a pour tout réel $t \geq 0$:

$$a(t) = 1; \quad b(t) = \lambda; \quad A(t) = \int \frac{b(t)}{a(t)} dt = \lambda t$$

D'après le théorème 9.2, la solution générale de (H) est : $\forall t \geq 0, y(t) = Ke^{-\lambda t}$ (avec $K \in \mathbb{R}$).⁴

La condition initiale permet de déterminer la valeur de K :

$$y(0) = N_0 \iff Ke^{-\lambda \times 0} = N_0 \iff K = N_0$$

On en déduit, en revenant aux notations initiales, que le nombre $N(t)$ de noyaux radioactifs présents dans le tissu à l'instant t est donné par : $\boxed{\forall t \geq 0, N(t) = N_0 e^{-\lambda t}}$.

Remarque. Dans ce contexte, on appelle période de *demi-vie* de l'élément radioactif la valeur de T telle que : $N(T) = N_0/2$. D'après ce qui précède, on a donc :

$$N_0 e^{-\lambda T} = \frac{N_0}{2} \iff e^{-\lambda T} = \frac{1}{2} \iff T = \frac{\ln(2)}{\lambda}$$

Exemple 2. Résolution dans $\mathbb{R}$ de (H) : $\cos(x)y' + \sin(x)y = 0$.

Dans la présente situation, on a : $a(x) = \cos(x)$; $b(x) = \sin(x)$ et $c(x) = 1$. Les trois fonctions a , b et c sont continues sur $\mathbb{R}$, mais a s'annule en $\frac{\pi}{2} + k\pi$ (avec $k \in \mathbb{Z}$). On peut donc choisir comme intervalle de résolution tout intervalle $I_k =]-\frac{\pi}{2} + k\pi; \frac{\pi}{2} + k\pi[$ (avec $k \in \mathbb{Z}$). Pour fixer les idées, choisissons de résoudre (H) sur $I_0 =]-\frac{\pi}{2}; \frac{\pi}{2}[$.

Calculons une primitive de b/a sur I_0 :

$$\forall x \in I_0, A(x) = \int \frac{\sin(x)}{\cos(x)} dx = \int \tan(x) dx = -\ln(|\cos(x)|) = \ln\left(\frac{1}{\cos(x)}\right)$$

la disparition des valeurs absolue étant rendue légitime par la positivité de la fonction $\cos$ sur l'intervalle I_0 .

On en déduit, d'après le théorème 9.2, que la solution générale de (H) sur I_0 est :

$$\forall x \in I_0, f(x) = Ke^{-A(x)} = Ke^{\ln(\cos(x))} \text{ soit : } \boxed{\forall x \in I_0, f(x) = K \cos(x) \text{ (avec } K \in \mathbb{R})}.$$

Exemple 3. Résolution dans $\mathbb{R}$ de (H) : $xy' + (x-1)y = 0$.

Commençons par choisir l'intervalle de résolution. Les trois fonctions a , b et c sont continues sur $\mathbb{R}$, mais a s'annule en 0. On peut choisir de résoudre (H) sur $]0; +\infty[$ ou sur $] -\infty; 0[$. Posons dans un premier temps $I =]0; +\infty[$. On a :

$$\forall x \in \mathbb{R}_+, \frac{b(x)}{a(x)} = \frac{x-1}{x} = 1 - \frac{1}{x}; \quad \text{une primitive de } \frac{b}{a} \text{ sur } \mathbb{R}_+^* \text{ est donc } A : x \in \mathbb{R}_+^* \mapsto x - \ln(x).$$

D'après le théorème 9.2, la solution générale de (H) est $f_K : x \in \mathbb{R}_+^* \mapsto Ke^{\ln(x)-x}$ soit

$$\boxed{f_K : x \in \mathbb{R}_+^* \mapsto Kxe^{-x} \text{ (} K \in \mathbb{R})}.$$

4. Dans le contexte, la fonction recherchée est évidemment à valeurs dans $\mathbb{R}$.

9.1.3 Recherche d'une solution particulière d'une EDL1 avec second membre

On présente ci-dessous deux méthodes générales pour déterminer une solution particulière de $ay' + by = c$.

9.1.3.1 Le flair !

Dans certaines situations, il est facile de deviner la forme d'une solution particulière.

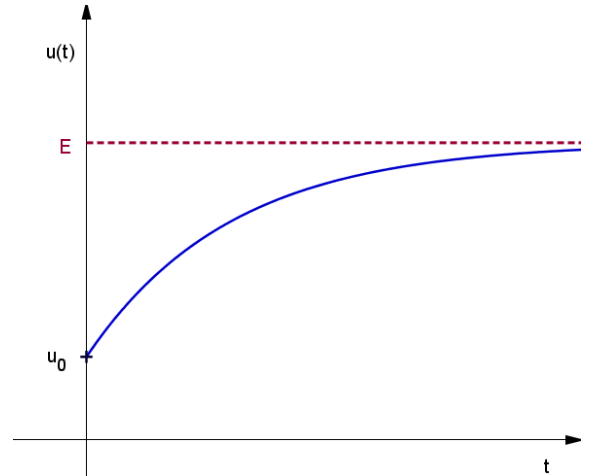
Par exemple, il est immédiat qu'une solution de l'EDL1 $\cos(x)y' + \sin(x)y = 1$ est la fonction sinus.

Par ailleurs, en Physique, les coefficients de l'EDL (les fonctions “ a , b et c ”) sont souvent des constantes ; dans ce cas, on pourra chercher (et trouver !) une solution particulière f_P constante elle aussi. On donne ci-dessous un exemple d'illustration.

Exemple (Charge d'un condensateur).

On peut étudier l'évolution de la tension $u(t)$ lors de la charge d'un condensateur à travers un résistor, sous une tension E . La tension $u(t)$ est dans ce cas régie par le problème de Cauchy suivant :

$$\begin{cases} \forall t \in [0; +\infty[, & \frac{du(t)}{dt} + \frac{u(t)}{RC} = \frac{E}{RC} \\ u(0) = u_0 & (u_0 \in \mathbb{R}_+) \end{cases}$$



Déterminons l'expression de $u(t)$ pour tout réel $t \geq 0$.

On commence par résoudre l'équation homogène associée : (H) $\frac{du(t)}{dt} + \frac{u(t)}{RC} = 0$.

On a : $\forall t \geq 0, \quad \frac{b(t)}{a(t)} = \frac{1}{RC} \quad \text{d'où : } \forall t \in \mathbb{R}_+, \quad A(t) = \int \frac{b(t)}{a(t)} dt = \frac{t}{RC}.$

La solution générale de (H) est donc : $\forall t \in \mathbb{R}_+, \quad u(t) = Ke^{-t/RC}$ (avec K réel).

Puisque les coefficients de l'équation avec second membre sont constants, on peut trouver une solution particulière u_P constante elle aussi. Explicitement, un calcul immédiat assure que la fonction constante égale à E est solution de l'équation complète.

D'après le théorème 9.1, la solution générale de $\frac{du(t)}{dt} + \frac{u(t)}{RC} = \frac{E}{RC}$ est :

$$\boxed{\forall t \in \mathbb{R}_+, \quad u(t) = E + Ke^{-t/RC} \quad (K \in \mathbb{R})}.$$

Pour achever les calculs et déterminer la valeur de K , on exploite la condition initiale :

$$u(0) = u_0 \iff E + Ke^{-0/RC} = u_0 \iff K = u_0 - E$$

Finalement : $\boxed{\forall t \in \mathbb{R}_+, \quad u(t) = E + (u_0 - E)e^{-t/RC}}$ (remarque : $\lim_{t \rightarrow +\infty} u(t) = E$).

9.1.3.2 Méthode de variation de la constante

Présentons le principe de cette méthode sur un exemple. Considérons l'EDL1 (E) : $xy' + (x-1)y = x^3$.

On a déjà résolu (voir exemple 3, page 200) l'équation homogène associée, et obtenu :

$$\forall x \in \mathbb{R}_+^*, f_H(x) = Cxe^{-x} \quad (C \in \mathbb{R})$$

La méthode *de variation de la constante* consiste à chercher une solution particulière f_P de l'équation avec second membre sous la forme :

$$\forall x \in \mathbb{R}_+^*, f_P(x) = C(x)xe^{-x} \quad \text{où } C \text{ est de classe } \mathcal{C}^1 \text{ sur } \mathbb{R}_+^*$$

Sous cette forme, f_P est dérivable sur $\mathbb{R}_+^*$ et : $\forall x \in \mathbb{R}_+^*, f'_P(x) = C'(x)xe^{-x} + C(x)(e^{-x} - xe^{-x})$

C'est-à-dire : $\forall x \in \mathbb{R}_+^*, f'_P(x) = C'(x)xe^{-x} + C(x)e^{-x}(1-x)$

Par suite : $\forall x \in \mathbb{R}_+^*, xf'_P(x) + (x-1)f_P(x) = C'(x)x^2e^{-x} + \underbrace{C(x)e^{-x}x(1-x) + C(x)e^{-x}x(x-1)}_{=0}$

D'où : $\forall x \in \mathbb{R}_+^*, xf'_P(x) + (x-1)f_P(x) = C'(x)x^2e^{-x}$

On en déduit que f_P est solution de (E) si et seulement si : $\forall x \in \mathbb{R}_+^*, C'(x)x^2e^{-x} = x^3$, soit : $\forall x \in \mathbb{R}_+^*, C'(x) = xe^x$. Il reste donc à déterminer une primitive sur $\mathbb{R}_+^*$ de $x \mapsto xe^x$, *via* une intégration par parties :

$$\int xe^x dx = xe^x - \int e^x dx = xe^x - e^x = e^x(x-1) \text{ donc : } \forall x \in \mathbb{R}_+^*, C(x) = e^x(x-1)$$

Une solution particulière de (E) est donc $f_P : x \in \mathbb{R}_+^* \mapsto x(x-1)$.

On déduit des calculs précédents et du théorème 9.1 que la solution générale de (E) est :

$$\forall x \in \mathbb{R}_+^*, f(x) = x(x-1) + Cxe^{-x} \quad (C \in \mathbb{R}) \quad \text{soit : } \boxed{\forall x \in \mathbb{R}_+^*, f(x) = x(x-1 + Ce^{-x}) \quad (C \in \mathbb{R})}$$

Remarque : très peu de modifications sont nécessaires pour résoudre cette EDL sur $\mathbb{R}_-^*$... De fait, la solution générale de (E) sur $\mathbb{R}_-^*$ est $g : x \in \mathbb{R}_-^* \mapsto x(x-1 + Ce^{-x}) \quad (C \in \mathbb{R})$.

Remarquons encore qu'au cours des calculs précédents, un "petit miracle" s'est produit dans la recherche d'une solution particulière de l'équation avec second membre : la "disparition" des termes en $C(x)$, qui nous a permis de d'achever nos calculs.

La propriété suivante assure que ce "miracle" n'en est pas un, et que la méthode de variation de la constante marche "à tous les coups".⁵

5. Ce qui ne signifie pas qu'il faut l'utiliser à tous les coups, voir paragraphe précédent.

THÉORÈME 9.3 - (Existence d'une solution pour une EDL1 avec second membre). Soient a, b et c trois fonctions de $\mathcal{C}^0(I, \mathbb{K})$.

On note (E) l'EDL1 : $a(x)y' + b(x)y = c(x)$.

On suppose que a ne s'annule pas sur I : $\forall x \in I, a(x) \neq 0$.

Alors il existe une fonction f_P de classe $\mathcal{C}^1$ sur I solution de (E) .

Plus précisément, il existe une fonction $K \in \mathcal{C}^1(I, \mathbb{K})$ telle que la fonction f_P définie par

$$\forall x \in I, \quad f_P(x) = K(x)e^{-A(x)} \quad (\text{où } A \text{ est une primitive de } \frac{b}{a} \text{ sur } I)$$

est solution de l'équation différentielle (E) .

PREUVE. Soit $K \in \mathcal{C}^1(I, \mathbb{K})$. Posons : $\forall x \in I, f_P(x) = K(x)e^{-A(x)}$. La fonction f_P est de classe $\mathcal{C}^1$ sur I (théorèmes généraux). A ce titre, elle est dérivable sur I et :

$$\forall x \in I, f'_P(x) = e^{-A(x)} (K'(x) - A'(x)K(x)) = e^{-A(x)} \left(K'(x) - \frac{b(x)}{a(x)} K(x) \right)$$

On en déduit que :

f_P est solution de (E)

$$\iff \forall x \in I, a(x)f'_P(x) + b(x)f_P(x) = c(x)$$

$$\iff \forall x \in I, a(x)e^{-A(x)} \left(K'(x) - \frac{b(x)}{a(x)} K(x) \right) + b(x)K(x)e^{-A(x)} = c(x)$$

$$\iff \forall x \in I, a(x)e^{-A(x)} K'(x) \underbrace{- b(x)K(x)e^{-A(x)} + b(x)K(x)e^{-A(x)}}_{=0} = c(x)$$

$$\iff \forall x \in I, a(x)e^{-A(x)} K'(x) = c(x)$$

$$\iff \forall x \in I, K'(x) = \frac{c(x)}{a(x)} e^{A(x)} \quad (\spadesuit)$$

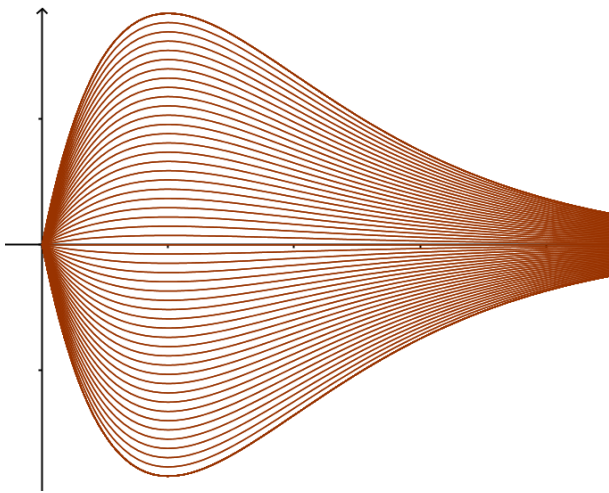
Les hypothèses du théorème et les théorèmes généraux sur la continuité assure que la fonction $x \mapsto \frac{c(x)}{a(x)} e^{A(x)}$ est continue sur I ; elle admet donc une primitive sur I . En d'autres termes, il existe une fonction K de classe $\mathcal{C}^1$ sur I satisfaisant l'identité $(\spadesuit)$.

On en déduit que la fonction $f_P : x \in I \mapsto e^{-A(x)} \left(\underbrace{\int \frac{c(x)}{a(x)} e^{A(x)} dx}_{=K(x)} \right)$ est solution de (E) . 

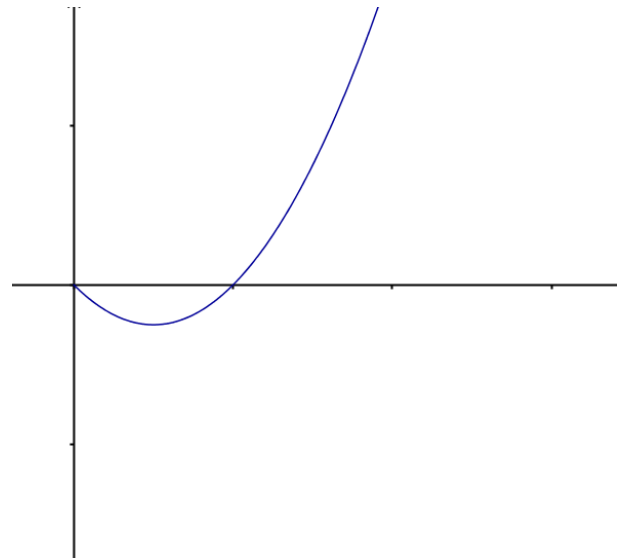
9.1.4 Conclusion - Problème de Cauchy

Il résulte du théorème 9.3 que toute EDL1 admet des solutions. On sait alors décrire complètement l'ensemble des solutions de (E) , grâce aux théorèmes 9.1 (décrivant l'ensemble des solutions) et 9.2 (donnant la solution générale de l'équation homogène associée). Un petit dessin valant mieux qu'un long discours, on illustre par quelques graphiques la méthode mise en œuvre précédemment pour résoudre (E) : $xy' + (x-1)y = x^3$ (page 202).

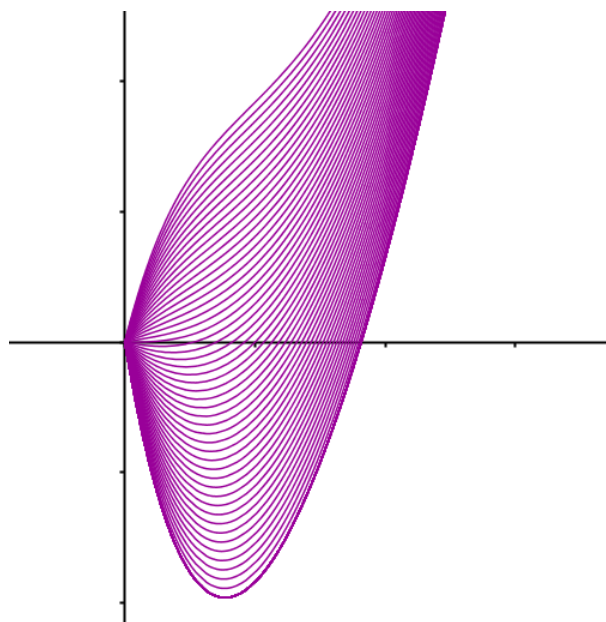
- Le premier graphique (“rouge”) montre quelques courbes représentatives de l’équation homogène, c’est-à-dire des fonctions $f_C : x \in \mathbb{R}_+^* \mapsto Cxe^{-x}$ (pour plusieurs valeurs de C).
- Le second correspond à la courbe représentative d’une solution particulière de l’équation avec second membre, la fonction $f_P : x \in \mathbb{R}_+^* \mapsto x(x-1)$.
- Sur le dernier graphique sont tracées les courbes de quelques solutions de l’équation avec second membre, c’est-à-dire quelques fonctions $g_C : x \in \mathbb{R}_+^* \mapsto x(x-1) + Cxe^{-x}$.



Solution générale de l’équation homogène



+ Une solution particulière de l’équation complète



= Solution générale de l’équation complète

Le dernier graphique de la page précédente suggère que les courbes représentatives des solutions de (E) “recouvrent le plan, et ne se coupent pas”. Cette intuition est précisée par l'énoncé ci-dessous.

THÉORÈME 9.4 - (Existence et unicité d'une solution pour un problème de Cauchy d'ordre 1).

Soient a , b et c trois fonctions de $\mathcal{C}^0(I, \mathbb{K})$.

On note (E) l'EDL1 : $a(x)y' + b(x)y = c(x)$.

On suppose que a ne s'annule pas sur I : $\forall x \in I, a(x) \neq 0$.

Pour tout couple $(x_0, y_0) \in I \times \mathbb{K}$, il existe une unique fonction $\varphi \in \mathcal{C}^1(I, \mathbb{K})$ telle que :

$$\begin{cases} \forall x \in I, & a(x)\varphi'(x) + b(x)\varphi(x) = c(x) \\ \varphi(x_0) = y_0 \end{cases}$$

Intuitivement : par tout point du plan, passe la courbe représentative d'une unique solution de (E).

PREUVE. Avec les notations de l'énoncé, le théorème 9.3 assure qu'il existe une solution f_P de (E).

Soit φ une solution arbitraire de (E). Il existe un scalaire K tel que : $\forall x \in I, \varphi(x) = f_P(x) + Ke^{-A(x)}$.

Considérons à présent un couple $(x_0, y_0) \in I \times \mathbb{K}$ arbitraire. On a :

$$\varphi(x_0) = y_0 \iff f_P(x_0) + Ke^{-A(x_0)} = y_0 \iff K = (y_0 - f_P(x_0))e^{A(x_0)}$$

On en déduit que la fonction φ définie en posant :

$$\forall x \in I, \varphi(x) = f_P(x) + (y_0 - f_P(x_0))e^{A(x_0) - A(x)}$$

est l'unique solution de (E) telle que $\varphi(x_0) = y_0$. ☕

9.2 EDL2 à coefficients constants

9.2.1 Généralités

DÉFINITION 9.2 - Soient a , b , c et d quatre fonctions de $\mathcal{C}^0(I, \mathbb{K})$. Une fonction $f \in \mathcal{C}^2(I, \mathbb{K})$ est dite **solution sur I de l'équation différentielle** $a(x)y'' + b(x)y' + c(x)y = d(x)$ si :

$$\forall x \in I, \quad a(x)f''(x) + b(x)f'(x) + c(x)f(x) = d(x)$$

Terminologie. (E) $a(x)y'' + b(x)y' + c(x)y = d(x)$ est appelée **équation** différentielle linéaire d'ordre 2 (ou EDL2) **avec second membre** ; et (H) $a(x)y'' + b(x)y' + c(x)y = 0$ est appelée **équation** (différentielle linéaire d'ordre 2) **homogène** (ou **sans second membre**) **associée à (E)**.

Par ailleurs, on note parfois simplement $ay'' + by' + cy = d$ l'équation différentielle (E).

Exemples. La fonction $\cos$ est solution sur $\mathbb{R}$ de l'EDL2 homogène $y'' + y = 0$. La fonction exponentielle $\mathbb{R}$ de l'EDL2 avec second membre $y'' + y' - y = e^x$.

Remarque. Tout au long de ce paragraphe, nous nous restreindrons au cas particulier des EDL2 à coefficients constants, c'est à dire de la forme

$$y'' + ay' + by = c(x) \quad \text{avec } a \text{ et } b \text{ dans } \mathbb{K}, \text{ et } c \in \mathcal{C}^0(I, \mathbb{K})$$

Cette remarque faite, on peut conjecturer assez facilement que les propriétés relatives aux EDL1 s'étendent aux EDL2. Commençons par le commencement :

THÉORÈME 9.5 - (Structure de l'ensemble des solutions d'une EDL2). Soient a et b deux scalaires, et $c \in \mathcal{C}^0(I, \mathbb{K})$.

On note (E) l'EDL2 : $y'' + ay' + by = c(x)$, et $(H) : y'' + ay' + by = 0$ l'équation homogène associée à (E) .

On suppose qu'il existe une solution φ de (E) sur I .

Alors, pour toute fonction $f \in \mathcal{C}^2(I, \mathbb{K})$:

$$[f \text{ est solution de } (E)] \iff [(f - \varphi) \text{ est solution de } (H)]$$

PREUVE.⁶ Soit f une fonction de classe $\mathcal{C}^2$ sur I à valeurs dans $\mathbb{K}$. Avec les notations et hypothèses de l'énoncé :

f est solution de (E)

$$\iff \forall x \in I, \quad f''(x) + af'(x) + bf(x) = c(x)$$

$$\iff \forall x \in I, \quad f''(x) + af'(x) + bf(x) = \varphi''(x) + a\varphi'(x) + b\varphi(x)$$

$$\iff \forall x \in I, \quad f''(x) - \varphi''(x) + a(f'(x) - \varphi'(x)) + b(f(x) - \varphi(x)) = 0$$

$$\iff \forall x \in I, \quad (f - \varphi)''(x) + a(f - \varphi)'(x) + b(f - \varphi)(x) = 0 \quad (\text{linéarité de la dérivation})$$

$$\iff (f - \varphi) \text{ est solution de } (H) \quad \text{☕}$$

En d'autres termes, et avec les notations précédemment introduites, toute solution f de (E) s'écrit $f_H + f_P$ où f_H (*resp.* f_P) désigne une solution quelconque de l'équation homogène associée à (E) (*resp.* une solution particulière de (E)).

La méthode (et le plan de ce paragraphe consacré aux EDL2) de résolution des EDL2 sera donc rigoureusement analogue à celui des EDL1.

9.2.2 Résolution d'une EDL 2 homogène

La petite difficulté dans ce paragraphe provient de ce que la solution générale d'une EDL2 pourra être différent selon que l'on cherche des solutions à valeurs réelles ou valeurs complexes d'une EDL2.

⁶. Les mêmes causes produisant les mêmes effets, la démonstration de ce théorème est analogue à celle du théorème analogue pour les EDL1.

THÉORÈME 9.6 - (Solution générale d'une EDL2 homogène - Cas COMPLEXE). Soient a et b deux complexes.

On note (H) l'EDL2 : $y'' + ay' + by = 0$.

On note encore (EC) : $X^2 + aX + b = 0$ **l'équation caractéristique** associée à (H) .

Enfin, on note $\Delta = a^2 - 4b$ le discriminant de l'équation caractéristique.

► Premier cas : $\Delta \neq 0$. L'équation (EC) possède deux racines complexes distinctes α et β . Les solutions de (H) sont alors les fonctions définies par

$$\forall x \in \mathbb{R}, \quad f(x) = C_1 e^{\alpha x} + C_2 e^{\beta x}$$

où C_1 et C_2 désignent deux complexes arbitraires.

► Second cas : $\Delta = 0$. L'équation (EC) possède une racine double α . Les solutions de (H) sont alors les fonctions définies par

$$\forall x \in \mathbb{R}, \quad f(x) = (C_1 + C_2 x) e^{\alpha x}$$

où C_1 et C_2 désignent deux complexes arbitraires.

PREUVE. Notons α et β les racines de l'équation caractéristique (même si éventuellement on peut avoir $\alpha = \beta$).

Posons : $\forall x \in \mathbb{R}, \varphi(x) = g(x)e^{\alpha x}$, où g est de classe $\mathcal{C}^2$ sur $\mathbb{R}$ (♣).

La fonction φ est de classe $\mathcal{C}^2$ sur $\mathbb{R}$ (en tant que produit de fonctions qui le sont); il est donc légitime de la dériver deux fois. On a :

$$\forall x \in \mathbb{R}, \varphi'(x) = e^{\alpha x} (g'(x) + \alpha g(x)) \text{ et } \forall x \in \mathbb{R}, \varphi''(x) = e^{\alpha x} (g''(x) + 2\alpha g'(x) + \alpha^2 g(x))$$

Par suite :

φ est solution de (H)

$$\iff \forall x \in \mathbb{R}, \varphi''(x) + a\varphi'(x) + b\varphi(x) = 0$$

$$\iff \forall x \in \mathbb{R}, e^{\alpha x} (g''(x) + 2\alpha g'(x) + \alpha^2 g(x) + ag'(x) + a\alpha g(x) + bg(x)) = 0$$

$$\iff \forall x \in \mathbb{R}, e^{\alpha x} \left(g''(x) + (2\alpha + a)g'(x) + \underbrace{(\alpha^2 + a\alpha + b)}_{=0} g(x) \right) = 0^7$$

On note alors que : $2\alpha + a = \alpha - \beta$.⁸ L'assertion précédente est donc encore équivalente à la suivante :

$$\forall x \in \mathbb{R}, e^{\alpha x} (g''(x) + (\alpha - \beta)g'(x)) = 0$$

$$\iff \forall x \in \mathbb{R}, g''(x) + (\alpha - \beta)g'(x) = 0^9$$

$$\iff g' \text{ est solution de l'équation différentielle : } Y' + (\alpha - \beta)Y = 0 \text{ (♠)}$$

7. On a $\alpha^2 + a\alpha + b = 0$ puisque α est racine de l'équation caractéristique.

8. Lorsque les racines sont distinctes, cela provient de l'égalité $\alpha + \beta = -a$ (relation coefficients racines pour un polynôme de degré 2). Lorsque α est racine double de l'équation caractéristique, on a : $2\alpha = -a$ donc $2\alpha + a = 0$; et d'autre part dans ce cas $\beta - \alpha = \alpha - \alpha = 0$. En résumé, dans les deux situations, on a : $2\alpha + a = \alpha - \beta$.

9. Puisque la fonction $x \mapsto e^{\alpha x}$ ne s'annule pas sur $\mathbb{R}$.

On distingue alors deux cas :

➤ Premier cas — Si $\alpha \neq \beta$: alors on déduit de ($\spadesuit$) que :

$$\exists \lambda \in \mathbb{C}, \forall x \in \mathbb{R}, g'(x) = \lambda e^{(\beta-\alpha)x}$$

$$\iff \exists (\lambda, \mu) \in \mathbb{C}^2, \forall x \in \mathbb{R}, g(x) = \frac{\lambda}{\beta - \alpha} e^{(\beta-\alpha)x} + \mu$$

En exploitant alors la relation ($\clubsuit$), on peut affirmer que :

$$\varphi \text{ est solution de } (H) \iff \exists (\lambda, \mu) \in \mathbb{C}^2, \forall x \in \mathbb{R}, \varphi(x) = \left(\frac{\lambda}{\beta - \alpha} e^{(\beta-\alpha)x} + \mu \right) e^{\alpha x}$$

$$\iff \exists (\lambda, \mu) \in \mathbb{C}^2, \forall x \in \mathbb{R}, \varphi(x) = \frac{\lambda}{\beta - \alpha} e^{\beta x} + \mu e^{\alpha x}$$

On en déduit donc que, dans le cas où $\Delta \neq 0$, toute solution est “de la forme $C_1 e^{\beta x} + C_2 e^{\alpha x}$ ” ; la réciproque est une vérification pas très rapide, mais aisée.

➤ Second cas — Si $\alpha = \beta$: alors on déduit de ($\spadesuit$) que :

$$\exists \lambda \in \mathbb{C}, \forall x \in \mathbb{R}, g'(x) = \lambda$$

$$\iff \exists (\lambda, \mu) \in \mathbb{C}^2, \forall x \in \mathbb{R}, g(x) = \lambda x + \mu$$

En exploitant alors la relation ($\clubsuit$), on peut affirmer que :

$$\varphi \text{ est solution de } (H) \iff \exists (\lambda, \mu) \in \mathbb{C}^2, \forall x \in \mathbb{R}, \varphi(x) = (\lambda x + \mu) e^{\alpha x}$$

On en déduit donc que, dans le cas où $\Delta = 0$, toute solution est “de la forme $(C_1 x + C_2) e^{\alpha x}$ ” ; la réciproque est, comme précédemment, aisée. Ce qui achève la preuve du théorème. ☕

Exemple 1. Résolution dans $\mathbb{C}$ de (E) : $y'' + 4y' + 13y = 0$.

L'équation caractéristique (EC) a pour discriminant $\Delta = -36$, et donc pour racines $-2 \pm 3i$. D'après le théorème 9.6, la solution générale de (E) est : $\forall x \in \mathbb{R}, f_H(x) = C_1 e^{(-2+3i)x} + C_2 e^{(-2-3i)x}$ (C_1, C_2 complexes).

Exemple 2. Résolution dans $\mathbb{C}$ de (E) : $y'' + 4y' + 4y = 0$.

L'équation caractéristique (EC) est $(X+2)^2 = 0$. Elle a donc une racine double : -2 . D'après le théorème 9.6, la solution générale de (E) est : $\forall x \in \mathbb{R}, f_H(x) = (C_1 + C_2 x) e^{-2x}$ (C_1, C_2 complexes).

On passe à présent au cas réel, c'est à dire que l'on détermine les fonctions $f \in \mathcal{C}^2(\mathbb{R}, \mathbb{R})$ (à valeurs réelles donc) solution de l'EDL2 $y'' + ay' + by = 0$.

THÉORÈME 9.7 - (Solution générale d'une EDL2 homogène - Cas REEL). Soient a et b deux réels.

On note (H) l'EDL2 : $y'' + ay' + by = 0$.

On note encore (EC) : $X^2 + aX + b = 0$ l'équation caractéristique associée à (H) .

Enfin, on note $\Delta = a^2 - 4b$ le discriminant de l'équation caractéristique.

► Premier cas : $\Delta > 0$. L'équation (EC) possède deux racines réelles distinctes α et β . Les solutions de (H) à valeurs réelles sont alors les fonctions définies par

$$\forall x \in \mathbb{R}, \quad f(x) = C_1 e^{\alpha x} + C_2 e^{\beta x}$$

où C_1 et C_2 désignent deux réels arbitraires.

► Second cas : $\Delta = 0$. L'équation (EC) possède une racine double α . Les solutions de (H) à valeurs réelles sont alors les fonctions définies par

$$\forall x \in \mathbb{R}, \quad f(x) = (C_1 + C_2 x) e^{\alpha x}$$

où C_1 et C_2 désignent deux réels arbitraires.

► Troisième cas : $\Delta < 0$. L'équation (EC) possède deux racines complexes (non réelles) conjuguées $u \pm i v$. Les solutions de (H) à valeurs réelles sont alors les fonctions définies par

$$\forall x \in \mathbb{R}, \quad f(x) = (C_1 \cos(vx) + C_2 \sin(vx)) e^{ux}$$

où C_1 et C_2 désignent deux réels arbitraires.

PREUVE. Mutatis mutandis, la preuve des deux premiers cas se déduit de la preuve du théorème 9.6.

Reste donc à traiter le cas où $\Delta < 0$. On note alors $u \pm i v$ les deux racines de l'équation caractéristique associée à (H) .

Soit $\varphi \in \mathcal{C}^2(\mathbb{R}, \mathbb{R})$ une solution de (H) . Alors φ est en particulier une solution à valeurs complexes de (H) (puisque $\mathbb{R} \subset \mathbb{C}$), et il existe donc deux nombres complexes C_1 et C_2 tels que :

$$\forall x \in \mathbb{R}, \quad \varphi(x) = C_1 e^{(u+iv)x} + C_2 e^{(u-iv)x}$$

Il reste à déterminer à quelle(s) condition(s) sur C_1 et C_2 la fonction φ est à valeurs réelles. D'après ce qui précède, on a :

$$\forall x \in \mathbb{R}, \quad \varphi(x) = C_1 e^{(u+iv)x} + C_2 e^{(u-iv)x}$$

$$\iff \forall x \in \mathbb{R}, \quad \varphi(x) = C_1 e^{ux} (\cos(vx) + i \sin(vx)) + C_2 e^{ux} (\cos(vx) - i \sin(vx))$$

$$\iff \forall x \in \mathbb{R}, \quad \varphi(x) = e^{ux} [(C_1 + C_2) \cos(vx) + i (C_1 - C_2) \sin(vx)]$$

$$\text{En particulier, on a : } \begin{cases} \varphi(0) \in \mathbb{R} \\ \varphi\left(\frac{\pi}{2v}\right) \in \mathbb{R} \end{cases} \iff \begin{cases} (C_1 + C_2) \in \mathbb{R} \\ i(C_1 - C_2) \in \mathbb{R} \end{cases}$$

En posant $K_1 = (C_1 + C_2)$ et $K_2 = i(C_1 - C_2)$, on a établi l'implication :

$$[\varphi \in \mathcal{C}^2(\mathbb{R}, \mathbb{R}) \text{ est solution de } (H)] \implies [\exists (K_1, K_2) \in \mathbb{R}^2, \forall x \in \mathbb{R}, \varphi(x) = (K_1 \cos(ux) + K_2 \sin(ux)) e^{ux}]$$

L'implication réciproque est une vérification immédiate qui achève la preuve de ce théorème. ☕

Exemple 1. Résolution dans $\mathbb{R}$ de (E) : $y'' + 4y' + 13y = 0$.

L'équation caractéristique (EC) a pour discriminant $\Delta = -36$, et donc pour racines $-2 \pm 3i$. D'après le théorème 9.7, la solution générale de (E) (à valeurs réelles) est :

$$\forall x \in \mathbb{R}, f_H(x) = (C_1 \cos(3x) + C_2 \sin(3x)) e^{-2x} \quad (C_1, C_2 \text{ réels}).$$

Exemple 2. Résolution dans $\mathbb{R}$ de (E) : $y'' + 4y' + 4y = 0$.

L'équation caractéristique (EC) est $(X + 2)^2 = 0$. Elle a donc une racine double : -2 .

D'après le théorème 9.7, la solution générale de (E) est :

$$\forall x \in \mathbb{R}, f_H(x) = (C_1 + C_2 x) e^{-2x} \quad (C_1, C_2 \text{ réels}).$$

9.2.3 Recherche d'une solution particulière

PROPRIÉTÉ 9.1 - Soient a et b deux réels, et α un scalaire.

On note (E) l'EDL2 : $y'' + ay' + by = e^{\alpha x}$.

On note encore (EC) : $X^2 + aX + b = 0$ l'équation caractéristique associée à (E) .

► Premier cas : α n'est pas racine de (EC) . L'équation (E) possède une solution particulière f_P avec : $\forall x \in \mathbb{R}, f_P(x) = Ke^{\alpha x}$, pour un certain $K \in \mathbb{C}$.

► Second cas : α est racine simple de (EC) . L'équation (E) possède une solution particulière f_P avec : $\forall x \in \mathbb{R}, f_P(x) = Kxe^{\alpha x}$, pour un certain $K \in \mathbb{C}$.

► Troisième cas : α est racine double de (EC) . L'équation (E) possède une solution particulière f_P avec : $\forall x \in \mathbb{R}, f_P(x) = Kx^2e^{\alpha x}$, pour un certain $K \in \mathbb{C}$.

PREUVE. ► Premier cas : α n'est pas racine de (EC) . Alors : $\alpha^2 + a\alpha + b \neq 0$.

Soit K un complexe. Posons : $\forall x \in \mathbb{R}, f_P(x) = Ke^{\alpha x}$. La fonction f_P est de classe $\mathcal{C}^2$ sur $\mathbb{R}$ et pour tout réel x on a : $f_P'(x) = \alpha Ke^{\alpha x}$ et $f_P''(x) = \alpha^2 Ke^{\alpha x}$.

Ainsi, pour tout réel x on a : $f_P''(x) + af_P'(x) + bf_P(x) = Ke^{\alpha x} (\alpha^2 + a\alpha + b)$.

Puisque $e^{\alpha x}$ est non nul pour tout réel x , on en déduit que :

$$[f_P \text{ est solution de } (E)] \iff [K(\alpha^2 + a\alpha + b) = 1] \iff \left[K = \frac{1}{\alpha^2 + a\alpha + b} \right]$$

La dernière égalité étant rendue légitime par le fait que α n'est pas racine de (EC) .

On peut alors conclure que la fonction f_P définie en posant : $\forall x \in \mathbb{R}, f_P(x) = \frac{1}{\alpha^2 + a\alpha + b} e^{\alpha x}$, est solution de (E) .

► Second cas : α est racine simple de (EC) . Alors : $\alpha^2 + a\alpha + b = 0$, et $2\alpha + a \neq 0$.¹⁰

Soit K un complexe. Posons : $\forall x \in \mathbb{R}, f_P(x) = Kxe^{\alpha x}$. La fonction f_P est de classe $\mathcal{C}^2$ sur $\mathbb{R}$ et pour tout réel x on a : $f_P'(x) = Ke^{\alpha x}(\alpha x + 1)$ et $f_P''(x) = Ke^{\alpha x}(\alpha^2 x + 2\alpha)$.

Ainsi, pour tout réel x on a : $f_P''(x) + af_P'(x) + bf_P(x) = Ke^{\alpha x}(\alpha^2 x + 2\alpha + a\alpha x + a + bx)$.

Puisque $e^{\alpha x}$ est non nul pour tout réel x , on en déduit que :

$$[f_P \text{ est solution de } (E)] \iff \left[K \left(\underbrace{(\alpha^2 + a\alpha + b)}_{=0} x + 2\alpha + a \right) = 1 \right] \iff \left[K = \frac{1}{2\alpha + a} \right]$$

La dernière égalité étant rendue légitime par le fait que α est racine simple de (EC) .

On peut alors conclure que la fonction f_P définie en posant : $\forall x \in \mathbb{R}, f_P(x) = \frac{1}{2\alpha + a} xe^{\alpha x}$, est solution de (E) .

► Troisième cas : α est racine double de (EC) . Alors : $\alpha^2 + a\alpha + b = 0$, et $2\alpha + a = 0$.

Soit K un complexe. Posons : $\forall x \in \mathbb{R}, f_P(x) = Kx^2 e^{\alpha x}$. La fonction f_P est de classe $\mathcal{C}^2$ sur $\mathbb{R}$ et pour tout réel x on a : $f_P'(x) = Ke^{\alpha x}(\alpha x^2 + 2x)$ et $f_P''(x) = Ke^{\alpha x}(\alpha^2 x^2 + 4\alpha x + 2)$.

Ainsi, pour tout réel x on a : $f_P''(x) + af_P'(x) + bf_P(x) = Ke^{\alpha x}(\alpha^2 x^2 + 4\alpha x + 2 + a\alpha x^2 + 2ax + bx^2)$.

Puisque $e^{\alpha x}$ est non nul pour tout réel x , on en déduit que :

$$[f_P \text{ est solution de } (E)] \iff \left[K \left(\underbrace{(\alpha^2 + a\alpha + b)}_{=0} x^2 + \underbrace{(2a + 4\alpha)}_{=0} x + 2 \right) = 1 \right] \iff \left[K = \frac{1}{2} \right]$$

On peut alors conclure que la fonction f_P définie en posant : $\forall x \in \mathbb{R}, f_P(x) = \frac{1}{2} x^2 e^{\alpha x}$, est solution de (E) ; ce qui achève la preuve de la propriété. ☕

Exemple 1. Considérons l'équation $(E) : y'' - 5y' + 6y = e^{ix}$.

L'équation caractéristique $(X^2 - 5X + 6 = 0)$ a pour racines 2 et 3.

Puisque i n'est pas racine de l'équation caractéristique, (E) admet une solution particulière de la forme Ke^{ix} (avec K complexe).

Posons donc : $\forall x \in \mathbb{R}, \varphi(x) = Ke^{ix}$. La fonction φ est de classe $\mathcal{C}^2$ sur $\mathbb{R}$ (et même de classe $\mathcal{C}^\infty$).

Pour tout réel x , on a : $\varphi'(x) = iKe^{ix}$ et $\varphi''(x) = -Ke^{ix}$. Il s'ensuit que φ est solution de (E) si et seulement si :

$$\begin{aligned} \forall x \in \mathbb{R}, -Ke^{ix} - 5iKe^{ix} + 6Ke^{ix} = e^{ix} &\iff \forall x \in \mathbb{R}, (-1 - 5i + 6)Ke^{ix} = e^{ix} \iff K(5 - 5i) = 1 \\ &\iff K = \frac{1}{5 - 5i} \iff K = \frac{1}{10} + \frac{1}{10}i \end{aligned}$$

On en déduit que la fonction $\varphi : x \in \mathbb{R} \mapsto \left(\frac{1}{10} + \frac{1}{10}i \right) e^{ix}$ est solution de (E) .

10. En effet, si α était racine double de (EC) , on aurait $\alpha = -a/2$, et donc $2\alpha + a = 0$.

Exemple 2. Considérons l'équation $(E) : y'' - 3y' + 2y = e^x$.

Puisque 1 est racine simple de l'équation caractéristique, (E) admet une solution particulière de la forme Kxe^x (avec K complexe).

Posons donc : $\forall x \in \mathbb{R}, \varphi(x) = Kxe^x$. La fonction φ est de classe $\mathcal{C}^2$ sur $\mathbb{R}$.

Pour tout réel x , on a : $\varphi'(x) = Ke^x(x+1)$ et $\varphi''(x) = Ke^x(x+2)$. Il s'ensuit que φ est solution de (E) si et seulement si :

$$\forall x \in \mathbb{R}, Ke^x(x+2-3x-3+2x) = e^x \iff K = -1$$

On en déduit que la fonction : $\varphi : x \in \mathbb{R} \mapsto -xe^x$ est solution de (E) .

Exemple 3. Considérons l'équation $(E) : y'' - 2y' + y = e^x$.

Puisque 1 est racine double de l'équation caractéristique, (E) admet une solution particulière de la forme Kx^2e^x (avec K complexe).

Posons donc : $\forall x \in \mathbb{R}, \varphi(x) = Kx^2e^x$. La fonction φ est de classe $\mathcal{C}^2$ sur $\mathbb{R}$.

Pour tout réel x , on a : $\varphi'(x) = Ke^x(x^2+2x)$ et $\varphi''(x) = Ke^x(x^2+4x+2)$. Il s'ensuit que φ est solution de (E) si et seulement si :

$$\forall x \in \mathbb{R}, Ke^x(x^2+4x+2-2x^2-4x+x^2) = e^x \iff K = \frac{1}{2}$$

On en déduit que la fonction : $\varphi : x \in \mathbb{R} \mapsto \frac{1}{2}x^2e^x$ est solution de (E) .

La propriété 9.1 peut sembler restrictive, dans le sens où elle ne s'applique qu'à un cas très particulier de second membre (" $e^{\alpha x}$ "). Une extension de ce résultat est fourni par l'énoncé suivant, qui repose sur la linéarité de l'application

$$\begin{aligned} \Delta : \mathcal{C}^2(\mathbb{R}, \mathbb{K}) &\longrightarrow \mathcal{C}^0(\mathbb{R}, \mathbb{K}) \\ f &\longmapsto f'' + af' + bf \end{aligned}$$

PROPRIÉTÉ 9.2 - (Principe de superposition des solutions, v1). Soient a et b deux scalaires ; c_1 et c_2 deux fonctions continues sur $\mathbb{R}$ à valeurs dans $\mathbb{K}$.

On suppose que f_1 (resp. f_2) est solution de $y'' + ay' + by = c_1$ (resp. $y'' + ay' + by = c_2$).

Alors pour tout couple de scalaires $(\lambda, \mu) \in \mathbb{K}^2$, la fonction $(\lambda f_1 + \mu f_2)$ est solution de l'équation différentielle $y'' + ay' + by = \lambda c_1 + \mu c_2$.

PREUVE. On renvoie à la preuve de la version générale, propriété 9.3 page 213. ☕

Exemple. Considérons l'équation $(E) : y'' - 3y' + 2y = 2\operatorname{ch}(x)$.

Pour tout réel x , on a : $2\operatorname{ch}(x) = e^x + e^{-x}$. D'après le principe de superposition, une solution particulière de (E) peut être obtenue en faisant la somme d'une solution particulière de (E_1) et d'une solution particulière de (E_2) , en ayant posé :

$$(E_1) : y'' - 3y' + 2y = e^{-x} \quad \text{et} \quad (E_2) : y'' - 3y' + 2y = e^x$$

➤ Recherche d'une solution particulière de (E_1) . On introduit l'équation $(E_1) : y'' - 3y' + 2y = e^{-x}$.

Puisque -1 n'est pas racine de l'équation caractéristique, (E_1) admet une solution particulière de la forme Ke^{-x} (avec K scalaire). Posons donc : $\forall x \in \mathbb{R}, f_1(x) = Ke^{-x}$. La fonction f_1 est de classe $\mathcal{C}^2$ sur $\mathbb{R}$ (et même de classe $\mathcal{C}^\infty$), et on peut donc calculer ses dérivées successives.

Pour tout réel x , on a clairement : $f_1'(x) = -Ke^{-x}$ et $f_1''(x) = Ke^{-x}$. Il s'ensuit que f_1 est solution de (E_1) si et seulement si :

$$\forall x \in \mathbb{R}, Ke^{-x}(1 + 3 + 2) = e^{-x} \iff K = \frac{1}{6}$$

On en déduit que la fonction $f_1 : x \in \mathbb{R} \mapsto \frac{1}{6}e^{-x}$ est solution de (E_1) ($\spadesuit$).

➤ Recherche d'une solution particulière de (E_2) . On a déjà établi page 212 qu'une solution de (E_2) est la fonction $f_2 : x \in \mathbb{R} \mapsto -xe^x$ ($\clubsuit$).

➤ D'après ($\clubsuit$), ($\spadesuit$) et le principe de superposition, on peut affirmer que la fonction :

$$f : x \in \mathbb{R} \mapsto \frac{1}{6}e^{-x} - xe^x \text{ est solution de } (E)$$

On énonce à présent le principe de superposition dans sa version générale.

PROPRIÉTÉ 9.3 - (Principe de superposition des solutions, v2). Soient a et b deux scalaires, et $n \in \mathbb{N}^*$.

Soient $c_1, \dots, c_n$ n fonctions continues sur $\mathbb{R}$ à valeurs dans $\mathbb{K}$.

Soient $f_1, \dots, f_n$ n fonctions de $\mathcal{C}^2(\mathbb{R}, \mathbb{K})$ telles que :

$\forall i \in \llbracket 1, n \rrbracket, f_i$ est solution de $(E_i) : y'' + ay' + by = c_i$.

Alors pour tout $(\lambda_1, \dots, \lambda_n) \in \mathbb{K}^n$,

$$\text{la fonction } \sum_{i=1}^n \lambda_i f_i \text{ est solution de l'équation différentielle } y'' + ay' + by = \sum_{i=1}^n \lambda_i c_i.$$

PREUVE. Sous les hypothèses de l'énoncé, on a :

$$\begin{aligned} \left(\sum_{i=1}^n \lambda_i f_i \right)'' + a \left(\sum_{i=1}^n \lambda_i f_i \right)' + b \sum_{i=1}^n \lambda_i f_i &= \sum_{i=1}^n \lambda_i f_i'' + a \sum_{i=1}^n \lambda_i f_i' + b \sum_{i=1}^n \lambda_i f_i \\ &= \sum_{i=1}^n \lambda_i (f_i'' + a f_i' + b f_i) = \sum_{i=1}^n \lambda_i c_i \end{aligned}$$

la première égalité provenant de la linéarité de la dérivation, la seconde de la linéarité de la somme, et la dernière de l'hypothèse suivant laquelle f_i est solution de $y'' + ay' + by = c_i$ pour tout $i \in \llbracket 1, n \rrbracket$. 

Le principe de superposition permet donc d'élargir la classe de seconds membres pour lesquels on peut déterminer une solution particulière. On peut en particulier l'appliquer lorsque le second membre est un cosinus hyperbolique ou un sinus hyperbolique.

On pourrait aussi l'appliquer lorsque le second membre est un cosinus ou un sinus ; mais il y a dans ce cas mieux à faire, comme le précise l'énoncé suivant.

LEMME 9.1 - (“Pont $\mathbb{R} \longleftrightarrow \mathbb{C}$ ” pour les EDL). Soient a et b deux réels.

Soient $c \in \mathcal{C}^0(\mathbb{R}, \mathbb{C})$ et $\varphi \in \mathcal{C}^2(\mathbb{R}, \mathbb{C})$. Alors :

$$[\varphi \text{ est solution de } y'' + ay' + by = c] \iff \begin{cases} \operatorname{Re}(\varphi) \text{ est solution de } y'' + ay' + by = \operatorname{Re}(c) \\ \operatorname{Im}(\varphi) \text{ est solution de } y'' + ay' + by = \operatorname{Im}(c) \end{cases}$$

PREUVE. Sous les hypothèses de l'énoncé, on a :

$$[\varphi \text{ est solution de } y'' + ay' + by = c]$$

$$\iff \forall x \in \mathbb{R}, \varphi''(x) + a\varphi'(x) + b\varphi(x) = c(x)$$

$$\iff \forall x \in \mathbb{R}, (\operatorname{Re}(\varphi''(x)) + a\operatorname{Re}(\varphi'(x)) + b\operatorname{Re}(\varphi(x))) + i(\operatorname{Im}(\varphi''(x)) + a\operatorname{Im}(\varphi'(x)) + b\operatorname{Im}(\varphi(x))) = \operatorname{Re}(c(x)) + i\operatorname{Im}(c(x))$$

$$\iff \forall x \in \mathbb{R}, ((\operatorname{Re}\varphi)''(x) + a(\operatorname{Re}\varphi)'(x) + b(\operatorname{Re}\varphi)(x)) + i((\operatorname{Im}\varphi)''(x) + a(\operatorname{Im}\varphi)'(x) + b(\operatorname{Im}\varphi)(x)) = (\operatorname{Re}c)(x) + i(\operatorname{Im}c)(x)$$

$$\iff \begin{cases} \forall x \in \mathbb{R}, (\operatorname{Re}\varphi)''(x) + a(\operatorname{Re}\varphi)'(x) + b(\operatorname{Re}\varphi)(x) = (\operatorname{Re}c)(x) \\ \forall x \in \mathbb{R}, (\operatorname{Im}\varphi)''(x) + a(\operatorname{Im}\varphi)'(x) + b(\operatorname{Im}\varphi)(x) = (\operatorname{Im}c)(x) \end{cases}$$

$$\iff \begin{cases} \operatorname{Re}(\varphi) \text{ est solution de } y'' + ay' + by = \operatorname{Re}(c) \\ \operatorname{Im}(\varphi) \text{ est solution de } y'' + ay' + by = \operatorname{Im}(c) \end{cases} \quad \text{☕}$$

Exemple. Considérons l'équation $(E) : y'' - 5y' + 6y = \cos(x)$.

On introduit l'équation $(E') : y'' - 5y' + 6y = e^{ix}$.

On a déjà établi (voir exemple 1, page 211) que la fonction $\varphi : x \in \mathbb{R} \mapsto \left(\frac{1}{10} + \frac{1}{10}i\right)e^{ix}$ est solution de (E') .

En vertu du lemme, il ne reste plus qu'à en déterminer la partie réelle pour avoir une solution de (E) .

Explicitement, la fonction $\varphi : x \in \mathbb{R} \mapsto \frac{1}{10}(\cos(x) - \sin(x))$ est une solution particulière de (E) .

Remarques. Pour conclure ce paragraphe, indiquons que l'énoncé de la propriété 9.1 peut être généralisé aux seconds membres de la forme $P(x)e^{\alpha x}$, où P désigne un polynôme. Explicitement, l'équation $y'' + ay' + by = P(x)e^{\alpha x}$ admet une solution de la forme $Q(x)e^{\alpha x}$ où Q est un polynôme, avec :

- $\deg(Q) = \deg(P)$ lorsque α n'est pas racine de (EC) ;
- $\deg(Q) = \deg(P) + 1$ lorsque α est racine simple de (EC) ;
- $\deg(Q) = \deg(P) + 2$ lorsque α est racine double de (EC) .

Autre astuce : lorsque le second membre est un “cos” ou un “sin”, on peut chercher une solution particulière sous la forme $\lambda \cos + \mu \sin$ avec λ et μ deux scalaires.

9.2.4 Conclusion

On donne ci-dessous un exemple de résolution complète d'un problème de Cauchy, pour faire le point sur la méthode générale et sur les résultats précédents.

Soient A , ω et y_0 trois réels strictement positifs, et α un réel strictement supérieur à 1.

Résoudre dans $\mathbb{R}$ (c'est-à-dire déterminer les solutions dans $\mathcal{C}^2(\mathbb{R}, \mathbb{R})$) le problème de Cauchy suivant :

$$\left\{ \begin{array}{l} \text{(E)} : \quad y'' + \omega^2 y = A \sin(\alpha x) \\ y(0) = y_0 \quad \text{et} \quad y'(0) = 0 \end{array} \right.$$

Commençons par résoudre (E) (EDL d'ordre 2 à coefficients constants).

► **Résolution de l'équation homogène associée.** Notons (H) $y'' + \omega^2 y = 0$ l'équation homogène associée à (E). L'équation caractéristique associée est $X^2 + \omega^2 = 0$, et elle a donc deux racines complexes conjuguées : $\pm i\omega$. On en déduit que la solution générale de (H) est

$$f : x \in \mathbb{R} \mapsto C_1 \cos(\omega x) + C_2 \sin(\omega x) \text{ (avec } C_1 \text{ et } C_2 \text{ réels)}$$

► **Recherche d'une solution particulière de l'équation complète.** On peut penser à déterminer une solution particulière de l'équation auxiliaire

$$(E') \quad y'' + \omega y = A e^{i\alpha x}$$

Mais il faut alors distinguer deux cas suivant que $\alpha = \omega$ (car alors $i\alpha$ est racine de l'équation caractéristique) ou $\alpha \neq \omega$.¹¹

► Cas $\alpha \neq \omega$. Ici, on cherche une solution particulière f_P de (E') en posant : $\forall x \in \mathbb{R}, f_P(x) = K e^{i\alpha x}$. On en déduit que : $\forall x \in \mathbb{R}, f_P''(x) = -\alpha^2 K e^{i\alpha x}$. Il s'ensuit que f_P est solution de (E') si et seulement si :

$$\forall x \in \mathbb{R}, K e^{i\alpha x} (\omega^2 - \alpha^2) = A e^{i\alpha x} \iff K = \frac{A}{\omega^2 - \alpha^2}$$

Ainsi la fonction $x \in \mathbb{R} \mapsto \frac{A}{\omega^2 - \alpha^2} e^{i\alpha x}$ est solution de (E'). On en déduit que sa partie imaginaire, la fonction : $f_P : x \in \mathbb{R} \mapsto \frac{A}{\omega^2 - \alpha^2} \sin(\alpha x)$ est une solution particulière de (E).

Conclusion. Dans le cas $\alpha \neq \omega$, la solution générale de (E) sur $\mathbb{R}$ est

$$y : x \in \mathbb{R} \mapsto \frac{A}{\omega^2 - \alpha^2} \sin(\alpha x) + C_1 \cos(\omega x) + C_2 \sin(\omega x) \text{ (} C_1 \text{ et } C_2 \text{ réels)}$$

11. Remarquons quand même, pour se remonter le moral, que l'on ne peut avoir $\alpha = -\omega$, puisque ces deux réels sont supposés strictement positifs dans l'énoncé.

➤ Cas $\alpha = \omega$. Dans cette situation, on cherche une solution particulière f_P de (E') en posant : $\forall x \in \mathbb{R}$, $f_P(x) = Kxe^{i\alpha x}$. On en déduit que : $\forall x \in \mathbb{R}$, $f'_P(x) = (1 + i\alpha x)Ke^{i\alpha x}$, puis : $\forall x \in \mathbb{R}$, $f''_P(x) = (2i\alpha - \alpha^2 x)Ke^{i\alpha x}$. Il s'ensuit que f_P est solution de (E') si et seulement si :

$$\forall x \in \mathbb{R}, Ke^{i\alpha x}((2i\alpha - \alpha^2 x) + \omega^2 x) = Ae^{i\alpha x} \iff \forall x \in \mathbb{R}, Ke^{i\alpha x}(2i\alpha) = Ae^{i\alpha x} \iff K = -\frac{A}{2\alpha}i$$

Ainsi la fonction $x \in \mathbb{R} \mapsto -\frac{A}{2\alpha}ixe^{i\alpha x}$ est solution de (E'). On en déduit que sa partie imaginaire, la fonction : $f_P : x \in \mathbb{R} \mapsto -\frac{A}{2\alpha}x \cos(\alpha x)$ est une solution particulière de (E).

Conclusion. Dans le cas $\alpha = \omega$, la solution générale de (E) sur $\mathbb{R}$ est

$$y : x \in \mathbb{R} \mapsto -\frac{A}{2\alpha}x \cos(\alpha x) + C_1 \cos(\alpha x) + C_2 \sin(\alpha x) \quad (C_1 \text{ et } C_2 \text{ réels})$$

Il ne reste “plus qu’à” résoudre le problème de Cauchy dans chaque cas :

➤ Dans le cas $\alpha \neq \omega$, la solution générale de (E) sur $\mathbb{R}$ est :

$$y : x \in \mathbb{R} \mapsto \frac{A}{\omega^2 - \alpha^2} \sin(\alpha x) + C_1 \cos(\omega x) + C_2 \sin(\omega x) \quad (C_1 \text{ et } C_2 \text{ réels})$$

Les conditions de Cauchy $y(0) = y_0$ et $y'(0) = 0$ imposent :

$$\begin{cases} C_1 = y_0 \\ \frac{A\alpha}{\omega^2 - \alpha^2} + \omega C_2 = 0 \end{cases} \quad \text{d'où :} \quad \begin{cases} C_1 = y_0 \\ C_2 = \frac{A\alpha}{\omega(\alpha^2 - \omega^2)} \end{cases}$$

Conclusion. Dans le cas $\alpha \neq \omega$, la solution du problème de Cauchy constitué de (E) et des conditions $y(0) = y_0$ et $y'(0) = 0$ est

$$y : x \in \mathbb{R} \mapsto \frac{A}{\omega^2 - \alpha^2} \sin(\alpha x) + y_0 \cos(\omega x) + \frac{A\alpha}{\omega(\alpha^2 - \omega^2)} \sin(\omega x)$$

➤ Dans le cas $\alpha = \omega$, la solution générale de (E) sur $\mathbb{R}$ est

$$y : x \in \mathbb{R} \mapsto -\frac{A}{2\alpha}x \cos(\alpha x) + C_1 \cos(\alpha x) + C_2 \sin(\alpha x) \quad (C_1 \text{ et } C_2 \text{ réels})$$

Les conditions de Cauchy $y(0) = y_0$ et $y'(0) = 0$ imposent : $\begin{cases} C_1 = y_0 \\ -\frac{A}{2\alpha} + \alpha C_2 = 0 \end{cases} \quad \text{d'où :} \quad \begin{cases} C_1 = y_0 \\ C_2 = \frac{A}{2\alpha^2} \end{cases}$

Conclusion. Dans le cas $\alpha = \omega$, la solution du problème de Cauchy constitué de (E) et des conditions $y(0) = y_0$ et $y'(0) = 0$ est

$$y : x \in \mathbb{R} \mapsto -\frac{A}{2\alpha}x \cos(\alpha x) + y_0 \cos(\alpha x) + \frac{A}{2\alpha^2} \sin(\alpha x)$$

9.3 Compléments sur les équations différentielles

9.3.1 EDL2 et oscillateur harmonique

En Physique, le comportement d'un oscillateur harmonique¹² est régi par une EDL2 de la forme

$$y'' + 2my' + \omega_0^2 y = c(t) \quad \text{avec } m \geq 0 \text{ et } \omega_0 > 0.$$

Plus précisément, on parle d'oscillateur harmonique *libre* lorsque $c = 0$; et d'oscillateur harmonique *forcé* sinon.

Enfin, lorsque $m = 0$, on parle d'oscillateur harmonique *sans frottements* lorsque $m = 0$, et *avec frottements* sinon.

On étudie ci-dessous les différents cas.

CAS N°1 — Etude de l'oscillateur harmonique libre, sans frottements (résolution de $y'' + \omega_0^2 y = 0$).
Solution g^{ale} peut s'écrire $f_{K,\varphi}(t) = K \cos(\omega_0 t - \varphi)$

On suppose que $\omega_0 \in \mathbb{R}_+^*$, et on note (H) l'équation différentielle linéaire du second ordre : $y'' + \omega_0^2 y = 0$. L'équation caractéristique associée possède deux racines complexes conjuguées $\pm i\omega_0$.¹³

On en déduit que la solution générale¹⁴ est : $\forall (t, \lambda, \mu) \in \mathbb{R}^3, f_{\lambda,\mu}(t) = \lambda \cos(\omega_0 t) + \mu \sin(\omega_0 t)$.

Fixons à présent les réels t , λ et μ , avec $(\lambda, \mu) \neq (0, 0)$. On écrit :

$$f_{\lambda,\mu}(t) = \sqrt{\lambda^2 + \mu^2} \left(\frac{\lambda}{\sqrt{\lambda^2 + \mu^2}} \cos(\omega_0 t) + \frac{\mu}{\sqrt{\lambda^2 + \mu^2}} \sin(\omega_0 t) \right)$$

Puis on pose : $X = \frac{\lambda}{\sqrt{\lambda^2 + \mu^2}}$ et $Y = \frac{\mu}{\sqrt{\lambda^2 + \mu^2}}$

Le couple de réels (X, Y) vérifie clairement la relation $X^2 + Y^2 = 1$. Par suite, le complexe $X + iY$ est un élément de $\mathbb{U}$, et il existe donc un unique réel $\varphi \in [0, 2\pi[$ tel que $X = \cos \varphi$ et $Y = \sin \varphi$. Par suite :

$$f_{\lambda,\mu}(t) = \sqrt{\lambda^2 + \mu^2} (\cos \varphi \cos(\omega_0 t) + \sin \varphi \sin(\omega_0 t)) \quad \text{d'où : } f_{\lambda,\mu}(t) = \sqrt{\lambda^2 + \mu^2} \cos(\omega_0 t - \varphi)$$

Le cas $(\lambda, \mu) = (0, 0)$ est trivial (prendre $K = 0$ et φ arbitraire).

Conclusion. La solution générale de (H) est : $\forall (t, K, \varphi) \in \mathbb{R}^3, f_{K,\varphi}(t) = K \cos(\omega_0 t - \varphi)$
(régime périodique)

12. Pour plus de renseignements sur l'oscillateur harmonique, consulter M Roveillo !

13. Qui sont effectivement distinctes puisque ω_0 est non nul par hypothèse.

14. Càd l'ensemble des fonctions solutions de (H) à valeurs *réelles*.

CAS N°2 — Etude de l'oscillateur harmonique libre, avec frottements (résolution de $y'' + 2my' + \omega_0^2 y = 0$ avec $m > 0$). Solution générale dans les trois cas.

On considère l'EDL2 : (H) $y'' + 2my' + \omega_0^2 y = 0$ avec m et ω_0 strictement positifs. L'équation caractéristique associée est : (EC) $X^2 + 2mX + \omega_0^2 = 0$, dont le discriminant est $\Delta = 4(m^2 - \omega_0^2)$.

On distingue alors trois cas :

► Si $\Delta < 0$ (càd si : $m < \omega_0$) : l'équation (EC) possède deux racines complexes conjuguées

$$-m \pm i\sqrt{\omega_0^2 - m^2}$$

Dans ce cas ("frottements faibles"), la solution générale de (H) est :

$$\forall (t, \lambda, \mu) \in \mathbb{R}^3, f_{\lambda, \mu}(t) = \left(\lambda \cos \left(\left(\sqrt{\omega_0^2 - m^2} \right) t \right) + \mu \sin \left(\left(\sqrt{\omega_0^2 - m^2} \right) t \right) \right) e^{-mt}$$

(régime pseudo-périodique)

► Si $\Delta = 0$ (càd si : $m = \omega_0$) : l'équation (EC) possède une racine double (réelle) : $-m$. Dans ce cas, la solution générale de (H) est :

$$\forall (t, \lambda, \mu) \in \mathbb{R}^3, f_{\lambda, \mu}(t) = (\lambda t + \mu) e^{-mt} \text{ (régime critique)}$$

► Si $\Delta > 0$ (càd si : $m > \omega_0$) : l'équation (EC) possède deux racines réelles $X_{1,2} = -m \pm \sqrt{m^2 - \omega_0^2}$. Dans ce cas ("frottements importants"), la solution générale de (H) est :

$$\forall (t, \lambda, \mu) \in \mathbb{R}^3, f_{\lambda, \mu}(t) = \lambda e^{X_1 t} + \mu e^{X_2 t} \text{ (régime apériodique)}$$

CAS N°3 — Exemple d'étude de l'oscillateur harmonique forcé, sans frottements :

résolution de $y'' + \omega_0^2 y = A \cos(\omega t)$.

Soient A , ω_0 et ω trois réels strictement positifs. Notons :

$$(E) : \quad y'' + \omega_0^2 y = A \cos(\omega t)$$

La résolution de l'équation homogène associée a déjà été faite précédemment : la solution générale de cette dernière est

$$\forall (t, K, \varphi) \in \mathbb{R}^3, f_{K, \varphi}(t) = K \cos(\omega_0 t - \varphi) \quad (\spadesuit)$$

Déterminons à présent une solution particulière de (E). En utilisant la méthode détaillée dans le paragraphe précédent, on s'intéresse en premier lieu à la résolution de

$$(E') : \quad y'' + \omega_0^2 y = A e^{i\omega t}$$

L'équation caractéristique associée possède deux racines complexes conjuguées $\pm i\omega_0$. On doit donc distinguer deux cas suivant que $\omega = \omega_0$ ou $\omega \neq \omega_0$.

➤ Premier cas : $\omega \neq \omega_0$. Dans ce cas, on cherche une solution particulière f_P de (E') en posant

$$\forall x \in \mathbb{R}, f_P(x) = Ke^{i\omega x}$$

On en déduit que : $\forall x \in \mathbb{R}, f_P''(x) = -\omega^2 Ke^{i\omega x}$. D'où f_P est solution de (E') si et seulement si :

$$\forall x \in \mathbb{R}, Ke^{i\omega x} (\omega_0^2 - \omega^2) = Ae^{i\omega x} \iff K = \frac{A}{\omega_0^2 - \omega^2}$$

Ainsi la fonction $x \in \mathbb{R} \mapsto \frac{A}{\omega_0^2 - \omega^2} e^{i\omega x}$ est solution de (E'). On en déduit que sa partie réelle, la fonction :

$f_P : x \in \mathbb{R} \mapsto \frac{A}{\omega_0^2 - \omega^2} \cos(\omega x)$ est une solution particulière de (E).

Conclusion. Dans le cas $\omega \neq \omega_0$, la solution générale de (E) sur $\mathbb{R}$ est

$$y : x \in \mathbb{R} \mapsto \frac{A}{\omega_0^2 - \omega^2} \cos(\omega x) + C_1 \cos(\omega_0 x) + C_2 \sin(\omega_0 x) \quad (C_1 \text{ et } C_2 \text{ réels})$$

➤ Second cas : $\omega = \omega_0$. Dans cette situation, on cherche une solution particulière f_P de (E') en posant :

$$\forall x \in \mathbb{R}, f_P(x) = Kxe^{i\omega_0 x}$$

On en déduit que : $\forall x \in \mathbb{R}, f_P'(x) = (1 + i\omega_0 x) Ke^{i\omega_0 x}$, puis : $\forall x \in \mathbb{R}, f_P''(x) = (2i\omega_0 - \omega_0^2 x) Ke^{i\omega_0 x}$.

Donc f_P est solution de (E') si et seulement si :

$$\forall x \in \mathbb{R}, Ke^{i\omega_0 x} ((2i\omega_0 - \omega_0^2 x) + \omega^2 x) = Ae^{i\omega_0 x} \iff \forall x \in \mathbb{R}, Ke^{i\omega_0 x} (2i\omega_0) = Ae^{i\omega_0 x} \iff K = -\frac{A}{2\omega_0} i$$

Ainsi la fonction $x \in \mathbb{R} \mapsto -\frac{A}{2\omega_0} ix e^{i\omega_0 x}$ est solution de (E'). On en déduit que sa partie réelle, la fonction

$f_P : x \in \mathbb{R} \mapsto \frac{A}{2\omega_0} x \sin(\omega_0 x)$ est une solution particulière de (E).

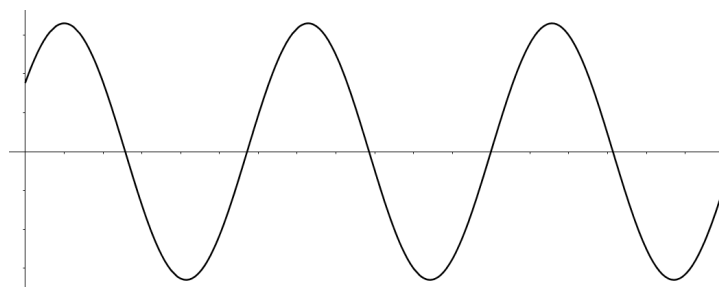
Conclusion. Dans le cas $\omega_0 = \omega$, la solution générale de (E) sur $\mathbb{R}$ est

$$y : x \in \mathbb{R} \mapsto \frac{A}{2\omega_0} x \sin(\omega_0 x) + C_1 \cos(\omega_0 x) + C_2 \sin(\omega_0 x) \quad (C_1 \text{ et } C_2 \text{ réels})$$

Remarque : dans cette dernière situation, on parle alors de phénomène de *résonance*.

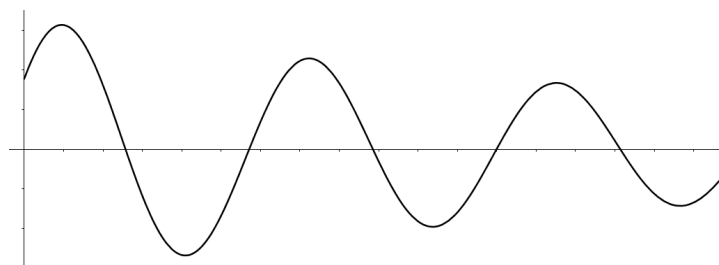
Les graphiques qui suivent font la synthèse des différents cas d'étude de l'oscillateur harmonique.

OSCILLATEUR HARMONIQUE LIBRE, SANS FROTTEMENTS



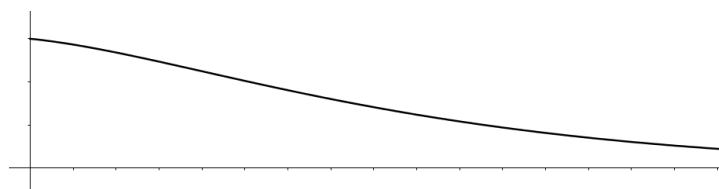
Régime périodique

OSCILLATEUR HARMONIQUE LIBRE, AVEC FROTTEMENTS (FAIBLES)



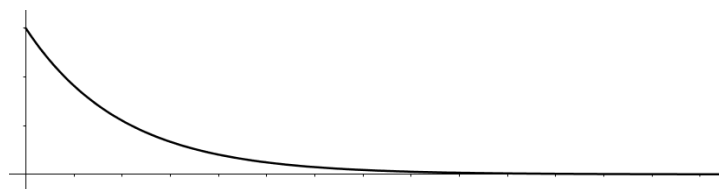
Régime pseudo-périodique

OSCILLATEUR HARMONIQUE LIBRE, AVEC FROTTEMENTS (“INTERMÉDIAIRES”)



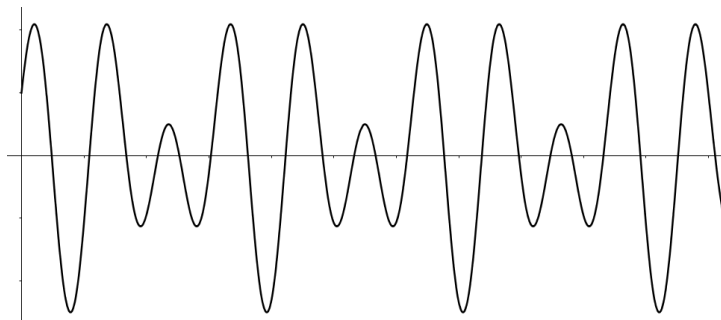
Régime critique

OSCILLATEUR HARMONIQUE LIBRE, AVEC FROTTEMENTS (IMPORTANTES)

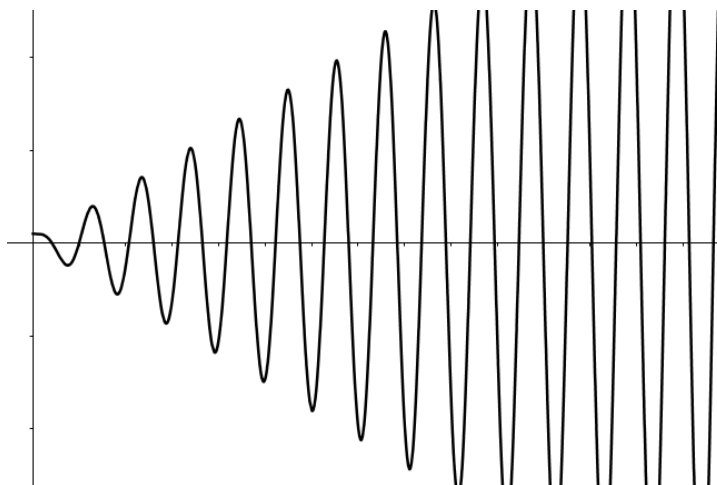


Régime aperiodique

OSCILLATEUR HARMONIQUE FORCÉ, SANS FROTTEMENTS, CAS GÉNÉRAL



OSCILLATEUR HARMONIQUE FORCÉ, SANS FROTTEMENTS, RÉSONANCE



9.3.2 Géométrie des équations différentielles

Soit $\vec{u}$ un vecteur non nul du plan réel. L'ensemble des multiples de $\vec{u}$:

$$\mathbb{R} \vec{u} = \{\lambda \vec{u} / \lambda \in \mathbb{R}\}$$

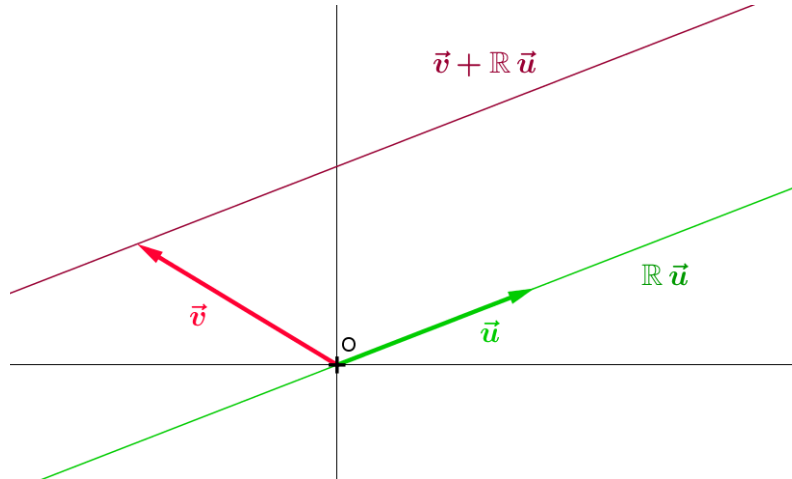
est une droite qui passe par l'origine du plan (on parle alors de droite *vectorielle*).

Considérons un second vecteur $\vec{v}$, non colinéaire à $\vec{u}$. L'image de $\mathbb{R} \vec{u}$ par la translation de vecteur $\vec{v}$ est une droite :

$$\vec{v} + \mathbb{R} \vec{u} = \{\vec{v} + \lambda \vec{u} / \lambda \in \mathbb{R}\}$$

qui bien évidemment ne passe plus par l'origine du plan (on parle alors de droite *affine*).

Ces considérations sont illustrées par le graphique de la page suivante.



Fin du premier acte.

Sous les hypothèses à présent usuelles, l'ensemble des solutions d'une EDL1 homogène ($ay' + by = 0$) est l'ensemble des multiples de e^{-A} , où A désigne une primitive de b/a . On peut donc noter cet ensemble :

$$\mathbb{R}e^{-A} = \{\lambda e^{-A} / \lambda \in \mathbb{R}\}$$

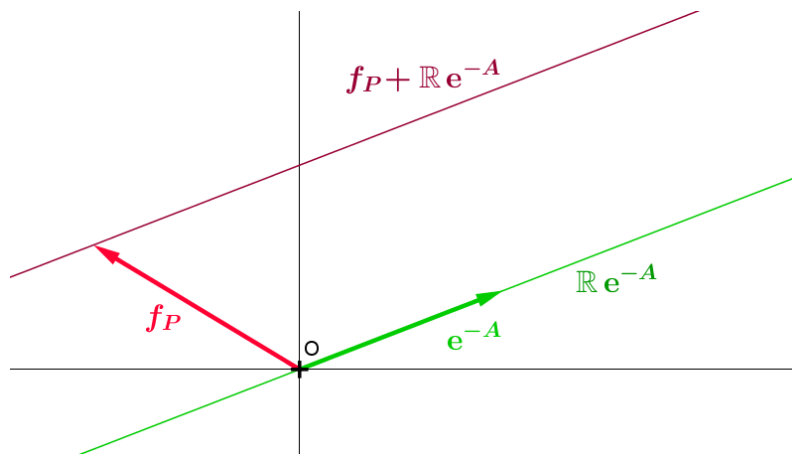
Par analogie avec la situation précédente, on dira que c'est une droite vectorielle (dans l'espace $\mathcal{C}^1(I, \mathbb{R})$).

Considérons à présent une solution particulière f_P de l'EDL1 $ay' + by = c$ (avec c non nulle). La fonction f_P est non colinéaire à e^{-A} (sinon elle serait solution de l'équation homogène). L'image de $\mathbb{R}e^{-A}$ par la translation de "vecteur" f_P

$$f_P + \mathbb{R}e^{-A} = \{f_P + \lambda e^{-A} / \lambda \in \mathbb{R}\}$$

est une droite affine (dans l'espace $\mathcal{C}^1(I, \mathbb{R})$).

Ces considérations sont illustrées par le graphique ci-dessous...



Remarque : de façon analogue, l'ensemble des solutions d'une EDL2 sans second membre est un plan vectoriel, et l'ensemble des solutions d'une EDL2 avec second membre est un plan affine.

9.3.3 Un exemple non linéaire : équation de Bernoulli

Une **équation de Bernoulli**¹⁵ est une équation différentielle de la forme

$$(E) \quad y' + a(x)y = b(x)y^\alpha$$

où α est un réel quelconque¹⁶ ; et a et b sont deux fonctions continues sur un certain intervalle de $\mathbb{R}$. Une telle équation est appelée une *équation différentielle ordinaire du premier ordre*. Elle n'est évidemment plus linéaire ; tout ce qui a été vu en cours sur les équas diffs (linéaires) ne s'applique donc plus.

Le principe de résolution est de se ramener à une équation différentielle linéaire *via* un *changement de fonction inconnue*, explicitement en posant : $z = y^{1-\alpha}$. On présente ci-dessous un exemple.

Exemple. Résolution de $(E) \quad y' + xy = xy^2$.

Supposant dans un premier temps que y ne s'annule pas, on réécrit $(E) : \frac{y'}{y^2} + x\frac{1}{y} = x$.

Le changement de fonction inconnue $Y = \frac{1}{y}$ conduit à résoudre : $-Y' + xY = x$.

Des calculs sans difficulté permettent d'en déduire que : $Y = 1 + Ke^{x^2/2}$ (avec $K \in \mathbb{R}$).

On en déduit qu'une solution de (E) est définie par une expression de la forme : $\varphi(x) = \frac{1}{1 + Ke^{x^2/2}}$.

9.3.4 Un problème : résolution d'une équation différentielle linéaire d'ordre 3

*L'objectif de ce problème est la résolution d'une équation différentielle linéaire d'ordre 3 ($y''' - y = 0$). Cette résolution se fait essentiellement dans la **partie C**. Pour y parvenir, on introduit dans la **partie B** des fonctions dont les dérivées possèdent des propriétés particulières, et qui seront utilisées dans la dernière partie.*

► **PARTIE A- Cadeaux.** Les deux questions de cette partie sont indépendantes.

1) On considère l'application :

$$\Delta : \mathcal{C}^\infty(\mathbb{R}, \mathbb{R}) \longrightarrow \mathcal{C}^\infty(\mathbb{R}, \mathbb{R})$$

$$f \longmapsto f'$$

qui à toute fonction de classe $\mathcal{C}^\infty$ sur $\mathbb{R}$ associe sa dérivée. L'application Δ est-elle injective ?

2) Résoudre dans $\mathbb{C}$ l'équation $z^3 - 1 = 0$. On précisera les formes exponentielles et algébriques des solutions.

► **PARTIE B - Calculs de dérivées**

On considère les trois fonctions définies sur $\mathbb{R}$ suivantes :

15. Ainsi nommées en référence à Jakob Bernoulli (1654-1705), mathématicien suisse, que l'on ne confondra pas avec son frère Jean Bernoulli (1667-1748), également mathématicien suisse, ni avec ses neveux Nicolas Bernoulli (1695-1726), Daniel Bernoulli (1700-1782) et Jean Bernoulli dit Jean II (1710-1790), tous mathématiciens suisses.

16. On supposera quand même que α est différent de 0 et 1, puisque dans ces deux cas l'équation (E) est une équation différentielle linéaire.

$$g_1 : x \in \mathbb{R} \mapsto e^x ; \quad g_2 : x \in \mathbb{R} \mapsto e^{-\frac{x}{2}} \sin \left(\frac{x\sqrt{3}}{2} \right) \quad \text{et} \quad g_3 : x \in \mathbb{R} \mapsto e^{-\frac{x}{2}} \cos \left(\frac{x\sqrt{3}}{2} \right)$$

- 3) Calculer g_1''' .¹⁷
- 4) Calculer g_2' . Vérifier qu'il existe deux réels a et b (que l'on précisera) tels que $g_2' = ag_2 + bg_3$.
- 5) Calculer g_3' . Vérifier qu'il existe deux réels c et d (que l'on précisera) tels que $g_3' = cg_2 + dg_3$.
- 6) Calculer g_2''' et g_3''' .

► **PARTIE C - Une équation différentielle**

Dans cette partie, on note (E_1) l'équation différentielle :

$$(E_1) \quad y''' = y$$

où y''' désigne la dérivée troisième de y . On note $\mathbf{S}$ l'ensemble des fonctions de classe $\mathcal{C}^\infty$ sur $\mathbb{R}$ et à valeurs réelles solutions de (E_1) .

Remarques : *Observez que comme l'équation différentielle (E_1) est linéaire d'ordre 3, les propriétés vues en cours ne vous permettent pas de la résoudre directement ; les questions ci-dessous donnent donc une méthode pour y parvenir.*

Par ailleurs dans cette partie, lorsque l'on demande de "résoudre l'équation différentielle...", on entend "déterminer les solutions à valeurs réelles de l'équation différentielle..."

- 7) Montrer que les fonctions g_1 , g_2 et g_3 définies au début de la partie B sont solutions de (E_1) .
- 8) On pose $\mathbf{F} = \{ag_1 + bg_2 + cg_3 / (a, b, c) \in \mathbb{R}^3\}$.¹⁸ Etablir que $\mathbf{F} \subset \mathbf{S}$.
- 9) Le but des questions suivantes est de montrer l'inclusion réciproque de celle établie dans la question précédente.
 - a) Soit f un élément de $\mathbf{S}$. On pose : $g = f'' + f' + f$. Montrer que g est solution de l'équation différentielle

$$(E_2) \quad y' - y = 0$$

- b) Donner la solution générale de l'équation différentielle (E_2) (*il n'est pas indispensable de détailler cette question*).
- c) Résoudre l'équation différentielle : $(E_3) \quad y'' + y' + y = 0$.
- d) Soit λ un nombre réel. Résoudre l'équation différentielle : $(E_4) \quad y'' + y' + y = \lambda e^x$.
- e) En déduire que $\mathbf{S} \subset \mathbf{F}$. Conclure.

17. On peut bien entendu noter $g_1^{(3)}$ en lieu et place et de g_1''' .

18. En d'autres termes, l'ensemble $\mathbf{F}$ est l'ensemble de ce que l'on appelle les *combinaisons linéaires* de g_1 , g_2 et g_3 , c'à-d des fonctions f s'écrivant sous la forme $f = ag_1 + bg_2 + cg_3$ avec a , b et c réels.

CORRIGÉ.**► PARTIE A - Cadeaux**

1) On a : $\Delta(\cos) = \Delta(\mathbb{1}_{\mathbb{R}} + \cos) = -\sin$. **Conclusion.** Δ n'est pas injective.

$$2) [z^3 - 1 = 0] \iff [z^3 = 1] \iff [z \in \mathbb{U}_3] \iff [z \in \{1, j, j^2\}] \iff \left[\begin{array}{l} z \in \{1, e^{2i\pi/3}, e^{-2i\pi/3}\} \\ \text{ou encore} \\ z \in \left\{1, \frac{-1 \pm i\sqrt{3}}{2}\right\} \end{array} \right]$$

► PARTIE B - Calculs de dérivées

3) Clairement : $g_1''' = g_1$.

4) La fonction g_2 est de classe $\mathcal{C}^\infty$ sur $\mathbb{R}$ (théorèmes généraux sur la dérivabilité) et : $g_2' = -\frac{1}{2}g_2 + \frac{\sqrt{3}}{2}g_3$.

5) La fonction g_3 est de classe $\mathcal{C}^\infty$ sur $\mathbb{R}$ (théorèmes généraux sur la dérivabilité) et : $g_3' = -\frac{1}{2}g_3 - \frac{\sqrt{3}}{2}g_2$.

6) On dérive la relation de la question 4 pour obtenir la dérivée seconde de g_2 :

$$\begin{aligned} g_2'' &= -\frac{1}{2}g_2' + \frac{\sqrt{3}}{2}g_3' \iff g_2'' = -\frac{1}{2}\left(-\frac{1}{2}g_2 + \frac{\sqrt{3}}{2}g_3\right) + \frac{\sqrt{3}}{2}\left(-\frac{1}{2}g_3 - \frac{\sqrt{3}}{2}g_2\right) \\ &\iff g_2'' = -\frac{1}{2}g_2 - \frac{\sqrt{3}}{2}g_3 \end{aligned}$$

Et on redérive :

$$g_2''' = -\frac{1}{2}\left(-\frac{1}{2}g_2 + \frac{\sqrt{3}}{2}g_3\right) - \frac{\sqrt{3}}{2}\left(-\frac{1}{2}g_3 - \frac{\sqrt{3}}{2}g_2\right) \iff \boxed{g_2''' = g_2}$$

Calculs analogues pour établir que $g_3''' = g_3$.

SOLUTION ALTERNATIVE — On peut calculer g'''_2 à l'aide de la formule de Leibniz. On obtient alors :

$$\begin{aligned} \forall x \in \mathbb{R}, g'''_2(x) &= \frac{1}{8} \left[-e^{-\frac{x}{2}} \sin\left(\frac{x\sqrt{3}}{2}\right) + 3\sqrt{3}e^{-\frac{x}{2}} \cos\left(\frac{x\sqrt{3}}{2}\right) + 9e^{-\frac{x}{2}} \sin\left(\frac{x\sqrt{3}}{2}\right) - 3\sqrt{3}e^{-\frac{x}{2}} \cos\left(\frac{x\sqrt{3}}{2}\right) \right] \\ &\iff \forall x \in \mathbb{R}, g'''_2(x) = e^{-\frac{x}{2}} \sin\left(\frac{x\sqrt{3}}{2}\right) \iff g'''_2 = g_2. \end{aligned}$$

► PARTIE C - Une équation différentielle

7) D'après les questions 3 et 6, on a : $\forall i \in \llbracket 1, 3 \rrbracket, g_i''' = g_i$.

8) Soit $f \in \mathbf{F}$. Il existe trois réels a, b et c tels que : $f = ag_1 + bg_2 + cg_3$. On a alors (essentiellement par linéarité de la dérivation) :

$$f''' = (ag_1 + bg_2 + cg_3)''' = ag_1''' + bg_2''' + cg_3''' = ag_1 + bg_2 + cg_3 = f. \text{ Il s'ensuit que } f \in \mathbf{S}.$$

En résumé : $[f \in \mathbf{F}] \implies [f \in \mathbf{S}]$. Par conséquent : $\boxed{\mathbf{F} \subset \mathbf{S}}$.

9) a) Avec les hypothèses de l'énoncé : $g' = f''' + f'' + f'$ et comme f est solution de (E_1) , on a $f''' = f$, d'où : $g' = f'' + f' + f$, soit : $g' = g$. Ainsi $\boxed{g \text{ est solution de l'équation } (E_2) : y' - y = 0}$.

b) $\boxed{\text{La solution générale sur } \mathbb{R} \text{ de } (E_2) \text{ est } f_\lambda : x \in \mathbb{R} \mapsto \lambda e^x \quad (\lambda \in \mathbb{R})}$.

c) L'équation caractéristique associée à l'équation (E_3) est : $r^2 + r + 1 = 0$. Cette équation possède deux solutions complexes conjuguées : $\frac{-1 \pm i\sqrt{3}}{2}$.¹⁹ Les solutions définies sur $\mathbb{R}$ et à valeurs réelles de l'équation différentielle $y'' + y' + y = 0$ sont donc les fonctions :

$$\boxed{f_{\lambda, \mu} : x \in \mathbb{R} \mapsto e^{-x/2} \left(\lambda \cos \left(\frac{x\sqrt{3}}{2} \right) + \mu \sin \left(\frac{x\sqrt{3}}{2} \right) \right) = \lambda g_3(x) + \mu g_2(x) \quad (\lambda, \mu \text{ réels})}.$$

d) On observe que $\frac{\lambda}{3} g_1$ est une solution particulière de (E_4) pour obtenir directement que les solutions définies sur $\mathbb{R}$ et à valeurs réelles de l'équation différentielle $y'' + y' + y = \lambda e^x$ sont les fonctions :

$$\boxed{f_{\alpha, \mu} : x \in \mathbb{R} \mapsto \frac{\lambda}{3} e^x + e^{-x/2} \left(\alpha \cos \left(\frac{x\sqrt{3}}{2} \right) + \mu \sin \left(\frac{x\sqrt{3}}{2} \right) \right) = \frac{\lambda}{3} g_1(x) + \alpha g_3(x) + \mu g_2(x) \quad (\alpha, \mu \text{ réels})}.$$

e) D'après les questions précédentes, si f est solution de (E_1) , alors il existe trois réels (λ, α, μ) tels que : $f = \frac{\lambda}{3} g_1 + \alpha g_3 + \mu g_2$. Donc : $f \in \mathbf{F}$. On vient donc de prouver l'inclusion : $\boxed{\mathbf{S} \subset \mathbf{F}}$.

On déduit de cette inclusion et de celle établie dans la question 8) que : $\mathbf{F} = \mathbf{S}$.

Conclusion : les solutions sur $\mathbb{R}$ de l'équation différentielle $y''' = y$ sont les fonctions $f_{\alpha_1, \alpha_2, \alpha_3}$ avec :

$$f_{\alpha_1, \alpha_2, \alpha_3} = \alpha_1 g_1 + \alpha_2 g_2 + \alpha_3 g_3 \quad (\alpha_1, \alpha_2 \text{ et } \alpha_3 \text{ réels})$$

19. Vous avez évidemment reconnu $j = e^{2i\pi/3}$ et $j^2 = e^{-2i\pi/3}$.

Chapitre 10

Nombres réels

10.1 Relations binaires, relations d'ordre, relations d'équivalence

DÉFINITION 10.1 - On appelle **relation binaire** sur un ensemble E est une partie R du produit cartésien $E \times E$.

Deux éléments x et y de E sont dits en relation (pour R) si $(x, y) \in R$; lorsque c'est le cas, on le note $x\mathcal{R}y$.

Exemples : la relation d'égalité sur un ensemble quelconque, la relation de congruence dans $\mathbb{Z}$, la relation de comparaison $\leq$ dans $\mathbb{R}$, la relation d'inclusion dans $\mathcal{P}(E)$, la relation de divisibilité dans $\mathbb{N}$ sont des relations binaires.

DÉFINITION 10.2 - Une relation binaire $\mathcal{R}$ sur un ensemble E est

- **réflexive** lorsque $\forall x \in E, x\mathcal{R}x$
- **symétrique** lorsque $\forall (x, y) \in E^2, (x\mathcal{R}y) \implies (y\mathcal{R}x)$
- **antisymétrique** lorsque $\forall (x, y) \in E^2, (x\mathcal{R}y) \wedge (y\mathcal{R}x) \implies (x = y)$
- **transitive** lorsque $\forall (x, y, z) \in E^3, (x\mathcal{R}y) \wedge (y\mathcal{R}z) \implies (x\mathcal{R}z)$

DÉFINITION 10.3 - Une relation binaire $\mathcal{R}$ sur un ensemble E est une

- une **relation d'ordre** si elle est réflexive, antisymétrique et transitive.
- une **relation d'équivalence** si elle est réflexive, symétrique et transitive.

Exemples : la relation d'égalité dans un ensemble quelconque, la relation de comparaison $\leq$ dans $\mathbb{R}$, la relation d'inclusion dans $\mathcal{P}(E)$, la relation de divisibilité dans $\mathbb{N}$ sont des relations d'ordre.

La relation de congruence dans $\mathbb{Z}$, la relation d'égalité dans un ensemble, la relation $[(a, b) \mathcal{R} (a', b') \text{ si } b' - a' = b - a]$ dans $\mathbb{N} \times \mathbb{N}$ sont des relations d'équivalence.

DÉFINITION 10.4 - Une relation d'ordre $\mathcal{R}$ est dite **totale** (on dit aussi que l'ordre est total) lorsque : $\forall (x, y) \in E^2, (x\mathcal{R}y) \vee (y\mathcal{R}x)$.

Exemples. La relation de comparaison $\leq$ dans $\mathbb{R}$ est une relation d'ordre totale ; la relation d'inclusion dans $\mathcal{P}(E)$ et celle de divisibilité dans $\mathbb{N}$ sont des relations d'ordre non totales.

DÉFINITION 10.5 - Un **ensemble ordonné** (*resp.* **totalelement ordonné**) est un couple $(E, \mathcal{R})$, où E est un ensemble et $\mathcal{R}$ une relation d'ordre (*resp.* ordre total) sur E .

Exemples : $(\mathbb{R}, \leq)$ est un ensemble totalement ordonné. $(\mathbb{Z}, |)$ et $(\mathcal{P}(E), \subset)$ sont des ensembles ordonnés (non totalement ordonnés).

DÉFINITION 10.6 - Soient $(E, \mathcal{R})$ un ensemble ordonné, et A une partie de E . On appelle :

- **majorant de A** un élément M de E tel que : $\forall a \in A, a\mathcal{R}M$
- **minorant de A** un élément m de E tel que : $\forall a \in A, m\mathcal{R}a$
- **plus grand élément de A** un élément M de A tel que : $\forall a \in A, a\mathcal{R}M$
- **plus petit élément de A** un élément m de A tel que : $\forall a \in A, m\mathcal{R}a$

La partie A est dite **majorée** (*resp.* **minorée**, *resp.* **bornée**) lorsqu'elle admet un majorant (*resp.* un minorant, *resp.* un majorant et un minorant).

Notations. Sous réserve d'existence, le plus grand (*resp.* petit) élément de A est noté $\max A$ (*resp.* $\min A$).

Exemples. Dans $(\mathbb{R}, \leq)$, $[0, 1[$ est bornée, admet un plus petit élément mais pas de plus grand élément.

Dans $(\mathbb{Z}, \leq)$ toute partie majorée (*resp.* minorée) non vide admet un plus grand (*resp.* petit) élément.

DÉFINITION 10.7 - Soient $(E, \mathcal{R})$ un ensemble ordonné, et A une partie de E . On appelle :

- **borne supérieure de A** le plus petit majorant de A (s'il existe)
- **borne inférieure de A** le plus grand minorant de A (s'il existe)

Notations. Sous réserve d'existence, la borne supérieure (*resp.* inférieure) de A est notée $\sup A$ (*resp.* $\inf A$).

Exemples. Dans $(\mathbb{R}, \leq)$: $\sup([0, 1[) = 1$, et $\inf([0, 1[) = \min([0, 1[) = 0$.

10.2 Bornes supérieures et inférieures dans l'ensemble des nombres réels

Jusqu'à la fin de ce chapitre, on se place dans $(\mathbb{R}, \leq)$.

Dans $\mathbb{Q}$, la partie $E = \{r \in \mathbb{Q}, r > 1 \wedge r^2 < 2\}$ ne possède ni plus grand élément, ni borne supérieure (alors qu'elle est non vide et majorée).¹

Puisque la construction de l'ensemble des nombres réels n'est pas à votre programme, signalons simplement que $\mathbb{R}$ est construit de telle manière que ce type de situation ne puisse se produire, et qu'il y ait donc "assez de nombres pour que les trous y soient comblés".

1. C'est vite écrit, mais c'est un peu technique à démontrer.

PROPRIÉTÉ 10.1 - (Caractérisation de la borne supérieure). Soit A une partie de $\mathbb{R}$, et soit S un réel.

$$S = \sup A \iff \begin{cases} \forall a \in A, a \leq S \\ \forall \varepsilon > 0, \exists a \in A, S - \varepsilon < a \leq S \end{cases}$$

Remarques. La première assertion (du membre de droite de l'équivalence) signifie que S est un majorant de A ; la seconde signifie que S est le plus petit, dans le sens où dès que l'on retire un "p'tit quelquechose" (strictement positif) à S , ce n'est plus un majorant de A .

Par ailleurs, on obtient mutatis mutandis un énoncé caractérisant la borne inférieure d'une partie A de $\mathbb{R}$.

PREUVE. Soit A une partie de $\mathbb{R}$, et soit S un réel.

► Supposons $S = \sup A$. Alors S est un majorant de A , d'où : $\forall a \in A, a \leq S$.

En outre, S est le plus petit des majorants, ce qui signifie que tout réel strictement inférieur à S n'est plus un majorant de A . Autrement dit :

$$[\forall \varepsilon > 0, S - \varepsilon \text{ n'est pas un majorant de } A] \iff \forall \varepsilon > 0, \exists a \in A, S - \varepsilon < a \leq S$$

$$\text{On a donc établi l'implication : } S = \sup A \implies \begin{cases} \forall a \in A, a \leq S \\ \forall \varepsilon > 0, \exists a \in A, S - \varepsilon < a \leq S \end{cases}$$

$$\text{► } \underline{\text{Réciproquement, supposons que : }} \begin{cases} \forall a \in A, a \leq S \\ \forall \varepsilon > 0, \exists a \in A, S - \varepsilon < a \leq S \end{cases}$$

L'assertion " $\forall a \in A, a \leq S$ " signifie déjà que S est un majorant de A . Montrons que c'est le meilleur !

Soit S' un réel tel que : $S' < S$. Posons $\varepsilon = S - S' > 0$. D'après la seconde assertion :

$$[\exists a \in A, S - (S - S') < a \leq S] \iff [\exists a \in A, S' < a \leq S]$$

En particulier, S' n'est pas un majorant de A . Ce qui prouve que S est le plus petit majorant de A , c-à-d : $S = \sup A$.

$$\text{On a ainsi établi l'implication réciproque : } \begin{cases} \forall a \in A, a \leq S \\ \forall \varepsilon > 0, \exists a \in A, S - \varepsilon < a \leq S \end{cases} \implies S = \sup A.$$

Ce qui achève la preuve de la propriété. ☕

Pour la suite des événements, le point-clef est la :

PROPRIÉTÉ 10.2 - (Propriété de la borne supérieure dans $\mathbb{R}$).

Toute partie non vide et majorée de $\mathbb{R}$ possède une borne supérieure.

PREUVE. Admise (snif!), car repose en grande partie sur la construction même de $\mathbb{R}$. ☕

Une conséquence en est que toute partie non vide et minorée de $\mathbb{R}$ possède une borne inférieure.

Remarques. Pour traduire l'énoncé ci-dessus, on dit parfois que $\mathbb{R}$ possède la propriété de la borne supérieure. Ce n'est pas le seul à jouir de cette propriété ; par exemple dans $\mathbb{N}$ ou $\mathbb{Z}$ (muni de la relation d'ordre " $\leq$ "), la propriété de la borne supérieure (et même "du max" !) tient aussi. Mais elle n'est pas valable dans $\mathbb{Q}$ (cf exemple donné au début de ce paragraphe).

PROPRIÉTÉ 10.3 - Soient A et B deux parties non vides et bornées de $\mathbb{R}$. Alors A et B possèdent chacun une borne inférieure et une borne supérieure, satisfaisant les propriétés ci-dessous :

$$1/ [A \subset B] \implies (\sup A \leq \sup B) \wedge (\inf B \leq \inf A)$$

$$2/ \sup(A \cup B) = \max(\sup A, \sup B)$$

$$3/ \sup(A + B) = \sup A + \sup B \text{ et } \inf(A + B) = \inf A + \inf B$$

(en ayant noté $A + B = \{a + b / a \in A, b \in B\}$).

PREUVE. Soient A et B deux parties non vides et bornées de $\mathbb{R}$. D'après la propriété de la borne supérieure (et celle de la borne inférieure), $\sup A$, $\sup B$, $\inf A$ et $\inf B$ existent.

► 1) Si $A \subset B$. Alors pour tout $a \in A$ on a $a \in B$, donc $a \leq \sup B$. Il s'ensuit que $\sup B$ est un majorant de A , donc $\sup A \leq \sup B$.²

► 3) Soit x un élément de $A + B$; il existe un couple (a, b) de $A \times B$ tel que $x = a + b$. Puisque $a \leq \sup A$ et $b \leq \sup B$, on a : $x \leq \sup A + \sup B$. Ce raisonnement tenant pour un élément arbitraire de $A + B$, on en déduit que $\sup A + \sup B$ est un majorant de $A + B$; d'où $\sup(A + B) \leq \sup A + \sup B$ (♠).

► Soient a et b deux éléments quelconques de A et B respectivement. On a : $a = (a + b) - b$. En particulier : $a \leq \sup(A + B) - b$.

On en déduit que pour tout $b \in B$, $\sup(A + B) - b$ est un majorant de A .

D'où : $\sup A \leq \sup(A + B) - b$ pour tout $b \in B$. D'où : $\forall b \in B, b \leq \sup(A + B) - \sup A$.

Par suite $\sup(A + B) - \sup A$ est un majorant de B , donc : $\sup B \leq \sup(A + B) - \sup A$.

D'où $\sup A + \sup B \leq \sup(A + B)$ (♣)

Conclusion : on déduit de (♠) et de (♣) que : $\sup(A + B) = \sup A + \sup B$

► 2) Puisque $A \subset (A \cup B)$ et $B \subset (A \cup B)$, on a déjà $\sup A \leq \sup(A \cup B)$ et $\sup(B) \leq \sup(A \cup B)$. Par suite : $\max(\sup A, \sup B) \leq \sup(A \cup B)$.

Réciproquement, soit $x \in A \cup B$. Alors $x \leq \sup A$ (si $x \in A$) ou $x \leq \sup B$ (si $x \in B$).

Donc $x \leq \max(\sup A, \sup B)$.

Il s'ensuit que $\max(\sup A, \sup B)$ est un majorant de $A \cup B$, d'où : $\sup(A \cup B) \leq \max(\sup A, \sup B)$.

On en déduit finalement que : $\sup(A \cup B) = \max(\sup A, \sup B)$. ☕

2. La réflexion puissante derrière ce raisonnement est qu'un majorant quelconque est supérieur ou égal au plus petit d'entre eux.

10.3 Conséquences de la propriété de la borne supérieure dans $\mathbb{R}$, notion de densité

THÉORÈME 10.1 - $\mathbb{R}$ est **archimédien**, càd : $\forall (x, y) \in \mathbb{R} \times \mathbb{R}_+^*, \exists n \in \mathbb{N}^*, ny > x$.

Traduction : “un voyage de 1000 lieues commence toujours par un premier pas” (Lao-Tseu).

PREUVE. Soit x un réel, et soit y un réel strictement positif. Si $x \leq 0$, $n = 1$ fait l'affaire.

On suppose donc à présent $x > 0$. Par l'absurde, supposons que $Ny \leq x$ pour tout $N \in \mathbb{N}^*$. Alors $A = \{Ny, N \in \mathbb{N}^*\}$ est une partie non vide (clair) et majorée (par x). Elle admet donc une borne supérieure, que nous noterons S .

On a alors : $\forall N \in \mathbb{N}^*, (N+1)y \leq S$, d'où $\forall N \in \mathbb{N}^*, Ny \leq S - y$. On en déduit que $S - y$ est un majorant de A . Or ce majorant est strictement inférieur à S (comme $y > 0$), qui est la borne supérieure de A : contradiction.

Il s'ensuit que l'hypothèse initiale est fautive d'où : $\boxed{\exists n \in \mathbb{N}^*, ny > x}$. ☕

Conséquence fun : entre deux rationnels, il existe une infinité d'irrationnels.

DÉFINITION 10.8 - Une partie A de $\mathbb{R}$ est **dense** dans $\mathbb{R}$ si :

$$\forall (x, y) \in \mathbb{R}^2, x < y, \exists a \in A, x < a < y$$

Remarques à un sou. $\mathbb{R}$ est dense dans $\mathbb{R}$; $\mathbb{N}$ et $\mathbb{Z}$ ne sont pas denses dans $\mathbb{R}$.

PROPRIÉTÉ 10.4 - (Caractérisation séquentielle de la densité).

Soit A une partie $\mathbb{R}$.

$$[A \text{ est dense dans } \mathbb{R}] \iff [\text{Tout réel est limite d'une suite d'éléments de } A]$$

PREUVE. Raisonnons par double implication.

► Supposons que A est dense dans $\mathbb{R}$: soit x un réel. Alors pour tout $n \in \mathbb{N}^*$, il existe un élément a_n de A tel que : $x < a_n < x + \frac{1}{n}$.

D'après le théorème d'encadrement, on en déduit que $\lim_{n \rightarrow +\infty} a_n = x$.

Par suite : $[A \text{ est dense dans } \mathbb{R}] \implies [\text{Tout réel est limite d'une suite d'éléments de } A]$.

► Supposons que tout réel est limite d'une suite d'éléments de A : soient x et y deux réels, tels que $x < y$. Puisque $(x+y)/2$ est limite d'une suite d'éléments de A , il existe au moins un élément a_n de A tel que :

$$\left| a_n - \frac{x+y}{2} \right| < \frac{y-x}{2}, \text{ càd tel que : } x < a_n < y.$$

Par suite : $[A \text{ est dense dans } \mathbb{R}] \impliedby [\text{Tout réel est limite d'une suite d'éléments de } A]$.

Ce qui établit l'équivalence de la propriété. ☕

Place à l'énoncé "star" de ce chapitre :

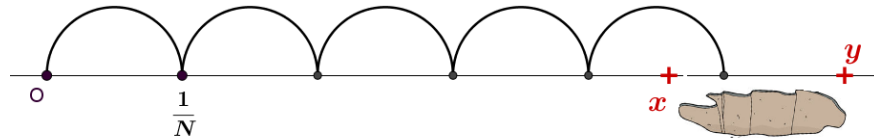
THÉORÈME 10.2 - $\mathbb{Q}$ est dense dans $\mathbb{R}$.

Traduction : entre deux nombres réels, il existe toujours un nombre rationnel.³

PREUVE. Il s'agit de prouver qu'il existe un nombre rationnels entre deux réels distincts. Considérons deux réels x et y , tels que $x < y$.

- Si $x < 0$ et $y > 0$: alors 0 (qui est rationnel !) est compris entre x et y .
- Si $0 \leq x < y$: puisque $\mathbb{R}$ est archimédien, il existe un entier naturel non nul N tel que : $N(y - x) > 1$ et il existe un entier naturel M tel que $M/N > x$.

Remarque : la figure ci-dessous illustre ces choix, dont on explique ici l'origine très informellement.



Entre x et y , il y a un trou, de longueur $y - x$. Pour être certain de tomber dans ce trou en partant de zéro, on doit s'assurer de deux conditions : ne pas faire des pas trop longs (sinon on pourrait passer au-dessus du trou), et en faire suffisamment.

Pour remplir la première, on choisit de faire des pas de longueur $1/N$ (qui est rationnel), de telle sorte que : $1/N < (y - x)$ (longueur d'un pas < longueur du trou). Ce qui équivaut à : $N(y - x) > 1$.

Pour remplir la seconde, on observe qu'il faut au moins dépasser x pour tomber dans le trou. Or, $\mathbb{R}$ étant archimédien, quitte à faire assez de pas de longueur $1/N$, on dépassera x . C'est-à-dire : il existe un entier M assez grand pour que $M \times (1/N) > x$.

Considérons l'ensemble : $E = \left\{ k \in \mathbb{N} / \frac{k}{N} \leq x \right\}$. Cet ensemble est non vide, puisque 0 y appartient. En outre, pour tout élément k de E , on a : $(k/N) \leq x < (M/N)$ d'où en particulier $k < M$. Donc E est majoré (par M).

En résumé : E est une partie de $\mathbb{N}$, non vide et majorée. A ce titre elle admet un plus grand élément, que nous noterons : N_0 . On a alors : $N_0 \in E$ et $N_0 + 1 \notin E$ (sinon N_0 ne serait pas le plus grand élément !).

D'où : $\boxed{\frac{N_0}{N} \leq x < \frac{N_0 + 1}{N}} \quad (\spadesuit).$

Par ailleurs, puisque $N(y - x) > 1$, on a : $y > x + \frac{1}{N}$. On déduit de cette inégalité et de $\frac{N_0}{N} \leq x$ que :

$\boxed{\frac{N_0 + 1}{N} < y} \quad (\clubsuit).$

On déduit de $(\spadesuit)$ et de $(\clubsuit)$ que : $\boxed{x < \frac{N_0 + 1}{N} < y}$, ce qui prouve l'existence d'un rationnel entre x et y .

3. Traduction plus libre : si on marche dans $\mathbb{R}$, on ne peut pas poser le pied sans écraser un rationnel (sauf si on a des pieds ponctuels).

► Si $x < y \leq 0$: posons $X = -x$ et $Y = -y$. D'après ce que nous venons d'établir plus haut, il existe un rationnel r tel que : $Y < r < X$. Donc $x < -r < y$, d'où la conclusion puisque l'opposé d'un rationnel est encore un rationnel.

Synthèse : $\forall (x, y) \in \mathbb{R}^2, x < y, \exists r \in \mathbb{Q}, x < r < y$ ($\mathbb{Q}$ est dense dans $\mathbb{R}$)



COROLLAIRE 10.1 - Tout nombre réel est limite d'une suite de rationnels.

PREUVE. C'est une conséquence immédiate du théorème 10.2 et de la propriété 10.4.



COROLLAIRE 10.2 - $\mathbb{R} \setminus \mathbb{Q}$ est dense dans $\mathbb{R}$.

Traduction : “entre deux nombres réels, il existe toujours un nombre irrationnel”.

PREUVE. Soit x un réel. Puisque $\mathbb{Q}$ est dense dans $\mathbb{R}$, il existe une suite de rationnels (ρ_n) convergeant vers x . Il suffit alors d'observer que la suite de terme général $\rho_n + \frac{\sqrt{2}}{n}$ est une suite d'irrationnels convergeant vers x ; via la caractérisation séquentielle de la densité, on peut en effet en déduire que $\mathbb{R} \setminus \mathbb{Q}$ est dense dans $\mathbb{R}$.



10.4 Partie entière d'un nombre réel

LEMME 10.1 - Pour tout réel x , il existe un unique entier relatif N tel que

$$N \leq x < N + 1$$

PREUVE. Soit x un réel positif ou nul. On considère l'ensemble $E = \{n \in \mathbb{N} / n \leq x\}$. L'ensemble E est une partie non vide (car $0 \in E$) et majorée de $\mathbb{N}$. En effet, puisque $\mathbb{R}$ est archimédien, il existe un entier n_0 tel que $n_0 + 1 > x$; tout élément de E doit être inférieur à n_0 .

On en déduit que E possède un plus grand élément, càd qu'il existe un entier N tel que : $N \in E$ et $N + 1 \notin E$. On en déduit que : $N \leq x < N + 1$.

Supposons à présent x strictement négatif. D'après le raisonnement précédent, il existe un unique entier naturel tel que : $N \leq -x < N + 1$. On en déduit que : $-N - 1 < x \leq -N$. Deux cas peuvent alors se présenter. Soit $x = -N$, et on a alors : $-N \leq x < -N + 1$; soit $x < -N$, et on a alors $-N - 1 \leq x < -N$.

Conclusion. Dans tous les cas, on a établi l'existence et l'unicité d'un entier relatif N tel que :

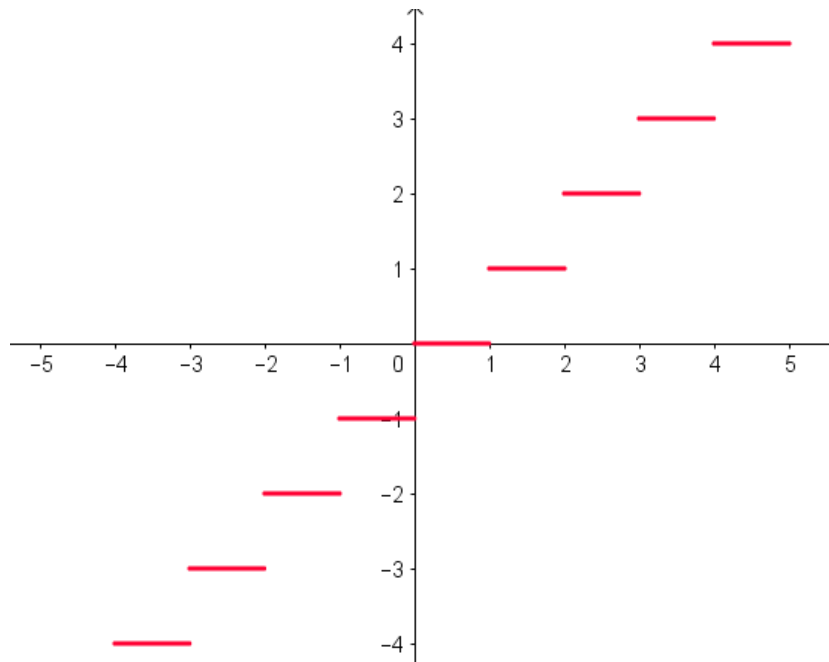
$$N \leq x < N + 1.$$



DÉFINITION 10.9 - L'unique entier relatif du corollaire ci-dessus est appelé **partie entière de x** , et est noté $\lfloor x \rfloor$.

Exemples. On a : $\lfloor \pi \rfloor = 3$; $\lfloor 0,5 \rfloor = 0$; $\lfloor \sqrt{2} \rfloor = 1$.

Mais attention : $\lfloor -\pi \rfloor = -4$; $\lfloor -0,5 \rfloor = -1$; $\lfloor -\sqrt{2} \rfloor = -2$.



COURBE REPRÉSENTATIVE DE LA FONCTION $x \in \mathbb{R} \mapsto [x]$

PROPRIÉTÉ 10.5 - (Propriétés de la partie entière).

1/ La fonction partie entière est croissante sur $\mathbb{R}$

2/ $\forall x \in \mathbb{R}, [x] \leq x < [x] + 1$

3/ $\forall x \in \mathbb{R}, x - 1 < [x] \leq [x]$

4/ $\forall x \in \mathbb{R}, \forall n \in \mathbb{N}^*, \left\lfloor \frac{[nx]}{n} \right\rfloor = [x]$

PREUVE. 1/ Soient $x \leq y$ deux nombres réels. Par définition de la partie entière, on a : $[x] \leq x$, d'où par hypothèse : $[x] \leq y$. Donc $[x]$ est un entier relatif inférieur ou égal à y , et à ce titre il est inférieur au plus grand entier relatif vérifiant cette propriété, c'est à dire : $[x] \leq [y]$.

Ce qui prouve que la fonction $x \in \mathbb{R} \mapsto [x]$ est croissante.

2/ Ce point est une conséquence immédiate de la définition de partie entière.

3/ D'après le point précédent : $\forall x \in \mathbb{R}, x - 1 < [x]$, l'autre inégalité de l'encadrement provenant du 1/.

4/ Soit x un réel, et soit n un entier naturel non nul. Alors $[nx] \leq nx$ (définition de partie entière), d'où

$\frac{[nx]}{n} \leq x$, d'où $\left\lfloor \frac{[nx]}{n} \right\rfloor = [x]$ ($\spadesuit$) (croissance de la partie entière).

Réciproquement, on a : $[x] \leq x$ (définition) donc $n[x] \leq nx$ donc $n[x] \leq [nx]$ (croissance) donc

$[x] \leq \left\lfloor \frac{[nx]}{n} \right\rfloor$ ($\clubsuit$) (re-croissance).

Conclusion : on déduit de ($\spadesuit$) et de ($\clubsuit$) que : $\forall x \in \mathbb{R}, \forall n \in \mathbb{N}^*, [x] = \left\lfloor \frac{[nx]}{n} \right\rfloor$



10.5 Application — $\mathbb{D}$ est dense dans $\mathbb{R}$

En préambule, commençons par une constatation : la partie entière de π vaut 3... OK !

On connaît aussi les premières approximations décimales par défaut de π , qui sont :

$$3,1 = \frac{31}{10}; \quad 3,14 = \frac{314}{100}; \quad 3,141 = \frac{3141}{1000}; \quad \dots$$

que l'on peut interpréter comme suit :

$$3,1 = \frac{\lfloor 10 \times \pi \rfloor}{10}; \quad 3,14 = \frac{\lfloor 10^2 \times \pi \rfloor}{10^2}; \quad 3,141 = \frac{\lfloor 10^3 \times \pi \rfloor}{10^3}; \quad \dots$$

On connaît encore les premières approximations décimales par excès de π , qui sont :

$$3,2 = \frac{32}{10}; \quad 3,15 = \frac{315}{100}; \quad 3,142 = \frac{3142}{1000}; \quad \dots$$

que l'on peut interpréter comme suit :

$$3,2 = \frac{\lfloor 10 \times \pi \rfloor + 1}{10}; \quad 3,15 = \frac{\lfloor 10^2 \times \pi \rfloor + 1}{10^2}; \quad 3,142 = \frac{\lfloor 10^3 \times \pi \rfloor + 1}{10^3}; \quad \dots$$

Il semble que les deux suites ainsi construites convergent toutes deux vers une limite commune. C'est ce que l'on s'attache à prouver dans ce paragraphe, dans une version plus générale.

Explicitement, pour tout réel x , on définit la suite u (*resp.* v) des approximations décimales par défaut (*resp.* par excès) de x en posant :

$$\forall n \in \mathbb{N}, \quad u_n = \frac{\lfloor 10^n x \rfloor}{10^n} \quad (\text{resp. } \forall n \in \mathbb{N}, \quad v_n = \frac{\lfloor 10^n x \rfloor + 1}{10^n})$$

L'énoncé ci-dessous affirme que ces deux suites convergent et ont x pour limite commune.

PROPRIÉTÉ 10.6 - Pour tout réel x on a : $\lim_{n \rightarrow +\infty} \frac{\lfloor 10^n x \rfloor}{10^n} = x$ et $\lim_{n \rightarrow +\infty} \frac{\lfloor 10^n x \rfloor + 1}{10^n} = x$.

PREUVE. Soit x un réel. L'idée de la preuve est d'établir que les suites u et v sont adjacentes, donc convergent vers une limite commune ; puis d'établir que cette limite commune est x .

Pour tout entier naturel n , on a $\lfloor 10^n x \rfloor \leq 10^n x \leq \lfloor 10^n x \rfloor + 1$, d'où :

$$\forall n \in \mathbb{N}, \quad \frac{\lfloor 10^n x \rfloor}{10^n} \leq x \leq \frac{\lfloor 10^n x \rfloor + 1}{10^n} \quad (\spadesuit).$$

Posons à présent pour tout entier naturel n : $u_n = \frac{\lfloor 10^n x \rfloor}{10^n}$ et $v_n = \frac{\lfloor 10^n x \rfloor + 1}{10^n}$.

On a : $\lfloor 10^n x \rfloor \leq 10^n x \Rightarrow 10 \lfloor 10^n x \rfloor \leq 10^{n+1} x$. Or $10 \lfloor 10^n x \rfloor$ est un entier, et à ce titre il est plus petit que le plus grand entier inférieur ou égal à $10^{n+1} x$. Autrement dit : $10 \lfloor 10^n x \rfloor \leq \lfloor 10^{n+1} x \rfloor$.

Par suite⁴ on obtient : $\frac{\lfloor 10^n x \rfloor}{10^n} \leq \frac{\lfloor 10^{n+1} x \rfloor}{10^{n+1}}$, soit $u_n \leq u_{n+1}$. Ainsi : la suite (u_n) est croissante. (♣)

Symétriquement : $10^n x < \lfloor 10^n x \rfloor + 1 \Rightarrow 10^{n+1} x < 10 (\lfloor 10^n x \rfloor + 1)$. Or $10 (\lfloor 10^n x \rfloor + 1)$ est un entier, et à ce titre il est supérieur ou égal au plus petit entier strictement supérieur à $10^{n+1} x$. Autrement dit : $10 (\lfloor 10^n x \rfloor + 1) \geq \lfloor 10^{n+1} x \rfloor + 1$.

D'où : $\frac{\lfloor 10^n x \rfloor + 1}{10^n} \geq \frac{\lfloor 10^{n+1} x \rfloor + 1}{10^{n+1}}$, soit $v_n \geq v_{n+1}$. Ainsi : la suite (v_n) est décroissante. (♡)

Enfin : $\forall n \in \mathbb{N}$, $v_n - u_n = 10^{-n}$. D'où : $\lim_{n \rightarrow +\infty} (v_n - u_n) = 0$. (◇)

► Il résulte de (♣), (♡) et (◇) que les suites u et v sont adjacentes. D'après le théorème du même nom, u et v convergent vers une limite commune ℓ . Il découle alors de (♠) et de la stabilité des inégalités larges par passage à la limite que $\ell = x$.

Conclusion : Pour tout réel x on a : $\lim_{n \rightarrow +\infty} \frac{\lfloor 10^n x \rfloor}{10^n} = x$ et $\lim_{n \rightarrow +\infty} \frac{\lfloor 10^n x \rfloor + 1}{10^n} = x$. ☕

COROLLAIRE 10.3 - $\mathbb{D}$ est dense dans $\mathbb{R}$.

PREUVE. D'après la propriété précédente, tout nombre réels est limite d'une suite de décimaux. Le résultat provient de cette constatation, et de la caractérisation séquentielle de la densité (cf propriété 10.4, page 231). ☕

10.6 Epilogue

10.6.1 $\mathbb{C}$ peut être totalement ordonné (si, si !)

On définit une relation binaire (notée $\prec$) sur $\mathbb{C}$ en posant :

$$\forall (z, z') \in \mathbb{C} \times \mathbb{C}, \quad z \prec z' \iff \begin{cases} \operatorname{Re}(z) < \operatorname{Re}(z') \\ \text{ou} \\ \operatorname{Re}(z) = \operatorname{Re}(z') \text{ et } \operatorname{Im}(z) \leq \operatorname{Im}(z') \end{cases}$$

PROPRIÉTÉ 10.7 - La relation $\prec$ est une relation d'ordre totale sur $\mathbb{C}$.

PREUVE. Soit z un nombre complexe. On a : $\operatorname{Re}(z) = \operatorname{Re}(z)$ et $\operatorname{Im}(z) \leq \operatorname{Im}(z)$. Il s'ensuit que $z \prec z$. Ainsi la relation $\prec$ est réflexive (♠).

Soient z et z' deux nombres complexes. On suppose que : $z \prec z'$ et $z' \prec z$. On commence par observer que l'on ne peut avoir $\operatorname{Re}(z) \neq \operatorname{Re}(z')$ (sinon l'une des deux hypothèses n'est pas vérifiée).

4. En divisant les deux termes de l'inégalité par 10^{n+1} .

On a donc $\operatorname{Re}(z) = \operatorname{Re}(z')$, et $\operatorname{Im}(z) \leq \operatorname{Im}(z')$ et $\operatorname{Im}(z') \leq \operatorname{Im}(z)$. Il s'ensuit que $\operatorname{Re}(z) = \operatorname{Re}(z')$ et $\operatorname{Im}(z) = \operatorname{Im}(z')$ d'où $z = z'$.

En résumé : $[(z \prec z') \wedge (z' \prec z)] \implies [z = z']$. Ainsi la relation $\prec$ est antisymétrique ($\clubsuit$).

Soient z, z' et z'' trois nombres complexes. On suppose que : $z \prec z'$ et $z' \prec z''$. Supposons que : $\operatorname{Re}(z) = \operatorname{Re}(z')$ et $\operatorname{Im}(z) \leq \operatorname{Im}(z')$, et $\operatorname{Re}(z') = \operatorname{Re}(z'')$ et $\operatorname{Im}(z') \leq \operatorname{Im}(z'')$.

Alors : $\operatorname{Re}(z) = \operatorname{Re}(z'')$ et $\operatorname{Im}(z) \leq \operatorname{Im}(z'')$. D'où $z \prec z''$.

Les autres cas se traitent de manière analogue, et permettent de conclure que :

$$[(z \prec z') \wedge (z' \prec z'')] \implies [z \prec z'']. \text{ Ainsi la relation } \prec \text{ est transitive } (\heartsuit).$$

Soient z et z' deux nombres complexes. Si leurs parties réelles sont distinctes, alors on a $\operatorname{Re}(z) < \operatorname{Re}(z')$ ou $\operatorname{Re}(z') < \operatorname{Re}(z)$, et donc : $z \prec z'$ ou $z' \prec z$.

Sinon on a $\operatorname{Re}(z) = \operatorname{Re}(z')$, mais au moins l'une des inégalités $\operatorname{Im}(z) \leq \operatorname{Im}(z')$ ou $\operatorname{Im}(z') \leq \operatorname{Im}(z)$ est vérifiée. Par suite, au moins l'une des relations $z \prec z'$ ou $z' \prec z$ est valide.

Dans tous les cas : $\forall (z, z') \in \mathbb{C}^2, (z \prec z') \vee (z' \prec z)$ ($\diamond$).

On déduit de ($\spadesuit$), ($\clubsuit$), ($\heartsuit$) et ($\diamond$) que la relation $\prec$ est une relation d'ordre totale sur $\mathbb{C}$.

Conclusion. $(\mathbb{C}, \prec)$ est totalement ordonné. ☕

Remarque. Evidemment, l'intérêt de cette relation d'ordre est à relativiser, notamment car elle n'est pas compatible avec les opérations usuelles dans $\mathbb{C}$. Par exemple, le carré d'un élément de $\mathbb{C}$ "positif" peut éventuellement être "négatif", ce qui fait un peu désordre. Explicitement, on a : $0 \prec i$ mais $i^2 \prec 0^2 \dots$

10.6.2 Un exemple-clef de relation d'équivalence : les congruences en arithmétique

DÉFINITION 10.10 - Soient a, b et n trois entiers relatifs. On dit que a est congru à b modulo n si $b - a$ est multiple de n , c'est à dire s'il existe $k \in \mathbb{Z}$ tel que : $b = a + kn$. On le note $a \equiv b [n]$.

Notons que : $a \equiv 0 [n]$ si et seulement si a est multiple de n .

PROPRIÉTÉ 10.8 - Soit n un entier, $n \geq 2$.

i/ $\forall a \in \mathbb{Z}, a \equiv a [n]$

ii/ $\forall (a, b) \in \mathbb{Z}^2, a \equiv b [n] \iff b \equiv a [n]$

iii/ $\forall (a, b, c) \in \mathbb{Z}^3, [a \equiv b [n] \text{ et } b \equiv c [n]] \implies [a \equiv c [n]]$

Autrement dit, la relation de congruence est une relation d'équivalence sur $\mathbb{Z}$.

PREUVE. Soit a un entier. On a : $a = a + 0 \times n$, d'où le point i/.

Si $a \equiv b [n]$, alors il existe $k \in \mathbb{Z}, a = b + kn$. D'où $b = a - kn$. En posant $K = -k$, on a donc : $a = b + Kn$ avec $k \in \mathbb{Z}$. Ainsi $b \equiv a [n]$, ce qui prouve l'implication : $a \equiv b [n] \implies b \equiv a [n]$; la réciproque vient du fait que l'on peut permuter a et b dans le raisonnement précédent, ce qui prouve l'équivalence du point ii/.

Sous les hypothèses de l'énoncé du iii, on a : $b = a + kn$ et $c = b + k'n$ d'où $c = a + (k + k')n$; ce qui prouve le transitivité de la relation de congruence. ☕

10.6.3 Une partie dense de $\mathbb{R}$

Exercice. On rappelle que l'ensemble $\mathbb{D}$ des nombres *décimaux* est l'ensemble des rationnels s'écrivant comme le quotient d'un entier relatif par une puissance de 10, explicitement :

$$\mathbb{D} = \left\{ \frac{n}{10^k} \mid n \in \mathbb{Z}, k \in \mathbb{N} \right\}$$

Dans cet exercice nous noterons $\mathbb{M}$ l'ensemble des nombres *millimaux*, c'est à dire l'ensemble des rationnels s'écrivant comme le quotient d'un entier relatif par une puissance de 1000. Explicitement :

$$\mathbb{M} = \left\{ \frac{n}{10^{3k}} \mid n \in \mathbb{Z}, k \in \mathbb{N} \right\}$$

Prouver que $\mathbb{M}$ est dense dans $\mathbb{R}$.

Corrigé. Montrons que l'ensemble

$$\mathbb{M} = \left\{ \frac{n}{10^{3k}} \mid n \in \mathbb{Z}, k \in \mathbb{N} \right\}$$

est dense dans $\mathbb{R}$ en prouvant que tout réel est limite d'une suite d'éléments de $\mathbb{M}$. Soit x un réel. Pour tout entier naturel n , on a $\lfloor 10^{3n}x \rfloor \leq 10^{3n}x \leq \lfloor 10^{3n}x \rfloor + 1$, d'où :

$$\forall n \in \mathbb{N}, \quad \frac{\lfloor 10^{3n}x \rfloor}{10^{3n}} \leq x \leq \frac{\lfloor 10^{3n}x \rfloor + 1}{10^{3n}} \quad (\spadesuit).$$

Posons à présent pour tout entier naturel n : $u_n = \frac{\lfloor 10^{3n}x \rfloor}{10^{3n}}$ et $v_n = \frac{\lfloor 10^{3n}x \rfloor + 1}{10^{3n}}$.

On a : $\lfloor 10^{3n}x \rfloor \leq 10^{3n}x \Rightarrow 10 \lfloor 10^{3n}x \rfloor \leq 10^{3n+1}x$. Or $10 \lfloor 10^{3n}x \rfloor$ est un entier, et à ce titre il est plus petit que le plus grand entier inférieur ou égal à $10^{3n+1}x$. Autrement dit : $10 \lfloor 10^{3n}x \rfloor \leq \lfloor 10^{3n+1}x \rfloor$.

Par suite⁵ on obtient : $\frac{\lfloor 10^{3n}x \rfloor}{10^{3n}} \leq \frac{\lfloor 10^{3n+1}x \rfloor}{10^{3n+1}}$, soit $u_n \leq u_{n+1}$. Ainsi : la suite (u_n) est croissante. $(\clubsuit)$

Symétriquement : $10^{3n}x < \lfloor 10^{3n}x \rfloor + 1 \Rightarrow 10^{3n+1}x < 10(\lfloor 10^{3n}x \rfloor + 1)$. Or $10(\lfloor 10^{3n}x \rfloor + 1)$ est un entier, et à ce titre il est supérieur ou égal au plus petit entier strictement supérieur à $10^{3n+1}x$. Autrement dit : $10(\lfloor 10^{3n}x \rfloor + 1) \geq \lfloor 10^{3n+1}x \rfloor + 1$.

D'où : $\frac{\lfloor 10^{3n}x \rfloor + 1}{10^{3n}} \geq \frac{\lfloor 10^{3n+1}x \rfloor + 1}{10^{3n+1}}$, soit $v_n \geq v_{n+1}$. Ainsi : la suite (v_n) est décroissante. $(\heartsuit)$

Enfin : $\forall n \in \mathbb{N}, v_n - u_n = 10^{-3n}$. D'où : $\lim_{n \rightarrow +\infty} (v_n - u_n) = 0$. $(\diamondsuit)$

► Il résulte de $(\clubsuit)$, $(\heartsuit)$ et $(\diamondsuit)$ que les suites u et v sont adjacentes. D'après le théorème du même nom, u et v convergent vers une limite commune ℓ . Il découle alors de $(\spadesuit)$ et de la stabilité des inégalités larges par passage à la limite que $\ell = x$.

5. En divisant les deux termes de l'inégalité par 10^{3n+1} .

Bilan : Pour tout réel x on a : $\lim_{n \rightarrow +\infty} \frac{\lfloor 10^{3n}x \rfloor}{10^{3n}} = x$ et $\lim_{n \rightarrow +\infty} \frac{\lfloor 10^{3n}x \rfloor + 1}{10^{3n}} = x$.

Conclusion. Tout réel est limite d'une suite d'éléments de $\mathbb{M}$. Par caractérisation séquentielle de la densité, ceci signifie que $\mathbb{M}$ est dense dans $\mathbb{R}$.

Remarque : il existe d'autres moyens pour obtenir résultat. Il est notamment beaucoup plus rapide d'utiliser le fait que $\mathbb{D}$ est dense dans $\mathbb{R}$, et d'utiliser les propriétés des suites extraites.

Explicitement : soit $x \in \mathbb{R}$. Puisque $\mathbb{D}$ est dense dans $\mathbb{R}$, il existe une suite de décimaux (u_n) qui converge vers x . Alors la suite extraite (u_{3n}) est une suite d'éléments de $\mathbb{M}$ qui converge vers x (théorème des suites extraites). Ainsi tout réel est limite d'une suite d'éléments de $\mathbb{M}$. Ce qui prouve (de nouveau, mais bien plus vite) que $\mathbb{M}$ est dense dans $\mathbb{R}$.

Chapitre 11

Suites

Notations et conventions : une suite réelle est une application de $\mathbb{N}$ dans $\mathbb{R}$; on notera donc :

- $\mathbb{R}^{\mathbb{N}}$ l'ensemble des suites réelles ;
- $\mathbb{C}^{\mathbb{N}}$ celui des suites complexes ;
- $\mathbb{K}^{\mathbb{N}}$ celui des suites à valeurs dans $\mathbb{K}$ (avec $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$).

Un élément de $\mathbb{R}^{\mathbb{N}}$ pourra être noté (u_n) ou $(u_n)_n$ ou simplement u . La notation u_n (sans parenthèse donc) sera réservée au terme général de la suite (u_n) .

Si u et v sont deux suites réelles, l'inégalité $u \geqslant v$ signifie que : $\forall n \in \mathbb{N}, u_n \geqslant v_n$.

En outre, on dira que l'inégalité $u \geqslant v$ est valable à partir *d'un certain rang* s'il existe $n_0 \in \mathbb{N}$ tel que : $\forall n \in \mathbb{N}, [n \geqslant n_0] \Rightarrow [u_n \geqslant v_n]$.

On rappelle également que pour deux réels a et x fixés, l'inégalité :

$$|x - a| < R$$

signifie que la distance entre a et x est strictement inférieure à R , c'est à dire que :

$$a - R < x < a + R \qquad \text{ou} \qquad x \in]a - R; a + R[$$

11.1 Généralités sur les suites réelles

DÉFINITION 11.1 - Soit u une suite réelle. u est :

- **majorée** si : $\exists M \in \mathbb{R}, \forall n \in \mathbb{N}, u_n \leq M$.
- **minorée** si : $\exists m \in \mathbb{R}, \forall n \in \mathbb{N}, u_n \geq m$.
- **bornée** si elle est majorée et minorée.
- **croissante** si : $\forall n \in \mathbb{N}, u_{n+1} \geq u_n$.
- **décroissante** si : $\forall n \in \mathbb{N}, u_{n+1} \leq u_n$.
- **monotone** si elle est croissante ou décroissante.
- **constante** si : $\forall n \in \mathbb{N}, u_{n+1} = u_n$.
- **stationnaire** si : $\exists n_0 \in \mathbb{N}, \forall n \geq n_0, u_{n+1} = u_n$.
- **périodique** si $\exists N \in \mathbb{N}^*, \forall n \in \mathbb{N}, u_{n+N} = u_n$.

Exemples :

→ La suite de terme général $u_n = \sum_{k=1}^n \left(\frac{1}{k} - \frac{1}{k+1} \right)$ est majorée par 1, puisque :

$$\forall n \in \mathbb{N}^*, u_n = 1 - \frac{1}{n+1} \leq 1$$

→ Observation pas chère ! Une suite réelle positive (*resp.* négative) est minorée (*resp.* majorée) par 0.

→ Si f est une fonction continue et positive sur $\mathbb{R}_+$, alors la suite de terme général $u_n = \int_0^n f(t) dt$ est croissante. En effet :

$$\forall n \in \mathbb{N}, u_{n+1} - u_n = \int_0^{n+1} f(t) dt - \int_0^n f(t) dt = \int_n^{n+1} f(t) dt \geq 0$$

la dernière inégalité provenant de la positivité de l'intégrale.

→ Si f est une fonction continue et positive sur $[0, 1]$, alors la suite de terme général $u_n = \int_0^n t^n f(t) dt$ est décroissante. En effet :

$$\forall n \in \mathbb{N}, u_{n+1} - u_n = \int_0^1 t^{n+1} f(t) dt - \int_0^1 t^n f(t) dt = \int_0^1 \underbrace{t^n (t-1)}_{\leq 0} f(t) dt \leq 0$$

la dernière inégalité provenant de nouveau de la positivité de l'intégrale.

→ Lorsque u est une suite réelle positive, la suite de terme général $S_n = \sum_{k=0}^n u_k$ est croissante, puisque :

$$\forall n \in \mathbb{N}, u_{n+1} - u_n = \sum_{k=0}^{n+1} u_k - \sum_{k=0}^n u_k = u_{n+1}$$

→ Notons I_n la n -ième intégrale de Wallis : $I_n = \int_0^{\pi/2} \cos^n(t) dt$. Pour tout entier naturel n , on a : $u_{n+1} = (n+2) I_{n+2} I_{n+1}$.

Or, il est par ailleurs connu que : $u_{n+1} = (n+2) \frac{n+1}{n+2} I_n I_{n+1} = (n+1) I_{n+1} I_n$, d'où : $u_{n+1} = u_n$. On en déduit que (u_n) est constante.

Par suite : $\forall n \in \mathbb{N}, u_n = u_0 = I_1 \times I_0 = \frac{\pi}{2}$. **Conclusion.** $\forall n \in \mathbb{N}, (n+1) I_{n+1} I_n = \frac{\pi}{2}$.

→ La suite de terme général $\left\lfloor \frac{3}{n} \right\rfloor$ est stationnaire (constante égale à 0 à partir de $n = 4$), tout comme celle de terme général $\binom{4}{n}$ (constante égale à 0 à partir de $n = 5$).

→ La suite de terme général $(-1)^n$ est périodique de période 2.

→ La suite (complexe) de terme général i^n (*resp.* j^n) est périodique de période 4 (*resp.* 3).

L'énoncé ci-dessous, démontré dans le cadre des fonctions, est encore valide pour les suites (et est très utile en pratique) :

LEMME 11.1 - u est bornée SSI $\exists M \in \mathbb{R}_+, \forall n \in \mathbb{N}, |u_n| \leq M$.

PREUVE. On rappelle simplement le principe de la preuve : si u est bornée, alors il existe deux réels a et b tels que pour tout entier naturel n , on a $a \leq u_n \leq b$. Ceci implique que :

$$\forall n \in \mathbb{N}, |u_n| \leq \max(-a, b)$$

Par suite : u bornée $\implies |u|$ majorée. L'implication réciproque est triviale. ☕

11.2 Limite d'une suite réelle

11.2.1 Généralités sur les limites

DÉFINITION 11.2 - soient u une suite et ℓ un réel. u **converge vers** ℓ (ou a pour limite ℓ) si :

$$\forall \varepsilon > 0, \exists n_0 \in \mathbb{N}, \forall n \in \mathbb{N}, (n \geq n_0 \implies |u_n - \ell| \leq \varepsilon)$$

Lorsque tel est le cas, on note : $\ell = \lim_{n \rightarrow +\infty} u_n$, ou simplement $\ell = \lim u$.

La suite u diverge si elle n'est pas convergente, c'est-à-dire si elle n'a pas de limite finie.

Exemples :

→ La suite de terme général $u_n = \sum_{k=0}^n \frac{1}{k!}$ converge vers e .

→ La suite de terme général $u_n = \sum_{k=1}^n \frac{1}{k^2}$ converge vers $\frac{\pi^2}{6}$.

→ La suite de terme général $u_n = \sum_{k=1}^n \frac{1}{k}$ diverge (pas de limite finie).

→ La suite de terme général $u_n = (-1)^n$ (pas de limite du tout).

LEMME 11.2 - u converge vers ℓ SSI $(u - \ell)$ converge vers 0.

PREUVE. Soient u une suite réelle et ℓ un réel. On a :

u converge vers $\ell \iff \forall \varepsilon > 0, \exists n_0 \in \mathbb{N}, \forall n \in \mathbb{N}, (n \geq n_0 \implies |(u_n - \ell) - 0| \leq \varepsilon) \dots$ ☕

PROPRIÉTÉ 11.1 - (unicité de la limite). Si u est une suite convergente, alors sa limite est unique.

PREUVE. Soit u une suite réelle convergente. Raisonnons par l'absurde et supposons qu'elle converge vers ℓ et ℓ' , deux réels distincts. Sans nuire à la généralité, on peut supposer que $\ell < \ell'$.

Posons : $\varepsilon = \frac{\ell' - \ell}{3}$ (notons que $\varepsilon > 0$).

► Comme u converge vers ℓ : $\exists n_0 \in \mathbb{N}, (n \geq n_0 \implies \ell - \varepsilon < u_n < \ell + \varepsilon)$

► Comme u converge vers ℓ' : $\exists n_1 \in \mathbb{N}, (n \geq n_1 \implies \ell' - \varepsilon < u_n < \ell' + \varepsilon)$

Posons $n_2 = \max(n_0, n_1)$. Pour tout entier $n \geq n_2$, les deux inégalités ci-dessus sont satisfaites, et on en déduit en particulier que :

$$\ell' - \varepsilon < \ell + \varepsilon \text{ d'où : } \ell' - \ell < 2\varepsilon \iff \ell' - \ell < 2\frac{\ell' - \ell}{3} \iff 1 < \frac{2}{3} \text{ ce qui est absurde.}$$

Conclusion. Si u converge, alors sa limite est unique. ☕

PROPRIÉTÉ 11.2 - Toute suite convergente est bornée (si u est une suite convergente, alors u est bornée).

PREUVE. Soit u une suite convergente de limite ℓ .

Alors : $\exists n_0 \in \mathbb{N}, (n \geq n_0 \implies \ell - 1 < u_n < \ell + 1)$ (♠).

On pose alors : $M = \max_{i=0}^{n_0-1} u_i$ et $m = \min_{i=0}^{n_0-1} u_i$. On a alors : $\forall n \in \mathbb{N}, (n < n_0 \implies m \leq u_n \leq M)$ (♣).

D'après (♠) et (♣) : $\forall n \in \mathbb{N}, \min(\ell - 1, m) \leq u_n \leq \max(\ell + 1, M)$. Par suite, u est bornée.

Conclusion. Toute suite convergente est bornée. ☕

Remarque : la réciproque de la propriété précédente est évidemment fausse. La suite de terme général $(-1)^n$ est bornée, mais ne converge pas pour autant (idem avec les suites de terme général $\cos(n)$, $\sin(n)$, ...).

LEMME 11.3 - Si u a pour limite un réel $\ell > 0$, alors $u > 0$ à partir d'un certain rang.

PREUVE. Soit u une suite réelle ayant pour limite un réel $\ell > 0$. Alors :

$$\exists n_0 \in \mathbb{N}, \forall n \in \mathbb{N}, [n \geq n_0] \implies \left[|u_n - \ell| < \frac{\ell}{2} \right]$$

En particulier : $\exists n_0 \in \mathbb{N}, \forall n \in \mathbb{N}, [n \geq n_0] \implies u_n > \frac{\ell}{2} > 0$. ☕

Remarque. Naturellement, le lemme précédent s'adapte au cas où $\ell < 0$. On peut donc généraliser la propriété précédente, et énoncer la suivante :

LEMME 11.4 - Si u a pour limite un réel $\ell \neq 0$, alors la suite u est de signe constant à partir d'un certain rang.

PREUVE. A la lumière de ce qui précède, triviale. ☕

11.2.2 Opérations algébriques sur les limites

PROPRIÉTÉ 11.3 - (**Propriétés algébriques des limites**). Soient u et v deux suites convergentes, de limites respectives ℓ et ℓ' .

1/ $u + v$ converge vers $\ell + \ell'$

2/ uv converge vers $\ell\ell'$

3/ $1/v$ converge vers $1/\ell'$ (sous réserve que $\ell' \neq 0$)

4/ u/v converge vers ℓ/ℓ' (sous réserve que $\ell' \neq 0$)

PREUVE. 1/ Soient u et v deux suites, de limites respectives ℓ et ℓ' .

Soit $\varepsilon > 0$.

Comme u a pour limite ℓ : $\exists n_0 \in \mathbb{N}, (n \geq n_0 \implies |u_n - \ell| < \frac{\varepsilon}{2})$ (♠)

Comme v a pour limite ℓ' : $\exists n_1 \in \mathbb{N}, (n \geq n_1 \implies |v_n - \ell'| < \frac{\varepsilon}{2})$ (♣)

Posons $n_2 = \max(n_0, n_1)$. Pour tout entier $n \geq n_2$, les deux inégalités ci-dessus sont satisfaites, et on a donc :

$|u_n + v_n - (\ell + \ell')| \leq |u_n - \ell| + |v_n - \ell'| < \varepsilon$ (la première inégalité provenant de l'inégalité triangulaire).

En résumé, on a établi que : $\forall \varepsilon > 0, \exists n_2 \in \mathbb{N}, (n \geq n_2 \implies |u_n + v_n - (\ell + \ell')| < \varepsilon)$

Conclusion. Si u et v convergent vers ℓ et ℓ' respectivement, alors $u + v$ converge vers $\ell + \ell'$

2/ Soient u et v deux suites, de limites respectives ℓ et ℓ' . On supposera dans cette preuve $\ell \neq 0$ et $\ell' \neq 0$.

Soit $\varepsilon > 0$.

► Pour tout entier naturel n , on a : $u_n v_n - \ell\ell' = u_n v_n - u_n \ell' + u_n \ell' - \ell\ell' = u_n (v_n - \ell') + \ell' (u_n - \ell)$, d'où grâce à l'inégalité triangulaire :

$$\forall n \in \mathbb{N}, |u_n v_n - \ell\ell'| \leq |u_n| \times |v_n - \ell'| + |\ell'| \times |u_n - \ell| \quad (\spadesuit)$$

► Puisque u est convergente, donc bornée : $\exists M > 0, \forall n \in \mathbb{N}, |u_n| \leq M$.

Et puisque v converge vers ℓ' : $\exists n_0 \in \mathbb{N}, \left(n \geq n_0 \implies |v_n - \ell'| < \frac{\varepsilon}{2M} \right)$.

D'où : $\forall n \geq n_0 \implies |u_n| \times |v_n - \ell'| < \frac{\varepsilon}{2}$ (♣)

► D'autre part, puisque u converge vers ℓ : $\exists n_1 \in \mathbb{N}, \left(n \geq n_1 \implies |u_n - \ell| < \frac{\varepsilon}{2|\ell'|} \right)$.

D'où : $\forall n \geq n_1 \implies |\ell'| \times |u_n - \ell| < \frac{\varepsilon}{2}$ (♡)

Posons $n_2 = \max(n_0, n_1)$. Pour tout entier $n \geq n_2$, les deux inégalités de (♡) et (♣) sont satisfaites, et on a donc, en utilisant (♠) :

$$\forall n \in \mathbb{N}, \left(n \geq n_2 \implies |u_n v_n - \ell \ell'| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \right)$$

En résumé, on a établi que : $\forall \varepsilon > 0, \exists n_2 \in \mathbb{N}, (n \geq n_2 \implies |u_n v_n - \ell \ell'| < \varepsilon)$

Conclusion. Si u et v convergent vers ℓ et ℓ' respectivement, alors uv converge vers $\ell \ell'$.

3/ Soit v une suite réelle, convergeant vers un réel $\ell' \neq 0$. SNALG, on peut supposer : $\ell' > 0$.

D'après le lemme 11.3, les termes de la suite v_n sont strictement positifs à partir d'un certain rang, disons n_0 .

Considérons à présent un entier naturel $n \geq n_0$. On a :

$$\left| \frac{1}{v_n} - \frac{1}{\ell'} \right| = \frac{|\ell' - v_n|}{\ell' v_n}. \quad 1 \quad (\spadesuit)$$

En outre, puisque la v converge vers ℓ' , il existe un entier naturel n_1 tel que pour tout $n \geq n_1$ on a : $v_n > \frac{\ell'}{2}$

Posons : $n_2 = \max(n_0, n_1)$. On a : $\forall n \in \mathbb{N}, \left[n \geq n_2 \implies 0 < \frac{1}{\ell' v_n} < \frac{2}{\ell'^2} \right]$ (♣)

Ces précautions prises, considérons un réel $\varepsilon > 0$. La suite v convergeant vers ℓ' :

$$\exists n_3 \in \mathbb{N}, \forall n \in \mathbb{N}, \left[n \geq n_3 \implies |v_n - \ell'| < \frac{\varepsilon \ell'^2}{2} \right] \quad (\heartsuit)$$

En posant $n_4 = \max(n_2, n_3)$, on déduit de (♠), (♣) et (♡) que :

$$\forall n \in \mathbb{N}, \left[n \geq n_4 \implies \left| \frac{1}{v_n} - \frac{1}{\ell'} \right| < \varepsilon \right]$$

Le réel $\varepsilon > 0$ étant arbitraire dans le raisonnement ci-dessus, on peut conclure que :

$$\forall \varepsilon > 0, \exists n_4 \in \mathbb{N}, \forall n \in \mathbb{N}, \left[n \geq n_4 \implies \left| \frac{1}{v_n} - \frac{1}{\ell'} \right| < \varepsilon \right]$$

Ce qui signifie exactement que : $\lim_{n \rightarrow +\infty} \frac{1}{v_n} = \frac{1}{\ell'}$.

Conclusion. Si v converge $\ell' \neq 0$, alors $1/v$ est définie à partir d'un certain rang et converge vers $\frac{1}{\ell'}$.

1. Puisque ℓ' et v_n sont strictement positifs, on a : $|\ell' v_n| = \ell' v_n$

4/ Conséquence immédiate des points 2/ et 3/. ☕

COROLLAIRE 11.1 - Toute combinaison linéaire de suites convergentes est convergente.

Plus explicitement et précisément, si u et v sont deux suites convergentes de limites respectives ℓ et ℓ' , alors pour tout couple de réels (λ, μ) , la suite $\lambda u + \mu v$ est convergente et

$$\lim (\lambda u + \mu v) = \lambda \lim u + \mu \lim v$$

PREUVE. Conséquence immédiate des points 1 et 2 de la propriété précédente. ☕

PROPRIÉTÉ 11.4 - Si u est bornée, et v converge vers 0, alors uv converge vers 0.

PREUVE. Soient v une suite convergeant vers 0, et u une suite bornée.

Puisque u est bornée, il existe un réel strictement positif M tel que $\forall n \in \mathbb{N}, |u_n| \leq M$ (♠).

Soit $\varepsilon > 0$.

Comme v a pour limite 0 : $\exists n_0 \in \mathbb{N}, (n \geq n_0 \implies |v_n| < \frac{\varepsilon}{M})$ (♣)

On déduit de (♠) et (♣) : $\exists n_0 \in \mathbb{N}, (n \geq n_0 \implies |u_n v_n| < M \times \frac{\varepsilon}{M} = \varepsilon)$

En résumé, on a établi que : $\forall \varepsilon > 0, \exists n_0 \in \mathbb{N}, (n \geq n_0 \implies |u_n v_n| < \varepsilon)$

Conclusion. Si u est bornée, et v tend vers 0, alors suite uv tend vers 0. ☕

Application. Considérons la suite de terme général $w_n = \frac{1}{n} \int_0^n \frac{1}{1+t^2} dt$.

On a : $\forall n \in \mathbb{N}, 0 \leq \int_0^n \frac{1}{1+t^2} dt \leq \frac{\pi}{2}$.

La suite de terme général $\int_0^n \frac{1}{1+t^2} dt$ est donc bornée, et puisqu'il est clair que celle de terme général $\frac{1}{n}$

tend vers 0, on peut conclure que : $\lim_{n \rightarrow +\infty} \frac{1}{n} \int_0^n \frac{1}{1+t^2} dt = 0$.

11.2.3 Suites réelles de limite infinie

DÉFINITION 11.3 - Soit u une suite. u **diverge vers** $+\infty$ (ou a pour limite $+\infty$) si : $\forall M > 0, \exists n_0 \in \mathbb{N}, \forall n \in \mathbb{N}, (n \geq n_0 \implies u_n \geq M)$. Et de manière analogue, u **diverge vers** $-\infty$ si : $\forall M < 0, \exists n_0 \in \mathbb{N}, \forall n \in \mathbb{N}, (n \geq n_0 \implies u_n \leq M)$.

Exemples : la suite de terme général $\sum_{k=1}^n \frac{1}{k}$ diverge vers $+\infty$. Il en va de même pour la suite de terme général $\sum_{k=2}^n \frac{1}{\ln(n)}$ ou encore $\sum_{k=1}^n \frac{1}{\sqrt{k}}$ ou encore $\sum_{k=0}^n \frac{1}{\cos^2(n)}$.

LEMME 11.5 - Si u diverge vers $+\infty$, alors $1/u$ existe à partir d'un certain rang et converge vers 0.

PREUVE. Si u diverge vers $+\infty$, il découle immédiatement de la définition que u est strictement positive à partir d'un certain rang ; ce qui permet de définir la suite $1/u$ au moins à partir de ce rang.

Soit à présent $\varepsilon > 0$.

Puisque u diverge vers $+\infty$:

$$\exists n_0 \in \mathbb{N}, \forall n \in \mathbb{N}, [n \geq n_0] \implies \left[u_n > \frac{1}{\varepsilon} \right]$$

Par suite :

$$\forall n \in \mathbb{N}, [n \geq n_0] \implies \left[0 < \frac{1}{u_n} < \varepsilon \right]$$

Le réel ε étant arbitraire dans le raisonnement ci-dessus, on peut conclure :

$$\forall \varepsilon > 0, \exists n_0 \in \mathbb{N}, \forall n \in \mathbb{N}, [n \geq n_0] \implies \left[0 < \frac{1}{u_n} < \varepsilon \right] \quad \text{càd : } \boxed{\lim_{n \rightarrow +\infty} \frac{1}{u_n} = 0} \quad \text{☕}$$

11.3 Suites réelles et inégalités

Toutes les suites considérées dans ce paragraphe sont REELLES.

11.3.1 Limites et inégalités

PROPRIÉTÉ 11.5 - Stabilité des inégalités larges par passage à la limite Soient u et v deux suites convergentes, de limites respectives ℓ et ℓ' . Si $u \leq v$ à partir d'un certain rang, alors $\ell \leq \ell'$.

PREUVE. Soient u et v deux suites réelles convergeant vers ℓ et ℓ' respectivement. On suppose en outre que :

$$\exists n_0 \in \mathbb{N}, \forall n \in \mathbb{N}, [n \geq n_0] \implies [u_n \leq v_n]$$

Raisonnons par l'absurde et supposons que $\ell > \ell'$ (♠).

Alors, puisque u converge vers ℓ :

$$\exists n_1 \in \mathbb{N}, \forall n \in \mathbb{N}, [n \geq n_1] \implies \left[|u_n - \ell| \leq \frac{\ell - \ell'}{3} \right]$$

Et puisque v converge vers ℓ' :

$$\exists n_2 \in \mathbb{N}, \forall n \in \mathbb{N}, [n \geq n_2] \implies \left[|v_n - \ell'| \leq \frac{\ell - \ell'}{3} \right]$$

Posons : $n_3 = \max(n_0, n_1, n_2)$. On a pour tout entier $n \geq n_3$:

$$\ell - \frac{\ell - \ell'}{3} < u_n \leq v_n < \ell' + \frac{\ell - \ell'}{3}$$

D'où en particulier : $\ell - \frac{\ell - \ell'}{3} < \ell' + \frac{\ell - \ell'}{3}$ soit : $\ell - \ell' < 2 \frac{\ell - \ell'}{3}$, càd : $1 < \frac{2}{3}$. Contradiction !

On en déduit que l'hypothèse ($\spadesuit$) est faussée. Par conséquent : $\ell \leq \ell'$. ☕

Remarque. Les inégalités strictes ne sont en revanche pas stables par passage à la limite ; il suffit pour cela de considérer la suite de terme général $u_n = 1/n$. On a évidemment $u_n > 0$, mais $\lim u_n = 0$.

THÉORÈME 11.1 - (comparaison). Soient u et v deux suites telles que :

- à partir d'un certain rang : $u \leq v$ (resp. $u \geq v$)
- $\lim u = +\infty$ (resp. $-\infty$)

Alors $\lim v = +\infty$ (resp. $-\infty$).

PREUVE. Soient u et v deux suites réelles. On suppose que : $\lim u = +\infty$, et que :

$$\exists n_0 \in \mathbb{N}, \forall n \in \mathbb{N}, [n \geq n_0] \implies [u_n \leq v_n]$$

Soit M un réel. Puisque u diverge vers $+\infty$, il existe un entier naturel n_1 tel que :

$$\forall n \in \mathbb{N}, [n \geq n_1] \implies [u_n \geq M]$$

Posons $n_2 = \max(n_0, n_1)$. Pour tout entier naturel $n \geq n_2$, on a : $v_n \geq u_n \geq M$.

Le réel M étant arbitraire dans ce qui précède, on a établi que :

$$\forall M \in \mathbb{R}, \exists n_2 \in \mathbb{N}, \forall n \in \mathbb{N}, [n \geq n_2] \implies [v_n \geq M]$$

Ce qui signifie exactement que : $\lim_{n \rightarrow +\infty} v_n = +\infty$.

La propriété est donc établie dans le cas “ $+\infty$ ”. L'autre cas en est une conséquence immédiate. ☕

Application : $\lim_{n \rightarrow +\infty} \left[\sum_{k=1}^n \frac{1}{k} \right] = +\infty$.

THÉORÈME 11.2 - (Encadrement ou “des gendarmes”). Soient u , v et w trois suites telles que :

- à partir d'un certain rang : $u \leq v \leq w$
- u et w sont convergentes
- $\lim u = \lim w$

Alors v converge et $\lim v = \lim u = \lim w$.

PREUVE. Soit $\varepsilon > 0$.

- Par hypothèse : $\exists n_0 \in \mathbb{N}, (n \geq n_0 \implies u_n \leq v_n \leq w_n)$
- Comme u a pour limite ℓ : $\exists n_1 \in \mathbb{N}, (n \geq n_1 \implies \ell - \varepsilon < u_n < \ell + \varepsilon)$
- Comme w a pour limite ℓ : $\exists n_2 \in \mathbb{N}, (n \geq n_2 \implies \ell - \varepsilon < w_n < \ell + \varepsilon)$

Posons : $n_3 = \max(n_0, n_1, n_2)$. Pour $n \geq n_3$, les trois conditions précédentes sont réalisées simultanément, et on a donc :

$$\ell - \varepsilon < u_n \leq v_n \leq w_n < \ell + \varepsilon \text{ d'où en particulier : } \ell - \varepsilon < v_n < \ell + \varepsilon \text{ soit : } |v_n - \ell| < \varepsilon.$$

En résumé, on a prouvé que : $\forall \varepsilon > 0, \exists n_3 \in \mathbb{N}, (n \geq n_3 \implies |v_n - \ell| < \varepsilon)$.

Donc la suite v est convergente et a pour limite $\ell = \lim u = \lim w$.



Exemple 1 : $\lim_{n \rightarrow +\infty} \int_0^1 t^{2n+1} \arctan(t) dt = 0$ puisque pour tout $n : 0 \leq \int_0^1 t^{2n+1} \arctan(t) dt \leq \frac{\pi}{8n+8}$.

Exemple 2 : pour tout $n \in \mathbb{N}^*$, on pose : $H_n = \sum_{k=1}^n \frac{1}{k}$.

On peut établir que pour tout $n \in \mathbb{N}^*$, $\ln(n+1) \leq H_n \leq \ln(n) + 1$. On en déduit que pour tout entier $n \geq 2$, on a :

$$\frac{\ln(n+1)}{\ln(n)} \leq \frac{H_n}{\ln(n)} \leq \frac{\ln(n)+1}{\ln(n)}$$

$$\text{Or : } \frac{\ln(n+1)}{\ln(n)} = \frac{\ln(n) + \ln(1+1/n)}{\ln(n)} = 1 + \frac{\ln(1+1/n)}{\ln(n)} \text{ d'où : } \lim_{n \rightarrow +\infty} \frac{\ln(n+1)}{\ln(n)} = 1.$$

$$\text{D'autre part : } \frac{\ln(n)+1}{\ln(n)} = 1 + \frac{1}{\ln(n)} \text{ d'où : } \lim_{n \rightarrow +\infty} \frac{\ln(n)+1}{\ln(n)} = 1.$$

On en déduit grâce au théorème d'encadrement que : $\lim_{n \rightarrow +\infty} \frac{H_n}{\ln(n)} = 1$ ²

11.3.2 Limites et monotonie

THÉORÈME 11.3 - (Limite monotone). Toute suite croissante et majorée converge. Toute suite décroissante et minorée converge.

PREUVE. Soit u une suite croissante et majorée : $\exists M \in \mathbb{R}, \forall n \in \mathbb{N}, u_n \leq M$.

Notons $E = \{u_n / n \in \mathbb{N}\}$ l'ensemble des termes de la suite u . Cet ensemble est une partie non vide (par définition) et majorée (par hypothèse) de $\mathbb{R}$. Il s'ensuit que E admet une borne supérieure, d'après la propriété du même nom (qui est valide dans $\mathbb{R}$).

Notons : $S = \sup E$, et considérons un réel $\varepsilon > 0$.

D'après la propriété caractérisant la borne supérieure, il existe un entier naturel n_0 tel que : $S - \varepsilon < u_{n_0} \leq S$.

Or, u étant croissante (par hypothèse) et majorée par S , on en déduit que :

$$\forall n \geq n_0, S - \varepsilon < u_{n_0} \leq u_n \leq S, \text{ d'où : } \forall n \geq n_0, S - \varepsilon < u_n \leq S \text{ ce qui implique : } \forall n \geq n_0, |u_n - S| < \varepsilon$$

En résumé, on a prouvé que : $\forall \varepsilon > 0, \exists n_0 \in \mathbb{N}, (n \geq n_0 \implies |u_n - S| < \varepsilon)$.

Donc la suite u est convergente et a pour limite S .

Conclusion. Si u est croissante et majorée, alors u est convergente.

2. Plus tard dans l'année, nous traduirons cette relation en disant que $\ln(n)$ est un *équivalent* de H_n au voisinage de $+\infty$.

Dans l'autre situation (celle où u est décroissante et minorée), il suffit d'observer que la suite $(-u)$ est croissante et majorée pour se ramener au cas précédent, et conclure en observant que $(-u)$ converge si et seulement si u converge. ☕

Exemple. Pour tout $n \in \mathbb{N}^*$, posons $S_n = \sum_{k=1}^n \frac{1}{k^2}$. La suite (S_n) est clairement croissante.

Par ailleurs : $\forall k \in \mathbb{N} \setminus \{0, 1\}, \frac{1}{k-1} - \frac{1}{k} = \frac{1}{k^2 - k} \geq \frac{1}{k^2}$.

Il s'ensuit que pour tout entier naturel n non nul, on a :

$$\sum_{k=1}^n \frac{1}{k^2} = 1 + \sum_{k=2}^n \frac{1}{k^2} \leq 1 + \sum_{k=2}^n \frac{1}{k-1} - \frac{1}{k} \text{ d'où : } \sum_{k=1}^n \frac{1}{k^2} \leq 1 + 1 - \frac{1}{n}$$

On en déduit que : $\forall n \in \mathbb{N}^*, 1 \leq \sum_{k=1}^n \frac{1}{k^2} \leq 2 - \frac{1}{n}$ donc : $\forall n \in \mathbb{N}^*, 1 \leq \sum_{k=1}^n \frac{1}{k^2} \leq 2$.

Conclusion : (S_n) étant croissante et majorée, elle converge. ³

DÉFINITION 11.4 - Deux suites u et v sont **adjacentes** si :

- u et v sont monotones, de monotonies opposées
- $\lim(v - u) = 0$

Exemple : les suites de termes généraux $u_n = 1 - 1/n$ et $v_n = 1 + e^{-n}$ sont adjacentes.

THÉORÈME 11.4 - (Suites adjacentes). Deux suites adjacentes convergent et ont la même limite.

PREUVE. Soient u et v deux suites adjacentes. SNALG, on peut supposer u croissante et v décroissante. En outre, l'hypothèse "adjacentes" impose que : $\lim(v - u) = 0$.

➤ **1ère étape** — Montrons que : $\forall n \in \mathbb{N}, u_n \leq v_n$.

Par l'absurde, supposons qu'il existe un entier n_0 tel que : $u_{n_0} > v_{n_0}$. Alors, pour tout entier n supérieur ou égal à n_0 , on a : $u_n \geq u_{n_0} > v_{n_0} \geq v_n$ (puisque u croît et que v décroît). Il s'ensuit que :

$$(n \geq n_0) \implies \left(v_n - u_n > \underbrace{v_{n_0} - u_{n_0}}_{\text{strict. pos.}} \right)$$

Cette assertion est clairement incompatible avec l'hypothèse $\lim(v - u) = 0$; contradiction. L'hypothèse fait quelques lignes plus haut est donc fausse, et on en déduit que : $\forall n \in \mathbb{N}, u_n \leq v_n$ (♠).

➤ **2ème étape** — Montrons que u et v convergent.

La suite v étant décroissante, elle est majorée par v_0 . Explicitement : $\forall n \in \mathbb{N}, v_n \leq v_0$ (♣).

On déduit de (♠) et (♣) que : $\forall n \in \mathbb{N}, u_n \leq v_0$. Ainsi u est majorée (par v_0). Comme en outre elle est croissante, le théorème de la limite monotone permet d'affirmer qu'elle converge.

Symétriquement, la suite u étant croissante, elle est minorée par u_0 . Explicitement : $\forall n \in \mathbb{N}, u_n \geq u_0$ (♡).

3. Et $\lim_{n \rightarrow +\infty} \sum_{k=1}^n \frac{1}{k^2} = \frac{\pi^2}{6}$.

On déduit de (♠) et (♥) que : $\forall n \in \mathbb{N}, u_0 \leq v_n$. Ainsi v est minorée (par u_0). Comme en outre elle est décroissante, le théorème de la limite monotone permet d'affirmer qu'elle converge.

Les suites u et v convergent.

► **3ème étape — Concluons.**

Puisque u et v sont convergentes, il est légitime d'utiliser la formule : $\lim (v - u) = \lim v - \lim u$. Or $\lim (v - u) = 0$ par hypothèse. D'où : $\lim u = \lim v$.

Conclusion : u et v convergent et $\lim u = \lim v$. ☕

Exemple. Les suites de termes généraux $u_n = \sum_{k=n}^{2n} \frac{1}{k}$ et $\forall n \in \mathbb{N}^*, v_n = u_n - \frac{1}{n}$ convergent et ont la même limite (exercice).

11.3.3 Applications des suites adjacentes

11.3.3.1 Irrationalité de e

On pose pour tout entier naturel n :

$$u_n = \sum_{k=0}^n \frac{1}{k!} \quad \text{et} \quad v_n = u_n + \frac{1}{n \times n!}$$

Il est clair que la suite de terme général $u_n = \sum_{k=0}^n \frac{1}{k!}$ est (strictement) croissante.

Par ailleurs, pour tout entier naturel n , on a :

$$\begin{aligned} v_{n+1} - v_n &= u_{n+1} - u_n + \frac{1}{(n+1) \times (n+1)!} - \frac{1}{n \times n!} = \frac{1}{(n+1)!} + \frac{1}{(n+1) \times (n+1)!} - \frac{1}{n \times n!} \\ &= \frac{n(n+1) + n - (n+1)^2}{n(n+1) \times (n+1)!} = -\frac{1}{n(n+1) \times (n+1)!} \end{aligned}$$

On en déduit que : $\forall n \in \mathbb{N}, v_{n+1} - v_n < 0$. La suite (v_n) est donc (strictement) décroissante.

Enfin : $\forall n \in \mathbb{N}, v_n - u_n = \frac{1}{n \times n!}$. D'où : $\lim_{n \rightarrow +\infty} (v_n - u_n) = 0$.

Il résulte de ce qui précède que les suites (u_n) et (v_n) sont adjacentes. D'après le théorème du même nom, elles convergent donc vers une limite commune.

De fait, on a déjà établi (cf chapitre portant sur le calcul intégral) que la limite de la suite u est e.

Montrons à présent l'irrationalité de e. D'après ce qui précède, la suite (u_n) est strictement croissante et converge vers e. D'où : $\forall n \in \mathbb{N}, u_n < e$.

D'après ce qui précède (bis!), la suite (v_n) est strictement décroissante et converge vers e. D'où : $\forall n \in \mathbb{N}, v_n > e$.

Ainsi : $\forall n \in \mathbb{N}, u_n < e < v_n$.

Raisonnons par l'absurde en supposant que e est rationnel. Il existe donc deux entiers naturels⁴ m et p (avec $p \neq 0$) tels que : $e = \frac{m}{p}$.

D'après ce qui précède, on a : $u_p < e < v_p$. En d'autres termes : $\sum_{k=0}^p \frac{1}{k!} < \frac{m}{p} < \sum_{k=0}^p \frac{1}{k!} + \frac{1}{p \times p!}$

On en déduit que : $p \times p! \times \sum_{k=0}^p \frac{1}{k!} < m \times p! < p \times p! \times \sum_{k=0}^p \frac{1}{k!} + 1$

Notons $N_0 = p \times p! \times \sum_{k=0}^p \frac{1}{k!}$. On a : $N_0 = p \times \sum_{k=0}^p \frac{p!}{k!}$. Pour tout entier k compris entre 0 et p , l'entier $p!$ est un multiple de $k!$, donc le rapport $\frac{p!}{k!}$ est un entier naturel. D'où N_0 est un entier naturel.

On a donc obtenu l'encadrement : $N_0 < m \times p! < N_0 + 1$. Il en résulte que $m \times p!$ est un entier naturel strictement compris entre deux entiers naturels consécutifs, ce qui est absurde.

Conclusion. e est irrationnel ($e \in \mathbb{R} \setminus \mathbb{Q}$).

11.3.3.2 Approximations décimales d'un réel, et densité de $\mathbb{D}$ dans $\mathbb{R}$

On rappelle que pour tout réel x , on définit la suite u (*resp.* v) des approximations décimales par défaut (*resp.* par excès) de x en posant :

$$\forall n \in \mathbb{N}, \quad u_n = \frac{\lfloor 10^n x \rfloor}{10^n} \quad (\text{resp. } \forall n \in \mathbb{N}, \quad v_n = \frac{\lfloor 10^n x \rfloor + 1}{10^n})$$

On a établi dans le chapitre précédent que ces deux suites convergent (car elles sont adjacentes) et ont x pour limite commune.

La conséquence remarquable de ce résultat est que $\mathbb{D}$ est dense dans $\mathbb{R}$.

11.3.3.3 Intervalles (segments) emboîtés

THÉORÈME 11.5 - (Intervalles (segments) emboîtés). Soit $(I_n)_n$ une suite d'intervalles non vides de $\mathbb{R}$ tels que :

- pour tout $n \in \mathbb{N}$, $I_n = [a_n, b_n]$ (avec $a_n < b_n$ réels) ;
- pour tout $n \in \mathbb{N}$, $I_{n+1} \subset I_n$;
- $\lim_{n \rightarrow +\infty} (b_n - a_n) = 0$

Alors : $\exists ! x \in \mathbb{R}, \forall n \in \mathbb{N}, x \in I_n$ (c-à-d : $\exists ! x \in \mathbb{R}, \bigcap_{n=0}^{+\infty} I_n = \{x\}$)

PREUVE. Sous les hypothèses du théorème, les suites (a_n) et (b_n) sont respectivement croissante et décroissante (d'après $I_{n+1} \subset I_n$), et $\lim_{n \rightarrow +\infty} (b_n - a_n) = 0$.

4. On sait en effet que $e \geq 0$, en tant que limite de suites à termes positifs.

Les suites (a_n) et (b_n) sont donc adjacentes, et par suite convergentes. Notons x leur limite commune. On a pour tout entier naturel n : $a_n \leq x \leq b_n$ puisque (a_n) est croissante (*resp.* (b_n) est décroissante) et de limite x . Par conséquent :

$$\forall n \in \mathbb{N}, x \in I_n \quad \text{càd} : \{x\} \subset \bigcap_{n=0}^{+\infty} I_n$$

La réciproque est aisée : si y est un réel distinct de x , alors on a $y > x$ ou $y < x$. Supposons que : $y > x$ (l'autre cas se traitant de manière analogue). Puisque b_n converge vers x , il existe un entier n_0 tel que : $x < b_{n_0} < y$.

Pour cet entier n_0 on a donc : $a_{n_0} < x < b_{n_0} < y$. Il s'ensuit que $y \notin I_{n_0}$. Par suite : $y \notin \bigcap_{n=0}^{+\infty} I_n$.

On vient donc d'établir que $x \in \bigcap_{n=0}^{+\infty} I_n$, et que l'intersection $\bigcap_{n=0}^{+\infty} I_n$ ne peut contenir d'autre réel que x .

Donc : $\bigcap_{n=0}^{+\infty} I_n = \{x\}$. ☕

Remarque A son tour, cette application (du théorème des suites adjacentes) en aura d'autres : un peu plus tard dans ce chapitre, le théorème de Bolzano-Weierstrass (qui affirme que de toute suite bornée on peut extraire une sous-suite convergente). Puis, lors de l'étude de la continuité, le théorème des bornes atteintes (qui affirme que toute fonction continue sur un segment est bornée et atteint ses bornes).

11.4 Suites extraites

DÉFINITION 11.5 - Soit $u = (u_n)_n$ une suite. Une suite $(v_n)_n$ est une **suite extraite de u** s'il existe une application $\varphi : \mathbb{N} \rightarrow \mathbb{N}$ strictement croissante telle que : $\forall n \in \mathbb{N}, v_n = u_{\varphi(n)}$.

Exemples : lorsque $\varphi(n) = 2n$, la suite extraite est celle des termes pairs. Lorsque $\varphi(n) = 2n + 1$, la suite extraite est celle des termes impairs. Lorsque $\varphi(n) = n + 1$, la suite extraite est la même que u , en ayant "oublié" le premier terme u_0 .

LEMME 11.6 - Si $\varphi : \mathbb{N} \rightarrow \mathbb{N}$ est strictement croissante, alors : $\forall n \in \mathbb{N}, \varphi(n) \geq n$.

PREUVE. Notons $P(n)$ l'assertion : $\varphi(n) \geq n$, et prouvons par récurrence sur $\mathbb{N}$ que l'assertion $P(n)$ est vraie pour tout $n \in \mathbb{N}$.

La véracité de $P(0)$ provient de l'ébouriffante observation qu'un entier naturel est positif ou nul...

Cette remarquable observation faite, héritons : on suppose $P(n)$ vraie pour un certain entier naturel n . Alors on a : $\varphi(n+1) > \varphi(n) \geq n$ d'où $\varphi(n+1) \geq n+1$.

Ce qui établit l'hérédité de $P(n)$ et achève cette récurrence. ☕

THÉORÈME 11.6 - (Suites extraites). Si u converge vers ℓ , toute suite extraite de u converge vers ℓ .

PREUVE. Soit (u_n) une suite convergente de limite $\ell \in \mathbb{K}$;⁵ et soit (v_n) une suite extraite de n ; il existe donc une application $\varphi : \mathbb{N} \rightarrow \mathbb{N}$ strictement croissante telle que $\forall n \in \mathbb{N}, v_n = u_{\varphi(n)}$.

Considérons un réel $\varepsilon > 0$.

Par hypothèse, il existe un entier n_0 tel que pour tout entier n on a : $(n \geq n_0) \implies (|u_n - \ell| < \varepsilon)$.

Or l'hypothèse faite sur φ implique que : $\forall n \in \mathbb{N}, \varphi(n) \geq n$.⁶

Il en résulte que : $(n \geq n_0) \implies (|u_{\varphi(n)} - \ell| < \varepsilon)$.

En recollant tous les morceaux du raisonnement précédent :

$$\forall \varepsilon > 0, \exists n_0 \in \mathbb{N}, \forall n \in \mathbb{N}, (n \geq n_0) \implies (|u_{\varphi(n)} - \ell| < \varepsilon).$$

Conclusion. La suite $(u_{\varphi(n)})$ est convergente et a pour limite $\ell = \lim u$. En d'autres termes, toute suite extraite de u converge vers ℓ . ☕

Exemple 1. Les suites de termes généraux $(-1)^n$ et $\cos(n)$ n'ont pas de limite.

Exemple 2. La propriété concernant la suite de terme général $\cos(n)$ peut être généralisée comme suit :

Exercice classique. La suite de terme général $\cos(n\theta)$ converge si et seulement si $\theta = 0 \pmod{2\pi}$.

Posons $u_n = \cos(n\theta)$. Puisque $(u_n)_n$ est bornée (entre -1 et 1), elle ne tend pas vers $+\infty$ ou $-\infty$.

Supposons que $(u_n)_n$ admette une limite finie ℓ .

➤ Considérons la suite (u_{2n}) (suite extraite des termes pairs). Puisque (u_n) converge vers ℓ , alors d'après la propriété des suites extraites, (u_{2n}) converge elle aussi vers ℓ .

En outre : $\forall n \in \mathbb{N}, u_{2n} = \cos(2n\theta) = 2\cos^2(n\theta) - 1 = 2u_n^2 - 1$. En passant à la limite lorsque n tend vers $+\infty$ dans cette relation, on obtient :

$$\ell = 2\ell^2 - 1 \iff 2\ell^2 - \ell - 1 = 0$$

Cette équation du second degré admet 1 comme racine évidente, et puisque le produit des racines est égal à " c/a ", on en déduit que la seconde racine est $-1/2$. D'où : $(\ell = 1) \quad \vee \quad \left(\ell = -\frac{1}{2}\right) \quad (\spadesuit)$.

➤ Par ailleurs, pour tout entier naturel $n \geq 1$ on a :

$$u_{n+1} + u_{n-1} = \cos((n+1)\theta) + \cos((n-1)\theta) = 2\cos(n\theta)\cos(\theta) = 2u_n\cos(\theta)$$

5. Où $\mathbb{K}$ désigne $\mathbb{R}$ ou $\mathbb{C}$. Dans le cours, seule la preuve dans le cas des suites réelles a été faite. Or cette preuve s'adapte sans aucun obstacle au cas des suites complexes; l'écriture de la démonstration est rigoureusement la même, seule la lecture des barres verticales sera différente (valeur absolue dans $\mathbb{R}$ et module dans $\mathbb{C}$).

6. Voir lemme 11.6 énoncé un peu plus haut.

En passant à la limite lorsque n tend vers $+\infty$ dans cette relation, on obtient :

$$2\ell = 2\ell \cos(\theta) \iff \ell (\cos(\theta) - 1) = 0.$$

➤ Si $\cos(\theta) \neq 1$, on en déduit que : $\boxed{\ell = 0 \quad (\clubsuit)}$. Les assertions $(\spadesuit)$ et $(\clubsuit)$ étant clairement incompatibles, la suite $(u_n)_n$ n'a donc pas de limite finie (dans ce cas).

➤ En revanche, si $\cos(\theta) = 1$, c'est-à-dire si $\theta = 0 \pmod{2\pi}$, alors la suite $(u_n)_n$ est convergente et de limite 1 (puisque'elle est alors constante égale à 1).

Conclusion. La suite de terme général $\cos(n\theta)$ admet une limite (égale à 1) SSI $\theta = 0 \pmod{2\pi}$.

THÉORÈME 11.7 - (Suites extraites, “version infinie”). Si u converge vers $+\infty$ (resp. $-\infty$), toute suite extraite de u converge vers $+\infty$ (resp. $-\infty$).

PREUVE. Soit (u_n) une suite divergeant vers $+\infty$, et soit (v_n) une suite extraite de n ; il existe donc une application $\varphi : \mathbb{N} \rightarrow \mathbb{N}$ strictement croissante telle que $\forall n \in \mathbb{N}, v_n = u_{\varphi(n)}$.

Considérons un réel M .

Par hypothèse, il existe un entier n_0 tel que pour tout entier n on a : $(n \geq n_0) \implies (u_n \geq M)$.

Or l'hypothèse faite sur φ implique que : $\forall n \in \mathbb{N}, \varphi(n) \geq n$.⁷

Il en résulte que : $(n \geq n_0) \implies (u_{\varphi(n)} \geq M)$.

En recollant tous les morceaux du raisonnement précédent :

$$\forall M \in \mathbb{R}, \exists n_0 \in \mathbb{N}, \forall n \in \mathbb{N}, (n \geq n_0) \implies (u_{\varphi(n)} \geq M).$$

Conclusion. La suite $(u_{\varphi(n)})$ a pour limite $+\infty$. En d'autres termes, toute suite extraite de u converge vers $+\infty$.

Le cas “ $-\infty$ ” est une conséquence immédiate du raisonnement précédent. ☕

L'énoncé ci-dessous fournit une forme de “réciproque” au théorème des suites extraites.

PROPRIÉTÉ 11.6 - Soit u une suite réelle, et soit ℓ un réel.

$$(u_n) \text{ converge vers } \ell \text{ SSI } (u_{2n}) \text{ et } (u_{2n+1}) \text{ convergent vers } \ell.$$

PREUVE. Supposons que (u_{2n}) et (u_{2n+1}) convergent vers ℓ .

Soit $\varepsilon > 0$.

Puisque (u_{2n}) converge vers ℓ : $\exists n_0 \in \mathbb{N}, \forall n \in \mathbb{N}, [n \geq n_0] \implies [|u_{2n} - \ell| < \varepsilon] \quad (\spadesuit)$

Puisque (u_{2n+1}) converge vers ℓ : $\exists n_1 \in \mathbb{N}, \forall n \in \mathbb{N}, [n \geq n_1] \implies [|u_{2n+1} - \ell| < \varepsilon] \quad (\clubsuit)$

Posons $n_2 = \max(2n_0, 2n_1 + 1)$. Pour tout entier naturel $n \geq n_2$, on a :

➔ si n est pair : il existe $p \in \mathbb{N}$ tel que $n = 2p$, avec $p \geq n_0$, et on a donc $|u_n - \ell| < \varepsilon$ d'après $(\spadesuit)$;

7. Voir lemme 11.6 énoncé un peu plus haut.

→ si n est impair : il existe $p \in \mathbb{N}$ tel que $n = 2p + 1$, avec $p \geq n_1$, et on a donc $|u_n - \ell| < \varepsilon$ d'après ($\clubsuit$).

Dans tous les cas, on a établi que :

$$\forall \varepsilon > 0, \exists n_2 \in \mathbb{N}, \forall n \in \mathbb{N}, [n \geq n_2] \implies [|u_n - \ell| < \varepsilon]$$

Ce qui signifie que u converge vers ℓ .

Finalement : $[(u_{2n}) \text{ et } (u_{2n+1}) \text{ convergent vers } \ell] \implies [(u_n) \text{ converge vers } \ell]$

La réciproque est une application immédiate du théorème 11.6 des suites extraites. ☕

Comme exemple d'application de cette propriété, on prend un peu d'avance sur la suite des événements, en établissant un résultat relatif aux *séries numériques*, qui seront étudiées au second semestre.

THÉORÈME 11.8 - (Critère spécial des séries alternées).

Soit (a_n) une suite réelle positive et décroissante, telle que : $\lim_{n \rightarrow +\infty} a_n = 0$.

La suite de terme général $S_n = \sum_{k=0}^n (-1)^k a_k$ est convergente.

PREUVE. Considérons les suites (S_{2n}) et (S_{2n+1}) ; ce sont les sommes partielles de rangs pairs et impairs respectivement.

Pour tout entier naturel n , on a :

- $S_{2n+2} - S_{2n} = a_{n+2} - a_{n+1} \leq 0$ (puisque (a_n) décroît). Ainsi (S_{2n}) est décroissante.
- $S_{2n+3} - S_{2n+1} = a_{n+2} - a_{n+3} \geq 0$ (puisque (a_n) décroît). Ainsi (S_{2n+1}) est croissante.
- $S_{2n+1} - S_{2n} = -a_{2n+1}$. D'où : $\lim_{n \rightarrow +\infty} S_{2n+1} - S_{2n} = 0$ (puisque $\lim_{n \rightarrow +\infty} a_n = 0$).

On en déduit que les suites (S_{2n}) et (S_{2n+1}) sont adjacentes. A ce titre, elles convergent et ont une limite commune, disons ℓ .

Il s'ensuit que la suite (S_n) converge vers ℓ , ce qu'il fallait démontrer. ☕

Application. La suite S de terme général

$$S_n = \sum_{k=0}^n \frac{(-1)^k}{k+1} \text{ est convergente.}^8$$

Remarquons encore que la propriété 11.6 est encore valable dans sa version "infinie". Explicitement :

PROPRIÉTÉ 11.7 - Soit u une suite réelle convergeant vers $\pm\infty$.

(u_n) converge vers $+\infty$ (resp. $-\infty$) SSI (u_{2n}) et (u_{2n+1}) convergent vers $+\infty$ (resp. $-\infty$)

PREUVE. Se déduit mutatis mutandis de la preuve de la propriété 11.6. ☕

8. En outre : $\lim_{n \rightarrow +\infty} \sum_{k=0}^n \frac{(-1)^k}{k+1} = \ln(2)$.

L'énoncé ci-dessous, clou du spectacle sur les suites extraites, peut être vu comme une application (un corollaire) du théorème des intervalles emboîtés.

THÉORÈME 11.9 - (Bolzano-Weierstrass). De toute suite bornée on peut extraire une sous-suite convergente.

Remarque : si l'énoncé est clair pour la suite de terme général $(-1)^n$, elle l'est nettement moins pour celle de terme général $\cos(n)$. Pour cette dernière, on pourra d'ailleurs extraire des sous-suites convergeant vers un réel $x \in [-1, 1]$ arbitraire (ce qui va encore plus loin que l'énoncé de Bolzano-Weierstrass).

PREUVE. Soit u une suite réelle bornée ; il existe deux réels m et M tels que : $\forall n \in \mathbb{N}, m \leq u_n \leq M$.

Appelons $c = \frac{m+M}{2}$ l'abscisse du milieu du segment $[m, M]$. On pose alors : $a_0 = m$, $c_0 = c$, $b_0 = M$ et $\varphi(0) = 0$.

Au moins l'une des deux moitiés de l'intervalle $[m, M]$ contient une infinité de termes.

- Si $[a_0, c_0]$ contient une infinité de termes, on pose : $a_1 = a_0$, $b_1 = c_0$ et $\varphi(1)$ le plus petit rang $> \varphi(0)$ d'un terme appartenant à $[a_1, b_1]$;
- Sinon, $[c_0, b_0]$ contient une infinité de termes, et on pose alors : $a_1 = c_0$, $b_1 = b_0$ et $\varphi(1)$ le plus petit rang $> \varphi(0)$ d'un terme appartenant à $[a_1, b_1]$.

En itérant ce processus, on construit une suite $(u_{\varphi(n)})$ (extraite de u), et deux suites (a_n) et (b_n) satisfaisant les conditions suivantes :

- $\forall n \in \mathbb{N}, u_{\varphi(n)} \in [a_n, b_n]$ ($\spadesuit$) ;
- $\forall n \in \mathbb{N}, [a_{n+1}, b_{n+1}] \subset [a_n, b_n]$ ($\clubsuit$) ;
- $\forall n \in \mathbb{N}, b_n - a_n = \frac{b_0 - a_0}{2^n}$ ($\heartsuit$).

D'après ($\heartsuit$) et ($\clubsuit$), la suite d'intervalles $([a_n, b_n])$ vérifie les hypothèses du théorème des segments emboîtés, et on peut donc conclure en particulier que les suites (a_n) et (b_n) sont adjacentes et convergent vers une limite commune, que nous noterons ℓ .

Or d'après ($\spadesuit$) : $\forall n \in \mathbb{N}, a_n \leq u_{\varphi(n)} \leq b_n$. Le théorème d'encadrement permet alors de conclure que la suite $(u_{\varphi(n)})$ converge (et a pour limite ℓ).

Conclusion : de toute suite réelle bornée on peut extraire une sous-suite convergente.



Applications (par anticipation) : toute fonction à valeurs réelles continue sur un intervalle fermé borné est bornée et atteint ses bornes (*théorème des bornes atteintes*). Avec des quantificateurs :

$$\forall f \in \mathcal{C}^0([a, b], \mathbb{R}), \exists (c, d) \in [a, b]^2, \forall x \in [a, b], f(c) \leq f(x) \leq f(d)$$

Une autre application est le *théorème de Heine*, qui assure que toute fonction continue sur un segment est uniformément continue (voir chapitre portant sur l'intégration).

11.5 Suites complexes

DÉFINITION 11.6 - Une suite complexe u est **bornée** si :

$$\exists M \in \mathbb{R}_+, \forall n \in \mathbb{N}, |u_n| \leq M$$

Exemple. Pour tout complexe q tel que $|q| \leq 1$, la suite de terme général q^n est bornée.

DÉFINITION 11.7 - Une suite complexe u **converge vers** $\ell \in \mathbb{C}$ si :

$$\forall \varepsilon > 0, \exists n_0 \in \mathbb{N}, \forall n \in \mathbb{N}, (n \geq n_0 \implies |u_n - \ell| < \varepsilon)$$

Exemple. La suite de terme général $u_n = \frac{e^{in\theta}}{2^n}$ converge vers 0, pour tout réel θ .

Remarque : la définition de convergence est strictement la même que pour les suites réelles. Il suffit de lire autrement les barres verticales (“module” au lieu de “valeur absolue”). Cette remarque faite, on démontre l’énoncé ci-dessous en “copiant-collant” la preuve dans le cas réel.

PROPRIÉTÉ 11.8 - Toute suite complexe convergente est bornée.

PREUVE. Voir preuve de la propriété 11.2, page 244. ☕

Arrivé à ce point des définitions, l’idée naturelle serait de reprendre les propriétés des suites réelles, et de les étendre aux suites complexes. Malheureusement, toutes les propriétés faisant intervenir de près ou de loin la relation d’ordre dans $\mathbb{R}$ (plus spécifiquement, tous les énoncés utilisant la propriété de la borne supérieure) ne pourront avoir d’analogue dans $\mathbb{C}$.

Explicitement, les théorèmes de la limite monotone, d’encadrement, des suites adjacentes ne s’étendent pas aux suites complexes.

On dispose néanmoins d’une caractérisation de la convergence qui permet de se ramener au cas réel par le biais de l’énoncé ci-dessous :

THÉORÈME 11.10 - (“pont $\mathbb{R} \longleftrightarrow \mathbb{C}$ ”). Soient u une suite complexe, et ℓ un nombre complexe.

$$u \text{ converge vers } \ell \text{ SSI } \operatorname{Re} u \text{ converge vers } \operatorname{Re} \ell \text{ et } \operatorname{Im} u \text{ converge vers } \operatorname{Im} \ell$$

PREUVE. ➤ Sens direct ($\implies$). Supposons que u converge vers ℓ .

Soit $\varepsilon > 0$. Alors il existe un entier naturel n_0 tel que : $\forall n \in \mathbb{N}, n \geq n_0 \implies |u_n - \ell| < \varepsilon$.

Ceci implique : $\forall n \in \mathbb{N}, n \geq n_0 \implies |\operatorname{Re}(u_n) - \operatorname{Re} \ell| < \varepsilon$ et $\forall n \in \mathbb{N}, n \geq n_0 \implies |\operatorname{Im}(u_n) - \operatorname{Im} \ell| < \varepsilon$.⁹

D’où : [les suites $(\operatorname{Re}(u_n))$ et $(\operatorname{Im}(u_n))$ convergent (vers $\operatorname{Re} \ell$ et $\operatorname{Im} \ell$ respectivement)].

9. Puisque pour tout nombre complexe z , on a : $|\operatorname{Re} z| \leq |z|$ et $|\operatorname{Im} z| \leq |z|$.

➤ Réciproque ($\Leftarrow$). Supposons que $(\operatorname{Re}(u_n))$ et $(\operatorname{Im}(u_n))$ convergent (vers $\operatorname{Re} \ell$ et $\operatorname{Im} \ell$ respectivement).

Soit $\varepsilon > 0$. Alors :

➤ il existe un entier naturel n_0 tel que : $\forall n \in \mathbb{N}, n \geq n_0 \implies |\operatorname{Re}(u_n) - \operatorname{Re} \ell| < \frac{\varepsilon}{2}$;

➤ il existe un entier naturel n_1 tel que : $\forall n \in \mathbb{N}, n \geq n_1 \implies |\operatorname{Im}(u_n) - \operatorname{Im} \ell| < \frac{\varepsilon}{2}$.

Posons : $n_2 = \max(n_0, n_1)$. Alors : $\forall n \in \mathbb{N}, n \geq n_2 \implies |u_n - \ell| \leq |\operatorname{Re}(u_n) - \operatorname{Re} \ell| + |\operatorname{Im}(u_n) - \operatorname{Im} \ell| < \varepsilon$ (la première inégalité provenant de l'inégalité triangulaire).

On en déduit que : $\forall \varepsilon > 0, \exists n_2 \in \mathbb{N}, \forall n \in \mathbb{N}, n \geq n_2 \implies |u_n - \ell| < \varepsilon$. Donc : la suite u converge vers ℓ .

Conclusion : u converge vers ℓ SSI $\operatorname{Re}(u)$ converge vers $\operatorname{Re}(\ell)$ et $\operatorname{Im}(u)$ converge vers $\operatorname{Im}(\ell)$.



COROLLAIRE 11.2 - La suite de terme général $e^{in\theta}$ converge SSI $\theta = 0 \pmod{2\pi}$.

PREUVE. Soit q un nombre complexe.

➤ Si $q = 0$: la propriété est évidente.

➤ Si $q \neq 0$: pour tout entier naturel n , on a : $|q^n| = |q|^n = e^{n \ln |q|}$.

➤ Si $|q| < 1$: alors $\ln |q| < 0$ donc $\lim_{n \rightarrow +\infty} n \ln |q| = -\infty$ donc $\lim_{n \rightarrow +\infty} |q^n| = 0$. On en déduit alors que :

Si $|q| < 1$, alors $\lim_{n \rightarrow +\infty} q^n = 0$

➤ Si $|q| > 1$: alors $\ln |q| > 0$ donc $\lim_{n \rightarrow +\infty} n \ln |q| = +\infty$ donc $\lim_{n \rightarrow +\infty} |q^n| = +\infty$. On en déduit alors que :

Si $|q| > 1$, alors $(q^n)_n$ diverge (est non bornée).

➤ Si $|q| = 1$: alors il existe un réel θ tel que $q = e^{in\theta}$. Or la suite $(e^{in\theta})$ converge si et seulement si les suites $(\cos(n\theta))$ et $(\sin(n\theta))$ convergent, ce qui ne se produit que lorsque $\theta = 0 \pmod{2\pi}$. Et dans ce cas, on a bien entendu $q = 1$.

Si $|q| = 1$, alors $(q^n)_n$ converge SSI $q = 1$.

Conclusion : la suite de terme général q^n converge SSI $|q| < 1$ ou $q = 1$.



Une autre application : $\mathbb{Q} + i\mathbb{Q} = \{r_1 + ir_2 / r_1 \in \mathbb{Q}, r_2 \in \mathbb{Q}\}$ est dense dans $\mathbb{C}$, dans le sens où tout nombre complexe est limite d'une suite d'éléments de $\mathbb{Q} + i\mathbb{Q}$. De la même façon, $\mathbb{D} + i\mathbb{D}$ est dense dans $\mathbb{C}$. Et $\mathbb{D} + i\mathbb{Q}$... et $\mathbb{Q} + i\mathbb{D}$...

THÉORÈME 11.11 - (Suites extraites). Si u converge vers $\ell \in \mathbb{C}$, toute suite extraite de u converge vers ℓ .

PREUVE. Voir preuve du théorème 11.6, page 255.



PROPRIÉTÉ 11.9 - Soit u une suite complexe, et soit ℓ un complexe.

(u_n) converge vers ℓ SSI (u_{2n}) et (u_{2n+1}) convergent vers ℓ .

PREUVE. Voir preuve de la propriété 11.6, page 256. ☕

THÉORÈME 11.12 - (**Bolzano-Weierstrass**). De toute suite complexe bornée on peut extraire une sous-suite convergente.

PREUVE. Reprendre la preuve du théorème de Bolzano-Weierstrass dans le cas réel (théorème 11.9, page 258), en remplaçant les intervalles emboîtés par des rectangles emboîtés. ☕

11.6 Suites particulières

Sauf mention explicite du contraire, les suites de ce paragraphe sont complexes.

11.6.1 Suites arithmétiques

DÉFINITION 11.8 - Une suite u est **arithmétique** s'il existe un complexe r tel que $\forall n \in \mathbb{N}, u_{n+1} = u_n + r$.

Lorsque tel est le cas, le complexe r est appelé **raison** de la suite u .

PROPRIÉTÉ 11.10 - Soit u une suite arithmétique de raison r . Alors :

$$1/ \forall n \in \mathbb{N}, u_n = u_0 + nr$$

$$2/ \forall N \in \mathbb{N}, \sum_{k=0}^N u_k = (N+1) \frac{u_0 + u_N}{2}$$

PREUVE. Ces résultats ont déjà été établis : voir propriété ?? page ?? pour le premier point, et propriété 1.8 page 26 pour le second. ☕

Application : $1 + 2 + \dots + 100 = 5050$.

Remarque à un sou. Une suite arithmétique u de raison r converge SSI $r = 0$.

11.6.2 Suites géométriques

DÉFINITION 11.9 - Une suite u est **géométrique** s'il existe un complexe q tel que $\forall n \in \mathbb{N}, u_{n+1} = qu_n$.

Lorsque tel est le cas, le complexe q est appelé **raison** de la suite u .

Exemple. La suite de terme général $e^{in\theta}$ est géométrique de raison $e^{i\theta}$.

PROPRIÉTÉ 11.11 - Soit u une suite géométrique de raison q . Alors :

$$1/ \forall n \in \mathbb{N}, u_n = u_0 q^n$$

$$2/ \underline{\text{Si } q \neq 1} : \forall N \in \mathbb{N}, \sum_{k=0}^N u_k = u_0 \frac{1 - q^{N+1}}{1 - q}$$

PREUVE. On renvoie pour cela à la propriété 2.2 page 36. ☕

Applications. Calculs de $\sum_{k=0}^N e^{kx}$, $\sum_{k=0}^N \cos(kx)$, $\sum_{k=0}^N \sin(kx)$, $\sum_{k=0}^N \operatorname{ch}(kx)$, $\sum_{k=0}^N \operatorname{sh}(kx)$.

PROPRIÉTÉ 11.12 - Soit $q \in \mathbb{C}$. La suite $(q^n)_n$ converge SSI $|q| < 1$ ou $q = 1$.

Un peu plus dans le détail :

- si $|q| < 1$, alors q^n tend vers 0 ;
- si $q = 1$, alors q^n tend vers 1 ;
- si $|q| = 1$ et $q \neq 1$, alors $(q^n)_n$ n'a pas de limite ;
- si $|q| > 1$, alors $(q^n)_n$ est non bornée.

PREUVE. Voir preuve du corollaire 11.2 page 260. ☕

11.6.3 Suites arithmético-géométriques

DÉFINITION 11.10 - Une suite u est **arithmético-géométrique** (SAG) s'il existe deux complexes a et b tels que $\forall n \in \mathbb{N}, u_{n+1} = au_n + b$.

Remarque. Pour éviter de “retomber” sur des cas triviaux ou déjà connus, on pourra raisonnablement supposer dans la suite des évènements que $a \neq 0$ (auquel cas u est stationnaire, car constante au moins à partir de $n = 1$), $a \neq 1$ (u est alors arithmétique de raison b) et $b \neq 0$ (u est alors géométrique de raison a).

PROPRIÉTÉ 11.13 - Soient a et b deux complexes, et u une suite complexe telle que : $\forall n \in \mathbb{N}, u_{n+1} = au_n + b$.

On suppose $a \neq 1$, et on pose $\ell = b/(1 - a)$.

Alors la suite $(u - \ell)$ est géométrique de raison a .

PREUVE. Soient a et b deux complexes (avec $a \neq 1$), et $u \in \mathbb{C}^{\mathbb{N}}$ telle que : $\forall n \in \mathbb{N}, u_{n+1} = au_n + b$. On pose $\ell = \frac{b}{1 - a}$.¹⁰ On pose : $\forall n \in \mathbb{N}, v_n = u_n - \ell$.

$$\text{Pour tout } n \in \mathbb{N} \text{ on a : } v_{n+1} = u_{n+1} - \ell = au_n + b - \frac{b}{1 - a} = au_n + \frac{b - ab - b}{1 - a} = a \left(u_n - \frac{b}{1 - a} \right) = av_n.$$

10. Attention : à ce stade, on ne peut pas encore affirmer que ℓ est la limite de la suite u , puisqu'on ne sait pas encore si u converge ou non. On prendra donc garde à ne pas appeler ℓ la limite (tout au plus peut-on l'appeler “limite potentielle”, ou “meilleure candidate imaginable pour la limite de u ”, ou plus simplement... ℓ).

Ainsi : La suite v est donc géométrique de raison a .

D'où : $\forall n \in \mathbb{N}, v_n = v_0 \times a^n$ et donc : $\forall n \in \mathbb{N}, u_n = \ell + (u_0 - \ell) \times a^n$. ☕

COROLLAIRE 11.3 - On conserve les notations et hypothèses de la propriété précédente.

On suppose en outre : $u_0 \neq \ell$.

Alors : la suite u converge et a pour limite ℓ SSI $|a| < 1$.

PREUVE. La propriété précédente nous conduit à considérer deux cas.

1er cas : si $u_0 - \ell = 0$ ($\iff v_0 = 0$). Alors la suite u est constante (égale à ℓ). Elle converge alors évidemment vers ℓ , quelle que soit la valeur de a .

2nd cas : si $u_0 - \ell \neq 0$ ($\iff v_0 \neq 0$). Alors la suite de terme général $(u_0 - \ell) \times a^n$ converge SSI $|a| < 1$.¹¹

On en déduit que : u converge et a pour limite ℓ SSI $|a| < 1$. ☕

11.6.4 Suites récurrentes linéaires d'ordre 2

DÉFINITION 11.11 - Une suite u est **récurrente linéaire d'ordre 2**, si :

$$\exists (a, b) \in \mathbb{C} \times \mathbb{C}^*, \forall n \in \mathbb{N}, u_{n+2} = au_{n+1} + bu_n \quad (\star)$$

Remarque. Une suite récurrente linéaire d'ordre 2 (SRL2) est donc entièrement déterminée par les complexes a et b , et ses deux premiers termes (u_0 et u_1).

DÉFINITION 11.12 - Avec les notations de la définition ci-dessus, on appelle **équation caractéristique** associée à la relation $(\star)$ l'équation (EC) $X^2 - aX - b = 0$.

Ces définitions posées, l'idée est de décrire le terme général d'une SRL2, à l'instar de ce que nous avons fait pour décrire la solution d'une EDL2 homogène à coefficients constants. Deux observations sont très utiles pour y parvenir.

Remarque-clef 1 : soit $q \in \mathbb{C}^*$. La suite $(q^n)_n$ vérifie $(\star)$ SSI q est racine de (EC).

Remarque-clef 2 : l'ensemble E des suites complexes vérifiant la relation $(\star)$ est **stable par somme**, et **par multiplication par un scalaire**. Explicitement :

- Si u et v sont dans E , alors $(u + v) \in E$;
- Si u est dans E , alors pour tout $\lambda \in \mathbb{C}$ on a $(\lambda u) \in E$.

Ces deux propriétés peuvent être synthétisés en une seule, que voici :

$$\forall (u, v) \in E^2, \forall (\lambda, \mu) \in \mathbb{C}^2, (\lambda u + \mu v) \in E$$

On dit que E est **stable par combinaison linéaire** (ce qui justifie *a posteriori* la terminologie).

Ces remarques faites, l'idée du théorème ci-dessous est que l'ensemble des SRL2 vérifiant une relation de récurrence donnée peut être identifié au plan $\mathbb{C}^2$; la correspondance entre ces deux ensembles est la bijection qui à une suite u associe le couple (u_0, u_1) de ses deux premiers termes.

11. Voir corollaire 11.2 page 260, en observant que a est supposé différent de 1 dans la présente démonstration.

THÉORÈME 11.13 - (Terme général d'une SRL2, cas COMPLEXE). Soit $(a, b) \in \mathbb{C} \times \mathbb{C}^*$.

On note $(\star)$ l'assertion $\forall n \in \mathbb{N}, u_{n+2} = au_{n+1} + bu_n$; E l'ensemble des suites complexes vérifiant $(\star)$; (EC) l'équation $X^2 - aX - b = 0$ et Δ le discriminant de (EC) .

► Si $\Delta \neq 0$: (EC) a deux racines distinctes r_1 et r_2 , et $u \in E$ SSI :

$$\exists (C_1, C_2) \in \mathbb{C}^2, \quad \forall n \in \mathbb{N}, u_n = C_1 r_1^n + C_2 r_2^n$$

► Si $\Delta = 0$: (EC) a une racine double r_0 , et $u \in E$ SSI :

$$\exists (C_1, C_2) \in \mathbb{C}^2, \quad \forall n \in \mathbb{N}, u_n = (C_1 + nC_2) r_0^n$$

PREUVE. La preuve du théorème repose essentiellement sur deux lemmes.

LEMME 11.7 - Soient u et v deux suites de E , et soit (λ, μ) un couple de complexes. Il existe une unique suite w de E telle que :

$$w_0 = \lambda u_0 + \mu v_0 \quad \text{et} \quad w_1 = \lambda u_1 + \mu v_1$$

Explicitement : $w = \lambda u + \mu v$.

PREUVE. (du lemme). On commence par observer qu'il existe une unique suite w de E telle que $w_0 = \lambda u_0 + \mu v_0$ et $w_1 = \lambda u_1 + \mu v_1$, par définition d'une SRL2.

En outre, pour une telle suite, une récurrence (double) immédiate permet d'établir que :

$$\forall n \in \mathbb{N}, w_n = \lambda u_n + \mu v_n$$

Ce qui achève la preuve du lemme. ☕

LEMME 11.8 - Soient $\vec{u}(a, b)$ et $\vec{v}(c, d)$ deux vecteurs non colinéaires de $\mathbb{C}^2$, c'àd tels que $ad - bc \neq 0$.

Alors tout vecteur de $\mathbb{C}^2$ est combinaison linéaire de $\vec{u}$ et de $\vec{v}$.

Explicitement :

$$\forall \vec{w}(x, y) \in \mathbb{C}^2, \quad \vec{w} = \frac{dx - cy}{ad - bc} \vec{u} + \frac{ay - bx}{ad - bc} \vec{v}$$

PREUVE. (du lemme). Soit $\vec{w}(x, y) \in \mathbb{C}^2$. On a :

$$\begin{aligned} \frac{dx - cy}{ad - bc} \vec{u} + \frac{ay - bx}{ad - bc} \vec{v} &= \frac{dx - cy}{ad - bc} \begin{pmatrix} a \\ b \end{pmatrix} + \frac{ay - bx}{ad - bc} \begin{pmatrix} c \\ d \end{pmatrix} = \frac{1}{ad - bc} \begin{pmatrix} adx - acy + acy - bcx \\ bdx - bcy + ady - bdx \end{pmatrix} \\ &= \begin{pmatrix} x \\ y \end{pmatrix} \end{aligned}$$

Ce qui prouve le lemme. ☕

RETOUR À LA PREUVE DU THÉORÈME. On distingue deux cas, suivant que l'équation caractéristique possède une ou deux racines.

→ Premier cas : $\Delta \neq 0$. Alors l'équation caractéristique possède deux racines r_1 et r_2 . Les suites u et v de termes généraux respectifs r_1^n et r_2^n sont des éléments de E (vérification immédiate).

En outre, les vecteurs $\vec{u}(1, r_1)$ et $\vec{v}(1, r_2)$ sont non-colinéaires (puisque $r_1 \neq r_2$) (♠).

Soit w une suite quelconque de E . Notons $\vec{w}(w_0, w_1)$. En vertu de (♠) et du lemme 11.8 :

$$\exists (\lambda, \mu) \in \mathbb{C}^2, \vec{w} = \lambda \vec{u} + \mu \vec{v}$$

On en déduit, grâce au lemme 11.7, que : $w = \lambda u + \mu v$. Autrement écrit :

$$\forall n \in \mathbb{N}, w_n = \lambda u_n + \mu v_n \quad \text{càd : } \forall n \in \mathbb{N}, w_n = \lambda r_1^n + \mu r_2^n.$$

On a donc établi dans ce cas que : $[w \in E] \implies [\exists (\lambda, \mu) \in \mathbb{C}^2, \forall n \in \mathbb{N}, w_n = \lambda r_1^n + \mu r_2^n]$

La réciproque est une vérification sans aucune difficulté.

→ Second cas : $\Delta = 0$. Alors l'équation caractéristique possède une unique racine r_0 (non nulle car $b \neq 0$). Les suites u et v de termes généraux respectifs r_0^n et nr_0^n sont des éléments de E ; la vérification est immédiate pour la première. Pour la seconde, observons que :

$$\forall n \in \mathbb{N}, (n+2)r_0^{n+2} + a(n_0+1)r_0^{n+1} + bnr_0^n = r_0^n \left[\underbrace{(r_0^2 + ar_0 + b)}_{=0 \text{ car } r_0 \text{ racine de (EC)}} + r_0 \underbrace{(2r_0 + a)}_{=0 \text{ car } \Delta = 0} \right] = 0$$

En outre, les vecteurs $\vec{u}(1, r_0)$ et $\vec{v}(0, r_0)$ sont non-colinéaires (puisque $r_0 \neq 0$) (♣).

Soit w une suite quelconque de E . Notons $\vec{w}(w_0, w_1)$. En vertu de (♣) et du lemme 11.8 :

$$\exists (\lambda, \mu) \in \mathbb{C}^2, \vec{w} = \lambda \vec{u} + \mu \vec{v}$$

On en déduit, grâce au lemme 11.7, que : $w = \lambda u + \mu v$. Autrement écrit :

$$\forall n \in \mathbb{N}, w_n = \lambda u_n + \mu v_n \quad \text{càd : } \forall n \in \mathbb{N}, w_n = \lambda r_0^n + \mu nr_0^n.$$

On a donc établi dans ce cas que : $[w \in E] \implies [\exists (\lambda, \mu) \in \mathbb{C}^2, \forall n \in \mathbb{N}, w_n = \lambda r_0^n + \mu nr_0^n]$

La réciproque est une nouvelle vérification sans aucune difficulté, qui achève la preuve du théorème. ☕

THÉORÈME 11.14 - (Terme général d'une SRL2, cas REEL). Soit $(a, b) \in \mathbb{R} \times \mathbb{R}^*$. Pour le reste, on conserve les notations du théorème précédent.

➤ Si $\Delta > 0$: (EC) a deux racines réelles distinctes r_1 et r_2 , et $u \in E$ SSI :

$$\exists (C_1, C_2) \in \mathbb{R}^2, \quad \forall n \in \mathbb{N}, u_n = C_1 r_1^n + C_2 r_2^n$$

➤ Si $\Delta = 0$: (EC) a une racine réelle double r_0 , et $u \in E$ SSI :

$$\exists (C_1, C_2) \in \mathbb{R}^2, \quad \forall n \in \mathbb{N}, u_n = (C_1 + nC_2) r_0^n$$

➤ Si $\Delta < 0$: (E) a deux racines complexes conjuguées $re^{\pm i\theta}$, et $u \in E$ SSI :

$$\exists (C_1, C_2) \in \mathbb{R}^2, \quad \forall n \in \mathbb{N}, u_n = r^n (\lambda \cos(n\theta) + \mu \sin(n\theta))$$

PREUVE. Les cas $\Delta > 0$ et $\Delta = 0$ se déduisent des deux cas du théorème précédent.

Traisons le cas où $\Delta < 0$: l'équation caractéristique (E) a deux racines complexes conjuguées $re^{\pm i\theta}$.

Soit $u \in E$. Alors u est en particulier à valeurs complexes, et d'après le théorème précédent :

$$\exists (\lambda, \mu) \in \mathbb{C}^2, \forall n \in \mathbb{N}, u_n = \lambda r^n e^{ni\theta} + \mu r^n e^{-ni\theta}$$

$$\implies \exists (\lambda, \mu) \in \mathbb{C}^2, \forall n \in \mathbb{N}, u_n = r^n [(\lambda + \mu) \cos(n\theta) + i(\lambda - \mu) \sin(n\theta)]$$

Puisque u est une suite réelle, on a : u_0 et u_1 réels. D'où :

$$\begin{cases} u_0 \in \mathbb{R} \\ u_1 \in \mathbb{R} \end{cases} \iff \begin{cases} \lambda + \mu \in \mathbb{R} \\ (\lambda + \mu) \cos(\theta) + i(\lambda - \mu) \sin(\theta) \in \mathbb{R} \end{cases}$$

En observant que $\theta \neq 0[\pi]$ (sinon l'équation (E) posséderait une des racines réelles), on en déduit que :

$$(\lambda + \mu) \in \mathbb{R} \quad \text{et} \quad i(\lambda - \mu) \in \mathbb{R}$$

En posant $\alpha = \lambda + \mu$ et $\beta = i(\lambda - \mu)$, on a donc établi que :

$$\exists (\alpha, \beta) \in \mathbb{R}^2, \forall n \in \mathbb{N}, u_n = r^n [\alpha \cos(n\theta) + i\beta \sin(n\theta)]$$

En résumé, dans le cas où $\Delta < 0$, on a établi que :

$$(u \in E) \implies (\exists (\alpha, \beta) \in \mathbb{R}^2, \forall n \in \mathbb{N}, u_n = r^n [\alpha \cos(n\theta) + i\beta \sin(n\theta)])$$

La réciproque est une vérification amusante qui permet de réviser le formulaire de trigonométrie. ☕

Application. La suite de Fibonacci est définie par ses deux premiers termes $u_0 = 1, u_1 = 1$, et la relation de récurrence $\forall n \in \mathbb{N}, u_{n+2} = u_{n+1} + u_n$.

L'équation caractéristique associée à $(u_n)_n$ est $X^2 - X - 1 = 0$. Elle a deux racines réelles : $x_1 = \frac{1 + \sqrt{5}}{2}$ et $x_2 = \frac{1 - \sqrt{5}}{2}$.¹²

Conclusion intermédiaire : $\exists (\lambda, \mu) \in \mathbb{R}^2, \forall n \in \mathbb{N}, u_n = \lambda \left(\frac{1 + \sqrt{5}}{2} \right)^n + \mu \left(\frac{1 - \sqrt{5}}{2} \right)^n$.

$$\begin{aligned} \rhd \text{De plus : } \begin{cases} u_0 &= 0 \\ u_1 &= 1 \end{cases} &\iff \begin{cases} \lambda + \mu &= 0 \\ \lambda \frac{1 + \sqrt{5}}{2} + \mu \frac{1 - \sqrt{5}}{2} &= 1 \end{cases} \iff \begin{cases} \mu &= -\lambda \\ \lambda ((1 + \sqrt{5}) - (1 - \sqrt{5})) &= 2 \end{cases} \\ &\iff \begin{cases} \mu &= -\lambda \\ \sqrt{5} \lambda &= 1 \end{cases} \iff \begin{cases} \mu &= -\lambda \\ \lambda &= \frac{\sqrt{5}}{5} \end{cases} \iff \begin{cases} \mu &= -\frac{\sqrt{5}}{5} \\ \lambda &= \frac{\sqrt{5}}{5} \end{cases} \end{aligned}$$

^{12.} Le réel x_1 est le célèbre *nombre d'or*, souvent noté φ ; et $x_2 = -1/\varphi$ (puisque le produit $x_1 x_2$ vaut -1).

Donc : $\forall n \in \mathbb{N}, u_n = \frac{\sqrt{5}}{5} \left(\left(\frac{1+\sqrt{5}}{2} \right)^n - \left(\frac{1-\sqrt{5}}{2} \right)^n \right).$

Soit encore : $\forall n \in \mathbb{N}, u_n = \frac{\sqrt{5}}{2^n 5} \left[(1+\sqrt{5})^n - (1-\sqrt{5})^n \right]$

11.6.5 Suites récurrentes non-linéaires “ $u_{n+1} = f(u_n)$ ”

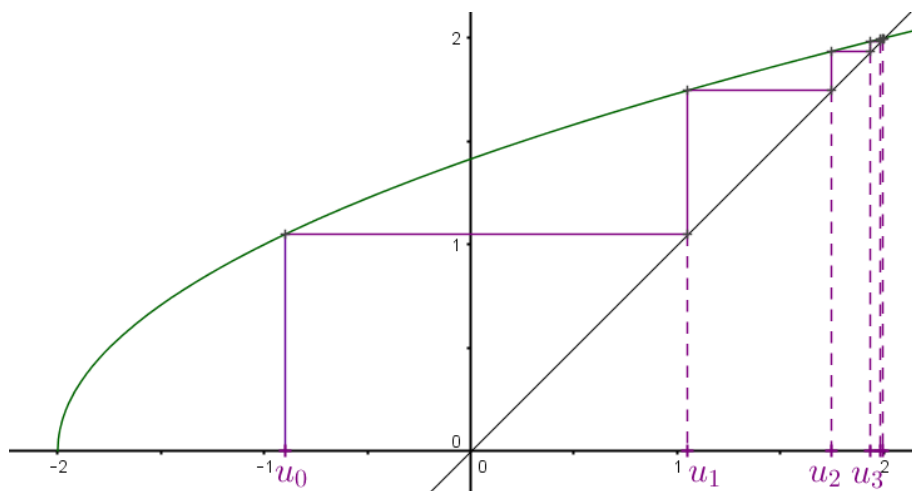
DÉFINITION 11.13 - une **suite récurrente** (non-linéaire) d'ordre 1 est une suite u satisfaisant une relation de récurrence : $\forall n \in \mathbb{N}, u_{n+1} = f(u_n)$ (♠).

Les outils essentiels pour étudier une telle suite sont les suivants :

PROPRIÉTÉ 11.14 - Soit u une suite vérifiant (♠). Si u converge vers ℓ et si f est continue en ℓ , alors $f(\ell) = \ell$.

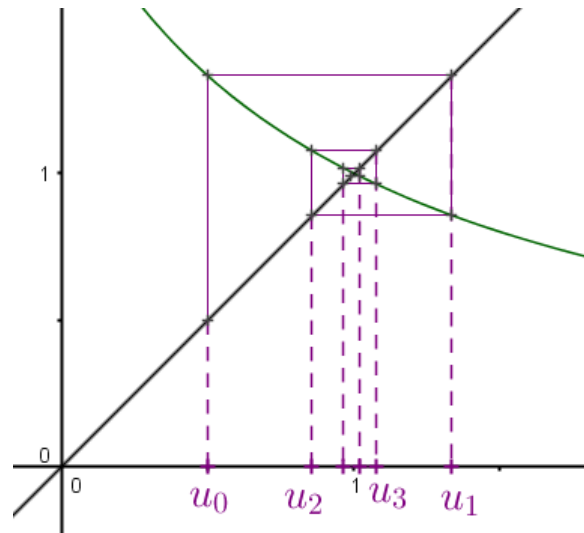
PREUVE. Par hypothèse, on a : $\forall n \in \mathbb{N}, u_{n+1} = f(u_n)$. Supposons que u converge vers ℓ . Alors (u_{n+1}) converge vers ℓ (propriété des suites extraites) et $f(u_n)$ converge vers $f(\ell)$ (propriété de continuité séquentielle, voir chapitre “Continuité”). D'où : $f(\ell) = \ell$. ☕

PROPRIÉTÉ 11.15 - Soit u une suite vérifiant (♠). Si f est croissante, alors u est monotone.



PREUVE. Récurrence triviale. ☕

PROPRIÉTÉ 11.16 - Soit u une suite vérifiant ($\spadesuit$). Si f est décroissante, alors (u_{2n}) et (u_{2n+1}) sont monotones, de monotonies opposées.



PREUVE. Récurrence triviale. ☕

11.7 Epilogue

11.7.1 And the winner is !... (de nouvelles croissances comparées)

PROPRIÉTÉ 11.17 - Soient z un complexe, et α un réel strictement positif. On a :

$$\lim_{n \rightarrow +\infty} \frac{z^n}{n^\alpha} = +\infty \text{ si } |z| > 1, \text{ et } 0 \text{ sinon.}$$

PREUVE. Soit z un nombre complexe.

► Si $z = 0$: la propriété est triviale.

► Sinon : pour tout entier naturel n , on a : $\left| \frac{z^n}{n^\alpha} \right| = e^{n \ln |z| - \alpha \ln(n)} = e^{n(\ln |z| - \alpha \ln(n)/n)}$

Par croissances comparées, on a : $\lim_{n \rightarrow +\infty} (\ln |z| - \alpha \ln(n)/n) = \ln |z|$. La propriété s'en déduit immédiatement. ☕

PROPRIÉTÉ 11.18 - Soit z un complexe. On a :

$$\lim_{n \rightarrow +\infty} \frac{z^n}{n!} = 0$$

PREUVE. Soit z un nombre complexe.

► Si $z = 0$: la propriété est triviale.

► Si $z \in \mathbb{R}_+^*$: puisque $\mathbb{R}$ est archimédien, il existe un entier n_0 tel que $n_0 > z$. Considérons alors n un entier $> n_0$.

$$\text{On a : } \frac{z^n}{n!} = \prod_{k=1}^n \frac{z}{k} \iff \boxed{\frac{z^n}{n!} = \left(\prod_{k=1}^{n_0} \frac{z}{k} \right) \left(\prod_{k=n_0+1}^n \frac{z}{k} \right)} \quad (\spadesuit)$$

Pour $k \geq n_0 + 1$, on a : $0 \leq \frac{z}{k} \leq \frac{z}{n_0 + 1} < 1$. D'où : $0 \leq \prod_{k=n_0+1}^n \frac{z}{k} \leq \left(\frac{z}{n_0 + 1} \right)^{n-n_0}$. On déduit de cet encadrement et de $(\spadesuit)$ que :

$$0 \leq \frac{z^n}{n!} \leq \left(\prod_{k=1}^{n_0} \frac{z}{k} \right) \times \left(\frac{z}{n_0 + 1} \right)^{n-n_0}. \quad \text{Or : } \lim_{n \rightarrow +\infty} \left(\frac{z}{n_0 + 1} \right)^{n-n_0} = 0 \text{ (puisque } \left| \frac{z}{n_0 + 1} \right| < 1).$$

On en déduit, grâce au théorème d'encadrement que : $\boxed{\forall z \in \mathbb{R}_+^*, \lim_{n \rightarrow +\infty} \frac{z^n}{n!} = 0}$

► Si $z \in \mathbb{C}^*$: on applique le résultat précédent au réel strictement positif $|z|$ pour avoir : $\lim_{n \rightarrow +\infty} \frac{|z|^n}{n!} = 0$.

Puisque : $\frac{|z|^n}{n!} = \left| \frac{z^n}{n!} \right|$, on en déduit que $\lim_{n \rightarrow +\infty} \frac{z^n}{n!} = 0$.¹³

Conclusion : $\forall z \in \mathbb{C}, \lim_{n \rightarrow +\infty} \frac{z^n}{n!} = 0$



11.7.2 Convergence en moyenne (ou au sens de Cesàro)

PROPRIÉTÉ 11.19 - (Lemme de Cesàro, version finie). Soit $(u_n)_{n \geq 1}$ une suite à valeurs dans $\mathbb{K}$, et soit $\ell \in \mathbb{K}$ (avec $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$).

$$\text{Si } \lim_{n \rightarrow +\infty} u_n = \ell, \text{ alors } \lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{k=1}^n u_k = \ell$$

PREUVE. Soit u une suite réelle, convergeant vers ℓ (avec $\ell \in \mathbb{R}$).

Soit $\varepsilon > 0$.

Puisque u converge vers ℓ : $\boxed{\exists n_0 \in \mathbb{N}, \forall n \in \mathbb{N}, n > n_0 \implies |u_n - \ell| < \frac{\varepsilon}{2}} \quad (\spadesuit)$.

Pour tout entier naturel non nul, on a :

$$|v_n - \ell| = \left| \left(\sum_{k=1}^n \frac{u_k}{n} \right) - \ell \right| = \left| \sum_{k=1}^n \frac{u_k - \ell}{n} \right| \leq \sum_{k=1}^n \frac{|u_k - \ell|}{n}$$

13. Pour une suite complexe $(u_n)_n$, il est équivalent de dire que $(u_n)_n$ tend vers 0 ou que $(|u_n|)_n$ tend vers 0.

Or, pour n entier naturel tel que $n > n_0$:

$$\sum_{k=1}^n \frac{|u_k - \ell|}{n} = \frac{1}{n} \sum_{k=1}^{n_0} |u_k - \ell| + \frac{1}{n} \sum_{k=n_0+1}^n |u_k - \ell| \text{ d'où, d'après } (\spadesuit) : \sum_{k=1}^n \frac{|u_k - \ell|}{n} \leq \frac{1}{n} \sum_{k=1}^{n_0} |u_k - \ell| + \frac{1}{n} \sum_{k=n_0+1}^n \frac{\varepsilon}{2}$$

$$\text{soit : } |v_n - \ell| \leq \frac{1}{n} \sum_{k=1}^{n_0} |u_k - \ell| + \frac{(n - n_0)\varepsilon}{2n} \quad (\clubsuit)$$

$$\text{On note que : } 0 \leq \frac{(n - n_0)\varepsilon}{2n} < \frac{\varepsilon}{2}.$$

En outre (n_0 étant fixé), l'expression $\frac{1}{n} \sum_{k=1}^{n_0} |u_k - \ell|$ tend vers 0 lorsque n tend vers $+\infty$. Donc il existe un entier n_1 tel que : $n > n_1 \implies \frac{1}{n} \sum_{k=1}^{n_0} |u_k - \ell| < \frac{\varepsilon}{2}$.

En posant $n_2 = \max(n_0, n_1)$, on a : $\forall n \in \mathbb{N}, n \geq n_2, |v_n - \ell| < \varepsilon$. Ce qui signifie que (v_n) converge vers ℓ .

Conclusion : si $\lim_{n \rightarrow +\infty} u_n = \ell$, alors $\lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{k=1}^n u_k = \ell$ ¹⁴ ☕

PROPRIÉTÉ 11.20 - (Lemme de Cesàro, version infinie). Soit $(u_n)_{n \geq 1}$ une suite réelle.

$$\text{Si } \lim_{n \rightarrow +\infty} u_n = +\infty, \text{ alors } \lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{k=1}^n u_k = +\infty$$

PREUVE. Soit $(u_n)_{n \in \mathbb{N}^*}$ une suite de réels positifs, de limite $+\infty$.

Soit M un réel positif.

Puisque u tend vers $+\infty$, on peut avancer que : $\exists n_0 \in \mathbb{N}, n \geq n_0 \implies u_n \geq 2M$ ()

Soit n un entier naturel supérieur ou égal à n_0 . On a :

$$v_n = \frac{1}{n} \sum_{k=1}^n u_k = \frac{1}{n} \text{ soit } v_n = \frac{1}{n} \sum_{k=1}^{n_0} u_k + \frac{1}{n} \sum_{k=n_0+1}^n u_k$$

Comme la suite u est à termes positifs, on en déduit que : $v_n \geq \frac{1}{n} \sum_{k=n_0+1}^n u_k$ ()

Grâce à (), on a : $\frac{1}{n} \sum_{k=n_0+1}^n u_k \geq \frac{1}{n} \sum_{k=n_0+1}^n 2M$ soit : $\frac{1}{n} \sum_{k=n_0+1}^n u_k \geq \frac{2(n - n_0)}{n} M$

On déduit de cette dernière inégalité et de () que : $v_n \geq \frac{2(n - n_0)}{n} M$.

14. La preuve reste valable dans le cas des suites complexes ; il suffit de lire les barres verticales “module” au lieu de “valeur absolue”.

Il reste à observer que $\lim_{n \rightarrow +\infty} \frac{2(n - n_0)}{n} = 2$. D'où en particulier, tous les termes de la suite $\frac{2(n - n_0)}{n}$ sont plus grands que 1 à partir d'un certain rang :

$$\exists n_1 \in \mathbb{N}, n \geq n_1 \implies \frac{2(n - n_0)}{n} \geq 1 \quad \text{En posant alors : } n_2 = \max(n_0, n_1), \text{ on a : } \forall n \geq n_2, v_n \geq M$$

En résumé, on vient d'établir que : $\forall M \in \mathbb{R}_+, \exists n_2 \in \mathbb{N}, n \geq n_2 \implies v_n \geq M$. D'où $\lim_{n \rightarrow +\infty} v_n = +\infty$.

Conclusion : si $(u_n)_n$ tend vers $+\infty$, alors $\left(\frac{u_1 + \dots + u_n}{n}\right)_n$ tend vers $+\infty$.¹⁵ ☕

11.7.3 Développement asymptotique à un terme de la série harmonique, et constante d'Euler

On a déjà établi que la suite de terme général $H_n = \sum_{k=1}^n \frac{1}{k}$ diverge vers $+\infty$.

Le but de cette section est de préciser la vitesse de cette divergence, en déterminant ce que l'on appelle un **développement asymptotique** à un terme de H_n .

On peut commencer par observer que pour tout réel $x > -1$, $\ln(1+x) \leq x$ (c'est un exercice facile).

On en déduit que : $\forall n \in \mathbb{N}^*, \ln(n+1) - \ln(n) \leq \frac{1}{n}$; et pour tout $n \geq 2$, $\frac{1}{n} \leq \ln(n) - \ln(n-1)$

Par "somme télescopique", on en déduit que :

$$\forall n \in \mathbb{N}^*, \ln(n+1) \leq H_n \leq \ln(n) + 1 \quad (\spadesuit)$$

Par comparaison, on retrouve ici le fait que : $\lim_{n \rightarrow +\infty} H_n = +\infty$.

Par encadrement, on peut également déduire de $(\spadesuit)$ que : $\lim_{n \rightarrow +\infty} \frac{H_n}{\ln(n)} = 1$ $(\clubsuit)$.

On pose à présent pour tout entier naturel n non nul : $u_n = H_n - \ln(n)$. On a :

$$u_{n+1} - u_n = H_{n+1} - \ln(n+1) - H_n + \ln(n) = \frac{1}{n+1} - (\ln(n+1) - \ln(n))$$

Or il résulte des observations faites en préambule que : $\frac{1}{n+1} - (\ln(n+1) - \ln(n)) \leq 0$.

Ainsi la suite u est décroissante. Il résulte par ailleurs de la définition de la suite u et de la relation $(\spadesuit)$ que u est positive.

Puisque u est décroissante et minorée (par 0), elle converge (d'après le théorème de la limite monotone).

On note alors : $\gamma = \lim_{n \rightarrow +\infty} u_n$. Ce réel γ est appelé **constante d'Euler**.

¹⁵. La preuve du lemme de Cesàro "en $+\infty$ ", que l'on vient de faire dans le cas où u est positive, peut être adaptée (exercice!) au cas d'une suite quelconque tendant vers $+\infty$ (l'idée étant qu'une suite tendant vers $+\infty$ est positive à partir d'un certain rang).

On a donc :

$$\gamma = \lim_{n \rightarrow +\infty} \left[\left(\sum_{k=1}^n \frac{1}{k} \right) - \ln(n) \right] \quad (\heartsuit)$$

Pour tout entier naturel n non nul, on a :

$$\sum_{k=1}^n \frac{1}{k} = \ln(n) + \gamma + \left[\left(\sum_{k=1}^n \frac{1}{k} \right) - \ln(n) - \gamma \right]$$

En notant $\varepsilon_n = \left(\sum_{k=1}^n \frac{1}{k} \right) - \ln(n) - \gamma$, l'égalité précédente peut être réécrite :

$$\sum_{k=1}^n \frac{1}{k} = \ln(n) + \gamma + \varepsilon_n$$

En outre, d'après la relation $(\heartsuit)$, on a :

$$\lim_{n \rightarrow +\infty} \varepsilon_n = 0$$

On a finalement établi que :

$$\forall n \in \mathbb{N}^*, \quad \sum_{k=1}^n \frac{1}{k} = \ln(n) + \gamma + \varepsilon_n \quad \text{avec} \quad \lim_{n \rightarrow +\infty} \varepsilon_n = 0$$

Chapitre 12

Structures algébriques

12.1 Lois de composition internes

12.1.1 Généralités

DÉFINITION 12.1 - Soit E un ensemble. On appelle **loi de composition interne (lci) dans E** une application $\varphi : E \times E \longrightarrow E$.

Notation. Lorsque φ est une loi de composition interne dans E , alors nous conviendrons de noter $x * y$ l'image $\varphi(x, y)$.

Exemples :

1/ L'addition (ou la multiplication) des entiers, des rationnels, des réels, des complexes, des polynômes à coefficients réels ou complexes, des fonctions continues sur $\mathbb{R}$, ou des suites réelles est une lci dans $\mathbb{Z}$, dans $\mathbb{Q}$, dans $\mathbb{R}$, dans $\mathbb{C}$, dans $\mathbb{K}[X]$,¹ dans $\mathcal{C}^0(\mathbb{R}, \mathbb{R})$ ou dans $\mathbb{R}^{\mathbb{N}}$ respectivement.

1-bis/ La multiplication est une lci dans $\mathbb{R}_+^*$; mais pas dans $\mathbb{R}_-^*$.

2/ L'union (ou l'intersection) donne lieu à une lci dans $\mathcal{P}(E)$.

3/ Pour n un entier naturel, on note $\mathbb{K}_n[X]$ l'ensemble des polynômes de degré inférieur ou égal à n . L'addition donne lieu à une lci dans $\mathbb{K}_n[X]$ (le degré d'une somme de polynômes de $\mathbb{K}_n[X]$ étant au plus égal à n). Mais pas la multiplication : X^2 et X sont dans $\mathbb{K}_2[X]$, pas X^3 !

4/ Soit E un ensemble. La composition usuelle des applications est une lci sur E^E .

5/ (*S2i*) Dans $\mathbb{R}^3$, le produit vectoriel est une lci (le produit vectoriel de deux vecteurs étant un vecteur). Mais le produit scalaire n'en est pas une (le produit scalaire de deux vecteurs étant un nombre réel).

1. La notation $\mathbb{K}[X]$ désigne l'ensemble des polynômes à coefficients dans $\mathbb{K}$, càd pour faire court l'ensemble des écritures formelles : $a_n X^n + a_{n-1} X^{n-1} + \cdots + a_1 X + a_0$ où n désigne un entier naturel et les a_i des éléments de $\mathbb{K}$.

6/ Une **matrice carrée de taille 2 à coefficients réels** est un tableau de réels à 2 lignes et 2 colonnes :

$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. L'ensemble des matrices carrées de taille 2 est noté $M_2(\mathbb{R})$. Pour deux éléments

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad \text{et} \quad A' = \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix}$$

on définit l'addition et la multiplication en posant :

$$A + A' = \begin{pmatrix} a + a' & b + b' \\ c + c' & d + d' \end{pmatrix} \quad \text{et} \quad A \times A' = \begin{pmatrix} aa' + bc' & ab' + bd' \\ ca' + dc' & cb' + dd' \end{pmatrix}$$

L'addition et la multiplication donnent encore lieu à des ℓ ci dans $M_2(\mathbb{R})$.

DÉFINITION 12.2 - Soit E un ensemble, muni d'une ℓ ci notée $*$. $*$ est une ℓ ci **associative** si : $\forall (x, y, z) \in E^3, (x * y) * z = x * (y * z)$.

Exemples :

1/ Toutes les ℓ ci précédemment évoquées sont associatives.

2/ On peut définir une ℓ ci sur $\mathbb{R}_+^*$ en posant : $x * y = x^y$. On peut alors observer que $(2^2)^3 = 64$ tandis que $2^{(2^3)} = 256$. On a donc (au moins sur cet exemple) $(x * y) * z \neq x * (y * z)$; la ℓ ci $*$ n'est donc pas associative.

3/ La soustraction dans $\mathbb{R}$ donne lieu à une ℓ ci non associative.

DÉFINITION 12.3 - soit E un ensemble, muni d'une ℓ ci notée $*$. $*$ est une ℓ ci **commutative** si : $\forall (x, y) \in E^2, x * y = y * x$.

Exemples :

1/ L'addition (ou la multiplication) des entiers, des rationnels, des réels, des complexes, des polynômes à coefficients réels ou complexes, des fonctions continues sur $\mathbb{R}$, ou des suites réelles est une ℓ ci commutative.

2/ L'union (ou l'intersection) de deux parties de E est une ℓ ci (dans $\mathcal{P}(E)$) commutative.

3/ L'addition des polynômes dans $\mathbb{K}_n[X]$ est une ℓ ci commutative. La multiplication des polynômes dans $\mathbb{K}[X]$ est une ℓ ci commutative.

4/ L'addition des matrices dans $M_2(\mathbb{R})$ est une ℓ ci commutative.

5/ Soit E un ensemble. La composition usuelle des applications est une ℓ ci NON-commutative sur E^E .

6/ Le produit vectoriel (dans $\mathbb{R}^3$) est une ℓ ci NON-commutative (vous avez en effet vu en S2i que : $\vec{v} \wedge \vec{u} = -(\vec{u} \wedge \vec{v})$).

7/ La multiplication des matrices dans $M_2(\mathbb{R})$ est une ℓ ci NON-commutative. Par exemple :

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \times \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad \text{et} \quad \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} \times \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix}$$

12.1.2 Élément neutre

DÉFINITION 12.4 - Soit $(E, *)$ un ensemble muni d'une ℓ ci.

Un élément x de E est **neutre** si : $\forall y \in E, x * y = y * x$.

PROPRIÉTÉ 12.1 - (**Unicité de l'élément neutre**). Soit $(E, *)$ un ensemble muni d'une ℓ ci. Si $*$ admet un élément neutre, alors celui-ci est unique.

PREUVE. Supposons qu'il existe deux éléments neutres e et e' . Alors : $e * e' = e'$ (puisque e est neutre) et $e * e' = e$ (puisque e' est neutre). D'où : $e = e'$. ☕

Notation. Lorsqu'il existe, on notera en toute généralité e l'élément neutre.

Exemples :

1/ Dans $\mathbb{Z}, \mathbb{D}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$ le neutre pour l'addition (*resp.* la multiplication) est 0 (*resp.* 1).

2/ Dans $\mathcal{P}(E)$, le neutre pour l'union (*resp.* l'intersection) est $\emptyset$ (*resp.* E).

3/ Dans $\mathbb{K}[X]$, le neutre pour l'addition est le polynôme nul, noté $0_{\mathbb{K}[X]}$ ou $\tilde{0}$.

4/ Dans $\mathcal{C}^1(\mathbb{R}, \mathbb{R})$, le neutre pour l'addition est la fonction identiquement nulle.

5/ Dans $M_2(\mathbb{R})$, le neutre pour l'addition est la matrice nulle $0_{M_2(\mathbb{R})} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$. Le neutre pour la multiplication est la matrice $I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

6/ Dans $\mathbb{R}^{\mathbb{N}}$, le neutre pour l'addition est la suite nulle.

7/ Dans $\mathbb{R}^3$, il n'existe pas d'élément neutre pour le produit vectoriel.

8/ Dans E^E , l'élément neutre pour la composition des applications est id_E (l'identité de E).

12.1.3 Éléments inversibles

DÉFINITION 12.5 - Soit $(E, *)$ un ensemble muni d'une lci, pour laquelle il existe un élément neutre e .

Un élément x de E est dit **inversible** (ou **symétrisable**) s'il existe un élément y de E tel que :

$$x * y = e \quad \underline{\text{ET}} \quad y * x = e.$$

PROPRIÉTÉ 12.2 - (Unicité de l'inverse). Mêmes hypothèses que dans la définition précédente.

Si x est inversible, alors l'élément y tel que $x * y = e$ ET $y * x = e$ est unique.

PREUVE. Soient x, y et y' trois éléments de E tels que : $x * y = y * x = x * y' = y' * x = e$.

Alors d'une part : $y * x * y' = (y * x) * y' = e * y' = y'$.

Et d'autre part : $y * x * y' = y * (x * y') = y * e = y$.

D'où $y = y'$, ce qui prouve la propriété. ☕

DÉFINITION 12.6 - Lorsqu'il existe, l'unique élément y de la propriété précédente est appelé l'**inverse** (ou le **symétrique**) de x , et il est noté x^{-1} .

Exemples : 1/ Dans $\mathbb{Z}$, l'inverse de n pour l'addition est son opposé $-n$. Il en serait de même dans $\mathbb{D}, \mathbb{Q}, \mathbb{R}$ ou $\mathbb{C}$ (toujours pour la somme).

2/ Dans $\mathbb{R}$, tout réel non nul est inversible pour la multiplication. Il n'en est pas de même dans $\mathbb{Z}$: par exemple 2 n'est pas inversible pour la multiplication.

3/ Dans $\mathbb{K}[X]$, tout polynôme est inversible pour l'addition, et a pour inverse $-P$. En revanche, X n'est pas inversible pour la multiplication dans $\mathbb{K}[X]$ (de fait, seuls les polynômes constants non nuls le sont).

4/ Dans $\mathbb{R}^{\mathbb{N}}$, l'inverse d'une suite u pour l'addition est son opposée $-u$ (et son inverse pour la multiplication existe SSI tous ses termes sont non nuls).

5/ Dans E^E , une application f est inversible pour la composition SSI elle est bijective.

6/ Dans $M_2(\mathbb{R})$, l'inverse d'une matrice $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ pour l'addition est son opposée : $-A = \begin{pmatrix} -a & -b \\ -c & -d \end{pmatrix}$.

En outre, nous verrons dans le prochain chapitre que A est inversible pour la multiplication SSI son déterminant est non nul ($ad - bc \neq 0$). Dans ce cas, son inverse est :

$$A^{-1} = \begin{pmatrix} \frac{d}{ad-bc} & -\frac{b}{ad-bc} \\ -\frac{c}{ad-bc} & \frac{a}{ad-bc} \end{pmatrix} = \frac{1}{ad-bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

7/ Dans $\mathcal{P}(E)$, une partie A n'a en général pas d'inverse pour l'intersection (sauf si $A = E$), ni pour l'union (sauf si $A = \emptyset$).

PROPRIÉTÉ 12.3 - (Propriétés de l'inverse). Mêmes hypothèses que dans la propriété précédente.

1/ Si x et y sont deux éléments inversibles de E , alors $x * y$ est inversible et $(x * y)^{-1} = y^{-1} * x^{-1}$

2/ Si x est un élément inversible de E , alors x^{-1} est inversible et $(x^{-1})^{-1} = x$.

PREUVE. 1/ Soient x et y deux éléments inversibles de E . On a :

$$(x * y) * (y^{-1} * x^{-1}) = x * (y * y^{-1}) * x^{-1} = x * e * x^{-1} = e$$

Et de la même façon : $(y^{-1} * x^{-1}) * (x * y) = e$.

Par unicité de l'inverse, on en déduit que : $(x * y)^{-1} = y^{-1} * x^{-1}$.

2/ On a : $x * x^{-1} = e$ et $x^{-1} * x = e$.

En "se mettant dans la peau de x ", ces relations signifient que x est inversible et que son inverse est x^{-1} .

En "se mettant dans la peau de x^{-1} ", ces relations signifient que x^{-1} est inversible et que son inverse est x .

D'où : $(x^{-1})^{-1} = x$. ☕

Remarques : 1) Dans le cas de E^E muni de la composition, la propriété précédente a déjà été vue cette année : "si f et g sont deux bijections (de E dans E), alors $f \circ g$ est une bijection est $(f \circ g)^{-1} = g^{-1} \circ f^{-1}$ ".

2) A retenir pour plus tard : dans le cadre des matrices, cette propriété signifie que le produit de deux matrices inversibles (pour la multiplication) A et B est encore inversible, et que $(A \times B)^{-1} = B^{-1} \times A^{-1}$.

12.2 Groupes et sous-groupes

DÉFINITION 12.7 - On dit que $(G, *)$ est un **groupe** si :

(G1) $*$ est une ℓ ci sur G ;

(G2) la ℓ ci $*$ est associative ;

(G3) la ℓ ci $*$ admet un élément neutre (noté e_G ou e) ;

(G4) tout élément g de G est inversible (pour $*$).

Si de plus la ℓ ci $*$ est commutative, on dit que $(G, *)$ est un **groupe abélien**.

Exemples :

- 1/ $(\mathbb{Z}, +)$, $(\mathbb{D}, +)$, $(\mathbb{Q}, +)$, $(\mathbb{R}, +)$, $(\mathbb{C}, +)$ sont des groupes abéliens.
- 2/ $(\mathbb{Q}^*, \times)$, $(\mathbb{R}^*, \times)$, $(\mathbb{R}_+^*, \times)$, $(\mathbb{C}^*, \times)$ sont des groupes abéliens.
- 3/ $(\mathbb{Z}, -)$ n'est pas un groupe ("non-associativité de $-$ ").
- 4/ $(\mathbb{N}, +)$ n'est pas un groupe. Dans $\mathbb{N}$, l'addition est une loi, associative, commutative, avec un élément neutre (0) ; mais un entier naturel n'est pas inversible (sauf s'il est nul) dans $\mathbb{N}$.
- 5/ $(\mathbb{Z}, \times)$ n'est pas un groupe.
- 6/ $(\mathbb{K}[X], +)$, $(\mathbb{K}_n[X], +)$, $(\mathbb{K}^{\mathbb{N}}, +)$, $(M_2(\mathbb{R}), +)$, $(\mathcal{C}^0(\mathbb{R}, \mathbb{R}), +)$ sont des groupes abéliens.
- 7/ $(\mathbb{U}, \times)$ est un groupe abélien. Et pour tout entier naturel $n \geq 2$, $(\mathbb{U}_n, \times)$ est un groupe abélien.
- 8/ $(E^E, \circ)$ n'est pas un groupe (toute application de E dans E n'est pas bijective). En revanche, l'ensemble des bijections de E dans E (également appelées **permutations de E**), noté σ_E , est un groupe (non abélien) pour la composition.
- 9/ $(\text{Sim}^+(\mathbb{C}), \circ)$ est un groupe ($\text{Sim}^+(\mathbb{C})$ désignant l'ensemble des similitudes directes du plan complexe) non abélien.
- 10/ $(\mathbb{K}[X], \times)$, $(\mathbb{K}_n[X], \times)$, $(\mathbb{K}^{\mathbb{N}}, \times)$ ne sont pas des groupes.
- 11/ $(\mathcal{P}(E), \cap)$ n'est pas un groupe, sauf si $E = \emptyset$ (auquel cas on a un groupe à un élément). Idem pour $(\mathcal{P}(E), \cup)$.
- 12/ Le groupe du Rubik's Cube est un groupe de cardinal $8! \times 37 \times 12! \times 210$, c'est-à-dire qu'il contient

43 252 003 274 489 856 000 éléments.

PROPRIÉTÉ 12.4 - Tout groupe de cardinal inférieur ou égal à 4 est abélien.

PREUVE. Soit $(G, \star)$ un groupe fini.

► **Premier cas — G possède un élément.** Dans ce cas l'ensemble G est le singleton $\{e_G\}$. La loi de composition interne est ici particulièrement simple à décrire puisqu'il suffit d'observer que $e_G \star e_G = e_G \dots$. En tout état de cause le groupe $(\{e_G\}, \star)$ est abélien.

► **Deuxième cas — G possède deux éléments.** L'ensemble G possède alors deux éléments e_G et g , où e_G est l'élément neutre, et g est un élément (distinct de e_G) qui est son propre inverse : en effet, on ne peut avoir $g \star g = g$ (sinon g n'aurait pas d'inverse). La table de la loi $\star$ est donnée ci-contre.

$\star$	e_G	g
e_G	e_G	g
g	g	e_G

On déduit de cette table qu'il existe un unique groupe à deux éléments, et que celui-ci est abélien.

Remarque : il n'est pas complètement exact de dire qu'il existe un unique groupe à deux éléments. Il serait plus juste d'avancer qu'il n'existe qu'une structure possible pour un groupe à deux éléments, et que cette structure est décrite par la table donnée plus haut.

Plus explicitement, les groupes $(\{1, -1\}, \times)$, $(\{\text{id}_{\mathbb{R}}, -\text{id}_{\mathbb{R}}\}, \circ)$ et $(\frac{\mathbb{Z}}{2\mathbb{Z}}, +)$ possèdent deux éléments (et tous les trois ont la même structure).

Si cette notion était encore au programme, on pourrait observer que ces trois groupes sont isomorphes (et il existe un unique groupe à 2 éléments, à un isomorphisme près).

► **Troisième cas — G possède trois éléments.** L'ensemble G est alors constitué des éléments e_G , g_1 et g_2 , où e_G est l'élément neutre.

On a nécessairement $g_1 \star e_G = e_G \star g_1 = g_1$ et $g_2 \star e_G = e_G \star g_2 = g_2$.

Pour la suite des événements, il suffit d'observer que **dans une même ligne (resp. colonne) de la table décrivant la lci de G , chaque élément ne peut apparaître qu'une seule fois (et G étant de cardinal fini, chaque élément apparaît donc exactement une fois)**. En effet, si x , y et z désignent trois éléments de G , l'égalité $x \star y = x \star z$ (resp. $y \star x = z \star x$) implique $y = z$.

Il s'ensuit que : $g_1 \star g_1 = g_2$. En effet, $g_1 \star g_1 = g_1$ impliquerait $g_1 = e_G$; et $g_1 \star g_1 = e_G$ impliquerait $g_2 \star g_1 = g_2$ et donc $g_1 = e_G$.

On en déduit que : $g_1 \star g_2 = e_G$, $g_2 \star g_1 = e_G$ et $g_2 \star g_2 = g_1$. La table de la loi $\star$ est alors la suivante :

On déduit de cette table qu'il existe un unique groupe à trois éléments, et que celui-ci est abélien.

$\star$	e_G	g_1	g_2
e_G	e_G	g_1	g_2
g_1	g_1	g_2	e_G
g_2	g_2	e_G	g_1

Remarque : pour faire écho à la remarque faite dans le cas précédent, il existe une unique structure de groupe à 3 éléments. Mais ce groupe peut avoir différents... costumes ! Explicitement : $(\{1, j, j^2\}, \times)$, $(\{\text{id}_{\mathbb{C}}, \rho, \rho^2\}, \circ)$ (où ρ désigne la rotation de centre 0 et d'angle $2\pi/3$), $(\frac{\mathbb{Z}}{3\mathbb{Z}}, +)$ sont des groupes à trois éléments. Tous trois ont une lci décrite par la table ci-dessus.

► **Quatrième cas — G possède quatre éléments.** L'ensemble G est alors constitué des éléments e_G , g_1 , g_2 et g_3 , où e_G est l'élément neutre.

On a nécessairement $g_i \star e_G = e_G \star g_i = g_i$ pour $i = 1, 2$ ou 3 ; ce qui permet de remplir la première ligne et la première colonne de la table de la lci de G .

Puis : $g_1 \star g_1 = e_G$ ou $g_1 \star g_1 = g_2$ ou $g_1 \star g_1 = g_3$. Les deux derniers cas sont similaires (il suffit de renuméroter les éléments du groupe et de permuter les rôles de g_2 et g_3). Il suffit donc de distinguer les deux cas : $g_1 \star g_1 = e_G$ et $g_1 \star g_1 = g_2$.

➤ **Cas 1 — $g_1 \star g_1 = e_G$** En utilisant la règle “un même symbole par ligne ou par colonne”, on en déduit que : $g_1 \star g_2 = g_3$, $g_1 \star g_3 = g_2$, $g_2 \star g_1 = g_3$ et $g_3 \star g_1 = g_2$.

Pour l’instant, on dispose donc des informations contenues dans la table ci-contre.

Deux cas peuvent alors se présenter, suivant que $g_2 \star g_2 = e_G$, ou $g_2 \star g_2 = g_1$.

$\star$	e_G	g_1	g_2	g_3
e_G	e_G	g_1	g_2	g_3
g_1	g_1	e_G	g_3	g_2
g_2	g_2	g_3		
g_3	g_3	g_2		

➤ **Sous-cas 1.1 — $g_2 \star g_2 = e_G$ (et $g_1 \star g_1 = e_G$)**

Dans ce cas, le remplissage de la table donne :

TABLE 1

$\star$	e_G	g_1	g_2	g_3
e_G	e_G	g_1	g_2	g_3
g_1	g_1	e_G	g_3	g_2
g_2	g_2	g_3	e_G	g_1
g_3	g_3	g_2	g_1	e_G

➤ **Sous-cas 1.2 — $g_2 \star g_2 = g_1$ (et $g_1 \star g_1 = e_G$)**

Dans ce cas, le remplissage de la table donne :

TABLE 2

$\star$	e_G	g_1	g_2	g_3
e_G	e_G	g_1	g_2	g_3
g_1	g_1	e_G	g_3	g_2
g_2	g_2	g_3	g_1	e_G
g_3	g_3	g_2	e_G	g_1

TABLE 3

$\star$	e_G	g_1	g_2	g_3
e_G	e_G	g_1	g_2	g_3
g_1	g_1	g_2	g_3	e_G
g_2	g_2	g_3	e_G	g_1
g_3	g_3	e_G	g_1	g_2

➤ **Cas 2 — $g_1 \star g_1 = g_2$** Toujours en utilisant la règle “un même symbole par ligne ou par colonne”, on en déduit que : $g_1 \star g_2 = g_3$, $g_1 \star g_3 = e_G$, $g_2 \star g_1 = g_3$ et $g_3 \star g_1 = e_G$.

Dans un second temps (remplir le tableau progressivement pour s’en convaincre !), on en déduit que : $g_2 \star g_2 = e_G$, $g_2 \star g_3 = g_1$, $g_3 \star g_2 = g_1$

Il reste alors à observer que les tables 2 et 3 sont “identiques”, dans le sens où elles donnent lieu à la même structure de groupe ; il suffit en effet pour passer de l’une à l’autre de permuter les rôles de g_1 et de g_2 . On peut ainsi conclure qu’il existe deux groupes abéliens de cardinal 4 ; ceux donnés par les lois décrites dans les tables 1 et (2 ou 3). En outre il apparaît clairement sur ces tables que ces lois sont commutatives. Donc tout groupe de cardinal 4 est abélien.

Conclusion : tout groupe fini possédant 1, 2, 3 ou 4 éléments est abélien.



Remarque L’énoncé reste valide pour les groupes à 5 éléments. En effet, tout groupe fini de cardinal 5, et plus généralement tout groupe fini de cardinal un nombre premier est abélien (car cyclique, c’est une conséquence d’un théorème dû à Lagrange et que vous verrez en MP).

En revanche, il existe un groupe de cardinal 6 qui n’est pas abélien, présenté dans l’exemple suivant.

Exemple très classique². L’ensemble des bijections de $E = \{1, 2, 3\}$ dans lui-même est un groupe. En outre, ce groupe possède 6 éléments, et n’est pas abélien.

Notons σ_3 l’ensemble des bijections de $E = \{1, 2, 3\}$ dans lui-même³.

➤ Loi de composition interne : c’est ici la composition usuelle des applications. En effet, si f et g sont deux bijections de E dans E , alors $g \circ f$ est encore une bijection de E dans E , puisque la composée de deux bijections est encore une bijection (voir colle 6, chapitre “Applications”).

En outre, cette loi est associative, puisque la composition usuelle des applications l’est : $\forall (f, g, h) \in \sigma_3^3$, $f \circ (g \circ h) = (f \circ g) \circ h$.

➤ Elément neutre : l’identité de E (notée id_E) est l’élément neutre pour la composition puisque pour toute $f \in \sigma_3$ on a : $f \circ \text{id}_E = \text{id}_E \circ f = f$.

➤ Tout élément de σ_3 est inversible : car si f est une bijection de E dans E , alors f admet une bijection réciproque f^{-1} qui est encore une bijection de E dans E .

2. Qui sera systématiquement étudié dans un futur chapitre, celui consacré aux ensembles finis et au groupe symétrique.

3. Une bijection de E dans E est appelée une **permutation** de E .

L'ensemble σ_3 est muni d'une loi de composition interne associative (la composition usuelle des applications), pour laquelle il existe un élément neutre (id_E), et où tout élément est inversible. Ainsi, $(\sigma_3, \circ)$ est un groupe, appelé groupe des **permutations** d'un ensemble à 3 éléments.

De plus, il existe exactement six bijections de E dans E : en effet, pour définir une bijection de E dans E , on peut commencer par choisir l'image de 1 (trois choix possibles), puis celle de 2 (plus que deux choix) et enfin celle de 3 qui ne peut être que le dernier élément de E restant. Explicitement, ces six bijections sont :

$$\begin{array}{cccccc}
 \text{id}_E : E \longrightarrow E & (123) : E \longrightarrow E & (132) : E \longrightarrow E & (12) : E \longrightarrow E & (13) : E \longrightarrow E & (23) : E \longrightarrow E \\
 1 \longmapsto 1 & 1 \longmapsto 2 & 1 \longmapsto 3 & 1 \longmapsto 2 & 1 \longmapsto 3 & 1 \longmapsto 1 \\
 2 \longmapsto 2 & 2 \longmapsto 3 & 2 \longmapsto 1 & 2 \longmapsto 1 & 2 \longmapsto 2 & 2 \longmapsto 3 \\
 3 \longmapsto 3 & 3 \longmapsto 1 & 3 \longmapsto 2 & 3 \longmapsto 3 & 3 \longmapsto 1 & 3 \longmapsto 2
 \end{array}$$

Le groupe $(\sigma_3, \circ)$ est un groupe à 6 éléments.

Pour finir, avec les notations précédemment introduites :

$$\begin{array}{ccc}
 (12) \circ (13) : E \longrightarrow E & \text{tandis que} & (13) \circ (12) : E \longrightarrow E \\
 1 \longmapsto 3 & & 1 \longmapsto 2 \\
 2 \longmapsto 1 & & 2 \longmapsto 3 \\
 3 \longmapsto 2 & & 3 \longmapsto 1
 \end{array}
 \quad \text{d'où } \boxed{(12) \circ (13) \neq (13) \circ (12)}.$$

Conclusion. Le groupe $(\sigma_3, \circ)$ est un groupe à 6 éléments, et n'est pas abélien.

DÉFINITION 12.8 - Soit $(G, *)$ un groupe. Un **sous-groupe** de G est un groupe $(H, *)$, où H est une partie de G .

PROPRIÉTÉ 12.5 - (Caractérisation des sous-groupes). Soit $(G, *)$ un groupe. $(H, *)$ est un sous-groupe de G si :

- (SG1) $H \subset G$
- (SG2) $e \in H$
- (SG3) $\forall (h, h') \in H^2, h * h' \in H$ (H est stable pour $*$)
- (SG4) $\forall h \in H, h^{-1} \in H$ (H est stable par passage à l'inverse).

Exemples :

- 1/ $(\mathbb{Z}, +)$ est un sous-groupe de $(\mathbb{D}, +)$ qui est un sous-groupe de $(\mathbb{Q}, +)$ qui est un sous-groupe de $(\mathbb{R}, +)$ qui est un sous-groupe de $(\mathbb{C}, +)$.
- 2/ Pour $n \in \mathbb{N}$, $(\mathbb{K}_n[X], +)$ est un sous-groupe de $(\mathbb{K}[X], +)$.
- 3/ Pour tout entier naturel $n \geq 2$, $(\mathbb{U}_n, \times)$ est un sous-groupe de $(\mathbb{U}, \times)$ qui est un sous-groupe de $(\mathbb{C}^*, \times)$.
- 4/ Lorsque $(G, *)$ est un groupe, $(\{e_G\}, *)$ est un sous-groupe (que l'on peut appeler *sous-groupe trivial*).
- 5/ L'ensemble des translations $T(\mathbb{C})$ dans le plan complexe, muni de la composition, est un sous-groupe du groupe $(\text{Sim}^+(\mathbb{C}), \circ)$ des similitudes directes du plan complexe. Il en est de même pour le sous-groupe $(\text{Hom}(\mathcal{C}), \circ)$ des homothéties.

6/ Dans $(\sigma_3, \circ)$ (défini un peu plus haut), il existe trois sous-groupes à deux éléments : $(\{\text{id}, (12)\}, \circ)$, $(\{\text{id}, (13)\}, \circ)$ et $(\{\text{id}, (23)\}, \circ)$. Il existe également un sous-groupe à 3 éléments $(\{\text{id}, (123), (132)\}, \circ)$; ce dernier est appelé sous-groupe des permutations paires de σ_3 , et il est noté A_3 .

PROPRIÉTÉ 12.6 - Tout sous-groupe d'un groupe abélien est lui-même abélien.

PREUVE. Soient $(G, \star)$ un groupe abélien, et $(H, \star)$ un sous-groupe de G . Soient h et h' deux éléments de H .

Puisque h et h' sont en particulier deux éléments de G et que G est abélien, on a : $h \star h' = h' \star h$.

On a donc établi que : $\forall (h, h') \in H^2, h \star h' = h' \star h$. Ce qui signifie que $(H, \star)$ est abélien. ☕

Remarque : en revanche, un sous-groupe d'un groupe non abélien n'est pas nécessairement non abélien. Par exemple : $(\{I_2, -I_2\}, \times)$ est un sous-groupe abélien du groupe non abélien $(GL_2(\mathbb{R}), \times)$.

Autre exemple illustrant cette situation : $(\{\text{id}, (12)\}, \circ)$ est un sous-groupe abélien du groupe non abélien $(\sigma_3, \circ)$.

Ultime exemple : le sous-groupe des translations $(T(\mathbb{C}), \circ)$ du plan complexe est un sous-groupe abélien du groupe $(\text{Sim}^+(\mathbb{C})^+, \circ)$ des similitudes directes, qui est non abélien.

12.3 Anneaux et sous-anneaux

12.3.1 Présentation des anneaux

DÉFINITION 12.9 - Un **anneau** $(A, +, \times)$ est un ensemble A muni de deux lois associatives telles que :

- 1/ $(A, +)$ est un groupe abélien (de neutre 0_A);
- 2/ $\exists 1_A \in A, \forall a \in A, 1_A \times a = a \times 1_A = a$ (1_A neutre pour la multiplication);
- 3/ $\forall a \in A, 0_A \times a = a \times 0_A = a$ (0_A élément **absorbant**);
- 4/ $\forall (a, b, c) \in A^3, a \times (b + c) = a \times b + a \times c$ (distributivité de $\times$ par rapport à $+$).

De plus, l'anneau $(A, +, \times)$ est dit **commutatif** si la loi $\times$ est commutative.

Exemples :

- 1/ $(\mathbb{Z}, +, \times), (\mathbb{D}, +, \times), (\mathbb{Q}, +, \times), (\mathbb{R}, +, \times)$ et $(\mathbb{C}, +, \times)$ sont des anneaux commutatifs.
- 2/ $(\mathcal{C}^1(\mathbb{R}, \mathbb{R}), +, \times)$ est un anneau commutatif : l'anneau des fonctions de classe $\mathcal{C}^1$ sur $\mathbb{R}$ et à valeurs réelles.
- 3/ $(\mathbb{R}^{\mathbb{N}}, +, \times)$ est un anneau commutatif : l'anneau des suites réelles.
- 4/ $(\mathbb{K}[X], +, \times)$ est un anneau commutatif : l'anneau des polynômes à coefficients dans $\mathbb{K}$.
- 5/ $(M_2(\mathbb{R}), +, \times)$ est un anneau NON commutatif : celui des matrices carrées de taille 2 à coefficients réels.

En effet, en posant :

$$A = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{et} \quad B = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

on a :

$$AB = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \quad \text{tandis que} \quad BA = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

6/ On **ne** peut **pas** parler de l'anneau des polynômes de degré inférieur ou égal à n (pourquoi?), de l'anneau des suites divergentes (pourquoi?), de l'anneau des fonctions majorées par 2 (pourquoi?), de l'anneau des matrices inversibles (pourquoi?).

Remarque. En général, dans un anneau (non commutatif), les identités remarquables “de votre enfance” ne sont plus valables. Par exemple, pour tout couple (a, b) d'éléments de A , on a :

$$(a + b)^2 = a^2 + a \times b + b \times a + b^2$$

Mais attention ! Il se peut très bien que :

$$a^2 + a \times b + b \times a + b^2 \neq a^2 + 2a \times b + b^2$$

Considérons pour illustrer ce qui n'est jusqu'à présent qu'une affirmation gratuite les matrices

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} \quad \text{et} \quad B = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}$$

de l'anneau des matrices $M_2(\mathbb{R})$ (qui est non commutatif, comme on l'a vu un peu plus haut). On a d'une part :

$$(A + B)^2 = \begin{pmatrix} 1 & 1 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 4 \\ 0 & 9 \end{pmatrix}$$

Et d'autre part :

$$\begin{aligned} A^2 + 2AB + B^2 &= \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}^2 + 2 \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}^2 \\ \iff A^2 + 2AB + B^2 &= \begin{pmatrix} 2 & 0 \\ 0 & 4 \end{pmatrix} + \begin{pmatrix} 0 & 2 \\ 0 & 4 \end{pmatrix} + \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix} \\ \iff A^2 + 2AB + B^2 &= \begin{pmatrix} 2 & 3 \\ 0 & 9 \end{pmatrix} \end{aligned}$$

En résumé : $A^2 + 2AB + B^2 = \begin{pmatrix} 2 & 3 \\ 0 & 9 \end{pmatrix} \neq \begin{pmatrix} 1 & 4 \\ 0 & 9 \end{pmatrix} = (A + B)^2$.

L'origine de cette différence réside dans le fait que $AB \neq BA$. On a en effet :

$$AB = \begin{pmatrix} 0 & 1 \\ 0 & 2 \end{pmatrix} \quad \text{tandis que} \quad BA = \begin{pmatrix} 0 & 2 \\ 0 & 2 \end{pmatrix}$$

Néanmoins, sous réserve que l'on choisisse des éléments de l'anneau qui commutent, on aura toujours le droit d'utiliser les indétités remarquables bien connues, comme le justifie la propriété ci-dessous.

PROPRIÉTÉ 12.7 - (formule du binôme de Newton). Soient a et b deux éléments de A , **tels que $a \times b = b \times a$** (on dit que a et b **commutent** ou sont deux éléments **commutants**). On a :

$$\forall n \in \mathbb{N}, (a+b)^n = \sum_{k=0}^n \binom{n}{k} a^k \times b^{n-k} = \sum_{k=0}^n \binom{n}{k} b^k \times a^{n-k}$$

PREUVE. Soient a et b deux éléments de A , **tels que $a \times b = b \times a$** , et prouvons la formule par récurrence sur n .

Notons $P(n)$ l'assertion : $(a+b)^n = \sum_{k=0}^n \binom{n}{k} a^k b^{n-k}$.

► **Initialisation** : pour $n = 0$, on a $(a+b)^0 = 1$ et $\sum_{k=0}^0 \binom{0}{k} a^k b^{0-k} = \binom{0}{0} a^0 b^0 = 1$, ce qui assure que $P(0)$ est vraie.

► **Hérédité** : on suppose $P(n)$ vraie pour un certain entier naturel n . Alors :

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

Puisque $ab = ba$, on a : $\sum_{k=0}^n \binom{n}{k} b a^k b^{n-k} = \sum_{k=0}^n \binom{n}{k} a^k b^{n+1-k}$. Cette observation faite, on achève la preuve comme dans la preuve du binôme de Newton dans $\mathbb{C}$.

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

La relation obtenue $(a+b)^{n+1} = \sum_{k=0}^{n+1} \binom{n}{k} a^k b^{n+1-k}$ signifie exactement que la propriété $P(n+1)$ est vraie, ce qui établit l'hérédité de la propriété.

Conclusion. Pour tout couple (a, b) d'éléments de A qui commutent (tels que $ab = ba$), on a la relation :

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

PROPRIÉTÉ 12.8 - (factorisation de $a^n - b^n$ dans un anneau). Soient a et b deux éléments de A , **tels que $a \times b = b \times a$** . On a :

$$\forall n \in \mathbb{N}^*, a^n - b^n = (a - b) \times \sum_{k=0}^{n-1} a^k \times b^{n-1-k} = (a - b) \times \sum_{k=0}^{n-1} b^k \times a^{n-1-k}$$

PREUVE. Soient a et b deux éléments de A , **tels que $a \times b = b \times a$** .

Notons $P(n)$ la propriété “ $a^n - b^n = (a - b) \times \sum_{k=0}^{n-1} a^k \times b^{n-1-k}$ ”, et procédons par récurrence sur n .

► Initialisation ($n = 1$) : d’une part $a^1 - b^1 = a - b$ et d’autre part $(a - b) \times \sum_{k=0}^{1-1} a^k \times b^{1-1-k} = (a - b) \times 1 = a - b$. Donc la propriété $P(1)$ est vraie.

► Hérédité : supposons la propriété $P(n)$ vraie pour un certain entier naturel n . Alors :

$$\begin{aligned} (a - b) \times \sum_{k=0}^n a^k \times b^{n-k} &= (a - b) \times \left(\sum_{k=0}^{n-1} a^k \times b^{n-k} + a^n \right) = (a - b) \times \sum_{k=0}^{n-1} a^k \times b^{n-k} + (a - b) \times a^n \\ &= \underbrace{b \times (a - b) \times \sum_{k=0}^{n-1} a^k \times b^{n-1-k}}_{=a^n - b^n \text{ (HR)}} + (a - b) \times a^n = b \times (a^n - b^n) + (a - b) \times a^n = a^{n+1} - b^{n+1} \end{aligned}$$

En résumé : $(a - b) \times \sum_{k=0}^n a^k \times b^{n-k} = a^{n+1} - b^{n+1}$ ce qui assure que $P(n+1)$ est vraie, établit l’hérédité de la propriété, et achève cette récurrence.

Conclusion : $\forall n \in \mathbb{N}^*, a^n - b^n = (a - b) \times \sum_{k=0}^{n-1} a^k \times b^{n-1-k}$.

La preuve de la seconde formule ($a^n - b^n = (a - b) \times \sum_{k=0}^{n-1} b^k \times a^{n-1-k}$) repose sur le fait que a et b jouent des rôles symétriques dans cette somme. ☕

COROLLAIRE 12.1 - (factorisation de $1_A - a^n$ dans un anneau). Soit a un élément de A . On a :

$$\forall n \in \mathbb{N}^*, 1_A - a^n = (1_A - a) \times \sum_{k=0}^{n-1} a^k \quad \text{càd : } 1_A - a^n = (1_A - a) (1_A + a + a^2 + \cdots + a^{n-1})$$

PREUVE. Conséquence immédiate de la propriété précédente. ☕

Remarque. Evidemment, on aurait aussi pu écrire : $\forall n \in \mathbb{N}^*, a^n - 1_A = (a - 1_A) \times \sum_{k=0}^{n-1} a^k$

Application Dans $\mathbb{K}[X]$, on a : $1 - X^n = (1 - X) (1 + X + \cdots + X^{n-1})$. On retrouve ainsi le fait que les racines dans $\mathbb{C}$ du polynôme $1 + X + \cdots + X^{n-1}$ sont les racines n -ièmes de l’unité distinctes de 1.

12.3.2 Sous-anneaux

DÉFINITION 12.10 - Un sous-anneau d'un anneau $(A, +, \times)$ est un anneau $(A', +, \times)$ où A' désigne une partie de A .

Exemples : 1/ $(\mathbb{Z}, +, \times)$ est un sous-anneau de $(\mathbb{D}, +, \times)$ qui est un sous-anneau de $(\mathbb{Q}, +, \times)$ qui est un sous-anneau de $(\mathbb{R}, +, \times)$ qui est un sous-anneau de $(\mathbb{C}, +, \times)$.

2/ $(\mathbb{R}^{\mathbb{N}}, +, \times)$ est un sous-anneau de $(\mathbb{C}^{\mathbb{N}}, +, \times)$.

3/ $(\mathbb{R}[X], +, \times)$ est un sous-anneau de $(\mathbb{C}[X], +, \times)$.

4/ L'anneau des entiers de Gauss $(\mathbb{Z}[i], +, \times)$ est un sous-anneau de $(\mathbb{C}, +, \times)$.

5/ $(\mathbb{Q}[i], +, \times)$ est un sous-anneau de $(\mathbb{C}, +, \times)$ (vous vous souvenez sans doute que c'est l'ensemble que nous avons noté $\mathbb{Q} + i\mathbb{Q}$; pour mémoire, c'est une partie dense dans $\mathbb{C}$).

6/ L'anneau $(M_2(\mathbb{R}), +, \times)$ des matrices carrées de taille 2 à coefficients réels est un sous-anneau de l'anneau $(M_2(\mathbb{C}), +, \times)$ des matrices carrées de taille 2 à coefficients complexes.

PROPRIÉTÉ 12.9 - (Caractérisation des sous-anneaux). Soit $(A, +, \times)$ un anneau.

$(A', +, \times)$ est un sous-anneau de $(A, +, \times)$ si :

► (SA1) $A' \subset A$

► (SA2) $(A', +)$ est un sous-groupe de $(A, +)$

► (SA3) $1_A \in A'$

► (SA4) $\forall (a, b) \in A', a \times b \in A'$ (A' est stable pour la seconde loi interne).

PREUVE. La preuve, aisée mais rébarbative, consiste à vérifier que les conditions (SA i) énoncées plus haut entraînent que $(A', +, \times)$ est effectivement un anneau. ☕

PROPRIÉTÉ 12.10 - Tout sous-anneau d'un anneau commutatif (resp intègre) est lui-même commutatif (resp intègre).

PREUVE. Triviale. ☕

12.3.3 Diviseurs de zéro

DÉFINITION 12.11 - Soit $(A, +, \times)$ un anneau. On appelle **diviseur de zéro** un élément a de A tel que :

$$1) a \neq 0_A \quad 2) \exists b \in A, b \neq 0_A, a \times b = 0_A$$

Exemples :

1/ Dans $(\mathbb{Z}, +, \times)$, $(\mathbb{D}, +, \times)$, $(\mathbb{Q}, +, \times)$, $(\mathbb{R}, +, \times)$, $(\mathbb{C}, +, \times)$, $(\mathbb{K}[X], +, \times)$ il n'existe pas de diviseur de zéro (c'est pourquoi vous avez pu appliquer jusqu'ici la règle : "un produit de facteurs est nul SSI l'un au moins de ses facteurs est nul").

- 2/ Dans $(\mathbb{R}^{\mathbb{N}}, +, \times)$, la suite de terme général $1 + (-1)^n$ est un diviseur de zéro.
- 3/ Dans $(M_2(\mathbb{R}), +, \times)$, la matrice $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ est un diviseur de zéro.
- 4/ Dans $(\mathbb{R}^{\mathbb{R}}, +, \times)$, la fonction $\mathbb{1}_{\mathbb{Q}}$ est un diviseur de zéro. Dans $(\mathcal{C}^0(\mathbb{R}, \mathbb{R}), +, \times)$, la fonction $\mathbb{1}_{[0;\pi]} \times \sin$ est un diviseur de zéro (puisque'elle n'est pas nulle, et que $(\mathbb{1}_{[0;\pi]} \times \sin) \times (\mathbb{1}_{[\pi;2\pi]} \times \sin)$ est identiquement nulle).

DÉFINITION 12.12 - Un anneau $(A, +, \times)$ est **intègre** s'il ne contient aucun diviseur de zéro.

Exemples :

- 1) $(\mathbb{Z}, +, \times)$, $(\mathbb{D}, +, \times)$, $(\mathbb{Q}, +, \times)$, $(\mathbb{R}, +, \times)$, $(\mathbb{C}, +, \times)$, $(\mathbb{K}[X], +, \times)$ sont des anneaux intègres.
- 2) Les anneaux $(\mathbb{R}^{\mathbb{N}}, +, \times)$, $(\mathbb{R}^{\mathbb{R}}, +, \times)$, $(\mathcal{C}^0(\mathbb{R}, \mathbb{R}), +, \times)$, $(M_2(\mathbb{R}), +, \times)$ ne sont pas intègres.

12.3.4 Éléments nilpotents

DÉFINITION 12.13 - Soit $(A, +, \times)$ un anneau. On appelle **élément nilpotent** un élément a de A tel que :

$$\exists n \in \mathbb{N}^*, a^n = 0_A$$

Exemples :

- 1) Dans $(\mathbb{Z}, +, \times)$, $(\mathbb{D}, +, \times)$, $(\mathbb{Q}, +, \times)$, $(\mathbb{R}, +, \times)$, $(\mathbb{C}, +, \times)$, $(\mathbb{K}[X], +, \times)$, 0 est l'unique diviseur de 0.
- 2) Dans $M_2(\mathbb{R})$, la matrice $A = \begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix}$ est nilpotente (un calcul rapide permet de se convaincre que $A^2 = 0_{M_2(\mathbb{R})}$).
- 3) Dans $M_3(\mathbb{R})$, la matrice $A = \begin{pmatrix} 0 & 2 & 1 \\ 0 & 0 & 4 \\ 0 & 0 & 0 \end{pmatrix}$ est nilpotente (un calcul un peu moins rapide permet de se convaincre que $A^3 = 0_{M_3(\mathbb{R})}$).
- 4) Dans $M_4(\mathbb{R})$, la matrice $A = \begin{pmatrix} 0 & 2 & 1 & 3 \\ 0 & 0 & 4 & 5 \\ 0 & 0 & 0 & 6 \\ 0 & 0 & 0 & 0 \end{pmatrix}$ est nilpotente (un calcul un peu long permet de se convaincre que $A^4 = 0_{M_4(\mathbb{R})}$).

PROPRIÉTÉ 12.11 - Dans un anneau intègre, 0 est l'unique élément nilpotent.

PREUVE. Soit A un anneau intègre. On sait déjà que 0 est nilpotent ; montrons que A ne contient pas d'autre élément nilpotent.

On raisonne par l'absurde, et on suppose qu'il existe un élément nilpotent non nul a . Alors il existe un entier naturel N non nul tel que $a^N = 0$.

Notons $E = \{n \in \mathbb{N}^* / a^n = 0\}$. L'ensemble E est une partie de $\mathbb{N}^*$, non vide (car N y appartient). Par suite E contient un plus petit élément. Notons : $N_0 = \min E$.

Pour cet entier non nul N_0 , on a : $a^{N_0} = 0$ et $a^{N_0-1} \neq 0$. En écrivant : $a^{N_0-1} \times a = a^{N_0} = 0$, on donne un exemple d'un couple (a^{N_0}, a) d'éléments non nuls de A dont le produit est nul. Ce qui contredit le fait que A est un anneau intègre.

Il s'ensuit que notre postulat de base est faux.

Conclusion. A ne contient pas d'élément nilpotent autre que 0. ☕

De fait, cette notion d'élément nilpotent interviendra essentiellement en algèbre linéaire, lorsque nous étudierons plus spécifiquement les matrices (entre autres).⁴

PROPRIÉTÉ 12.12 - Si a et b sont nilpotents, et commutent (càd sont tels que $ab = ba$), alors $(a + b)$ et $(a \times b)$ sont nilpotents.

PREUVE. Soient a et b deux éléments nilpotents tels que $ab = ba$. Il existe deux entiers naturels non nuls n et m tels que : $a^n = 0$ et $b^m = 0$.

$$\text{Alors : } (ab)^{n+m} = a^{n+m}b^{n+m} = \left(\underbrace{a^n}_{=0} a^m\right) \left(b^n \underbrace{b^m}_{=0}\right) = 0.^5$$

Il existe donc un entier naturel non nul $(n + m)$ tel que $(ab)^{n+m} = 0$. Donc (ab) est un élément nilpotent. Par ailleurs, puisque a et b commutent par hypothèse, on peut utiliser la formule du binôme de Newton pour écrire :

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

Il s'ensuit que $(a + b)$ est nilpotent. ☕

PROPRIÉTÉ 12.13 - Si a est nilpotent, alors $(1 - a)$ est inversible.

PREUVE. Soit a un élément nilpotent. Il existe un entier naturel n non nul tel que : $a^n = 0$.

Par ailleurs, d'après la propriété de factorisation dans un anneau énoncée un peu plus haut, on a :

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

On en déduit que : $(1 - a) \sum_{k=0}^{n-1} a^k = 1.$

4. On a déjà vu un peu plus haut quelques exemples de matrices carrées non nulles mais nilpotentes.

5. Le fait que a et b commutent est indispensable pour écrire la première égalité. Sans cette hypothèse, on a simplement : $(ab)^{n+m} = (ab)(ab) \cdots (ab).$

Par conséquent, $(1 - a)$ est inversible, et $(1 - a)^{-1} = \sum_{k=0}^{n-1} a^k$. ☕

Remarque. Evidemment, on aurait aussi pu énoncer la propriété : “si a est nilpotent, alors $(a - 1)$ est inversible”.

Exemple. On a déjà vu précédemment que la matrice $A = \begin{pmatrix} 0 & 2 & 1 \\ 0 & 0 & 4 \\ 0 & 0 & 0 \end{pmatrix}$ est nilpotente (car $A^3 = 0_{M_3(\mathbb{R})}$).

On déduit alors de la propriété ci-dessus que la matrice : $I_3 - A = \begin{pmatrix} 1 & -2 & -1 \\ 0 & 1 & -4 \\ 0 & 0 & 1 \end{pmatrix}$ est inversible, et que $(I_3 - A)^{-1} = I_3 + A + A^2$.

12.4 Corps et sous-corps

PROPRIÉTÉ 12.14 - Soit $(A, +, \times)$ un anneau (*non nul*). On note A^* l'ensemble des éléments inversibles de A .

Alors : $(A^*, \times)$ est un groupe. En outre, c'est un groupe abélien dès que A est commutatif.

PREUVE. Soit A un anneau (*non nul*), et A^* l'ensemble des éléments inversibles de A .

La loi “ $\times$ ” de A est une ℓ ci dans A^* . En effet, si a et b sont inversibles, alors ab l'est (et pour mémoire $(ab)^{-1} = b^{-1}a^{-1}$).

Cette ℓ ci est associative (puisque'elle l'est déjà dans A).

En outre, 1_A est inversible... puisque $1_A \times 1_A = 1_A$.

Enfin, si a est inversible, alors a^{-1} est encore inversible (d'inverse a).

En résumé, la loi “ $\times$ ” restreinte à A^* est une ℓ ci associative, possédant un élément neutre, et pour laquelle tout élément de A^* possède un inverse dans A^* .

Conclusion. $(A^*, \times)$ est un groupe.

Naturellement, lorsque A est commutatif, alors la loi $\times$ fait de A^* un groupe abélien. ☕

Terminologie. On dit parfois que $(A^*, \times)$ est le groupe (multiplicatif) des éléments inversibles de A .

Exemples :

1/ $\mathbb{Q}^* = \mathbb{Q} \setminus \{0\}$; $\mathbb{R}^* = \mathbb{R} \setminus \{0\}$; $\mathbb{C}^* = \mathbb{C} \setminus \{0\}$.

2/ $\mathbb{Z}^* = \{1, -1\}$; ainsi $\mathbb{Z}^* \subsetneq \mathbb{Z} \setminus \{0\}$

3/ $\mathbb{K}[X]^* = \mathbb{K}^*$; ainsi $\mathbb{K}[X]^* \subsetneq \mathbb{K}[X] \setminus \{0\}$

4/ $(\mathcal{C}^0(\mathbb{R}, \mathbb{R}))^*$ est l'ensemble des fonctions continues sur $\mathbb{R}$ qui ne s'annulent pas. Ainsi : $(\mathcal{C}^0(\mathbb{R}, \mathbb{R}))^* \subsetneq (\mathcal{C}^0(\mathbb{R}, \mathbb{R})) \setminus \{0\}$

5/ $M_2(\mathbb{R})^* = GL_2(\mathbb{R})$; ainsi $M_2(\mathbb{R})^* \subsetneq M_2(\mathbb{R}) \setminus \{0\}$

6/ $\mathbb{Q}[i]^* = \mathbb{Q}[i] \setminus \{0\}$; $\mathbb{Q}[\sqrt{2}]^* = \mathbb{Q}[\sqrt{2}] \setminus \{0\}$.

7/ $\mathbb{Z}[i]^* = \{\pm 1; \pm i\}$; ainsi $\mathbb{Z}[i]^* \subsetneq \mathbb{Z}[i] \setminus \{0\}$

DÉFINITION 12.14 - Un anneau commutatif $(A, +, \times)$ est un **corps** si $A^* = A \setminus \{0_A\}$.

Il revient au même dire qu'un corps est un anneau commutatif (*non nul*) où tout élément non nul est inversible.

Enfonçons une porte ouverte : un corps est en particulier un anneau intègre.

Exemples. Grâce aux exemples précédents, on peut affirmer que :

1/ $\mathbb{R}$ est un corps (le corps des réels) ; $\mathbb{C}$ est un corps (le corps des complexes) ; $\mathbb{Q}$ est un corps (le corps des rationnels).

2/ L'anneau des entiers $\mathbb{Z}$ n'est pas un corps.

3/ L'anneau des entiers de Gauss $\mathbb{Z}[i]$ n'est pas un corps.

4/ $\mathbb{Q}[i]$ et $\mathbb{Q}[\sqrt{2}]$ sont des corps.

5/ L'anneau $\mathbb{K}[X]$ des polynômes à coefficients dans $\mathbb{K}$ n'est pas un corps.

6/ L'anneau $\mathcal{C}^0(\mathbb{R}, \mathbb{R})$ des fonctions continues sur $\mathbb{R}$ et à valeurs réelles n'est pas un corps.

7/ L'anneau $\mathbb{C}^{\mathbb{N}}$ des suites complexes n'est pas un corps.

DÉFINITION 12.15 - Soit $(K, +, \times)$ un corps. Un **sous-corps de $\mathbb{K}$** est un sous-anneau $(K', +, \times)$ de K qui est un corps.

Exemple. $\mathbb{Q}$ est un sous-corps de $\mathbb{R}$ qui est un sous-corps de $\mathbb{C}$.

12.5 Epilogue

12.5.1 $\mathbb{Q}$ est le plus petit sous-corps de $\mathbb{R}$

PROPRIÉTÉ 12.15 - $\mathbb{Q}$ est le plus petit sous-corps de $\mathbb{R}$.

Explicitement, si K est un sous-corps de $\mathbb{R}$, alors : $\mathbb{Q} \subset K$.

PREUVE. Soit K un sous-corps de $\mathbb{R}$. Par définition, 0 et 1 appartiennent à K .

Puisque K est en particulier un sous-anneau de $\mathbb{R}$, il est en stable pour l'addition. Par suite : $2 = 1+1 \in K$. Une récurrence immédiate permet alors d'établir que : $\mathbb{N} \subset K$.

De plus, tout élément de K est inversible (dans K) pour l'addition, de nouveau car K est un sous-anneau de $\mathbb{R}$ (et donc un sous-groupe additif de $(\mathbb{R}, +)$). Grâce au résultat précédent, on en déduit que : $\mathbb{Z} \subset K$.

Enfin, tout élément non nul de K admet un inverse pour la multiplication dans K (puisque K est un corps), et K étant stable par produit, on en déduit que :

$$\forall (p, q) \in \mathbb{Z} \times \mathbb{N}^*, \quad \frac{p}{q} = \underbrace{p}_{\in K} \times \underbrace{\frac{1}{q}}_{\in K} \in K$$

Autrement écrit : $\mathbb{Q} \subset K$. On a donc établi que tout sous-corps de $\mathbb{R}$ contient $\mathbb{Q}$. ☕

12.5.2 Un théorème de Lagrange sur les groupes finis

Rappelons qu'un groupe fini est un groupe G de cardinal fini.

Notations. Soit $(G, *)$ un groupe. Pour tout élément g de G , on note $g^2 = g * g$; $g^3 = g * g * g$ et plus généralement pour tout entier $n \in \mathbb{N}$: $g^n = \underbrace{g * g * \cdots * g}_{n \text{ termes}}$.

Cette définition est étendue à un exposant entier relatif, en posant pour tout $n \in \mathbb{Z}_-$: $g^n = (g^{-1})^{-n}$.

On note par ailleurs : $\langle g \rangle = \{g^k / k \in \mathbb{Z}\}$.

PROPRIÉTÉ 12.16 - Soient $(G, *)$ un groupe.

Pour tout élément g de G , $\langle g \rangle$ est un sous-groupe de G , appelé **sous-groupe engendré par g** .

PREUVE. Soit g un élément de G . Montrons que $\langle g \rangle$ est un sous-groupe de G .

SG1 : $\langle g \rangle \subset G$ par définition.

SG2 : $\text{id} = g^0 \in \langle g \rangle$.

SG3 : soient γ et γ' deux éléments de $\langle g \rangle$. Par définition de $\langle g \rangle$, il existe deux entiers k et k' tels que : $\gamma = g^k$ et $\gamma' = g^{k'}$. Il s'ensuit que : $\gamma * \gamma' = g^{k+k'}$, d'où en particulier $\gamma * \gamma' \in \langle g \rangle$.

En résumé : $[(\gamma, \gamma') \in \langle g \rangle^2] \implies [(\gamma * \gamma') \in \langle g \rangle]$

SG4 : soit γ un élément de $\langle g \rangle$. Par définition de $\langle g \rangle$, il existe un entier k tel que : $\gamma = g^k$. Il s'ensuit que : $\gamma^{-1} = g^{-k}$, d'où en particulier $\gamma^{-1} \in \langle g \rangle$. En résumé : $[\gamma \in \langle g \rangle] \implies [\gamma^{-1} \in \langle g \rangle]$

Conclusion. Pour tout élément g de G , $\langle g \rangle$ est un sous-groupe de G . ☕

Deux exemples. Dans S_3 , considérons l'élément $g = (123)$.

On a : $(123)^0 = \text{id}$; $(123)^1 = (123)$; $(123)^2 = (132)$ et $(123)^3 = \text{id}$. Il s'ensuit que $\langle (123) \rangle = A_3$.

Dans $(\mathbb{Z}, +)$, le sous-groupe engendré par 2 est $2\mathbb{Z}$, le sous-groupe des entiers pairs.

DÉFINITION 12.16 - On dit qu'un élément g de G est **d'ordre fini** s'il existe un entier naturel N non nul tel que $g^N = e$.

Lorsque tel est le cas, on appelle **ordre de g** le plus petit N non nul tel que $g^N = e$.

Exemple 1. Dans $(S_4, \circ)$, on considère le 4-cycle $\sigma = (1234)$.

On a : $(1234)^1 = (1234)$; $(1234)^2 = (13)(24)$; $(1234)^3 = (1432)$ et $(1234)^4 = \text{id}$.

Il s'ensuit que le 4-cycle (1234) est d'ordre 4.

Exemple 2. Dans $(\text{GL}_2(\mathbb{R}), \times)$, on considère $g = 2I_2$. La matrice g est telle que : $\forall n \in \mathbb{N}$, $g^n = 2^n I_2$. Par conséquent : $\forall n \in \mathbb{N}^*$, $g^n \neq I_2$. Il s'ensuit que la matrice g n'est pas d'ordre fini dans $\text{GL}_2(\mathbb{R})$.

PROPRIÉTÉ 12.17 - Dans un groupe fini, tout élément est d'ordre fini.

PREUVE. Soient G un groupe fini, et g un élément de G . Notons $n = \text{Card}(G)$. On considère l'ensemble : $\{g^k / k \in \llbracket 0, n \rrbracket\}$. Si tous les éléments de cet ensemble étaient distincts, alors son cardinal serait $(n + 1)$; ceci est absurde, puisqu'il s'agit d'une partie de G , qui est de cardinal n par hypothèse.

Il existe donc deux entiers k et k' distincts dans $\llbracket 0, n \rrbracket$ tels que : $g^k = g^{k'}$. Sans nuire à la généralité on peut supposer que $k < k'$ (si ce n'est pas le cas, on permute les rôles de k et k'). On a alors : $g^0 = g^{k'-k}$, c'est à dire : $g^{k'-k} = e$.

En observant que $k' - k$ est un entier naturel non nul (puisque $k < k'$), on a établi l'existence d'un élément N de $\mathbb{N}^*$ tel que $g^N = e$. Ce qui signifie que g est d'ordre fini.

Conclusion. Tout élément d'un groupe fini est d'ordre fini . ☕

PROPRIÉTÉ 12.18 - Soit g un élément de G d'ordre N , et soit m un entier. On a :

$$g^m = e \text{ si et seulement si } N \text{ divise } m$$

PREUVE. Soit g un élément de G d'ordre N , et soit m un entier.

Supposons que N divise m : alors il existe un entier k tel que $m = kN$. Il s'ensuit que : $g^m = g^{kN} = (g^N)^k = e^k = e$. D'où : $[N \text{ divise } m] \implies [g^m = e]$.

Réciproquement, supposons que $g^m = e$. D'après le théorème de la division euclidienne, il existe un (unique) couple (q, r) tel que $m = Nq + r$ avec $q \in \mathbb{Z}$ et $r \in \llbracket 0, N - 1 \rrbracket$. On a alors : $g^{Nq+r} = g^{Nq} * g^r = e * g^r = g^r$; et donc $g^r = e$.

Si r était non nul, on aurait alors prouvé l'existence d'un élément r de $\mathbb{N}^*$ tel que $g^r = e$, avec r strictement inférieur à l'ordre de g : contradiction.

On en déduit que $r = 0$, ce qui implique que $m = Nq$, et donc que N divise m . D'où : $[g^m = e] \implies [N \text{ divise } m]$.

Conclusion. Pour g un élément d'ordre N on a : $[g^m = e] \iff [N \text{ divise } m]$. ☕

Jusqu'à la fin de ce paragraphe, on considérera un groupe fini G , et H un sous-groupe de G .

LEMME 12.1 - Pour tout élément g de G , on note : $gH = \{g * h / h \in H\}$. Alors :

1/ Pour tout élément g de G , on a : $\text{Card}(gH) = \text{Card}(H)$.

2/ Soient g et g' deux éléments de G . Montrer que les ensembles gH et $g'H$ sont soit égaux, soit disjoints.

PREUVE. 1/ Soit H un sous-groupe d'un groupe fini G . Observons que H est de cardinal fini, en tant que partie d'un ensemble fini.

Soit g un élément arbitraire de G .

Les applications $\varphi : H \longrightarrow gH$ et $\psi : gH \longrightarrow H$ sont clairement réciproques l'une de l'autre.

$$h \longmapsto g * h \qquad x \longmapsto g^{-1} * x$$

En particulier φ et ψ sont bijectives ; les ensembles H et gH sont donc équipotents. D'où : $\text{Card}(H) = \text{Card}(gH)$.

Conclusion. Pour tout élément g de G on a : $\text{Card}(H) = \text{Card}(gH)$.

2/ Soient g et g' deux éléments de G . Supposons que gH et $g'H$ ne sont pas disjoints. Alors il existe deux éléments h et k dans H tels que : $g * h = g' * k$. En particulier : $g^{-1} * g' = h * k^{-1} \in H$.⁶

Soit alors γ un élément quelconque de $g'H$. Il existe un élément h_0 de H tel que : $\gamma = g' * h_0$. On écrit alors judicieusement : $\gamma = g * g^{-1} * g' * h_0$. On en déduit que : $\gamma = g * \underbrace{h * k^{-1} * h_0}_{\in H}$. Par suite : $\gamma \in gH$.

On a ainsi établi que : $g'H \subset gH$.

En réécrivant le même raisonnement en permutant g et g' , ou en observant que g et g' jouent des rôles symétriques dans le raisonnement précédent, on obtient l'autre inclusion : $gH \subset g'H$. Finalement, lorsque gH et $g'H$ ne sont pas disjoints, ils sont égaux.

Conclusion. Pour tout couple (g, g') d'éléments de G , on a $[gH = g'H]$ ou $[gh \cap g'H = \emptyset]$. ☕

THÉORÈME 12.1 - (de Lagrange). Soient G un groupe fini, et H un sous-groupe de G .

Le cardinal de H divise le cardinal de G .

PREUVE. Le groupe G est fini par hypothèse. On peut donc noter : $G = \{g_1, \dots, g_n\}$ (avec $n = \text{Card}(G)$). On a alors :

$$G = \bigcup_{i=1}^n \{g_i\} \text{ et donc } G = \bigcup_{i=1}^n g_i H. \text{ La première égalité provient de l'observation puissante selon laquelle}$$

un ensemble fini est la réunion des singletons qui le composent ; la seconde découle de la première et du fait que pour tout entier i on a : $\{g_i\} \subset g_i H \subset G$.

D'après la question précédente, il peut exister parmi les $g_i H$ des ensembles égaux. Notons alors k le nombre de parties disjointes parmi ces $g_i H$. Quitte à renuméroter les g_i , on peut supposer qu'il s'agit des k premières parties, et on a alors :

$$G = \bigcup_{i=1}^k g_i H. \text{ Cette union étant disjointe, on a donc : } \text{Card}(G) = \sum_{i=1}^k \text{Card}(g_i H). \text{ Or}$$

d'après la question 6, on a : $\text{Card}(g_i H) = \text{Card}(H)$ (pour tout g_i).

On en déduit que : $\text{Card}(G) = \sum_{i=1}^k \text{Card}(H)$ d'où : $\text{Card}(G) = k \text{Card}(H)$. Finalement, le cardinal de G est multiple de celui de H .

Conclusion. Si G est un groupe fini, le cardinal de tout sous-groupe de G divise celui de G . ☕

COROLLAIRE 12.2 - Dans un groupe fini G , tout élément g est d'ordre fini, et l'ordre de g divise le cardinal de G .

PREUVE. Soit g un élément d'un groupe fini G . Alors g est d'ordre fini (résultat établi un peu plus haut), et son ordre est le cardinal de $\langle g \rangle$. Or $\langle g \rangle$ est un sous-groupe de G , donc son ordre divise celui de G (d'après le théorème de Lagrange).

Conclusion. Dans un groupe fini G , tout élément g est d'ordre fini, et l'ordre de g divise le cardinal de G . ☕

6. En effet, k est un élément de H . Puisque H est un sous-groupe, k^{-1} est un élément de H . Par ailleurs, h est un autre élément de H . En utilisant une nouvelle fois le fait que H est un sous-groupe, on en déduit que : $h * k^{-1} \in H$.

12.5.3 Les groupes finis de cardinal premier sont abéliens

Dans cette section, on prouve que les groupes finis dont le cardinal est un nombre premier sont abéliens, en utilisant de manière essentielle le théorème de Lagrange.

DÉFINITION 12.17 - Un groupe G fini est appelé **cyclique** s'il existe un élément g de G tel que $G = \langle g \rangle$.

Exemple. D'après la définition, il est équivalent de dire que G est cyclique ou qu'il existe un élément g de G dont l'ordre est égal au cardinal de G .

Dans le groupe S_3 , les éléments peuvent être d'ordre 1 (l'identité), d'ordre 2 (les trois transpositions) ou d'ordre 3 (les deux 3-cycles). Aucun élément n'a donc un ordre égal au cardinal de S_3 , qui vaut 6. Donc S_3 n'est pas cyclique.

En revanche le groupe A_3 est cyclique, puisque $A_3 = \langle (123) \rangle$ (ou $A_3 = \langle (132) \rangle$).

PROPRIÉTÉ 12.19 - Tout groupe cyclique est abélien.

PREUVE. Si G est un groupe cyclique, alors il existe un élément g de G tel que $G = \{g^k / k \in \mathbb{Z}\}$. Il est alors clair que G est abélien (essentiellement car $g^k * g^{k'} = g^{k+k'} = g^{k'} * g^k$).

Conclusion. $[G \text{ cyclique}] \implies [G \text{ abélien}]$. ☕

COROLLAIRE 12.3 - Soit G un groupe fini. Si le cardinal de G est un nombre premier, alors G est cyclique.

PREUVE. Soit G un groupe de cardinal p , avec p premier. Soit g un élément de G , distinct de l'élément neutre. On sait que g est d'ordre fini, et que cet ordre divise p . Ainsi l'ordre de g pourrait valoir 1 ou p ; mais $g \neq e$, donc g n'est pas d'ordre 1. Il s'ensuit que g est d'ordre p , ce qui signifie que : $G = \langle g \rangle$. D'où le groupe G est cyclique. ☕

COROLLAIRE 12.4 - Soit G un groupe fini. Si le cardinal de G est un nombre premier, alors G est abélien.

PREUVE. Soit G un groupe de cardinal p , avec p premier. D'après le corollaire précédent, G est cyclique, donc abélien d'après la propriété 12.19. ☕

COROLLAIRE 12.5 - Tout groupe de cardinal inférieur ou égal à 5 est abélien.

PREUVE. On a déjà établi en début de chapitre que tout groupe de cardinal au plus 4 est abélien, et le corollaire précédent assure que tout groupe de cardinal 5 est abélien, ce qui établit l'énoncé. ☕

12.5.4 Arithmétique dans l'anneau des entiers de Gauss

Dans cette section, on présente un sous-anneau de $\mathbb{C}$ particulier : l'anneau des entiers de Gauss. Cet anneau et celui des entiers relatifs⁷ possèdent beaucoup de points communs. L'objectif de cet énoncé est de montrer que, comme dans l'anneau des entiers relatifs, on peut construire dans $\mathbb{Z}[i]$ une division euclidienne. Cette propriété implique que tous les idéaux de $\mathbb{Z}[i]$ sont monogènes, ce qui permet de définir la notion de pgcd de deux entiers de Gauss puis d'entiers de Gauss premiers entre eux...

DÉFINITION 12.18 - L'anneau des entiers de Gauss $\mathbb{Z}[i]$ est l'anneau des nombres complexes à parties réelle et imaginaire entières. Explicitement :

$$\mathbb{Z}[i] = \{a + ib \mid (a, b) \in \mathbb{Z}^2\}$$

Par ailleurs, on utilise dans le problème ci-dessous une “norme”, l'application
$$\left[\begin{array}{l} N : \mathbb{C} \longrightarrow \mathbb{R}_+ \\ z \longmapsto z\bar{z} \end{array} \right]$$

Partie I — Quelques propriétés de l'anneau des entiers de Gauss

- 1) **Description des inversibles de $\mathbb{Z}[i]$.** Un entier de Gauss u est **inversible** s'il existe $v \in \mathbb{Z}[i]$ tel que : $uv = 1$.
 - a) Soient u et v deux éléments de $\mathbb{Z}[i]$.⁸ Montrer que : $N(uv) = N(u)N(v)$, et justifier que pour tout $u \in \mathbb{Z}[i]$, $N(u) \in \mathbb{N}$.
 - b) Montrer que si u est inversible, alors $N(u) = 1$.
 - c) Déterminer l'ensemble des éléments inversibles de $\mathbb{Z}[i]$.
- 2) **Divisibilité dans $\mathbb{Z}[i]$.** Soient u et v deux éléments de $\mathbb{Z}[i]$. On dit que u **divise** v dans $\mathbb{Z}[i]$ s'il existe $s \in \mathbb{Z}[i]$ tel que : $v = us$. Lorsque c'est le cas, on le note : $u|v$.
 - a) Montrer que si $u|v$, alors $N(u) | N(v)$ dans $\mathbb{Z}$.
 - b) Déterminer les diviseurs de $v_1 = 1 + i$ dans $\mathbb{Z}[i]$, puis ceux de $v_2 = 1 + 3i$.
- 3) **Division euclidienne dans $\mathbb{Z}[i]$.**
 - a) Montrer que pour tout $z \in \mathbb{C}$, il existe $u \in \mathbb{Z}[i]$ tel que $N(u - z) < 1$. Ce u est-il unique ?
 - b) Montrer que pour tout $u \in \mathbb{Z}[i]$ et pour tout $v \in \mathbb{Z}[i]^*$, il existe $(q, r) \in \mathbb{Z}[i] \times \mathbb{Z}[i]$ tel que $u = vq + r$ avec $N(r) < N(v)$.

7. Et l'anneau des polynômes en une indéterminée $\mathbb{K}[X]$.

8. Il reviendrait au même de dire : “Soient u et v deux entiers de Gauss”.

Partie II — Idéaux dans l'anneau des entiers de Gauss

Soient u et v dans $\mathbb{Z}[i]$. On note $I(u, v) = \{uz + vz' / z \text{ et } z' \in \mathbb{Z}[i]\}$.⁹

4) On suppose $(u, v) \neq (0, 0)$. Montrer que l'ensemble $A = \{N(w) / w \in I(u, v) \setminus \{0\}\}$ possède un plus petit élément $d > 0$.

5) Soit δ un élément de $I(u, v)$ tel que $N(\delta) = d$. Etablir que : $I(u, v) = \delta\mathbb{Z}[i]$.

On pourra utiliser la division euclidienne présentée en I.2.b.

6) Montrer que $\delta|u$ et $\delta|v$. Puis établir l'équivalence : $\forall w \in \mathbb{Z}[i], [w|u \wedge w|v] \iff [w|\delta]$

Terminologie : on dit que δ est **un pgcd** de u et v (dans $\mathbb{Z}[i]$).

Partie III — Entiers de Gauss premiers entre eux

Soient u et v dans $\mathbb{Z}[i]$, avec $u \neq 0$ ou $v \neq 0$.

On dit que u et v sont **premiers entre eux** si le nombre δ défini dans la partie précédente appartient à $\{\pm 1; \pm i\}$.

Dans les deux questions ci-dessous, on suppose que u et v sont deux entiers de Gauss premiers entre eux.

7) Justifier qu'il existe deux entiers de Gauss z et z' tels que : $1 = uz + vz'$.

8) Soit $w \in \mathbb{Z}[i]$. Montrer que si $u|vw$, alors $u|w$.

Partie IV — Entiers de Gauss irréductibles

Soit $u \in \mathbb{Z}[i] \setminus \{0; \pm 1; \pm i\}$.

On dit que u est **irréductible** si ses seuls diviseurs dans $\mathbb{Z}[i]$ sont $\pm 1, \pm i, \pm u$ et $\pm iu$.¹⁰

9) Soit $v \in \mathbb{Z}[i]$. On suppose que u est irréductible et ne divise pas v . Montrer que u et v sont premiers entre eux.

10) Soient v et $w \in \mathbb{Z}[i]$. On suppose que u est irréductible et divise vw . Montrer que $u|v$ ou $u|w$.

CORRIGÉ.

1) **Description des inversibles de $\mathbb{Z}[i]$.** a) Soient u et v dans $\mathbb{Z}[i]$. On a : $N(uv) = uv\overline{uv} = u\overline{u}v\overline{v} = N(u)N(v)$.

En outre : $\forall u \in \mathbb{Z}[i], \exists (a, b) \in \mathbb{Z}^2, u = a + ib$. D'où : $N(u) = a^2 + b^2 \in \mathbb{N}$.

Conclusion. $\forall (u, v) \in \mathbb{Z}[i] \times \mathbb{Z}[i], N(uv) = N(u)N(v)$ et $\forall u \in \mathbb{Z}[i], N(u) \in \mathbb{N}$.

b) Si $u \in \mathbb{Z}[i]$ est inversible, alors il existe $v \in \mathbb{Z}[i]$ tel que : $N(uv) = 1$. D'après la question précédente, on en déduit que $N(u)$ est un diviseur entier naturel de 1, ce qui implique : $N(u) = 1$.

c) D'après la question précédente, les entiers de Gauss u inversibles sont à rechercher parmi ceux vérifiant : $N(u) = 1$. Or $N(u) = 1 \implies \operatorname{Re}^2(u) + \operatorname{Im}^2(u) = 1$. Par suite : u est inversible si et seulement si $\operatorname{Re}(u) = \pm 1$ et $\operatorname{Im}(u) = 0$; ou $\operatorname{Re}(u) = 0$ et $\operatorname{Im}(u) = \pm 1$. En résumé : $\mathbb{Z}[i]^* = \{\pm 1; \pm i\}$.

9. Ou ce qui revient au même : $I(u, v) = u\mathbb{Z}[i] + v\mathbb{Z}[i]$.

10. La notion d'entier de Gauss irréductible étend à $\mathbb{Z}[i]$ celle de nombre premier dans $\mathbb{Z}$.

2) **Divisibilité dans $\mathbb{Z}[i]$.** a) Soient u et v dans $\mathbb{Z}[i]$, tels que u divise v dans $\mathbb{Z}[i]$. Alors il existe $s \in \mathbb{Z}[i]$ tel que : $v = us$. D'après la question 1-a, on en déduit que : $N(v) = N(u)N(s)$, ce qui signifie que $N(u)$ divise $N(v)$ dans $\mathbb{Z}$.

Conclusion. $\forall (u, v) \in \mathbb{Z}[i] \times \mathbb{Z}[i], (u|v) \implies (N(u) | N(v))$.

b) Soit $u \in \mathbb{Z}[i]$; u est un diviseur de $v_1 = 1 + i$ si $N(u) | N(v_1)$. Or $N(v_1) = 2$. Une condition nécessaire pour que $u|v_1$ est donc que $N(u) = 1$ ou 2 . Les entiers de Gauss vérifiant cette condition sont $\pm 1, \pm i$ et $\pm 1 \pm i$. Réciproquement, tous ces nombres sont des diviseurs de v_1 puisque :

$$v_1 = 1v_1; v_1 = (-1)(-v_1); v_1 = i(-iv_1); v_1 = (-i)(iv_1); v_1 = (1+i)1; v_1 = (1-i)i; v_1 = (-1-i)(-1); v_1 = (-1+i)(-i).$$

Conclusion. Les diviseurs de $1 + i$ dans $\mathbb{Z}[i]$ sont : $\pm 1, \pm i, \pm 1 \pm i$.

De façon analogue. Soit $u \in \mathbb{Z}[i]$; u est un diviseur de $v_2 = 1 + 3i$ si $N(u) | N(v_2)$. Or $N(v_2) = 10$. Une condition nécessaire pour que $u|v_2$ est donc que $N(u) = 1, 2, 5$ ou 10 . Les entiers de Gauss vérifiant cette condition sont $\pm 1, \pm i, \pm 1 \pm i, \pm 1 \pm 2i, \pm 2 \pm i, \pm 1 \pm 3i, \pm 3 \pm i$. Réciproquement, tous ces nombres sont des diviseurs de v_2 .

3) a) Soit $z \in \mathbb{C}$. Posons $a = \lfloor \operatorname{Re}(z) \rfloor$ et $a' = a + 1$. On a alors : $|\operatorname{Re}(z) - a| \leq \frac{1}{2}$ ou (non exclusif) $|\operatorname{Re}(z) - a'| \leq \frac{1}{2}$. Appelons A un des deux entiers relatifs (a ou a') vérifiant cette condition.

De même, appelons B l'un des deux entiers relatifs b ou b' vérifiant $|\operatorname{Im}(z) - b| \leq \frac{1}{2}$ ou $|\operatorname{Im}(z) - b'| \leq \frac{1}{2}$.

Alors le nombre complexe $u = A + iB$ est un entier de Gauss tel que : $N(z - u) \leq \left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^2 = \frac{1}{2}$ d'où en particulier : $N(z - u) < 1$. **Conclusion.** $\forall z \in \mathbb{C}, \exists u \in \mathbb{Z}[i], N(z - u) < 1$.

En général, on n'a pas unicité de l'entier de Gauss u : pour le complexe $z = \frac{1+i}{2}$, les quatre entiers de Gauss $0, 1, i$ et $1 + i$ conviennent.

b) Soient $u \in \mathbb{Z}[i]$ et $v \in \mathbb{Z}[i]^*$. On pose $z = \frac{u}{v}$, et on choisit $q \in \mathbb{Z}[i]$ tel que : $N\left(\frac{u}{v} - q\right) < 1$.¹¹ On pose alors : $r = u - qv$. On a : $u = qv + r$, et $N(r) = N\left(\frac{u}{v} - q\right)N(v)$ donc $N(r) < N(v)$ (l'inégalité est stricte puisque $N(v)$ ne peut être nul).

Conclusion. $\forall (u, v) \in \mathbb{Z}[i] \times \mathbb{Z}[i]^*, \exists (q, r) \in \mathbb{Z}[i] \times \mathbb{Z}[i], [(u = vq + r) \wedge (N(r) < N(v))]$.

Remarque : on a ainsi prouvé qu'il existe une "pseudo"-division euclidienne dans $\mathbb{Z}[i]$. L'originalité de cette division par rapport à celle de $\mathbb{Z}$ ou de $\mathbb{K}[X]$ est que l'on n'a pas unicité du quotient ni du reste, en vertu de la question précédente.

Partie II — Idéaux dans l'anneau des entiers de Gauss

4) L'ensemble $A = \{N(w) / w \in I(u, v) \setminus \{0\}\}$ est non vide sous réserve que $(u, v) \neq (0, 0)$. En effet, dès lors que $u \neq 0$ on a $u \in I(u, v)$ et $N(u) > 0$ (de même si $v \neq 0$). De plus, A est une partie de $\mathbb{N}^*$; elle possède donc un plus petit élément $\min A$. On pose alors : $d = \min A > 0$.

11. On continue de noter $N(Z)$ l'image d'un complexe Z quelconque par l'application N , c'est à dire : $\forall Z \in \mathbb{C}, N(Z) = Z\bar{Z}$.

5) Soient u et v deux entiers de Gauss, non simultanément nuls. Notons δ un élément de $I(u, v)$ tel que $N(\delta) = d$. Notons que $\delta \neq 0$ puisque $(u, v) \neq (0, 0)$.

Montrons que : $\delta\mathbb{Z}[i] \subset I(u, v)$. Puisque δ est un élément de $I(u, v)$, il existe un couple d'entiers de Gauss (a, b) tel que : $\delta = au + bv$. Considérons alors un élément x quelconque de $\delta\mathbb{Z}[i]$. Il existe $c \in \mathbb{Z}[i]$ tel que : $x = c\delta$. Par suite : $x = acu + bcv$, d'où : $x \in I(u, v)$. D'où : $\delta\mathbb{Z}[i] \subset I(u, v)$.

Réciproquement, montrons que : $I(u, v) \subset \delta\mathbb{Z}[i]$. Soit x un élément de $I(u, v)$. D'après la question 3-b, il existe un couple d'entiers de Gauss (q, r) tel que : $x = q\delta + r$, avec $N(r) < N(\delta)$.

Raisonnons par l'absurde et supposons que $r \neq 0$. Alors r est un élément de $I(u, v)$ (puisque $r = x - \delta q$ et que x et δ sont dans $I(u, v)$), pour lequel : $0 < N(r) < N(\delta)$. Ce qui contredit la minimalité de $N(\delta)$. Par suite : $r = 0$. Donc : $x = q\delta$, ce qui assure que $x \in \delta\mathbb{Z}[i]$. D'où : $I(u, v) \subset \delta\mathbb{Z}[i]$.

D'après la règle de double inclusion : $I(u, v) = \delta\mathbb{Z}[i]$.

Remarque : ce résultat permet de montrer que tout idéal de $\mathbb{Z}[i]$ est monogène ; on dit que l'anneau des entiers de Gauss est principal. On peut observer que le résultat tient encore dans le cas où u et v sont simultanément nuls, puisqu'alors : $I(u, v) = 0\mathbb{Z}[i] = \{0\}$.

6) Notons que $u \in I(u, v)$. D'après la question précédente, on a donc : $u \in \delta\mathbb{Z}[i]$. Par suite, il existe un entier de Gauss a tel que : $u = \delta a$. En particulier : $\delta|u$. Le raisonnement s'appliquant également à v , on peut conclure : $\delta|u$ et $\delta|v$.

Soit w un entier de Gauss. Supposons que $w|\delta$. Puisqu'en outre $\delta|u$, la transitivité de la divisibilité permet d'affirmer que : $w|u$. On montre de façon analogue que $w|v$. Par suite : $[w|\delta] \implies [w|u \wedge w|v]$.

Réciproquement, supposons que $w|u$ et $w|v$. Alors il existe deux entiers de Gauss a et b tels que : $u = aw$ et $v = bw$. De plus, $\delta \in \delta\mathbb{Z}[i]$ donc d'après la question précédente, $\delta \in I(u, v)$; il existe donc deux entiers de Gauss c et d tels que $\delta = cu + dv$. Par suite : $\delta = caw + dbw$ d'où $\delta = w(ca + db)$ donc : $w|\delta$. En résumé : $[w|u \wedge w|v] \implies [w|\delta]$.

Conclusion. $[w|u \wedge w|v] \iff [w|\delta]$.

Partie III — Entiers de Gauss premiers entre eux

7) Supposons que z et z' soient deux entiers de Gauss premiers entre eux. En reprenant les notations de la partie précédente, on a donc $\delta = \pm 1$ ou $\pm i$.

Par ailleurs, puisque $I(u, v) = \delta\mathbb{Z}[i]$ et que $\delta \in \delta\mathbb{Z}[i]$, il existe deux entiers de Gauss a et b tels que : $\delta = au + bv$.

En multipliant cette dernière égalité par $\bar{\delta}$ on obtient : $a(\bar{\delta}u) + b(\bar{\delta}v) = 1$. Il reste à voir que $\bar{\delta}u$ et $\bar{\delta}v$ sont deux entiers de Gauss, d'une part car l'anneau $\mathbb{Z}[i]$ est stable par conjugaison, et d'autre part car il est stable par produit (le produit étant une loi de composition interne dans $\mathbb{Z}[i]$).

En posant donc : $z = \bar{\delta}u$ et $z' = \bar{\delta}v$, on peut conclure :

Si u et v sont deux entiers de Gauss premiers entre eux, alors : $\exists (z, z') \in \mathbb{Z}[i]^2, uz + vz' = 1$.

8) Soit w un entier de Gauss tel que $u|vw$. D'après la question précédente, il existe un couple (z, z') d'entiers de Gauss tel que : $uz + vz' = 1$. D'où : $uwz + vwz' = w$. Il est immédiat que $u|uwz$, et par hypothèse on peut affirmer que $u|vwz'$. Donc u divise la somme de ces deux entiers, c'est à dire : $u|w$.

Conclusion. Si u et v sont deux entiers de Gauss premiers entre eux, alors : $[u|vw] \implies [u|w]$.

Remarque : ce résultat est à mettre en parallèle avec le lemme de Gauss dans $\mathbb{Z}$: "soient a, b et c trois entiers. Si $a|bc$ et a est premier avec b , alors $a|c$ ".

Partie IV — Entiers de Gauss irréductibles

9) Soit d un entier de Gauss, diviseur commun à u et v . Puisque d est en particulier un diviseur de u , et que u est irréductible, on a : $d = \pm 1$ ou $\pm i$ ou $\pm u$ ou $\pm iu$.

Supposons que : $d = \pm u$ ou $\pm iu$. Alors $\pm u$ ou $\pm iu$ serait un diviseur de v (puisque $d|v$), ce qui impliquerait $u|v$: contradiction. Par suite : $d = \pm 1$ ou $\pm i$; u et v sont donc premiers entre eux.

Conclusion. Si u est un entier de Gauss irréductible, et si u ne divise pas $v \in \mathbb{Z}[i]$, alors u et v sont premiers entre eux.

Remarque : énoncé à mettre en parallèle avec celui connu dans $\mathbb{Z}$: “un nombre premier est premier avec tout entier qu’il ne divise pas”.

10) Soit u un entier de Gauss irréductible, et soit v un entier de Gauss. Prouvons la propriété de l’énoncé par disjonction des cas.

Si $u|v$: l’assertion est prouvée.

Si u ne divise pas v : d’après la question précédente, u et v sont premiers entre eux. D’après la question 7, il existe alors deux entiers de Gauss z et z' tels que : $uz + vz' = 1$. D’où : $uwz + vwz' = w$. Il est alors immédiat que $u|uwz$, et $u|vwz'$ (par hypothèse). Donc $u|uwz + vwz'$, soit : $u|w$.

Conclusion. Soient v et w dans $\mathbb{Z}[i]$. Si u est un entier de Gauss irréductible, et si $u|vw$, alors $u|v$ ou $u|w$.

Remarque : énoncé à mettre en parallèle avec celui connu dans $\mathbb{Z}$: “soient a et b deux entiers, et p un nombre premier. Si $p|ab$, alors $p|a$ ou $p|b$ ”.

Chapitre 13

Calcul matriciel

Sauf mention explicite du contraire, n et p désignent dans ce chapitre deux entiers naturels non nuls, et $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$.

13.1 Les ensembles de matrices

13.1.1 Généralités

DÉFINITION 13.1 - Une **matrice à n lignes et p colonnes à coefficients dans $\mathbb{K}$** est une famille $(a_{ij})_{i \in \llbracket 1, n \rrbracket; j \in \llbracket 1, p \rrbracket}$ d'éléments de $\mathbb{K}$.

L'ensemble de ces matrices est noté $M_{np}(\mathbb{K})$.

On peut se représenter une matrice $A = (a_{ij})_{i \in \llbracket 1, n \rrbracket; j \in \llbracket 1, p \rrbracket}$ de $M_{np}(\mathbb{K})$ comme un tableau de scalaires à n lignes et p colonnes :

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{np} \end{pmatrix}$$

Exemples et cas particuliers :

$$1/ \quad A = \begin{pmatrix} 1 & 2 & 4 \\ 3 & 5 & 6 \end{pmatrix} \in M_{23}(\mathbb{K}); \quad B = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \\ 7 & 8 \end{pmatrix} \in M_{42}(\mathbb{K}); \quad \text{et } C = \begin{pmatrix} 1 & 2 & 3 \\ 5 & 6 & 7 \\ 7 & 8 & 9 \end{pmatrix} \in M_3(\mathbb{K}).^1$$

2/ La **matrice nulle** de $M_{np}(\mathbb{K})$ est celle dont tous les coefficients sont nuls ; pour éviter les confusions, elle est notée $0_{M_{np}(\mathbb{K})}$.

1. Lorsque n et p sont égaux, on notera $M_n(\mathbb{K})$ plutôt que $M_{nn}(\mathbb{K})$.

3/ La matrice $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ est appelée **matrice identité** de $M_2(\mathbb{K})$, et est notée I_2 . La matrice $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ est appelée **matrice identité** de $M_3(\mathbb{K})$, et est notée I_3 . Plus généralement, la matrice identité de $M_n(\mathbb{K})$, notée I_n , est une matrice carrée à n lignes et n colonnes dont les coefficients diagonaux sont égaux à 1, et les autres sont nuls.

4/ Une **matrice colonne** (*resp.* **ligne**) est un élément de $M_{n,1}(\mathbb{K})$ (*resp.* $M_{1,p}(\mathbb{K})$).

13.1.2 Opérations sur les matrices

Remarque. Jusqu'à la fin de ce chapitre, pour une matrice $A \in M_{np}(\mathbb{K})$, on notera $A = (a_{ij})$ plutôt que $A = (a_{ij})_{i \in [1,n], j \in [1,p]}$.

13.1.2.1 Somme de deux matrices de $M_{np}(\mathbb{K})$

DÉFINITION 13.2 - Soient $A = (a_{ij})$ et $B = (b_{ij})$ deux matrices de $M_{np}(\mathbb{K})$. On définit la **somme des matrices A et B** et on note $A + B$ la matrice de $M_{np}(\mathbb{K})$ telle que :

$$\forall (i, j) \in [1, n] \times [1, p], (A + B)_{ij} = a_{ij} + b_{ij}$$

Exemple : $\begin{pmatrix} 1 & 2 & 4 \\ 3 & 5 & 6 \end{pmatrix} + \begin{pmatrix} 0 & 3 & -1 \\ 2 & 1 & -2 \end{pmatrix} = \begin{pmatrix} 1 & 5 & 3 \\ 5 & 6 & 4 \end{pmatrix}$

PROPRIÉTÉ 13.1 - Soient A, B et C trois matrices quelconques de $M_{np}(\mathbb{K})$.

1/ **LCI** : $A + B \in M_{np}(\mathbb{K})$

2/ **Associativité** : $(A + B) + C = A + (B + C)$

3/ **Elément neutre** : $A + 0_{M_{np}(\mathbb{K})} = A = 0_{M_{np}(\mathbb{K})} + A$

4/ **Inverse** : $A + (-A) = 0_{M_{np}(\mathbb{K})} = (-A) + A$ (en ayant noté $(-A) = (-a_{ij}) \in M_{np}(\mathbb{K})$)

5/ **Commutativité** : $A + B = B + A$

PREUVE. Ces différentes propriétés proviennent soit de la définition de l'addition dans $M_{np}(\mathbb{K})$ (pour la première), soit de vérifications immédiates provenant des propriétés de l'addition dans $\mathbb{K}$. ☕

En d'autres termes :

PROPRIÉTÉ 13.2 - $(M_{np}(\mathbb{K}), +)$ est un groupe abélien.

Remarque : pour $A \in M_{np}(\mathbb{K})$, la matrice $(-A)$ est appelée **matrice opposée** de A .

13.1.2.2 Multiplication par un scalaire dans $M_{np}(\mathbb{K})$

DÉFINITION 13.3 - Soient $A = (a_{ij})$ et $\lambda \in \mathbb{K}$. On définit la matrice λA comme la matrice de $M_{np}(\mathbb{K})$ telle que :

$$\forall (i, j) \in \llbracket 1, n \rrbracket \times \llbracket 1, p \rrbracket, (\lambda A)_{ij} = \lambda a_{ij}$$

Exemple : $3 \begin{pmatrix} 1 & 2 & 4 \\ 3 & 5 & 6 \end{pmatrix} = \begin{pmatrix} 3 & 6 & 12 \\ 9 & 15 & 18 \end{pmatrix}$

PROPRIÉTÉ 13.3 - Soient A et B dans $M_{np}(\mathbb{K})$, λ et μ dans $\mathbb{K}$:

1/ $\lambda A \in M_{np}(\mathbb{K})$

2/ $1A = A$; $0A = 0_{M_{np}(\mathbb{K})}$; $(-1)A = -A$

3/ $(\lambda + \mu)A = \lambda A + \mu A$

4/ $\lambda(\mu A) = (\lambda\mu)A$

5/ $\lambda(A + B) = \lambda A + \lambda B$

PREUVE. De nouveau, ces différentes propriétés proviennent soit de la définition de la multiplication par un scalaire (pour la première), soit de vérifications immédiates provenant des propriétés de la multiplication dans $\mathbb{K}$. ☕

Remarque (“pour l’avenir”) : les propriétés de la somme et de la multiplication par un scalaire évoquées plus haut permettront d’affirmer que $M_{np}(\mathbb{K})$ est un $\mathbb{K}$ -**espace vectoriel**.

➤ **Un mot sur les combinaisons linéaires.** Si $(A_k)_{k \in \llbracket 1, m \rrbracket}$ est une famille d’éléments de $M_{np}(\mathbb{K})$, une **combinaison linéaire** de ces matrices s’écrit $\sum_{k=1}^m \lambda_k A_k$, les λ_k étant des scalaires quelconques. Par exemple, toute matrice de $M_{2,2}(\mathbb{K})$ s’écrit comme combinaison linéaire de quatre matrices (appelées matrices *élémentaires*), explicitement :

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = a \underbrace{\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}}_{E_{11}} + b \underbrace{\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}}_{E_{12}} + c \underbrace{\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}}_{E_{21}} + d \underbrace{\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}}_{E_{22}}$$

Dans cette situation, nous dirons un peu plus tard que la famille $\{E_{11}, E_{12}, E_{21}, E_{22}\}$ est une *base* du $\mathbb{K}$ -espace vectoriel $M_2(\mathbb{K})$.

13.1.2.3 Produit matriciel

DÉFINITION 13.4 - Soient $A \in M_{np}(\mathbb{K})$ et $B \in M_{pq}(\mathbb{K})$.

On définit la **matrice produit** $A \times B$ (ou AB) comme la matrice de $M_{nq}(\mathbb{K})$ telle que :

$$\forall (i, j) \in \llbracket 1, n \rrbracket \times \llbracket 1, q \rrbracket, (AB)_{ij} = \sum_{k=1}^p a_{ik} b_{kj}$$

Remarque 1. Dans un sens qui sera rendu plus précis au second semestre, le coefficient $(AB)_{ij}$ est le produit scalaire de la i -ème ligne de A et de la j -ème colonne de B .

Remarque 2. Le produit de deux matrices $A \times B$ est défini si et seulement si le nombre de colonnes de A est égal au nombre de lignes de B . Explicitement, en posant :

$$A = \begin{pmatrix} 1 & 2 & 4 \\ 3 & 5 & 6 \end{pmatrix} \quad \text{et} \quad B = \begin{pmatrix} 1 & 1 & 0 \\ 0 & -1 & 1 \\ 1 & 0 & 0 \end{pmatrix}$$

le produit AB est défini tandis que le produit BA ne l'est pas !

Au passage d'ailleurs, on a : $AB = \begin{pmatrix} 5 & -1 & 2 \\ 9 & -2 & 5 \end{pmatrix}$.

Il se peut encore que les produits AB et BA soient simultanément définis, mais ne soient pas de même format. Pour illustrer cette affirmation, considérons les matrices :

$$A = \begin{pmatrix} 1 & 2 & 4 \end{pmatrix} \quad \text{et} \quad B = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$$

Dans cette situation, les produits AB et BA existent, mais : $AB \in M_1(\mathbb{K})$ tandis que $BA \in M_3(\mathbb{K})$. Explicitement :

$$AB = \begin{pmatrix} 5 \end{pmatrix} \quad \text{et} \quad BA = \begin{pmatrix} 1 & 2 & 4 \\ 0 & 0 & 0 \\ 1 & 2 & 4 \end{pmatrix}$$

En résumé, dans un produit matriciel, il faut **faire très attention à l'ordre!!!**

Remarque 3. Lorsque le produit AB est défini, la matrice AB hérite du nombre de lignes de A et du nombre de colonnes de B .

PROPRIÉTÉ 13.4 - (**Propriétés du produit matriciel**). Sous réserve que les produits suivants soient définis, on a :

1/ $A(BC) = (AB)C$ (associativité)

2/ $A(B + C) = AB + AC$ (distributivité par rapport à la somme)

3/ $(\lambda A)B = \lambda(AB) = A(\lambda B)$ (compatibilité avec la multiplication par un scalaire)

PREUVE. Les preuves de ces propriétés reposent encore une fois sur des vérifications triviales, mais lourdes à écrire. Pour ne pas botter une nouvelle fois en touche, on donne ci-dessous la preuve de l'associativité (les deux autres propriétés se démontrent de manière analogue, mais un peu plus "légère").

Soient $A \in M_{np}(\mathbb{K})$, $B \in M_{pq}(\mathbb{K})$, et $C \in M_{qr}(\mathbb{K})$.

Dans ce contexte, le produit AB est défini, et $AB \in M_{nq}(\mathbb{K})$. Par suite, le produit $(AB)C$ est défini (puisque le nombre de colonnes de (AB) est égal au nombre de lignes de C), et $(AB)C \in M_{nr}(\mathbb{K})$.

De même, le produit BC est défini, et $BC \in M_{pr}(\mathbb{K})$. Par suite, le produit $A(BC)$ est défini (puisque le nombre de colonnes de A est égal au nombre de lignes de (BC)), et $A(BC) \in M_{nr}(\mathbb{K})$.

Posons $P = (AB)C$ et $Q = A(BC)$, et montrons que $P = Q$ ($\in M_{nr}(\mathbb{K})$).

Soient $i \in \llbracket 1, n \rrbracket$ et $j \in \llbracket 1, r \rrbracket$ deux entiers arbitraires. On a d'une part :

$$P_{ij} = \sum_{k=1}^q (AB)_{ik} c_{kj} = \sum_{k=1}^q \sum_{m=1}^p a_{im} b_{mk} c_{kj}$$

Et d'autre part :

$$Q_{ij} = \sum_{k=1}^p a_{ik} (BC)_{kj} = \sum_{k=1}^p a_{ik} \sum_{m=1}^q b_{km} c_{mj} = \sum_{k=1}^p \sum_{m=1}^q a_{ik} b_{km} c_{mj}$$

Le changement d'indice " $k \longleftrightarrow m$ " permet de conclure que : $P_{ij} = Q_{ij}$. En résumé, on a établi que :

$$(AB)C \in M_{nr}(\mathbb{K}) \text{ et } A(BC) \in M_{nr}(\mathbb{K}), \text{ et } : \forall (i, j) \in \llbracket 1, n \rrbracket \times \llbracket 1, r \rrbracket, [(AB)C]_{ij} = [A(BC)]_{ij}$$

Ce qui prouve l'associativité du produit matriciel (encore une fois, sous réserve de définition). ☕

Remarque. Attention !!! Le produit matriciel n'est pas commutatif.

Un type d'exemple-classique à connaître pour illustrer cette affirmation : on pose

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} \quad \text{et} \quad B = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

On a alors :

$$AB = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad \text{tandis que} \quad BA = \begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix}$$

La raison pour laquelle cette situation est classique est que la matrice A est diagonale avec des coefficients distincts sur la diagonale, et que les seules matrices $M \in M_2(\mathbb{K})$ commutant avec A (càd telles que $AM = MA$) sont les matrices diagonales (comme nous le verrons un peu plus tard).

13.1.2.4 Transposée d'une matrice

DÉFINITION 13.5 - Soit $A \in M_{np}(\mathbb{K})$. La **transposée** de A , notée ${}^tA$ est la matrice de $M_{pn}(\mathbb{K})$ définie en posant : $\forall i \in \llbracket 1; p \rrbracket, \forall j \in \llbracket 1; n \rrbracket, ({}^tA)_{ij} = A_{ji}$.

Informellement, la transposée de A est obtenue en “écrivant en colonnes les lignes de A ”.

Exemple. En posant $A = \begin{pmatrix} 1 & 2 & 4 \\ 3 & 5 & 6 \end{pmatrix}$, on a : ${}^tA = \begin{pmatrix} 1 & 3 \\ 2 & 5 \\ 4 & 6 \end{pmatrix}$.

PROPRIÉTÉ 13.5 -

1/ $\forall (\lambda, \mu) \in \mathbb{K}^2, \forall (A, B) \in M_{np}(\mathbb{K})^2, {}^t(\lambda A + \mu B) = \lambda {}^tA + \mu {}^tB$ (linéarité)

2/ $\forall A \in M_{np}(\mathbb{K}), {}^t({}^tA) = A$

3/ $\forall A \in M_{np}(\mathbb{K}), \forall B \in M_{pq}(\mathbb{K}), {}^t(AB) = {}^tB {}^tA$.

PREUVE. Les preuves des points 1 et 2 sont immédiates.

Etablissons le point 3. Soient $A \in M_{np}(\mathbb{K})$ et $B \in M_{pq}(\mathbb{K})$. Le produit AB est bien défini, et appartient à $M_{nq}(\mathbb{K})$. De même, le produit ${}^tB {}^tA$ est défini, et appartient à $M_{qn}(\mathbb{K})$.

Soient $i \in \llbracket 1, q \rrbracket$ et $j \in \llbracket 1, n \rrbracket$ deux entiers arbitraires. On a d'une part :

$$[{}^t(AB)]_{ij} = (AB)_{ji} = \sum_{k=1}^p a_{jk} b_{ki}$$

Et d'autre part :

$$({}^tB {}^tA)_{ij} = \sum_{k=1}^p ({}^tB)_{ik} ({}^tA)_{kj} = \sum_{k=1}^p b_{ki} a_{jk}$$

D'où : $\forall (i, j) \in \llbracket 1, q \rrbracket \times \llbracket 1, n \rrbracket, [{}^t(AB)]_{ij} = ({}^tB {}^tA)_{ij}$.

Ce qui prouve que : ${}^t(AB) = {}^tB {}^tA$ (égalité à lire dans $M_{qn}(\mathbb{K})$). ☕

13.2 Matrices carrées

13.2.1 Généralités

DÉFINITION 13.6 - Une matrice est **carrée** lorsqu'elle a autant de lignes que de colonnes (“ $n = p$ ”).

Le groupe abélien des matrices carrées à n lignes et n colonnes (de taille n) à coefficients dans $\mathbb{K}$ est noté $M_n(\mathbb{K})$ (plutôt que $M_{n,n}(\mathbb{K})$).

DÉFINITION 13.7 - La matrice **identité** de $M_n(\mathbb{K})$, notée I_n , est la matrice n'ayant que des 1 sur sa diagonale, des 0 partout ailleurs. Plus précisément :

$$\forall (i, j) \in ([1; n])^2, (I_n)_{ij} = 1 \text{ si } i = j, \text{ et } 0 \text{ sinon.}$$

Notation. Pour faire plus court, on peut réécrire la définition précédente en utilisant le **symbole de Kronecker** :

$$\forall (i, j) \in ([1; n])^2, (I_n)_{ij} = \delta_{ij}$$

Le symbole de Kronecker δ_{ij} est précisément défini comme l'entier égal à 1 lorsque $i = j$, nul sinon.

PROPRIÉTÉ 13.6 - La matrice I_n est l'élément neutre pour la multiplication dans $M_n(\mathbb{K})$, càd :

$$\forall A \in M_n(\mathbb{K}), AI_n = I_n A = A$$

PREUVE. Soient n un entier naturel non nul, et $A = (a_{ij})$ une matrice de $M_n(\mathbb{K})$.

Notons $P = (p_{ij})$ la matrice produit $A \times I_n$. On rappelle que $\forall (i, j) \in [1, n]^2, (I_n)_{ij} = \delta_{ij}$.

Soient i et j deux entiers de $[1, n]$. On a : $P_{ij} = \sum_{k=1}^n a_{ik} (I_n)_{kj}$.

Or $(I_n)_{kj} = 0$ pour $k \neq j$, et $(I_n)_{jj} = 1$. Il s'ensuit que : $P_{ij} = a_{ij}$.

En résumé : $\forall (i, j) \in [1, n]^2, P_{ij} = a_{ij}$. Donc $P = A$, et donc : $A \times I_n = A$.

“Din l'aut'sinse” : notons $Q = (q_{ij})$ la matrice produit $I_n \times A$.

Soient i et j deux entiers de $[1, n]$. On a : $Q_{ij} = \sum_{k=1}^n (I_n)_{ik} a_{kj}$.

Or $(I_n)_{ik} = 0$ pour $k \neq i$, et $(I_n)_{ii} = 1$. Il s'ensuit que : $Q_{ij} = a_{ij}$.

En résumé : $\forall (i, j) \in [1, n]^2, Q_{ij} = a_{ij}$. Donc $Q = A$, et donc : $I_n \times A = A$. ☕

COROLLAIRE 13.1 - La matrice I_n , et plus généralement la matrice λI_n ($\lambda \in \mathbb{K}$) commutent avec toute matrice de $M_n(\mathbb{K})$, dans le sens où :

$$\forall \lambda \in \mathbb{K}, \forall A \in M_n(\mathbb{K}), A(\lambda I_n) = (\lambda I_n)A = \lambda A$$

PREUVE. C'est une conséquence immédiate de la propriété précédente. ☕

PROPRIÉTÉ 13.7 - $(M_n(\mathbb{K}), +, \times)$ est un anneau (**non commutatif et non intègre** pour $n \geq 2$).

PREUVE. On a déjà établi que $(M_n(\mathbb{K}), +, \times)$ est un groupe abélien. En outre, la multiplication dans $M_n(\mathbb{K})$ est distributive par rapport à l'addition (voir propriété 13.4 page 304) ; et elle possède un élément neutre, qui est I_n d'après une récente propriété.

Il reste à vérifier que pour toute matrice $A \in M_n(\mathbb{K})$, on a : $0_{M_n(\mathbb{K})} A = 0_{M_n(\mathbb{K})} = A 0_{M_n(\mathbb{K})}$, ce qui est une démonstration encore plus facile qu'inintéressante, à moins que ce ne soit le contraire.

On en déduit que $(M_n(\mathbb{K}), +, \times)$ est un anneau.

Finalement, il est plus intéressant de s'intéresser aux “anti-propriétés” de l'énoncé : l'anneau $M_n(\mathbb{K})$ n'est ni commutatif, ni intègre (pour $n \geq 2$). Considérons pour illustrer ces affirmations les matrices suivantes :

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \quad \text{et} \quad B = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

On a alors :

$$A \neq 0_{M_2(\mathbb{K})}; \quad B \neq 0_{M_2(\mathbb{K})}; \quad AB = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}; \quad BA = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

On peut en conclure que $M_2(\mathbb{K})$ n'est ni commutatif ($AB \neq BA$), ni intègre ($BA = 0_{M_2(\mathbb{K})}$ avec A et B non nulles). Cet exemple se généralise à $M_n(\mathbb{K})$ avec $n \geq 2$ (il suffit de rajouter suffisamment de zéros!). ☕

On achève ce paragraphe en insistant une dernière fois sur les dangers du produit matriciel.

ATTENTION

Pour deux matrices A et B carrées (et de même taille) :

$$AB \neq BA \quad (\text{en général})$$

$$\text{On peut avoir } A \neq 0, B \neq 0 \text{ mais } AB = 0$$

$$\text{On peut avoir } AB = 0 \text{ et } BA \neq 0$$

Voir plus haut pour des illustrations de chaque situation.

Conséquences : les simplifications sont interdites ($AC = BC$ n'implique pas $A = B$), et les identités remarquables ne sont plus valides en général $((A + B)^2 = A^2 + AB + BA + B^2)$.

13.2.2 (Quelques) sous-anneaux remarquables de $M_n(\mathbb{K})$

DÉFINITION 13.8 - Une matrice carrée est **diagonale** lorsque tous ses coefficients sont nuls en dehors de la diagonale.

Explicitement, une matrice diagonale est une matrice $D \in M_n(\mathbb{K})$ (pour un certain entier naturel n non nul) telle que :

$$\forall (i, j) \in \llbracket 1, n \rrbracket^2, [i \neq j] \implies [d_{ij} = 0]$$

Exemples. Dans $M_n(\mathbb{K})$, la matrice nulle et la matrice identité sont diagonales.

Dans $M_3(\mathbb{K})$, les matrices $\begin{pmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 3 \end{pmatrix}$ et $\begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ sont diagonales.

DÉFINITION 13.9 - Une matrice carrée D est **scalaire** s'il existe un scalaire λ tel que $D = \lambda I_n$. Explicitement, une matrice scalaire est une matrice $D \in M_n(\mathbb{K})$ (pour un certain entier naturel n non nul) telle que :

$$\exists \lambda \in \mathbb{K}, \forall (i, j) \in \llbracket 1, n \rrbracket^2, d_{ij} = \lambda \delta_{ij}$$

Notation. Si $\lambda_1, \dots, \lambda_n$ désignent n scalaires, on notera $\text{diag}(\lambda_1, \dots, \lambda_n)$ la matrice diagonale de $M_n(\mathbb{K})$ ayant pour coefficients diagonaux les λ_i .

Explicitement : $\text{diag}(\lambda_1, \dots, \lambda_n) = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$

Avec cette notation, une matrice scalaire est donc de la forme : $\text{diag}(\lambda, \dots, \lambda) = \begin{pmatrix} \lambda & 0 & \cdots & 0 \\ 0 & \lambda & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda \end{pmatrix}$

DÉFINITION 13.10 - Une matrice carrée $A \in M_n(\mathbb{K})$ est **triangulaire supérieure** (*resp.* **triangulaire inférieure**) lorsque :

$$\forall (i, j) \in \llbracket 1; n \rrbracket^2, (i > j) \implies (a_{ij}) = 0 \quad (\text{resp. } (i < j) \implies (a_{ij}) = 0)$$

Exemples. Informellement, une matrice triangulaire supérieure (*resp.* inférieure) est une matrice carrée dont tous les coefficients situés strictement en-dessous (*resp.* au-dessus) de la diagonale sont nuls.

Dans $M_3(\mathbb{K})$, les matrices : $\begin{pmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 3 \end{pmatrix}$, $\begin{pmatrix} 2 & 1 & 4 \\ 0 & -1 & 2 \\ 0 & 0 & 3 \end{pmatrix}$, $\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$ sont triangulaires supérieures.

Dans $M_3(\mathbb{K})$, les matrices : $\begin{pmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 3 \end{pmatrix}$, $\begin{pmatrix} 2 & 0 & 0 \\ 4 & -1 & 0 \\ 7 & 2 & 3 \end{pmatrix}$, $\begin{pmatrix} 0 & 0 & 0 \\ -1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$ sont triangulaires inférieures.

La matrice identité I_n et la matrice nulle $0_{M_n(\mathbb{K})}$ sont des matrices triangulaires supérieures et inférieures. Plus généralement, toute matrice diagonale est une matrice triangulaire supérieure et inférieure.

Enfin, en général, une matrice triangulaire supérieure s'écrit :

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix}$$

Tandis qu'une matrice triangulaire inférieure s'écrit :

$$\begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

Notations : notons $D_n(\mathbb{K})$, $T_n^+(\mathbb{K})$ et $T_n^-(\mathbb{K})$ les ensembles des matrices diagonales, triangulaires supérieures et triangulaires inférieures de $M_n(\mathbb{K})$ respectivement.

On peut observer que : $T_n^+(\mathbb{K}) \cap T_n^-(\mathbb{K}) = D_n(\mathbb{K})$.²

PROPRIÉTÉ 13.8 - $(D_n(\mathbb{K}), +, \times)$, $(T_n^+(\mathbb{K}), +, \times)$ et $(T_n^-(\mathbb{K}), +, \times)$ sont des sous-anneaux de $(M_n(\mathbb{K}), +, \times)$.

En outre, l'anneau $(D_n(\mathbb{K}), +, \times)$ est commutatif pour tout $n \in \mathbb{N}^*$.

Enfin, pour $n \geq 2$, aucun de ces anneaux n'est intègre, et les anneaux $T_n^+(\mathbb{K})$ et $T_n^-(\mathbb{K})$ ne sont pas commutatifs.

PREUVE. Montrons que $T_n^+(\mathbb{K})$ est un sous-anneau de $M_n(\mathbb{K})$. En premier lieu, il s'agit de vérifier que $(T_n^+(\mathbb{K}), +)$ est un sous-groupe abélien de $(M_n(\mathbb{K}), +)$, ce qui passe par quatre vérifications immédiates :

(SG1) : $T_n^+(\mathbb{K}) \subset M_n(\mathbb{K})$ (par définition)

(SG2) : $T_n^+(\mathbb{K})$ est stable par somme (vérification aisée)

(SG3) : $0_{M_n(\mathbb{K})} \in T_n^+(\mathbb{K})$ (idem)

(SG4) : l'opposée d'une matrice de $T_n^+(\mathbb{K})$ est encore dans $T_n^+(\mathbb{K})$ (idem).

Ainsi $(T_n^+(\mathbb{K}), +)$ est un sous-groupe abélien de $(M_n(\mathbb{K}), +)$.

Pour achever la preuve de la propriété concernant $T_n^+(\mathbb{K})$, il suffit d'observer que $I_n \in T_n^+(\mathbb{K})$, et que $T_n^+(\mathbb{K})$ est stable par produit, ce qui est le seul point non trivial de cette série de vérifications.

Soient donc $A = (a_{ij})$ et $B = (b_{ij})$ deux matrices triangulaires supérieures de $M_n(\mathbb{K})$. Par définition (de $T_n^+(\mathbb{K})$), on a :

$$\forall (i, j) \in \llbracket 1, n \rrbracket^2, (i > j) \implies (a_{ij} = 0 \wedge b_{ij} = 0) \quad (\spadesuit)$$

Notons $P = A \times B$. Soient i et j deux entiers de $\llbracket 1, n \rrbracket$, avec $i > j$ ().

$$\text{On a : } P_{ij} = \sum_{k=1}^n a_{ik}b_{kj} = \sum_{k=1}^{i-1} a_{ik}b_{kj} + \sum_{k=i}^n a_{ik}b_{kj}.$$

2. Petit exercice.

Dans la somme : $\sum_{k=1}^{i-1} a_{ik} b_{kj}$ on a $k < i$, donc $a_{ik} = 0$ (d'après ($\spadesuit$)). Ainsi : $\sum_{k=1}^{i-1} a_{ik} b_{kj} = 0$.

Dans la somme : $\sum_{k=i}^n a_{ik} b_{kj}$ on a $k \geq i$ d'où $k > j$ (d'après ($\clubsuit$)), donc $b_{kj} = 0$ (d'après ($\spadesuit$)). Ainsi : $\sum_{k=i}^n a_{ik} b_{kj} = 0$.

Finalement $P_{ij} = 0$. En résumé, on a établi que : $\forall (i, j) \in \llbracket 1, n \rrbracket^2, (i > j) \implies (P_{ij} = 0)$; ce qui signifie exactement que $P \in T_n^+(\mathbb{K})$.

Conclusion : $\forall (A, B) \in T_n^+(\mathbb{K})^2, A \times B \in T_n^+(\mathbb{K})$ (la matrice produit de deux matrices triangulaires supérieures est une matrice triangulaire supérieure).

Synthèse. On déduit de ce qui précède que $(T_n^+(\mathbb{K}), +, \times)$ est un sous-anneau de $(M_n(\mathbb{K}), +, \times)$.

Ceci établi, on démontre de façon analogue que les matrices triangulaires inférieures dans $M_n(\mathbb{K})$ constituent un sous-anneau de $M_n(\mathbb{K})$; il suffit essentiellement de modifier ce qui doit l'être dans la preuve de la stabilité par produit. Ainsi :

$(T_n^-(\mathbb{K}), +, \times)$ est un sous-anneau de $(M_n(\mathbb{K}), +, \times)$.

Montrons que $D_n(\mathbb{K})$ est un sous-anneau de $M_n(\mathbb{K})$. De nouveau, il s'agit de vérifier que $(D_n(\mathbb{K}), +)$ est un sous-groupe abélien de $(M_n(\mathbb{K}), +)$, ce qui re-passe par quatre vérifications immédiates³ :

- (SG1) : $D_n(\mathbb{K}) \subset M_n(\mathbb{K})$ (par définition)
- (SG2) : $D_n(\mathbb{K})$ est stable par somme (vérification aisée)
- (SG3) : $0_{M_n(\mathbb{K})} \in D_n(\mathbb{K})$ (idem)
- (SG4) : l'opposée d'une matrice de $D_n(\mathbb{K})$ est encore dans $D_n(\mathbb{K})$ (idem).

Ainsi $(D_n(\mathbb{K}), +)$ est un sous-groupe abélien de $(M_n(\mathbb{K}), +)$.

Pour achever la preuve de la propriété concernant $D_n(\mathbb{K})$, il suffit d'observer que $I_n \in D_n(\mathbb{K})$, et que $D_n(\mathbb{K})$ est stable par produit. Établissons ce dernier point.

Soient D et D' deux matrices dans $D_n(\mathbb{K})$. Par définition, il existe $(2n)$ scalaires $\alpha_1, \beta_1, \dots, \alpha_n, \beta_n$ tels que $D = \text{diag}(\alpha_1, \dots, \alpha_n)$ et $D' = \text{diag}(\beta_1, \dots, \beta_n)$.

Notons leur produit : $P = DD'$.

Puisque les matrices D et D' sont toutes deux triangulaires supérieures et inférieures, leur produit est une matrice triangulaire supérieure et inférieure, c'est-à-dire une matrice diagonale.

Ainsi : $\forall (i, j) \in \llbracket 1, n \rrbracket^2, (i \neq j) \implies (P_{ij} = 0)$.

Il reste donc à calculer les coefficients diagonaux de la matrice P . Soit i un entier quelconque dans $\llbracket 1, n \rrbracket$. On a :

$$P_{ii} = \sum_{k=1}^n \underbrace{d_{ik}}_{=\alpha_i \delta_{ik}} \underbrace{\{d'\}_{ki}}_{=\beta_i \delta_{ik}} = \alpha_i \beta_i$$

3. Ou par l'utilisation de ce qui précède, et du fait que $T_n^+(\mathbb{K}) \cap T_n^-(\mathbb{K}) = D_n(\mathbb{K})$.

On en déduit finalement que :

$$\text{diag}(\alpha_1, \dots, \alpha_n) \times \text{diag}(\beta_1, \dots, \beta_n) = \text{diag}(\alpha_1\beta_1, \dots, \alpha_n\beta_n) \quad (b)$$

Ce résultat permet de faire d'une pierre deux coups, dans le sens où il permet d'abord de conclure que le produit de deux matrices diagonales (dans $M_n(\mathbb{K})$) est encore une matrice diagonale. D'où, d'après ce qui précède :

$$(D_n(\mathbb{K}), +, \times) \text{ est un sous-anneau de } (M_n(\mathbb{K}), +, \times).$$

En outre, il découle de la formule (b) que la multiplication est commutative dans $D_n(\mathbb{K})$. Ainsi :

$$(D_n(\mathbb{K}), +, \times) \text{ est un anneau commutatif.}$$

► Pour finir, il reste à vérifier les “anti-propriétés” de l'énoncé, à savoir que pour $n \geq 2$, les anneaux $D_n(\mathbb{K})$, $T_n^+(\mathbb{K})$ et $T_n^-(\mathbb{K})$ ne sont pas intègres. A cette fin, il suffit de considérer les matrices :

$$A = \text{diag}(1, 0, \dots, 0) \quad \text{et} \quad B = I_n - A = \text{diag}(0, 1, 1, \dots, 1)$$

On a alors $A \in D_n(\mathbb{K})$ (donc $A \in T_n^+(\mathbb{K})$ et $A \in T_n^-(\mathbb{K})$); $B \in D_n(\mathbb{K})$ (donc $B \in T_n^+(\mathbb{K})$ et $B \in T_n^-(\mathbb{K})$); A et B sont non nulles, mais $AB = 0_{M_n(\mathbb{K})}$. On en déduit que les anneaux $D_n(\mathbb{K})$, $T_n^+(\mathbb{K})$ et $T_n^-(\mathbb{K})$ ne sont pas intègres.

On renvoie enfin à la preuve de la non-commutativité de $M_n(\mathbb{K})$ pour un exemple permettant d'affirmer que les anneaux $T_n^+(\mathbb{K})$ et $T_n^-(\mathbb{K})$ ne sont pas commutatifs dès que $n \geq 2$ (voir page 308). ☕

Remarques. Il résulte en particulier de cette propriété que le produit de deux matrices triangulaires supérieures (ou inférieures) est une matrice triangulaire supérieure (ou inférieure). Et c'est un bon petit exercice de savoir retrouver la démonstration de cette affirmation.

Il résulte également de la propriété précédente que l'ensemble des matrices diagonales $D_n(\mathbb{K})$ est un anneau (muni de l'addition et de la multiplication), tout comme $T_n^+(\mathbb{K})$ et $T_n^-(\mathbb{K})$. On peut donc parler de l'**anneau $D_n(\mathbb{K})$ des matrices diagonales**, de l'**anneau $T_n^+(\mathbb{K})$ des matrices triangulaires supérieures** et de l'**anneau $T_n^-(\mathbb{K})$ des matrices triangulaires inférieures**.

Pour en finir avec ces remarques, l'ensemble des matrices scalaires de $M_n(\mathbb{K})$ (muni de l'addition et de la multiplication) est un sous-anneau de l'anneau $D_n(\mathbb{K})$ des matrices diagonales ; il est commutatif (sans surprise puisque $D_n(\mathbb{K})$ l'est) et intègre (trivial).

► Outre le fait d'être stable par somme et par produit, l'ensemble des matrices diagonales (ou triangulaires) est également stable par multiplication par un scalaire, et donc par combinaison linéaire. Cette remarque est formalisée par l'énoncé suivant.

PROPRIÉTÉ 13.9 - Soient A et B deux matrices de $D_n(\mathbb{K})$ (*resp.* $T_n^+(\mathbb{K})$, *resp.* $T_n^-(\mathbb{K})$). Alors :

$$\forall (\lambda, \mu) \in \mathbb{K}^2, (\lambda A + \mu B) \in D_n(\mathbb{K}) \text{ (resp. } T_n^+(\mathbb{K}), \text{ resp. } T_n^-(\mathbb{K}))$$

PREUVE. Puisque l'on a déjà établi que $D_n(\mathbb{K})$ (*resp.* $T_n^+(\mathbb{K})$, *resp.* $T_n^-(\mathbb{K})$) est stable par somme, il suffit de vérifier que pour toute matrice $A \in D_n(\mathbb{K})$ (*resp.* $T_n^+(\mathbb{K})$, *resp.* $T_n^-(\mathbb{K})$), et pour tout scalaire λ , la matrice (λA) est encore dans $D_n(\mathbb{K})$ (*resp.* $T_n^+(\mathbb{K})$, *resp.* $T_n^-(\mathbb{K})$), ce qui est une vérification sans difficulté. ☕

Remarque (“pour l’avenir”) : les propriétés de la somme et de la multiplication par un scalaire évoquées plus haut permettront d’affirmer que $D_n(\mathbb{K})$, $T_n^+(\mathbb{K})$ et $T_n^-(\mathbb{K})$ sont des $\mathbb{K}$ -**espaces vectoriels**, et plus précisément que ce sont des **sous-espaces vectoriels de $M_n(\mathbb{K})$** .

13.2.3 Commutant d’une matrice carrée

DÉFINITION 13.11 - Soit $A \in M_n(\mathbb{K})$ une matrice carrée. On appelle **commutant** de la matrice A et on note $\text{COM}(A)$ l’ensemble des matrices $B \in M_n(\mathbb{K})$ qui commutent avec A , càd telles que $AB = BA$. Explicitement :

$$\text{COM}(A) = \{B \in M_n(\mathbb{K}) / AB = BA\}$$

Exemple 1. Puisque toute matrice de $M_n(\mathbb{K})$ commute avec la matrice identité, on a : $\text{COM}(I_n) = M_n(\mathbb{K})$. Dans le même registre, on a aussi : $\text{COM}(0_{M_n(\mathbb{K})}) = M_n(\mathbb{K})$.

Exemple 2. Pour généraliser l’exemple 1, notons que pour tout scalaire λ , on a : $\text{COM}(\lambda I_n) = M_n(\mathbb{K})$.

Exemple 3 (exercice). On pose : $A = \text{diag}(1, 2, 3) \in M_3(\mathbb{K})$. On a : $\text{COM}(A) = D_3(\mathbb{K})$.

Exemple 4. La situation de l’exemple précédent peut être généralisée de la façon suivante.

Exercice classique — Soient n un entier ≥ 2 , et soient $\lambda_1, \dots, \lambda_n$ n scalaires deux à deux distincts. Le commutant de $D = \text{diag}(\lambda_1, \dots, \lambda_n)$ est l’ensemble des matrices diagonales de $M_n(\mathbb{K})$.

PREUVE. Soient $\lambda_1, \dots, \lambda_n$ n scalaires distincts. Notons $D = \text{diag}(\lambda_1, \dots, \lambda_n)$.

Une matrice $M = (m_{ij})_{1 \leq i, j \leq n}$ est dans le commutant de D si et seulement si $MD = DM$.

Or : $MD = DM \iff \forall (i, j) \in \llbracket 1, n \rrbracket^2, (MD)_{ij} = (DM)_{ij}$

$$\iff \forall (i, j) \in \llbracket 1, n \rrbracket^2, \sum_{k=1}^n m_{ik} d_{kj} = \sum_{k=1}^n d_{ik} m_{kj}$$

$$\iff \forall (i, j) \in \llbracket 1, n \rrbracket^2, m_{ij} d_{jj} = d_{ii} m_{ij}$$

$$\iff \forall (i, j) \in \llbracket 1, n \rrbracket^2, m_{ij} \lambda_j = \lambda_i m_{ij}$$

$$\iff \forall (i, j) \in \llbracket 1, n \rrbracket^2, (\lambda_j - \lambda_i) m_{ij} = 0$$

D’où, puisque l’on a supposé les λ_i distincts : $\forall (i, j) \in \llbracket 1, n \rrbracket^2, i \neq j \implies m_{ij} = 0$

On en déduit que les matrices M commutant avec D sont celles dont tous les coefficients situés en dehors de la diagonale ($i \neq j$) sont nuls ; en d’autres termes, ce sont exactement les matrices diagonales de $M_n(\mathbb{K})$.

Conclusion : le commutant de D est l’ensemble des matrices diagonales de $M_n(\mathbb{K})$.



PROPRIÉTÉ 13.10 - Pour toute $A \in M_n(\mathbb{K})$, $(\text{COM}(A), +, \times)$ est un sous-anneau de $(M_n(\mathbb{K}), +, \times)$.

En outre : $\forall (\lambda, \mu) \in \mathbb{K}^2, \forall (B, C) \in \text{COM}(A)^2, (\lambda B + \mu C) \in \text{COM}(A)$

PREUVE. Montrons que $\text{COM}(A)$ est un sous-anneau de $M_n(\mathbb{K})$. En premier lieu, il s’agit de vérifier que $(\text{COM}(A), +)$ est un sous-groupe abélien de $(M_n(\mathbb{K}), +)$, ce qui passe par quatre vérifications aisées :

(SG1) : $\text{COM}(A) \subset M_n(\mathbb{K})$ (par définition)

(SG2) : $\text{COM}(A)$ est stable par somme. Soient en effet B et C deux matrices de $\text{COM}(A)$. On a :

$$(B + C)A = BA + CA = AB + AC = A(B + C)$$

La deuxième égalité provenant de l'hypothèse suivant laquelle B et C appartiennent au commutant de A ; les autres égalités provenant des propriétés du produit matriciel dans $M_n(\mathbb{K})$.

(SG3) : $0_{M_n(\mathbb{K})} \in \text{COM}(A)$ (trivial)

(SG4) : l'opposée d'une matrice de $\text{COM}(A)$ est encore dans $\text{COM}(A)$. En effet, si B est dans $\text{COM}(A)$ alors :

$$(-B) \times A = -(BA) = -(AB) = A \times (-B) \quad \text{Ce qui signifie que : } (-B) \in \text{COM}(A).$$

Ainsi $(\text{COM}(A), +)$ est un sous-groupe abélien de $(M_n(\mathbb{K}), +)$.

Pour achever la preuve de la propriété, il reste à observer que $I_n \in \text{COM}(A)$ (puisque la matrice identité commute avec "tout le monde"), et que $\text{COM}(A)$ est stable par produit, ce que l'on prouve ci-dessous.

Soient donc B et C deux matrices de $\text{COM}(A)$. Par définition (de $\text{COM}(A)$), et en utilisant à plein régime l'associativité du produit dans $M_n(\mathbb{K})$, on a :

$$(BC)A = B(CA) = B(AC) = (BA)C = (AB)C = A(BC) \quad \text{Ce qui signifie que : } (BC) \in \text{COM}(A).$$

Conclusion : $\forall (B, C) \in \text{COM}(A)^2, B \times C \in \text{COM}(A)$.

Synthèse. On déduit de ce qui précède que $(\text{COM}(A), +, \times)$ est un sous-anneau de $(M_n(\mathbb{K}), +, \times)$.

La vérification de la dernière partie de la propriété (stabilité de $\text{COM}(A)$ par combinaison linéaire) est aisée (et laissée en exercice). ☕

Remarque ("pour l'avenir") : les propriétés de la somme et de la multiplication par un scalaire évoquées plus haut permettront d'affirmer que $\text{COM}(A)$ est un $\mathbb{K}$ -**espace vectoriel**, et plus précisément que c'est un **sous-espace vectoriel de $M_n(\mathbb{K})$** .

13.2.4 Puissances d'une matrice carrée, et matrices carrées nilpotentes

DÉFINITION 13.12 - (Puissances d'une matrice carrée). Soit $A \in M_n(\mathbb{K})$, et soit N un entier naturel. On définit la matrice A^N en posant :

$$A^0 = I_n, \text{ et pour tout } N \in \mathbb{N}^*, A^N = A \times A^{N-1}$$

Enfonçons trois portes ouvertes :

PROPRIÉTÉ 13.11 - $\forall N \in \mathbb{N}, I_n^N = I_n$

PROPRIÉTÉ 13.12 - $\forall A \in M_n(\mathbb{K}), \forall (N, M) \in \mathbb{N}^2, A^N \times A^M = A^{N+M} = A^M \times A^N$

PROPRIÉTÉ 13.13 - $\forall N \in \mathbb{N}, [\text{diag}(\lambda_1, \dots, \lambda_n)]^N = \text{diag}(\lambda_1^N, \dots, \lambda_n^N)$

Les preuves de ces trois dernières propriétés reposent sur des récurrences aisées.

Exemple 1 - Exercice classique. Calculer A^n pour tout $n \in \mathbb{N}$ avec $A = \begin{pmatrix} 1 & \alpha \\ 0 & 1 \end{pmatrix}$ où α est un scalaire quelconque.

On peut commencer par calculer les premières puissances de A , dans l'idée de conjecturer l'expression de A^n .

$$\text{On a : } A^0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}; A = \begin{pmatrix} 1 & \alpha \\ 0 & 1 \end{pmatrix}; A^2 = \begin{pmatrix} 1 & 2\alpha \\ 0 & 1 \end{pmatrix}; A^3 = \begin{pmatrix} 1 & 3\alpha \\ 0 & 1 \end{pmatrix}$$

A ce stade du raisonnement, on se propose donc de noter pour tout n entier naturel $P(n)$ l'assertion :
 “ $A^n = \begin{pmatrix} 1 & n\alpha \\ 0 & 1 \end{pmatrix}$ ”.

Démontrons que $P(n)$ est vraie pour tout entier naturel n par récurrence sur n . Il est immédiat que $P(0)$ est vraie, ce qui fournit l'initialisation de notre raisonnement.

Passons à l'hérédité : supposons que $P(n)$ est vraie pour un certain entier naturel n . Alors :

$$A^{n+1} = A \times A^n = \begin{pmatrix} 1 & \alpha \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & n\alpha \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & n\alpha + \alpha \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & (n+1)\alpha \\ 0 & 1 \end{pmatrix}$$

On en déduit que $P(n+1)$ est vraie, ce qui établit l'hérédité de la propriété.

Il s'ensuit que : $\forall n \in \mathbb{N}, \begin{pmatrix} 1 & \alpha \\ 0 & 1 \end{pmatrix}^n = \begin{pmatrix} 1 & n\alpha \\ 0 & 1 \end{pmatrix}$.

Exemple 2 - Exercice classique. Calculer $A(\theta)^n$ pour tout $n \in \mathbb{N}$ avec $A(\theta) = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$ où θ désigne un réel quelconque.

L'idée ici est de faire une observation qui sera justifiée par la suite du cours de cette année⁴ :

$$\forall (\theta, \varphi) \in \mathbb{R}^2, A(\theta)A(\varphi) = A(\theta + \varphi) \quad (\spadesuit)$$

En effet, pour tout couple de réels θ et φ on a :

$$\begin{aligned} A(\theta)A(\varphi) &= \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \begin{pmatrix} \cos(\varphi) & -\sin(\varphi) \\ \sin(\varphi) & \cos(\varphi) \end{pmatrix} \\ &= \begin{pmatrix} \cos(\theta)\cos(\varphi) - \sin(\theta)\sin(\varphi) & -\cos(\theta)\sin(\varphi) - \sin(\theta)\cos(\varphi) \\ \cos(\theta)\sin(\varphi) + \sin(\theta)\cos(\varphi) & \cos(\theta)\cos(\varphi) - \sin(\theta)\sin(\varphi) \end{pmatrix} \end{aligned}$$

En reconnaissant les formules d'addition pour le cosinus et le sinus, on en déduit que :

$$\forall (\theta, \varphi) \in \mathbb{R}^2, A(\theta)A(\varphi) = \begin{pmatrix} \cos(\theta + \varphi) & -\sin(\theta + \varphi) \\ \sin(\theta + \varphi) & \cos(\theta + \varphi) \end{pmatrix}$$

4. Plus précisément, par l'interprétation géométrique que nous ferons de la matrice $A(\theta)$.

Ce qui établit la formule (♠).

Cet outil en poche, on vérifie (encore à l'aide d'une récurrence facile) que :

$$\forall \theta \in \mathbb{R}, \forall n \in \mathbb{N}, [A(\theta)]^n = A(n\theta)$$

Autrement écrit :

$$\forall \theta \in \mathbb{R}, \forall n \in \mathbb{N}, \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}^n = \begin{pmatrix} \cos(n\theta) & -\sin(n\theta) \\ \sin(n\theta) & \cos(n\theta) \end{pmatrix}$$

Exemple 3. Le petit jeu des deux exemples précédents, qui consiste finalement à conjecturer l'expression de A^n , puis à démontrer cette conjecture par récurrence, trouve assez vite ses limites. Grossièrement, elle fonctionne bien pour des matrices carrées (évidemment) de taille 2, mais devient plus pénible à mettre en œuvre, voire carrément inextricable pour de “plus grandes” matrices.

Explicitement, il faut un peu de courage pour calculer les puissances de A avec : $A = \begin{pmatrix} 2 & 1 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$.

Exemple 4. Comme dans l'exemple précédent, il faut cette fois-ci beaucoup de courage pour tenter de conjecturer l'expression de A^n avec $A = \begin{pmatrix} 0 & 3/4 & 0 \\ 3/4 & 0 & 1 \\ 1/4 & 1/4 & 0 \end{pmatrix}$.

Nous verrons d'ici la fin de ce chapitre deux nouvelles méthodes pour calculer les puissances d'une matrice carrée, qui fourniront les réponses aux exemples 3 et 4.

La notion introduite ci-dessous est un premier pas pour y parvenir.

DÉFINITION 13.13 - Une matrice carrée $N \in M_n(\mathbb{K})$ est **nilpotente** s'il existe un entier naturel m tel que $N^m = 0_{M_n(\mathbb{K})}$.

Remarque. Cette définition n'est pas nouvelle, puisque c'est la traduction dans l'anneau $M_n(\mathbb{K})$ de la définition plus générale d'élément nilpotent dans un anneau quelconque énoncée dans le chapitre précédent.

Exemples. Les matrices de la forme $\begin{pmatrix} 0 & a \\ 0 & 0 \end{pmatrix}$ dans $M_2(\mathbb{K})$ et $\begin{pmatrix} 0 & a & b \\ 0 & 0 & c \\ 0 & 0 & 0 \end{pmatrix}$ dans $M_3(\mathbb{K})$ sont nilpotentes (preuves laissées en exercice).

Plus généralement d'ailleurs :

PROPRIÉTÉ 13.14 - Dans $M_n(\mathbb{K})$, toute matrice triangulaire (supérieure ou inférieure) ayant une diagonale nulle est nilpotente.

PREUVE. Voir au second semestre, chapitres d'algèbre linéaire. ☕

PROPRIÉTÉ 13.15 - (**Formule du binôme de Newton dans $M_n(\mathbb{K})$**). Soient A et B dans $M_n(\mathbb{K})$, telles que $AB = BA$, et soit $N \in \mathbb{N}$. Alors :

$$(A + B)^N = \sum_{k=0}^N \binom{N}{k} A^k B^{N-k} = \sum_{k=0}^N \binom{N}{k} A^{N-k} B^k$$

PREUVE. C'est un cas particulier d'une formule démontrée plus généralement dans le cadre d'un anneau quelconque (voir propriété 12.7, page 284). ☕

Application pratique importante. Calcul des puissances d'une matrice A se décomposant comme : $A = \lambda I_n + N$, avec $\lambda \in \mathbb{K}$ et N matrice nilpotente.

Exemple. On reprend la situation de l'exemple 3, page 316, avec : $A = \begin{pmatrix} 2 & 1 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$.

Observons que : $A = 2I_3 + B$ avec $B = \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$.

► On a : $B^2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ et $B^3 = 0_{M_3(\mathbb{K})}$. Ainsi : $\forall k \in \mathbb{N}, (k \geq 3) \implies (B^k = 0_{M_3(\mathbb{K})})$ (♠).⁵

► On a : $(2I_3) \times B = B \times (2I_3)$ (♣). En effet, toute matrice de la forme $(\lambda I_3)^6$ commute avec toute matrice de $M_3(\mathbb{K})$.

► Soit N un entier naturel. On a : $A^N = (2I_3 + B)^N$. Grâce à (♣), on peut utiliser la formule du binôme de Newton pour écrire :

$$A^N = \sum_{k=0}^N \binom{N}{k} B^k (2I_3)^{N-k} = \sum_{k=0}^N \binom{N}{k} 2^{N-k} B^k$$

D'après (♠), on a encore :

$$A^N = \sum_{k=0}^2 \binom{N}{k} 2^{N-k} B^k = 2^N \underbrace{B^0}_{=I_3} + N 2^{N-1} B + \frac{N(N-1)}{2} 2^{N-2} B^2$$

Explicitement :

$$A^N = \begin{pmatrix} 2^N & 0 & 0 \\ 0 & 2^N & 0 \\ 0 & 0 & 2^N \end{pmatrix} + \begin{pmatrix} 0 & N 2^{N-1} & N 2^{N-1} \\ 0 & 0 & N 2^{N-1} \\ 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & N(N-1) 2^{N-3} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

D'où finalement :

$$\forall N \in \mathbb{N}, A^N = \begin{pmatrix} 2^N & N 2^{N-1} & N(N+3) 2^{N-3} \\ 0 & 2^N & N 2^{N-1} \\ 0 & 0 & 2^N \end{pmatrix}$$

5. Ainsi la matrice B est nilpotente.

6. Une telle matrice est appelée matrice **scalaire**.

13.2.5 Matrices symétriques et antisymétriques

Dans tout ce paragraphe, on ne considère de nouveau que des matrices carrées.

DÉFINITION 13.14 - Une matrice carrée A de $M_n(\mathbb{K})$ est **symétrique** (*resp.* **antisymétrique**) lorsque

$${}^tA = A \text{ (resp. } {}^tA = -A\text{)}$$

Notations. Dans ce chapitre, nous conviendrons de noter $S_n(\mathbb{K})$ et $A_n(\mathbb{K})$ les ensembles des matrices de $M_n(\mathbb{K})$ symétriques et antisymétriques respectivement, càd :

$$S_n(\mathbb{K}) = \{A \in M_n(\mathbb{K}) \mid {}^tA = A\} \quad \text{et} \quad A_n(\mathbb{K}) = \{A \in M_n(\mathbb{K}) \mid {}^tA = -A\}$$

Exemples.

1/ La matrice nulle $0_{M_n(\mathbb{K})}$ est symétrique et antisymétrique.

2/ La matrice identité I_n est symétrique, mais pas antisymétrique.

3/ La matrice $\begin{pmatrix} 1 & 2 & 3 \\ 2 & 0 & 4 \\ 3 & 4 & 5 \end{pmatrix}$ est symétrique mais pas antisymétrique.

4/ La matrice $\begin{pmatrix} 0 & 2 & 3 \\ -2 & 0 & 4 \\ -3 & -4 & 0 \end{pmatrix}$ est antisymétrique mais pas symétrique.

5/ La matrice $\begin{pmatrix} 1 & 2 & 3 \\ -2 & 0 & 4 \\ -3 & -4 & 5 \end{pmatrix}$ n'est ni symétrique ni antisymétrique.

6/ La matrice $\begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$ est symétrique (*resp.* antisymétrique) si $\theta = 0 \text{ } [\pi]$ (*resp.* $\theta = \frac{\pi}{2} \text{ } [\pi]$);
ni symétrique ni antisymétrique lorsque $\theta \neq 0 \text{ } \left[\frac{\pi}{2}\right]$.

PROPRIÉTÉ 13.16 - (Propriétés des matrices symétriques et antisymétriques).

1/ $(S_n(\mathbb{K}), +)$ et $(A_n(\mathbb{K}), +)$ sont des sous-groupes de $(M_n(\mathbb{K}), +)$.

2/ $S_n(\mathbb{K})$ est stable par combinaison linéaire, càd :

$$\forall (\lambda, \mu) \in \mathbb{K}^2, \forall (A, B) \in S_n(\mathbb{K})^2, (\lambda A + \mu B) \in S_n(\mathbb{K})$$

3/ $A_n(\mathbb{K})$ est stable par combinaison linéaire, càd :

$$\forall (\lambda, \mu) \in \mathbb{K}^2, \forall (A, B) \in A_n(\mathbb{K})^2, (\lambda A + \mu B) \in A_n(\mathbb{K})$$

4/ Une matrice antisymétrique ne comporte que des zéros sur sa diagonale.

5/ La matrice nulle $0_{M_n(\mathbb{K})}$ est la seule matrice de $M_n(\mathbb{K})$ simultanément symétrique et antisymétrique, càd :

$$S_n(\mathbb{K}) \cap A_n(\mathbb{K}) = \{0_{M_n(\mathbb{K})}\}$$

PREUVE. 1/ Montrons que $(S_n(\mathbb{K}), +)$ est un sous-groupe de $(M_n(\mathbb{K}), +)$.

(SG1) : $S_n(\mathbb{K}) \subset M_n(\mathbb{K})$ (par définition)

(SG2) : $S_n(\mathbb{K})$ est stable par somme. En effet, si A et B sont deux matrices symétriques, alors⁷ :

$${}^t(A + B) = {}^tA + {}^tB = A + B \quad \text{D'où : } (A + B) \in S_n(\mathbb{K}).$$

(SG3) : $0_{M_n(\mathbb{K})} \in S_n(\mathbb{K})$ car ${}^t0_{M_n(\mathbb{K})} = 0_{M_n(\mathbb{K})}$ (!!!)

(SG4) : l'opposée d'une matrice de $S_n(\mathbb{K})$ est encore dans $S_n(\mathbb{K})$, puisque si $A \in S_n(\mathbb{K})$, on a

$${}^t(-A) = -{}^tA = -A \quad \text{D'où : } (-A) \in S_n(\mathbb{K}).$$

Ainsi $(S_n(\mathbb{K}), +)$ est un sous-groupe abélien de $(M_n(\mathbb{K}), +)$.

La démonstration est analogue pour $A_n(\mathbb{K})$ (il suffit de rajouter des signes “-” aux places idoines).

2/ Puisqu'on a déjà établi que $S_n(\mathbb{K})$ est stable par somme, il suffit de vérifier que $S_n(\mathbb{K})$ est stable par multiplication par un scalaire pour conclure. Considérons donc $A \in S_n(\mathbb{K})$ et $\lambda \in K$. On a (encore une fois par linéarité de la transposée) :

$${}^t(\lambda A) = \lambda {}^tA = \lambda A \quad \text{D'où : } (\lambda A) \in S_n(\mathbb{K}).$$

On en déduit, avec le point 1, que : $\forall (\lambda, \mu) \in \mathbb{K}^2, \forall (A, B) \in S_n(\mathbb{K})^2, (\lambda A + \mu B) \in S_n(\mathbb{K})$.

3/ Puisqu'on a déjà établi que $A_n(\mathbb{K})$ est stable par somme, il suffit de vérifier que $A_n(\mathbb{K})$ est stable par multiplication par un scalaire pour conclure. Considérons donc $A \in A_n(\mathbb{K})$ et $\lambda \in K$. On a :

$${}^t(\lambda A) = \lambda {}^tA = \lambda(-A) = -\lambda A \quad \text{D'où : } (\lambda A) \in A_n(\mathbb{K}).$$

7. Par linéarité de la transposée, voir propriété 13.5 page 306.

On en déduit, avec le point 1, que : $\forall (\lambda, \mu) \in \mathbb{K}^2, \forall (A, B) \in A_n(\mathbb{K})^2, (\lambda A + \mu B) \in A_n(\mathbb{K})$.

4/ Soit $A \in A_n(\mathbb{K})$ une matrice antisymétrique (on a donc ${}^tA = -A$). Soit i un entier dans $\llbracket 1, n \rrbracket$. On a :

$$({}^tA)_{ii} = a_{ii} \text{ (par définition de transposée)} \quad \text{et} \quad ({}^tA)_{ii} = -a_{ii} \text{ (par hypothèse)}$$

D'où : $[\forall i \in \llbracket 1, n \rrbracket, a_{ii} = -a_{ii}] \implies [\forall i \in \llbracket 1, n \rrbracket, a_{ii} = 0]$

On a ainsi établi que : $[A \in A_n(\mathbb{K})] \implies [\text{La diagonale de } A \text{ est nulle}]$.

5/ Soit $A \in M_n(\mathbb{K})$. On a :

$$[A \in S_n(\mathbb{K}) \cap A_n(\mathbb{K})] \iff [{}^tA = A \wedge {}^tA = -A] \iff [A = -A] \iff [A = 0_{M_n(\mathbb{K})}]$$

D'où : $S_n(\mathbb{K}) \cap A_n(\mathbb{K}) = \{0_{M_n(\mathbb{K})}\}$. ☕

Remarque (“pour l’avenir”) : les points 1 et 2 de la propriété énoncée plus haut permettront d'affirmer que $S_n(\mathbb{K})$ et $A_n(\mathbb{K})$ sont des **sous-espaces vectoriels de $M_n(\mathbb{K})$** .

“Anti-propriété” : il n'est pas vrai en général que le produit de deux matrices symétriques est encore une matrice symétrique. Considérons par exemple les matrices :

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} \quad \text{et} \quad B = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

Il est clair que A et B sont **symétriques**. En revanche :

$$AB = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 2 & 0 \end{pmatrix}$$

AB n'est donc pas symétrique.

Tout le malheur dans cet exemple provient encore une fois de la non commutativité du produit matriciel. Explicitement, lorsque A et B sont symétriques, on a :

$${}^t(AB) = {}^tB {}^tA = BA$$

Mais on ne peut conclure que AB est symétrique que si $BA = AB$, càd seulement si A et B commutent.

Il est donc correct d'affirmer que le produit de deux matrices symétriques **qui commutent** est une matrice symétrique.

Les mêmes remarques peuvent être adaptées aux matrices antisymétriques.

THÉORÈME 13.1 - (Décomposition “Symétrique + Antisymétrique”). Toute matrice de $M_n(\mathbb{K})$ s’écrit de manière unique (à l’ordre des facteurs près) comme somme d’une matrice symétrique et d’une matrice antisymétrique. Formellement :

$$\forall M \in M_n(\mathbb{K}), \exists! (S, A) \in S_n(\mathbb{K}) \times A_n(\mathbb{K}), M = S + A$$

Explicitement : $S = \frac{1}{2}(M + {}^tM)$ et $A = \frac{1}{2}(M - {}^tM)$.

PREUVE. On raisonne par analyse-synthèse. Soit $M \in M_n(\mathbb{K})$.

► Analyse : supposons qu’il existe une matrice symétrique S et une matrice antisymétrique A telles que : $M = S + A$.

Alors : ${}^tM = {}^tS + {}^tA$. Or par hypothèse ${}^tS = S$ et ${}^tA = -A$. On a donc : ${}^tM = S - A$.

Il s’ensuit que S et A sont solutions du système :
$$\begin{cases} S + A = M \\ S - A = {}^tM \end{cases}$$

La résolution aisée de celui-ci donne : $S = \frac{1}{2}(M + {}^tM)$ et $A = \frac{1}{2}(M - {}^tM)$.

► Synthèse : il ne reste plus qu’à vérifier que le couple (S, A) obtenu précédemment convient. A cette fin, on commence par s’assurer que $M = S + A$ (trivial). En outre :

► en posant $S = \frac{1}{2}(M + {}^tM)$, on a : ${}^tS = \frac{1}{2}({}^tM + {}^t({}^tM)) = \frac{1}{2}({}^tM + M) = \frac{1}{2}(M + {}^tM) = S$; donc S est symétrique ;

► et en posant $A = \frac{1}{2}(M - {}^tM)$, on a : ${}^tA = \frac{1}{2}({}^tM - M) = -\frac{1}{2}(M - {}^tM) = -A$; donc A est antisymétrique.

Conclusion (partielle) : nous venons d’établir l’existence, pour toute matrice carrée M d’un couple (S, A) (avec S symétrique et A antisymétrique) tel que : $M = S + A$.

En outre, ce couple est explicitement donné par les formules : $S = \frac{1}{2}(M + {}^tM)$ et $A = \frac{1}{2}(M - {}^tM)$.

Pour achever la preuve du théorème, il reste à établir l’unicité du couple (S, A) (pour une matrice M fixée).

► Unicité : supposons qu’il existe deux couples (S, A) et (S', A') tels que $S + A = M$ et $S' + A' = M$, avec S et S' symétriques, et A et A' antisymétriques. Dans ce cas : $S + A = S' + A'$, d’où $S - S' = A' - A$.

Dans cette dernière égalité la matrice de gauche $(S - S')$ est symétrique (puisque S et S' le sont) ; mais elle est aussi antisymétrique, car égale à $A' - A$ (et A et A' sont antisymétriques).

Or la seule matrice symétrique et antisymétrique est la matrice nulle. On en déduit donc que $S - S' = 0_{M_n(\mathbb{K})}$ d’où $S = S'$; et $A' - A = 0_{M_n(\mathbb{K})}$ d’où $A = A'$. Par suite les couples (S, A) et (S', A') sont égaux, ce qui prouve l’unicité et complète la démonstration du théorème. ☕

Remarque : ne pas rater (c’est inratable !) la fantastique analogie avec la propriété concernant les fonctions paires et impaires. Rappelons que si I est une partie de $\mathbb{R}$ symétrique par rapport à zéro, toute fonction f définie sur I et à valeurs dans $\mathbb{K}$ peut s’écrire de manière unique (à l’ordre près) comme somme d’une fonction paire et d’une fonction impaire.

Exemple. Soit $M = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 4 & -2 \\ 6 & 0 & -2 \end{pmatrix}$. On a : ${}^tM = \begin{pmatrix} 1 & 0 & 6 \\ 2 & 4 & 0 \\ 3 & -2 & 2 \end{pmatrix}$.

$$\text{D'où : } \frac{1}{2}(M + {}^tM) = \begin{pmatrix} 1 & 1 & 9/2 \\ 1 & 4 & -1 \\ 9/2 & -1 & 0 \end{pmatrix} \text{ et } \frac{1}{2}(M - {}^tM) = \begin{pmatrix} 0 & 1 & -3/2 \\ -1 & 0 & -1 \\ 3/2 & 1 & 0 \end{pmatrix}.$$

La décomposition de M comme somme d'une matrice symétrique et d'une matrice antisymétrique est donc :

$$\begin{pmatrix} 1 & 2 & 3 \\ 0 & 4 & -2 \\ 6 & 0 & -2 \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & 1 & 9/2 \\ 1 & 4 & -1 \\ 9/2 & -1 & 0 \end{pmatrix}}_{\text{symétrique}} + \underbrace{\begin{pmatrix} 0 & 1 & -3/2 \\ -1 & 0 & -1 \\ 3/2 & 1 & 0 \end{pmatrix}}_{\text{antisymétrique}}$$

13.2.6 Trace d'une matrice carrée

Dans tout ce paragraphe, on ne considère que des matrices carrées.

DÉFINITION 13.15 - Soit $A \in M_n(\mathbb{K})$. La **trace** de A , notée $\text{tr}(A)$, est le scalaire :

$$\text{tr}(A) = \sum_{i=1}^n a_{ii} \quad (\text{somme des éléments diagonaux})$$

Exemples : $\text{tr}(I_n) = n$; $\text{tr}(0_{M_n(\mathbb{K})}) = 0$; $\text{tr}\begin{pmatrix} 1 & 2 & 3 \\ 0 & 4 & -2 \\ 6 & 0 & -2 \end{pmatrix} = 3$; $\text{tr}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = 0$.

PROPRIÉTÉ 13.17 - (Propriétés de la trace d'une matrice carrée).

1/ $\forall (A, B) \in M_n(\mathbb{K})^2$, $\forall (\lambda, \mu) \in \mathbb{K}^2$, $\text{tr}(\lambda A + \mu B) = \lambda \text{tr}(A) + \mu \text{tr}(B)$ (**linéarité de la trace**)

2/ $\forall (A, B) \in M_n(\mathbb{K})^2$, $\text{tr}(AB) = \text{tr}(BA)$

PREUVE. 1/ Soient A et B deux matrices de $M_n(\mathbb{K})$, et λ et μ deux scalaires. On a :

$$\text{tr}(\lambda A + \mu B) = \sum_{i=1}^n (\lambda A + \mu B)_{ii} = \sum_{i=1}^n (\lambda a_{ii} + \mu b_{ii}) = \lambda \sum_{i=1}^n a_{ii} + \mu \sum_{i=1}^n b_{ii} = \lambda \text{tr}(A) + \mu \text{tr}(B)$$

2/ Soient $A = (a_{ij})$ et $B = (b_{ij})$ deux matrices de $M_n(\mathbb{K})$.

$$\text{D'une part : } \text{tr}(AB) = \sum_{i=1}^n (AB)_{ii} = \sum_{i=1}^n \sum_{k=1}^n a_{ik} b_{ki}.$$

$$\text{D'autre part : } \text{tr}(BA) = \sum_{i=1}^n (BA)_{ii} = \sum_{i=1}^n \sum_{k=1}^n b_{ik} a_{ki} = \sum_{k=1}^n \sum_{i=1}^n b_{ik} a_{ki} = \sum_{i=1}^n \sum_{k=1}^n b_{ki} a_{ik}.$$

Conclusion : $\forall (A, B) \in M_n(\mathbb{K})^2$, $\text{tr}(AB) = \text{tr}(BA)$



13.3 Matrices inversibles

Dans ce paragraphe on ne considère encore une fois que des matrices carrées.

13.3.1 Généralités

DÉFINITION 13.16 - Une matrice carrée $A = (a_{ij})_{1 \leq i, j \leq n}$ de $M_n(\mathbb{K})$ est **inversible** s'il existe une matrice $B \in M_n(\mathbb{K})$ telle que : $AB = BA = I_n$.

Notation : lorsque A est inversible, on note A^{-1} son inverse (qui est unique en vertu de la propriété 12.2, page 275).

Quelques exemples de matrices inversibles (ou pas)

➤ I_n est inversible et $I_n^{-1} = I_n$. Plus généralement, pour tout scalaire λ non nul, la matrice λI_n est inversible, et $(\lambda I_n)^{-1} = \frac{1}{\lambda} I_n$.

➤ La matrice $\text{diag}(\lambda_1, \dots, \lambda_n)$ est inversible SSI tous les λ_i sont non nuls. Lorsque tel est le cas, on a : $[\text{diag}(\lambda_1, \dots, \lambda_n)]^{-1} = \text{diag}\left(\frac{1}{\lambda_1}, \dots, \frac{1}{\lambda_n}\right)$.

➤ Pour tout scalaire α , la matrice $A = \begin{pmatrix} 1 & \alpha \\ 0 & 1 \end{pmatrix}$ est inversible et on a $A^{-1} = \begin{pmatrix} 1 & -\alpha \\ 0 & 1 \end{pmatrix}$.

➤ La matrice nulle n'est pas inversible, mais il ne suffit pas qu'une matrice soit non nulle pour être inversible : par exemple, les matrices $\begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix}$ et $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ ne sont pas inversibles (bien que non nulles).

PROPRIÉTÉ 13.18 - (provisoirement admise). Une matrice A de $M_n(\mathbb{K})$ est inversible si et seulement si il existe une matrice B telle que $AB = I_n$ ou $BA = I_n$.

PREUVE. On renvoie au chapitre portant sur le rang d'une application linéaire pour la preuve de cette propriété. Observons que la différence entre cet énoncé et la définition de matrice inversible repose sur l'utilisation du "ou" en lieu et place d'un "et". Sa conséquence pratique est qu'il suffit de vérifier qu'une matrice est inversible à gauche ou à droite pour pouvoir affirmer qu'elle est inversible tout court. ☕

PROPRIÉTÉ 13.19 - Soient A et B deux matrices carrées de $M_n(\mathbb{K})$.

1/ Si A est inversible, alors A^{-1} est inversible et $(A^{-1})^{-1} = A$.

2/ Si A et B sont inversibles, alors AB est inversible et $(AB)^{-1} = B^{-1}A^{-1}$.

3/ Si A est inversible, alors A^N est inversible pour tout entier naturel N ; et $(A^N)^{-1} = (A^{-1})^N$. Dans ce cas, on notera A^{-N} pour désigner $(A^N)^{-1}$.

4/ Si A est inversible, alors ${}^tA$ est inversible et $({}^tA)^{-1} = {}^t(A^{-1})$.

PREUVE. Les points 1 et 2 sont des cas particuliers des propriétés plus générales de l'inverse démontrées au cours du chapitre 11 (voir propriété 12.3, page 276) ; le point 3 se déduit par récurrence du point 2.

Prouvons le point 4 : supposons que $A \in M_n(\mathbb{K})$ est inversible. Alors on a :

$${}^t(AA^{-1}) = {}^t(A^{-1}) {}^tA \quad \text{et} \quad {}^t(AA^{-1}) = {}^tI_n = I_n$$

Il s'ensuit que : ${}^t(A^{-1}) {}^tA = I_n$. On en déduit, à l'aide de la propriété 13.18, que ${}^tA$ est inversible et que : $({}^tA)^{-1} = {}^t(A^{-1})$. ☕

Notation. On note $GL_n(\mathbb{K})$ l'ensemble des matrices de $M_n(\mathbb{K})$ inversibles (pour le produit...). Avec les notations du chapitre précédent, il revient au même d'écrire que : $GL_n(\mathbb{K}) = M_n(\mathbb{K})^*$.

PROPRIÉTÉ 13.20 - $(GL_n(\mathbb{K}), *)$ est un groupe (appelé **groupe linéaire** des matrices inversibles).

PREUVE. Encore une fois, c'est un cas particulier d'une propriété plus générale démontrée dans le cadre des anneaux : lorsque $(A, +, \times)$ est un anneau. Explicitement, lorsque $(A^*, \times)$ est un groupe (voir propriété 12.14, page 289). ☕

13.3.2 Cas particulier des matrices inversibles dans $M_2(\mathbb{K})$

THÉORÈME 13.2 - (Description de $GL_2(\mathbb{K})$). Dans $M_2(\mathbb{K})$, la matrice $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ est inversible si et seulement si $ad - bc \neq 0$. Son inverse est alors : $A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$.

PREUVE. Soit $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$.

Supposons que $ad - bc \neq 0$. Alors la matrice $B = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$ est bien définie, et on vérifie sans difficulté que $BA = I_2$. En vertu de la propriété 13.18, on en déduit que A est inversible et que $A^{-1} = B$.

On a prouvé l'implication : $[ad - bc \neq 0] \implies [A \in GL_n(\mathbb{K})]$.

Montrons l'implication réciproque, en établissant sa contraposée : $[ad - bc = 0] \implies [A \notin GL_n(\mathbb{K})]$.

On suppose donc : $ad - bc = 0$.

Raisonnons par l'absurde et supposons que A soit inversible : il existerait alors une matrice $B = \begin{pmatrix} e & f \\ g & h \end{pmatrix}$ telle que $AB = I_2$.

Or, puisque $ad - bc = 0$, les vecteurs $\begin{pmatrix} a \\ c \end{pmatrix}$ et $\begin{pmatrix} b \\ d \end{pmatrix}$ sont colinéaires. Deux cas peuvent se présenter : soit le vecteur $\begin{pmatrix} a \\ c \end{pmatrix}$ est nul, soit il existe un scalaire λ tel que : $\begin{pmatrix} b \\ d \end{pmatrix} = \lambda \begin{pmatrix} a \\ c \end{pmatrix}$.

Premier cas : le vecteur $\begin{pmatrix} a \\ c \end{pmatrix}$ est nul. Dans ce cas :

$$AB = I_2 \iff \begin{pmatrix} 0 & b \\ 0 & d \end{pmatrix} \begin{pmatrix} e & f \\ g & h \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \iff \begin{pmatrix} bg & bh \\ dg & dh \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$\text{D'où : } AB = I_2 \iff \begin{cases} bg = 1 & (L1) \\ bh = 0 & (L2) \\ dg = 0 & (L3) \\ dh = 1 & (L4) \end{cases}$$

On déduit de (L2) que $b = 0$ ou $h = 0$. Dans le premier cas, l'égalité (L1) ne peut être satisfaite, et dans le second, c'est l'égalité (L4) qui ne l'est pas. Contradiction.

Second cas : il existe un scalaire λ tel que : $\begin{pmatrix} b \\ d \end{pmatrix} = \lambda \begin{pmatrix} a \\ c \end{pmatrix}$.

Par suite : $A = \begin{pmatrix} a & \lambda a \\ c & \lambda c \end{pmatrix}$. On a alors :

$$AB = I_2 \iff \begin{pmatrix} a & \lambda a \\ c & \lambda c \end{pmatrix} \begin{pmatrix} e & f \\ g & h \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \iff \begin{pmatrix} a(e + \lambda g) & a(f + \lambda h) \\ c(e + \lambda g) & c(f + \lambda h) \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$\text{D'où : } AB = I_2 \iff \begin{cases} a(e + \lambda g) = 1 & (L1) \\ a(f + \lambda h) = 0 & (L2) \\ c(e + \lambda g) = 0 & (L3) \\ c(f + \lambda h) = 1 & (L4) \end{cases}$$

On déduit de (L2) que $a = 0$ ou $(f + \lambda h) = 0$. Or dans le premier cas, l'égalité (L1) n'est pas satisfaite, et dans le second, c'est (L4) qui ne l'est pas. Contradiction.

Il s'ensuit que A n'est pas inversible lorsque $ad - bc = 0$.

En d'autres termes : $[ad - bc = 0] \implies [A \notin \text{GL}_n(\mathbb{K})]$. Soit encore : $[A \in \text{GL}_n(\mathbb{K})] \implies [ad - bc \neq 0]$.

Ce qui achève la preuve de ce théorème. ☕

DÉFINITION 13.17 - Mêmes notations que précédemment. Le scalaire $ad - bc$ est appelé **déterminant** de la matrice A et est noté $\det A$.

En utilisant le déterminant, le théorème précédent peut être énoncé comme suit :

$$\left[A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \text{GL}_2(\mathbb{K}) \right] \iff [\det(A) \neq 0]$$

Et lorsque $\det A \neq 0$ on a : $A^{-1} = \frac{1}{\det(A)} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$

Exemples.

1/ La matrice $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ est inversible puisque $\det A = -2$, et $A^{-1} = -\frac{1}{2} \begin{pmatrix} 4 & -2 \\ -3 & 1 \end{pmatrix}$.

2/ La matrice $A = \begin{pmatrix} 1 & 2 \\ 3 & 6 \end{pmatrix}$ n'est pas inversible puisque $\det A = 0$.

3/ Pour tout réel θ , la matrice $A(\theta) = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$ est inversible (car son déterminant vaut 1), et son inverse est la matrice $A(-\theta)$.⁸

Enfin, pour anticiper sur la suite des événements, on peut citer les propriétés suivantes du déterminant dans $M_2(\mathbb{K})$.

PROPRIÉTÉ 13.21 - (Propriétés du déterminant dans $M_2(\mathbb{K})$).

1/ $\forall (A, B) \in M_2(\mathbb{K})^2, \det(AB) = \det(A) \det(B)$

2/ $\forall A \in M_2(\mathbb{K}), \forall \lambda \in \mathbb{K}, \det(\lambda A) = \lambda^2 \det(A)$

3/ $\forall A \in GL_2(\mathbb{K}), \det(A^{-1}) = \frac{1}{\det(A)}$

4/ $\forall A \in M_2(\mathbb{K}), \det({}^t A) = \det(A)$

PREUVE. Toutes ces propriétés peuvent être établies à l'aide de la définition de déterminant dans $M_2(\mathbb{K})$ donnée plus haut ; il s'agit alors de calculs mécaniques somme toute assez froids... Nous établirons ces propriétés en les généralisant et en les interprétant lors du chapitre consacré au déterminant au second semestre. ☕

13.3.3 Matrices inversibles et systèmes linéaires

THÉORÈME 13.3 - (Caractérisation des éléments de $GL_n(\mathbb{K})$). Soit $A \in M_n(\mathbb{K})$.

A est inversible si et seulement si pour tout $B \in \mathbb{K}^n$, le système $AX = B$ admet une (unique) solution. Lorsque tel est le cas, l'unique solution du système $AX = B$ est $A^{-1}B$ (qui est en effet un élément de $\mathbb{K}^n$).

Terminologie. Lorsque A est inversible, le système $AX = B$ est dit **de Cramer**.

PREUVE. On raisonne par double implication. Soit $A \in M_n(\mathbb{K})$.

► Supposons que A est inversible. Soit $B \in \mathbb{K}^n$ un vecteur quelconque.

Alors $A(A^{-1}B) = (AA^{-1})B = I_n B = B$. Donc le vecteur $X_0 = A^{-1}B$ est solution du système $AX = B$.

Établissons son unicité : soit $X_1 \in \mathbb{K}^n$ un vecteur solution du système $AX = B$. Alors : $AX_0 = AX_1$. Puisque A est inversible, on en déduit que : $X_0 = X_1$ (il suffit de multiplier à gauche par A^{-1} les deux termes de l'égalité précédente).

En résumé : si A est inversible, alors pour tout $B \in \mathbb{K}^n$, le système $AX = B$ admet une unique solution.

8. Cette propriété matricielle devient très intuitive dès que l'on sait que la matrice $A(\theta)$ "correspond" à une rotation de centre O et d'angle θ dans le plan. Cette correspondance sera précisée au second semestre.

► Réciproquement, supposons que pour tout $B \in \mathbb{K}^n$, le système $AX = B$ admet une unique solution.

$$\text{Notons alors : } e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, e_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \dots, e_n = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}.$$

Soit j un entier quelconque entre 1 et n . Par hypothèse, il existe un unique vecteur X_j tel que : $AX_j = e_j$.

$$\text{En notant } X_j = \begin{pmatrix} x_{1j} \\ \vdots \\ x_{nj} \end{pmatrix}, \text{ ceci signifie que : } A \begin{pmatrix} x_{1j} \\ \vdots \\ \vdots \\ x_{nj} \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \quad \longleftarrow j\text{-ème place}$$

Formons alors une matrice C en concaténant les colonnes $X_1, \dots, X_n$ (dans l'ordre) :

$$C = \begin{pmatrix} x_{11} & \cdots & x_{1j} & \cdots & x_{1n} \\ \vdots & \ddots & & & \vdots \\ \vdots & & \ddots & & \vdots \\ \vdots & & & \ddots & \vdots \\ \underbrace{x_{n1}}_{X_1} & \cdots & \underbrace{x_{nj}}_{X_j} & \cdots & \underbrace{x_{nn}}_{X_n} \end{pmatrix}$$

Alors, par construction des vecteurs X_j , on a :

$$AC = A \begin{pmatrix} x_{11} & \cdots & x_{1j} & \cdots & x_{1n} \\ \vdots & \ddots & & & \vdots \\ \vdots & & \ddots & & \vdots \\ \vdots & & & \ddots & \vdots \\ \underbrace{x_{n1}}_{X_1} & \cdots & \underbrace{x_{nj}}_{X_j} & \cdots & \underbrace{x_{nn}}_{X_n} \end{pmatrix} = \begin{pmatrix} 1 & \cdots & 0 & \cdots & 1 \\ \vdots & \ddots & & & \vdots \\ \vdots & & 1 & & \vdots \\ \vdots & & & \ddots & \vdots \\ \underbrace{1}_{e_1} & \cdots & \underbrace{0}_{e_j} & \cdots & \underbrace{1}_{e_n} \end{pmatrix} = I_n$$

Une nouvelle application de la propriété 13.18 permet de conclure que A est inversible et que $A^{-1} = C$.

En résumé : si pour tout $B \in \mathbb{K}^n$, le système $AX = B$ admet une (unique) solution, alors A est inversible.

Ce qui achève la preuve de la réciproque, et donc de l'équivalence du théorème. ☕

Remarque. Cet énoncé est d'une **très grande** utilité pratique, car il fournit une méthode pour déterminer concrètement si une matrice carrée de $M_n(\mathbb{K})$ est inversible ou non ; et dans l'affirmative, on peut expliciter son inverse (voir exemples traités en cours, avec la méthode " $AX = B$ ").

Exemple 1. Soit la matrice $A = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}$.

On veut déterminer si A est inversible, et dans l'affirmative obtenir A^{-1} .

A cette fin, on introduit un vecteur $B = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$ et on résout le système : $AX = B$ où $X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ désigne un vecteur inconnu.

Pratiquement :

$$AX = B \iff \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} \iff \begin{cases} x_1 - x_2 = b_1 \\ x_2 - x_3 = b_2 \\ x_3 = b_3 \end{cases} \iff \begin{cases} x_1 = b_1 + b_2 + b_3 \\ x_2 = b_2 + b_3 \\ x_3 = b_3 \end{cases}$$

Il s'ensuit que le système $AX = B$ admet une unique solution (pour tout choix de B), puisque le triplet (x_1, x_2, x_3) est déterminé de manière unique en fonction de b_1, b_2 et b_3 . On en déduit déjà que A **est inversible**.

Puisqu'en outre on sait que dans ce cas l'unique solution de $AX = B$ est $A^{-1}B$, la lecture des coefficients intervenant devant b_1, b_2 et b_3 dans les expressions de x_1, x_2 et x_3 permet d'obtenir A^{-1} , à savoir :

$$A^{-1} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}.$$

Exemple 2. On considère la matrice : $B = \begin{pmatrix} 1 & -1 & 0 & 0 \\ -1 & 1 & -1 & 0 \\ 0 & -1 & 1 & -1 \\ 0 & 0 & -1 & 1 \end{pmatrix} \in M_4(\mathbb{K})$.

On veut déterminer si B est inversible, et dans l'affirmative obtenir B^{-1} .

On fixe donc C un vecteur arbitraire de $\mathbb{K}^4$, et on cherche $X \in \mathbb{K}^4$ tel que $BX = C$.

$$BX = C \iff \begin{cases} x_1 - x_2 = c_1 \\ -x_1 + x_2 - x_3 = c_2 \\ -x_2 + x_3 - x_4 = c_3 \\ -x_3 + x_4 = c_4 \end{cases} \iff \begin{cases} x_2 = -c_3 - c_4 \\ x_1 = c_1 - c_3 - c_4 \\ x_3 = -c_1 - c_2 \\ x_4 = -c_1 - c_2 + c_4 \end{cases}$$

Conclusion : B est inversible et $B^{-1} = \begin{pmatrix} 1 & 0 & -1 & -1 \\ 0 & 0 & -1 & -1 \\ -1 & -1 & 0 & 0 \\ -1 & -1 & 0 & 1 \end{pmatrix}$.

Remarque. Puisque B est clairement symétrique, elle est égale à sa transposée. D'où $({}^t B)^{-1} = B^{-1}$.

Exemple 3. En suivant le même plan, on peut montrer que la matrice $P = \begin{pmatrix} 1 & 3 & 12 \\ -1 & -1 & 16 \\ 0 & -2 & 7 \end{pmatrix}$ est

inversible, et que $P^{-1} = \begin{pmatrix} \frac{5}{14} & -\frac{9}{14} & \frac{6}{7} \\ \frac{1}{10} & \frac{1}{10} & -\frac{2}{5} \\ \frac{1}{35} & \frac{1}{35} & \frac{1}{35} \end{pmatrix}$.

Exemple 4. Considérons la matrice $A = \begin{pmatrix} 1 & 1 & 2 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix}$.

On veut déterminer si A est inversible, et dans l'affirmative obtenir A^{-1} .

A cette fin, on introduit un vecteur $B = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$ et on résout le système : $AX = B$ où $X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ désigne un vecteur inconnu.

$$\text{Or : } AX = B \iff \begin{pmatrix} 1 & 1 & 2 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} \iff \begin{cases} x_1 + x_2 + 2x_3 = b_1 \\ x_2 - x_3 = b_2 \\ 0 = b_3 \end{cases}$$

Clairement, la dernière équation n'est vérifiée que lorsque b_3 est nul, et le système $AX = B$ n'admet donc pas de solution pour tout $B \in \mathbb{K}^3$. Il s'ensuit que **la matrice A n'est pas inversible**.

Remarque. Plus généralement d'ailleurs, une matrice possédant une ligne ou une colonne nulle n'est pas inversible.⁹

Exemple 5. Une matrice peut encore être non inversible de manière “moins évidente” que sur l'exemple précédent.

Considérons la matrice $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$.

On veut déterminer si A est inversible, et dans l'affirmative obtenir A^{-1} .

A cette fin, on introduit un vecteur $B = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$ et on résout le système : $AX = B$ où $X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ désigne un vecteur inconnu.

9. Cela encore, nous l'établirons au second semestre lorsque nous étudierons le rang des matrices et des applications linéaires.

$$\begin{aligned} \text{Or : } AX = B &\iff \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} \iff \begin{cases} x_1 + 2x_2 + 3x_3 = b_1 \\ 4x_1 + 5x_2 + 6x_3 = b_2 \\ 7x_1 + 8x_2 + 9x_3 = b_3 \end{cases} \\ &\iff \begin{cases} x_1 + 2x_2 + 3x_3 = b_1 \\ -3x_2 - 6x_3 = b_2 - 4b_1 \\ -6x_2 - 12x_3 = b_3 - 7b_1 \end{cases} \iff \begin{cases} x_1 + 2x_2 + 3x_3 = b_1 \\ -3x_2 - 6x_3 = b_2 - 4b_1 \\ 0 = b_1 - 2b_2 + b_3 \end{cases} \end{aligned}$$

Encore une fois, la dernière équation n'est vérifiée que de manière conditionnelle, lorsque $b_1 - 2b_2 + b_3 = 0$. Le système $AX = B$ n'admet donc pas de solution pour tout $B \in \mathbb{K}^3$. On peut donc conclure que **la matrice A n'est pas inversible**.

On présente ci-dessous une méthode de résolution des systèmes linéaires, celle dite du *pivot de Gauss*. Elle consiste de fait à présenter différemment les calculs faits lors de la résolution du système $AX = B$.

► Une autre présentation des calculs.

Le système linéaire que l'on est conduit à résoudre lors du calcul de l'inverse d'une matrice peut¹⁰ être effectué systématiquement grâce à un algorithme, appelé algorithme du pivot de Gauss. Celui-ci consiste essentiellement à faire des combinaisons linéaires sur les équations du système que l'on doit résoudre, de manière à se ramener à un système triangulaire (c'est ce que nous avons fait lors des exemples précédents).

En outre, on peut présenter les calculs en faisant apparaître uniquement les coefficients des différentes inconnues dans le système, et non les équations dans leur intégralité.

Comme un petit exemple vaut mieux qu'un long discours, on présente sur l'exemple ci-dessous la :

Méthode du pivot de Gauss pour la résolution des systèmes linéaires

Considérons la matrice $A = \begin{pmatrix} 2 & -1 & -3 \\ 1 & -1 & -1 \\ -3 & 3 & 4 \end{pmatrix}$. Montrons que A est inversible et calculons son inverse par la méthode du pivot de Gauss.

► **Etape 1 : construction de ce que l'on appelle la “matrice augmentée”,** c'est-à-dire la matrice obtenue en “concaténant” A et la matrice I_3 :

$$M = \begin{pmatrix} 2 & -1 & -3 & 1 & 0 & 0 \\ 1 & -1 & -1 & 0 & 1 & 0 \\ -3 & 3 & 4 & 0 & 0 & 1 \end{pmatrix}$$

► **Etape 2 : on fait apparaître des zéros dans la première colonne**

$$M = \begin{pmatrix} 2 & -1 & -3 & 1 & 0 & 0 \\ 0 & 1 & -1 & 1 & -2 & 0 \\ 0 & 3 & -1 & 3 & 0 & 2 \end{pmatrix} \begin{array}{l} L_1 \\ L_2 \leftarrow L_1 - 2L_2 \\ L_3 \leftarrow 3L_1 + 2L_3 \end{array}$$

► **Etape 3 : on fait apparaître un zéro à la dernière ligne de la seconde colonne**

$$M = \begin{pmatrix} 2 & -1 & -3 & 1 & 0 & 0 \\ 0 & 1 & -1 & 1 & -2 & 0 \\ 0 & 0 & -4 & 0 & 6 & 2 \end{pmatrix} \begin{array}{l} L_1 \\ L_2 \\ L_3 \leftarrow L_3 - 3L_2 \end{array}$$

10. Remarque pour l'avenir : “peut et non “doit”.

➤ **Etape 4 : et on remonte ! On fait apparaître des zéros sur la dernière colonne du système**

$$M = \begin{pmatrix} 4 & -2 & 0 & 2 & 18 & 6 \\ 0 & 2 & 0 & 2 & 2 & 2 \\ 0 & 0 & 2 & 0 & 6 & 2 \end{pmatrix} \begin{array}{l} L_1 \leftarrow 2L_1 + 3L_3 \\ L_2 \leftarrow L_3 + 2L_2 \\ L_3 \end{array}$$

➤ **Etape 5 : Enfin, on fait apparaître un zéro à la première ligne, deuxième colonne**

$$M = \begin{pmatrix} 4 & 0 & 0 & 4 & 20 & 8 \\ 0 & 1 & 0 & 2 & 10 & 2 \\ 0 & 0 & 2 & 0 & 6 & 2 \end{pmatrix} \begin{array}{l} L_1 \leftarrow L_1 + L_2 \\ L_2 \\ L_3 \end{array}$$

➤ **Etape 6 : il ne reste plus qu'à faire apparaître des 1 sur la diagonale**

$$M = \begin{pmatrix} 1 & 0 & 0 & 1 & 5 & 2 \\ 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 3 & 1 \end{pmatrix} \begin{array}{l} L_1 \leftarrow L_1/4 \\ L_2 \leftarrow L_2/2 \\ L_3 \leftarrow L_3/2 \end{array}$$

Conclusion : A est inversible et $A^{-1} = \begin{pmatrix} 1 & 5 & 2 \\ 1 & 1 & 1 \\ 0 & 3 & 1 \end{pmatrix}$.

13.3.4 Synthèse (nécessaire et suffisante ?) sur les matrices inversibles

Pour commencer, insistons sur le fait que la question “la matrice A est-elle inversible ?” n’a de sens que pour les matrices carrées. Cette observation faite, rappelons la définition :

Définition : une matrice carrée $A = (a_{ij})_{1 \leq i, j \leq n}$ de $M_n(\mathbb{K})$ est **inversible** s’il existe une matrice $B \in M_n(\mathbb{K})$ telle que : $AB = BA = I_n$.

Cette définition fournit une première stratégie pour déterminer si une matrice est inversible et le cas échéant pour obtenir son inverse. Explicitement :

Méthode 1 : pour déterminer si une matrice carrée donnée A est inversible, on cherche une matrice B (à coefficients inconnus) telle que $AB = I_n$ et $BA = I_n$.

Exemple 1 : soit la matrice $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$. Pour établir que A est inversible, on cherche une matrice $B = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ telle que $AB = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ **et** $BA = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. Chacune de ces égalités donne lieu à un système de 4 équations à quatre inconnues, dont la résolution n’est pas insurmontable, mais le processus est un peu long et rébarbatif.

Evidemment, les choses se compliquent si l’on souhaite utiliser cette méthode pour inverser une matrice 3×3 , puisqu’il s’agira cette fois de résoudre deux systèmes de 9 équations à 9 inconnues... Et plus généralement :

Inconvénient de la méthode 1 : pour inverser une matrice carrée de taille n par la méthode ci-dessus, il faut résoudre **deux** systèmes de n^2 équations à n^2 inconnues¹¹.

11. Ce qui fait beaucoup.

Heureusement, la propriété suivante permet de simplifier la méthode 1.

Propriété : une matrice carrée $A = (a_{ij})_{1 \leq i, j \leq n}$ de $M_n(\mathbb{K})$ est **inversible** s'il existe une matrice $B \in M_n(\mathbb{K})$ telle que : $AB = I_n$ ou $BA = I_n$.

On déduit de cette propriété une spectaculaire(!) amélioration de la première méthode :

Méthode 2 : pour déterminer si une matrice carrée donnée A est inversible, on cherche une matrice B (à coefficients inconnus) telle que $AB = I_n$ ou $BA = I_n$.

Pour reprendre la situation de l'exemple 1, il suffira donc de résoudre un seul des deux systèmes de quatre équations à quatre inconnues ; ce qui est un peu moins long, mais qui reste lourd. Pour donner un autre

exemple, le calcul de l'inverse de $A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 2 & -1 & 1 \\ 0 & 0 & -1 & 3 \\ 0 & 0 & 0 & -2 \end{pmatrix}$ passera par la sympathique résolution d'un

système de 16 équations à autant d'inconnues. Et re-plus généralement :

Inconvénient de la méthode 2 : pour inverser une matrice carrée de taille n par la méthode présentée ci-dessus, il faut résoudre **un** système de n^2 équations à n^2 inconnues¹².

La véritable avancée est donnée par le théorème fondamental de ce chapitre, à savoir :

Théorème : soit $A \in M_n(\mathbb{K})$. La matrice A est inversible si et seulement si pour tout $B \in \mathbb{K}^n$, le système $AX = B$ admet une (unique) solution. Et lorsque A est inversible, l'unique solution du système $AX = B$ est $A^{-1}B$.

On en déduit la

Méthode 3 : pour déterminer si une matrice carrée donnée $A \in M_n(\mathbb{K})$ est inversible, on résout le système $AX = B$, où B désigne un vecteur quelconque de $\mathbb{K}^n$.

Exemple 2 : voir l'exemple 1, page 328 pour le calcul de l'inverse de $A = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}$.

Atout de la méthode 3 : pour inverser une matrice carrée de taille n , il suffit à présent de résoudre un système de n équations à n inconnues¹³.

Pour être complet et tenter d'être impartial, disons que cette dernière méthode peut encore comporter des inconvénients. En premier lieu, il n'est pas toujours rapide de résoudre un système de n équations à n inconnues. C'est évident lorsque n devient grand ; mais même pour $n = 3$, les calculs peuvent être un peu pénibles. Il n'y a toutefois pas grand chose à faire pour surmonter cet inconvénient : à un moment ou un autre, il faudra bien se lancer dans les calculs !

En second lieu, cette belle stratégie ne tient la route que si la matrice A est donnée explicitement, c'est-à-dire si ses coefficients sont connus. Si A est donnée implicitement, par exemple : "soit A une matrice de $M_n(\mathbb{K})$ telle que $A^3 - 2A^2 + 5A - I_n = 0_{M_n(\mathbb{K})}$ ", la méthode 3 ne peut¹⁴ pas être appliquée. On privilégiera dans ce cas la

12. Deux fois moins que beaucoup, c'est encore beaucoup (trop en tous les cas).

13. Ce qui fait beaucoup moins que précédemment, mais encore pas mal.

14. Et ne doit.

Méthode 4 : pour déterminer si une matrice carrée $A \in M_n(\mathbb{K})$ donnée implicitement est inversible, on tente d'utiliser la propriété qui caractérise cette matrice pour faire apparaître une relation du type $A \times B = I_n$ (ou $B \times A = I_n$). Si l'on y parvient, on peut conclure que la matrice A est inversible et $A^{-1} = B$.

Exemple 3 : reprenons la situation évoquée plus haut. Soit $A \in M_n(\mathbb{K})$ telle que $A^3 - 2A^2 + 5A - I_n = 0_{M_n(\mathbb{K})}$. On réécrit cette égalité : $A^3 - 2A^2 + 5A = I_n$, puis : $A(A^2 - 2A + 5I_n) = I_n$. On en déduit que A est inversible, et que $A^{-1} = A^2 - 2A + 5I_n$.

Pour sortir du cas général (n quelconque), indiquons pour finir que dans la situation particulière des matrices carrées de taille 2, on dispose cette fois d'une formule explicite pour traiter les questions d'inversibilité.

Théorème : $\left[A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \text{GL}_2(\mathbb{K}) \right] \iff [\det(A) \neq 0]$.

Et lorsque $\det(A) \neq 0$, on a : $A^{-1} = \frac{1}{\det(A)} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$.

D'où la

Méthode 5 : pour déterminer si une matrice carrée $A \in M_2(\mathbb{K})$ est inversible, on calcule son déterminant. S'il est non nul, la matrice A est inversible et on applique la formule du théorème pour obtenir A^{-1} .

Remarque (culturelle) : la formule donnant l'inverse de A dans $M_2(\mathbb{K})$ admet une généralisation à $M_n(\mathbb{K})$ avec n entier quelconque. Cette formule généraliste n'est plus au programme de Sup (ni de Spé), et présente essentiellement un intérêt théorique. Sans aller trop loin dans le détail, disons qu'une matrice $A \in M_n(\mathbb{K})$ est inversible SSI son déterminant est non nul¹⁵, et lorsque tel est le cas, l'inverse de A est : $A^{-1} = \frac{1}{\det(A)} {}^t\text{co}(A)$, où $\text{co}(A)$ désigne la célèbre **comatrice** de A , dont les coefficients ne sont autres que les non moins célèbres **cofacteurs** de A . Je ne souhaite pas rentrer davantage dans le détail (de la définition et du calcul de ces cofacteurs), mais juste indiquer que le calcul de l'inverse d'une matrice " 3×3 " par cette méthode demande de calculer 9 déterminants " 2×2 "; le calcul de l'inverse d'une matrice " 4×4 " exige de calculer 144 déterminants " 2×2 "; et celui d'une matrice " $n \times n$ " exige d'en calculer $(n!)^2/4 \dots$. Et le carré de $n!$, même divisé par 4, c'est très vite un très très gros nombre.¹⁶

Pour conclure, disons qu'il existe quelques cas particuliers triviaux où le recours à l'une des méthodes précédemment exposées serait très maladroit. Pensez en particulier au cas de la matrice $\text{diag}(\lambda_1, \dots, \lambda_n)$, qui est inversible SSI tous les λ_i sont non nuls ; et dans ce cas, on a : $[\text{diag}(\lambda_1, \dots, \lambda_n)]^{-1} = \text{diag}\left(\frac{1}{\lambda_1}, \dots, \frac{1}{\lambda_n}\right)$.

15. Cette propriété, en particulier la notion de déterminant pour une matrice carrée de taille n , sera en revanche vue cette année.

16. Pour $n = 7$, ce nombre est de l'ordre de 6 millions.

Résumé — Comment montrer que A est inversible ? — Comment calculer A^{-1} ?

ETAPE 1 : A est-elle donnée explicitement (par ses coefficients) ou implicitement (par une relation) ?

➤ Dans le premier cas, passer à l'ETAPE 2.

➤ Dans le second, on essaye d'exploiter la relation pour se ramener à une équation du type $A \times B = I_n$, ce qui permet de conclure (voir méthode 4 et exemple 3 un peu plus haut dans le présent texte).

ETAPE 2 : A est-elle “évidemment” inversible (ou non inversible) ?

➤ Si $A = \lambda I_n$, ou si $A = \text{diag}(\lambda_1, \dots, \lambda_n)$, on peut conclure directement. Notez que votre catalogue de “matrices standards évidemment inversibles” (ou pas) s'enrichira au long de l'année.

➤ Si l'on n'est pas dans une situation “évidente”, on passe à l'ETAPE 3.

ETAPE 3 : A est-elle dans $M_2(\mathbb{K})$?

➤ Si oui, on dispose d'un outil (le déterminant) pour étudier l'inversibilité de A , et d'une formule pour obtenir A^{-1} .

➤ Sinon, on résout le système “ $AX = B$ ” et on conclut en fonction du résultat obtenu.

Remarque : le plan d'étude que l'on vient de présenter donne un algorithme pour déterminer si une matrice est inversible ou non. Vous verrez dans la suite du cours (de cette année et de la prochaine) d'autres outils, qui permettront de perfectionner cet algorithme.

En attendant, considérez-le comme une bouée de sauvetage, à laquelle vous pourrez vous raccrocher en cas de manque d'inspiration. En effet, ce qui n'est pas indiqué dans les lignes précédentes, c'est que l'application de cet algo n'est pas obligatoire. Il y aura toujours une place pour l'originalité et l'esprit d'initiative ; d'ailleurs, une question mathématique ne se ramène pas systématiquement à l'application froide et sans âme d'une recette (ce serait désespérant).

13.4 Matrices semblables

DÉFINITION 13.18 - Deux matrices A et B de $M_n(\mathbb{K})$ sont **semblables** s'il existe une matrice $P \in GL_n(\mathbb{K})$ telle que : $B = P^{-1}AP$.

Exemples.

1/ La matrice identité I_n n'est semblable qu'à elle-même.

2/ La matrice nulle $0_{M_n(\mathbb{K})}$ n'est semblable qu'à elle-même.

3/ Considérons les matrices $A = \begin{pmatrix} 0 & 3/4 \\ -1/4 & -1 \end{pmatrix}$ et $P = \begin{pmatrix} -3 & 1 \\ 1 & -1 \end{pmatrix}$. La matrice P est inversible (car $\det(P) \neq 0$) et : $P^{-1} = -\frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & 3 \end{pmatrix}$.

On a alors $P^{-1}AP = -\frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} 0 & 3/4 \\ -1/4 & -1 \end{pmatrix} \begin{pmatrix} -3 & 1 \\ 1 & -1 \end{pmatrix} = -\frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} 3/4 & -3/4 \\ -1/4 & 3/4 \end{pmatrix}$

Soit : $P^{-1}AP = -\frac{1}{2} \begin{pmatrix} 1/2 & 0 \\ 0 & 3/2 \end{pmatrix}$ d'où finalement : $P^{-1}AP = \begin{pmatrix} -1/4 & 0 \\ 0 & -3/4 \end{pmatrix}$

Les matrices $A = \begin{pmatrix} 0 & 3/4 \\ -1/4 & -1 \end{pmatrix}$ et $D = \begin{pmatrix} -1/4 & 0 \\ 0 & -3/4 \end{pmatrix}$ sont donc **semblables**.

4/ Considérons les matrices $A = \begin{pmatrix} 0 & 3/4 & 0 \\ 3/4 & 0 & 1 \\ 1/4 & 1/4 & 0 \end{pmatrix}$ et $P = \begin{pmatrix} 1 & 3 & 12 \\ -1 & -1 & 16 \\ 0 & -2 & 7 \end{pmatrix}$. La matrice P est inversible

et $P^{-1} = \begin{pmatrix} \frac{5}{14} & -\frac{9}{14} & \frac{6}{7} \\ \frac{1}{10} & \frac{1}{10} & -\frac{2}{5} \\ \frac{1}{35} & \frac{1}{35} & \frac{1}{35} \end{pmatrix}$ (c'est la situation évoquée dans l'exemple 3, page 329).

Dans cette situation, on peut vérifier que : $P^{-1}AP = \begin{pmatrix} -\frac{3}{4} & 0 & 0 \\ 0 & -\frac{1}{4} & 0 \\ 0 & 0 & 1 \end{pmatrix}$.

Notation. Nous conviendrons de noter $A \simeq B$ lorsque A est semblable à B , et nous appellerons relation de similitude la relation binaire “ A est semblable à B ”.

PROPRIÉTÉ 13.22 - La relation de similitude est une relation d'équivalence sur $M_n(\mathbb{K})$, càd :

1/ **Réflexivité** : $\forall A \in M_n(\mathbb{K}), A \simeq A$

2/ **Symétrie** : $\forall (A, B) \in M_n(\mathbb{K})^2, [A \simeq B] \iff [B \simeq A]$

3/ **Transitivité** : $\forall (A, B, C) \in M_n(\mathbb{K})^3, [(A \simeq B) \wedge (B \simeq C)] \implies [A \simeq C]$

PREUVE. 1/ Soit $A \in M_n(\mathbb{K})$. Alors : $A = (I_n)^{-1}AI_n$, donc $A \simeq A$, d'où la réflexivité.

2/ Soient A et B dans $M_n(\mathbb{K})$. Si $A \simeq B$, il existe une matrice $P \in GL_n(\mathbb{K})$ telle que $B = P^{-1}AP$. D'où $A = PBP^{-1}$.

On pose $Q = P^{-1}$. On peut alors affirmer que $Q \in GL_n(\mathbb{K})$ (l'inverse d'une matrice inversible est inversible!) et $Q^{-1} = (P^{-1})^{-1} = P$. Il existe donc une matrice $Q \in GL_n(\mathbb{K})$ telle que $A = Q^{-1}BQ$, d'où : $B \simeq A$, ce qui établit la symétrie.

3/ Soient A, B et C dans $M_n(\mathbb{K})$. Supposons que $A \simeq B$ et $B \simeq C$. Alors il existe deux matrices P et Q dans $GL_n(\mathbb{K})$ telles que $B = P^{-1}AP$ et $C = Q^{-1}BQ$.

On en déduit que : $C = Q^{-1}P^{-1}APQ$, c'est-à-dire : $C = (PQ)^{-1}A(PQ)$. En posant $R = PQ$, on peut ainsi affirmer que $R \in GL_n(\mathbb{K})$, et $C = R^{-1}AR$. D'où $C \simeq A$, ce qui établit la transitivité.

Conclusion : la relation de similitude est une relation d'équivalence dans $M_n(\mathbb{K})$. ☕

Remarque. Au cours de cette démonstration, on a utilisé le fait que le produit matriciel est une loi de composition interne dans $GL_n(\mathbb{K})$ (le produit de deux matrices inversibles est inversible, dans le point 3), associative (pour effectuer les produits “dans l’ordre que l’on souhaite”, preuve du point 3 encore une fois), qu’il existe un élément neutre (qui est I_n) pour cette loi dans $GL_n(\mathbb{K})$ (preuve du point 1), et que l’inverse d’un élément de $GL_n(\mathbb{K})$ est encore dans $GL_n(\mathbb{K})$ (preuve du point 2). Bref on a utilisé toutes les propriétés permettant d’affirmer que $(GL_n(\mathbb{K}), \times)$ est un groupe (pour mémoire, qui n’est pas abélien pour $n \geq 2$).

PROPRIÉTÉ 13.23 - Deux matrices semblables ont même trace, càd :

$$\forall (A, B) \in M_n(\mathbb{K})^2, [A \simeq B] \implies [\text{tr}(A) = \text{tr}(B)]$$

Dans ce contexte, on dit que la trace est un **invariant de similitude**.

PREUVE. Soient A et B deux matrices semblables dans $M_n(\mathbb{K})$. Il existe une matrice $P \in GL_n(\mathbb{K})$ telle que : $B = P^{-1}AP$.

Il s’ensuit que : $\text{tr}(B) = \text{tr}(P^{-1}AP) = \text{tr}(PP^{-1}A) = \text{tr}(A)$.

Conclusion : deux matrices semblables ont même trace . ☕

On peut observer que la réciproque de l’implication précédente est fautive : il ne suffit pas que deux matrices aient même trace pour qu’elles soient semblables. Pour s’en convaincre, on peut considérer une matrice A non nulle et de trace nulle ; dans ce cas, A a une trace nulle, mais n’est pas semblable à la matrice nulle puisque cette dernière n’est semblable qu’à elle-même.

Explicitement : les matrices $A = \begin{pmatrix} 1 & 4 \\ 2 & -1 \end{pmatrix}$ et $0_{M_2(\mathbb{K})} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ ont même trace, mais ne sont pas semblables.

Suivant le même type de raisonnement : $A = \begin{pmatrix} 1 & 4 & 2 \\ 2 & -1 & 3 \\ 5 & 0 & 1 \end{pmatrix}$ et $I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ ont même trace, mais ne sont pas semblables.

On achève ce chapitre par deux “propriétés” extrêmement utiles en pratique. Le plus souvent, **on vous demandera de ne pas utiliser ces propriétés telles quelles**, mais de savoir les redémontrer. Dans le contexte des Concours, il s’agit donc de “questions classiques” plutôt que de propriétés.

PROPRIÉTÉ 13.24 - (Exercice classique). Soient A et D deux matrices de $M_n(\mathbb{K})$.

Si A et D sont semblables, alors A^n et D^n sont semblables pour tout $n \in \mathbb{N}$.

Plus précisément, s’il existe $P \in GL_n(\mathbb{K})$ telle que $D = P^{-1}AP$, alors :

$$\forall n \in \mathbb{N}, \quad D^n = P^{-1}A^nP \quad \text{ou, ce qui revient au même :} \quad \forall n \in \mathbb{N}, \quad A^n = PD^nP^{-1}$$

PREUVE. Soient A et D deux matrices semblables de $M_n(\mathbb{K})$. Il existe donc $P \in GL_n(\mathbb{K})$ telle que : $D = P^{-1}AP$.

Notons $\mathcal{P}(n)$ la propriété “ $D^n = P^{-1}A^nP$ ”. L’initialisation (pour $n = 0$) est immédiate puisque $D^0 = I_n$ et $P^{-1}A^0P = P^{-1}I_nP = P^{-1}P = I_n$. Reste à établir l’hérédité.

Supposons que $\mathcal{P}(n)$ soit vraie pour un certain entier naturel n . Alors $D^{n+1} = D^n \times D$, et en utilisant la définition de D et l'hypothèse de récurrence, on peut écrire : $D^{n+1} = P^{-1}D^nPP^{-1}DP = P^{-1}D^{n+1}P$. Ainsi $\mathcal{P}(n+1)$ est vraie, ce qui prouve l'hérédité.

Conclusion : $\forall n \in \mathbb{N}, D^n = P^{-1}A^nP$. ☕

Cette "propriété" est particulièrement utile pour calculer les puissances d'une matrice, comme l'illustrent les exemples ci-dessous.

Exemple 1. Calcul des puissances de $A = \begin{pmatrix} 0 & 3/4 \\ -1/4 & -1 \end{pmatrix}$.

On a déjà établi (voir exemple 3, page 334) que $P^{-1}AP = D$ avec $P = \begin{pmatrix} -3 & 1 \\ 1 & -1 \end{pmatrix}$ et $D = \begin{pmatrix} -1/4 & 0 \\ 0 & -3/4 \end{pmatrix}$.

D'après ce qui précède, on a donc : $\forall n \in \mathbb{N}, A^n = PD^nP^{-1}$.

Explicitement, pour tout entier naturel n on a :

$$\begin{aligned} A^n &= \begin{pmatrix} -3 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} (-1/4)^n & 0 \\ 0 & (-3/4)^n \end{pmatrix} \begin{pmatrix} -1/2 & 1 \\ 1 & 3 \end{pmatrix} \\ \Leftrightarrow A^n &= -\frac{1}{2} \begin{pmatrix} -3 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} (-1/4)^n & 0 \\ 0 & (-3/4)^n \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 3 \end{pmatrix} \\ \Leftrightarrow A^n &= -\frac{1}{2} \begin{pmatrix} -3 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} (-1/4)^n & (-1/4)^n \\ (-3/4)^n & 3(-3/4)^n \end{pmatrix} \\ \Leftrightarrow A^n &= -\frac{1}{2} \begin{pmatrix} -3\left(-\frac{1}{4}\right)^n + \left(-\frac{3}{4}\right)^n & -3\left(-\frac{1}{4}\right)^n + 3\left(-\frac{3}{4}\right)^n \\ \left(-\frac{1}{4}\right)^n - \left(-\frac{3}{4}\right)^n & \left(-\frac{1}{4}\right)^n - 3\left(-\frac{3}{4}\right)^n \end{pmatrix} \end{aligned}$$

Conclusion : $\forall n \in \mathbb{N}, A^n = \begin{pmatrix} \frac{3}{2}\left(-\frac{1}{4}\right)^n - \frac{1}{2}\left(-\frac{3}{4}\right)^n & \frac{3}{2}\left(-\frac{1}{4}\right)^n - \frac{3}{2}\left(-\frac{3}{4}\right)^n \\ -\frac{1}{2}\left(-\frac{1}{4}\right)^n + \frac{1}{2}\left(-\frac{3}{4}\right)^n & -\frac{1}{2}\left(-\frac{1}{4}\right)^n + \frac{3}{2}\left(-\frac{3}{4}\right)^n \end{pmatrix}$

Exemple 2. Calcul des puissances de $A = \begin{pmatrix} 0 & 3/4 & 0 \\ 3/4 & 0 & 1 \\ 1/4 & 1/4 & 0 \end{pmatrix}$.¹⁷

On a déjà établi (cf ex 4, page 335) que $P^{-1}AP = D$ avec $P = \begin{pmatrix} 1 & 3 & 12 \\ -1 & -1 & 16 \\ 0 & -2 & 7 \end{pmatrix}$ et $D = \text{diag}(-3/4, -1/4, 1)$.

D'après ce qui précède, on a $\forall n \in \mathbb{N}, A^n = PD^nP^{-1}$

¹⁷. Il s'agit ici d'une question que nous avons laissée sans réponse page 316.

$$\begin{aligned} \iff \forall n \in \mathbb{N}, A^n &= \begin{pmatrix} 1 & 3 & 12 \\ -1 & -1 & 16 \\ 0 & -2 & 7 \end{pmatrix} \begin{pmatrix} \left(-\frac{3}{4}\right)^n & 0 & 0 \\ 0 & \left(-\frac{1}{4}\right)^n & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \frac{5}{14} & -\frac{9}{14} & \frac{6}{7} \\ \frac{1}{10} & \frac{1}{10} & -\frac{2}{5} \\ \frac{1}{35} & \frac{1}{35} & \frac{1}{35} \end{pmatrix} \\ \iff \forall n \in \mathbb{N}, \\ A^n &= \begin{pmatrix} \frac{5}{14} \times \left(-\frac{3}{4}\right)^n + \frac{3}{10} \times \left(-\frac{1}{4}\right)^n + \frac{12}{35} & \frac{3}{10} \times \left(-\frac{1}{4}\right)^n - \frac{9}{14} \times \left(-\frac{3}{4}\right)^n + \frac{12}{35} & \frac{6}{7} \times \left(-\frac{3}{4}\right)^n - \frac{6}{5} \times \left(-\frac{1}{4}\right)^n + \frac{12}{35} \\ -\frac{5}{14} \times \left(-\frac{3}{4}\right)^n - \frac{1}{10} \times \left(-\frac{1}{4}\right)^n + \frac{16}{35} & \frac{9}{14} \times \left(-\frac{3}{4}\right)^n - \frac{1}{10} \times \left(-\frac{1}{4}\right)^n + \frac{16}{35} & \frac{2}{5} \times \left(-\frac{1}{4}\right)^n - \frac{6}{7} \times \left(-\frac{3}{4}\right)^n + \frac{16}{35} \\ \frac{1}{5} - \frac{1}{5} \times \left(-\frac{1}{4}\right)^n & \frac{1}{5} - \frac{1}{5} \times \left(-\frac{1}{4}\right)^n & \frac{1}{5} + \frac{4}{5} \times \left(-\frac{1}{4}\right)^n \end{pmatrix} \end{aligned}$$

Exemple 3. On pose $A = \begin{pmatrix} 4 & -5 & 2 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$.

1. On pose $P = \begin{pmatrix} 4 & 1 & 1 \\ 2 & 1 & 0 \\ 1 & 1 & -1 \end{pmatrix}$. Montrer que P est inversible et déterminer son inverse.

2. Soit $T = P^{-1}AP$. Vérifier que $T = D + N$ où $D = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ et $N = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$.

3. Etablir que : $\forall n \in \mathbb{N}, A^n = PT^nP^{-1}$.

4. Calculer T^n pour tout entier naturel n .

Corrigé. 1. On peut par exemple montrer que P est inversible en résolvant le système $PX = B$ d'inconnue $X \in \mathbb{R}^3$ (et de second membre arbitraire $B = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$). En avant :

$$\begin{aligned} PX = B &\iff \begin{pmatrix} 4 & 1 & 1 \\ 2 & 1 & 0 \\ 1 & 1 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} \iff \begin{cases} 4x_1 + x_2 + x_3 = b_1 \\ 2x_1 + x_2 = b_2 \\ x_1 + x_2 - x_3 = b_3 \end{cases} \\ &\iff \begin{cases} 4x_1 + x_2 + x_3 = b_1 \\ x_2 - x_3 = 2b_2 - b_1 \\ 3x_2 - 5x_3 = 4b_3 - b_1 \end{cases} \iff \begin{cases} 4x_1 + x_2 + x_3 = b_1 \\ x_2 - x_3 = 2b_2 - b_1 \\ -2x_3 = 4b_3 + 2b_1 - 6b_2 \end{cases} \end{aligned}$$

$$\text{D'où : } PX = B \iff \begin{cases} x_3 = -b_1 + 3b_2 - 2b_3 \\ x_2 = -2b_1 + 5b_2 - 2b_3 \\ 4x_1 = 4b_1 - 8b_2 + 4b_3 \end{cases} \iff \begin{cases} x_1 = b_1 - 2b_2 + b_3 \\ x_2 = -2b_1 + 5b_2 - 2b_3 \\ x_3 = -b_1 + 3b_2 - 2b_3 \end{cases}$$

Conclusion. P est inversible et $P^{-1} = \begin{pmatrix} 1 & -2 & 1 \\ -2 & 5 & -2 \\ -1 & 3 & -2 \end{pmatrix}$.

$$\begin{aligned} 2. \text{ On a : } T = P^{-1}AP &= \begin{pmatrix} 1 & -2 & 1 \\ -2 & 5 & -2 \\ -1 & 3 & -2 \end{pmatrix} \begin{pmatrix} 4 & -5 & 2 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 4 & 1 & 1 \\ 2 & 1 & 0 \\ 1 & 1 & -1 \end{pmatrix} \\ &= \begin{pmatrix} 2 & -4 & 2 \\ -3 & 8 & -4 \\ -1 & 3 & -2 \end{pmatrix} \begin{pmatrix} 4 & 1 & 1 \\ 2 & 1 & 0 \\ 1 & 1 & -1 \end{pmatrix} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \end{aligned}$$

Conclusion. $P^{-1}AP = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} = D + N$ avec $D = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ et $N = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$.

3. Notons $\mathcal{P}(n)$ la propriété " $A^n = PT^nP^{-1}$ ". L'initialisation (pour $n = 0$) est immédiate puisque $A^0 = I_3$ et $PT^0P^{-1} = PI_3P^{-1} = PP^{-1} = I_3$.

Etablissons à présent l'hérédité de la propriété. Supposons que $\mathcal{P}(n)$ soit vraie pour un certain entier naturel n . Alors $A^{n+1} = A^n \times A$, et en utilisant la définition de T (à savoir $T = P^{-1}AP$), on peut écrire : $A = PTP^{-1}$. De plus, d'après l'hypothèse de récurrence : $A^n = PT^nP^{-1}$.

Par suite : $A^{n+1} = PT^nP^{-1}PTP^{-1} = PT^nI_3TP^{-1} = PT^nTP^{-1} = PT^{n+1}P^{-1}$, et cette dernière écriture signifie que $\mathcal{P}(n+1)$ est vraie, ce qui prouve l'hérédité.

Conclusion. $\forall n \in \mathbb{N}, A^n = PT^nP^{-1}$.

4. Soit n un entier naturel. On a : $T^n = (D + N)^n$. Observons que : $N^2 = 0$ (N est donc nilpotente). Il s'ensuit que : $\forall k \in \mathbb{N}, k \geq 2, N^k = 0$.

Observons encore que : $DN = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$ et $ND = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$. Puisque D et N commutent, on peut légitimement utiliser la formule du binôme de Newton pour écrire :

$$T^n = \sum_{k=0}^n \binom{n}{k} N^k D^{n-k} = \sum_{k=0}^1 \binom{n}{k} N^k D^{n-k} = D^n + nND^{n-1} = \begin{pmatrix} 2^n & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \underbrace{n \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 2^{n-1} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{=N}$$

D'où finalement : $\forall n \in \mathbb{N}, T^n = \begin{pmatrix} 2^n & 0 & 0 \\ 0 & 1 & n \\ 0 & 0 & 1 \end{pmatrix}$.

On en déduit que : $\forall n \in \mathbb{N}, A^n = P \begin{pmatrix} 2^n & 0 & 0 \\ 0 & 1 & n \\ 0 & 0 & 1 \end{pmatrix} P^{-1}$.

Comme promis, on conclut ce chapitre par une “propriété - exemple classique” :

PROPRIÉTÉ 13.25 - (Exercice classique). Soient A et D deux matrices semblables de $M_n(\mathbb{K})$. Il existe donc $P \in GL_n(\mathbb{K})$ telle que : $D = P^{-1}AP$.

Alors pour toute $M \in M_n(\mathbb{K})$ on a :

$$[M \in \text{COM}(A)] \iff [P^{-1}MP \in \text{COM}(D)]$$

PREUVE. Avec les notations de la propriété :

$$\begin{aligned} [P^{-1}MP \in \text{COM}(D)] &\iff [P^{-1}MPD = DP^{-1}MP] \iff [P^{-1}MPP^{-1}AP = P^{-1}APP^{-1}MP] \\ &\iff [P^{-1}MAP = P^{-1}AMP] \iff [MA = AM] \iff [M \in \text{COM}(A)] \end{aligned}$$

Conclusion : $\forall M \in M_n(\mathbb{K}), [M \in \text{COM}(A)] \iff [P^{-1}MP \in \text{COM}(D)]$. ☕

Exemple d'application 1. Déterminons le commutant de $A = \begin{pmatrix} 0 & 3/4 \\ -1/4 & -1 \end{pmatrix}$.

Dans cette situation, l'utilisation de la propriété précédente est un peu “le marteau pour écraser une mouche”, dans le sens où il n'est pas très compliqué de déterminer par coefficients inconnus les matrices $M \in M_2(\mathbb{K})$ telles que : $AM = MA$.

On sait que $A = \begin{pmatrix} 0 & 3/4 \\ -1/4 & -1 \end{pmatrix}$ est semblable à $D = \begin{pmatrix} -1/4 & 0 \\ 0 & -3/4 \end{pmatrix}$ (cf exemple 3, page 334).

Explicitement : $D = P^{-1}AP$ avec $P = \begin{pmatrix} -3 & 1 \\ 1 & -1 \end{pmatrix}$ (et donc $P^{-1} = -\frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & 3 \end{pmatrix}$).

On sait (c'est un autre exercice classique, voir page 313) que le commutant de D est l'ensemble des matrices diagonales de $M_2(\mathbb{K})$. En d'autres termes :

$$\text{COM}(D) = \left\{ \begin{pmatrix} \alpha & 0 \\ 0 & \beta \end{pmatrix} / \alpha \text{ et } \beta \text{ dans } \mathbb{K} \right\}$$

D'après ce qui précède : $[M \in \text{COM}(A)] \iff [P^{-1}MP \in \text{COM}(D)]$.

D'où : $[M \in \text{COM}(A)] \iff \left[\exists (\alpha, \beta) \in \mathbb{K}^2, P^{-1}MP = \begin{pmatrix} \alpha & 0 \\ 0 & \beta \end{pmatrix} \right]$.

Ainsi : $[M \in \text{COM}(A)] \iff \left[\exists (\alpha, \beta) \in \mathbb{K}^2, M = P \begin{pmatrix} \alpha & 0 \\ 0 & \beta \end{pmatrix} P^{-1} \right]$ (♠)

$$\begin{aligned} \text{Or : } P \begin{pmatrix} \alpha & 0 \\ 0 & \beta \end{pmatrix} P^{-1} &= -\frac{1}{2} \begin{pmatrix} -3 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} \alpha & 0 \\ 0 & \beta \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 3 \end{pmatrix} = -\frac{1}{2} \begin{pmatrix} -3 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} \alpha & \alpha \\ \beta & 3\beta \end{pmatrix} \\ &= -\frac{1}{2} \begin{pmatrix} \beta - 3\alpha & 3(\beta - \alpha) \\ \alpha - \beta & \alpha - 3\beta \end{pmatrix} \end{aligned}$$

Conclusion. $\text{COM}(A) = \left\{ \begin{pmatrix} \beta - 3\alpha & 3(\beta - \alpha) \\ \alpha - \beta & \alpha - 3\beta \end{pmatrix} / (\alpha, \beta) \in \mathbb{K}^2 \right\}$.

Exemple d'application 2. Calcul du commutant de $A = \begin{pmatrix} 5 & 5 & -14 \\ 6 & 6 & -16 \\ 5 & 5 & -14 \end{pmatrix}$.

1. On note P la matrice : $P = \begin{pmatrix} 1 & 1 & 1 \\ 2 & -1 & 1 \\ 1 & 0 & 1 \end{pmatrix}$.

Montrer que P est inversible et calculer son inverse.

2. Calculer la matrice produit $D = P^{-1}AP$ (on vérifiera en particulier que D est diagonale¹⁸).

3. Déterminer $\text{COM}(D)$.

4. En déduire $\text{COM}(A)$.

Corrigé.

1. On utilise la méthode " $PX = B$ " pour établir que P est inversible, et que : $P^{-1} = \begin{pmatrix} 1 & 1 & -2 \\ 1 & 0 & -1 \\ -1 & -1 & 3 \end{pmatrix}$.

2. $D = P^{-1}AP \iff D = \begin{pmatrix} 1 & 1 & -2 \\ 1 & 0 & -1 \\ -1 & -1 & 3 \end{pmatrix} \begin{pmatrix} 5 & 5 & -14 \\ 6 & 6 & -16 \\ 5 & 5 & -14 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 2 & -1 & 1 \\ 1 & 0 & 1 \end{pmatrix}$

$$\iff \begin{pmatrix} 1 & 1 & -2 \\ 1 & 0 & -1 \\ -1 & -1 & 3 \end{pmatrix} \begin{pmatrix} 1 & 0 & -4 \\ 2 & 0 & -4 \\ 1 & 0 & -4 \end{pmatrix} \iff D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -4 \end{pmatrix}.$$

3. Soit N une matrice de $M_3(\mathbb{K})$; il existe 9 scalaires a, b, c, d, e, f, g, h et i tels que $N = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$.

Avec ces notations, on a d'une part : $ND = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -4 \end{pmatrix}$ soit : $ND = \begin{pmatrix} a & 0 & -4c \\ d & 0 & -4f \\ g & 0 & -4i \end{pmatrix}$

Et d'autre part : $DN = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -4 \end{pmatrix} \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$ soit : $DN = \begin{pmatrix} a & b & c \\ 0 & 0 & 0 \\ -4g & -4h & -4i \end{pmatrix}$

Ainsi : $[ND = DN] \iff \left[\begin{pmatrix} a & 0 & -4c \\ d & 0 & -4f \\ g & 0 & -4i \end{pmatrix} = \begin{pmatrix} a & b & c \\ 0 & 0 & 0 \\ -4g & -4h & -4i \end{pmatrix} \right] \iff [b = c = d = f = g = h = 0]$

Il s'ensuit qu'une matrice N de $M_3(\mathbb{R})$ vérifie $[ND = DN]$ si et seulement si il existe trois scalaires a, e et

i tels que $N = \begin{pmatrix} a & 0 & 0 \\ 0 & e & 0 \\ 0 & 0 & i \end{pmatrix}$, c'est-à-dire si et seulement si N est diagonale.

18. Explicitement $\text{diag}(1, 0, -4)$.

Conclusion : l'ensemble des matrices N de $M_3(\mathbb{K})$ telles que $[ND = DN]$ est le sous-espace vectoriel des matrices diagonales de $M_3(\mathbb{K})$.

4. Soit $M \in M_3(\mathbb{K})$. D'après ce que nous avons observé précédemment : $[MA = AM] \iff [ND = DN]$ en ayant posé : $N = P^{-1}MP$, c'est-à-dire $M = PNP^{-1}$. D'après cette observation et la question précédente, M appartient au commutant de A si et seulement s'il existe trois scalaires a , e et i tels que : $M =$

$$P \begin{pmatrix} a & 0 & 0 \\ 0 & e & 0 \\ 0 & 0 & i \end{pmatrix} P^{-1}.$$

$$\begin{aligned} \text{Or : } P \begin{pmatrix} a & 0 & 0 \\ 0 & e & 0 \\ 0 & 0 & i \end{pmatrix} P^{-1} &= \begin{pmatrix} 1 & 1 & 1 \\ 2 & -1 & 1 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} a & 0 & 0 \\ 0 & e & 0 \\ 0 & 0 & i \end{pmatrix} \begin{pmatrix} 1 & 1 & -2 \\ 1 & 0 & -1 \\ -1 & -1 & 3 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 1 & 1 \\ 2 & -1 & 1 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} a & a & -2a \\ e & 0 & -e \\ -i & -i & 3i \end{pmatrix} = \begin{pmatrix} a+e-i & a-i & 3i-2a-e \\ 2a-e-i & 2a-i & e+3i-4a \\ a-i & a-i & 3i-2a \end{pmatrix} \end{aligned}$$

Conclusion : l'ensemble des matrices M de $M_3(\mathbb{K})$ telles que $[MA = AM]$ est :

$$\left\{ \begin{pmatrix} a+e-i & a-i & 3i-2a-e \\ 2a-e-i & 2a-i & e+3i-4a \\ a-i & a-i & 3i-2a \end{pmatrix} / (a, e, i) \in \mathbb{K}^3 \right\}$$

Autrement écrit, $\text{COM}(A)$ est l'ensemble des combinaisons linéaires des matrices :

$$\begin{pmatrix} 1 & 1 & -2 \\ 2 & 2 & -4 \\ 1 & 1 & -2 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 & -1 \\ -1 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{et} \quad \begin{pmatrix} -1 & -1 & 3 \\ -1 & -1 & 3 \\ -1 & -1 & 3 \end{pmatrix}$$

13.5 Epilogue

13.5.1 Un problème classique de réduction

Dans $M_3(\mathbb{R})$, on considère les matrices

$$A = \begin{pmatrix} -3 & -1 & -3 \\ 2 & 3 & 0 \\ 2 & 1 & 2 \end{pmatrix} \quad \text{et} \quad P = \begin{pmatrix} 2 & -1 & -1 \\ -1 & 1 & 2 \\ -1 & 1 & 1 \end{pmatrix}$$

► **PARTIE A - Calcul des puissances de A.**

- 1) Montrer que P est inversible et calculer son inverse.
- 2) On pose $D = P^{-1}AP$. Calculer D (on vérifiera que D est une matrice diagonale).
- 3) Soit n un entier naturel. Exprimer D^n en fonction de n .
- 4) Montrer que pour tout entier naturel n , on a $A^n = PD^nP^{-1}$.

► **PARTIE B - Etude du commutant de A.** On rappelle que pour une matrice $N \in M_3(\mathbb{R})$, le commutant de N désigne l'ensemble noté $\text{COM}(N)$ des matrices $Q \in M_3(\mathbb{R})$ qui commutent avec N , c'est-à-dire telles que $NQ = QN$.

- 5) Soit $M \in M_3(\mathbb{R})$. Etablir l'équivalence : $[M \in \text{COM}(A)] \iff [P^{-1}MP \in \text{COM}(D)]$
- 6) Déterminer $\text{COM}(D)$.
- 7) Dédire de ce qui précède $\text{COM}(A)$.
- 8) Etablir l'existence de trois matrices B_1, B_2 et B_3 dans $M_3(\mathbb{R})$ que l'on explicitera telles que :

$$\forall M \in \text{COM}(A), \exists (\alpha_1, \alpha_2, \alpha_3) \in \mathbb{R}^3, M = \alpha_1 B_1 + \alpha_2 B_2 + \alpha_3 B_3$$

► **PARTIE C - Application à l'étude de trois suites imbriquées.** On définit trois suites réelles u, v et w en posant :

$$u_0 = 1, v_0 = 1, w_0 = 0 \text{ et : } \quad \forall n \in \mathbb{N}, \begin{cases} u_{n+1} = -3u_n - v_n - 3w_n \\ v_{n+1} = 2u_n + 3v_n \\ w_{n+1} = 2u_n + v_n + 2w_n \end{cases}$$

Pour tout entier naturel n , on pose : $X_n = \begin{pmatrix} u_n \\ v_n \\ w_n \end{pmatrix}$.

Soit n un entier naturel quelconque.

- 9) Etablir une relation de récurrence entre X_{n+1} et X_n .
- 10) Dédire de ce qui précède une relation entre X_n et X_0 .
- 11) Déterminer les expressions des termes généraux u_n, v_n et w_n en fonction de n .

Corrigé.

► **PARTIE A - Calcul des puissances de A .**

1) Soit $B = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} \in \mathbb{R}^3$ fixé, et soit $X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{R}^3$ arbitraire. Résolvons le système $PX = B$ par la méthode du pivot de Gauss.

$$[PX = B] \iff \begin{cases} 2x_1 - x_2 - x_3 = b_1 & (L_1) \\ -x_1 + x_2 + 2x_3 = b_2 & (L_2) \\ -x_1 + x_2 + x_3 = b_3 & (L_3) \end{cases} \iff \begin{cases} 2x_1 - x_2 - x_3 = b_1 & (L_1) \\ x_2 + 3x_3 = b_1 + 2b_2 & (L_2) \leftarrow (L_1) + 2(L_2) \\ x_2 + x_3 = b_1 + 2b_3 & (L_3) \leftarrow (L_1) + 2(L_3) \end{cases}$$

$$\iff \begin{cases} 2x_1 - x_2 - x_3 = b_1 & (L_1) \\ x_2 + 3x_3 = b_1 + 2b_2 & (L_2) \\ 2x_3 = 2b_2 - 2b_3 & (L_3) \leftarrow (L_2) - (L_3) \end{cases} \iff \begin{cases} 2x_1 - x_2 - x_3 = b_1 & (L_1) \\ x_2 + 3x_3 = b_1 + 2b_2 & (L_2) \\ x_3 = b_2 - b_3 & (L_3) \end{cases}$$

$$\iff \begin{cases} 2x_1 = 2b_1 + 2b_3 & (L_1) \\ x_2 = b_1 - b_2 + 3b_3 & (L_2) \\ x_3 = b_2 - b_3 & (L_3) \end{cases} \iff \begin{cases} x_1 = b_1 + b_3 & (L_1) \\ x_2 = b_1 - b_2 + 3b_3 & (L_2) \\ x_3 = b_2 - b_3 & (L_3) \end{cases}$$

Conclusion. $P \in \text{GL}_3(\mathbb{R})$ et $P^{-1} = \begin{pmatrix} 1 & 0 & 1 \\ 1 & -1 & 3 \\ 0 & 1 & -1 \end{pmatrix}$

2) Après calculs : $D = P^{-1}AP = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$

3) $\forall n \in \mathbb{N}, D^n = \begin{pmatrix} (-1)^n & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2^n \end{pmatrix}.$

4) Notons $\mathcal{P}(n)$ la propriété " $A^n = PD^nP^{-1}$ ". L'initialisation (pour $n = 0$) est immédiate puisque $A^0 = I_n$ et $PD^0P^{-1} = PI_nP^{-1} = PP^{-1} = I_n$. Reste à établir l'hérédité.

Supposons que $\mathcal{P}(n)$ soit vraie pour un certain entier naturel n . Alors $A^{n+1} = A^n \times A$, et en utilisant la définition de D et l'hypothèse de récurrence, on peut écrire : $A^{n+1} = PD^nP^{-1}PDP^{-1} = PD^{n+1}P^{-1}$. Ainsi $\mathcal{P}(n+1)$ est vraie, ce qui prouve l'hérédité.

Conclusion : $\forall n \in \mathbb{N}, A^n = PD^nP^{-1}$.

► **PARTIE B - Etude du commutant de A .**

5) Soit $M \in \text{M}_3(\mathbb{R})$. On a : $[P^{-1}MP \in \text{COM}(D)] \iff [P^{-1}MPD = DP^{-1}MP]$

$$\iff [P^{-1}MPP^{-1}AP = P^{-1}APP^{-1}MP] \iff [P^{-1}MAP = P^{-1}AMP] \iff [MA = AM] \iff [M \in \text{COM}(A)]$$

Conclusion : $\forall M \in \text{M}_3(\mathbb{R}), [M \in \text{COM}(A)] \iff [P^{-1}MP \in \text{COM}(D)]$.

6) Soit $M \in \text{M}_3(\mathbb{R})$. Il existe 9 réels $a, b, \dots, i$ tels que : $M = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$. La matrice M est dans le commutant de D si et seulement si $MD = DM$.

$$\text{Or : } MD = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} = \begin{pmatrix} -a & b & 2c \\ -d & e & 2f \\ -g & h & 2i \end{pmatrix}$$

$$\text{Et : } DM = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} = \begin{pmatrix} -a & -b & -c \\ d & e & f \\ 2g & 2h & 2i \end{pmatrix}$$

$$\text{On déduit de ces calculs que : } [MD = DM] \iff \begin{cases} b = -b \\ 2c = -c \\ d = -d \\ 2f = f \\ 2g = -g \\ 2h = h \end{cases} \iff \begin{cases} b = 0 \\ c = 0 \\ d = 0 \\ f = 0 \\ g = 0 \\ h = 0 \end{cases}$$

Conclusion : $[M \in \text{COM}(D)] \iff [\exists (a, e, i) \in \mathbb{R}^3, M = \text{diag}(a, e, i)]$. En d'autres termes, le commutant de la matrice D est l'ensemble des matrices diagonales de $M_3(\mathbb{R})$.

7) Soit M une matrice de $M_3(\mathbb{R})$; M est dans le commutant de A si et seulement si $P^{-1}MP$ est dans le commutant de D d'après la question 5. Donc d'après la question précédente :

$$\begin{aligned} [M \in \text{COM}(A)] &\iff [P^{-1}MP \in \text{COM}(D)] \iff [\exists (a, e, i) \in \mathbb{R}^3, P^{-1}MP = \text{diag}(a, e, i)] \\ &\iff [\exists (a, e, i) \in \mathbb{R}^3, M = P \text{diag}(a, e, i) P^{-1}] \end{aligned}$$

$$\text{Or : } P \text{diag}(a, e, i) P^{-1} = \begin{pmatrix} 2a - e & e - i & 2a - 3e + i \\ e - a & 2i - e & 3e - a - 2i \\ e - a & i - e & 3e - a - i \end{pmatrix}$$

Conclusion : $\text{COM}(A) = \left\{ \begin{pmatrix} 2a - e & e - i & 2a - 3e + i \\ e - a & 2i - e & 3e - a - 2i \\ e - a & i - e & 3e - a - i \end{pmatrix} / (a, e, i) \in \mathbb{R}^3 \right\}$

8) D'après la question précédente : $[M \in \text{COM}(A)] \iff [\exists (\alpha_1, \alpha_2, \alpha_3) \in \mathbb{R}^3, M = \alpha_1 B_1 + \alpha_2 B_2 + \alpha_3 B_3]$ en ayant posé :

$$B_1 = \begin{pmatrix} 2 & 0 & 2 \\ -1 & 0 & -1 \\ -1 & 0 & -1 \end{pmatrix}, B_2 = \begin{pmatrix} -1 & 1 & -3 \\ 1 & -1 & 3 \\ 1 & -1 & 3 \end{pmatrix} \text{ et } B_3 = \begin{pmatrix} 0 & -1 & 1 \\ 0 & 2 & -2 \\ 0 & 1 & -1 \end{pmatrix}$$

9) D'après l'énoncé : $[\forall n \in \mathbb{N}, X_{n+1} = AX_n]$.

10) Posons $\mathcal{P}(n)$: " $X_n = A^n X_0$ ", et montrons par récurrence que $\mathcal{P}(n)$ est vraie pour entier naturel n .

Initialisation : pour $n = 0$, on a $A^0 X_0 = I_3 X_0 = X_0$. Ainsi $\mathcal{P}(0)$ est vraie.

Hérédité : on suppose $\mathcal{P}(n)$ vraie pour un certain entier naturel n . Alors :

$$X_{n+1} = AX_n = A(A^n X_0) = (AA^n) X_0 = A^{n+1} X_0$$

Ce qui assure que $\mathcal{P}(n+1)$ est vraie, et établit l'hérédité de la propriété.

Conclusion : $\forall n \in \mathbb{N}, X_n = A^n X_0$.

11) Soit n un entier naturel. D'après la question précédente, on a : $X_n = A^n X_0$. D'où grâce à la question

$$4 : X_n = PD^n P^{-1} X_0. \text{ Or, d'après l'énoncé : } X_0 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}.$$

Calculons :

$$P^{-1}X_0 = \begin{pmatrix} 1 & 0 & 1 \\ 1 & -1 & 3 \\ 0 & 1 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \text{ d'où : } P^{-1}X_0 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$$

$$\text{D'où : } D^n P^{-1}X_0 = \begin{pmatrix} (-1)^n & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2^n \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \text{ soit : } D^n P^{-1}X_0 = \begin{pmatrix} (-1)^n \\ 0 \\ 2^n \end{pmatrix}$$

$$\text{Soit enfin : } PD^n P^{-1}X_0 = \begin{pmatrix} 2 & -1 & -1 \\ -1 & 1 & 2 \\ -1 & 1 & 1 \end{pmatrix} \begin{pmatrix} (-1)^n \\ 0 \\ 2^n \end{pmatrix} \text{ càd : } PD^n P^{-1}X_0 = \begin{pmatrix} 2(-1)^n - 2^n \\ (-1)^{n+1} + 2^{n+1} \\ (-1)^{n+1} + 2^n \end{pmatrix}$$

$$\text{Ainsi : } \forall n \in \mathbb{N}, X_n = \begin{pmatrix} 2(-1)^n - 2^n \\ (-1)^{n+1} + 2^{n+1} \\ (-1)^{n+1} + 2^n \end{pmatrix}.$$

$$\text{On en déduit que : } \forall n \in \mathbb{N}, \begin{cases} u_n = 2(-1)^n - 2^n \\ v_n = (-1)^{n+1} + 2^{n+1} \\ w_n = (-1)^{n+1} + 2^n \end{cases}.$$

13.5.2 Interprétation matricielle du corps des nombres complexes

Préambule

Les nombres complexes ont été introduits dans les mathématiques il y a environ quatre siècles, dans les travaux de Jérôme Cardan¹⁹ et Raphaël Bombelli²⁰. Pour faire court, ces deux mathématiciens ont défini le nombre i comme une solution de l'équation $x^2 = -1$, et c'est cette définition qui a traversé les âges pour intervenir encore dans votre cours de Terminale.

Mais une autre approche des nombres complexes aurait été possible, en utilisant le calcul matriciel. Cette approche constitue l'objet du présent problème, dont le but est d'essayer de vous convaincre que l'on aurait pu définir les nombres complexes autrement. Explicitement, au lieu de voir ici un complexe comme un nombre s'écrivant $a + ib$, on définira un complexe comme une matrice de la forme $aI + bJ$ (où I est la matrice identité I_2 , et J une matrice telle que $J^2 = -I$).

L'intérêt de cette approche est géométrique, dans un sens qui sera rendu explicite dans la dernière partie de ce problème.

L'ensemble des "nombres complexes" que nous définirons dans cet énoncé sera donc un ensemble de matrices, et les différentes parties de ce problème auront pour but de prouver que toutes les propriétés que vous avez vues dans le chapitre "Complexes" de cette année s'adaptent à notre nouveau point de vue.

Pour finir, n'ayez pas peur de ce saut dans l'inconnu ! Un grand nombre des questions de ce problème vous paraîtront évidentes dès lors que vous avez bien révisé votre cours sur les complexes, le calcul matriciel et les groupes... Enfin, les différentes parties de ce problème ne sont pas du tout indépendantes ; mais vous pouvez bien entendu utiliser les résultats d'une question dans une autre.

19. Qui n'était pas biélorusse.

20. Lui non plus.

Notations

Dans $M_2(\mathbb{R})$, on considère les matrices

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{et} \quad J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

On pourra se convaincre aisément que $J^2 = -I$ (★).

On définit les deux sous-ensembles $\widehat{\mathbf{R}}$ et $\widehat{\mathbf{C}}$ de $M_2(\mathbb{R})$ en posant :

$$\widehat{\mathbf{R}} = \{aI / a \in \mathbb{R}\} \quad \text{et} \quad \widehat{\mathbf{C}} = \{aI + bJ / (a, b) \in \mathbb{R}^2\}$$

càd :

$$\widehat{\mathbf{R}} = \left\{ \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix} / a \in \mathbb{R} \right\} \quad \text{et} \quad \widehat{\mathbf{C}} = \left\{ \begin{pmatrix} a & -b \\ b & a \end{pmatrix} / (a, b) \in \mathbb{R}^2 \right\}$$

On pourra observer que $\widehat{\mathbf{R}} \subsetneq \widehat{\mathbf{C}}$.

Pour tout couple (a, b) de nombres réels, on notera $Z(a, b)$ l'élément $aI + bJ = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$ de $\widehat{\mathbf{C}}$.

On notera encore : $\overline{Z(a, b)} = Z(a, -b)$.

Enfin, pour tout élément $Z(a, b)$ de $\widehat{\mathbf{C}}$ on définit son **module** en posant : $|Z(a, b)| = \sqrt{a^2 + b^2}$.

► PARTIE A - Quelques propriétés algébriques de $\widehat{\mathbf{R}}$ et de $\widehat{\mathbf{C}}$

1) Justifier brièvement que l'application

$$\begin{aligned} \varphi : \mathbb{R} &\longrightarrow \widehat{\mathbf{R}} \\ x &\longmapsto Z(x, 0) \end{aligned}$$

est une bijection.

2) Montrer que $(\widehat{\mathbf{C}}, +)$ est un sous-groupe de $(M_2(\mathbb{R}), +)$. En déduire que $(\widehat{\mathbf{C}}, +)$ est un groupe abélien.

3) Montrer que $(\widehat{\mathbf{R}}, +)$ est un sous-groupe de $(\widehat{\mathbf{C}}, +)$. En déduire que $(\widehat{\mathbf{R}}, +)$ est un groupe abélien.

4) Soit $Z(a, b) \in \widehat{\mathbf{C}}$. Vérifier que : $Z(a, b) + \overline{Z(a, b)} = 2aI$ et $Z(a, b) - \overline{Z(a, b)} = 2bJ$.

5) Soient $Z(a, b)$ et $Z(a', b')$ dans $\widehat{\mathbf{C}}$. Calculer : $Z(a, b) \times Z(a', b')$.

6) Soient x un réel et $Z(a, b) \in \widehat{\mathbf{C}}$. Vérifier que : $Z(x, 0) \times Z(a, b) = Z(xa, xb)$.

7) Propriétés du module.

a) Soit $Z(a, b) \in \widehat{\mathbf{C}}$. Montrer que : $Z(a, b) \times \overline{Z(a, b)} = |Z(a, b)|^2 \times I$.

b) Montrer que :

$$\forall (Z(a, b), Z(a', b')) \in \widehat{\mathbf{C}}^2, |Z(a, b) \times Z(a', b')| = |Z(a, b)| \times |Z(a', b')|$$

c) Soit $Z(a, b)$ un élément de $\widehat{\mathbf{C}}$. Etablir que : $\forall n \in \mathbb{N}, \quad |Z(a, b)^n| = |Z(a, b)|^n$.

► PARTIE B - Les corps $\widehat{\mathbf{R}}$ et $\widehat{\mathbf{C}}$

8) Le corps $\widehat{\mathbf{R}}$.

a) Montrer que le produit matriciel (la loi “ $\times$ ”) est une LCI sur $\widehat{\mathbf{R}}$. Justifier brièvement que cette LCI possède un élément neutre dans $\widehat{\mathbf{R}}$, et est commutative.

Par la suite, on pourra admettre que la loi $\times$ est associative, distributive par rapport à l'addition, et que pour tout $Z(x, 0)$ dans $\widehat{\mathbf{R}}$ on a : $\mathbb{Z}(x, 0) \times 0_{M_2(\mathbb{R})} = 0_{M_2(\mathbb{R})} = 0_{M_2(\mathbb{R})} \times Z(x, 0)$. Les propriétés énoncées ci-dessus permettent alors d'affirmer que $(\widehat{\mathbf{R}}, +, \times)$ est un anneau commutatif.

b) Montrer que l'anneau $(\widehat{\mathbf{R}}, +, \times)$ est intègre, et que tout élément non nul de $(\widehat{\mathbf{R}}, +, \times)$ est inversible²¹ dans $\widehat{\mathbf{R}}$.

9) Le corps $\widehat{\mathbf{C}}$.

a) Montrer que le produit matriciel (la loi “ $\times$ ”) est une LCI sur $\widehat{\mathbf{C}}$.

b) Justifier brièvement que cette LCI possède un élément neutre dans $\widehat{\mathbf{C}}$.

Par la suite, on pourra admettre que la loi $\times$ est associative, distributive par rapport à l'addition, et que pour tout $Z(a, b)$ dans $\widehat{\mathbf{C}}$ on a : $\mathbb{Z}(a, b) \times 0_{M_2(\mathbb{R})} = 0_{M_2(\mathbb{R})} = 0_{M_2(\mathbb{R})} \times Z(a, b)$. Les propriétés énoncées ci-dessus permettent alors d'affirmer que $(\widehat{\mathbf{C}}, +, \times)$ est un anneau.

c) Montrer que l'anneau $(\widehat{\mathbf{C}}, +, \times)$ est commutatif.

d) Montrer que tout élément non nul de $\widehat{\mathbf{C}}$ est inversible²² dans $\widehat{\mathbf{C}}$.

e) Dédurre de ce qui précède que l'anneau $(\widehat{\mathbf{C}}, +, \times)$ est intègre, puis que l'anneau $(\widehat{\mathbf{C}}, +, \times)$ est un corps.

f) Soit $Z(a, b) \in \widehat{\mathbf{C}}$, avec $Z(a, b) \neq \mathbf{I}$. Montrer que pour tout entier naturel n on a :

$$\sum_{k=0}^n Z(a, b)^k = (\mathbf{I} - Z(a, b)^{n+1}) \times (\mathbf{I} - Z(a, b))^{-1}$$

► PARTIE C - Notation exponentielle

Pour tout réel θ , on pose :

$$E^{\theta J} = Z(\cos \theta, \sin \theta) \quad \text{càd} \quad E^{\theta J} = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$$

10) Vérifier que pour tout réel θ on a : $|E^{\theta J}| = 1$.

11) Montrer que pour tout couple $(\theta, \varphi) \in \mathbb{R}^2$ on a : $E^{\theta J} \times E^{\varphi J} = E^{(\theta+\varphi)J}$.

21. Pour la loi $\times$.

22. Pour la loi $\times$.

- 12) Montrer que pour tout réel θ on a : $(E^{\theta J})^{-1} = E^{(-\theta)J}$.²³
- 13) Justifier que pour tout couple $(\theta, \varphi) \in \mathbb{R}^2$ on a : $[E^{\theta J} = E^{\varphi J}] \iff [\theta = \varphi \mod{2\pi}]$.

► PARTIE D - Forme exponentielle d'un élément non nul de $\widehat{\mathbf{C}}$

A partir de maintenant, on note $\widehat{\mathbf{C}}^*$ l'ensemble des éléments non nuls de $\widehat{\mathbf{C}}$, càd : $\widehat{\mathbf{C}}^* = \widehat{\mathbf{C}} \setminus \{0_{M_2(\mathbb{R})}\}$.

- 14) Soit $Z(a, b) \in \widehat{\mathbf{C}}^*$. Montrer qu'il existe deux réels α et β que l'on exprimera en fonction de a et b tels que :

$$Z(a, b) = |Z(a, b)| \times Z(\alpha, \beta)$$

- 15) Avec les notations de la question précédente, vérifier que : $\alpha^2 + \beta^2 = 1$.

- 16) En déduire que pour tout $Z(a, b) \in \widehat{\mathbf{C}}^*$, il existe un réel θ tel que

$$Z(a, b) = |Z(a, b)| \times E^{\theta J}$$

Par la suite, ce réel θ sera appelé **argument** de $Z(a, b)$, et il sera noté : $\text{ARG}(Z(a, b))$.

- 17) **Méthode d'identification dans $\widehat{\mathbf{C}}^*$** . Soient $Z(a, b)$ et $Z(a', b')$ dans $\widehat{\mathbf{C}}^*$. Montrer que

$$(Z(a, b) = Z(a', b')) \iff \begin{cases} |Z(a, b)| = |Z(a', b')| \\ \text{ARG}(Z(a, b)) = \text{ARG}(Z(a', b')) \end{cases} \quad [2\pi]$$

- 18) Soient $Z(a, b)$ et $Z(a', b')$ dans $\widehat{\mathbf{C}}$. Montrer que :

$$\text{ARG}(Z(a, b) \times Z(a', b')) = \text{ARG}(Z(a, b)) + \text{ARG}(Z(a', b'))$$

A ce point du raisonnement, on peut alors démontrer toutes les propriétés algébriques attendues des arguments (celles concernant l'argument du conjugué, de l'inverse, d'un quotient). On n'en démontre qu'une seule (question suivante) qui sera d'une importance capitale pour la suite des événements.

- 19) Soit $Z(a, b) \in \widehat{\mathbf{C}}^*$. Montrer que : $\forall n \in \mathbb{N}, \text{ARG}(Z(a, b)^n) = n \text{ARG}(Z(a, b))$

► PARTIE E - Racines carrées dans $\widehat{\mathbf{C}}$

- 20) Résoudre dans $\widehat{\mathbf{C}}$ l'équation $X^2 = I$.²⁴
- 21) Résoudre dans $\widehat{\mathbf{C}}$ l'équation $X^2 = -I$.
- 22) Résoudre dans $\widehat{\mathbf{C}}$ l'équation $X^4 = I$.
- 23) Soit à présent $Z(a, b) \in \widehat{\mathbf{C}}^*$. Montrer que l'équation $X^2 = Z(a, b)$ admet exactement deux solutions dans $\widehat{\mathbf{C}}$ (que l'on pourra appeler racines carrées de $Z(a, b)$ dans $\widehat{\mathbf{C}}$).

23. On a également $(E^{\theta J})^{-1} = \overline{E^{\theta J}}$, mais on ne demande pas de vérifier cette dernière relation.

24. On cherchera donc tous les $X \in \widehat{\mathbf{C}}$ tels que $X^2 = I$. Indication : on pourra utiliser une identité remarquable, et utiliser le fait que $\widehat{\mathbf{C}}$ est en particulier un anneau intègre.

- 24) **Application.** Déterminer les racines carrées dans $\widehat{\mathbf{C}}$ de $A = \begin{pmatrix} 1 & -\sqrt{3} \\ \sqrt{3} & 1 \end{pmatrix}$, c'est-à-dire déterminer toutes les matrices $X \in \widehat{\mathbf{C}}$ telles que $X^2 = A$.

► **PARTIE F - Equations du second degré dans $\widehat{\mathbf{C}}$**

Soient A, B et C dans $\widehat{\mathbf{C}}$, avec $A \neq 0_{\mathbf{M}_2(\mathbb{R})}$.

Dans cette partie on notera (E) l'équation suivante d'inconnue $X \in \widehat{\mathbf{C}}$:

$$(E) : \quad AX^2 + BX + C = 0_{\mathbf{M}_2(\mathbb{R})}$$

On appelle **discriminant de l'équation** (E) l'élément de $\widehat{\mathbf{C}}$ défini en posant

$$\Delta = B^2 - 4AC$$

- 25) Etablir que l'équation (E) possède exactement deux solutions dans $\widehat{\mathbf{C}}$ lorsque $\Delta \neq 0_{\mathbf{M}_2(\mathbb{R})}$; et exactement une solution dans $\widehat{\mathbf{C}}$ lorsque $\Delta = 0_{\mathbf{M}_2(\mathbb{R})}$. Dans les deux cas, on explicitera la ou les solutions.
- 26) **Application.** Résoudre dans $\widehat{\mathbf{C}}$ l'équation :

$$X^2 - 2JX - J = I$$

► **PARTIE G - Racines n -ièmes de l'unité dans $\widehat{\mathbf{C}}$**

Pour tout entier naturel $n \geq 2$, on note :

$$\widehat{\mathbf{U}}_n = \left\{ Z(a, b) \in \widehat{\mathbf{C}} / Z(a, b)^n = I \right\}$$

et on note $\widehat{\mathbf{U}}$ l'ensemble des éléments de $\widehat{\mathbf{C}}$ de module égal à 1, soit :

$$\widehat{\mathbf{U}} = \left\{ Z(a, b) \in \widehat{\mathbf{C}} / |Z(a, b)| = 1 \right\}$$

Dans les questions de cette partie, n désigne un entier naturel $n \geq 2$.

- 27) Soit n un entier naturel $n \geq 2$. Etablir que : $\widehat{\mathbf{U}}_n \subset \widehat{\mathbf{U}}$.
- 28) Etablir que $(\widehat{\mathbf{U}}_n, \times)$ est un sous-groupe de $(\widehat{\mathbf{U}}, \times)$.
- 29) Montrer que : $\widehat{\mathbf{U}}_n = \left\{ \mathbf{E}^{\frac{2k\pi}{n} J} / k \in \llbracket 0, n-1 \rrbracket \right\}$.
- 30) Montrer que : $\sum_{W \in \widehat{\mathbf{U}}_n} W = 0_{\mathbf{M}_2(\mathbb{R})}$.
- 31) **Application.** Déterminer toutes les matrices $X \in \widehat{\mathbf{C}}$ telles que

$$X^4 = \begin{pmatrix} 8\sqrt{2} & -8\sqrt{2} \\ 8\sqrt{2} & 8\sqrt{2} \end{pmatrix}$$

► PARTIE H - Finalement

On rappelle que $\text{Sim}^+(\mathbb{C})$ désigne l'ensemble des similitudes directes du plan complexe.

On note $\text{Sim}_0^+(\mathbb{C})$ la partie de $\text{Sim}^+(\mathbb{C})$ constituée des similitudes directes de centre O .

32) Justifier brièvement qu'un élément de $\text{Sim}_0^+(\mathbb{C})$ est une transformation f_a ayant pour écriture complexe :
 $z \mapsto az$ avec $a \in \mathbb{C}^*$, et que $(\text{Sim}_0^+(\mathbb{C}), \circ)$ est un groupe abélien.

33) Montrer que l'application

$$\begin{aligned} \psi : \text{Sim}_0^+(\mathbb{C}) &\longrightarrow \widehat{\mathbf{C}}^* \\ f_a &\longmapsto |a| \text{E}^{\arg(a)} \end{aligned}$$

est bijective.

34) Montrer que ψ est compatible avec les structures de groupes de $\text{Sim}_0^+(\mathbb{C})$ et de $\widehat{\mathbf{C}}^*$ dans le sens où :

$$\forall (f, g) \in \text{Sim}_0^+(\mathbb{C})^2, \psi(f \circ g) = \psi(f) \times \psi(g)$$

Une telle application (compatible avec les lois de groupes au départ et à l'arrivée) s'appelle un morphisme de groupes. Un morphisme de groupes bijectif est appelé un isomorphisme de groupes ; enfin deux groupes G et G' sont dits isomorphes lorsqu'il existe un morphisme de groupes entre G et G' .

Dans cette dernière question, vous avez donc établi que les groupes $(\text{Sim}_0^+(\mathbb{C}), \circ)$ et $(\widehat{\mathbf{C}}^*, \times)$ sont isomorphes. Cela signifie que non seulement il existe une correspondance bijective entre les éléments des ensembles $\text{Sim}_0^+(\mathbb{C})$ et $\widehat{\mathbf{C}}^*$, mais aussi que cette correspondance est compatible avec les LCI de ces deux groupes.

► PARTIE I - Bonus

35) Déterminer toutes les matrices $M \in \widehat{\mathbf{C}}$ telles que

$$(M + \text{I})^5 = (M - \text{I})^5$$

Corrigé.

1) Il est immédiat que l'application $\psi : \widehat{\mathbf{R}} \longrightarrow \mathbb{R}$ qui à tout $Z(x, 0) \in \widehat{\mathbf{R}}$ associe le réel x est telle que $\psi \circ \varphi = \text{id}_{\mathbb{R}}$ et $\varphi \circ \psi = \text{id}_{\widehat{\mathbf{R}}}$. Il s'ensuit que φ est bijective (et $\varphi^{-1} = \psi$).

2) On utilise la méthode bien connue :

(SG1) : $\widehat{\mathbf{C}} \subset \text{M}_2(\mathbb{R})$, par définition.

(SG2) : l'élément neutre de $(\text{M}_2(\mathbb{R}), +)$, qui est la matrice nulle $0_{\text{M}_2(\mathbb{R})}$, appartient à $\widehat{\mathbf{C}}$ puisque : $0_{\text{M}_2(\mathbb{R})} = Z(0, 0)$.

(SG3) : soient $Z(a, b)$ et $Z(a', b')$ deux éléments de $\widehat{\mathbf{C}}$. Il est clair que $Z(a, b) + Z(a', b') = Z(a + a', b + b')$. L'addition est donc une LCI sur $\widehat{\mathbf{C}}$.

(SG4) : soit $Z(a, b) \in \widehat{\mathbf{C}}$. Son inverse pour l'addition est clairement $Z(-a, -b)$, qui est encore dans $\widehat{\mathbf{C}}$. L'ensemble $\widehat{\mathbf{C}}$ est donc stable par passage à l'inverse.

Conclusion : $(\widehat{\mathbf{C}}, +)$ est un sous-groupe de $(M_2(\mathbb{R}), +)$, abélien car sous-groupe d'un groupe abélien.

3) On ré-utilise la méthode bien connue :

(SG1) : $\widehat{\mathbf{R}} \subset \widehat{\mathbf{C}}$, par définition.

(SG2) : l'élément neutre de $(\widehat{\mathbf{C}}, +)$, qui est la matrice nulle $0_{M_2(\mathbb{R})}$, appartient à $\widehat{\mathbf{R}}$ puisque : $0_{M_2(\mathbb{R})} = Z(0, 0)$.

(SG3) : soient $Z(a, 0)$ et $Z(a', 0)$ deux éléments de $\widehat{\mathbf{R}}$. Il est clair que $Z(a, 0) + Z(a', 0) = Z(a + a', 0)$. L'addition est donc une LCI sur $\widehat{\mathbf{R}}$.

(SG4) : soit $Z(a, 0) \in \widehat{\mathbf{R}}$. Son inverse pour l'addition est clairement $Z(-a, 0)$, qui est encore dans $\widehat{\mathbf{R}}$. L'ensemble $\widehat{\mathbf{R}}$ est donc stable par passage à l'inverse.

Conclusion : $(\widehat{\mathbf{R}}, +)$ est un sous-groupe de $(\widehat{\mathbf{C}}, +)$, abélien car tout sous-groupe d'un groupe abélien l'est.

4) Il est immédiat que : $\forall Z(a, b) \in \widehat{\mathbf{C}}, Z(a, b) + \overline{Z(a, b)} = 2aI$ et $Z(a, b) - \overline{Z(a, b)} = 2bJ$.²⁵

5) Soient $Z(a, b)$ et $Z(a', b')$ dans $\widehat{\mathbf{C}}$. On a :

$$Z(a, b) \times Z(a', b') = \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \begin{pmatrix} a' & -b' \\ b' & a' \end{pmatrix} = \begin{pmatrix} aa' - bb' & -a'b - ab' \\ a'b + ab' & aa' - bb' \end{pmatrix} = Z(aa' - bb', a'b + ab')$$

Conclusion. $\forall (Z(a, b), Z(a', b')) \in \widehat{\mathbf{C}}^2, Z(a, b) \times Z(a', b') = Z(aa' - bb', a'b + ab')$

6) D'après la question 5 : **Conclusion.** $\forall (Z(x, 0), Z(a, b)) \in \widehat{\mathbf{R}} \times \widehat{\mathbf{C}}, Z(x, 0) \times Z(a, b) = Z(xa, xb)$

7) Propriétés du module. a) Soit $Z(a, b) \in \widehat{\mathbf{C}}$. On a :

$$Z(a, b) \times \overline{Z(a, b)} = Z(a, b) \times Z(a, -b) = Z(a^2 + b^2, 0) = (a^2 + b^2) Z(1, 0) = |Z(a, b)|^2 \times I$$

Conclusion. $\forall Z(a, b) \in \widehat{\mathbf{C}}, Z(a, b) \times \overline{Z(a, b)} = |Z(a, b)|^2 \times I$

b) Soient $Z(a, b)$ et $Z(a', b')$ dans $\widehat{\mathbf{C}}$.

$$\text{D'une part : } |Z(a, b) \times Z(a', b')| = |Z(aa' - bb', a'b + ab')| = \sqrt{(aa' - bb')^2 + (a'b + ab')^2}$$

$$= \sqrt{a^2a'^2 - 2aa'bb' + b^2b'^2 + a'^2b^2 + 2aa'bb' + a^2b'^2} = \sqrt{a^2a'^2 + b^2b'^2 + a'^2b^2 + a^2b'^2}$$

$$\text{D'autre part : } |Z(a, b)| \times |Z(a', b')| = \sqrt{a^2 + b^2} \times \sqrt{a'^2 + b'^2} = \sqrt{a^2a'^2 + b^2b'^2 + a'^2b^2 + a^2b'^2}.$$

Conclusion. $\forall (Z(a, b), Z(a', b')) \in \widehat{\mathbf{C}}^2, |Z(a, b) \times Z(a', b')| = |Z(a, b)| \times |Z(a', b')|$

c) Soit $Z(a, b) \in \widehat{\mathbf{C}}$. Notons $\mathcal{P}(n)$ la propriété " $|Z(a, b)^n| = |Z(a, b)|^n$ ". L'initialisation (pour $n = 0$) est immédiate puisque $|Z(a, b)^0| = |I| = 1$ et $|Z(a, b)|^0 = 1$. Reste à établir l'hérédité.

Supposons que $\mathcal{P}(n)$ soit vraie pour un certain entier naturel n . Alors :

$$|Z(a, b)^{n+1}| = |Z(a, b)^n \times Z(a, b)| = |Z(a, b)^n| \times |Z(a, b)| = |Z(a, b)|^n \times |Z(a, b)| = |Z(a, b)|^{n+1}$$

Ainsi $\mathcal{P}(n + 1)$ est vraie, ce qui prouve l'hérédité et achève cette récurrence.

Conclusion. $\forall Z(a, b) \in \widehat{\mathbf{C}}, \forall n \in \mathbb{N}, |Z(a, b)^n| = |Z(a, b)|^n$.

^{25.} Formules à rapprocher de celles que vous connaissez dans le cadre des nombres complexes usuels : $\forall z \in \mathbb{C}, z + \bar{z} = 2\text{Re}(z)$ et $z - \bar{z} = 2i\text{Im}(z)$.

► PARTIE B - Les corps $\widehat{\mathbf{R}}$ et $\widehat{\mathbf{C}}$

8) Le corps $\widehat{\mathbf{R}}$.

a) D'après la question 5 : $\forall (Z(x, 0), Z(y, 0)) \in \widehat{\mathbf{R}}^2$, $Z(x, 0) \times Z(y, 0) = Z(xy, 0) \in \widehat{\mathbf{R}}$. Le produit matriciel est donc une LCI sur $\widehat{\mathbf{R}}$, possédant un élément neutre ($Z(1, 0) = I$), et évidemment commutative.

b) Soient $Z(x, 0)$ et $Z(y, 0)$ dans $\widehat{\mathbf{R}}$. On a : $Z(x, 0) \times Z(y, 0) = 0_{M_2(\mathbb{R})}$ si et seulement si $xy = 0$, càd si et seulement si $x = 0$ ou $y = 0$.

Donc : $[Z(x, 0) \times Z(y, 0) = 0_{M_2(\mathbb{R})}] \iff [(Z(x, 0) = 0_{M_2(\mathbb{R})}) \vee (Z(y, 0) = 0_{M_2(\mathbb{R})})]$

Ce qui signifie que l'anneau $(\widehat{\mathbf{R}}, +, \times)$ est intègre.

Soit $Z(x, 0)$ un élément non nul de $\widehat{\mathbf{R}}$. $Z(x, 0)$ est trivialement inversible d'inverse $Z(1/x, 0)$.

Conclusion. Puisque $(\widehat{\mathbf{R}}, +, \times)$ est un anneau commutatif où tout élément non nul est inversible, on peut affirmer que $(\widehat{\mathbf{R}}, +, \times)$ est un corps.

9) Le corps $\widehat{\mathbf{C}}$.

a) D'après la question 5, le produit de deux éléments de $\widehat{\mathbf{C}}$ est encore un élément de $\widehat{\mathbf{C}}$. Donc le produit matriciel est une LCI sur $\widehat{\mathbf{C}}$.

b) L'élément neutre pour le produit matriciel est la matrice I_2 , qui appartient à $\widehat{\mathbf{C}}$ puisque $I_2 = Z(1, 0)$.

c) Il résulte de la question 5 que le produit matriciel dans $\widehat{\mathbf{C}}$ est commutatif.

d) Soit $Z(a, b)$ un élément non nul de $\widehat{\mathbf{C}}$. Alors : $\det(Z(a, b)) = a^2 + b^2$. D'où $\det(Z(a, b)) \neq 0$ puisque $(a, b) \neq (0, 0)$ par hypothèse. Il s'ensuit que $Z(a, b)$ est inversible, d'inverse : $Z(a, b)^{-1} = \frac{1}{a^2 + b^2} \overline{Z(a, b)}$.

Conclusion. Tout élément non nul de $\widehat{\mathbf{C}}$ est inversible dans $\widehat{\mathbf{C}}$.

e) Soient $Z(a, b)$ et $Z(a', b')$ dans $\widehat{\mathbf{C}}$. Supposons que $Z(a, b)Z(a', b') = 0$.

Distinguons deux cas. Si $Z(a, b) = 0$, l'égalité précédente est satisfaite et il n'y a plus rien à faire.

Si $Z(a, b) \neq 0$, alors d'après la question précédente $Z(a, b)$ est inversible dans $\widehat{\mathbf{C}}$. En multipliant²⁶ l'égalité $Z(a, b)Z(a', b') = 0$ par $Z(a, b)^{-1}$, on obtient $Z(a', b') = 0$.

En résumé : $[Z(a, b)Z(a', b') = 0] \implies [Z(a, b) = 0 \vee Z(a', b') = 0]$. Ainsi : l'anneau $(\widehat{\mathbf{C}}, +, \times)$ est intègre.

Conclusion. Puisque $(\widehat{\mathbf{C}}, +, \times)$ est un anneau commutatif où tout élément non nul est inversible, on peut affirmer que $(\widehat{\mathbf{C}}, +, \times)$ est un corps.

f) Soit $Z(a, b)$ un élément de $\widehat{\mathbf{C}}$ distinct de I , et soit n un entier naturel. Puisque $\widehat{\mathbf{C}}$ est un anneau commutatif, on a :

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

26. A gauche ou à droite, cela n'a pas d'importance puisque le produit est commutatif dans $\widehat{\mathbf{C}}$.

Notons alors que $(I - Z(a, b))$ est un élément de $\widehat{\mathbf{C}}$, non nul par hypothèse. D'après la question précédente, il est donc inversible dans $\widehat{\mathbf{C}}$, et on peut donc conclure :

$$\forall Z(a, b) \in \widehat{\mathbf{C}}, \forall n \in \mathbb{N}, \sum_{k=0}^n Z(a, b)^k = (I - Z(a, b)^{n+1}) \times (I - Z(a, b))^{-1}.$$

► **PARTIE C - Notation exponentielle**

10) Pour tout réel θ : $|E^{\theta J}| = \cos^2 \theta + \sin^2 \theta = 1$.

11) Soient θ et φ deux réels. On a :

$$E^{\theta J} \times E^{\varphi J} = Z(\cos \theta, \sin \theta) Z(\cos \varphi, \sin \varphi) = Z(\cos \theta \cos \varphi - \sin \varphi \sin \theta, \sin \theta \cos \varphi + \sin \varphi \cos \theta)$$

$$= Z(\cos(\theta + \varphi), \sin(\theta + \varphi)).$$

Conclusion. $\forall (\theta, \varphi) \in \mathbb{R}^2, E^{\theta J} \times E^{\varphi J} = E^{(\theta + \varphi)J}$.

12) Soit θ un réel. On a : $\det(E^{\theta J}) = 1$, donc : $\forall \theta \in \mathbb{R}, (E^{\theta J})^{-1} = E^{(-\theta)J}$

13) Par définition des “matrices exponentielles”, on a : $E^{\theta J} = E^{\varphi J}$ si et seulement si $\cos \theta = \cos \varphi$ et $\sin \theta = \sin \varphi$. Ces deux conditions sont simultanément satisfaites si et seulement si $\theta = \varphi [2\pi]$.

Conclusion. $\forall (\theta, \varphi) \in \mathbb{R}^2, [E^{\theta J} = E^{\varphi J}] \iff [\theta = \varphi [2\pi]]$.

► **PARTIE D - Forme exponentielle d'un élément non nul de $\widehat{\mathbf{C}}$**

14) Clairement : $\forall Z(a, b) \in \widehat{\mathbf{C}}^*, Z(a, b) = |Z(a, b)| \times Z\left(\underbrace{\frac{a}{|Z(a, b)|}}_{=\alpha}, \underbrace{\frac{b}{|Z(a, b)|}}_{=\beta}\right)$.

15) Par définition des réels α et β , on a : $\alpha^2 + \beta^2 = 1$

16) Puisque α et β vérifient $\alpha^2 + \beta^2 = 1$, le couple (α, β) est celui des coordonnées d'un point du cercle trigonométrique. Il existe donc un réel θ défini modulo 2π tel que : $\alpha = \cos \theta$ et $\beta = \sin \theta$. D'où : $Z(a, b) = |Z(a, b)| \times Z(\cos \theta, \sin \theta)$ donc : $Z(a, b) = |Z(a, b)| \times E^{\theta J}$.

Conclusion. $\forall Z(a, b) \in \widehat{\mathbf{C}}^*, \exists \theta \in \mathbb{R}, Z(a, b) = |Z(a, b)| \times E^{\theta J}$.

17) **Méthode d'identification dans $\widehat{\mathbf{C}}^*$.** Soient $Z(a, b)$ et $Z(a', b')$ dans $\widehat{\mathbf{C}}^*$.

On suppose que : $Z(a, b) = Z(a', b')$. Il est immédiat que $|Z(a, b)| = |Z(a', b')|$. Il en résulte, grâce à la question précédente que : $E^{\text{ARG}(Z(a, b))J} = E^{\text{ARG}(Z(a', b'))J}$.

D'après la question 13, ceci implique que : $\text{ARG}(Z(a, b)) = \text{ARG}(Z(a', b')) [2\pi]$.

En résumé, on a établi l'implication : $(Z(a, b) = Z(a', b')) \implies \begin{cases} |Z(a, b)| = |Z(a', b')| \\ \text{ARG}(Z(a, b)) = \text{ARG}(Z(a', b')) [2\pi] \end{cases}$

La réciproque est triviale.

Conclusion. $(Z(a, b) = Z(a', b')) \iff \begin{cases} |Z(a, b)| = |Z(a', b')| \\ \text{ARG}(Z(a, b)) = \text{ARG}(Z(a', b')) [2\pi] \end{cases}$

18) Provient de la définition d'argument et de la question 11.

19) Provient de 18 et d'une récurrence immédiate sur n .

► PARTIE E - Racines carrées dans $\widehat{\mathbf{C}}$

20) Dans $\widehat{\mathbf{C}}$: $X^2 = I \iff X^2 - I = 0 \iff (X - I)(X + I) = 0$.

Puisque $\widehat{\mathbf{C}}$ est un ANNEAU INTEGRE, on peut alors affirmer que : $X^2 = I \iff X = \pm I$.

Conclusion. Dans $\widehat{\mathbf{C}}$, l'équation $X^2 = I$ possède exactement deux solutions : $\pm I$.

21) Dans $\widehat{\mathbf{C}}$: $X^2 = -I \iff X^2 + I = 0 \iff X^2 - J^2 = 0 \iff (X - J)(X + J) = 0$.

Puisque $\widehat{\mathbf{C}}$ est un ANNEAU INTEGRE, on peut alors affirmer que : $X^2 = -I \iff X = \pm J$.

Conclusion. Dans $\widehat{\mathbf{C}}$, l'équation $X^2 = -I$ possède exactement deux solutions : $\pm J$.

22) Dans $\widehat{\mathbf{C}}$: $X^4 = I \iff X^4 - I = 0 \iff (X^2 - I)(X^2 + I) = 0$.

Puisque $\widehat{\mathbf{C}}$ est un ANNEAU INTEGRE, on peut alors affirmer que : $X^4 = I \iff X^2 = \pm I$.

Conclusion. Dans $\widehat{\mathbf{C}}$, l'équation $X^4 = I$ possède exactement quatre solutions : $\pm I$ et $\pm J$.

23) Soit $Z(a, b) \in \widehat{\mathbf{C}}^*$. D'après la question 16, on peut écrire : $Z(a, b) = |Z(a, b)| \times e^{i\theta}$ où $\theta = \text{ARG}(Z(a, b))$.

Soit $X \in \widehat{\mathbf{C}}$ tel que $X^2 = Z(a, b)$. Nécessairement $X \neq 0$, puisque $Z(a, b)$ est supposé non nul. On peut donc écrire : $X = |X| \times e^{i\varphi}$ où $\varphi = \text{ARG}(X)$.

L'équation $X^2 = Z(a, b)$ est alors équivalente à : $|X|^2 \times e^{2i\varphi} = |Z(a, b)| \times e^{i\theta}$. D'après la méthode d'identification, on en déduit que :

$$\begin{aligned}
 [X^2 = Z(a, b)] &\iff \begin{cases} |X|^2 = |Z(a, b)| \\ 2\varphi = \theta \pmod{2\pi} \end{cases} \iff \begin{cases} |X| = |Z(a, b)|^{1/2} \\ \varphi = \theta/2 \pmod{\pi} \end{cases} \\
 &\iff \left[X = |Z(a, b)|^{1/2} e^{i\frac{\theta}{2}} \vee X = |Z(a, b)|^{1/2} e^{i(\frac{\theta}{2} + \pi)} \right] \iff X = \pm |Z(a, b)|^{1/2} e^{i\frac{\theta}{2}}
 \end{aligned}$$

Conclusion. Pour tout $Z(a, b) \in \widehat{\mathbf{C}}^*$, l'équation $X^2 = Z(a, b)$ possède exactement deux solutions dans $\widehat{\mathbf{C}}$, qui sont : $\pm |Z(a, b)|^{1/2} e^{i\frac{\theta}{2}}$ (avec $\theta = \text{ARG}(Z(a, b))$).

Il revient au même de dire que tout élément non nul de $\widehat{\mathbf{C}}$ possède exactement deux racines carrées dans $\widehat{\mathbf{C}}$.

24) **Application.** D'après la question précédente, la matrice $A = \begin{pmatrix} 1 & -\sqrt{3} \\ \sqrt{3} & 1 \end{pmatrix}$, qui est un élément de $\widehat{\mathbf{C}}^*$, possède exactement deux racines carrées dans $\widehat{\mathbf{C}}$ qui sont $\pm |A|^{1/2} e^{i\frac{\text{ARG}(A)}{2}}$.

Deux calculs directs permettent d'obtenir : $|A| = 2$ et $\text{ARG}(A) = \frac{\pi}{3}$.

Conclusion. L'équation $X^2 = \begin{pmatrix} 1 & -\sqrt{3} \\ \sqrt{3} & 1 \end{pmatrix}$ admet exactement deux solutions dans $\widehat{\mathbf{C}}$ qui sont $\pm \sqrt{2} e^{i\frac{\pi}{6}}$, c'est à dire explicitement : $\frac{1}{4} \begin{pmatrix} \sqrt{6} & -\sqrt{2} \\ \sqrt{2} & \sqrt{6} \end{pmatrix}$ et $\frac{1}{4} \begin{pmatrix} -\sqrt{6} & \sqrt{2} \\ -\sqrt{2} & -\sqrt{6} \end{pmatrix}$

► PARTIE F - Equations du second degré dans $\widehat{\mathbf{C}}$

25) Puisque A est un élément non nul de $\widehat{\mathbf{C}}$, et que $\widehat{\mathbf{C}}$ est un corps, A est inversible dans $\widehat{\mathbf{C}}$. On peut donc multiplier²⁷ tous les termes de (E) par A^{-1} pour débiter la suite d'équivalences suivante :

$$\begin{aligned}
 & AX^2 + BX + C = 0_{\mathbf{M}_2(\mathbb{R})} \\
 \iff & A^{-1}(AX^2 + BX + C) = 0_{\mathbf{M}_2(\mathbb{R})} \\
 \iff & X^2 + BA^{-1}X + CA^{-1} = 0_{\mathbf{M}_2(\mathbb{R})} \\
 \iff & \left(X + \frac{1}{2}BA^{-1}\right)^2 - \frac{1}{4}B^2A^{-2} + CA^{-1} = 0_{\mathbf{M}_2(\mathbb{R})} \\
 \iff & \left(X + \frac{1}{2}BA^{-1}\right)^2 - \frac{1}{4}B^2A^{-2} + \frac{1}{4}4ACA^{-2} = 0_{\mathbf{M}_2(\mathbb{R})} \\
 \iff & \left(X + \frac{1}{2}BA^{-1}\right)^2 + \frac{1}{4}(4AC - B^2)A^{-2} = 0_{\mathbf{M}_2(\mathbb{R})} \\
 \iff & \left(X + \frac{1}{2}BA^{-1}\right)^2 - \frac{1}{4}\Delta A^{-2} = 0_{\mathbf{M}_2(\mathbb{R})}
 \end{aligned}$$

► Si $\Delta = 0$: l'équation (E) équivaut alors à $\left(X + \frac{1}{2}BA^{-1}\right)^2 = 0_{\mathbf{M}_2(\mathbb{R})}$ d'où : $X = -\frac{1}{2}BA^{-1}$.

► Si $\Delta \neq 0$: d'après la question précédente, Δ possède alors deux racines carrées dans $\widehat{\mathbf{C}}$. Notons δ l'une d'entre elles, de telle sorte que $\delta \in \widehat{\mathbf{C}}$ et $\delta^2 = \Delta$. L'équation (E) équivaut alors à :

$$\begin{aligned}
 & \left(X + \frac{1}{2}BA^{-1}\right)^2 - \frac{1}{4}\delta^2 A^{-2} = 0_{\mathbf{M}_2(\mathbb{R})} \\
 \iff & \left(X + \frac{1}{2}BA^{-1}\right)^2 - \left(\frac{1}{2}\delta A^{-1}\right)^2 = 0_{\mathbf{M}_2(\mathbb{R})} \\
 \iff & \left(\left(X + \frac{1}{2}BA^{-1}\right) - \left(\frac{1}{2}\delta A^{-1}\right)\right) \times \left(\left(X + \frac{1}{2}BA^{-1}\right) + \left(\frac{1}{2}\delta A^{-1}\right)\right) = 0_{\mathbf{M}_2(\mathbb{R})}
 \end{aligned}$$

En utilisant une nouvelle fois le fait que $\widehat{\mathbf{C}}$ est intègre, on peut continuer notre raisonnement par équivalences :

$$\begin{aligned}
 & \left[\left(X + \frac{1}{2}BA^{-1}\right) - \left(\frac{1}{2}\delta A^{-1}\right) = 0_{\mathbf{M}_2(\mathbb{R})}\right] \vee \left[\left(X + \frac{1}{2}BA^{-1}\right) + \left(\frac{1}{2}\delta A^{-1}\right) = 0_{\mathbf{M}_2(\mathbb{R})}\right] \\
 \iff & \left[X + \frac{1}{2}BA^{-1} = \frac{1}{2}\delta A^{-1}\right] \vee \left[X + \frac{1}{2}BA^{-1} = -\frac{1}{2}\delta A^{-1}\right] \\
 \iff & \left[X = \frac{1}{2}(-B + \delta)A^{-1}\right] \vee \left[X = \frac{1}{2}(-B - \delta)A^{-1}\right]
 \end{aligned}$$

27. A gauche ou à droite, puisqu'en outre l'anneau $\widehat{\mathbf{C}}$ est commutatif.

Dans le cas où $\Delta \neq 0$, l'équation (E) possède exactement deux solutions dans $\widehat{\mathbf{C}}$ qui sont : $\frac{1}{2}(-B \pm \delta)A^{-1}$.

Conclusion. Soient A , B et C dans $\widehat{\mathbf{C}}$, avec $A \neq 0_{M_2(\mathbb{R})}$. On note (E) l'équation

$$AX^2 + BX + C = 0_{M_2(\mathbb{R})} \quad \text{d'inconnue } X \in \widehat{\mathbf{C}}$$

Enfin on note Δ (discriminant de l'équation (E)) l'élément de $\widehat{\mathbf{C}}$ défini par : $\Delta = B^2 - 4AC$.

Deux cas peuvent se présenter :

► Si $\Delta = 0_{M_2(\mathbb{R})}$: alors l'équation (E) possède une unique solution dans $\widehat{\mathbf{C}}$, qui est : $-\frac{1}{2}BA^{-1}$.

► Si $\Delta \neq 0_{M_2(\mathbb{R})}$: alors l'équation (E) possède exactement deux solutions dans $\widehat{\mathbf{C}}$ qui sont : $\frac{1}{2}(-B \pm \delta)A^{-1}$ (où δ est un élément de $\widehat{\mathbf{C}}$ tel que $\delta^2 = \Delta$).

26) **Application.** On a : $X^2 - 2JX - J = I \iff X^2 - 2JX - J - I = 0_{M_2(\mathbb{R})}$. Cette équation étant du second degré à coefficients dans $\widehat{\mathbf{C}}$, on peut lui appliquer le résultat de la question précédente pour la résoudre dans $\widehat{\mathbf{C}}$.

Son discriminant est : $\Delta = (-2J)^2 - 4(-J - I) = 4J^2 + 4J + 4I$ d'où $\Delta = 4J$ (ainsi $\Delta \neq 0_{M_2(\mathbb{R})}$).

En outre : $4J = 4E^{\frac{\pi}{2}J}$, donc $\delta = 2E^{\frac{\pi}{4}J}$ est une racine carrée de Δ dans $\widehat{\mathbf{C}}$.

D'après la question précédente, l'équation possède donc deux solutions dans $\widehat{\mathbf{C}}$ qui sont : $J \pm E^{\frac{\pi}{4}J}$.

Conclusion. L'équation $X^2 - 2JX - J = I$ possède exactement deux solutions dans $\widehat{\mathbf{C}}$, qui sont $J \pm E^{\frac{\pi}{4}J}$; c'est à dire les matrices

$$\begin{pmatrix} \frac{\sqrt{2}}{2} & -1 - \frac{\sqrt{2}}{2} \\ 1 + \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{pmatrix} \quad \text{et} \quad \begin{pmatrix} -\frac{\sqrt{2}}{2} & -1 + \frac{\sqrt{2}}{2} \\ 1 - \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \end{pmatrix}$$

► PARTIE G - Racines n -ièmes de l'unité dans $\widehat{\mathbf{C}}$

27) Soit $Z \in \widehat{\mathbf{U}}_n$. Alors $Z^n = I$ donc $|Z^n| = |I|$ donc $|Z|^n = 1$ d'où : $|Z| = 1$ et donc $Z \in \widehat{\mathbf{U}}$. En résumé : $[Z \in \widehat{\mathbf{U}}_n] \implies [Z \in \widehat{\mathbf{U}}]$. **Conclusion.** $\widehat{\mathbf{U}}_n \subset \widehat{\mathbf{U}}$.

28) Comme d'habitude :

► (SG1) $\widehat{\mathbf{U}}_n \subset \widehat{\mathbf{U}}$, d'après la question précédente.

► (SG2) $I \in \widehat{\mathbf{U}}_n$ puisque $I^n = I$.

► (SG3) $\forall (Z, Z') \in \widehat{\mathbf{U}}_n^2$, $Z \times Z' \in \widehat{\mathbf{U}}_n$ puisque $(Z \times Z')^n = Z^n \times Z'^n = I$.

► (SG4) $\forall Z \in \widehat{\mathbf{U}}_n$, $Z^{-1} \in \widehat{\mathbf{U}}_n$ puisque $(Z^{-1})^n = (Z^n)^{-1} = I$.

Conclusion : $(\widehat{\mathbf{U}}_n, \times)$ est un sous-groupe de $(\widehat{\mathbf{U}}, \times)$.

29) Notons $E = \left\{ E^{\frac{2k\pi}{n}J} / k \in \llbracket 0, n-1 \rrbracket \right\}$, et prouvons que $E = \widehat{\mathbf{U}}_n$.

Soit $Z \in \widehat{\mathbf{U}}_n$. On a déjà montré que $\widehat{\mathbf{U}}_n \subset \widehat{\mathbf{U}}$, donc $|Z| = 1$. Il existe donc un réel θ tel que : $Z = E^{\theta J}$. Puisque $Z \in \widehat{\mathbf{U}}_n$, il s'ensuit que $E^{n\theta J} = I$, et donc : $n\theta = 0 [2\pi]$, c'est-à-dire : $\theta = 0 [2\pi/n]$.

Par suite il existe un entier relatif k tel que : $\theta = 2k\pi/n$. Donc : $\widehat{\mathbf{U}}_n = \left\{ \mathbf{E}^{\frac{2k\pi}{n} \mathbf{J}} / k \in \mathbb{Z} \right\}$. Il reste à établir que : $\widehat{\mathbf{U}}_n = E$.

Il est déjà immédiat que : $\boxed{E \subset \widehat{\mathbf{U}}_n \text{ (♠)}}$.

Réciproquement, soit Z un élément de $\widehat{\mathbf{U}}_n$. Il existe un entier relatif k tel que $Z = \mathbf{E}^{\frac{2k\pi}{n} \mathbf{J}}$. En effectuant la division euclidienne de k par n , on a : $k = nq + r$, avec $r \in \llbracket 0, n-1 \rrbracket$. D'où :

$$Z = \mathbf{E}^{\frac{2k\pi}{n} \mathbf{J}} = \mathbf{E}^{\frac{2(nq+r)\pi}{n} \mathbf{J}} = \underbrace{\mathbf{E}^{\frac{2nq\pi}{n} \mathbf{J}}}_{=\mathbf{I}} \mathbf{E}^{\frac{2r\pi}{n} \mathbf{J}} = \mathbf{E}^{\frac{2r\pi}{n} \mathbf{J}}.$$

Par suite : $Z = \mathbf{E}^{\frac{2r\pi}{n} \mathbf{J}}$ avec $r \in \llbracket 0, n-1 \rrbracket$, ce qui prouve que $Z \in E$, et l'inclusion : $\boxed{\widehat{\mathbf{U}}_n \subset E \text{ (♣)}}$.

Conclusion : d'après la règle de double inclusion on a $\widehat{\mathbf{U}}_n = \left\{ \mathbf{E}^{\frac{2k\pi}{n} \mathbf{J}} / k \in \llbracket 0, n-1 \rrbracket \right\}$.

30) D'après la question précédente : $\sum_{W \in \widehat{\mathbf{U}}_n} W = \sum_{k=0}^{n-1} \mathbf{E}^{\frac{2k\pi}{n} \mathbf{J}} = \sum_{k=0}^{n-1} \left(\mathbf{E}^{\frac{2\pi}{n} \mathbf{J}} \right)^k$.

D'après la question 9-f on a : $\sum_{k=0}^{n-1} \left(\mathbf{E}^{\frac{2\pi}{n} \mathbf{J}} \right)^k = \underbrace{\left(\mathbf{I} - \left(\mathbf{E}^{\frac{2\pi}{n} \mathbf{J}} \right)^n \right)}_{=0_{M_2(\mathbb{R})}} \left(\mathbf{I} - \mathbf{E}^{\frac{2\pi}{n} \mathbf{J}} \right)^{-1} = 0_{M_2(\mathbb{R})}.$

Conclusion : $\sum_{W \in \widehat{\mathbf{U}}_n} W = 0_{M_2(\mathbb{R})}.$

31) **Application.** Déterminons toutes les matrices $X \in \widehat{\mathbf{C}}$ telles que

$$X^4 = \begin{pmatrix} 8\sqrt{2} & -8\sqrt{2} \\ 8\sqrt{2} & 8\sqrt{2} \end{pmatrix}$$

Notons A la matrice de droite de cette égalité. On a : $X^4 = A \iff \begin{cases} |X| = |A|^{1/4} \\ \text{ARG}(X) = \frac{\text{ARG}(A)}{4} [\pi/2] \end{cases}.$

Or : $|A| = \sqrt{(8\sqrt{2})^2 + (8\sqrt{2})^2} = \sqrt{256} = 16$. Il s'ensuit que : $|X| = 2$.

Par ailleurs : $\text{ARG}(A) = \frac{\pi}{4} [2\pi]$. D'où : $\text{ARG}(X) = \frac{\pi}{16} [\pi/2]$.

Conclusion : l'équation $X^4 = \begin{pmatrix} 8\sqrt{2} & -8\sqrt{2} \\ 8\sqrt{2} & 8\sqrt{2} \end{pmatrix}$ possède exactement quatre solutions dans $\widehat{\mathbf{C}}$:

$$\pm 2\mathbf{E}^{\frac{\pi}{16} \mathbf{J}} \text{ et } \pm 2\mathbf{J}\mathbf{E}^{\frac{\pi}{16} \mathbf{J}}$$

Comme il est d'ailleurs bien connu que : $\cos\left(\frac{\pi}{16}\right) = \frac{\sqrt{2 + \sqrt{2 + \sqrt{2}}}}{2}$ et $\sin\left(\frac{\pi}{16}\right) = \frac{\sqrt{2 - \sqrt{2 + \sqrt{2}}}}{2}$, on peut également conclure explicitement :

Conclusion : l'équation $X^4 = \begin{pmatrix} 8\sqrt{2} & -8\sqrt{2} \\ 8\sqrt{2} & 8\sqrt{2} \end{pmatrix}$ possède exactement quatre solutions dans $\widehat{\mathbf{C}}$:

$$\begin{pmatrix} \sqrt{2 + \sqrt{2 + \sqrt{2}}} & -\sqrt{2 - \sqrt{2 + \sqrt{2}}} \\ \sqrt{2 - \sqrt{2 + \sqrt{2}}} & \sqrt{2 + \sqrt{2 + \sqrt{2}}} \end{pmatrix}; \quad \begin{pmatrix} -\sqrt{2 + \sqrt{2 + \sqrt{2}}} & \sqrt{2 - \sqrt{2 + \sqrt{2}}} \\ -\sqrt{2 - \sqrt{2 + \sqrt{2}}} & -\sqrt{2 + \sqrt{2 + \sqrt{2}}} \end{pmatrix}; \\ \begin{pmatrix} -\sqrt{2 - \sqrt{2 + \sqrt{2}}} & -\sqrt{2 + \sqrt{2 + \sqrt{2}}} \\ \sqrt{2 + \sqrt{2 + \sqrt{2}}} & -\sqrt{2 - \sqrt{2 + \sqrt{2}}} \end{pmatrix}; \quad \begin{pmatrix} \sqrt{2 - \sqrt{2 + \sqrt{2}}} & \sqrt{2 + \sqrt{2 + \sqrt{2}}} \\ -\sqrt{2 + \sqrt{2 + \sqrt{2}}} & \sqrt{2 - \sqrt{2 + \sqrt{2}}} \end{pmatrix}$$

32) Une similitude directe ayant un centre Ω d'affixe ω admet une écriture complexe de la forme : $z \mapsto a(z - \omega) + \omega$ (avec $a \in \mathbb{C}^*$). Lorsque ce centre est le point O , on a $\omega = 0$.

Il s'ensuit qu'un élément de $\text{Sim}_0^+(\mathbb{C})$ est une transformation f_a ayant pour écriture complexe : $z \mapsto az$ ($a \in \mathbb{C}^*$).

En outre $(\text{Sim}_0^+(\mathbb{C}), \circ)$ est un groupe abélien puisque c'est un sous-groupe du groupe des similitudes directes, et que deux similitudes de même centre commutent.

33) Il est immédiat que l'application $\rho : \widehat{\mathbf{C}}^* \rightarrow \text{Sim}_0^+(\mathbb{C})$ qui à tout $Z = |Z|e^{i\theta} \in \widehat{\mathbf{C}}^*$ associe la similitude d'écriture complexe $z \mapsto |Z|e^{i\theta}z$ est telle que $\psi \circ \rho = \text{id}_{\widehat{\mathbf{C}}^*}$ et $\rho \circ \psi = \text{id}_{\text{Sim}_0^+(\mathbb{C})}$.

Il s'ensuit que ψ est bijective (et $\psi^{-1} = \rho$).

34) Il suffit de voir que la composée des similitudes d'écritures complexes $z \mapsto az$ et $z \mapsto a'z$ est la similitude d'écriture complexe $z \mapsto (aa')z$ pour conclure que : $\forall (f, g) \in \text{Sim}_0^+(\mathbb{C})^2, \psi(f \circ g) = \psi(f) \times \psi(g)$.

35) **BONUS.** Déterminer toutes les matrices $M \in \widehat{\mathbf{C}}$ telles que $(M + I)^5 = (M - I)^5$.

Observons déjà que I n'est pas solution de l'équation. Ainsi $M \neq I$, et $M - I$ est donc inversible, en tant qu'élément non nul de $\widehat{\mathbf{C}}$. On a donc $M \neq I$ et :

$$[(M + I)^5 = (M - I)^5] \iff [((M + I)(M - I)^{-1})^5 = I] \iff [(M + I)(M - I)^{-1} \in \widehat{\mathbf{U}}_5]$$

Or : $(M + I)(M - I)^{-1} \in \widehat{\mathbf{U}}_5$ et $M \neq I$

$$\iff \exists k \in \llbracket 1, 4 \rrbracket, (M + I)(M - I)^{-1} = e^{\frac{2k\pi}{5}J}$$

$$\iff \exists k \in \llbracket 1, 4 \rrbracket, M + I = e^{\frac{2k\pi}{5}J}(M - I)$$

$$\iff \exists k \in \llbracket 1, 4 \rrbracket, \left(I - e^{\frac{2k\pi}{5}J}\right)M = -I - e^{\frac{2k\pi}{5}J}$$

$$\iff \exists k \in \llbracket 1, 4 \rrbracket, \left(e^{\frac{2k\pi}{5}J} - I\right)M = e^{\frac{2k\pi}{5}J} + I$$

$$\iff \exists k \in \llbracket 1, 4 \rrbracket, e^{\frac{k\pi}{5}J} \left(e^{\frac{k\pi}{5}J} - e^{-\frac{k\pi}{5}J}\right)M = e^{\frac{k\pi}{5}J} \left(e^{\frac{k\pi}{5}J} + e^{-\frac{k\pi}{5}J}\right)$$

$$\Longleftrightarrow \exists k \in \llbracket 1, 4 \rrbracket, \left(E^{\frac{k\pi}{5} J} - E^{-\frac{k\pi}{5} J} \right) M = E^{\frac{k\pi}{5} J} + E^{-\frac{k\pi}{5} J}$$

$$\Longleftrightarrow \exists k \in \llbracket 1, 4 \rrbracket, \left(E^{\frac{k\pi}{5} J} - \overline{E^{\frac{k\pi}{5} J}} \right) M = E^{\frac{k\pi}{5} J} + \overline{E^{\frac{k\pi}{5} J}}$$

$$\Longleftrightarrow \exists k \in \llbracket 1, 4 \rrbracket, 2 \sin \left(\frac{k\pi}{5} \right) JM = 2 \cos \left(\frac{k\pi}{5} \right)$$

$$\Longleftrightarrow \exists k \in \llbracket 1, 4 \rrbracket, M = -\cotan \left(\frac{k\pi}{5} \right) J$$

Conclusion : l'équation $(M + I)^5 = (M - I)^5$ possède exactement quatre solutions dans $\widehat{\mathbf{C}}$:

$$-\cotan \left(\frac{\pi}{5} \right) J; \quad -\cotan \left(\frac{2\pi}{5} \right) J; \quad -\cotan \left(\frac{3\pi}{5} \right) J \quad \text{et} \quad -\cotan \left(\frac{4\pi}{5} \right) J$$

Chapitre 14

Limites et continuité

14.1 Limites des fonctions à valeurs réelles

Dans ce paragraphe, on ne considère que des fonctions à valeurs réelles.

14.1.1 Définitions

L'objet de ce paragraphe est de donner les définitions de " $\lim_{x \rightarrow a} f(x) = \ell$ " avec a et ℓ réels ou $\pm\infty$ (càd a et ℓ dans $\overline{\mathbb{R}}$).

DÉFINITION 14.1 - Soit a un réel.

- 1) On appelle **voisinage** (ouvert) de a un intervalle ouvert contenant a .
- 2) On appelle **voisinage** (ouvert) de $+\infty$ un intervalle ouvert de la forme $]a, +\infty[$.
- 3) On appelle **voisinage** (ouvert) de $-\infty$ un intervalle ouvert de la forme $] - \infty, a[$.

Exemples. $\mathbb{R}$ est un voisinage de chacun de ses points, de $+\infty$ et de $-\infty$. Pour tout réel a et pour tout réel strictement positif ε , l'intervalle $]a - \varepsilon, a + \varepsilon[$ est un voisinage de a .

DÉFINITION 14.2 - Soient a et ℓ deux réels, et soit f une fonction définie sur un voisinage I de a (et à valeurs réelles, en vertu de la remarque préliminaire).

La fonction f a pour limite ℓ en a si :

$$\forall \varepsilon > 0, \exists \alpha > 0, \forall x \in I, [|x - a| < \alpha] \implies [|f(x) - \ell| < \varepsilon]$$

Interprétation. Dire que f a pour limite ℓ en a , c'est dire que l'on peut enfermer les valeurs de f dans un rectangle de largeur arbitrairement petite (mais non nulle!) centré en ℓ , pour peu que l'on soit assez proche de a .

Notation. $\lim_{x \rightarrow a} f(x) = \ell$

Exemple. La fonction $x \mapsto |x|$ a pour limite 0 en 0.

DÉFINITION 14.3 - Soit a un réel, et soit f une fonction définie sur $I \setminus \{a\}$ (où I est un voisinage de a).

La fonction f a pour limite $+\infty$ en a si :

$$\forall M \in \mathbb{R}, \exists \alpha > 0, \forall x \in I \setminus \{a\}, [|x - a| < \alpha] \implies [f(x) > M]$$

Interprétation. Dire que f a pour limite $+\infty$ en a , c'est dire que les valeurs de f sont situées au-dessus d'une barre que l'on peut fixer arbitrairement haut, pour peu que l'on soit assez proche de a .

Notation. $\lim_{x \rightarrow a} f(x) = +\infty$

Exemple. La fonction $x \mapsto \frac{1}{x^2}$ a pour limite $+\infty$ en 0.

DÉFINITION 14.4 - Soit a un réel, et soit f une fonction définie sur $I \setminus \{a\}$ (où I est un voisinage de a).

La fonction f a pour limite $-\infty$ en a si :

$$\forall M \in \mathbb{R}, \exists \alpha > 0, \forall x \in I \setminus \{a\}, [|x - a| < \alpha] \implies [f(x) < M]$$

Interprétation. Dire que f a pour limite $-\infty$ en a , c'est dire que les valeurs de f sont situées en-dessous d'une barre que l'on peut fixer arbitrairement bas, pour peu que l'on soit assez proche de a .

Notation. $\lim_{x \rightarrow a} f(x) = -\infty$

Exemple. La fonction $x \mapsto \ln |x|$ a pour limite $-\infty$ en 0.

DÉFINITION 14.5 - Soit ℓ un réel, et soit f une fonction définie sur un voisinage I de $+\infty$.

La fonction f a pour limite ℓ en $+\infty$ si :

$$\forall \varepsilon > 0, \exists x_0 \in \mathbb{R}, \forall x \in I, [x > x_0] \implies [|f(x) - \ell| < \varepsilon]$$

Interprétation. Dire que f a pour limite ℓ en $+\infty$, c'est dire que l'on peut enfermer les valeurs de f dans un rectangle de largeur arbitrairement petite centré en ℓ , pour peu que l'on prenne des valeurs de x assez grandes.

Notation. $\lim_{x \rightarrow +\infty} f(x) = \ell$

Exemple. La fonction $x \mapsto 1 - e^{-x}$ a pour limite 1 en $+\infty$.

DÉFINITION 14.6 - Soit ℓ un réel, et soit f une fonction définie sur un voisinage I de $-\infty$.

La fonction f a pour limite ℓ en $-\infty$ si :

$$\forall \varepsilon > 0, \exists x_0 \in \mathbb{R}, \forall x \in I, [x < x_0] \implies [|f(x) - \ell| < \varepsilon]$$

Interprétation. Dire que f a pour limite ℓ en $-\infty$, c'est dire que l'on peut enfermer les valeurs de f dans un rectangle de largeur arbitrairement petite centré en ℓ , pour peu que l'on prenne des valeurs de x assez petites.

Notation. $\lim_{x \rightarrow -\infty} f(x) = \ell$

Exemple. La fonction $x \mapsto e^x$ a pour limite 0 en $-\infty$.

DÉFINITION 14.7 - Soit f une fonction définie sur un voisinage I de $+\infty$.

La fonction f a pour limite $+\infty$ en $+\infty$ si :

$$\forall M > 0, \exists x_0 \in \mathbb{R}, \forall x \in I, [x > x_0] \implies [f(x) > M]$$

Interprétation. Dire que f a pour limite $+\infty$ en $+\infty$, c'est dire que les valeurs de f sont situées au-dessus d'une barre que l'on peut fixer arbitrairement haut, pour peu que l'on choisisse des valeurs de x suffisamment grandes.

Notation. $\lim_{x \rightarrow +\infty} f(x) = +\infty$

Exemple. La fonction $x \mapsto e^x$ a pour limite $+\infty$ en $+\infty$.

On obtient mutatis mutandis les définitions de $\lim_{x \rightarrow +\infty} f(x) = -\infty$, $\lim_{x \rightarrow -\infty} f(x) = +\infty$ et $\lim_{x \rightarrow -\infty} f(x) = -\infty$.

Remarque — limites à gauche et limites à droite. Les définitions des limites de f en un réel a peuvent être “latéralisées”, si l'on ne s'intéresse qu'au comportement de la fonction à gauche ou à droite de a . Par exemple, on pourra dire que :¹

DÉFINITION 14.8 - La fonction f a pour limite à droite (*resp.* à gauche) ℓ en a si :

$$\forall \varepsilon > 0, \exists \alpha > 0, \forall x \in I, [x < a + \alpha] \implies [|f(x) - \ell| < \varepsilon]$$

$$(\text{resp. } \forall \varepsilon > 0, \exists \alpha > 0, \forall x \in I, [x > a - \alpha] \implies [|f(x) - \ell| < \varepsilon])$$

Notation. $\lim_{x \rightarrow a^+} f(x) = \ell$ (limite à droite) et $\lim_{x \rightarrow a^-} f(x) = \ell$ (limite à gauche).

Exemple. La fonction $x \mapsto \lfloor x \rfloor$ a pour limite 1 à droite de 1, et a pour limite 0 à gauche de 1.

Remarques. Un raisonnement sans difficulté permet d'affirmer que f admet pour limite ℓ en a SSI f a pour limite ℓ à gauche et à droite de a .

Par ailleurs, on peut également adapter la définition à celle de fonction f ayant pour limite $+\infty$ (ou $-\infty$) à gauche ou à droite de a (par exemple : la fonction inverse a pour limite $+\infty$ à droite de 0, et $-\infty$ à gauche de 0).

14.1.2 Propriétés des limites

➤ Les **propriétés algébriques** (faisant intervenir la somme, le produit, *etc.*) sur les limites de fonction sont analogues à celles vues sur les suites. De même, la plupart des énoncés du chapitre sur les suites sont encore valides dans ce chapitre (par exemple : unicité de la limite, théorème de comparaison, théorème des gendarmes). Leurs démonstrations sont des adaptations des preuves vues alors (et c'est un bon petit exercice de les retrouver).

1. Pour une fonction f définie sur un voisinage à droite I de a , c-à-d un intervalle I de la forme $[a, b[$ avec $b > a$.

PROPRIÉTÉ 14.1 - Si f est définie au voisinage d'un réel a , et $\lim_{x \rightarrow a} f(x) = \ell$ (avec $\ell \in \mathbb{R}$), alors f est bornée au voisinage de a .

PREUVE. La preuve est analogue à l'énoncé correspondant pour les suites.² Sous les hypothèses de l'énoncé, on peut en effet affirmer que :

$$\exists \alpha > 0, [|x - a| < \alpha] \implies [|f(x) - \ell| < 1]$$

En particulier f est bornée (entre $\ell - 1$ et $\ell + 1$) sur un voisinage de a (l'intervalle $]a - \alpha, a + \alpha[$). ☕

Remarque. L'énoncé précédent tient encore si l'on choisit $a = +\infty$ (ou $a = -\infty$). Il suffit dans la preuve ci-dessus de remplacer l'assertion " $\exists \alpha > 0, |x - a| < \alpha \dots$ " par " $\exists x_0, x > x_0 \dots$ " (ou $\dots$).

On peut donc compléter la propriété précédente :

PROPRIÉTÉ 14.2 - Si f est définie au voisinage de $a \in \overline{\mathbb{R}}$, et $\lim_{x \rightarrow a} f(x) = \ell$ (avec $\ell \in \mathbb{R}$), alors f est bornée au voisinage de a .

Sur le même thème :

PROPRIÉTÉ 14.3 - Si f est définie au voisinage de $a \in \overline{\mathbb{R}}$, et $\lim_{x \rightarrow a} f(x) = \ell$ avec $\ell > 0$ (resp. $\ell < 0$), alors f est strictement positive (resp. négative) au voisinage de a .

PREUVE. Supposons que $\lim_{x \rightarrow a} f(x) = \ell$, avec $\ell > 0$.

► **Premier cas** — a est réel. Alors³ : $\exists \alpha > 0, [|x - a| < \alpha] \implies \left[f(x) > \frac{\ell}{2}\right]$. En particulier, f est à valeurs strictement positives sur $]a - \alpha, a + \alpha[$ (qui est un voisinage de a).

► **Second cas** — $a = +\infty$. Alors : $\exists x_0 \in \mathbb{R}, [x > x_0] \implies \left[f(x) > \frac{\ell}{2}\right]$. En particulier, f est à valeurs strictement positives sur $]x_0, +\infty[$ (qui est un voisinage de $+\infty$).

Les autres cas (avec $a = -\infty$, et $\ell < 0$) se traitent de manière analogue. ☕

Pour en finir avec ces propriétés génériques sur les limites, indiquons aussi que l'on a toujours l'équivalence (dont le pendant a été vu pour les suites) :

$$\lim_{x \rightarrow a} f(x) = \ell \iff \lim_{x \rightarrow a} |f(x) - \ell| = 0$$

où a est dans $\overline{\mathbb{R}}$ et ℓ est un réel.

PROPRIÉTÉ 14.4 - (**Limite séquentielle**). Soit f une fonction définie au voisinage de a , avec $a \in \overline{\mathbb{R}}$, et soit $(x_n)_n$ une suite réelle convergeant vers a .

On suppose que $\lim_{x \rightarrow a} f(x) = \ell$ (avec $\ell \in \overline{\mathbb{R}}$).

Alors :

$$\lim_{n \rightarrow +\infty} f(x_n) = \ell$$

2. Explicitement : toute suite convergente est bornée.

3. Prendre $\varepsilon = \frac{\ell}{2}$ dans la définition de limite.

PREUVE. Juste pour rire, on peut distinguer 9 cas (suivant que avec a et ℓ sont réels, ou égaux à $\pm\infty$).

➤ **Cas 1 :** $\lim_{n \rightarrow +\infty} x_n = a$ et $\lim_{x \rightarrow a} f(x) = \ell$ (avec a et ℓ réels).

Fixons $\varepsilon > 0$.

Puisque : $\lim_{x \rightarrow a} f(x) = \ell$, on peut affirmer que : $\exists \alpha > 0, |x - a| < \alpha \implies |f(x) - \ell| < \varepsilon$ (♠).

Comme par ailleurs : $\lim_{n \rightarrow +\infty} x_n = a$, on peut affirmer que : $\exists n_0 \in \mathbb{N}, n \geq n_0 \implies |x_n - a| < \alpha$ (♣).

D'après (♠) et (♣), et puisque ε est un réel strictement positif arbitraire dans ce qui précède, on a établi que :

$$\forall \varepsilon > 0, \exists n_0 \in \mathbb{N}, n \geq n_0 \implies |f(x_n) - \ell| < \varepsilon, \text{ c'est-à-dire : } \boxed{\lim_{n \rightarrow +\infty} f(x_n) = \ell}$$

➤ **Cas 2 :** $\lim_{n \rightarrow +\infty} x_n = +\infty$ et $\lim_{x \rightarrow +\infty} f(x) = \ell$ (avec ℓ réel).

Fixons $\varepsilon > 0$.

Puisque : $\lim_{x \rightarrow +\infty} f(x) = \ell$, on peut affirmer que : $\exists x_0 \in \mathbb{R}, x \geq x_0 \implies |f(x) - \ell| < \varepsilon$ (♠).

Comme par ailleurs : $\lim_{n \rightarrow +\infty} x_n = +\infty$, on peut affirmer que : $\exists n_0 \in \mathbb{N}, n \geq n_0 \implies x_n \geq x_0$ (♣).

D'après (♠) et (♣), et puisque ε est un réel strictement positif arbitraire dans ce qui précède, on a établi que :

$$\forall \varepsilon > 0, \exists n_0 \in \mathbb{N}, n \geq n_0 \implies |f(x_n) - \ell| < \varepsilon, \text{ c'est-à-dire : } \boxed{\lim_{n \rightarrow +\infty} f(x_n) = \ell}$$

➤ **Cas 3 :** $\lim_{n \rightarrow +\infty} x_n = -\infty$ et $\lim_{x \rightarrow -\infty} f(x) = \ell$ (avec ℓ réel).

Fixons $\varepsilon > 0$.

Puisque : $\lim_{x \rightarrow -\infty} f(x) = \ell$, on peut affirmer que : $\exists x_0 \in \mathbb{R}, x \leq x_0 \implies |f(x) - \ell| < \varepsilon$ (♠).

Comme par ailleurs : $\lim_{n \rightarrow +\infty} x_n = -\infty$, on peut affirmer que : $\exists n_0 \in \mathbb{N}, n \geq n_0 \implies x_n \leq x_0$ (♣).

D'après (♠) et (♣), et puisque ε est un réel strictement positif arbitraire dans ce qui précède, on a établi que :

$$\forall \varepsilon > 0, \exists n_0 \in \mathbb{N}, n \geq n_0 \implies |f(x_n) - \ell| < \varepsilon, \text{ c'est-à-dire : } \boxed{\lim_{n \rightarrow +\infty} f(x_n) = \ell}$$

➤ **Cas 4 :** $\lim_{n \rightarrow +\infty} x_n = a$ et $\lim_{x \rightarrow a} f(x) = +\infty$ (avec a réel).

Fixons $M \in \mathbb{R}$.

Puisque : $\lim_{x \rightarrow a} f(x) = +\infty$, on peut affirmer que : $\exists \alpha > 0, |x - a| < \alpha \implies f(x) \geq M$ (♠).

Comme par ailleurs : $\lim_{n \rightarrow +\infty} x_n = a$, on peut affirmer que : $\exists n_0 \in \mathbb{N}, n \geq n_0 \implies |x_n - a| < \alpha$ (♣).

D'après (♠) et (♣), et puisque M est un réel arbitraire dans ce qui précède, on a établi que :

$$\forall M \in \mathbb{R}, \exists n_0 \in \mathbb{N}, n \geq n_0 \implies f(x_n) \geq M, \text{ c'est-à-dire : } \boxed{\lim_{n \rightarrow +\infty} f(x_n) = +\infty}$$

➤ **Cas 5 :** $\lim_{n \rightarrow +\infty} x_n = +\infty$ et $\lim_{x \rightarrow +\infty} f(x) = +\infty$.

Fixons $M \in \mathbb{R}$.

Puisque : $\lim_{x \rightarrow +\infty} f(x) = +\infty$, on peut affirmer que : $\exists x_0 \in \mathbb{R}, x \geq x_0 \implies f(x) \geq M$ (♠).

Comme par ailleurs : $\lim_{n \rightarrow +\infty} x_n = +\infty$, on peut affirmer que : $\exists n_0 \in \mathbb{N}, n \geq n_0 \implies x_n \geq x_0$ (♣).

D'après (♠) et (♣), et puisque M est un réel arbitraire dans ce qui précède, on a établi que :

$$\forall M \in \mathbb{R}, \exists n_0 \in \mathbb{N}, n \geq n_0 \implies f(x_n) \geq M, \text{ c'est-à-dire : } \boxed{\lim_{n \rightarrow +\infty} f(x_n) = +\infty}$$

➤ **Cas 6 :** $\lim_{n \rightarrow +\infty} x_n = -\infty$ et $\lim_{x \rightarrow -\infty} f(x) = +\infty$.

Fixons $M \in \mathbb{R}$.

Puisque : $\lim_{x \rightarrow -\infty} f(x) = +\infty$, on peut affirmer que : $\exists x_0 \in \mathbb{R}, x \leq x_0 \implies f(x) \geq M$ (♠).

Comme par ailleurs : $\lim_{n \rightarrow +\infty} x_n = -\infty$, on peut affirmer que : $\exists n_0 \in \mathbb{N}, n \geq n_0 \implies x_n \leq x_0$ (♣).

D'après (♠) et (♣), et puisque M est un réel arbitraire dans ce qui précède, on a établi que :

$$\forall M \in \mathbb{R}, \exists n_0 \in \mathbb{N}, n \geq n_0 \implies f(x_n) \geq M, \text{ c'est-à-dire : } \boxed{\lim_{n \rightarrow +\infty} f(x_n) = +\infty}$$

➤ **Cas 7 :** $\lim_{n \rightarrow +\infty} x_n = a$ et $\lim_{x \rightarrow a} f(x) = -\infty$ (avec a réel).

Fixons $M \in \mathbb{R}$.

Puisque : $\lim_{x \rightarrow a} f(x) = -\infty$, on peut affirmer que : $\exists \alpha > 0, |x - a| < \alpha \implies f(x) \leq M$ (♠).

Comme par ailleurs : $\lim_{n \rightarrow +\infty} x_n = a$, on peut affirmer que : $\exists n_0 \in \mathbb{N}, n \geq n_0 \implies |x_n - a| < \alpha$ (♣).

D'après (♠) et (♣), et puisque M est un réel arbitraire dans ce qui précède, on a établi que :

$$\forall M \in \mathbb{R}, \exists n_0 \in \mathbb{N}, n \geq n_0 \implies f(x_n) \leq M, \text{ c'est-à-dire : } \boxed{\lim_{n \rightarrow +\infty} f(x_n) = -\infty}$$

➤ **Cas 8 :** $\lim_{n \rightarrow +\infty} x_n = +\infty$ et $\lim_{x \rightarrow +\infty} f(x) = -\infty$.

Fixons $M \in \mathbb{R}$.

Puisque : $\lim_{x \rightarrow +\infty} f(x) = -\infty$, on peut affirmer que : $\exists x_0 \in \mathbb{R}, x \geq x_0 \implies f(x) \leq M$ (♠).

Comme par ailleurs : $\lim_{n \rightarrow +\infty} x_n = +\infty$, on peut affirmer que : $\exists n_0 \in \mathbb{N}, n \geq n_0 \implies x_n \geq x_0$ (♣).

D'après (♠) et (♣), et puisque M est un réel arbitraire dans ce qui précède, on a établi que :

$$\forall M \in \mathbb{R}, \exists n_0 \in \mathbb{N}, n \geq n_0 \implies f(x_n) \leq M, \text{ c'est-à-dire : } \boxed{\lim_{n \rightarrow +\infty} f(x_n) = -\infty}$$

➤ **Cas 9 :** $\lim_{n \rightarrow +\infty} x_n = -\infty$ et $\lim_{x \rightarrow -\infty} f(x) = -\infty$.

Fixons $M \in \mathbb{R}$.

Puisque : $\lim_{x \rightarrow -\infty} f(x) = -\infty$, on peut affirmer que : $\exists x_0 \in \mathbb{R}, x \leq x_0 \implies f(x) \leq M$ (♠).

Comme par ailleurs : $\lim_{n \rightarrow +\infty} x_n = -\infty$, on peut affirmer que : $\exists n_0 \in \mathbb{N}, n \geq n_0 \implies x_n \leq x_0$ (♣).

D'après (♠) et (♣), et puisque M est un réel arbitraire dans ce qui précède, on a établi que :

$$\forall M \in \mathbb{R}, \exists n_0 \in \mathbb{N}, n \geq n_0 \implies f(x_n) \leq M, \text{ c'est-à-dire : } \boxed{\lim_{n \rightarrow +\infty} f(x_n) = -\infty}$$

Conclusion : dans tous les cas on a montré que : $\lim_{n \rightarrow +\infty} f(x_n) = \ell$ (avec ℓ réel, ou $\ell = \pm\infty$). ☕

Exemple d'application 1. $\lim_{n \rightarrow +\infty} \arctan(n) = \frac{\pi}{2}$ (application on ne peut plus directe!).

Exemple d'application 2. L'exemple précédent peut évidemment être généralisé. Si ℓ désigne un réel et f une fonction définie au voisinage de $+\infty$, alors selon la propriété de limite séquentielle :

$$\left[\lim_{x \rightarrow +\infty} f(x) = \ell \right] \implies \left[\lim_{n \rightarrow +\infty} f(n) = \ell \right]$$

Observation : la réciproque est **FAUSSE** ! Considérer à cet effet la fonction $x \mapsto \sin(2\pi x)$.

Exemple d'application 3. La fonction $x \mapsto \sin(x)$ n'a pas de limite en $+\infty$ (raisonner par l'absurde et considérer les suites de TG $x_n = 2n\pi$ et $y_n = x_n + \pi/2$).

Exemple d'application 4. Si $f : \mathbb{R} \longrightarrow \mathbb{R}$ est une fonction T -périodique, ayant une limite finie en $+\infty$, alors f est constante.

En effet, pour tout réel x , la suite de TG $x_n = x + nT$ tend vers $+\infty$. On a d'une part $f(x_n) = f(x)$ pour tout n , puisque f est T -périodique. D'autre part, $f(x_n)$ tend vers ℓ lorsque n tend vers $+\infty$, d'après la propriété de limite séquentielle. Il s'ensuit que $f(x) = \ell$. Ce raisonnement étant valable pour tout réel x , on en déduit que f est constante égale à ℓ .

14.2 Continuité des fonctions à valeurs réelles

Dans ce paragraphe (et le suivant), les fonctions considérées sont définies sur une partie (souvent un intervalle non-vidé I) de $\mathbb{R}$, et sont à valeurs dans $\mathbb{R}$.

DÉFINITION 14.9 - Soient a un réel et f une fonction définie au voisinage de a . La fonction f est **continue en a** si $\lim_{x \rightarrow a} f(x)$ existe et vaut $f(a)$, c'est-à-dire si :

$$\forall \varepsilon > 0, \exists \alpha > 0, |x - a| < \alpha \implies |f(x) - f(a)| < \varepsilon$$

ou

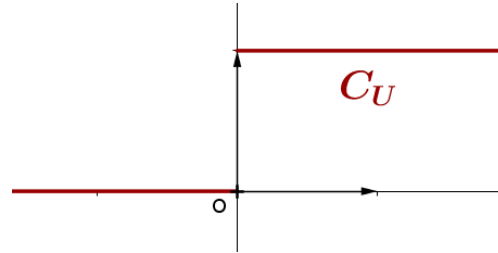
$$\forall \varepsilon > 0, \exists \alpha > 0, x \in]a - \alpha; a + \alpha[\implies f(x) \in]f(a) - \varepsilon; f(a) + \varepsilon[$$

Traduction informelle. “ f est continue en a si $f(x)$ peut être rendu arbitrairement proche de $f(a)$, sous réserve que l’on impose à x d’être suffisamment proche de a ”.

DÉFINITION 14.10 - Soient I un intervalle et $f \in \mathbb{R}^I$. La fonction f est **continue sur I** si f est continue en tout réel a de I .

Exemples. Les fonctions $x \mapsto x^n$ (avec $n \in \mathbb{N}$), polynômes, exponentielle, valeur absolue, cosinus, sinus, ch, sh, th, arctan sont continues sur $\mathbb{R}$. Les autres fonctions usuelles ($\ln$, $\tan$, $\arccos$, $\arcsin$, racine carrée) sont continues sur leurs ensembles de définition respectifs.

En revanche, la fonction U de Heaviside (ou échelon-unité, ou fonction indicatrice de $\mathbb{R}_+$) constante égale à 0 sur $\mathbb{R}_-$ et constante égale à 1 sur $\mathbb{R}_+$ est discontinue en 0.



En effet, si elle admet une limite à gauche en 0 et une limite à droite en 0, ses limites sont distinctes (respectivement égales à 0 et 1) et donc U n’admet pas de limite (tout court) en 0. Donc U n’est pas continue en 0.

De même, la fonction partie entière est discontinue en tout entier relatif. En effet, si $n \in \mathbb{Z}$, cette fonction admet $(n - 1)$ comme limite à gauche de n , et n comme limite à droite de n ; par suite, elle est discontinue en n .

Enfin, le fait précédent (la fonction partie entière admet une limite à droite en tout réel x , égale à $\lfloor x \rfloor$) signifie que la fonction partie entière est **semi-continue supérieurement** (avouez que cela fait son petit effet).

DÉFINITION 14.11 - Soient a un réel et f une fonction définie sur $I \setminus \{a\}$, où I est un intervalle contenant a . On dit que f peut être **prolongée par continuité en a** si f admet en a une limite finie.

Exemples. La fonction $x \in \mathbb{R}_+^* \mapsto \frac{\sin x}{x}$ peut être prolongée par continuité en 0 (limite égale à 1); la fonction $x \in \mathbb{R}_+^* \mapsto \frac{1}{x^2}$ ne peut pas être prolongée par continuité en 0 (limite égale à $+\infty$); la fonction $x \in \mathbb{R}_+^* \mapsto \cos\left(\frac{1}{x}\right)$ ne peut pas être prolongée par continuité en 0 (pas de limite).

14.3 Propriétés des fonctions continues à valeurs réelles

14.3.1 Propriétés générales

THÉORÈME 14.1 - “**Théorème général 1**”. Soient f et g deux fonctions définies au voisinage de a , et continues en a . Alors :

1) $f + g$ est continue en a	3) $1/g$ est continue en a , sous réserve que $g(a) \neq 0$
2) fg est continue en a	4) f/g est continue en a , sous réserve que $g(a) \neq 0$

THÉORÈME 14.2 - “**Théorème général 2**”. Soient $f : I \rightarrow J$ et $g : J \rightarrow \mathbb{R}$ deux fonctions; et a un réel dans I . On suppose que f est continue en a , et que g est continue en $f(a)$. Alors $g \circ f$ est continue en a .

On pourrait rajouter à ces énoncés un troisième affirmant que (presque) toutes les fonctions usuelles sont continues sur leurs ensembles de définition respectifs.

Les preuves de ces théorèmes généraux reposent pour l'essentiel sur les propriétés algébriques des limites, évoquées en début du présent chapitre.

PROPRIÉTÉ 14.5 - Soient $a \in \overline{\mathbb{R}}$, et f une fonction continue en a . Si $f(a) > 0$, alors f est strictement positive sur un voisinage de a .

PREUVE. C'est une conséquence immédiate de la propriété 14.3 page 364 et de la définition de continuité en a ($\lim_{x \rightarrow a} f(x) = f(a)$). ☕

14.3.2 Propriété de continuité séquentielle

PROPRIÉTÉ 14.6 - Soient $a \in \overline{\mathbb{R}}$ et f une fonction continue en a .

Soit $(x_n)_n$ une suite réelle convergeant vers a .

Alors :

$$\lim_{n \rightarrow +\infty} f(x_n) = f(a)$$

PREUVE. C'est une conséquence immédiate de la propriété de limite séquentielle 14.4 page 364 et de la définition de continuité. ☕

Exemple. Soit $f \in \mathbb{R}^{\mathbb{R}}$ une fonction continue en 0. On suppose que : $\forall x \in \mathbb{R}, f(x) = f\left(\frac{x}{2}\right)$. Alors f est constante.

En effet, soit x un réel quelconque. La suite de TG $x/2^n$ converge vers 0. D'après la propriété de continuité séquentielle, la suite de TG $f(x/2^n)$ converge vers $f(0)$. Par ailleurs, une récurrence immédiate à partir de l'hypothèse de l'énoncé permet de montrer que $f(x/2^n) = f(x)$ pour tout $n \in \mathbb{N}$. On en déduit que $f(x) = f(0)$, d'où la conclusion.

Pour conclure ce paragraphe, mentionnons que cette propriété de continuité séquentielle est un ingrédient essentiel dans la preuve du célèbre théorème des valeurs intermédiaires, faisant l'objet de la section suivante.

14.3.3 Lemme de Cauchy et théorème des valeurs intermédiaires

THÉORÈME 14.3 - Soit f une fonction continue sur $[a, b]$ et à valeurs réelles.

On suppose que $f(a)f(b) \leq 0$. Alors : $\exists c \in [a, b], f(c) = 0$.

PREUVE. La preuve du lemme de Cauchy repose sur l'algorithme de dichotomie, dont voici une illustration et une brève présentation.

L'hypothèse $f(a)f(b) \leq 0$ signifie que les réels $f(a)$ et $f(b)$ sont de signes opposés. Deux cas peuvent se présenter : ($f(a) \leq 0$ et $f(b) \geq 0$) ou ($f(a) \geq 0$ et $f(b) \leq 0$).

Sans nuire à la généralité, on peut supposer ($f(a) \leq 0$ et $f(b) \geq 0$) ; dans l'autre cas, il suffira de remplacer f par $(-f)$.

Supposons $f(a) \leq 0$ et $f(b) \geq 0$.

L'idée est alors de construire deux suites (a_n) et (b_n) satisfaisant les conditions :

- $\forall n \in \mathbb{N}, \quad f(a_n) \leq 0 \text{ et } f(b_n) \geq 0;$
- $\forall n \in \mathbb{N}, \quad [a_{n+1}, b_{n+1}] \subset [a_n, b_n];$
- $\forall n \in \mathbb{N}, \quad b_n - a_n = \frac{b-a}{2^n}$

par récurrence comme indiqué ci-dessous.

► Initialisation : $a_0 = a$ et $b_0 = b$.

► Hérédité : on suppose construits les réels $a_0, a_1, \dots, a_n, b_0, b_1, \dots, b_n$ satisfaisant les conditions de la preuve. On veut construire a_{n+1} et b_{n+1} vérifiant les conditions au rang supérieur.

Pour cela, on “coupe l'intervalle $[a_n, b_n]$ en deux, et on regarde ce qui se passe au milieu”. Plus sérieusement et précisément :

$$\text{On pose : } m = \frac{a_n + b_n}{2}$$

Si $f(m) \leq 0$ alors

$$a_{n+1} = \frac{a_n + b_n}{2} \text{ et } b_{n+1} = b_n$$

Sinon ⁴

$$a_{n+1} = a_n \text{ et } b_{n+1} = \frac{a_n + b_n}{2}$$

Dans les deux cas, les réels a_{n+1} et b_{n+1} vérifient les trois conditions :

- i) $f(a_{n+1}) \leq 0$ et $f(b_{n+1}) \geq 0;$
- ii) $[a_{n+1}, b_{n+1}] \subset [a_n, b_n];$
- iii) $b_{n+1} - a_{n+1} = \frac{b_n - a_n}{2}$

Il ne “reste plus qu'à” établir que les suites (a_n) et (b_n) ont une limite commune, et que celle-ci est solution de l'équation $f(x) = 0$, ce que l'on fait à présent.

Pour le premier point, il suffit d'appliquer le théorème des segments emboîtés (ou, ce qui revient au même, observer que les suites $(a_n)_n$ et $(b_n)_n$ sont adjacentes). Posons alors : $c = \lim_{n \rightarrow +\infty} a_n = \lim_{n \rightarrow +\infty} b_n$.

Par continuité séquentielle, on a : $f(c) = \lim_{n \rightarrow +\infty} f(a_n) = \lim_{n \rightarrow +\infty} f(b_n)$.

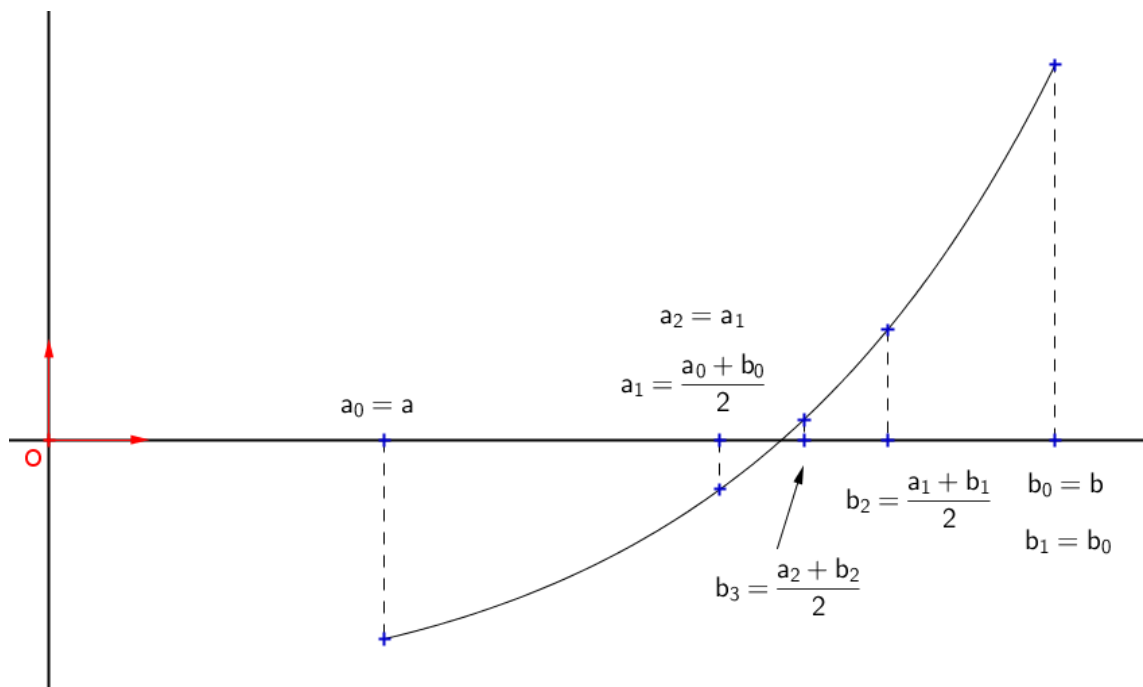
Et par stabilité des inégalités larges par passage à la limite, on a : $f(c) \leq 0$ et $f(c) \geq 0$.

Il s'ensuit que $f(c) = 0$. ☕

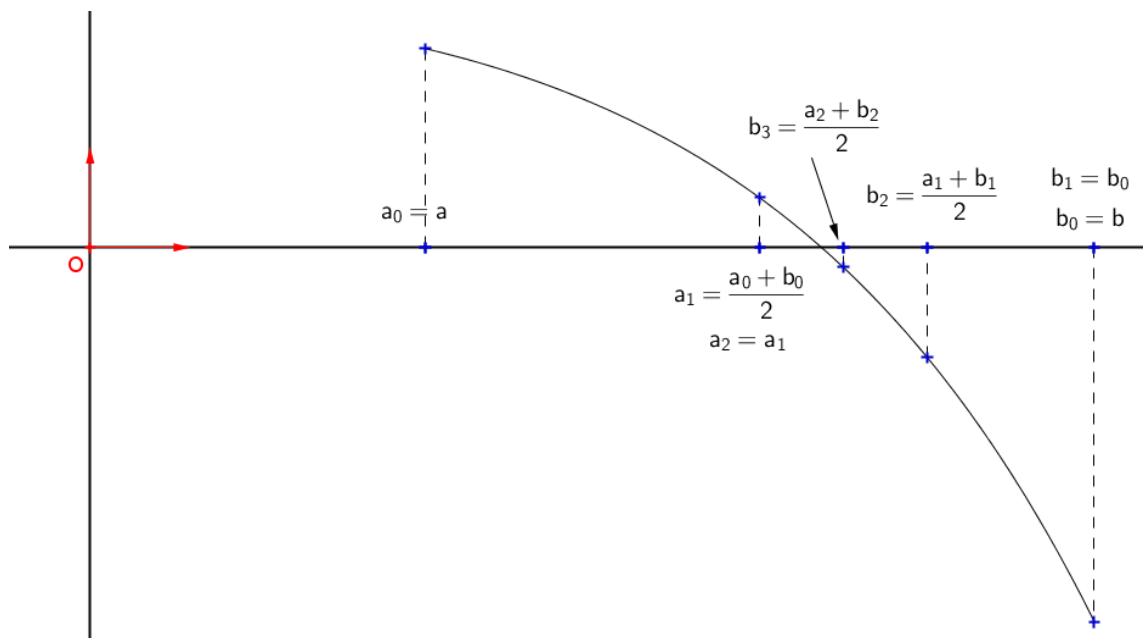
4. C'est-à-dire si $f(m) > 0$.

Illustrations de la méthode de dichotomie

Dans le cas où $f(a) \leq 0$ et $f(b) \geq 0$



Dans le cas où $f(a) \geq 0$ et $f(b) \leq 0$



Remarques. Que l'on ne s'y trompe pas ! S'il peut arriver qu'un *lemme* désigne un résultat technique de "moindre importance", ou un énoncé intermédiaire dans une preuve plus élaborée, le lemme de Cauchy est d'un intérêt crucial puisqu'il s'agit déjà du théorème des valeurs intermédiaires. Plus explicitement, le "TVI" tel que vous le connaissez depuis l'an dernier en est corollaire immédiat.

La seconde remarque concerne le "mode de fonctionnement" du lemme de Cauchy : s'il permet d'affirmer que sous ses hypothèses, une équation admet une solution, il ne permet pas d'établir l'unicité de cette solution (lorsqu'elle est unique), ni *a contrario* s'il en existe plusieurs.

THÉORÈME 14.4 - Soit f une fonction continue sur $[a, b]$ et à valeurs réelles.
Soit k un réel compris entre $f(a)$ et $f(b)$. Alors : $\exists c \in [a, b], f(c) = k$.

PREUVE. Il suffit d'appliquer le lemme de Cauchy à la fonction continue $x \in [a, b] \mapsto f(x) - k$. ☕

14.3.4 TVI : applications et conséquences

14.3.4.1 Résolutions d'équations

Il est clair que le TVI permet de prouver qu'une équation donnée possède (au moins) une solution. Ce qui peut constituer une précieuse information, lorsque l'équation ne peut être résolue explicitement (ce qui est génériquement le cas...).

Exemple. L'équation $\cos(x) = x$ admet une solution dans l'intervalle $[0, \pi/2]$.

14.3.4.2 Théorème du point fixe

THÉORÈME 14.5 - (théorème (du point fixe)). Soit $f : [a, b] \rightarrow [a, b]$ une fonction continue sur $[a, b]$ et à valeurs dans $[a, b]$.

Alors : $\exists c \in [a, b], f(c) = c$.

PREUVE. Par hypothèse, la fonction $g : x \mapsto f(x) - x$ est continue sur $[a, b]$. Et par hypothèse encore, on a $g(a) \geq 0$ et $g(b) \leq 0$. La conclusion provient de ces observations et du lemme de Cauchy appliqué à la fonction g . ☕

14.3.4.3 Image continue d'un intervalle

THÉORÈME 14.6 - (Image continue d'un intervalle). Soit f une fonction continue sur un intervalle I .

Alors $f(I)$ est un intervalle.

PREUVE. Soit f une fonction continue sur un intervalle I .

Soient y et y' deux éléments de $f(I)$. S'ALG, on peut supposer $y \leq y'$. En outre, par définition de $f(I)$, il existe deux réels x et x' tels que $y = f(x)$ et $y' = f(x')$.

Considérons à présent un réel Y compris entre y et y' . Alors Y est une valeur intermédiaire de la fonction f (sur $[x, x']$ ou sur $[x', x]$) et il existe donc (d'après le TVI) un réel c compris entre x et x' tel que $f(c) = Y$.

Puisque c est compris entre x et x' (qui sont deux éléments de I) et que I est un intervalle, on peut affirmer que $c \in I$. Ce fait, combiné avec l'égalité $Y = f(c)$ permet d'affirmer que : $Y \in f(I)$.

En résumé, on a établi : $\forall (y, y') \in (f(I))^2, y \leq y', [y, y'] \subset f(I)$.

Cette assertion signifie exactement que $f(I)$ est un intervalle, ce qui prouve le théorème. ☕

Remarques. Le point-clef pour cette preuve est évidemment la caractérisation des intervalles de $\mathbb{R}$. Un intervalle de $\mathbb{R}$ est une partie $I \subset \mathbb{R}$ “sans trous”, dans le sens plus précis suivant : $\forall (x, x') \in I^2, [x, x'] \subset I$.

Par ailleurs, le théorème est très grossièrement faux pour les fonctions non continues. Par exemple, en notant U la fonction “échelon-unité”, on a $U([-1; 1]) = \{0, 1\}$.

COROLLAIRE 14.1 - (“Théorème de la bijection”) Soit f une fonction définie sur un segment $[a, b]$ et à valeurs réelles.

Si f est continue et strictement croissante (*resp.* strictement décroissante), alors f réalise une bijection de $[a, b]$ sur $[f(a), f(b)]$ (*resp.* de $[a, b]$ sur $[f(b), f(a)]$).

PREUVE. C’est une conséquence directe du TVI et de l’hypothèse de stricte monotonie faite sur la fonction continue f . ☕

A noter que l’on a déjà fait grand usage de ce théorème lors du chapitre sur les applications, notamment pour définir les fonctions circulaires réciproques.

14.3.5 Théorème des bornes atteintes

THÉORÈME 14.7 - (Théorème des bornes atteintes, TBA). Toute fonction définie et continue sur un segment est bornée et atteint ses bornes.

Explicitement : soit f une fonction continue sur un intervalle fermé borné $[a, b]$ (avec a et b réels, $a < b$). Alors : $\exists (c, d) \in [a, b]^2, \forall x \in [a, b], f(c) \leq f(x) \leq f(d)$

PREUVE. Soit f une fonction continue sur le segment $[a, b]$ (avec a et b réels, $a < b$).

On pose $M = \sup \{f(x) / x \in [a, b]\} \in \mathbb{R} \cup \{+\infty\}$.

D’après la caractérisation séquentielle de la borne supérieure : $\exists (x_n) \in [a, b]^{\mathbb{N}}, \lim_{n \rightarrow +\infty} f(x_n) = M$.

La suite (x_n) étant bornée⁵, le théorème de Bolzano-Weierstrass (théorème 11.9 page 258) permet d’affirmer que l’on peut en extraire une sous-suite $(x_{\varphi(n)})$ convergente. Notons : $c = \lim_{n \rightarrow +\infty} x_{\varphi(n)}$.

Puisque la suite $(x_{\varphi(n)})$ est une suite d’éléments de $[a, b]$, on a : $c \in [a, b]$ (par stabilité des inégalités larges par passage à la limite).

Il est alors légitime d’appliquer la propriété de continuité séquentielle à la fonction f , qui est continue en c car elle est en particulier continue sur $[a, b]$, pour obtenir : $f(c) = \lim_{n \rightarrow +\infty} f(x_{\varphi(n)})$ (♠).

Il reste à observer que la suite de terme général $f(x_{\varphi(n)})$ est extraite de la suite de terme général $f(x_n)$ pour affirmer que : $\lim_{n \rightarrow +\infty} f(x_{\varphi(n)}) = M$ (♣).

Par unicité de la limite, on déduit de (♠) et de (♣) que : $M = f(c)$. Ainsi la fonction f admet un maximum sur $[a, b]$, égal à $f(c)$.

On établit de façon analogue que f admet un minimum sur $[a, b]$. ☕

5. C’est une suite de réels de $[a, b]$.

Une application particulièrement importante du TBA est le théorème de Rolle, qui entraîne à son tour le théorème des accroissements finis, qui implique à son tour toutes les propriétés dont vous avez eu besoin en Analyse jusqu'au Bac.

Une application plus directe est une formule appréciée des Physiciens :

COROLLAIRE 14.2 - (Formule de la moyenne). Soit $f : [a; b] \longrightarrow \mathbb{R}$ une fonction continue. Alors :

$$\exists c \in [a; b], \quad f(c) = \frac{1}{b-a} \int_a^b f(x) \, dx$$

Interprétation. Le terme de droite de l'égalité est ce que l'on appelle **valeur moyenne** de la fonction f sur le segment $[a, b]$. Le corollaire signifie donc que cette valeur moyenne est une valeur prise par la fonction sur le segment $[a, b]$.

PREUVE. Puisque f est continue sur le segment $[a, b]$, le théorème des bornes atteintes permet d'affirmer que :

$$\exists (t, u) \in [a, b]^2, \quad \forall x \in [a, b], \quad f(t) \leq f(x) \leq f(u)$$

$$\text{Par suite : } \int_a^b f(t) \, dx \leq \int_a^b f(x) \, dx \leq \int_a^b f(u) \, dx$$

$$\text{D'où : } (b-a) f(t) \leq \int_a^b f(x) \, dx \leq (b-a) f(u)$$

$$\text{Soit encore : } f(t) \leq \frac{1}{b-a} \int_a^b f(x) \, dx \leq f(u)$$

Ce dernier encadrement fait apparaître le terme du milieu comme une valeur intermédiaire de la fonction f sur $[a, b]$. Le théorème du même nom permet d'affirmer que : $\exists c \in [a; b], \quad f(c) = \frac{1}{b-a} \int_a^b f(x) \, dx$, ce qui prouve la formule. ☕

14.4 Extension aux fonctions à valeurs complexes

Les fonctions considérées dans ce paragraphe sont définies sur une partie de $\mathbb{R}$ et sont à valeurs dans $\mathbb{C}$.

DÉFINITION 14.12 - Soient $f : I \longrightarrow \mathbb{C}$ une fonction, et $a \in I$ un réel. La fonction f est **continue en** a si $\lim_{x \rightarrow a} f(x)$ existe et vaut $f(a)$; ce qui se traduit avec des quantificateurs par :

$$\forall \varepsilon > 0, \quad \exists \alpha > 0, \quad |x - a| < \alpha \implies |f(x) - f(a)| < \varepsilon$$

On pourra observer que cette définition est analogue à celle donnée dans le cadre des fonctions à valeurs réelles ; encore une fois, seule la lecture des barres verticales est différente (module au lieu de valeur absolue).

Par ailleurs, comme lors du passage des suites réelles aux suites complexes, certaines définitions et propriétés concernant les limites et la continuité s'adaptent aussi bien aux fonctions à valeurs dans $\mathbb{R}$ que dans $\mathbb{C}$. C'est le cas notamment des énoncés concernant les propriétés algébriques des limites et des fonctions continues, ainsi que les propriétés séquentielles (limite et continuité).

Mais bien évidemment, toutes les propriétés faisant intervenir de près ou de loin la relation d'ordre (en particulier le sens de variation) ne pourront s'étendre aux fonctions à valeurs complexes. En particulier, on doit faire une croix sur le théorème des valeurs intermédiaires, qui n'est absolument plus valide pour les fonctions à valeurs dans $\mathbb{C}$.

Il reste néanmoins un moyen permettant d'utiliser une partie des résultats du chapitre précédent, dont l'énoncé est donné ci-dessous.

PROPRIÉTÉ 14.7 - (“Pont $\mathbb{R} \longleftrightarrow \mathbb{C}$ ” pour les fonctions continues) Soit $f \in \mathcal{C}^I$. Alors :

$$[f \in \mathcal{C}^0(I, \mathbb{C})] \iff [(\operatorname{Re} f) \in \mathcal{C}^0(I, \mathbb{R})] \wedge [(\operatorname{Im} f) \in \mathcal{C}^0(I, \mathbb{R})]$$

En d'autres termes, une fonction à valeurs complexes est continue SSI sa partie réelle et sa partie imaginaire sont des fonctions continues (à valeurs réelles).

PREUVE. Il s'agit de prouver que pour tout réel a dans I , il est équivalent de dire que $f(x)$ tend vers $f(a)$ lorsque x tend vers a , ou que $\operatorname{Re}(f(x))$ et $\operatorname{Im}(f(x))$ tendent respectivement vers $\operatorname{Re}(f(a))$ et $\operatorname{Im}(f(a))$ lorsque x tend vers a . A cette fin, il suffit d'adapter la preuve du “pont $\mathbb{R} \longleftrightarrow \mathbb{C}$ pour les suites”. Explicitement, on raisonne par double implication : l'une repose sur l'inégalité triangulaire, et la réciproque est triviale. ☕

Exemple. La fonction $f : t \in [0, \pi] \mapsto e^{it}$ est continue sur $[0, \pi]$, puisque les parties réelle et imaginaire de f (les fonctions cosinus et sinus) sont continues sur cet intervalle.

C'est l'occasion d'illustrer notre précédente affirmation concernant l'absence d'un “TVI” pour les fonctions à valeurs complexes. La fonction f ci-dessus est continue sur $[0, \pi]$, et vérifie $f(0)f(\pi) \leq 0$. Néanmoins, f ne s'annule pas sur $[0, \pi]$, puisque le module de f est constant égal à 1.

14.5 Epilogue

14.5.1 Quelles sont les fonctions continues telles que $f(x + y) = f(x) + f(y)$?

Fait. Soit f une fonction continue sur $\mathbb{R}$ et à valeurs réelles.

LASSE :

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

$$(2) \quad f \text{ est linéaire, c\`ad : } \exists a \in \mathbb{R}, \quad \forall x \in \mathbb{R}, \quad f(x) = ax$$

PREUVE. Supposons que f soit continue sur $\mathbb{R}$, et vérifie (1). Alors $f(0) = 0$.

Une récurrence aisée permet de prouver que $f(n) = nf(1)$ pour tout entier naturel n .

En utilisant le fait que $f(-x) = -f(x)$ pour tout réel x , on en déduit que $f(n) = nf(1)$ pour tout entier relatif n .

La propriété précédente s'étend facilement aux rationnels : $f(r) = rf(1)$ pour tout nombre rationnel r .

Le fait que $\mathbb{Q}$ est dense dans $\mathbb{R}$, la caractérisation séquentielle de la densité, et la propriété de continuité séquentielle permettent d'en déduire que : $f(x) = xf(1)$ pour tout réel x .

Ainsi f est linéaire, ce qui prouve l'implication (1) $\implies$ (2).

L'implication réciproque est triviale.

14.5.2 Ensemble des points de discontinuité d'une fonction à valeurs réelles

Fait. L'ensemble des points de discontinuité d'une fonction monotone sur $\mathbb{R}$ est au plus dénombrable (c\`ad fini ou dénombrable).

PREUVE. Soit f une fonction monotone sur $\mathbb{R}$. SNALG, on peut supposer f croissante.

Notons E l'ensemble des points de discontinuité de f .

Pour tout $a \in E$, on a nécessairement : $\lim_{x \rightarrow a^-} f(x) < \lim_{x \rightarrow a^+} f(x)$. On peut alors choisir un rationnel ρ_a tel que $\lim_{x \rightarrow a^-} f(x) < \rho_a < \lim_{x \rightarrow a^+} f(x)$ (puisque $\mathbb{Q}$ est dense dans $\mathbb{R}$).

Il est clair que l'hypothèse de croissance faite sur f et la définition de l'ensemble E imposent que pour a et a' deux éléments de E , on a : $[\implies a < a'] \implies [\rho_a < \rho_{a'}]$.

Il s'ensuit que l'application $F : a \in E \longmapsto \rho_a \in \mathbb{Q}$ est injective. Puisque $\mathbb{Q}$ est dénombrable, on en déduit que E est au plus dénombrable.

Remarque. L'hypothèse de monotonie est évidemment indispensable : la fonction indicatrice de $\mathbb{Q}$ est discontinue en tout réel (car $\mathbb{Q}$ et $\mathbb{R} \setminus \mathbb{Q}$ sont denses dans $\mathbb{R}$). L'ensemble de ses points de discontinuité ($\mathbb{R}$ tout entier) est donc non dénombrable.

14.5.3 Etude d'une suite récurrente

ENONCÉ. Soient a un réel strictement positif, et f la fonction définie sur $\mathbb{R}$ en posant :

$$\forall x \in \mathbb{R}, f(x) = \frac{x^3 + 3ax}{3x^2 + a}$$

Soit (u_n) la suite définie par : $u_0 \in \mathbb{R}$ et $\forall n \in \mathbb{N}, u_{n+1} = f(u_n)$.

1/ Etablir que l'équation $f(x) = x$ admet exactement trois racines réelles, que l'on précisera.

2/ Etudier le signe de $f(x) - x$, ainsi que la monotonie de f .

3/ Préciser le comportement de (u_n) , en discutant plusieurs cas suivant la valeur de u_0 .

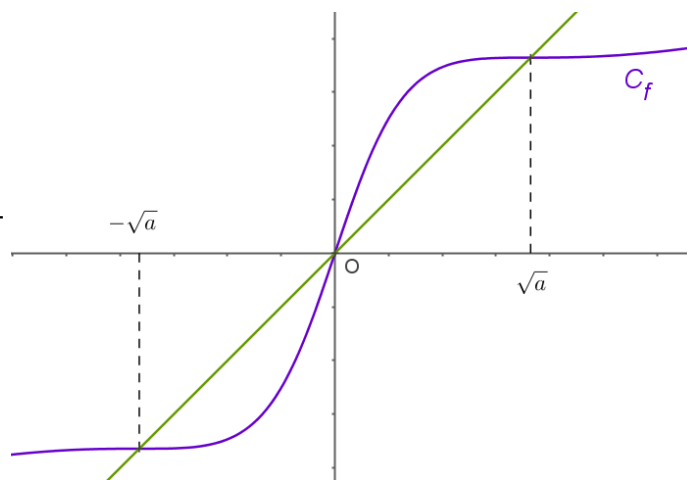
CORRIGÉ. 1/ Soit x un réel. On a :

$$\begin{aligned} [f(x) = x] &\iff \left[\frac{x^3 + 3ax}{3x^2 + a} = x \right] \iff \left[\frac{x^3 + 3ax - 3x^3 - ax}{3x^2 + a} = 0 \right] \iff [2(x^3 - ax) = 0] \\ &\iff [x = 0 \vee x = \sqrt{a} \vee x = -\sqrt{a}] \end{aligned}$$

Conclusion. L'équation $f(x) = x$ possède exactement trois racines réelles : 0, $\sqrt{a}$ et $-\sqrt{a}$.

2/ Soit x un réel. D'après les calculs de la question précédente, le réel $f(x) - x$ est du signe de $2(ax - x^3)$, c'est à dire du signe de : $x(a - x^2)$. On en déduit le tableau de signes de l'expression $f(x) - x$.

x	$-\infty$	$-\sqrt{a}$	0	$\sqrt{a}$	$+\infty$	
$f(x) - x$		+	0	-	0	-



Par ailleurs, la fonction f est de classe $\mathcal{C}^\infty$ sur $\mathbb{R}$, selon les théorèmes généraux. En particulier, f est dérivable sur $\mathbb{R}$, et pour tout réel x on a :

$$f'(x) = \frac{3(x^2 + a)(3x^2 + a) - 6x(x^3 + 3ax)}{(3x^2 + a)^2}$$

Il s'ensuit que le réel $f'(x)$ est du signe de son numérateur $N(x)$. Or :

$$N(x) = 9x^4 + 12ax^2 + 3a^2 - 6x^4 - 18ax^2 = 3(x^4 - 2ax^2 + a^2) = 3(x^2 + a)^2$$

On en déduit que pour tout réel x , le réel $N(x)$, et par suite le réel $f'(x)$, est strictement positif.

Conclusion. La fonction f est strictement croissante sur $\mathbb{R}$.

3/ Puisque la fonction f est (strictement) croissante sur $\mathbb{R}$, la suite (u_n) est monotone pour toute valeur de u_0 . Le tableau de signes de l'expression $f(x) - x$ nous conduit alors à distinguer plusieurs cas :

- Cas 1 : si $u_0 < -\sqrt{a}$. Alors pour tout entier naturel n , on a : $u_n < -\sqrt{a}$, comme le prouve un rapide raisonnement par récurrence. En notant $P(n)$ l'assertion " $u_n < -\sqrt{a}$ ", on obtient l'initialisation grâce à notre hypothèse de départ, et l'hérédité provient du fait que si $u_n < -\sqrt{a}$, alors $f(u_n) < f(-\sqrt{a})$ (puisque f est strictement croissante), d'où $u_{n+1} < -\sqrt{a}$ (par définition de la suite (u_n) , et car $-\sqrt{a}$ est un point fixe de f).

En outre : $u_1 \geq u_0$ d'après le tableau de signes de la question 2. Puisque l'on sait que la suite u est monotone, on en déduit que u est croissante.

En résumé, la suite u est croissante, et majorée par $-\sqrt{a}$. Elle est donc convergente d'après le théorème de la limite monotone. Comme tous les termes de u appartiennent à l'intervalle $] -\infty, -\sqrt{a}]$,⁶ que cet intervalle contient un unique point fixe de f (qui est $-\sqrt{a}$), et enfin que f est continue en $-\sqrt{a}$, on peut conclure que u converge et a pour limite $-\sqrt{a}$.

- Cas 2, 2-bis et 2-ter : si $u_0 = -\sqrt{a}$ ou $u_0 = 0$ ou $u_0 = \sqrt{a}$. Dans chacun de ces cas, la suite u est constante puisque les réels $-\sqrt{a}$, 0 et $\sqrt{a}$ sont les points fixes de f . En particulier, la suite u converge.
- Cas 3 : si $u_0 \in] -\sqrt{a}, 0[$. Alors pour tout entier naturel n , on a : $u_n \in] -\sqrt{a}, 0[$.⁷

En outre, la suite u est alors décroissante (cf question 2). Puisqu'elle est minorée (cf ci-dessus), elle converge. La fonction f étant continue sur $\mathbb{R}$, la suite u converge vers un des points fixes de f . Mais u ne peut converger vers $\sqrt{a}$ (immédiat), ni vers 0 puisque $u_0 < 0$ et u est décroissante.

On en déduit que u converge vers $-\sqrt{a}$.

- Cas 4 : si $u_0 \in]0, \sqrt{a}[$. Alors pour tout entier naturel n , on a : $u_n \in]0, \sqrt{a}[$.

En outre, la suite u est alors croissante (cf question 2). Puisqu'elle est majorée (cf ci-dessus), elle converge. La fonction f étant continue sur $\mathbb{R}$, la suite u converge vers un des points fixes de f . Mais u ne peut converger vers $-\sqrt{a}$ (immédiat), ni vers 0 puisque $u_0 > 0$ et u est croissante.

On en déduit que u converge vers $\sqrt{a}$.

- Cas 5 : si $u_0 > \sqrt{a}$. Alors pour tout entier naturel n , on a : $u_n > \sqrt{a}$.

En outre : $u_1 \leq u_0$ d'après le tableau de signes de la question 2. Puisque l'on sait que la suite u est monotone, on en déduit que u est décroissante.

En résumé, la suite u est décroissante, et minorée par $\sqrt{a}$. Elle est donc convergente d'après le théorème de la limite monotone. Comme tous les termes de u appartiennent à l'intervalle $[\sqrt{a}, +\infty[$, que cet intervalle contient un unique point fixe de f (qui est $\sqrt{a}$), et enfin que f est continue en $\sqrt{a}$, on peut conclure que u converge et a pour limite $\sqrt{a}$.

Conclusion. Pour toute valeur de u_0 , la suite u est convergente, et plus précisément :

- si $u_0 = 0$, alors $\lim_{n \rightarrow +\infty} u_n = 0$ (la suite u est alors constante égale à 0).
- si $u_0 > 0$, alors $\lim_{n \rightarrow +\infty} u_n = \sqrt{a}$.
- si $u_0 < 0$, alors $\lim_{n \rightarrow +\infty} u_n = -\sqrt{a}$.

6. Tous les termes de la suite u appartiennent à l'intervalle $] -\infty, -\sqrt{a}[$ et qui peut le plus peut le moins...

7. Cette affirmation peut être prouvée par récurrence sur n comme dans le cas 1, en utilisant les mêmes arguments (la croissance de f , et ses points fixes).

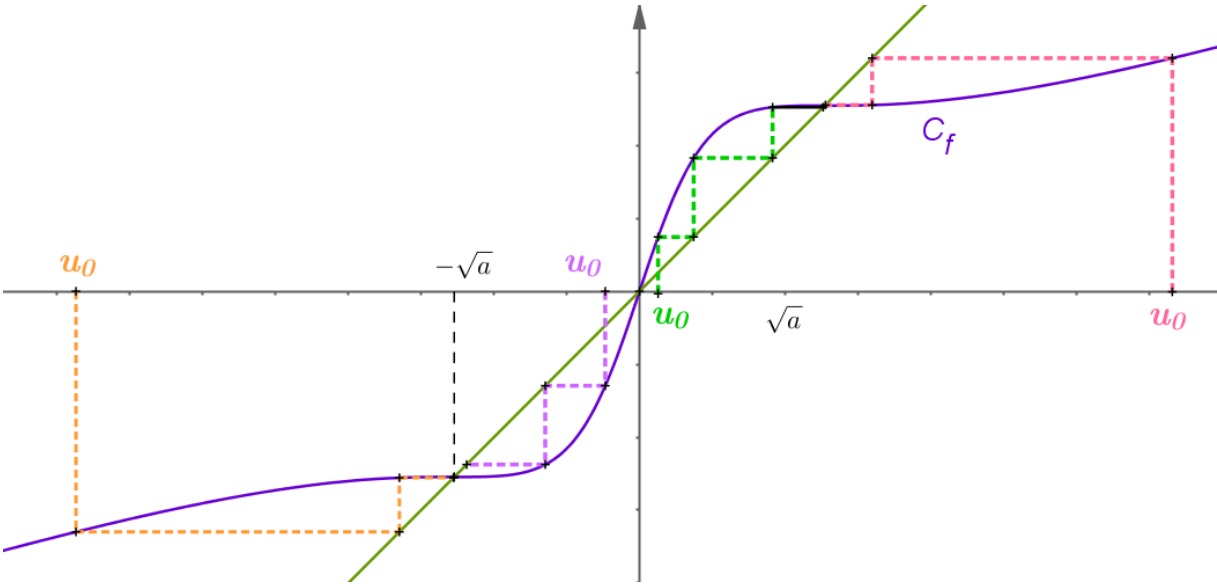


ILLUSTRATION — Pour tout $u_0 \in \mathbb{R}$, la suite u converge.

Chapitre 15

Ensembles finis et groupe symétrique

15.1 Ensembles finis

Notation. Soit n un entier naturel non nul. On note $\mathbb{N}_n = \llbracket 1, n \rrbracket$. Par convention : $\mathbb{N}_0 = \emptyset$.

DÉFINITION 15.1 - Soient E et F deux ensembles.

On dit que E **est équipotent à** F s'il existe une bijection de E dans F .

Remarque. La relation binaire “être équipotent à” (équipotence) est une **relation d'équivalence** (pour mémoire : réflexive, symétrique, transitive) sur les ensembles. Cette observation faite, nous pourrions dire par la suite que deux ensembles E et F sont équipotents s'il existe une bijection entre E et F .

LEMME 15.1 - La relation binaire d'équipotence est une relation d'équivalence.

PREUVE. Ecrivons $E \sim F$ lorsque E est équipotent à F .

- Réflexivité : pour tout ensemble E , l'application id_E est une bijection de E dans E . Donc $E \sim E$.
- Symétrie : soient E et F deux ensembles équipotents. Il existe donc une bijection $f : E \longrightarrow F$. Dans ce cas, la bijection réciproque f^{-1} est une bijection de F dans E . Par suite : $(E \sim F) \implies (F \sim E)$.
- Transitivité : soient E , F et G trois ensembles tels que $E \sim F$ et $F \sim G$. Alors il existe une bijection $f : E \longrightarrow F$ et une bijection $g : F \longrightarrow G$. La composée $g \circ f$ est une bijection de E dans G , puisque la composée de deux bijections en est encore une (cf colle 6). D'où $E \sim G$. En résumé : $[(E \sim F) \wedge (F \sim G)] \implies (E \sim G)$

Conclusion. La relation d'équipotence est une relation d'équivalence.



Exemples : les ensembles $E = \{1, 2, 3\}$ et $F = \{a, b, c\}$ sont équipotents ; $\mathbb{N}$ et $\mathbb{N}^*$ sont équipotents ; plus surprenant $\mathbb{N}$ et $\mathbb{Z}$ sont équipotents ; encore plus surprenant, $\mathbb{N}$ et $\mathbb{Q}$ sont équipotents (Cantor-Bernstein...). En revanche, $\mathbb{N}$ et $\mathbb{R}$ ne le sont pas (vous verrez l'an prochain que le premier est *dénombrable*, mais pas le second).

PROPRIÉTÉ 15.1 - (Exercice classique). $\mathbb{N}$ est équipotent à $\mathbb{N}^*$ et à $\mathbb{Z}$.

PREUVE. ➤ Pour le premier point : l'application $f : k \in \mathbb{N} \mapsto k + 1 \in \mathbb{N}^*$ est clairement bijective, de bijection réciproque $f^{-1} : k \in \mathbb{N}^* \mapsto k - 1 \in \mathbb{N}$. Ainsi : $\boxed{\mathbb{N} \sim \mathbb{N}^*}$.

➤ Pour le second point : on construit une application $f : \mathbb{N} \longrightarrow \mathbb{Z}$ en posant :

$$\forall n \in \mathbb{N}, f(n) = \begin{cases} \frac{n}{2} & \text{si } n \text{ est pair} \\ -\frac{n+1}{2} & \text{si } n \text{ est impair} \end{cases}$$

Montrons l'injectivité de f : soient n et n' deux entiers naturels tels que : $f(n) = f(n')$ (♠).

Si $f(n)$ et $f(n')$ sont tous deux positifs ou nuls, alors n et n' sont pairs, et $f(n) = n/2$ et $f(n') = n'/2$. D'où (♠) implique $n = n'$.

Si $f(n)$ et $f(n')$ sont tous deux strictement négatifs, alors n et n' sont impairs, et $f(n) = -(n+1)/2$ et $f(n') = -(n'+1)/2$. D'où (♠) implique encore $n = n'$.

Dans les deux cas : $f(n) = f(n') \implies n = n'$, ce qui prouve que f est injective.

Montrons la surjectivité de f : soit N un entier relatif. Si N est positif ou nul, alors N admet pour antécédent $2N$ par f . Et si N est strictement négatif, alors il admet pour antécédent $(-2N - 1)$ par f . On a donc établi que : $\forall N \in \mathbb{Z}, \exists n \in \mathbb{N}, f(n) = N$. Ce qui assure que f est surjective.

Par suite, l'application f est bijective, d'où : $\boxed{\mathbb{N} \sim \mathbb{Z}}$. ☕

DÉFINITION 15.2 - Un ensemble E est **fini** s'il existe un entier naturel n tel que E et $\mathbb{N}_n$ soient équipotents.

LEMME 15.2 - Soient n et m deux entiers naturels.

- 1) S'il existe une injection de $\mathbb{N}_n$ dans $\mathbb{N}_m$, alors $n \leq m$.
- 2) S'il existe une surjection de $\mathbb{N}_n$ dans $\mathbb{N}_m$, alors $n \geq m$.
- 3) S'il existe une bijection de $\mathbb{N}_n$ dans $\mathbb{N}_m$, alors $n = m$.

PREUVE. ➤ Montrons le 1). Raisonnons par récurrence sur l'entier naturel n en posant :

$\mathcal{P}(n)$: "s'il existe une injection de $\mathbb{N}_n$ dans $\mathbb{N}_m$, alors $n \geq m$ "

➡ Initialisation ($n = 0$) : OK. ¹

➡ Hérédité : supposons la propriété établie jusqu'à un certain entier naturel n , et montrons que $\mathcal{P}(n+1)$ est vraie.

Considérons $f : \mathbb{N}_{n+1} \longrightarrow \mathbb{N}_m$ une application injective. Deux cas peuvent alors se présenter ; soit $f(n+1) = m$, soit $f(n+1) \neq m$.

Premier cas — Si $f(n+1) = m$: on introduit alors la restriction de f à $\mathbb{N}_n$, que nous noterons g . Pour mémoire, il s'agit de l'application $g = f|_{\mathbb{N}_n} : \mathbb{N}_n \longrightarrow \mathbb{N}_m$ définie en posant : $\forall k \in \mathbb{N}_n, g(k) = f(k)$. ²

L'application g est encore injective, puisque f l'est et que la restriction d'une application injective l'est encore.

1. On pourra puissamment observer que tout entier naturel est supérieur ou égal à zéro.

2. " g est définie par la même formule que f , sur un ensemble de définition plus petit".

En outre g est à valeurs dans $\mathbb{N}_{m-1}$, puisque m ne peut avoir d'antécédent par g . En effet, s'il existait un entier $k \in \mathbb{N}_n$ tel que $g(k) = m$, alors on aurait $f(k) = m$ (par définition de g) et $f(n+1) = m$ (par hypothèse). L'injectivité de f impliquerait alors $k = n+1$, ce qui serait absurde puisque $k < n+1$.

En résumé, l'application g induit une application injective de $\mathbb{N}_n$ dans $\mathbb{N}_{m-1}$. Par hypothèse de récurrence, on en déduit que $n \leq m-1$, d'où $n+1 \leq m$.

Second cas — Si $f(n+1) \neq m$: on introduit alors l'application :

$$\begin{array}{ccc} \tau : \mathbb{N}_m & \longrightarrow & \mathbb{N}_m \\ m & \longmapsto & f(n+1) \\ f(n+1) & \longmapsto & m \\ k & \longmapsto & k \text{ si } k \neq f(n+1) \text{ et } k \neq m \end{array}$$

L'application τ est bijective, car c'est clairement une involution (Toronto est l'identité de $\mathbb{N}_m$!).³

Cette remarque faite, on introduit l'application : $F = \tau \circ f : \mathbb{N}_n \longrightarrow \mathbb{N}_m$. L'application F est injective, car c'est la composée de τ (injective car bijective) et de f (injective par hypothèse).

De plus, $F(n+1) = \tau(f(n+1)) = m$. On est ainsi ramenés au premier cas, et on peut donc conclure que $n+1 \leq m$.

Dans les deux cas, on a établi (sous l'hypothèse que $\mathcal{P}(n)$ est vraie) que s'il existe une application injective de $\mathbb{N}_{n+1}$ dans $\mathbb{N}_m$, alors $n+1 \leq m$. Ce qui prouve que la propriété $\mathcal{P}(n+1)$ est vraie. La propriété est donc initialisée et héréditaire, et on peut donc conclure.

Conclusion. S'il existe une injection de $\mathbb{N}_n$ dans $\mathbb{N}_m$, alors $n \leq m$.

► Montrons le 2). Soient n et m deux entiers naturels, et supposons qu'il existe une surjection de $\mathbb{N}_n$ dans $\mathbb{N}_m$. Notons alors $f : \mathbb{N}_n \longrightarrow \mathbb{N}_m$ une application surjective.

Soit $k \in \mathbb{N}_m$. Puisque f est surjective, $f^{-1}(\{k\})$ est non vide. Plus précisément, c'est une partie non vide de $\mathbb{N}_n$, et à ce titre elle admet un plus petit élément, disons N_k . Ce procédé, qui consiste à associer à un entier $k \in \mathbb{N}_m$ le plus petit de ses antécédents par f , donne lieu à une application :

$$\begin{array}{ccc} g : \mathbb{N}_m & \longrightarrow & \mathbb{N}_n \\ k & \longmapsto & N_k \end{array}$$

Par construction même, on a : $f \circ g = \text{id}_{\mathbb{N}_m}$. Il s'ensuit que $f \circ g$ est injective (car bijective), et donc que g est injective⁴.

D'après le point 1), puisque $g : \mathbb{N}_m \longrightarrow \mathbb{N}_n$ est injective, on peut conclure que $m \leq n$.

Conclusion. S'il existe une surjection de $\mathbb{N}_n$ dans $\mathbb{N}_m$, alors $n \geq m$.

3. Nous verrons plus tard que τ est un exemple de **transposition**.

4. Car : $f \circ g$ injective implique g injective. Preuve : supposons $f \circ g$ injective. Soient x et x' tels que $g(x) = g(x')$. Alors $f(g(x)) = f(g(x'))$ d'où $x = x'$ puisque $f \circ g$ est injective. En résumé : $g(x) = g(x')$ implique $x = x'$, ce qui prouve que g est injective.

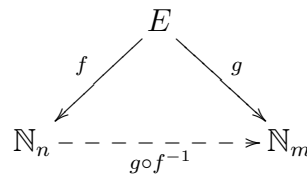
► Montrons le 3) (enfin un peu de repos). Soient n et m deux entiers naturels, et supposons qu'il existe une bijection de $f : \mathbb{N}_n \longrightarrow \mathbb{N}_m$. Puisque f est en particulier injective, on a $n \leq m$; et puisque f est en particulier surjective, on a $n \geq m$. Donc : $n = m$.

Conclusion. S'il existe une bijection de $\mathbb{N}_n$ dans $\mathbb{N}_m$, alors $n = m$.



Remarque : cette propriété est particulièrement importante, puisqu'elle permet de définir (de manière unique !) le cardinal d'un ensemble fini E . En effet, un ensemble fini E est un ensemble pour lequel il existe un entier n tel que E et $\mathbb{N}_n$ sont équipotents. Le "3)" de la propriété assure l'unicité de cet entier n , ce qui autorise à l'appeler cardinal de l'ensemble E .

En effet, supposons qu'il existe deux entiers naturels n et m et deux bijections $f : E \longrightarrow \mathbb{N}_n$ et $g : E \longrightarrow \mathbb{N}_m$:



Alors l'application $g \circ f^{-1}$ est bijective (composée de deux bijections) de $\mathbb{N}_n$ dans $\mathbb{N}_m$, d'où $n = m$ d'après la propriété.

DÉFINITION 15.3 - Soit E un ensemble fini. On appelle **cardinal de E** l'unique entier naturel n tel que E et $\mathbb{N}_n$ sont équipotents.

Notation. On note $\text{Card}(E)$, ou $\#E$ ou encore $|E|$ le cardinal d'un ensemble fini E .

LEMME 15.3 - Soient E un ensemble fini de cardinal $n \in \mathbb{N}^*$, et $a \in E$. Alors $E \setminus \{a\}$ est un ensemble fini de cardinal $n - 1$.

PREUVE. Considérons E un ensemble fini de cardinal non nul, et a un élément de E .

Puisque E est de cardinal n , il existe une bijection de $f : E \longrightarrow \mathbb{N}_n$. Deux cas peuvent alors se présenter ; soit $f(a) = n$, soit $f(a) \neq n$.

Premier cas — Si $f(a) = n$: on introduit alors la restriction de f à $E \setminus \{a\}$, que nous noterons $g : E \setminus \{a\} \longrightarrow \mathbb{N}_n$.

L'application g est encore injective, puisque f l'est et que la restriction d'une application injective l'est encore.

Par ailleurs : soit k un élément de $\mathbb{N}_{n-1}$. Puisqu'alors k appartient également à $\mathbb{N}_n$, k admet un unique antécédent x par f dans E . Or cet unique antécédent ne peut être a , puisque a est l'unique antécédent de n et que $k \neq n$. Il s'ensuit que :

$$\forall k \in \mathbb{N}_{n-1}, \exists! x \in E \setminus \{a\}, g(x) = k$$

Par suite, l'application g induit une bijection de $E \setminus \{a\}$ dans $\mathbb{N}_{n-1}$. D'où : $E \setminus \{a\} \sim \mathbb{N}_{n-1}$.

Second cas — Si $f(a) \neq n$: on introduit alors l'application :

$$\begin{array}{ccc} \tau : \mathbb{N}_n & \longrightarrow & \mathbb{N}_n \\ n & \longmapsto & f(a) \\ f(a) & \longmapsto & n \\ k & \longmapsto & k \text{ si } k \neq f(a) \text{ et } k \neq n \end{array}$$

L'application τ est bijective, car c'est une involution (Toronto...)

Donc l'application : $F = \tau \circ f : E \longrightarrow \mathbb{N}_n$ est bijective (c'est la composée de deux bijections) et a le bon goût de satisfaire la condition : $F(a) = n$. On est ainsi ramenés au premier cas, et on peut donc conclure que $E \setminus \{a\} \sim \mathbb{N}_{n-1}$.

Dans les deux cas, on a établi que $E \setminus \{a\}$ et $\mathbb{N}_{n-1}$ sont équipotents.

Conclusion. Si E est un ensemble fini de cardinal $n \in \mathbb{N}^*$, et $a \in E$, alors $E \setminus \{a\}$ est un ensemble fini de cardinal $n - 1$. ☕

THÉORÈME 15.1 - Soient E un ensemble fini, et $A \in \mathcal{P}(E)$.

Alors A est un ensemble fini, et $\text{Card}(A) \leq \text{Card}(E)$ (avec égalité SSI $A = E$).

PREUVE. Prouvons par récurrence sur n que toute partie d'un ensemble fini à n éléments est elle-même finie, en posant :

$\mathcal{P}(n)$: "si E est fini de cardinal n , alors toute partie A de E est finie et $\text{Card}(A) \leq \text{Card}(E)$ ".

➡ Initialisation ($n = 0$) : OK. ⁵

➡ Hérédité : supposons la propriété établie jusqu'à un certain entier naturel n , et montrons que $\mathcal{P}(n+1)$ est vraie.

Soit E un ensemble à $n+1$ éléments, et a un élément de E .

Soit A une partie de E . Deux cas peuvent alors se présenter : soit $a \in A$, soit $a \notin A$.

Premier cas — Si $a \in A$: alors l'ensemble $A \setminus \{a\}$ est une partie de l'ensemble $E \setminus \{a\}$. D'après la propriété précédente : $\text{Card}(E \setminus \{a\}) = \text{Card}(E) - 1 = n$. Ainsi $A \setminus \{a\}$ est une partie d'un ensemble à n éléments. Par hypothèse de récurrence, elle est donc finie, et : $\text{Card}(A \setminus \{a\}) \leq n$. Il s'ensuit que : $\text{Card}(A) \leq n+1$.

Second cas — Si $a \notin A$: alors l'ensemble A est une partie de l'ensemble $E \setminus \{a\}$, qui possède n éléments. Il s'ensuit, par hypothèse de récurrence, que A est finie, et que $\text{Card}(A) \leq n$. Puisque qui peut le plus peut le moins, on a aussi : $\text{Card}(A) \leq n+1$.

Dans les deux cas, on a établi que toute partie d'un ensemble à $n+1$ éléments est finie, de cardinal au plus égal à $n+1$, ce qui établit l'hérédité de la propriété.

Conclusion. Toute partie A d'un ensemble fini E est elle-même un ensemble fini, et $\text{Card}(A) \leq \text{Card}(E)$.

Cas d'égalité : montrons l'implication $[\text{Card}(A) = \text{Card}(E)] \implies [A = E]$, en établissant la contraposée. Si $A \neq E$, alors il existerait $x \in E \setminus A$. On aurait donc $A \subset E \setminus \{x\}$, et par suite : $\text{Card}(A) \leq (n-1)$, d'où en particulier : $\text{Card}(A) \neq \text{Card}(E)$. En résumé : $[A \subsetneq E] \implies [\text{Card}(A) \neq \text{Card}(E)]$, ce qui équivaut à : $[\text{Card}(A) = \text{Card}(E)] \implies [A = E]$. Comme il est par ailleurs clair que : $[A = E] \implies [\text{Card}(A) = \text{Card}(E)]$.

En résumé : $[A = E] \iff [\text{Card}(A) = \text{Card}(E)]$. ☕

5. La description des parties de l'ensemble vide étant relativement rapide.

15.2 Opérations sur les ensembles finis

PROPRIÉTÉ 15.2 - Soient A et B deux parties disjointes d'un même ensemble fini E .

Alors $A \cup B$ est fini, et $\text{Card}(A \cup B) = \text{Card}(A) + \text{Card}(B)$

PREUVE. Soient A et B deux parties disjointes d'un même ensemble fini E . On peut déjà affirmer que A et B sont des ensembles finis d'après la propriété précédente.

Débarrassons nous des cas extrêmes (et extrêmement triviaux) dans lesquels A et/ou B est vide. Si par exemple $A = \emptyset$, alors $A \cup B = B$ et $\text{Card}(A \cup B) = \text{Card}(B) = \text{Card}(A) + \text{Card}(B)$ puisque $\text{Card}(A) = 0 \dots$

Dans la suite des événements, on suppose donc que A et B sont non vides. Puisque ce sont des ensembles finis (cf supra), il existe deux entiers naturels non nuls n et m et deux bijections $f : A \rightarrow \mathbb{N}_n$ et $g : B \rightarrow \mathbb{N}_m$.

On construit alors judicieusement une application

$$\begin{aligned} F : A \cup B &\longrightarrow \mathbb{N}_{n+m} \\ x &\longmapsto f(x) && \text{si } x \in A \\ x &\longmapsto g(x) + n && \text{si } x \in B \end{aligned}$$

- L'application F est bien définie car A et B sont disjointes. Explicitement, si x appartient à $A \cup B$, alors $x \in A$ ou (exclusif) $x \in B$, et son image par F est donc définie de manière unique.

- Montrons l'injectivité de F : soient x et y deux éléments de $A \cup B$ tels que $F(x) = F(y)$ ($\spadesuit$). Trois cas peuvent se présenter :

- ➔ Si $x \in A$ et $y \in A$: alors ($\spadesuit$) implique $f(x) = f(y)$ d'où $x = y$ puisque f est injective.

- ➔ Si $x \in B$ et $y \in B$: alors ($\spadesuit$) implique $g(x) + n = g(y) + n$ d'où $g(x) = g(y)$ d'où $x = y$ puisque g est injective.

- ➔ Si $x \in A$ et $y \in B$: alors ($\spadesuit$) implique $f(x) = g(y) + n$. Or $f(x) \leq n$ (par construction de f) tandis que $g(y) + n > n$. L'égalité $f(x) = g(y) + n$ ne peut donc en aucun cas être vérifiée.

En résumé : $\forall (x, y) \in (A \cup B)^2, [F(x) = F(y)] \implies [x = y]$. L'application F est donc injective.

- Montrons la surjectivité de F : soit N un entier de $\mathbb{N}_{n+m}$. Deux cas se présentent : soit $N \in \mathbb{N}_n$, soit $N \in \llbracket n+1, n+m \rrbracket$.

- ➔ Si $N \in \mathbb{N}_n$: l'application f étant bijective, il existe un élément $a \in A$ tel que $f(a) = N$, d'où $F(a) = N$.

- ➔ Si $N \in \llbracket n+1, n+m \rrbracket$: alors $N - n \in \mathbb{N}_m$. L'application g étant bijective, il existe un élément $b \in B$ tel que $g(b) = N - n$, d'où $g(b) + n = N$ d'où encore $F(b) = N$.

On a donc établi que : $\forall N \in \mathbb{N}_{n+m}, \exists x \in A \cup B, F(x) = N$. Ce qui assure que F est surjective.

- Par suite, l'application F est bijective, donc $A \cup B$ et $\mathbb{N}_{n+m}$ sont équipotents, donc $\text{Card}(A \cup B) = n+m$.

Conclusion. Soient A et B deux parties disjointes d'un même ensemble fini E . Alors l'ensemble $A \cup B$ est fini et $\text{Card}(A \cup B) = \text{Card}(A) + \text{Card}(B)$. ☕

COROLLAIRE 15.1 - Soit A une partie d'un ensemble fini E .

Alors $\bar{A}$ est un ensemble fini, et : $\text{Card}((E \setminus A)) = \text{Card}(E) - \text{Card}(A)$.

PREUVE. Il résulte encore une fois du théorème 15.1 que les ensembles A et $(E \setminus A)$ sont finis.

Par ailleurs, on peut observer que A et $(E \setminus A)$ sont disjoints, et que $E = A \cup (E \setminus A)$. On déduit de la première observation que $\text{Card}(A \cup (E \setminus A)) = \text{Card}(A) + \text{Card}((E \setminus A))$ et de la seconde que $\text{Card}(A \cup (E \setminus A)) = \text{Card}(E)$. La conclusion provient directement de ces deux égalités. ☕

COROLLAIRE 15.2 - Soient $n \in \mathbb{N}^*$ et $(A_i)_{i \in [1, n]}$ une famille de parties de E , deux à deux disjointes (c'est à dire : $\forall (i, j) \in \mathbb{N}_n^2, (i \neq j) \Rightarrow (A_i \cap A_j = \emptyset)$). On a :

$$\text{Card} \left(\bigcup_{i=1}^n A_i \right) = \sum_{i=1}^n \text{Card}(A_i)$$

PREUVE. Prouvons la propriété par récurrence sur n , en posant :

$\mathcal{P}(n)$: "si $(A_i)_{i \in [1, n]}$ est une famille de parties de E deux à deux disjointes, alors : $\text{Card} \left(\bigcup_{i=1}^n A_i \right) = \sum_{i=1}^n \text{Card}(A_i)$ ".

➡ Initialisation ($n = 1$) : OK.

➡ Hérédité : supposons la propriété établie jusqu'à un certain entier naturel non nul n , et montrons que $\mathcal{P}(n+1)$ est vraie.

Soit $(A_i)_{i \in [1, n+1]}$ une famille de parties de E deux à deux disjointes. On a : $\bigcup_{i=1}^{n+1} A_i = \left(\bigcup_{i=1}^n A_i \right) \cup A_{n+1}$.

Cette union est disjointe, car :

$$\left(\bigcup_{i=1}^n A_i \right) \cap A_{n+1} = \left(\bigcup_{i=1}^n \underbrace{A_i \cap A_{n+1}}_{=\emptyset \text{ (par H)}} \right) = \emptyset$$

On a donc : $\text{Card} \left(\left(\bigcup_{i=1}^n A_i \right) \cup A_{n+1} \right) = \text{Card} \left(\bigcup_{i=1}^n A_i \right) + \text{Card}(A_{n+1})$ (propriété 3) puis :

$$\text{Card} \left(\left(\bigcup_{i=1}^n A_i \right) \cup A_{n+1} \right) = \left[\sum_{i=1}^n \text{Card}(A_i) \right] + \text{Card}(A_{n+1}) \text{ (HR).}$$

D'où : $\text{Card} \left(\bigcup_{i=1}^{n+1} A_i \right) = \sum_{i=1}^{n+1} \text{Card}(A_i)$, ce qui prouve que $\mathcal{P}(n+1)$ est vraie, et achève la preuve de l'hérédité.

Conclusion. Soient $n \in \mathbb{N}^*$ et $(A_i)_{i \in [1, n]}$ une famille de parties de E , deux à deux disjointes.

$$\text{On a : } \text{Card} \left(\bigcup_{i=1}^n A_i \right) = \sum_{i=1}^n \text{Card}(A_i) \quad \text{☕}$$

COROLLAIRE 15.3 - Soient A et B deux parties d'un même ensemble E . On a :

$$\text{Card}(A \cup B) = \text{Card}(A) + \text{Card}(B) - \text{Card}(A \cap B)$$

PREUVE. Soient A et B deux parties d'un ensemble fini. On a : $A \cup B = (A \setminus B) \cup B$. Dans le terme de droite de cette égalité, l'union est disjointe, et la propriété 3 permet donc d'affirmer que : $\text{Card}(A \cup B) = \text{Card}(A \setminus B) + \text{Card}(B)$ (♠).

Par ailleurs : $A = (A \setminus B) \cup (A \cap B)$, et puisque cette union est encore disjointe, une nouvelle application de la propriété 15.2 page 386 donne : $\text{Card}(A) = \text{Card}(A \setminus B) + \text{Card}(A \cap B)$ (♣).

On déduit de (♠) et (♣) que : $\boxed{\text{Card}(A \cup B) = \text{Card}(A) + \text{Card}(B) - \text{Card}(A \cap B)}$. ☕

PROPRIÉTÉ 15.3 - Soient E et F deux ensembles finis. On a :

1/ $E \times F$ est fini, et $\text{Card}(E \times F) = \text{Card}(E) \times \text{Card}(F)$

2/ Pour tout $n \in \mathbb{N}^*$, E^n est fini et $\text{Card}(E^n) = [\text{Card}(E)]^n$

3/ F^E est fini, et $\text{Card}(F^E) = \text{Card}(F)^{\text{Card}(E)}$

PREUVE. Preuve de 1/ : puisque nous sommes entre nous, on peut supposer que $E = \{x_1, \dots, x_n\}$ et $F = \{y_1, \dots, y_m\}$ où l'on a noté n et m les cardinaux respectifs de E et F .

Alors $E \times F$ est l'ensemble des couples (x_i, y_j) obtenus en faisant parcourir $\mathbb{N}_n$ à l'indice i , et $\mathbb{N}_m$ à l'indice j . Il est clair que l'on obtient ainsi nm éléments distincts. D'où $\text{Card}(E \times F) = nm$, c'est à dire :

$$\boxed{\text{Card}(E \times F) = \text{Card}(E) \times \text{Card}(F)}$$

Preuve de 2/ : la preuve peut s'obtenir via une récurrence aisée sur n , l'hérédité provenant du fait que $E^{n+1} = E^n \times E$ et du point 1/.

Preuve de 3/ : notons encore une fois n le cardinal de E et m celui de F . Une application f de E dans F est uniquement déterminée par les images des éléments de E . En d'autres termes, f est définie de manière unique par $f(x_i)$ (avec $i \in \mathbb{N}_n$). Or pour chacune de ces n valeurs (qui sont des éléments de F), on a m choix possibles. On obtiendra donc m^n façons différentes de définir une application de E dans F . D'où : $\text{Card}(F^E) = m^n$, c'est à dire : $\boxed{\text{Card}(F^E) = \text{Card}(F)^{\text{Card}(E)}}$. ☕

THÉORÈME 15.2 - Soient E et F deux ensembles finis, et $f : E \rightarrow F$ une application.

Alors $f(E)$ est un ensemble fini, et $\text{Card}(f(E)) \leq \text{Card}(E)$.

En outre, f est injective SSI $\text{Card}(f(E)) = \text{Card}(E)$.

PREUVE. Notons $E = \{x_1, \dots, x_n\}$ où $n \in \mathbb{N}^*$ désigne le cardinal de E (on pourra laisser de côté le cas extrême où $E = \emptyset$).

Alors $f(E) = \{f(x_1), \dots, f(x_n)\}$ et il est clair que cet ensemble possède au plus n éléments, ce qui assure déjà que : $\text{Card}(f(E)) \leq \text{Card}(E)$.

Plus précisément, $\text{Card}(f(E)) = n$ si et seulement si les $f(x_i)$ sont deux à deux distincts, c'est à dire si et seulement si l'application f est injective, ce qui achève la preuve du théorème. ☕

Remarques : dans cet énoncé, seule la finitude de l'ensemble E intervient, et l'hypothèse “ F fini” est donc superflue.

Par ailleurs, le théorème tient toujours dans le cas particulier où $E = \emptyset$, même si elle peut paraître un peu “étrange”. Explicitement, un ensemble F étant donné il existe une unique application $f : \emptyset \longrightarrow F$ (parfois appelée application vide), l'application qui “n'associe rien à personne”. Dans ce contexte, $f(\emptyset) = \emptyset$, et on a donc : $\text{Card}(f(\emptyset)) = \text{Card}(\emptyset) = 0$; ce qui ne doit pas nous surprendre, puisque l'application vide est injective.⁶

COROLLAIRE 15.4 - Soient E et F deux ensembles finis, et $f : E \longrightarrow F$ une application. On suppose $\text{Card}(E) = \text{Card}(F)$. L'ASSE :

- 1/ f est bijective
- 2/ f est injective
- 3/ f est surjective

Remarque : il est bon d'observer que l'énoncé ci-dessus est en général complètement faux sans l'hypothèse d'égalité des cardinaux.

PREUVE. Montrons $1) \implies 2) \implies 3) \implies 1)$; la boucle sera ainsi bouclée !

Supposons E et F finis de même cardinal $n \in \mathbb{N}^*$.

- $1) \implies 2)$: trivial.
- $2) \implies 3)$: supposons f injective. On peut commencer par noter que $f(E)$ est une partie de F , ça ne mange pas de pain. En outre, puisque f est injective, $f(E)$ est un ensemble fini, de cardinal égal à celui de E (en vertu du théorème 15.2). Or $\text{Card}(E) = \text{Card}(F)$ (par hypothèse). Donc l'ensemble $f(E)$ est une partie de F , de cardinal égal à celui de F ; il s'ensuit que $f(E) = F$ (théorème 15.1 page 385). Donc f est surjective, ce qui prouve l'implication.
- $3) \implies 1)$: supposons à présent f surjective. Alors $f(E) = F$, donc $\text{Card}(f(E)) = \text{Card}(F)$, et par conséquent $\text{Card}(f(E)) = \text{Card}(E)$ (puisque $\text{Card}(E) = \text{Card}(F)$ par hypothèse). On en déduit (grâce au théorème 15.2 page 388) que f est injective, et donc bijective. Ce qui prouve l'implication, et achève la preuve du théorème. ☕

15.3 Injections et permutations

LEMME 15.4 - Soient E et F deux ensembles finis, de cardinaux respectifs p et n . Le nombre d'injections de E dans F est l'entier :

$$A_n^p = \frac{n!}{(n-p)!} \text{ si } p \in \llbracket 0, n \rrbracket, \text{ et } A_n^p = 0 \text{ si } p > n.$$

PREUVE. Commençons par observer que $f(E)$ étant une partie de F , on a $\text{Card}(f(E)) \leq n = \text{Card}(F)$. Par ailleurs, f est injective si et seulement si $\text{Card}(f(E)) = \text{Card}(E) = p$. Il s'ensuit qu'il ne peut pas exister d'application injective de E dans F si $p > n$, ce qui prouve déjà la seconde partie de la proposition.

6. En effet, pour tout couple (x, y) de $\emptyset \times \emptyset$, la condition $f(x) = f(y)$ implique $x = y$ est vérifiée...

Supposons maintenant $p \leq n$, et notons $E = \{x_1, \dots, x_p\}$. Pour définir une application f , on peut choisir successivement les valeurs de $f(x_1), \dots, f(x_p)$. Si l'on souhaite que l'application f soit injective, on doit prendre ces éléments distincts dans F . On a alors n choix pour $f(x_1)$, $n-1$ choix pour $f(x_2)$, $\dots$, $(n-p+1)$ choix pour $f(x_p)$. Il existe donc : $n(n-1) \cdots (n-p+1)$ injections de E dans F , d'où : pour tout $p \in \llbracket 0, n \rrbracket$, il existe $\frac{n!}{(n-p)!}$ injections de E dans F . ☕

DÉFINITION 15.4 - Soient n un entier naturel non nul, $p \in \llbracket 1, n \rrbracket$ et E un ensemble de cardinal n . Un **arrangement** de p éléments de E est un p -uplet d'éléments distincts de E .

Exemples. Soit $E = \{1, 2, 3, 4\}$. Les triplets $(4, 2, 1)$ et $(1, 4, 2)$ sont des arrangements de E . Bien qu'ils contiennent les mêmes éléments, ils sont distincts, puisque dans un triplet l'ordre doit être pris en compte. Ici réside la différence fondamentale entre arrangements et combinaisons : les premiers sont ordonnés, tandis que les secondes ne le sont pas. Pour enfoncer le clou : les arrangements $(2, 3)$ et $(3, 2)$ sont distincts ; alors que $\{2, 3\}$ et $\{3, 2\}$ sont la même combinaison.

DÉFINITION 15.5 - Soit E un ensemble. Une **permutation** de E est une bijection de E dans E .

Exemples. L'application $x \mapsto x^3$ est une permutation de $\mathbb{R}$. L'application $k \mapsto 6-k$ est une permutation de $\mathbb{N}_5$. L'application qui à 1 associe 2, à 2 associe 3, et à 3 associe 1 est une permutation de $\mathbb{N}_3$.

Notation. Soit E un ensemble. On note S_E l'ensemble des permutations de E .

THÉORÈME 15.3 - Si E est un ensemble fini de cardinal n , alors $(S_E, \circ)$ est un groupe fini de cardinal $n!$.

PREUVE. Un élément de S_E est une bijection de E dans E . En vertu du corollaire 15.4 page 389, il existe autant de bijections de E dans E que d'injections de E dans E (puisque E est fini). Or d'après le lemme 15.4 page 389, il existe exactement $n!$ injections de E dans E .

Par suite, il existe $n!$ bijections de E dans E , ce qui assure déjà que : $\text{Card}(S_E) = n!$

En outre, la loi de composition (" $\circ$ ") est une *loi de composition interne* dans S_E , puisque la composée de deux bijections est une bijection. Cette loi est *associative* car plus généralement la composition des applications l'est. Elle possède un *élément neutre* qui est l'identité de E . Enfin, tout élément f de S_E est *inversible* (pour la loi " $\circ$ ") dans S_E , car si f est une bijection de E dans E , sa réciproque f^{-1} est elle aussi une bijection de E dans E .

Ce qui prouve que $(S_E, \circ)$ est un groupe, et achève cette preuve. ☕

15.4 Groupe symétrique

15.4.1 Généralités

DÉFINITION 15.6 - Soit n un entier naturel non nul. On note S_n (ou $(S_n, \circ)$) et on appelle **groupe symétrique** l'ensemble des permutations de $\mathbb{N}_n$.

Une conséquence directe du théorème 15.3 est la :

PROPRIÉTÉ 15.4 - Pour tout entier naturel non nul n , $(S_n, \circ)$ est un groupe fini de cardinal $n!$.

PREUVE. C'est un cas particulier du théorème 15.3, appliqué à l'ensemble $E = \mathbb{N}_n$. ☕

Convention. Dans la suite des évènements, on notera **multiplicativement** la loi de S_n . Explicitement, si σ et τ désignent deux permutations, on parlera du produit $\sigma\tau$ pour désigner la composée $\sigma \circ \tau$.

Exemples. $S_1 = \{\text{id}\}$ est le groupe à un élément ; $S_2 = \{\text{id}_{\mathbb{N}_2}, (12)\}$ est le groupe à deux éléments. Tous deux sont abéliens.

En revanche, $S_3 = \{\text{id}, (12), (13), (23), (123), (132)\}$ est un groupe à 6 éléments, non abélien (voir plus loin ou chapitre 11).

15.4.2 Transpositions et p -cycles

DÉFINITION 15.7 - Soit $n \geq 2$ un entier, et soient i et j deux éléments distincts de $\mathbb{N}_n$. On définit une application de $\mathbb{N}_n$ dans $\mathbb{N}_n$, que l'on note (ij) en posant :

$$\forall k \in \mathbb{N}_n, (ij)(k) = \begin{cases} j & \text{si } k = i \\ i & \text{si } k = j \\ k & \text{si } k \neq i \text{ et } k \neq j \end{cases}$$

Une telle application est appelée **transposition**.

Des exemples de transpositions de $\mathbb{N}_3$ ont déjà été utilisés dans le chapitre 11 sur les groupes, pour établir que le groupe S_3 (de cardinal 6) est non abélien.

PROPRIÉTÉ 15.5 - Une transposition est une involution (et donc une permutation).

PREUVE. Soient n un entier naturel supérieur ou égal à 2, et i et j deux entiers distincts de $\mathbb{N}_n$. Considérons la transposition (ij) , c'est à dire pour mémoire l'application définie sur $\mathbb{N}_n$ et à valeurs dans $\mathbb{N}_n$ en posant :

$$\forall k \in \mathbb{N}_n, (ij)(k) = \begin{cases} j & \text{si } k = i \\ i & \text{si } k = j \\ k & \text{si } k \neq i \text{ et } k \neq j \end{cases}$$

Soit k un élément de $\mathbb{N}_n$.

➡ Si $k \neq i$ et $k \neq j$: alors $[(ij)(ij)](k) = (ij)((ij)(k)) = (ij)(k) = k$.

➡ Si $k = i$: alors $[(ij)(ij)](i) = (ij)((ij)(i)) = (ij)(j) = i$.

➡ Si $k = j$: alors $[(ij)(ij)](j) = (ij)((ij)(j)) = (ij)(i) = j$.

On a donc établi que : $\forall k \in \mathbb{N}_n, [(ij)(ij)](k) = k$, ce qui signifie que $(ij)(ij) = \text{id}_{\mathbb{N}_n}$.

Ainsi la transposition (ij) est une involution, et par conséquent une permutation de $\mathbb{N}_n$ (donc un élément de S_n). ☕

LEMME 15.5 - Soient $n \geq 3$ un entier ; et i, j et k trois entiers distincts dans $\mathbb{N}_n$. Alors :

$$(ij)(jk) \neq (jk)(ij)$$

PREUVE. Calculons l'image de i par chacune des deux permutations intervenant dans l'énoncé.

D'une part : $[(ij)(jk)](i) = (ij)(i) = j$ et d'autre part $[(jk)(ij)](i) = (jk)(j) = k$.

Puisque les images de i par $(ij)(jk)$ et $(jk)(ij)$ sont distinctes ($i \neq k$ par hyp), on peut affirmer que les applications $(ij)(jk)$ et $(jk)(ij)$ sont distinctes, ce qu'il fallait établir. ☕

COROLLAIRE 15.5 - Le groupe symétrique S_n est non abélien dès que $n \geq 3$.

PREUVE. Conséquence directe du lemme précédent. ☕

DÉFINITION 15.8 - Soient $n \geq 2$ un entier, et p un entier tel que $2 \leq p \leq n$. Soient encore $a_1, \dots, a_p$ p éléments distincts de $\mathbb{N}_n$. On définit une application de $\mathbb{N}_n$ dans $\mathbb{N}_n$, que l'on note $(a_1 a_2 \cdots a_p)$ en posant :

$$\begin{cases} \forall i \in \mathbb{N}_{p-1}, (a_1 a_2 \cdots a_p)(a_i) = a_{i+1} \\ (a_1 a_2 \cdots a_p)(a_p) = a_1 \\ \forall k \in \mathbb{N}_n \setminus \{a_1, \dots, a_p\}, (a_1 a_2 \cdots a_p)(k) = k \end{cases}$$

Une telle application est appelée un **p -cycle** ; l'ensemble $\{a_1, a_2, \dots, a_p\}$ est appelé son **support**.

PROPRIÉTÉ 15.6 - Un p -cycle est une permutation. En outre, avec les notations introduites plus haut :

$$(a_1 a_2 \cdots a_p)^{-1} = (a_p a_{p-1} \cdots a_1)$$

PREUVE. Soit k un élément de $\mathbb{N}_n$.

➡ Si $k \notin \{a_1, \dots, a_p\}$: alors $[(a_p a_{p-1} \cdots a_1)(a_1 a_2 \cdots a_p)](k) = k$ puisque k est laissé invariant par les p -cycles $(a_p a_{p-1} \cdots a_1)$ et $(a_1 a_2 \cdots a_p)$.

➡ Si $k \in \{a_1, \dots, a_{p-1}\}$: il revient au même de dire qu'il existe $i \in \mathbb{N}_{p-1}$ tel que $k = a_i$. Alors :

$$[(a_p a_{p-1} \cdots a_1)(a_1 a_2 \cdots a_p)](k) = (a_p a_{p-1} \cdots a_1)((a_1 a_2 \cdots a_p)(a_i)) = (a_p a_{p-1} \cdots a_1)(a_{i+1}) = a_i$$

➡ Si $k = a_p$: Alors :

$$[(a_p a_{p-1} \cdots a_1)(a_1 a_2 \cdots a_p)](a_p) = (a_p a_{p-1} \cdots a_1)(a_1) = a_p$$

Conclusion intermédiaire : $\forall k \in \mathbb{N}_n$,

$$[(a_p a_{p-1} \cdots a_1)(a_1 a_2 \cdots a_p)](k) = k, \text{ soit : } (a_p a_{p-1} \cdots a_1)(a_1 a_2 \cdots a_p) = \text{id}_{\mathbb{N}_n}.$$

A'r'cominche, din l'aut'sens :

➡ Si $k \notin \{a_1, \dots, a_p\}$: comme dans le cas précédent $[(a_1 a_2 \cdots a_p)(a_p a_{p-1} \cdots a_1)](k) = k$.

➡ Si $k \in \{a_2, \dots, a_p\}$: il revient au même de dire qu'il existe $i \in \llbracket 2, p \rrbracket$ tel que $k = a_i$. Alors :

$$[(a_1 a_2 \cdots a_p)(a_p a_{p-1} \cdots a_1)](k) = (a_1 a_2 \cdots a_p)(a_p a_{p-1} \cdots a_1)((a_i)) = (a_1 a_2 \cdots a_p)(a_{i-1}) = a_i$$

➡ Si $k = a_1$: Alors :

$$[(a_1 a_2 \cdots a_p) (a_p a_{p-1} \cdots a_1)](k) = (a_1 a_2 \cdots a_p) (a_p a_{p-1} \cdots a_1) ((a_1)) = (a_1 a_2 \cdots a_p) (a_p) = a_1$$

Conclusion intermédiaire bis : $\forall k \in \mathbb{N}_n$,

$$[(a_1 a_2 \cdots a_p) (a_p a_{p-1} \cdots a_1)](k) = k, \text{ soit : } (a_1 a_2 \cdots a_p) (a_p a_{p-1} \cdots a_1) = \text{id}_{\mathbb{N}_n}.$$

Conclusion. D'après les calculs précédents : $(a_p a_{p-1} \cdots a_1) (a_1 a_2 \cdots a_p) = \text{id}_{\mathbb{N}_n}$ et $(a_1 a_2 \cdots a_p) (a_p a_{p-1} \cdots a_1) = \text{id}_{\mathbb{N}_n}$. Par suite, l'application $(a_1 a_2 \cdots a_p)$ est une bijection de $\mathbb{N}_n$ dans $\mathbb{N}_n$; c'est donc une permutation, d'inverse $(a_p a_{p-1} \cdots a_1)$ dans S_n . ☕

Remarque. Deux cycles sont à **supports disjoints** lorsque... leurs supports sont disjoints (pas d'applaudissements!). Deux cycles à supports disjoints commutent.

Exemples. Une transposition est un 2-cycle. Tous les éléments (distincts de l'identité) du groupe S_3 sont des cycles. En revanche, dans S_4 , il existe des permutations qui ne sont pas des cycles, comme par exemple le produit (12) (34).

PROPRIÉTÉ 15.7 - Mêmes notations que plus haut. On a :

$$(a_1 a_2 \cdots a_p) = \prod_{i=1}^{p-1} (a_i a_{i+1})$$

PREUVE. Soit k un élément de $\mathbb{N}_n$.

➡ Si $k \notin \{a_1, \dots, a_p\}$: alors $(a_1 a_2 \cdots a_p)(k) = k$ puisque k n'appartient pas au support du cycle, et $\left[\prod_{i=1}^{p-1} (a_i a_{i+1}) \right](k) = k$ puisque k n'appartient au support d'aucune transposition $(a_i a_{i+1})$ du produit.

$$\text{D'où : } \forall k \in \mathbb{N}_n, k \notin \{a_1, \dots, a_p\}, \left[\prod_{i=1}^{p-1} (a_i a_{i+1}) \right](k) = (a_1 a_2 \cdots a_p)(k)$$

➡ Si $k \in \{a_1, \dots, a_{p-2}\}$: il revient au même de dire qu'il existe $m \in \mathbb{N}_{p-1}$ tel que $k = a_m$. Alors d'une part :

$$(a_1 a_2 \cdots a_p)(k) = (a_1 a_2 \cdots a_p)(a_m) = a_{m+1}$$

Et d'autre part :

$$\begin{aligned} \left[\prod_{i=1}^{p-1} (a_i a_{i+1}) \right](a_m) &= \left[\left(\prod_{i=1}^{m-1} (a_i a_{i+1}) \right) (a_m a_{m+1}) \left(\prod_{i=m+1}^{p-1} (a_i a_{i+1}) \right) \right](a_m) \\ &= \left[\left(\prod_{i=1}^{m-1} (a_i a_{i+1}) \right) (a_m a_{m+1}) \right] \left(\left(\prod_{i=m+1}^{p-1} (a_i a_{i+1}) \right) (a_m) \right) \end{aligned}$$

Puisque a_m n'appartient au support d'aucune transposition $(a_i a_{i+1})$ pour $i \geq m+1$, on a :

$$\left(\prod_{i=m+1}^{p-1} (a_i a_{i+1}) \right) (a_m) = a_m$$

Il s'ensuit que :

$$\left[\prod_{i=1}^{p-1} (a_i a_{i+1}) \right] (a_m) = \left[\left(\prod_{i=1}^{m-1} (a_i a_{i+1}) \right) (a_m a_{m+1}) \right] (a_m) = \left[\left(\prod_{i=1}^{m-1} (a_i a_{i+1}) \right) \right] (a_{m+1})$$

De nouveau, puisque a_{m+1} n'appartient au support d'aucune transposition $(a_i a_{i+1})$ pour $i \leq m-1$, on a :

$$\left(\prod_{i=1}^{m-1} (a_i a_{i+1}) \right) (a_{m+1}) = a_{m+1}. \text{ D'où finalement : } \left[\prod_{i=1}^{p-1} (a_i a_{i+1}) \right] (a_m) = a_{m+1}.$$

$$\text{D'où : } \forall k \in \{a_1, \dots, a_{p-2}\}, \left[\prod_{i=1}^{p-1} (a_i a_{i+1}) \right] (k) = (a_1 a_2 \cdots a_p) (k)$$

➡ Si $k = a_{p-1}$: Alors d'une part : $(a_1 a_2 \cdots a_p) (a_{p-1}) = a_p$

Et d'autre part :

$$\left[\prod_{i=1}^{p-1} (a_i a_{i+1}) \right] (a_{p-1}) = \left[\left(\prod_{i=1}^{p-2} (a_i a_{i+1}) \right) (a_{p-1} a_p) \right] (a_{p-1}) = \left[\left(\prod_{i=1}^{p-2} (a_i a_{i+1}) \right) \right] (a_p) = a_p$$

La dernière égalité provenant de ce que a_p n'appartient au support d'aucune transposition $(a_i a_{i+1})$ pour $i \leq p-2$.

$$\text{D'où : } \left[\prod_{i=1}^{p-1} (a_i a_{i+1}) \right] (a_{p-1}) = (a_1 a_2 \cdots a_p) (a_{p-1})$$

➡ Si $k = a_p$: Alors d'une part : $(a_1 a_2 \cdots a_p) (a_p) = a_1$

Et d'autre part :

$$\left[\prod_{i=1}^{p-1} (a_i a_{i+1}) \right] (a_p) = \left[\left(\prod_{i=1}^{p-2} (a_i a_{i+1}) \right) (a_{p-1} a_p) \right] (a_p) = \left[\left(\prod_{i=1}^{p-2} (a_i a_{i+1}) \right) \right] (a_{p-1}) = \left[\left(\prod_{i=1}^{p-3} (a_i a_{i+1}) \right) \right] (a_{p-2})$$

$$= \cdots = \left[\left(\prod_{i=1}^2 (a_i a_{i+1}) \right) \right] (a_3) = (a_1 a_2) (a_2) = a_1 \quad \text{D'où : } \left[\prod_{i=1}^{p-1} (a_i a_{i+1}) \right] (a_p) = (a_1 a_2 \cdots a_p) (a_p)$$

Conclusion. $\forall k \in \mathbb{N}_n, \left[\prod_{i=1}^{p-1} (a_i a_{i+1}) \right] (k) = (a_1 a_2 \cdots a_p) (k)$ soit : $\prod_{i=1}^{p-1} (a_i a_{i+1}) = (a_1 a_2 \cdots a_p)$ ☕

Cette dernière propriété signifie que tout p -cycle s'écrit comme un produit d'exactly $(p-1)$ transpositions.

Les transpositions constituent donc les “briques élémentaires de S_n ”, dans un sens rendu plus précis par le théorème suivant :

THÉORÈME 15.4 - (Version courte). Les transpositions engendrent le groupe symétrique.

Ou, le même en plus explicite :

THÉORÈME 15.5 - (Version longue du théorème 15.4). Soit n un entier naturel supérieur ou égal à 2, et soit $\sigma \in S_n$.

Il existe un entier $k \leq n$, et k transpositions $\tau_1, \dots, \tau_k$ de S_n tels que : $\sigma = \prod_{i=1}^k \tau_i$.

PREUVE. Raisonnons par récurrence sur l'entier naturel n en posant pour tout $n \in \mathbb{N} \setminus \{0, 1\}$:

$\mathcal{P}(n)$: “Pour toute permutation $\sigma \in S_n$ il existe un entier $k \leq n$ et k transpositions $\tau_1, \dots, \tau_k$ de S_n tels que : $\sigma = \prod_{i=1}^k \tau_i$ ”

➔ Initialisation ($n = 2$) : les permutations de S_2 sont l'identité (produit de 0 transposition) et (12) (produit d'une transposition), ce qui assure que $\mathcal{P}(2)$ est vraie.

➔ Hérédité : supposons la propriété établie jusqu'à un certain entier naturel n , et montrons que $\mathcal{P}(n+1)$ est vraie.

Soit $\sigma \in S_{n+1}$. On peut distinguer deux cas : $\sigma(n+1) = n+1$ et $\sigma(n+1) \neq n+1$.

Premier cas — Si $\sigma(n+1) = n+1$: dans ce cas, la restriction $\sigma|_{\mathbb{N}_n}$ est une permutation de $\mathbb{N}_n$. Par hypothèse de récurrence, il existe un entier $k \leq n$ et k transpositions $\tau_1, \dots, \tau_k$ de S_n tels que : $\sigma|_{\mathbb{N}_n} = \prod_{i=1}^k \tau_i$.

Il reste à voir qu'une transposition dans S_n est aussi une transposition dans S_{n+1} ; en effet, une transposition de S_n s'écrit (ab) avec a et b deux entiers distincts dans $\mathbb{N}_n$; ce sont donc naturellement deux entiers distincts dans $\mathbb{N}_{n+1}$. Et puisque $\sigma(n+1) = n+1$, on a donc : $\sigma = \prod_{i=1}^k \tau_i$ (égalité dans S_{n+1}). Ainsi, sous l'hypothèse $\sigma(n+1) = n+1$, il existe un entier $k \leq n$ (en particulier $k \leq n+1$) et k transpositions $\tau_1, \dots, \tau_k$ de S_{n+1} tels que : $\sigma = \prod_{i=1}^k \tau_i$.

Second cas — Si $\sigma(n+1) \neq n+1$: dans ce cas, on introduit la permutation $\rho = (\sigma(n+1) \ n+1) \sigma$.⁷

ρ est un élément de S_{n+1} , qui vérifie : $\rho(n+1) = n+1$. En vertu de l'étude faite dans le premier cas : il existe un entier $k \leq n$ et k transpositions $\tau_1, \dots, \tau_k$ de S_{n+1} tels que : $\rho = \prod_{i=1}^k \tau_i$. Par suite, on a :

$\sigma = (\sigma(n+1) \ n+1) \prod_{i=1}^k \tau_i$. La permutation σ est ainsi écrite comme produit de $k+1$ transpositions (et $k+1 \leq n+1$).

Synthèse : dans les deux cas, on a montré que $\sigma \in S_{n+1}$ peut s'écrire comme produit de K transpositions, avec $K \leq n+1$. Ce qui assure que $\mathcal{P}(n+1)$ est vraie, et achève la preuve de l'hérédité.

7. La permutation ρ est le produit de σ , et de la transposition échangeant $\sigma(n+1)$ et $n+1$.

Conclusion. Pour tout entier naturel $n \geq 2$, et pour toute permutation $\sigma \in S_n$ il existe un entier $k \leq n$ et k transpositions $\tau_1, \dots, \tau_k$ de S_n tels que : $\sigma = \prod_{i=1}^k \tau_i$. En d'autres termes, les transpositions engendrent le groupe symétrique. ☕

Remarque : le théorème est encore valable pour $n = 1$, puisque tout élément de S_1 (c'est à dire la seule identité) s'écrit comme produit de zéro transposition...

THÉORÈME 15.6 - Toute permutation peut s'écrire comme un produit de cycles à supports disjoints.

PREUVE. Preuve par récurrence sur $n \in \mathbb{N}$, $n \geq 2$. On pose $P(n)$: “Dans S_n , toute permutation de S_n s'écrit comme produit de cycles à supports disjoints”.

► **Initialisation (pour $n = 2$).** Les éléments de S_2 sont $\text{id}_{\mathbb{N}_2}$ (produit de zéro cycle) et (12) (produit d'un cycle).

► **Hérédité.** Supposons $P(n)$ vraie pour un entier naturel $n \geq 2$. Soit $\sigma \in S_{n+1}$; on distingue deux cas. Premier cas : si $\sigma(n+1) = n+1$. Alors $\sigma|_{\mathbb{N}_n}$ est une permutation de S_n . Par hypothèse de récurrence, il

existe k cycles $c_1, \dots, c_k$ à supports disjoints tels que : $\sigma|_{\mathbb{N}_n} = \prod_{i=1}^k c_i$. Or ces k -cycles c_i peuvent être vus

comme des éléments de S_{n+1} , et puisque σ laisse $(n+1)$ invariant, l'égalité $\sigma = \prod_{i=1}^k c_i$ est valide dans S_{n+1} .

Second cas : si $\sigma(n+1) \neq n+1$. On considère alors la transposition $\tau = ((n+1), \sigma(n+1))$. Alors $\tau\sigma$ est un élément de S_{n+1} tel que : $\tau\sigma(n+1) = n+1$.

D'après l'étude faite dans le premier cas, il existe k -cycles $c_1, \dots, c_k$ de S_{n+1} tels que : $\tau\sigma = \prod_{i=1}^k c_i$.

Par suite⁸ : $\sigma = \tau c_1 \cdots c_k$.

Si $\sigma(n+1) \notin \bigcup_{i=1}^k \text{supp}(c_i)$, alors $\tau, c_1, \dots, c_k$ sont à supports disjoints, et c'est gagné.

Sinon : $\exists ! j \in [1, k]$, $c_j = (\sigma(n+1), x_1, \dots, x_p)$.

L'entier x_p est alors l'unique antécédent de $n+1$ par σ .

On pose alors : $\widehat{c}_j = (\sigma(n+1), x_1, \dots, x_p, n+1)$.

Les cycles $c_1, \dots, \widehat{c}_j, \dots, c_k$ sont à supports disjoints, et par construction : $\sigma = c_1 \cdots \widehat{c}_j \cdots c_k$.

On en déduit que dans les deux cas, tout élément de S_{n+1} est produit de cycles à supports disjoints, d'où $P(n+1)$ est vraie.

Conclusion. Pour tout entier $n \geq 2$, tout élément de S_n est produit de cycles à supports disjoints. ☕

8. En multipliant à gauche par τ les deux termes de l'égalité précédente.

15.4.3 Signature et groupe alterné

DÉFINITION 15.9 - Soit $n \geq 2$ un entier, et $(i, j) \in \mathbb{N}_n^2$ avec $i < j$. Une permutation $\sigma \in S_n$ réalise une **inversion** du couple (i, j) si $\sigma(i) > \sigma(j)$.

Notation. On note $I(\sigma)$ le nombre de couples sur lesquels σ réalise une inversion.

Exemples. $I(\text{id}) = 0$.

Dans S_3 , considérons la transposition $\tau(12)$. On a : $\tau(1) > \tau(2)$ (une inversion) ; $\tau(1) < \tau(3)$ (pas d'inversion) ; $\tau(2) < \tau(3)$ (pas d'inversion) ; par suite $I((12)) = 1$.

Toujours dans S_3 , considérons à présent la transposition $\tau(13)$. On a : $\tau(1) > \tau(2)$ (une inversion) ; $\tau(1) > \tau(3)$ (une inversion) ; $\tau(2) > \tau(3)$ (une inversion) ; par suite $I((13)) = 3$.

DÉFINITION 15.10 - Soit $\sigma \in S_n$. La **signature** de σ est le réel $\varepsilon(\sigma) = (-1)^{I(\sigma)}$

DÉFINITION 15.11 - Soit $\sigma \in S_n$. Il résulte de la définition précédente que la signature de σ vaut 1 ou (-1) . La permutation σ sera dite **paire** (*resp.* **impaire**) lorsque $\varepsilon(\sigma) = 1$ (*resp.* $\varepsilon(\sigma) = -1$)

THÉORÈME 15.7 - $\forall (\sigma, \tau) \in S_n^2, \varepsilon(\sigma\tau) = \varepsilon(\sigma)\varepsilon(\tau)$

PREUVE. Soient σ et τ deux éléments de S_n .

Soit (i, j) un couple d'entiers de $\mathbb{N}_n$ avec : $1 \leq i < j \leq n$.

La permutation $\sigma\tau$ réalise une inversion si et seulement si :

- ➡ τ réalise une inversion sur le couple (i, j) , et σ ne réalise pas d'inversion sur le couple $(\tau(i), \tau(j))$;
- ➡ τ ne réalise pas d'inversion sur le couple (i, j) , et σ réalise une inversion sur le couple $(\tau(i), \tau(j))$

En sommant $I(\sigma)$ et $I(\tau)$, on compte les inversions décrites ci-dessus, ainsi que deux fois les cas où τ réalise une inversion sur le couple (i, j) et σ réalise une inversion sur le couple $(\tau(i), \tau(j))$.

Il s'ensuit que $I(\sigma\tau)$ et $I(\sigma) + I(\tau)$ ont même parité. Donc :

$$(-1)^{I(\sigma\tau)} = (-1)^{I(\sigma)+I(\tau)} \iff (-1)^{I(\sigma\tau)} = (-1)^{I(\sigma)}(-1)^{I(\tau)} \iff \boxed{\varepsilon(\sigma\tau) = \varepsilon(\sigma)\varepsilon(\tau)} \quad \text{☕}$$

Remarque. Si l'on n'avait pas peur d'utiliser des gros mots (en tout cas, des mots hors programme de Sup), on pourrait énoncer le théorème ci-dessus en affirmant que le théorème ci-dessus signifie que l'application $\varepsilon : (S_n, \circ) \longrightarrow (\{\pm 1\}, \times)$ est un morphisme de groupes, dans le sens où elle est compatible avec les lois de groupes au départ et à l'arrivée ; l'image par ε de $\sigma \circ \tau$ est égal au produit des images $\varepsilon(\sigma) \times \varepsilon(\tau)$.

PROPRIÉTÉ 15.8 - La signature d'une transposition vaut -1 .

PREUVE. Soient n un entier supérieur ou égal à 2, et τ une transposition dans S_n . Par définition de transposition, il existe deux entiers i et j tels que $1 \leq i < j \leq n$ et $\tau = (ij)$. On peut représenter τ de la manière suivante :

$$\begin{pmatrix} 1 & 2 & \cdots & i & \cdots & j & \cdots & n \\ 1 & 2 & \cdots & j & \cdots & i & \cdots & n \end{pmatrix}$$

On compte alors le nombre d'inversions réalisées sur les couples (k, m) où k allant de 1 à $n-1$ et $m > k$:

$$I(\tau) = \underbrace{0}_{k=1} + \cdots + \underbrace{0}_{k=i-1} + \underbrace{j-i}_{k=i} + \underbrace{1}_{k=i+1} + \cdots + \underbrace{1}_{k=j-1} + \underbrace{0}_{k=j} + \cdots + \underbrace{0}_{k=n}$$

D'où : $I(\tau) = j-i + j-i-1 = 2(j-i) - 1$. Puisque $I(\tau)$ est impair, on a $(-1)^{I(\tau)} = -1$, c'à d $\varepsilon(\tau) = -1$.

Conclusion. La signature d'une transposition est égale à -1 . ☕

Et puisque tout p -cycle est en particulier un produit de $p-1$ transpositions, on peut affirmer que :

COROLLAIRE 15.6 - La signature d'un p -cycle vaut $(-1)^{p-1}$.

PREUVE. Soit $C = (a_1 a_2 \cdots a_p)$ un p -cycle. D'après le théorème 15.4, C est produit de $p-1$ transpositions ; il existe $\tau_1, \dots, \tau_{p-1}$ tels que $C = \prod_{i=1}^{p-1} \tau_i$.

Donc, d'après le théorème 15.7⁹ : $\varepsilon(C) = \prod_{i=1}^{p-1} \varepsilon(\tau_i)$

D'où, d'après la propriété 15.8 $\varepsilon(C) = \prod_{i=1}^{p-1} (-1)$ soit : $\varepsilon((a_1 a_2 \cdots a_p)) = (-1)^{p-1}$



Enfin :

COROLLAIRE 15.7 - Pour toute permutation $\sigma \in S_n$, on a : $\varepsilon(\sigma^{-1}) = \varepsilon(\sigma)$.

PREUVE. Pour toute permutation σ de S_n , $\sigma^{-1}\sigma = \text{id}_{\mathbb{N}_n}$. D'où : $\varepsilon(\sigma^{-1}\sigma) = \varepsilon(\text{id}_{\mathbb{N}_n})$. Par suite : $\varepsilon(\sigma^{-1})\varepsilon(\sigma) = 1$. On déduit de cette égalité que $\varepsilon(\sigma)$ et $\varepsilon(\sigma^{-1})$ sont simultanément égaux à 1 ou simultanément égaux à -1 .

En tous les cas : $\varepsilon(\sigma^{-1}) = \varepsilon(\sigma)$. ☕

La conséquence la plus importante de ce dernier corollaire étant l'énoncé ci-dessous.

PROPRIÉTÉ 15.9 - (Propriété - Définition). On note A_n l'ensemble des permutations paires de S_n .

$(A_n, \circ)$ est un sous-groupe de $(S_n, \circ)$, appelé **groupe alterné**.

C'est en particulier un groupe fini, et : $\text{Card}(A_n) = \frac{n!}{2}$.

PREUVE. ➡ **(SG1)** : par définition même de A_n (c'est l'ensemble des permutations paires), on a

$$A_n \subset S_n \quad (\spadesuit)$$

➡ **(SG2)** : $\text{id}_{\mathbb{N}_n}$, qui est l'élément neutre de $(S_n, \circ)$, est une permutation paire puisque $\varepsilon(\text{id}_{\mathbb{N}_n}) = 1$. Ainsi :

$$\text{id}_{\mathbb{N}_n} \in A_n \quad (\heartsuit)$$

9. Plus précisément, le théorème 15.7 permet d'établir par une récurrence facile sur n que pour une famille de n permutations $(\sigma_i)_{i \in \mathbb{N}_n}$, on a : $\varepsilon\left(\prod_{i=1}^n \sigma_i\right) = \prod_{i=1}^n \varepsilon(\sigma_i)$.

➡ **(SG3)** : soient σ et τ deux permutations paires. Alors $\varepsilon(\sigma\tau) \underbrace{=}_{\text{th. 15.7}} \underbrace{\varepsilon(\sigma)}_{=1 \text{ car } \sigma \in A_n} \underbrace{\varepsilon(\tau)}_{=1 \text{ car } \tau \in A_n} = 1$. D'où :

$\sigma\tau \in A_n$.

Par conséquent : $\boxed{\forall (\sigma, \tau) \in A_n^2, \sigma\tau \in A_n \ (\diamond)}$

➡ **(SG4)** : soit $\sigma \in A_n$. Alors $\varepsilon(\sigma) = 1$, donc $\varepsilon(\sigma^{-1}) = 1$ en vertu du corollaire 15.7. Par suite : $\sigma^{-1} \in A_n$.

D'où : $\boxed{\forall \sigma \in A_n, \sigma^{-1} \in A_n \ (\clubsuit)}$

On déduit de $(\spadesuit)$, $(\heartsuit)$, $(\diamond)$ et $(\clubsuit)$ que $\boxed{(A_n, \circ) \text{ est un sous-groupe de } (S_n, \circ)}$.

Assertion concernant le cardinal : par définition de A_n (et de S_n), l'ensemble S_n est la réunion disjointe de A_n et de $S_n \setminus A_n$. D'où : $\text{Card}(S_n) = \text{Card}(A_n) + \text{Card}(S_n \setminus A_n)$, ainsi : $\boxed{\text{Card}(A_n) + \text{Card}(S_n \setminus A_n) = n! \ (\spadesuit)}$.

D'autre part, l'application : $F : \sigma \in A_n \mapsto (12)\sigma \in S_n \setminus A_n$ et l'application : $G : \sigma \in S_n \setminus A_n \mapsto (12)\sigma \in A_n$ sont clairement réciproques l'une de l'autre. En particulier, elles sont bijectives ce qui implique : $\boxed{\text{Card}(A_n) = \text{Card}(S_n \setminus A_n) \ (\clubsuit)}$.

On déduit de $(\spadesuit)$ et de $(\clubsuit)$ que : $\boxed{\text{Card}(A_n) = \frac{n!}{2}}$. ☕

Remarque : une preuve (hors-programme de Sup, pas de MP) de la propriété précédente : A_n est l'ensemble des antécédents de l'élément neutre (1) par le morphisme de groupes $\varepsilon : (S_n, \circ) \longrightarrow (\{\pm 1\}, \times)$. C'est (par définition) le noyau de ce morphisme. Or le noyau d'un morphisme de groupes est toujours un sous-groupe du groupe de départ. D'où la propriété.

15.5 Epilogue

15.5.1 Description de A_6

L'objet de ce paragraphe est de s'adonner à un petit exercice de dénombrement, et de décrire les éléments du groupe alterné A_6 , c'est à dire de lister les permutations paires de S_6 .

On sait déjà que : $\text{card}(A_6) = \frac{6!}{2}$, soit : $\text{card}(A_6) = 360$.

Puisque on sait encore que $\text{id}_{\mathbb{N}_6} \in A_6$, il nous reste 359 éléments de A_6 à préciser...

Pour ce faire, on utilise le fait que toute permutation est produit de cycles à supports disjoints, et que la signature d'un p -cycle est égale à $(-1)^{p-1}$.

Il s'ensuit qu'un élément de A_6 est :

- soit l'identité ;
- soit un 3-cycle ;
- soit un 5-cycle ;
- soit le produit de 2 transpositions à supports disjoints ;
- soit le produit d'une transposition et d'un 4-cycle à supports disjoints ;
- soit le produit de deux 3-cycles à supports disjoints.

Comptons les éléments de chaque famille.

→ Il n'y a qu'une seule identité :-)

→ Pour construire un 3-cycle, on commence par choisir 3 entiers $a_1, \dots, a_3$ de $\mathbb{N}_6$: il y a $\binom{6}{3} = 20$ manières de le faire. Le support du cycle choisi, on peut toujours convenir d'écrire le plus petit entier a_i en premier ; ceci fait, il reste 2 choix possibles pour le second entier du cycle, et le troisième est alors connu.

On en déduit qu'il existe $\boxed{40 \text{ 3-cycles dans } A_6}$.

→ Pour construire un 5-cycle, on commence par choisir 5 entiers $a_1, \dots, a_5$ de $\mathbb{N}_6$: il y a $\binom{6}{5} = 6$ manières de le faire. Le support du cycle choisi, on peut toujours convenir d'écrire le plus petit entier a_i en premier ; ceci fait, il reste 4 choix possibles pour le second entier du cycle, 3 pour le troisième, 2 pour le quatrième, et... 1 pour le dernier.

On en déduit qu'il existe $\boxed{144 \text{ 5-cycles dans } A_6}$.

→ Pour construire un produit de deux transpositions à supports disjoints, on commence par choisir quatre entiers $\{a_1, a_2, a_3, a_4\}$. Il y a $\binom{6}{4} = 15$ manières de le faire. Ce choix fait, on peut toujours convenir d'écrire le produit sous la forme $(a_1 \star)(\star \star)$. Le produit est alors uniquement déterminé par l'image de a_1 , que l'on peut choisir parmi a_2, a_3 et a_4 .

Finalement, il existe $\boxed{3 \times 15 = 45 \text{ produits de deux transpositions à supports disjoints dans } A_6}$.

→ Pour construire un produit d'une transposition et d'un 4-cycle à supports disjoints, on commence par choisir le support de la transposition : $\{a_1, a_2\}$ (le support du second est alors connu !). Il y a $\binom{6}{2} = 15$ manières de le faire. Ce choix fait, la transposition est connue. Reste à construire le 4-cycle : on peut toujours convenir d'écrire dans celui-ci le plus petit entier en premier. Ceci fait, il reste 3 choix pour le second, 2 choix pour le troisième (et 1 pour le dernier).

Finalement, il existe $\boxed{6 \times 15 = 90 \text{ produits d'une transposition et d'un 4-cycle à supports disjoints dans } A_6}$.

→ Pour construire un produit de deux 3-cycles à supports disjoints, on commence par choisir le support du premier : $\{a_1, a_2, a_3\}$ (le support du second est alors connu !). Il y a $\binom{6}{3} = 20$ manières de le faire. Ce choix fait, on peut toujours convenir d'écrire dans chacun des 2 cycles le plus petit entier a_i en premier ; il reste 2 choix possibles pour le second entier de chaque cycle (et 1 pour le second). Pour un support donné, on a donc a priori 4 produits de deux 3-cycles distincts. Mais il faut tenir compte du fait que le produit $(a_1 a_2 a_3)(a_4 a_5 a_6)$ sera alors compté deux fois : une fois lorsque le support du premier cycle choisi est $\{a_1, a_2, a_3\}$, et une seconde lorsque le support choisi est $\{a_4, a_5, a_6\}$.

Finalement, il existe $\boxed{\frac{4 \times \binom{6}{3}}{2} = 40 \text{ produits de deux 3-cycles à supports disjoints dans } A_6}$.

Conclusion. Outre l'identité, le groupe A_6 est constitué de :

- 40 3-cycles ;
- 144 5-cycles ;
- 45 produits de 2 transpositions à supports disjoints ;
- 90 produits d'une transposition et d'un 4-cycle à supports disjoints ;
- 40 produits de deux 3-cycles à supports disjoints.

15.5.2 Simplicité de A_5

L'énoncé présenté ci-dessous est constitué des deux dernières parties d'un problème donné en MPSI en 2016. Le début du problème était constitué de questions destinées à établir le théorème de Lagrange sur les groupes finis, et le fait que tout groupe de cardinal premier est cyclique donc abélien (voir sections 11.5.2 et 11.5.3).

► **PARTIE E - Conjugaison dans un groupe.**

Soit n un entier supérieur ou égal à 2. On dit que deux permutations τ_1 et τ_2 de S_n sont **conjuguées** s'il existe une permutation σ de S_n telle que : $\tau_2 = \sigma\tau_1\sigma^{-1}$.

- 13) Montrer que la relation de conjugaison définie ci-dessus est une relation d'équivalence sur S_n .
- 14) Un sous-groupe H de S_n est dit **distingué** s'il est stable par conjugaison dans le sens suivant :

$$\forall h \in H, \forall \sigma \in S_n \quad \sigma h \sigma^{-1} \in H$$

- a) Etablir que $\langle (12) \rangle$ n'est pas un sous-groupe distingué de S_3 .
- b) Montrer que (pour tout entier $n \geq 2$) A_n est un sous-groupe distingué de S_n .

► **PARTIE F - Simplicité de A_5 .**

- 15) Quel est le cardinal de A_5 ?
- 16) Montrer que tout élément de A_5 distinct de l'identité est soit un 3-cycle, soit un 5-cycle, soit une composée de deux transpositions à supports disjoints.
- 17) Intermède : un peu de dénombrement.
- a) Combien existe-t-il de 5-cycles dans A_5 ?
- b) Combien existe-t-il de 3-cycles dans A_5 ?
- c) Combien existe-t-il de produits de deux transpositions à supports disjoints dans A_5 ?
- 18) Temporairement, on se replace dans le cas général de S_n , avec $n \geq 5$.

- a) Soient $a_1, \dots, a_{n-2}$ ($n-2$) éléments de $\llbracket 1, n \rrbracket$; on note a_{n-1} et a_n les deux éléments restants.

Et soient de même $b_1, \dots, b_{n-2}$ ($n-2$) éléments de $\llbracket 1, n \rrbracket$; on note b_{n-1} et b_n les deux éléments restants.

Montrer qu'il existe une permutation paire σ telle que :

$$\forall i \in \llbracket 1, n-2 \rrbracket, \sigma(a_i) = b_i$$

On pourra éventuellement utiliser une composition par une certaine transposition pour obtenir la bonne parité.

- b) Soit $\sigma \in S_n$. On suppose que $\tau_1 = (i_1 \cdots i_k)$ est un k -cycle (avec $k \in \llbracket 2, n \rrbracket$), et on pose $\tau_2 = \sigma \tau_1 \sigma^{-1}$.
Etablir que :

$$\tau_2 = (\sigma(i_1) \cdots \sigma(i_k))$$

- 19) Dédurre de la question précédente que les 3-cycles sont deux à deux conjugués dans A_5 , càd que si c_1 et c_2 sont deux 3-cycles, il existe $\sigma \in A_5$ telle que : $c_2 = \sigma c_1 \sigma^{-1}$.
- 20) Montrer de même que les composées de deux transpositions à supports disjoints sont conjuguées dans A_5 .
- 21) Un sous-groupe H de A_5 est dit **distingué** si

$$\forall h \in H, \forall \sigma \in A_5, \sigma h \sigma^{-1} \in H$$

Soit H un sous-groupe distingué (donc stable par conjugaison) de A_5 .

- a) Montrer que si H contient un 3-cycle, alors il les contient tous. De même, montrer que si H contient un produit de deux transpositions à supports disjoints, alors il contient tous les produits de deux transpositions à supports disjoints.

On pourra admettre par la suite qu'il en va de même pour les 5-cycles : si H contient un 5-cycle, alors il les contient tous.

- b) Etablir que $H = \{\text{id}\}$ ou $H = A_5$.

Commentaire. Cet ultime résultat signifie que le groupe A_5 ne possède pas de sous-groupe distingué non trivial (c'est-à-dire distinct de lui-même et de $\{\text{id}\}$) ; un tel groupe est appelé un groupe **simple**. Comme indiqué dans le préambule du problème, c'est cette simplicité qui implique qu'une équation polynomiale de degré 5 n'est pas résoluble par radicaux, càd qu'il n'existe pas (contrairement au cas du second degré) de formule "avec des racines" donnant les solutions d'une équation de degré 5 en général. Ce résultat peut être généralisé (Galois l'a fait) à $n \geq 5$; en effet, pour tout entier $n \geq 5$, le groupe A_n est également simple.

CORRIGÉ.

► **PARTIE E - Conjugaison dans un groupe.**

Soit n un entier supérieur ou égal à 2. On dit que deux permutations τ_1 et τ_2 de S_n sont **conjuguées** s'il existe une permutation σ de S_n telle que : $\tau_2 = \sigma\tau_1\sigma^{-1}$.

13) Soit $\tau \in S_n$. Alors : $\tau = \text{id}\tau\text{id}^{-1}$. Donc τ est conjuguée à elle-même, ce qui prouve la **réflexivité** de la relation de conjugaison.

Soient τ_1 et τ_2 deux permutations (dans S_n) conjuguées. Par définition, il existe une permutation σ de S_n telle que : $\tau_2 = \sigma\tau_1\sigma^{-1}$. On a donc : $\tau_1 = \sigma^{-1}\tau_2\sigma$ d'où $\tau_1 = \sigma'\tau_2\sigma'^{-1}$ (en ayant posé $\sigma' = \sigma^{-1}$). Il s'ensuit que la relation de conjugaison est **symétrique**.

Soient enfin τ_1 , τ_2 et τ_3 trois permutations (dans S_n) telles que τ_1 et τ_2 d'une part, τ_2 et τ_3 d'autre part sont conjuguées. Par définition, il existe deux permutations σ et ρ de S_n telles que : $\tau_2 = \sigma\tau_1\sigma^{-1}$ et $\tau_3 = \rho\tau_2\rho^{-1}$. On en déduit que :

$$\tau_3 = \rho\sigma\tau_1\sigma^{-1}\rho^{-1} = \rho\sigma\tau_1(\rho\sigma)^{-1} \quad \text{d'où : } \tau_3 = \zeta\tau_1\zeta^{-1} \quad (\text{en ayant posé : } \zeta = \rho\sigma).$$

On en déduit que τ_1 et τ_3 sont conjuguées ; ce qui établit la **transitivité** de la relation de conjugaison.

Conclusion. La relation de conjugaison dans S_n est réflexive, symétrique et transitive. C'est donc une relation d'équivalence sur S_n .

14) a) On a : $(13)(12)(13)^{-1} = (13)(12)(13) = (23)$. Or : $(23) \notin \langle (12) \rangle$.

Il s'ensuit que : $\langle (12) \rangle$ n'est pas un sous-groupe distingué de S_3 .

b) Soit τ un élément arbitraire de A_n . Pour tout élément σ de S_n , on a : $\varepsilon(\sigma\tau\sigma^{-1}) = \varepsilon(\sigma)\varepsilon(\tau)\varepsilon(\sigma^{-1})$.

Puisque $\varepsilon(\sigma^{-1}) = \varepsilon(\sigma)$, on en déduit que : $\varepsilon(\sigma\tau\sigma^{-1}) = \varepsilon(\tau)$ d'où : $\varepsilon(\sigma\tau\sigma^{-1}) = 1$ puisque τ est paire par hypothèse.

On en déduit que pour toute permutation σ de S_n , la permutation $\sigma\tau\sigma^{-1}$ est encore un élément de A_n . Le sous-groupe A_n est donc stable par conjugaison. Ainsi, A_n est un sous-groupe distingué de S_n .

► **PARTIE F - Simplicité de A_5 .**

15) D'après le cours : $\text{Card}(A_5) = \frac{5!}{2}$ d'où : $\text{Card}(A_5) = 60$.

16) Une première idée. Soit $\sigma \in A_n$, avec $\sigma \neq \text{id}$. On sait que σ s'écrit comme un produit de 2 ou 4 transpositions. En effet, tout élément de S_5 s'écrit comme un produit d'au plus 5 transpositions (résultat général du cours) ; pour une permutation paire de S_5 , le nombre de transpositions intervenant dans ce produit peut donc être égal à 0, 2 ou 4. Mais la valeur 0 est exclue, puisque l'on suppose que $\sigma \neq \text{id}$.

Supposons que σ soit produit de deux transpositions : alors d'après la discussion faite dans la question 2-a, σ est soit égale au produit de deux transpositions à supports disjoints (1er cas), soit égale à un 3-cycle (2ème cas).¹⁰ Jusque-là tout va bien.

Supposons que σ soit produit de 4 transpositions : aïe ! Même avec la discussion faite dans la question 2-a, et même en décidant par exemple de regrouper deux par deux les transpositions, le nombre de cas à étudier suivant les supports des transpositions est ici très dissuasif !

Une seconde idée. Une autre stratégie consiste à observer que toute permutation peut s'écrire comme un produit de cycles à supports disjoints ; c'est en tout cas clair sur les exemples vus précédemment dans ce devoir (et en exercices).

10. Le 3ème cas est exclu puisque $\sigma \neq \text{id}$.

On pose donc pour tout entier naturel $n \geq 2$:

$P(n)$: “Toute permutation de S_n peut s’écrire comme un produit de cycles à supports disjoints.”

Prouvons $P(n)$ par récurrence sur n .

► **Initialisation.** Pour $n = 2$, le groupe S_2 est constitué de id (produit de 0 cycle), et de (12) (produit d’un cycle). La propriété $P(2)$ est donc vraie.

► **Hérédité.** Supposons $P(n)$ vraie pour un certain entier $n \geq 2$, et montrons que $P(n+1)$ l’est.

Soit $\sigma \in S_{n+1}$. On distingue deux cas :

☞ Si $\sigma(n+1) = n+1$. Alors $\sigma|_{\mathbb{N}_n}$ est un élément de S_n . D’après l’hypothèse de récurrence, il existe alors k cycles à supports disjoints $c_1, \dots, c_k$ (dans S_n) tels que : $\sigma|_{\mathbb{N}_n} = \prod_{i=1}^k c_i$. Or un cycle de S_n induit naturellement un cycle de S_{n+1} , et la relation précédente implique que σ est le produit des k cycles c_i vus comme éléments de S_{n+1} .

☞ Si $\sigma(n+1) \neq n+1$. Il existe alors un entier naturel non nul p tel que les entiers $n+1, \sigma(n+1), \dots, \sigma^{p-1}(n+1)$ sont distincts, et $\sigma^p(n+1) = n+1$. On définit alors un cycle c en posant :

$$c = ((n+1) \sigma(n+1) \cdots \sigma^{p-1}(n+1))$$

Par construction, la permutation $c^{-1}\sigma$ laisse invariants les entiers : $n+1, \sigma(n+1), \dots, \sigma^{p-1}(n+1)$.

En particulier, puisque $c^{-1}\sigma$ laisse invariant $n+1$, on peut (d’après l’étude faite dans le premier cas) affirmer que $c^{-1}\sigma = \prod_{i=1}^k c_i$ où les c_i sont k cycles à supports disjoints. En outre, les supports des c_i sont disjoints de $\{n+1, \sigma(n+1), \dots, \sigma^{p-1}(n+1)\}$ puisque le produit des c_i laisse invariants ces éléments.

Par suite : $\sigma = c \prod_{i=1}^k c_i$, et on a écrit σ comme produit de cycles à supports disjoints.

Dans les deux cas, on a montré que $P(n+1)$ est vraie.

► **Conclusion.** Toute permutation de S_n peut s’écrire comme un produit de cycles à supports disjoints.

En particulier, tout élément de A_5 peut s’écrire comme un produit de cycles à supports disjoints. Mais le nombre de décompositions est alors limité ! En effet, pour des raisons de parité, un élément de A_5 ne pourra être que l’identité, ou un produit de deux transpositions à supports disjoints, ou un 3-cycle ou un 5-cycle.

Conclusion. Tout élément de A_5 distinct de l’identité est soit un 3-cycle, soit un 5-cycle, soit un produit de deux transpositions à supports disjoints.

17) a) Pour définir un 5-cycle $(a_1 a_2 a_3 a_4 a_5)$ dans A_5 , on peut toujours convenir de commencer par $a_1 = 1$. En d’autres termes, tout 5-cycle admet une écriture de la forme : $(1 a_2 a_3 a_4 a_5)$. Pour créer un tel cycle, on a 4 choix pour a_2 ; 3 choix pour a_3 ; 2 choix pour a_4 et plus de choix du tout pour a_5 .

Conclusion. Il existe 24 5-cycles dans A_5 .

b) Pour définir un 3-cycle, on commence par choisir trois éléments de $\llbracket 1, 5 \rrbracket$. Il y a $\binom{5}{3} = 10$ façons de le faire. On observe alors que chaque combinaison $\{a_1, a_2, a_3\}$ donne naissance à exactement deux 3-cycles : $(a_1 a_2 a_3)$ et $(a_1 a_3 a_2)$.

Conclusion. Il existe 20 3-cycles dans A_5 .

c) D'après la question 16, le groupe A_5 est constitué de l'identité, des 5-cycles, des 3-cycles et des produits de deux transpositions à supports disjoints dans A_5 . On en déduit, d'après les questions 15, 17-a et 17-b qu'il existe :

$60 - (1 - 24 - 20) = 15$ produits de deux transpositions à supports disjoints dans A_5 .

18) a) Soient $a_1, \dots, a_{n-2}$ (n-2) éléments de $\llbracket 1, n \rrbracket$; on note a_{n-1} et a_n les deux éléments restants. Et soient de même $b_1, \dots, b_{n-2}$ (n-2) éléments de $\llbracket 1, n \rrbracket$; on note b_{n-1} et b_n les deux éléments restants.

On définit un élément τ de S_n en posant : $\forall i \in \llbracket 1, n \rrbracket, \tau(a_i) = b_i$.

Si τ est paire, on pose $\sigma = \tau$. Sinon, on pose $\sigma = (a_{n-1}a_n)\tau$.

Dans les deux cas, σ est paire et vérifie : $\forall i \in \llbracket 1, n-2 \rrbracket, \sigma(a_i) = b_i$.

b) Etablir l'égalité $\tau_2 = (\sigma(i_1) \cdots \sigma(i_k))$, c'est établir que pour tout $\sigma(i_j)$ on a $\tau_2(\sigma(i_j)) = \sigma(i_{j+1})$ (avec la convention que $\sigma(i_{k+1}) = i_1$); et que tout les autres entiers de $\mathbb{N}_n$ ne sont pas dans le support de τ_2 .

On sépare donc les entiers de $\mathbb{N}_n$ en deux parties disjointes :

ceux qui appartiennent à l'ensemble $\{\sigma(i_1), \dots, \sigma(i_k)\}$ et les autres.

► Soit m un entier dans $\{\sigma(i_1), \dots, \sigma(i_k)\}$. Alors il existe un entier j compris entre 1 et k tel que : $m = \sigma(i_j)$.

D'une part on a alors : $\tau_2(m) = \sigma\tau_1\sigma^{-1}(\sigma(i_j)) = \sigma\tau_1(i_j) = \sigma(i_{j+1})$

Et d'autre part : $(\sigma(i_1) \cdots \sigma(i_k))(m) = (\sigma(i_1) \cdots \sigma(i_k))(\sigma(i_j)) = \sigma(i_{j+1})$

On en déduit que pour tout m un entier dans $\{\sigma(i_1), \dots, \sigma(i_k)\}$ on a : $\tau_2(m) = (\sigma(i_1) \cdots \sigma(i_k))(m)$.

► Soit m un entier dans $\mathbb{N}_n \setminus \{\sigma(i_1), \dots, \sigma(i_k)\}$. Alors il est déjà clair que : $(\sigma(i_1) \cdots \sigma(i_k))(m) = m$.

Par ailleurs, l'hypothèse faite sur m impose que $\sigma^{-1}(m) \notin \{i_1, \dots, i_k\}$. En effet, dans le cas contraire, m serait égal à un $\sigma(i_j)$ pour un certain entier j , ce qui contredirait notre hypothèse. Il s'ensuit que $\sigma^{-1}(m)$ est laissé invariant par τ_1 , d'où :

$$\tau_2(m) = \sigma\tau_1\sigma^{-1}(m) = \sigma\tau_1(\sigma^{-1}(m)) = \sigma(\sigma^{-1}(m)) = m$$

On en déduit que pour tout m un entier dans $\mathbb{N}_n \setminus \{\sigma(i_1), \dots, \sigma(i_k)\}$ on a : $\tau_2(m) = (\sigma(i_1) \cdots \sigma(i_k))(m)$.

► En conclusion, on a établi que : $\forall m \in \mathbb{N}_n, \tau_2(m) = (\sigma(i_1) \cdots \sigma(i_k))(m)$. D'où : $\tau_2 = (\sigma(i_1) \cdots \sigma(i_k))$

19) Soient c_1 et c_2 deux 3-cycles de A_5 . Il existe trois entiers a_1, a_2 et a_3 (resp. b_1, b_2 et b_3) dans $\mathbb{N}_5$ tels que $c_1 = (a_1a_2a_3)$ (resp. $c_2 = (b_1b_2b_3)$). D'après la question 18-a, il existe une permutation paire $\sigma \in A_5$ telle que : $\forall i \in \llbracket 1, 3 \rrbracket, \sigma(a_i) = b_i$. D'après la question 18-b, on a alors $c_2 = \sigma c_1 \sigma^{-1}$ ce qui signifie que c_1 et c_2 sont conjugués.

Conclusion. Les 3-cycles sont deux à deux conjugués dans A_5 .

20) Soient $\sigma_1 = (i_1j_1)(k_1l_1)$ et $\sigma_2 = (i_2j_2)(k_2l_2)$ deux produits de transpositions à supports disjoints. On note m_1 (resp. m_2) l'unique entier de $\mathbb{N}_5$ tel que $\mathbb{N}_5 = \{i_1, j_1, k_1, l_1, m_1\}$ (resp. ...). On définit une permutation α dans S_5 en posant : $\alpha(i_1) = i_2, \alpha(j_1) = j_2, \dots$

Si α est paire, on pose $\rho = \alpha$ et on a alors : $\rho\sigma_1\rho^{-1} = (\rho(i_1)\rho(j_1))(\rho(k_1)\rho(l_1)) = (i_2j_2)(k_2l_2) = \sigma_2$

Si α est impaire, on pose $\rho = \alpha(i_1j_1)$ et on a alors : $\rho\sigma_1\rho^{-1} = (\rho(j_1)\rho(i_1))(\rho(k_1)\rho(l_1)) = (j_2i_2)(k_2l_2) = \sigma_2$

Dans les deux cas, on a établi l'existence d'une permutation $\rho \in A_5$ telle que : $\rho\sigma_1\rho^{-1} = \sigma_2$.

Conclusion. Les produits de 2 transpositions à supports disjoints sont deux à deux conjugués dans A_5 .

21) a) Soit H un sous-groupe distingué de A_5 . Si H contient un 3-cycle c , alors H contient tout élément conjugué à c (puisque H est stable par conjugaison). Or d'après la question 19, les 3-cycles sont deux à deux conjugués dans A_5 ; donc H contient tous les 3-cycles.

Le même raisonnement s'applique aux permutations qui sont produits de deux transpositions à supports disjoints, en vertu de la question 20 cette fois-ci.

Conclusion. Lorsque H est un sous-groupe distingué (càd stable par conjugaison) de H_5 , il contient tous les 3-cycles dès lors qu'il en contient un. De même, si H contient un produit de deux transpositions à supports disjoints, alors il contient tous les produits de deux transpositions à supports disjoints.

b) Il est clair que $\{\text{id}\}$ et A_5 sont des sous-groupes distingués de A_5 . Le but est de prouver qu'il n'existe que ceux-là.

Soit H un sous-groupe distingué de A_5 . Si H n'est pas réduit à l'identité, alors H contient au moins un 3-cycle (ou un 5-cycle, ou un produit de deux transpositions à supports disjoints). Mais alors, d'après la question précédente, H contient tous les 3-cycles (ou tous les 5-cycles, ou...), qui sont au nombre de 20 (ou 24 ou 15).

Or H ne peut contenir (en plus de l'identité évidemment) que les 3-cycles, sinon il serait de cardinal 21; or son cardinal doit diviser celui de A_5 (égal à 60) d'après le théorème de Lagrange (question 8). Il doit donc contenir au moins un 5-cycle ou un produit de deux transpositions à supports disjoints; mais alors il contient toutes les permutations de ce type et on obtient donc : $\text{Card}(H) > 30$. Une nouvelle application du théorème de Lagrange implique que $\text{Card}(H) = 60$, càd $H = A_5$.

De même H ne peut contenir que les 5-cycles, sinon il serait de cardinal 25; et H ne peut contenir que les produits de deux transpositions à supports disjoints, sinon il serait de cardinal 16. Dans ces deux cas, on raisonne comme précédemment pour conclure finalement que $H = A_5$.

Conclusion. Le groupe A_5 ne possède aucun sous-groupe non trivial stable par conjugaison (les seuls sous-groupes distingués de A_5 sont $\{\text{id}\}$ et A_5).

15.5.3 Dénombrabilité de $\mathbb{Q}$

Numérotation des nombres rationnels

Un ensemble E est dit **dénombrable** lorsqu'il est équipotent à $\mathbb{N}$, c'est à dire s'il existe une bijection entre $\mathbb{N}$ et E ; cette bijection permet de "numéroter" les éléments de E , ce qui justifie a posteriori la terminologie. Vous avez par exemple déjà établi que $\mathbb{Z}$ est dénombrable (c'était une question de cours au programme de la semaine passée). Plus explicitement :

Propriété (*) — L'application $f : \mathbb{N} \longrightarrow \mathbb{Z}$ définie en posant :

$$\forall n \in \mathbb{N}, f(n) = \begin{cases} \frac{n}{2} & \text{si } n \text{ est pair} \\ -\frac{n+1}{2} & \text{si } n \text{ est impair} \end{cases}$$

est bijective.

Ce résultat pourra être utilisé au besoin au cours du problème, sans qu'il soit nécessaire de le redémontrer.

L'objectif principal de ce problème est d'aller plus loin, et de prouver que l'ensemble $\mathbb{Q}$ des nombres rationnels est dénombrable. Ceci passera par des étapes intermédiaires (notamment : établir que $\mathbb{N}^2$ est dénombrable) et l'utilisation du théorème de Cantor-Bernstein.

Partie I — Proche du cours

- 1) Notons $E_1 = \mathbb{N} \setminus \{0, 1\}$. Montrer que E_1 est dénombrable.
- 2) Notons E_2 l'ensemble des entiers naturels pairs, soit : $E = \{2k / k \in \mathbb{N}\}$. Montrer que E_2 est dénombrable.

Partie II — Dénombrabilité de $\mathbb{N}^2$, dénombrabilité de $\mathbb{Z}^2$

On définit l'application

$$\begin{aligned} \varphi : \mathbb{N}^2 &\longrightarrow \mathbb{N}^* \\ (p, q) &\longmapsto 2^p (2q + 1) \end{aligned}$$

- 3) Montrer que φ est bien définie et qu'elle est injective.
- 4) Soit $n \in \mathbb{N}^*$. Justifier qu'il existe un entier N_0 tel que 2^{N_0} divise n , et 2^{N_0+1} ne divise pas n . Quel est le nom de cet entier N_0 ?
- 5) En utilisant la question précédente, établir que φ est surjective.
- 6) Conclure que $\mathbb{N}^2$ est dénombrable ; puis que $\mathbb{Z}^2$ est dénombrable.

Partie III — Dénombrabilité de $\mathbb{Q}$

- 7) Construire une application injective ("simple") de $\mathbb{N}$ dans $\mathbb{Q}$.

- 8) On appelle **représentant irréductible** d'un nombre rationnel non nul r l'unique fraction irréductible p/q égale à r avec $p \in \mathbb{Z}$ et $q \in \mathbb{N}^*$.¹¹ Nous conviendrons que le représentant irréductible de 0 est 0/1.
- a) Justifier brièvement que l'application $\psi : \mathbb{Q} \longrightarrow \mathbb{Z} \times \mathbb{N}^*$ qui à $r \in \mathbb{Q}$ associe le couple $(p, q) \in \mathbb{Z} \times \mathbb{N}^*$ est injective. Est-elle surjective?
- b) Construire une application injective de $\mathbb{Q}$ dans $\mathbb{N}$.

CORRIGÉ.

- 1) Considérons les applications $f : n \in E_1 \longrightarrow (n-2) \in \mathbb{N}$ et $g : n \in \mathbb{N} \longrightarrow (n+2) \in E_1$. Il est clair que $f \circ g = \text{id}_{\mathbb{N}}$ et $g \circ f = \text{id}_{E_1}$. Donc f (tout comme g) est bijective, donc E_1 est dénombrable.
- 2) Considérons les applications $f : n \in E_2 \longrightarrow (n/2) \in \mathbb{N}$ et $g : n \in \mathbb{N} \longrightarrow (2n) \in E_2$. Il est clair que $f \circ g = \text{id}_{\mathbb{N}}$ et $g \circ f = \text{id}_{E_2}$. Donc f (tout comme g) est bijective, donc E_2 est dénombrable.

Partie II — Dénombrabilité de $\mathbb{N}^2$, dénombrabilité de $\mathbb{Z}^2$

- 3) Pour tout couple (p, q) d'entiers naturels, $2^p(2q+1)$ est un entier naturel non nul, donc φ est bien définie.

Soient (p, q) et (p', q') deux couples d'entiers naturels tels que $\varphi(p, q) = \varphi(p', q')$. Alors p et p' sont égaux, car tous deux sont égaux à la valuation 2-adique de $\varphi(p, q)$. On en déduit que $2q+1 = 2q'+1$ d'où évidemment $q = q'$.

Conclusion. φ est injective.

Remarque : sans utiliser la valuation 2-adique, on peut répondre à cette question comme suit. Si $2^p(2q+1) = 2^{p'}(2q'+1)$, alors $2^p | 2^{p'}(2q'+1)$. Si $p > 0$, cette relation entraîne que $2^p | 2^{p'}$ d'où $p \leq p'$. Dans "l'autre sens", on a : $2^{p'} | 2^p(2q+1)$ donc $2^{p'} | 2^p$ donc $p' \leq p$; finalement $p = p'$. Si $p = 0$, alors $2^{p'}(2q'+1) = 2q+1$ est impair, ce qui implique $p' = 0$, d'où encore une fois $p = p'$. Une fois cette égalité établie, celle entre q et q' en est une conséquence triviale.

- 4) Soit $n \in \mathbb{N}^*$. On pose $E = \{k \in \mathbb{N}, 2^k | n\}$. L'ensemble E est une partie de $\mathbb{N}$ qui est non vide ($0 \in E$) et majorée : en effet, 2^k tend vers $+\infty$ lorsque k tend vers $+\infty$, et il existe donc un entier K tel que : $(k \geq K) \implies (2^k > n)$. Par suite, E admet un plus grand élément, et on peut légitimement poser $N_0 = \max E$.

Par définition de plus grand élément, $N_0 \in E$ et $(N_0+1) \notin E$, c'est-à-dire : 2^{N_0} divise n et 2^{N_0+1} ne divise pas n . Cet entier N_0 est la valuation 2-adique de n (on note : $N_0 = v_2(n)$).

- 5) Soit $n \in \mathbb{N}^*$. D'après la question précédente, on peut écrire : $n = 2^{v_2(n)}m$, où m est impair. En effet, si m était pair, alors $2^{v_2(n)+1}$ diviserait n ce qui contredirait la définition de $v_2(n)$. Il existe donc un entier naturel q tel que : $m = 2q+1$ et on peut écrire : $n = 2^{v_2(n)}(2q+1)$, d'où : $n = \varphi(v_2(n), q)$. Il s'ensuit que n est dans l'image de φ , et puisque n est arbitraire dans le raisonnement précédent, on peut conclure que φ est surjective.

- 6) D'après les questions 3 et 5, $\varphi : \mathbb{N}^2 \longrightarrow \mathbb{N}^*$ est bijective. En outre, $g : n \in \mathbb{N}^* \longrightarrow (n-1) \in \mathbb{N}$ est bijective (question de cours d'une colle récente). La composée de deux bijections étant encore bijective, on en déduit que l'application $g \circ \varphi : \mathbb{N}^2 \longrightarrow \mathbb{N}$ est bijective. Conclusion. $\mathbb{N}^2$ est dénombrable.

¹¹. Par exemple, le représentant irréductible de 10/15 est 2/3; celui de -9/12 est (-3)/4.

Par ailleurs, d'après la propriété (*), il existe une bijection $f : \mathbb{N} \rightarrow \mathbb{Z}$, donc il existe une bijection $h : \mathbb{Z} \rightarrow \mathbb{N}$ (avec $h = f^{-1}$). On définit alors une bijection H de $\mathbb{Z}^2$ dans $\mathbb{N}^2$ en posant :

$$\begin{aligned} H : \mathbb{Z}^2 &\longrightarrow \mathbb{N}^2 \\ (n, m) &\longmapsto (h(n), h(m)) \end{aligned}$$

Il s'ensuit que l'application $g \circ \varphi \circ H : \mathbb{Z}^2 \rightarrow \mathbb{N}$ est bijective¹², d'où : $\mathbb{Z}^2$ est dénombrable.

Partie III — Dénombrabilité de $\mathbb{Q}$

7) L'application $\iota : n \in \mathbb{N} \rightarrow n \in \mathbb{Q}$ est clairement injective.¹³

8) a) Soient x et x' deux rationnels. Si x et x' ont la même image (p, q) par ψ , alors $x = x' = p/q$. Donc ψ est injective. En revanche, ψ n'est pas surjective puisque par exemple $(2, 4)$ n'appartient pas à l'image de ψ , comme plus généralement tout couple d'entiers non nuls non premiers entre eux.

b) On dispose d'une application injective (et même bijective) $h = f^{-1} : \mathbb{Z} \rightarrow \mathbb{N}$, et d'une application injective (et même bijective) $g : \mathbb{N}^* \rightarrow \mathbb{N}$. On peut donc construire une application injective :

$$\begin{aligned} G : \mathbb{Z} \times \mathbb{N}^* &\longrightarrow \mathbb{N}^2 \\ (n, m) &\longmapsto (h(n), g(m)) \end{aligned}$$

On dispose en outre d'une application injective (et même bijective) $g \circ \varphi : \mathbb{N}^2 \rightarrow \mathbb{N}$. La composition de deux injections étant encore une injection, l'application :

$$\begin{aligned} g \circ \varphi \circ G : \mathbb{Z} \times \mathbb{N}^* &\longrightarrow \mathbb{N} \\ (n, m) &\longmapsto g(\varphi(h(n), g(m))) \end{aligned}$$

est injective. Une nouvelle application de la stabilité de l'injectivité par composition permet enfin de conclure que l'application :

$$\begin{aligned} g \circ \varphi \circ G \circ \psi : \mathbb{Q} &\longrightarrow \mathbb{N} \\ r &\longmapsto g(\varphi(G(\psi(r)))) \end{aligned}$$

est injective. Conclusion. Il existe une injection de $\mathbb{Q}$ dans $\mathbb{N}$.

D'après les questions 7 et 8-b, il existe une injection de $\mathbb{N}$ dans $\mathbb{Q}$ et une injection de $\mathbb{Q}$ dans $\mathbb{N}$. Le théorème de Cantor-Bernstein (dont l'énoncé est donné ci-dessous) permet alors d'affirmer qu'il existe une bijection entre $\mathbb{N}$ et $\mathbb{Q}$. Par suite : $\mathbb{Q}$ est dénombrable.

THÉORÈME (CANTOR-BERNSTEIN) — Soient E et F deux ensembles. S'il existe une injection de E dans F et une injection de F dans E , alors il existe une bijection entre E et F .

12. C'est la composée de 3 bijections.

13. C'est l'injection *canonique* de $\mathbb{N}$ dans $\mathbb{Q}$.

Chapitre 16

Arithmétique

Au cas où je l'oublierais en cours de route, le terme "entier" désignera dans ce chapitre un entier relatif.

16.1 Divisibilité dans $\mathbb{Z}$

DÉFINITION 16.1 - Soient a et b deux entiers ; b **divise** a (ce que l'on note $b|a$) s'il existe un entier relatif k tel $a = kb$. Dans cette situation, on dit que b est un **diviseur** de a ou que a est **multiple** de b .

Exemples : $2|64834$; $3|4527$; $5|875$; $4|2324$; $11|12327656$. Pour tout $n \in \mathbb{Z}$, $1|n$ et $(-1)|n$; si $n \in \mathbb{Z} \setminus \{\pm 1\}$, n admet au moins quatre diviseurs : ± 1 et $\pm n$. 0 ne divise aucun entier non nul, et est multiple de tout entier.

PROPRIÉTÉ 16.1 - Soient a et b entiers, avec $b \neq 0$. Alors : $a|b \implies |a| \leq |b|$.

PREUVE. Soient a et b comme dans l'énoncé : $\exists k \in \mathbb{Z} \setminus \{0\}$, $b = ka$. D'où : $|a| = \frac{|b|}{|k|}$. Puisque $|k| \geq 1$, on en déduit que : $|a| \leq |b|$. ☕

PROPRIÉTÉ 16.2 - Soient a, b, c et d entiers.

- i/ $(\forall (n, m) \in \mathbb{Z}^2, a|b \text{ et } a|c) \implies (a|nb + mc)$
- ii/ $(a|b \text{ et } a|c) \implies (a|b + c)$
- iii/ $(a|c \text{ et } b|d) \implies (ab|cd)$
- iv/ $a|b \iff a|(-b)$

PREUVE. Soient a, b, c et d entiers.

i/ Supposons que a divise b et c , et soient n et m deux entiers.

Par hypothèse : $\exists (k, k') \in \mathbb{Z}^2$, $b = ka \wedge c = k'a$. Par suite : $nb + mc = (nk + mk')a$. Ce qui signifie que : $a|nb + mc$.

ii/ Cas particulier de i/ avec $n = m = 1$.

iii/ Par hypothèse : $\exists (k, k') \in \mathbb{Z}^2$, $c = ka \wedge d = k'b$. D'où : $cd = kk'(ab)$. Ainsi : $ab|cd$.

iv/ Trivial. ☕

Exemples d'application.

Exemple 1. “Déterminer tous les entiers n tels que $(n+1)|(2n+7)$ ”.

Soit n un tel entier. Alors $(n+1)$ divise $(n+1)$ et $(2n+7)$. Donc : $(n+1)$ divise $(2n+7) - 2(n+1)$. Autrement dit, $(n+1)$ divise 5, c'ad : $n+1 = \pm 1, \pm 5$. D'où finalement : $n = -2, 0, -6, 4$. Réciproquement, on vérifie aisément que ces entiers sont tels que $(n+1)|(2n+7)$.

Exemple 2. “Le produit de 100 entiers consécutifs est multiple de 83”.

Dans un produit P de 100 entiers consécutifs, au moins l'un (disons n_1) est multiple de 83. Donc 83 divise n_1 qui divise P , d'où la conclusion.

Notation. Soit a un entier. On note : $a\mathbb{Z} = \{ka / k \in \mathbb{Z}\}$

COROLLAIRE 16.1 - Soit $a \in \mathbb{Z}$. $(a\mathbb{Z}, +)$ est un sous-groupe de $(\mathbb{Z}, +)$.

PREUVE. Par définition, $a\mathbb{Z}$ est une partie de $\mathbb{Z}$ (SG1), qui contient 0 (car $0 = 0 \times a$, SG2), stable par somme (trivial, SG3) et par passage à l'opposé (re-trivial, SG4). Il s'ensuit que $(a\mathbb{Z}, +)$ est un sous-groupe de $(\mathbb{Z}, +)$. ☕

Pour alléger les notations d'une part, et puisqu'il n'y a pas vraiment d'ambiguïté sur la loi d'autre part, on peut énoncer le corollaire précédent en écrivant que $a\mathbb{Z}$ est un **sous-groupe additif** de $\mathbb{Z}$. Par ailleurs, ce sous-groupe est évidemment abélien¹.

PROPRIÉTÉ 16.3 - i/ $\forall a \in \mathbb{Z}, a|a$

ii/ $\forall (a, b) \in \mathbb{Z}^2, [a|b \text{ et } b|a] \implies |a| = |b|$

iii/ $\forall (a, b, c) \in \mathbb{Z}^3, [a|b \text{ et } b|c] \implies a|c$

PREUVE. i/ $a = 1 \times a \dots$

ii/ Soient a et b deux entiers, tels que : $a|b$ et $b|a$. Alors il existe deux entiers k et k' tels que : $a = kb$ et $b = k'a$. D'où : $a = kk'a$, c'est à dire : $a(1 - kk') = 0$.

On distingue alors deux cas. Si a est nul, alors b est nul (puisque $a|b$ par hypothèse), et la propriété est en particulier vraie. Sinon, on a : $kk' = 1$; puisque k et k' sont entiers, ceci n'est possible que si $k = k' = 1$ ou $k = k' = -1$, c'est à dire si et seulement si a et b sont égaux ou opposés.

Dans les deux cas, on en déduit que : $|a| = |b|$.

iii/ Sous les hypothèses de l'assertion, il existe deux entiers relatifs k et k' tels que : $b = ka$ et $c = k'b$. Il s'ensuit que : $c = kk'a$. D'où : $a|c$. ☕

Le petit i/ de la propriété ci-dessus traduit la réflexivité de la relation de divisibilité, et le iii/ sa transitivité. Il s'en faut donc “de peu” pour que la relation binaire de divisibilité soit une relation d'ordre (les barres de valeur absolue dans l'assertion ii). En revanche, si l'on ne considère que des entiers naturels, le ii/ fournit l'antisymétrie. Par suite :

COROLLAIRE 16.2 - La relation de divisibilité est une relation d'ordre dans $\mathbb{N}$.

PREUVE. Conséquence immédiate des observations précédentes. ☕

1. Puisque tout sous-groupe d'un groupe abélien l'est encore.

16.2 Division euclidienne dans $\mathbb{Z}$ et PGCD

THÉORÈME 16.1 - Poétiquement :

$$\forall (a, b) \in \mathbb{Z} \times \mathbb{N}^*, \exists! (q, r) \in \mathbb{Z} \times \mathbb{N}, [a = bq + r] \wedge [0 \leq r < b]$$

Terminologie : l'entier q est appelé le **quotient**, et r le **reste** dans la division euclidienne de a par b .

PREUVE. Supposons dans un premier temps $a \in \mathbb{N}$, et $b \in \mathbb{N}^*$.

Considérons l'ensemble : $E = \{k \in \mathbb{N} / kb \leq a\}$.

L'ensemble E est une partie non vide ($0 \in E$) et majorée de $\mathbb{N}$, et admet à ce titre un plus grand élément. On pose : $q = \max E$, et $r = a - bq$. De la sorte, on a évidemment : $a = bq + r$.

En outre, puisque $q \in E$, on a $bq \leq a$. Et comme q est le plus grand élément de E , on a $q + 1 \notin E$, càd : $(q + 1)b > a$. En résumé : $a - bq < b$, soit : $0 \leq r < b$.

On a donc établi l'**existence** d'un couple (q, r) d'entiers tels que : $a = bq + r$ et $0 \leq r < b$.

Prouvons son **unicité** : soient (q, r) et (q', r') deux couples d'entiers tels que : $a = bq + r$, $a' = bq' + r'$ et $0 \leq r, r' < b$. Alors : $bq + r = bq' + r'$ d'où : $r' - r = b(q - q')$. Ainsi $r' - r$ est multiple de b .

Or il résulte des conditions satisfaites par r et r' que : $-b < r' - r < b$.

Comme l'unique multiple de b strictement compris entre $-b$ et b est 0, on en déduit que $r = r'$, puis aisément $q = q'$.

Conclusion. Soient $a \in \mathbb{N}$ et $b \in \mathbb{N}^*$. Il existe un unique couple (q, r) d'entiers tels que : $a = bq + r$ et $0 \leq r < b$.

Pour achever la preuve, il reste à traiter le cas où $a \in \mathbb{Z}_-$ (et b est toujours dans $\mathbb{N}^*$). Dans ce cas, on applique le théorème de la division euclidienne à $(-a)$ (celui que l'on vient d'établir) : il existe un unique couple (q, r) d'entiers tels que : $(-a) = bq + r$ et $0 \leq r < b$. D'où : $a = -bq - r$. Hic : le reste r n'est alors plus compris entre 0 et b .

On distingue alors deux cas : si $r = 0$, alors $a = -bq$ et le couple $(-q, 0)$ répond au problème.

Si en revanche $r \neq 0$, alors $-b < -r < 0$. Dans ce cas, on réécrit : $a = -b(q + 1) + b - r$. On a alors : $0 < b - r < b$ (d'où en particulier $0 \leq b - r < b$). Le couple $(-(q + 1), b - r)$ répond alors au problème.

Dans les deux cas, on a établi l'existence d'un couple (q, r) d'entiers tels que : $a = bq + r$ et $0 \leq r < b$. Quant à l'unicité, elle a déjà été établie dans le cas $a \in \mathbb{N}$.

Conclusion. Soient $a \in \mathbb{Z}$ et $b \in \mathbb{N}^*$. Il existe un unique couple (q, r) d'entiers tels que : $a = bq + r$ et $0 \leq r < b$.



Application. Pour mémoire, on a utilisé de manière essentielle ce théorème pour décrire l'ensemble $\mathbb{U}_n$ des racines n -ièmes de l'unité (voir chapitre "Nombres complexes").

Une conséquence immédiate de l'unicité du reste est la suivante :

PROPRIÉTÉ 16.4 - Soient a et b deux entiers : $b|a \iff r = 0$.

PREUVE. Si $b|a$, alors il existe q tel que $b = aq + 0$. Par unicité du reste dans la division euclidienne de a par b , on en déduit que $r = 0$. Réciproque triviale. ☕

DÉFINITION 16.2 - Soient a et b deux entiers non nuls. Le **plus grand commun diviseur (PGCD)** de a et b est le plus grand (au sens de la relation d'ordre $\leq$) entier divisant à la fois a et b . Il sera noté $a \wedge b$.

Par convention, $0 \wedge 0 = 0$.

PROPRIÉTÉ 16.5 - Soient a et b deux entiers.

- i/ $\forall (a, b) \in \mathbb{Z}^2, a \wedge b = b \wedge a$
- ii/ $\forall (a, b) \in \mathbb{Z}^2, a \wedge b = |b| \wedge |a|$
- iii/ $\forall a \in \mathbb{Z}, a \wedge 0 = |a|$

PREUVE. Le point i/ provient de la symétrie de la définition de PGCD ; le point ii/ provient du fait que a divise b SSI $|a|$ divise $|b|$ comme nous l'avons vu au début du chapitre ; le point iii/ provient de ce que tout entier divise 0, et que le PGCD est par convention un entier naturel. ☕

L'algorithme d'Euclide est une méthode qui permet de faire coup double, dans le sens où il permet non seulement de calculer le PGCD de deux entiers, mais aussi de trouver ce que nous appellerons plus tard les coefficients de Bezout de ces entiers. Cet algorithme consiste essentiellement à appliquer itérativement la propriété suivante :

PROPRIÉTÉ 16.6 - Soient a et b deux entiers, $b \neq 0$. On a :

$$a \wedge b = b \wedge r$$

où r désigne le reste dans la division euclidienne de a par b .

PREUVE. D'après le théorème de la division euclidienne : $\exists! (q, r) \in \mathbb{Z}^2, a = bq + r$ et $0 \leq r < b$.

Il suffit de prouver que les diviseurs communs à a et b sont les diviseurs communs à b et r .

► Soit d un diviseur commun à a et b . Alors d divise b (si, si!) et d divise $a - bq = r$. Donc d est un diviseur commun à b et r .

► Réciproquement : soit d un diviseur commun à b et r . Alors d divise b (si, si!) et d divise $bq + r = a$. Donc d est un diviseur commun à a et b .

En résumé : les entiers a et b d'une part, et les entiers b et r d'autre part, ont les mêmes diviseurs communs.

En particulier : $a \wedge b = b \wedge r$. ☕

La conséquence la plus importante de cette propriété est que le PGCD de a et b est le dernier reste non nul obtenu dans la suite de divisions euclidiennes effectuées dans l'algorithme d'Euclide appliqué à a et b .

Exemple d'utilisation de l'algorithme d'Euclide. Calcul du PGCD de 25 et 72.

$$72 = 25 \times 2 + 22$$

$$25 = 22 \times 1 + 3$$

$$22 = 3 \times 7 + 1$$

$$3 = 3 \times 1 + 0$$

On en déduit que $\boxed{25 \wedge 72 = 1}$.

► En outre, “en remontant” cet algorithme, on peut établir le :

THÉORÈME 16.2 - (Bezout)

$$\forall (a, b) \in \mathbb{Z}^2, \exists (u, v) \in \mathbb{Z}^2, au + bv = a \wedge b$$

PREUVE. La propriété est immédiate lorsqu'au moins un des deux entiers a ou b est nul. On suppose donc dans la suite de la preuve que a et b sont non nuls.

On considère l'ensemble : $E = \{m \in \mathbb{N}^* / \exists (u, v) \in \mathbb{Z}^2, m = au + bv\}$.

Il est assez facile de se convaincre que E est non vide. En effet, puisque a est non nul, E contiendra au moins l'élément a (*resp.* $-a$) si a est positif (*resp.* négatif). Ainsi E est une partie non vide de $\mathbb{N}^*$; elle admet à ce titre un plus petit élément. Notons $d = \min E$.

Observons que d étant un élément de $E : \exists (u, v) \in \mathbb{Z}^2, au + bv = d$ ($\spadesuit$).

Montrons que : $d = a \wedge b$.

► On note déjà que $(a \wedge b) | a$ et $(a \wedge b) | b$, donc : $(a \wedge b) | au + bv$, c-à-d : $(a \wedge b) | d$.

Il s'ensuit que : $(a \wedge b) \leq d$ ($\clubsuit$).

► Montrons que d divise a . D'après le théorème de la division euclidienne :

$$\exists! (q, r) \in \mathbb{Z}^2, a = dq + r \text{ avec } 0 \leq r < d.$$

On a : $r = a - dq = a - (au + bv)$ d'où $r = a(1 - uq) + b(vq) \in E$. Si r était non nul, alors r serait un élément de E strictement inférieur au minimum d de E : contradiction. Il s'ensuit que $r = 0$. Par conséquent : d divise a .

Un raisonnement analogue permet d'affirmer que : d divise b .

Ainsi d est un diviseur commun à a et b , et à ce titre d est inférieur au plus grand d'entre eux.

Explicitement : $d \leq (a \wedge b)$ ($\heartsuit$).

On déduit alors de ($\clubsuit$) et de ($\heartsuit$) que : $d = (a \wedge b)$.

D'où, grâce à ($\spadesuit$) : $\boxed{\exists (u, v) \in \mathbb{Z}^2, au + bv = a \wedge b}$. 

Remarque. Les entiers u et v ci-dessus peuvent être appelés *coefficients de Bezout* des entiers a et b . Il n'y a pas unicité de ce couple. Pour finir, on peut déterminer de tels entiers en “remontant” l'algorithme d'Euclide.

A titre d'exemple, on détermine ci-dessous des coefficients pour les entiers 25 et 72.

$$1 = 22 - 7 \times 3$$

$$1 = 22 - 7 \times (25 - 22) = 8 \times 22 - 7 \times 25$$

$$1 = 8 \times (72 - 2 \times 25) - 7 \times 25 \text{ soit finalement :}$$

$$1 = 8 \times 72 - 23 \times 25$$

On a donc obtenu une relation de Bezout pour les entiers 25 et 72, c'est à dire deux entiers u et v tels que $25u + 72v = 25 \wedge 72$ avec $u = -23$ et $v = 8$.

► Première conséquence : tout diviseur commun à a et b divise $a \wedge b$.

COROLLAIRE 16.3 - Soient a et b deux entiers. On a : $[d|a \text{ et } d|b] \iff [d|a \wedge b]$

Remarque. Ainsi, $a \wedge b$ est aussi le plus grand commun diviseur à a et b pour la relation (d'ordre) de divisibilité.

PREUVE. Sens direct : soit d un entier qui divise a et b . D'après le théorème de Bezout, il existe un couple d'entiers (u, v) tels que : $au + bv = a \wedge b$. Puisque d divise a et b , alors d divise $au + bv$, donc d divise $a \wedge b$.

Réciproquement, si d divise $a \wedge b$, alors d divise a et b puisque $a \wedge b$ divise a et b (+ transitivité de la relation de divisibilité). ☕

COROLLAIRE 16.4 - Soient a et b deux entiers. On a :

$$a\mathbb{Z} + b\mathbb{Z} = (a \wedge b)\mathbb{Z}$$

PREUVE. ► Si au moins l'un des deux entiers a ou b est nul, la propriété est triviale.

► Dans la suite des événements, on suppose donc a et b non nuls. Montrons l'égalité de l'énoncé par double inclusion.

Prouvons $a\mathbb{Z} + b\mathbb{Z} \subset (a \wedge b)\mathbb{Z}$: par définition du PGCD, $(a \wedge b) | a$ et $(a \wedge b) | b$. Il existe donc deux entiers a' et b' tels que $a = (a \wedge b)a'$ et $b = (a \wedge b)b'$.

Soit $N \in a\mathbb{Z} + b\mathbb{Z}$. Alors il existe deux entiers k et k' tels que : $N = ka + k'b$. D'où : $N = (a \wedge b)(ka' + k'b')$, et $N \in (a \wedge b)\mathbb{Z}$. Donc : $a\mathbb{Z} + b\mathbb{Z} \subset (a \wedge b)\mathbb{Z}$ (♠)

Prouvons $(a \wedge b)\mathbb{Z} \subset a\mathbb{Z} + b\mathbb{Z}$: d'après le théorème de Bezout, il existe deux entiers u et v tels que : $(a \wedge b) = au + bv$.

Soit $N \in (a \wedge b)\mathbb{Z}$. Alors il existe un entier k tel que : $N = k(a \wedge b)$. Donc : $N = a(ku) + b(kv)$. Donc : $N \in a\mathbb{Z} + b\mathbb{Z}$. D'où : $(a \wedge b)\mathbb{Z} \subset a\mathbb{Z} + b\mathbb{Z}$ (♣)

Conclusion. On déduit de (♠) et de (♣) que : $(a \wedge b)\mathbb{Z} = a\mathbb{Z} + b\mathbb{Z}$. ☕

COROLLAIRE 16.5 - Soient a et b deux entiers.

S'il existe deux entiers u et v tels que $au + bv = 1$, alors $a \wedge b = 1$.

PREUVE. Soit d un diviseur commun à a et b . Alors d divise $au + bv$, donc d divise 1. On en déduit que les seuls diviseurs communs à a et b sont 1 et -1 . D'où : $a \wedge b = 1$. ☕

16.3 Congruences

DÉFINITION 16.3 - Soient a, b et n trois entiers relatifs. On dit que a **est congru à b modulo n** si $b - a$ est multiple de n , c'est à dire s'il existe $k \in \mathbb{Z}$ tel que : $b = a + kn$. On le note $a \equiv b [n]$.

Remarques : pour tout entier n , les multiples de n sont les multiples de $(-n)$. On pourra donc se restreindre à étudier des congruences modulo un entier naturel n . En outre, on n'étudiera guère les congruences modulo 0 (puisqu'alors tout entier a n'est congru qu'à lui-même) ni modulo 1 (deux entiers sont toujours congrus modulo 1). En pratique, on supposera donc souvent n entier et $n \geq 2$.

Notons aussi que : $a \equiv 0 [n]$ si et seulement si a est multiple de n .

PROPRIÉTÉ 16.7 - Soit n un entier, $n \geq 2$.

i/ $\forall a \in \mathbb{Z}, a \equiv a [n]$

ii/ $\forall (a, b) \in \mathbb{Z}^2, a \equiv b [n] \iff b \equiv a [n]$

iii/ $\forall (a, b, c) \in \mathbb{Z}^3, [a \equiv b [n] \text{ et } b \equiv c [n]] \implies [a \equiv c [n]]$

Autrement dit, la relation de congruence est une relation d'équivalence sur $\mathbb{Z}$.

PREUVE. Soit a un entier. On a : $a = a + 0 \times n$, d'où le point i/.

Si $a \equiv b [n]$, alors il existe $k \in \mathbb{Z}, a = b + kn$. D'où $b = a - kn$. En posant $K = -k$, on a donc : $a = b + Kn$ avec $k \in \mathbb{Z}$. Ainsi $b \equiv a [n]$, ce qui prouve l'implication : $a \equiv b [n] \implies b \equiv a [n]$; la réciproque vient du fait que l'on peut permuter a et b dans le raisonnement précédent, ce qui prouve l'équivalence du point ii/.

Sous les hypothèses de l'énoncé du iii, on a : $b = a + kn$ et $c = b + k'n$ d'où $c = a + (k + k')n$; ce qui prouve le transitivité de la relation de congruence. ☕

En outre, la relation de congruence est compatible avec la structure d'anneau de $\mathbb{Z}$, dans le sens plus explicite suivant :

PROPRIÉTÉ 16.8 - Soit n un entier, $n \geq 2$; et soient a, b, c et d quatre entiers tels que $a \equiv b [n]$ et $c \equiv d [n]$. Alors :

i/ $a + c \equiv b + d [n]$

ii/ $ac \equiv bd [n]$

iii/ $\forall N \in \mathbb{N}, a^N \equiv b^N [n]$

PREUVE. Triviale. ☕

Exemples d'application. “Montrer que $2^{3000} - 1$ est multiple de 7” ; “Montrer que l'équation $5x^2 + 2y^4 = 9999$ n'a pas de solution dans $\mathbb{Z}^2$ ” ; “Montrer que $7|n$ et $7|m$ SSI $7|n^2 + m^2$ ”...

En pratique, on utilise souvent la propriété suivante :

PROPRIÉTÉ 16.9 - Soit n un entier, $n \geq 2$, et a un entier quelconque.
On a : $a \equiv r [n]$, où r désigne le reste dans la division euclidienne de a par n .

PREUVE. D'après le théorème de la division euclidienne : $a = nq + r$, d'où $a \equiv r [n]$. ☕

DÉFINITION 16.4 - Soit a et n deux entiers, avec $n \geq 2$. On appelle **inverse de a modulo n** un entier b tel que $ab \equiv 1 [n]$.

Exemples. 3 a pour inverse 5 modulo 7. Mais 12 et -2 sont aussi des inverses de 3 modulo 7. Plus généralement :

LEMME 16.1 - Si b est un inverse de a modulo n , alors $b + kn$ est un inverse de a modulo n (pour tout entier k).

PREUVE. Triviale. ☕

En d'autres termes, un entier admet une infinité d'inverses modulo n dès qu'il en admet un. Mais à l'opposé, il peut très bien ne pas exister d'inverse du tout : par exemple, 8 n'a pas d'inverse modulo 4, 6 n'a pas d'inverse modulo 9... Plus généralement :

PROPRIÉTÉ 16.10 - Soit n un entier, $n \geq 2$, et $a \in \mathbb{Z}$. L'entier a est inversible modulo n si et seulement si $a \wedge n = 1$.

PREUVE. ► Supposons a inversible modulo n : alors il existe un entier m tel que $am \equiv 1 [n]$. Par définition de congruence, ceci implique qu'il existe un entier k tel que $am = 1 + kn$, d'où : $am - kn = 1$. On en déduit que : $a \wedge n = 1$.

► Réciproquement, si $a \wedge n = 1$, alors il existe d'après le théorème de Bezout un couple d'entiers (u, v) tels que : $au + bn = 1$. On en déduit que : $au \equiv 1 [n]$. Par suite a est inversible modulo n (et u est un inverse de a modulo n).

Conclusion. a est inversible modulo n si et seulement si $a \wedge n = 1$. ☕

16.4 Entiers premiers entre eux, Lemme de Gauss

DÉFINITION 16.5 - Deux entiers a et b sont **premiers entre eux** si $a \wedge b = 1$.

Exemples : 3 et 8 sont premiers entre eux ; 16 et 33 sont premiers entre eux ; pour $n \geq 2$, les entiers $2n - 1$ et $2n + 1$ sont premiers entre eux.

PROPRIÉTÉ 16.11 - Soient a et b deux entiers non nuls. On note $d = a \wedge b$. Il existe un unique couple d'entiers relatifs tels que $a = da'$, $b = db'$ et $a' \wedge b' = 1$.

PREUVE. Puisque d est un diviseur commun à a et b , il existe un couple d'entiers (a', b') tels que $a = da'$ et $b = db'$; ce couple est clairement unique.

Reste à établir que a' et b' sont premiers entre eux. Si tel n'était pas le cas, il existerait un entier $k \geq 2$ divisant a' et b' . On aurait alors : $a' = ka''$ et $b' = kb''$. D'où : $a = dka''$ et $b = dkb''$. Dans cette situation, dk serait donc un diviseur commun à a et b , strictement plus grand que d : contradiction !

Il s'ensuit que a' et b' sont premiers entre eux, ce qui achève la preuve. ☕

Application. Déterminer tous les couples d'entiers relatifs tels que :
$$\begin{cases} x \wedge y = 15 \\ xy = 900 \end{cases}$$

THÉORÈME 16.3 - (Lemme de Gauss) Soient a, b et c trois entiers.

Si $a|bc$ et $a \wedge b = 1$ alors $a|c$.

PREUVE. Puisque $a|bc$, il existe un entier k tel que : $bc = ka$.

Par ailleurs puisque $a \wedge b = 1$, il existe d'après le théorème de Bezout un couple (u, v) d'entiers tels que : $au + bv = 1$.

On en déduit que : $bcv = kav$ d'où : $(1 - au)c = a(kv)$ soit encore : $a(kv + uc) = c$. Puisque a divise le terme de gauche de cette égalité, il divise également celui de droite, càd : $a|c$.

Conclusion. Si $a|bc$ et $a \wedge b = 1$ alors $a|c$. ☕

PROPRIÉTÉ 16.12 - Soient a et b deux entiers tels que $a \wedge b = 1$.

Si $a|c$ et $b|c$, alors $ab|c$.

PREUVE. Puisque $a \wedge b = 1$, il existe d'après le théorème de Bezout un couple (u, v) d'entiers tels que : $au + bv = 1$.

Par ailleurs, puisque $a|c$ et $b|c$, il existe deux entiers k et k' tels que $c = ka$ et $c = k'b$. On en déduit que : $c = ka(au + bv)$ d'où : $c = ka^2u + kabv$. L'entier b divise c (par hypothèse) et $kabv$ (immédiat). Donc : $b|ka^2u$. Or a et b sont premiers entre eux, donc a^2 et b sont également premiers entre eux. On en déduit grâce au lemme de Gauss que : $b|ku$. Il existe donc un entier m tel que : $ku = mb$.

Finalement : $c = a^2mb + kabv$, donc : $c = ab(am + kv)$, ce qui prouve que c est multiple de ab .

Conclusion. Si $a \wedge b = 1$, alors $a|c$ et $b|c$ implique $ab|c$. ☕

Application : le produit de trois entiers consécutifs est multiple de 6.

Soient $n, n+1$ et $n+2$ trois entiers consécutifs, dont on note le produit P . Au moins un de ces entiers est multiple de 3, et au moins l'un d'entre eux est pair. Donc 2 et 3 divisent P . Puisque $2 \wedge 3 = 1$, on conclut grâce à ce qui précède que P est multiple de 6.

16.5 L'équation diophantienne $ax + by = c$

Dans ce paragraphe, on considère l'équation (E) : $ax + by = c$ avec a, b et c entiers.

16.5.1 Equation homogène associée

PROPRIÉTÉ 16.13 - On note $d = a \wedge b$, $a = da'$ et $b = db'$.

L'ensemble des solutions de l'équation (H) : $ax + by = 0$ est : $\{(kb', -ka') / k \in \mathbb{Z}\}$.

PREUVE. Notons S l'ensemble des solutions de l'équation (H), et E l'ensemble $\{(kb', -ka') / k \in \mathbb{Z}\}$.

► Montrons que $S \subset E$. Soit $(x, y) \in \mathbb{Z}^2$ un couple solution de (H). Alors : $ax + by = 0$, donc $da'x + db'y = 0$. Puisque $d \neq 0$ (a et b étant non nuls), on en déduit : $a'x + b'y = 0$, puis : $b'y = -a'x$. L'entier b' divise donc $a'x$, et comme il est premier avec a' , on en déduit par le lemme de Gauss que b' divise x . Il existe donc un entier k tel que : $x = kb'$.

Alors : $b'y = -a'kb'$ d'où : $y = -ka'$. Ainsi : $(x, y) = (kb', -ka')$, càd : $(x, y) \in E$. Par suite : $\boxed{S \subset E}$.

► L'inclusion $E \subset S$ est immédiate.

Conclusion. $S = \{(kb', -ka') / k \in \mathbb{Z}\}$.



Exemple. Soit (H) l'équation : $30x + 21y = 0$. L'ensemble des solutions de (H) est : $\{(7k, -10k) / k \in \mathbb{Z}\}$.

16.5.2 Equation complète

PROPRIÉTÉ 16.14 - (Structure de l'ensemble des solutions de (E)). Si $(X, Y) \in \mathbb{Z}^2$ est un couple solution de (E), alors l'ensemble des solutions de (E) est :

$$\{(X + kb', Y - ka') / k \in \mathbb{Z}\}$$

PREUVE. Soit (x, y) un couple d'entiers. On a :

$$(x, y) \text{ solution de (E)}$$

$$\iff ax + by = c$$

$$\iff ax + by = aX + bY$$

$$\iff a(x - X) + b(y - Y) = 0$$

$$\iff \exists k \in \mathbb{Z}, x - X = kb' \wedge y - Y = kb'$$

$$\iff \exists k \in \mathbb{Z}, x = X + kb' \wedge y = Y + kb' \quad \text{☕}$$

On dispose donc d'une formule explicite donnant la solution générale de (E), sous réserve qu'il en existe au moins une (et que l'on puisse la trouver!). Deux cas peuvent se présenter :

► Si $a \wedge b | c$: alors l'équation (E) possède des solutions. On en trouve une en déterminant les coefficients de Bezout de a et b , puis toutes à l'aide de la propriété ci-dessus.

► Sinon (si c n'est pas multiple de $a \wedge b$) : alors l'équation (E) ne possède aucune solution.

16.6 PPCM

DÉFINITION 16.6 - Soient a et b deux entiers non nuls. Le **plus petit commun multiple (PPCM)** de a et b est le plus petit (au sens de la relation d'ordre $\leq$) entier naturel multiple à la fois de a et de b . Le PPCM de a et b est noté $a \vee b$.

PROPRIÉTÉ 16.15 - (**Propriétés du PPCM**) Soient a et b deux entiers.

- i) $\forall (a, b) \in \mathbb{Z}^2, a \vee b = b \vee a$
- ii) $\forall (a, b) \in \mathbb{Z}^2, a \vee b = |b| \vee |a|$
- iii) $\forall a \in \mathbb{Z}, a \vee 0 = 0$

PREUVE. Triviale. ☕

PROPRIÉTÉ 16.16 - Soient a et b deux entiers. Si $a \wedge b = 1$, alors $a \vee b = |ab|$.

PREUVE. Il est clair que $|ab|$ est un multiple commun à a et b . Montrons que c'est le plus petit.

Notons $E = \text{Mul}(a, b) = \{k \in \mathbb{N}^*, a|k \text{ et } b|k\}$.

E est une partie non vide et minorée de $\mathbb{N}^*$. A ce titre, E admet un plus petit élément. Notons : $m = \min E$.

Puisque $m \in E$, alors a et b divisent m . Puisqu'en outre a et b sont premiers entre eux, on en déduit (conséquence du lemme de Gauss) que ab divise m . Il existe donc un entier k non nul (puisque m est non nul) tel que : $m = kab$. On en déduit que $m = |k| \times |ab|$, d'où la conclusion puisque $|k| > 1$. ☕

PROPRIÉTÉ 16.17 - Soient a et b deux entiers. On note $d = a \wedge b$, et (a', b') l'unique couple d'entiers relatifs tels que $a = da'$ et $b = db'$. Alors : $a \vee b = d|a'b'|$.

PREUVE. Avec les notations et hypothèses de l'énoncé, il est clair que $d|a'b'|$ est un multiple commun à a et b .

Montrons que c'est le plus petit (dans $\mathbb{N}^*$). Soit m un multiple commun à a et b . Il existe alors deux entiers k et k' tels que : $m = ka$ et $m = k'b$. On a donc : $m = kda'$ et $m = k'db'$. Puisque m est multiple de d , on peut considérer l'entier $\frac{m}{d}$ et écrire que :

$$\frac{m}{d} = ka' = k'b'$$

Il s'ensuit que $\frac{m}{d}$ est un multiple commun à a' et b' . On en déduit, d'après la propriété précédente (et puisque a' et b' sont premiers entre eux), que : $a'b'|\frac{m}{d}$. Il existe donc un entier p tel que : $\frac{m}{d} = pa'b'$, et par suite : $m = |p| d|a'b'|$.

Ainsi : $a \vee b = d|a'b'|$. ☕

PROPRIÉTÉ 16.18 - (Lien PGCD $\longleftrightarrow$ PPCM).

$$\forall (a, b) \in \mathbb{Z}^2, (a \wedge b) \times (a \vee b) = |ab|$$

PREUVE. On peut déjà supposer que a et b sont non nuls (sinon il n'y a pas grand chose à dire), et qu'ils sont tous les deux positifs (puisque $(a \wedge b) = (|a| \wedge |b|)$ et $(a \vee b) = (|a| \vee |b|)$).

Par ailleurs, il existe un unique couple (a', b') d'entiers premiers entre eux tels que : $a = (a \wedge b) a'$ et $b = (a \wedge b) b'$.

A ce point du raisonnement, si l'on admet la propriété suivant laquelle $a \vee b = (a \wedge b) a' b'$, le résultat est immédiat. On a en effet :

$$(a \wedge b) \times (a \vee b) = (a \wedge b) \times (a \wedge b) a' b' = \underbrace{[(a \wedge b) \times a']}_{=a} \times \underbrace{[(a \wedge b) \times b']}_{=b}, \text{ d'où : } (a \wedge b) \times (a \vee b) = ab.$$

Conclusion. $\forall (a, b) \in \mathbb{Z}^2, (a \wedge b) \times (a \vee b) = |ab|.$



16.7 Nombres premiers

16.7.1 Définition, caractérisation

DÉFINITION 16.7 - Un nombre **premier** est un entier admettant exactement deux diviseurs dans $\mathbb{N}$: 1 et lui-même. Un entier non premier est appelé **composé**.

Notation. On note $\mathcal{P}$ l'ensemble des nombres premiers.

Exemples. 2, 3, 5, 7, 11, 13, 59 sont premiers. 0, 1, 1341, 745, 3883 sont composés.

PROPRIÉTÉ 16.19 - Soit $p \in \mathbb{N}$. LASSE :

- 1) p est premier
- 2) p est premier avec tout entier qu'il ne divise pas.
- 3) p est premier avec tout entier compris entre 2 et $p - 1$.

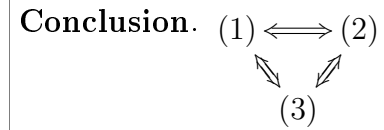
PREUVE. Il suffit de prouver la suite d'implications $1) \implies 2) \implies 3) \implies 1)$ pour que la boucle soit bouclée !

Prouvons que $1) \implies 2)$. Soit p un nombre premier, et soit n un entier tel que p ne divise pas n . Raisonnons par l'absurde, en supposant que p et n ne sont pas premiers entre eux. Alors il existe un entier naturel $d > 1$ divisant p et n . En particulier, d divise p , et puisque p est premier (et $d > 1$), on a nécessairement : $d = p$. Il s'ensuit que p divise n : contradiction. Par conséquent, si p est premier, alors p est premier avec tout entier qu'il ne divise pas.

Prouvons que $2) \implies 3)$. Evident, après avoir observé que p ne divise aucun entier compris entre 2 et $p - 1$.

Prouvons que $3) \implies 1)$. Supposons que p soit premier avec tout entier compris entre 2 et $p - 1$, et montrons que p est premier. Par l'absurde encore une fois, on suppose que p n'est pas premier. Il admet alors un diviseur entier naturel d distinct de 1 et de lui-même, càd $d \in \llbracket 2, p - 1 \rrbracket$. On en déduit que d n'est pas premier avec p : contradiction.

Il s'ensuit que si p est premier avec tout entier compris entre 2 et $p-1$, alors p est premier.



16.7.2 Petit théorème de Fermat

PROPRIÉTÉ 16.20 - Soit p un nombre premier. Alors :

$$\forall k \in \llbracket 1, p-1 \rrbracket, \quad p \mid \binom{p}{k}$$

Interprétation. Sur la “ligne p ” du triangle de Pascal, p divise tous les coefficients binomiaux (à l’exception des deux extrêmes, qui sont égaux à 1!).

PREUVE. Soit p un nombre premier. Pour tout entier $k \in \llbracket 1, p-1 \rrbracket$, on a : $\binom{p}{k} = \frac{p}{k} \binom{p-1}{k-1}$. Il s’ensuit que : $k \binom{p}{k} = p \binom{p-1}{k-1}$. D’où $p \mid k \binom{p}{k}$. Puisque p et k sont premiers entre eux², on en déduit par le lemme de Gauss que $p \mid \binom{p}{k}$.

Conclusion. $\forall p \in \mathcal{P}, \forall k \in \llbracket 1, p-1 \rrbracket, p \mid \binom{p}{k}$.

THÉORÈME 16.4 - (“Petit” théorème de Fermat) Soit p un nombre premier.

Alors : $\forall m \in \mathbb{Z}, m^p \equiv m \pmod{p}$.

Autrement écrit : $[p \in \mathcal{P}] \implies [\forall m \in \mathbb{Z}, m^p \equiv m \pmod{p}]$

PREUVE. Fixons p un nombre premier, et prouvons l’assertion $m^p \equiv m \pmod{p}$ par récurrence sur m , pour tout $m \in \mathbb{N}$.

L’initialisation (pour $m = 0$) est évidente.

Supposons donc la propriété établie à un certain rang m , avec $m \in \mathbb{N}$, et établissons la au rang $m+1$.

On a : $(m+1)^p = \sum_{k=0}^p \binom{p}{k} m^k = m^p + 1 + \sum_{k=1}^{p-1} \binom{p}{k} m^k$. Or, p étant premier, on a : $\forall k \in \llbracket 1, p-1 \rrbracket$

, $p \mid \binom{p}{k}$. Par conséquent : $\sum_{k=1}^{p-1} \binom{p}{k} m^k \equiv 0 \pmod{p}$. Il s’ensuit que : $(m+1)^p \equiv m^p + 1 \pmod{p}$. L’hypothèse de récurrence permet alors d’affirmer que $(m+1)^p \equiv m+1 \pmod{p}$, ce qui prouve l’hérédité.

Conclusion intermédiaire : $\forall p \in \mathcal{P}, \forall m \in \mathbb{N}, m^p \equiv m \pmod{p}$ (♠)

2. L’entier p étant premier, il est premier avec tout entier qu’il ne divise pas.

Montrons à présent la propriété pour $m \in \mathbb{Z}_-$. Si m est un entier négatif, alors $(-m) \in \mathbb{N}$ et d'après ce qui précède, on a donc : $(-m)^p \equiv -m \pmod{p}$. Deux cas sont alors à distinguer, suivant que p est impair ou $p = 2$.

Si p est impair, a $(-m)^p = -m^p$, et on peut alors conclure que $m^p \equiv m \pmod{p}$.

Et si $p = 2$, alors $(-m)^2 = m^2$. D'où : $m^2 \equiv -m \pmod{2}$, et on conclut en observant³ que $1 \equiv -1 \pmod{2}$.

Conclusion : $\forall p \in \mathcal{P}, \forall m \in \mathbb{Z}, m^p \equiv m \pmod{p}$ (♠) 

COROLLAIRE 16.6 - Soit p un nombre premier. Alors :

$$\forall m \in \mathbb{Z} \setminus p\mathbb{Z}, m^{p-1} \equiv 1 \pmod{p}$$

PREUVE. Soient p un nombre premier, et m un entier non multiple de p (càd $m \in \mathbb{Z} \setminus p\mathbb{Z}$). D'après le petit théorème de Fermat : $m^p \equiv m \pmod{p}$. Il est équivalent de dire que p divise $m^p - m$, soit encore : $p \mid m \times (m^{p-1} - 1)$. Or p est premier avec m (puisque'il ne le divise pas), d'où grâce au lemme de Gauss : $p \mid (m^{p-1} - 1)$. Par suite : $m^{p-1} \equiv 1 \pmod{p}$.

Conclusion. $\forall p \in \mathcal{P}, \forall m \in \mathbb{Z} \setminus p\mathbb{Z}, m^{p-1} \equiv 1 \pmod{p}$ 

16.7.3 Décomposition en facteurs premiers d'un entier

PROPRIÉTÉ 16.21 - Tout entier $n \geq 2$ admet au moins un diviseur premier.

PREUVE. Prouvons la propriété $P(n)$: “ n admet un diviseur premier” par récurrence forte sur n .

L'initialisation (pour $n = 2$) est immédiate ; passons à l'hérédité. On suppose donc qu'il existe un entier $n \geq 2$ tel que $P(k)$ est vraie pour tout entier k compris entre 2 et n , puis on distingue deux cas :

Premier cas : $(n + 1)$ est premier. Alors $P(n + 1)$ est vérifiée.

Second cas : $(n + 1)$ est composé. Alors $(n + 1)$ admet un diviseur d dans $\mathbb{N}$, avec $d \in \llbracket 2, n \rrbracket$. Par hypothèse de récurrence (appliquée à l'entier d), il existe un nombre premier p tel que p divise d . Par transitivité de la relation de divisibilité, on en déduit que p divise $(n + 1)$. Ainsi $(n + 1)$ admet un diviseur premier, et $P(n + 1)$ est donc vraie.

Conclusion. Tout entier $n \geq 2$ admet un diviseur premier. 

PROPRIÉTÉ 16.22 - Il existe une infinité de nombres premiers.

PREUVE. Notons $\mathcal{P}$ l'ensemble des nombres premiers, et supposons que $\mathcal{P}$ soit fini. Il existe alors un entier naturel non nul $n \in \mathbb{N}^*$ tels que $\mathcal{P} = \{p_1, \dots, p_n\}$.

On pose $N = 1 + \prod_{i=1}^n p_i$. Supposons que N ne soit pas premier. Alors N admet un diviseur premier (propriété 21) ; il existe donc un p_j (pour un certain $j \in \llbracket 1, n \rrbracket$) qui divise N . Dans ce cas, p_j divise N et clairement $\prod_{i=1}^n p_i$, donc p_j divise 1 : absurde. Il s'ensuit que N est premier. Or par construction : $N > \max_{p \in \mathcal{P}} p$, donc $N \notin \mathcal{P}$.

3. Ou en observant que m et $-m$ ont la même parité, ce qui revient au même.

Ainsi, $N \in \mathcal{P}$ et $N \notin \mathcal{P}$: contradiction. Il s'ensuit que $\mathcal{P}$ est infini.



COROLLAIRE 16.7 - Tout entier $n \geq 2$ composé admet au moins un diviseur compris entre 2 et $\lfloor \sqrt{n} \rfloor$.

PREUVE. Soit n comme dans l'énoncé. Il existe deux entiers p et $q \geq 2$ tels que $pq = n$. L'un des entiers (au moins) p ou q est inférieur ou égal à $\sqrt{n}$ (par l'absurde, si les deux étaient $> \sqrt{n}$...). Sans nuire à la généralité, on peut supposer que $p \leq \sqrt{n}$; et puisque p est entier, on a : $2 \leq p \leq \lfloor \sqrt{n} \rfloor$.

THÉORÈME 16.5 - (**Décomposition en facteurs premiers**). Soit N un entier ≥ 2 .

Il existe n nombres premiers $p_1, \dots, p_n$ et n entiers naturels $\alpha_1, \dots, \alpha_n$ tels que :

$$n = \prod_{i=1}^N p_i^{\alpha_i}$$

Cette écriture, unique à l'ordre des facteurs près, est appelée **décomposition en facteurs premiers** de N .

PREUVE. Prouvons la propriété $P(n)$: “ n admet une décomposition en facteurs premiers” par récurrence forte sur n .

L'initialisation (pour $n = 2$) est immédiate ; passons à l'hérédité. On suppose donc qu'il existe un entier $n \geq 2$ tel que $P(k)$ est vraie pour tout entier k compris entre 2 et n , puis on distingue deux cas :

Premier cas : $(n + 1)$ est premier. Alors $P(n + 1)$ est vérifiée.

Second cas : $(n + 1)$ est composé. Alors $(n + 1)$ admet un diviseur premier p , avec $p \in \llbracket 2, n \rrbracket$; et il existe alors un entier naturel k tel que : $n + 1 = kp$ (avec k également dans $\llbracket 2, n \rrbracket$). Par hypothèse de récurrence (appliquée à l'entier k), il existe N nombres premiers $p_1, \dots, p_N$ et N entiers naturels $\alpha_1, \dots, \alpha_N$ tels que :

$$k = \prod_{i=1}^N p_i^{\alpha_i}.$$

On en déduit que : $n + 1 = p \prod_{i=1}^N p_i^{\alpha_i}$, et on a ainsi écrit $(n + 1)$ comme produit de facteurs premiers. Il s'ensuit que $P(n + 1)$ est vraie.

Conclusion. Tout entier n supérieur ou égal à 2 admet une décomposition en facteurs premiers.

Unicité : soit $n \geq 2$ un entier. Supposons que l'on puisse écrire $n = \prod_{i=1}^N p_i^{\alpha_i} = \prod_{j=1}^M q_j^{\beta_j}$ où les p_i et les q_j sont

premiers (les p_i deux à deux distincts, les q_j deux à deux distincts). L'entier p_1 divise n , donc p_1 divise $\prod_{j=1}^M q_j^{\beta_j}$. Si tous les q_j étaient distincts de p_1 , alors p_1 serait premier avec chaque q_j , donc avec chaque $q_j^{\beta_j}$, et donc avec leur produit (appliquer le lemme de Gauss à répétition pour chaque “donc”) ; dans ce cas, il ne pourrait diviser n : contradiction. Par suite, il existe un q_{j_0} tel que p_1 divise q_{j_0} , et donc $p_1 = q_{j_0}$ (puisque p_1 et q_{j_0} sont premiers). Quitte à renuméroter les q_j , on peut supposer que : $q_{j_0} = q_1$. On en déduit alors que : $p_1 = q_1$.

Un raisonnement (par l'absurde encore une fois) permet alors d'obtenir $\alpha_1 = \beta_1$, puis : $\prod_{i=2}^N p_i^{\alpha_i} = \prod_{j=2}^M q_j^{\beta_j}$. En réitérant le processus, on obtient $p_2 = q_2$ et $\alpha_2 = \beta_2, \dots$. Si l'on veut faire proprement les choses, on écrit une jolie récurrence pour conclure que : $N = M$, et quitte à renuméroter les q_i , $q_i = p_i$ pour tout $i \in \llbracket 1, N \rrbracket$, et $\alpha_i = \beta_i$ pour tout $i \in \llbracket 1, N \rrbracket$, ce qui établit l'unicité. ☕

16.7.4 Valuations p -adiques

DÉFINITION 16.8 - Soit n un entier relatif non nul et p un nombre premier.

On appelle **valuation p -adique** de n l'exposant de n dans la décomposition en facteurs premiers de p .

Exemple : $360 = 2^3 \times 3^2 \times 5$. On a $v_2(360) = 3$, $v_3(360) = 2$, $v_5(360) = 1$ et pour tout $p \in \mathcal{P} \setminus \{2, 3, 5\}$, $v_p(360) = 0$.

PROPRIÉTÉ 16.23 - (Propriétés de la valuation) Avec les notations de la définition, la valuation p -adique de n est la plus grande puissance de n divisant p . Explicitement, $p^{v_p(n)}$ divise n , et $p^{v_p(n)+1}$ ne divise pas n . En outre :

- La suite $(v_p(n))_{p \in \mathcal{P}}$ est nulle à partir d'un certain rang (c'est une conséquence du fait que tout entier non nul est divisible par un nombre fini de nombres premiers, d'après le théorème 6 de décomposition en facteurs premiers).
- Pour tout entier naturel n non nul, $n = \prod_{p \in \mathcal{P}} p^{v_p(n)}$ (c'est une réécriture du théorème 6 utilisant la définition de valuation).
- $\forall (m, n) \in (\mathbb{N}^*)^2$, $v_p(mn) = v_p(m) + v_p(n)$
- PGCD, PPCM et valuations : $\forall (m, n) \in (\mathbb{N}^*)^2$,

$$m \wedge n = \prod_{p \in \mathcal{P}} p^{\min(v_p(m), v_p(n))} \text{ et } m \vee n = \prod_{p \in \mathcal{P}} p^{\max(v_p(m), v_p(n))}$$

PREUVE. Comme indiqué en cours d'énoncé, il s'agit ici de traduire avec la définition de valuation les propriétés de la décomposition en facteurs premiers d'un entier. ☕

16.8 Epilogue

16.8.1 Compléments historiques

- Nombres de Mersenne : nombres premiers de la forme $2^n - 1$. Exercice : $2^n - 1 \in \mathcal{P} \implies n \in \mathcal{P}$; réciproque (malheureusement !) fausse puisque : $2^{11} - 1 = 23 \times 89$). Le plus grand nombre premier connu (en mars 2018) est un nombre de Mersenne ($2^{74\,207\,281} - 1$).

- Nombres de Carmichael : ce sont les entiers naturels N non premiers tel que pour tout entier a on a : $a^N \equiv a \pmod{N}$. 561 est le plus petit d'entre eux, puis viennent 1105, 1729, 2465... C'est un bon exercice de programmation d'essayer de les trouver, en un temps "raisonnable". Il existe une infinité de nombres de Carmichael (1994, Alford, Granville, Pomerance).
- Polynôme de Sato-Jones-Wada-Wiens : polynôme P de degré 25 en 26 variables à coefficients dans $\mathbb{Z}$ tel que $P(\mathbb{N}^{26}) \cap \mathbb{N} = \mathcal{P} \dots$

16.8.2 Nombres de Carmichael

DÉFINITION 16.9 - Un entier n est dit **de Carmichael** si :

- 1/ $n \notin \mathcal{P}$
- 2/ $\forall m \in \mathbb{Z}, m^n \equiv m \pmod{n}$

En d'autres termes, un entier de Carmichael est un entier fournissant un contre-exemple à la réciproque du petit théorème de Fermat.

FAIT 1. 561 est un nombre de Carmichael : $[561 \notin \mathcal{P}]$ et $[\forall n \in \mathbb{Z}, n^{561} \equiv n \pmod{561}]$.

Soit n un entier. Observons que $561 = 3 \times 11 \times 17$ et que 3, 11 et 17 sont premiers. On peut donc appliquer le petit théorème de Fermat à chacun de ces trois entiers pour obtenir :

- Modulo 3 : $n^{561} \equiv n^{187} \equiv n^{63} \equiv n^{21} \equiv n^7 \equiv n^3 \equiv n$
- Modulo 11 : $n^{561} \equiv n^{51} \equiv n^{11} \equiv n$
- Modulo 17 : $n^{561} \equiv n^{33} \equiv n^{17} \equiv n$

Puisque 3 et 11 divisent $n^{561} - n$ (d'après ce qui précède), et que 3 et 11 sont premiers entre eux, on en déduit par le lemme de Gauss que $n^{561} - n$ est multiple de 33.

Puisque 33 et 17 divisent $n^{561} - n$ (d'après ce qui précède), et que 33 et 17 sont premiers entre eux, une nouvelle application du lemme de Gauss permet d'affirmer que $n^{561} - n$ est multiple de 561.

Conclusion : $\forall n \in \mathbb{Z}, n^{561} \equiv n \pmod{561}$

La propriété du petit théorème de Fermat est donc vérifiée pour l'entier 561. En revanche, l'assertion du corollaire (du petit Fermat) ne le sera en général pas. En effet, pour passer de la relation $n^p \equiv n \pmod{p}$ à $n^{p-1} \equiv 1 \pmod{p}$, il est nécessaire que n soit inversible modulo p . Lorsque p est premier, ceci équivaut à dire que n n'est pas multiple de p ; ce n'est plus vrai si l'on remplace p par 561. Il ne suffit plus que n ne soit pas multiple de 561 pour être premier avec 561 (considérer l'entier 6 par exemple). Illustrations en Python :

```

1  # La fonction ListPuisMod(N, n) ci-dessous retourne la liste des a^n modulo N,
2  # avec a compris entre 0 et (N-1)
3
4  def ListPuisMod(N,n):
5      L=[]
6      for k in range(N):
7          L=L+[(k**n)%N]
8      return L

```

Avec 19 (qui est premier), on a $n^{19} \equiv n [19]$ pour tout n (petit Fermat), comme le confirment les calculs rassurants ci-dessous :

```
1 >>>ListPuisMod(19,19)
2 [ 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18 ]
```

Et toujours avec 19, on a $n^{18} \equiv 1 [19]$ pour tout n non multiple de 19 (corollaire du petit Fermat), comme le confirment encore les calculs ci-dessous :

```
1 >>>ListPuisMod(19,18)
2 [ 0, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1 ]
```

Avec 561 (qui n'est premier, mais qui est de Carmichael), on a $n^{561} \equiv n [561]$ pour tout n (démontré ci-dessus), comme le confirment les calculs suivants :

```
1 >>>ListPuisMod(561,561)
2 [ 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22,
3 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43,
4 ... 535, 536, 537, 538, 539, 540, 541, 542, 543, 544, 545, 546, 547, 548, 549, 550,
5 551, 552, 553, 554, 555, 556, 557, 558, 559, 560 ]
```

Mais avec 561, on n'a plus en général $n^{560} \equiv 1 [561]$ pour n non multiple de 561, comme l'illustrent les calculs ci-dessous (on observera que les n tels que $n^{560} \not\equiv 1 [561]$ sont les multiples de 3, de 11 et de 17) :

```
1 >>>ListPuisMod(561,560)
2 [ 0, 1, 1, 375, 1, 1, 375, 1, 1, 375, 1, 154, 375, 1, 1, 375, 1, 34, 375, 1, 1, 375,
3 154, 1, 375, 1, 1, 375, 1, 1, 375, 1, 1, 528, 34, 1, 375, 1, 1, 375, 1, 1, 375, 1,
4 154, ... 1, 1, 375, 1, 1, 375, 1, 1, 375, 1, 154, 375, 1, 1, 375, 34, 1, 375, 1, 1,
5 375, 154, 1, 375, 1, 1, 375, 1, 1, 375, 1, 1 ]
```

Pour être complet, on donne ci-dessous des fonctions écrites en Python permettant de déterminer si un entier est de Carmichael ou non.

```
1 # TEST DE PRIMALITE
2
3 def TPRIM(n):
4     BOOL = True
5     k=2
6     borne = sqrt(n)
7     while (k<=borne) and BOOL:
8         reste = n%k
9         if reste ==0:
10             BOOL =not BOOL
11         k =k+1
12     return BOOL
```

```

1  # TEST DE "CARMICHAELITE"...
2
3  def CARMIORNOT(n):
4      BOOL = True
5      if TPRIM(n):
6          BOOL = not BOOL
7      if BOOL:
8          k=2
9          reste = 0
10         while BOOL and k<n:
11             reste = (-k+k**n)%n
12             if reste != 0:
13                 BOOL = not BOOL
14             k = k + 1
15     return BOOL

1  # CREATION DE LA LISTE DES ENTIERS DE CARMICHAEL INFERIEURS A N
2
3  from time import *
4
5  def ListeCARMI(N):
6      k = 100
7      Liste = []
8      t1 = time()
9      while k < N+1:
10         if CARMIORNOT(k):
11             Liste = Liste + [k]
12         k = k+1
13     t2 = time()
14     return Liste, t2-t1 # renvoie la liste voulue, et affiche la durée d'exécution

```

16.8.3 Codage RSA

Le problème ci-dessous (issu d'un ancien DS) décrit le principe de ce que l'on appelle le codage RSA (du nom de ses inventeurs Rivest, Shamir et Adleman), qui est une méthode très utilisée en cryptographie pour coder les données numériques. L'intérêt de cette méthode de codage est qu'elle est très difficile à "pirater" en un temps raisonnable (même équipé d'un ordinateur performant !)..

ENONCÉ

1) Congruences et exponentiation rapide

- Déterminer le PGCD et les coefficients de Bezout des entiers 27 et 112.
- On pose $a = 12$. Calculer a^2 , puis a^4 , puis a^8 et enfin a^{16} modulo 145.
- Déduire de ce qui précède la valeur de 12^{27} modulo 145.

2) Questions de cours

- Rappeler la définition d'entiers premiers entre eux. Rappeler la définition de nombre premier.

- b) Soient a et b deux entiers. A quelle condition sur a et b l'entier a est-il inversible modulo b ? Donner, en justifiant brièvement votre réponse, un inverse de 27 modulo 112, ainsi qu'un inverse de 112 modulo 27.
- c) Etablir que si p et q sont deux nombres premiers distincts, alors ils sont premiers entre eux.
- d) Soit p un nombre premier. Le petit théorème de Fermat affirme que : $\forall a \in \mathbb{Z}, a^p \equiv a [p]$.
Redémontrer alors que : $\forall a \in \mathbb{Z} \setminus p\mathbb{Z}, a^{p-1} \equiv 1 [p]$.
- 3) **Une nouvelle conséquence du théorème de Fermat.** Soient p et q deux nombres premiers distincts, et soit $N = pq$.
- a) Montrer que a est premier avec N si et seulement si a est premier avec p et avec q .
- b) On suppose que a est premier avec N . Montrer que $a^{(p-1)(q-1)} \equiv 1 [p]$ et $a^{(p-1)(q-1)} \equiv 1 [q]$.
- c) Dédurre de la question précédente que $a^{(p-1)(q-1)} \equiv 1 [N]$.
- 4) On pose $\varphi(N) = (p-1)(q-1)$. Soit e un entier naturel premier avec $\varphi(N)$. Justifier qu'il existe un entier d tel que $ed \equiv 1 [\varphi(N)]$.
- 5) Soit m un entier naturel.
- a) Montrer que si m est premier avec N , alors $m^{ed} \equiv m [N]$.⁴
- b) Montrer que pour tout entier m , on a $m^{ed} \equiv m [N]$.
- c) **“Vérification”**. Dans cette question, on prend $p = 5$, $q = 29$, $m = 12$ et $e = 27$. Donner sans justification la valeur de d . Déterminer la valeur de m^e modulo N , puis détailler le calcul de $(m^e)^d$ modulo N .

CORRIGÉ

- 1) a) On applique l'algorithme d'Euclide aux entiers 27 et 112.

$$112 = 27 \times 4 + 4$$

$$27 = 4 \times 6 + 3$$

$$4 = 3 \times 1 + 1$$

$$3 = 3 \times 1 + 0$$

On en déduit déjà que $\boxed{27 \wedge 112 = 1}$.

$$1 = 4 - 3 \times 1$$

$$1 = 4 - \times(27 - 4 \times 6) = 7 \times 4 - 1 \times 27$$

$$1 = 7 \times (112 - 4 \times 27) - 1 \times 27 \text{ soit finalement :}$$

$$\boxed{1 = 7 \times 112 - 29 \times 27}$$

On a donc obtenu une relation de Bezout pour les entiers 27 et 112, c'est-à-dire deux entiers u et v tels que $27u + 112v = 27 \wedge 112$ avec $\boxed{u = -29 \text{ et } v = 7}$.

4. Où e et d sont ceux définis dans la question 4, et $N = pq$ avec p et q premiers distincts.

b) $\boxed{a = 12 \text{ donc } a \equiv 12 \text{ [145]}}; a^2 = 144 \text{ donc } \boxed{a^2 \equiv -1 \text{ [145]}}.$

Il s'ensuit que $\boxed{a^4 \equiv a^8 \equiv a^{16} \equiv 1 \text{ [145]}}.$

c) On peut observer que : $27 = 16 + 8 + 2 + 1$ pour écrire : $a^{27} = a^{16} \times a^8 \times a^2 \times a$ d'où : $a^{27} \equiv 12 \times -1 \text{ [145]}.$
Par suite : $\boxed{12^{27} \equiv -12 \text{ [145]}}.$

2) a) Deux entiers a et b sont premiers entre eux si leur PGCD est égal à 1 (il est équivalent de dire que leurs seuls diviseurs communs sont 1 et -1 , ou encore que leur seul diviseur commun dans $\mathbb{N}$ est 1).

Un entier naturel n est premier s'il admet exactement deux diviseurs dans $\mathbb{N}$: 1 et lui-même.

b) D'après le cours, un entier a est inversible modulo b si et seulement si a et b sont premiers entre eux.

D'après la question 1-a, un inverse de 27 modulo 112 est -29 , et un inverse de 112 modulo 27 est 7.

c) Puisque p (*resp.* q) est premier, il admet exactement 1 et p (*resp.* 1 et q) comme diviseurs dans $\mathbb{N}$. Par conséquent, p et q admettent 1 comme unique diviseur commun dans $\mathbb{N}$. Donc $\boxed{p \text{ et } q \text{ sont premiers entre eux}}.$

d) Soient p un nombre premier, et a un entier tel que p ne divise pas a .

D'après le théorème de Fermat : $a^p \equiv a \text{ [} p \text{]}$. Il est équivalent de dire que : $p | a^p - a$, soit : $p | a(a^{p-1} - 1)$. Comme p ne divise pas a et que p est premier, p est premier avec a (en vertu de la propriété : "tout nombre premier est premier avec tout entier qu'il ne divise pas"). D'après le lemme de Gauss, il s'ensuit que $p | a^{p-1} - 1$, c'est à dire : $\boxed{a^{p-1} \equiv 1 \text{ [} p \text{]}}.$

3) a) Sens direct : supposons a premier avec pq . Alors (théorème de Bezout), il existe deux entiers u et v tels que : $au + vpq = 1$. En écrivant cette relation $au + p(vq) = 1$, on en déduit que a et p sont premiers entre eux ; et en l'écrivant $au + q(vp) = 1$, on en déduit que a et q sont premiers entre eux.

Conclusion : (a premier avec $N = pq$) implique (a premier avec p et q).

Réciproquement : supposons a premier avec p et a premier avec q . Les diviseurs de pq dans $\mathbb{N}$ sont 1, p , q et pq . On sait déjà que p ne divise pas a (a et p étant premiers entre eux) et que q ne divise pas a (a et q étant premiers entre eux). A fortiori, pq ne divise pas a , puisque si tel était le cas p et q diviseraient a . Il s'ensuit que a et pq ont un unique diviseur commun dans $\mathbb{N}$ (qui est 1). Donc a et pq sont premiers entre eux.

Conclusion : (a premier avec p et q) implique (a premier avec $N = pq$).

$\boxed{\text{En résumé : (} a \text{ premier avec } p \text{ et } q \text{) SSI (} a \text{ premier avec } N = pq \text{)}}.$

b) Puisque a est premier avec N , alors d'après la question précédente, a est premier avec p et q .

Par suite : $a^{p-1} \equiv 1 \text{ [} p \text{]}$ et $a^{q-1} \equiv 1 \text{ [} q \text{]}$.

On en déduit que : $a^{(p-1)(q-1)} \equiv (a^{p-1})^{q-1} \equiv 1^{q-1} \equiv 1 \text{ [} p \text{]}$; et symétriquement : $a^{(p-1)(q-1)} \equiv 1 \text{ [} q \text{]}$.

En déduit de ces deux congruences que : $p | a^{(p-1)(q-1)} - 1$ et $q | a^{(p-1)(q-1)} - 1$. Or les entiers p et q sont premiers entre eux, puisque ce sont deux nombres premiers distincts. Donc : $(pq) | a^{(p-1)(q-1)} - 1$, ce qui se traduit par la congruence : $\boxed{a^{(p-1)(q-1)} \equiv 1 \text{ [} pq \text{]}}.$

4) Puisque e est premier avec $\varphi(N)$, il existe d'après le théorème de Bezout deux entiers d et v tels que : $ed + v\varphi(N) = 1$. Cette égalité implique la congruence : $\boxed{ed \equiv 1 \text{ [} \varphi(N) \text{]}}$, d'où la conclusion.

5) a) Par définition des entiers e et d , il existe un entier k tel que : $ed = 1 + k\varphi(N)$, soit encore : $1 + k(p-1)(q-1)$.

On peut ainsi écrire : $m^{ed} = m^{1+k(p-1)(q-1)}$ soit : $m^{ed} = (m^{(p-1)(q-1)})^k \times m$. Or $m^{(p-1)(q-1)} \equiv 1 [pq]$ d'après la question 3-b et puisque m est premier avec N . On en déduit que $\boxed{m^{ed} \equiv m [N]}$.

b) Supposons à présent que m ne soit pas premier avec N . D'après la question 3-a, il revient au même de dire que $p|m$ ou $q|m$. Supposons donc que m soit divisible par p . Alors $m \equiv 0 [p]$ donc $m^{1+(p-1)(q-1)} \equiv 0 [p]$. en particulier $m^{ed} \equiv m [p]$. Si m est également divisible par q , alors le même raisonnement entraîne : $m^{ed} \equiv m [q]$. Et si q ne divise pas m , alors $m^{ed} \equiv m [q]$ (conséquence du théorème de Fermat).

En résumé, dans tous les cas, on a $m^{ed} \equiv m [p]$ et $m^{ed} \equiv m [q]$. Il s'ensuit que p et q divisent $m^{ed} - m$, donc (p et q étant premiers entre eux) pq divise $m^{ed} - m$. D'où $\boxed{\forall m \in \mathbb{Z}, m^{ed} \equiv m [N]}$.

c) Avec les notations et données de l'énoncé : $p-1 = 4$ et $q-1 = 28$ d'où $\varphi(N) = 112$. En outre, on a $e = 27$ et l'entier d recherché doit vérifier $ed \equiv 1 [\varphi(N)]$, c-à-d : $27d \equiv 1 [112]$. D'après la question 1-a), $d = -29$ fait l'affaire, tout comme $d = 83$ (puisque $-29 \equiv 83 [112]$).

En posant $m = 12$, on a : $m^e = 12^{27}$, et on sait d'après la question $12^{27} \equiv -12 [145]$.

Calculons à présent $(m^e)^d$ modulo 145. D'après ce qui précède $(m^e)^d \equiv (-12)^{83} [145]$.

$$\text{Or : } (-12)^{83} \equiv \underbrace{[(-12)^4]^{20}}_{\equiv 1 [145]} (-12)^3 \equiv -12^3 \equiv 12^2 \times (-12) \equiv (-1) \times (-12) \equiv 12 [145]$$

Par suite : $\boxed{(12^{27})^{83} \equiv 12 [145]}$

16.8.4 Théorème de Wilson

PROPRIÉTÉ 16.24 - (Théorème de Wilson). Soit p un entier supérieur ou égal à 2.

$$[p \in \mathcal{P}] \iff [(p-1)! \equiv -1 [p]]$$

PREUVE. Supposons que p soit premier. Si $p = 2$ ou 3 , alors l'égalité $(p-1)! \equiv -1 [p]$ est satisfaite (vérification immédiate). Par la suite, on pourra donc supposer $p \geq 5$.

Le point-clef consiste à voir que tous les entiers compris entre 1 et $p-1$ admettent un unique inverse modulo p , lui aussi compris entre 1 et $p-1$.

Plus précisément, si n est un entier compris entre 1 et $p-1$, alors n et p sont premiers entre eux (caractérisation des nombres premiers), donc n est inversible modulo p . Il existe un entier k tel que $nk \equiv 1 [p]$. Qui à remplacer k par son reste dans la division euclidienne par p , on peut supposer que $k \in \mathbb{N}_{p-1}$. Dans ces conditions, l'entier $k \in \mathbb{N}_{p-1}$ tel que $nk \equiv 1 [p]$ est unique, car s'il existait un second entier $k' \in \mathbb{N}_{p-1}$ tel que $nk' \equiv 1 [p]$, alors on en déduirait que p divise $n(k-k')$, d'où (par le lemme de Gauss) p diviserait $k-k'$. Puisque k et k' appartiennent au même intervalle d'entiers d'amplitude $p-1$, ceci entraîne que $k = k'$.

On a donc établi que : tout entier n compris entre 1 et $p-1$ possède un unique inverse modulo p compris entre 1 et $p-1$.

En outre, exactement deux entiers de $\mathbb{N}_{p-1}$ sont leur propre inverse modulo p : 1 et $p-1$. En effet, l'équation $x^2 \equiv 1 [p]$ ne possède que deux solutions modulo p , qui sont 1 et -1 (car $x^2 \equiv 1 [p]$ SSI $p \mid (x-1)(x+1)$, et p est premier).

On en déduit que :

$$(p-1)! \equiv 1 \times (p-1) \times \prod_{k=2}^{p-2} k \equiv - \prod_{k=2}^{p-2} k [p]$$

Or, d'après la discussion précédente, le produit $\prod_{k=2}^{p-2} k$ est celui de $(p-3)/2$ couples d'entiers qui sont inverses l'un de l'autre modulo p . Par suite : $\prod_{k=2}^{p-2} k \equiv 1 [p]$.

On en déduit que : $(p-1)! \equiv -1 [p]$.

On a ainsi établi que : $[p \in \mathcal{P}] \implies [(p-1)! \equiv -1 [p]]$

Réciproquement, supposons que $(p-1)! \equiv -1 [p]$. Il existe un entier relatif k tel que : $(p-1)! - kp = -1$. Soit d un entier compris entre 2 et $p-1$. Si 2 divisait p , alors d diviserait -1 ce qui est absurde. On en déduit que p est premier avec tout entier compris entre 2 et $p-1$. Par suite, p est premier.

On a ainsi établi que : $[p \in \mathcal{P}] \iff [(p-1)! \equiv -1 [p]]$, d'où l'équivalence du théorème. ☕

16.8.5 Sous-groupes de $(\mathbb{Z}, +)$

PROPRIÉTÉ 16.25 - (Description des sous-groupes additifs de $\mathbb{Z}$).

Soit H un sous-groupe additif de $\mathbb{Z}$.

Il existe un entier naturel n tel que : $H = n\mathbb{Z}$.

PREUVE. Soit H un sous-groupe additif de $\mathbb{Z}$. Par définition de sous-groupe, on a déjà : $0 \in H$.

Si $H = \{0\}$, alors $H = 0\mathbb{Z}$, et la propriété est établie.

Sinon, H contient au moins un élément non nul, que nous noterons m . Puisque H est stable par passage à l'opposé, on peut supposer que $m > 0$.

On introduit alors l'ensemble : $E = \{h \in H, h > 0\}$.

L'ensemble E est une partie non vide (elle contient au moins l'entier m) de $\mathbb{N}^*$. A ce titre, elle possède un plus petit élément. Notons judicieusement :

$$m = \min E$$

Soit à présent h un élément de H . Réalisant la division euclidienne de h par n , on peut affirmer que :

$$\exists! (q, r) \in \mathbb{Z} \times \mathbb{N}, h = nq + r \wedge 0 \leq r < n$$

Supposons que $r \neq 0$. Alors on déduit de la relation $r = \underbrace{h}_{\in H} - \underbrace{nq}_{\in H}$ et du fait que H est un sous-groupe de $\mathbb{Z}$ que $r \in H$.

Par suite, r est un élément de H , strictement positif et strictement inférieur au plus petit élément de E : contradiction.

On en déduit que $r = 0$. D'où : $h = nq$. Ainsi : $\forall h \in H, \exists q \in \mathbb{Z}, h = nq$.

Il revient au même d'écrire : $H \subset n\mathbb{Z}$. Par ailleurs, l'inclusion $H \supset n\mathbb{Z}$ provient immédiatement du fait que n est un élément de H et que H est un sous-groupe additif de $\mathbb{Z}$. Donc : $H = n\mathbb{Z}$.

Conclusion. Pour tout sous-groupe additif H de $\mathbb{Z}$, il existe un entier naturel n tel que $H = n\mathbb{Z}$ ☕

Chapitre 17

Applications de la dérivation

17.1 Généralités

DÉFINITION 17.1 - (Rappel) Soit $f : I \longrightarrow \mathbb{K}$ une fonction, et soit $a \in I$. La fonction f est **dérivable en a** si la limite quand h tend vers 0 du taux d'accroissement de f entre a et $(a + h)$ existe et est finie. Lorsque c'est le cas, on note

$$f'(a) = \lim_{h \rightarrow 0} \frac{f(a+h) - f(a)}{h} \text{ (ou encore } f'(a) = \lim_{h \rightarrow a} \frac{f(x) - f(a)}{x - a} \text{)}$$

DÉFINITION 17.2 - f est **dérivable à droite** de a si $\lim_{h \xrightarrow{h>0} 0} \frac{f(a+h) - f(a)}{h}$ existe et est finie ; f est **dérivable à gauche** de a si $\lim_{h \xrightarrow{h<0} 0} \frac{f(a+h) - f(a)}{h}$ existe et est finie.

PROPRIÉTÉ 17.1 - f est dérivable en a SSI elle est dérivable à droite et à gauche de a et

$$\lim_{h \xrightarrow{h>0} 0} \frac{f(a+h) - f(a)}{h} = \lim_{h \xrightarrow{h<0} 0} \frac{f(a+h) - f(a)}{h} (= f'(a))$$

DÉFINITION 17.3 - Une fonction $f : I \longrightarrow \mathbb{K}$ est **dérivable sur I** si elle est dérivable en tout $a \in I$.

THÉORÈME 17.1 - Soit $f : I \longrightarrow \mathbb{K}$ une fonction, et soit $a \in I$. La fonction f est dérivable en a SSI il existe un scalaire ℓ tel que $f(a+h) = f(a) + h\ell + h\varepsilon(h)$ avec $\lim_{h \rightarrow 0} \varepsilon(h) = 0$.

Remarque. Si f est dérivable en a , on a donc $f(a+h) = f(a) + hf'(a) + h\varepsilon(h)$ (avec $\lim_{h \rightarrow 0} \varepsilon(h) = 0$) ; cette expression est appelée **développement limité à l'ordre 1** en a de f .

PREUVE. On raisonne par double implication.

► **Sens direct** : supposons que f soit dérivable en a .

On écrit, pour tout réel h (tel que $(a+h) \in I$) :

$$f(a+h) = f(a) + hf'(a) + f(a+h) - f(a) - hf'(a)^1$$

1. Diabolique, n'est-ce pas ?

D'où : $f(a+h) = f(a) + hf'(a) + h \left(\frac{f(a+h) - f(a)}{h} - f'(a) \right)$ d'où : $f(a+h) = f(a) + hf'(a) + h\varepsilon(h)$ (♠)

en ayant posé : $\varepsilon(h) = \frac{f(a+h) - f(a)}{h} - f'(a)$.

Or, f étant dérivable en a , on a : $\lim_{h \rightarrow 0} \frac{f(a+h) - f(a)}{h} = f'(a)$ et donc $\lim_{h \rightarrow 0} \varepsilon(h) = 0$.

En résumé, on a établi l'implication :

$$[f \text{ est dérivable en } a] \implies \left[\forall h \in \mathbb{R}, (a+h) \in I, f(a+h) = f(a) + \ell h + h\varepsilon(h) \text{ avec } \lim_{h \rightarrow 0} \varepsilon(h) = 0 \right] \quad (\heartsuit)$$

► **Réciproquement** : supposons qu'il existe un réel ℓ tel que f vérifie : $f(a+h) = f(a) + \ell h + h\varepsilon(h)$ avec $\lim_{h \rightarrow 0} \varepsilon(h) = 0$.

Alors pour tout réel h non nul tel que $(a+h) \in I$, on a : $\frac{f(a+h) - f(a)}{h} = \ell + \varepsilon(h)$.

D'où : $\lim_{h \rightarrow 0} \frac{f(a+h) - f(a)}{h} = \ell$. Ce qui signifie que la fonction f est dérivable en a (et que $f'(a) = \ell$). Ce qui assure que :

$$\left[\forall h \in \mathbb{R}, (a+h) \in I, f(a+h) = f(a) + \ell h + h\varepsilon(h) \text{ avec } \lim_{h \rightarrow 0} \varepsilon(h) = 0 \right] \implies [f \text{ est dérivable en } a] \quad (\clubsuit)$$

Conclusion. f est dérivable en a SSI il existe un réel ℓ et une fonction ε définie au voisinage de zéro tels que :

$$\forall h \in \mathbb{R}, (a+h) \in I, f(a+h) = f(a) + \ell h + h\varepsilon(h) \text{ avec } \lim_{h \rightarrow 0} \varepsilon(h) = 0 \quad \text{☕}$$

COROLLAIRE 17.1 - Si f est dérivable en a , alors f est continue en a .

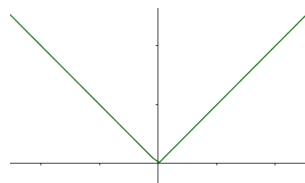
PREUVE. D'après la propriété précédente, si f est dérivable en a , alors pour h au voisinage de 0 :

$$f(a+h) = f(a) + h\ell + h\varepsilon(h) \text{ avec } \lim_{h \rightarrow 0} \varepsilon(h) = 0.$$

Il s'ensuit que $\lim_{h \rightarrow 0} f(a+h) = f(a)$, ce qui prouve la continuité de f en a . ☕

Réciproque archi-fausse : la fonction valeur absolue est

continue en 0, MAIS n'est pas dérivable en 0.



Remarque : c'est ici l'occasion de vous rappeler que nous avons vu en début d'année les dérivées des fonctions usuelles (cf formulaire), les dérivées d'ordre supérieur (définition, linéarité, dérivées n-ièmes "usuelles" et formule de Leibniz), ainsi que les développements limités à l'ordre 1 en 0 des fonctions usuelles, et qu'il serait bon que vous n'ayez pas complètement oublié ces notions.

17.2 Dérivation et extremums — Théorème de Rolle

DÉFINITION 17.4 - Soient I un intervalle de $\mathbb{R}$, $f : I \rightarrow \mathbb{R}$ une fonction à valeurs réelles, et a un réel de I .

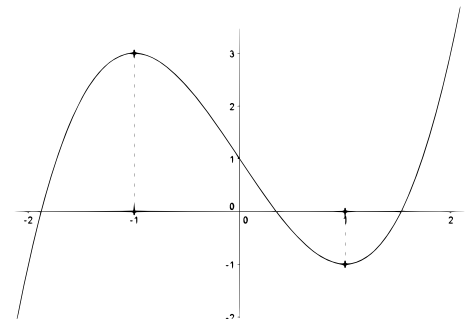
La fonction f admet un **maximum local en a** (*resp.* **minimum local en a**) s'il existe un intervalle ouvert J tel que $a \in J \subset I$, et tel que $f(a)$ soit le maximum (*resp.* minimum) de $f|_J$ (la restriction de f à J).

La fonction f admet un **extremum local en a** si $f(a)$ est un maximum local ou un minimum local de la fonction f .

Exemples.

1/ La fonction cosinus (définie sur $\mathbb{R}$) admet un maximum local en $2k\pi$ pour tout $k \in \mathbb{Z}$; et un minimum local en $(2k+1)\pi$ pour tout $k \in \mathbb{Z}$.

2/ La fonction $f : x \in \mathbb{R} \mapsto (x+1)(x-1)^2 - 1$, représentée sur la figure ci-contre, admet un maximum local en -1 et un minimum local en 1 . Il faut observer que la fonction f ne possède en revanche pas de maximum ni de minimum (global), puisque ses limites en $+\infty$ et $-\infty$ sont $+\infty$ et $-\infty$ respectivement.



PROPRIÉTÉ 17.2 - Soit f une fonction dérivable sur I (intervalle ouvert non vide de $\mathbb{R}$), et soit $c \in I$. Si f admet un extremum local en c , alors $f'(c) = 0$.

PREUVE. Puisque f admet un extremum local en c , il existe un voisinage ouvert de c tel que $f(c)$ soit un extremum de f sur ce voisinage. De plus, f étant dérivable en c , on a :

$$\lim_{h \rightarrow 0^+} \frac{f(c+h) - f(c)}{h} = f'(c) = \lim_{h \rightarrow 0^-} \frac{f(c+h) - f(c)}{h}$$

Or, sur un voisinage ouvert de c , l'expression $\frac{f(c+h) - f(c)}{h}$ possède des signes opposés suivant que h est strictement positif ou strictement négatif. Par stabilité des inégalités large par passage à la limite, on en déduit que : $\lim_{h \rightarrow 0^+} \frac{f(c+h) - f(c)}{h}$ et $\lim_{h \rightarrow 0^-} \frac{f(c+h) - f(c)}{h}$ sont de signes opposés. Par conséquent :

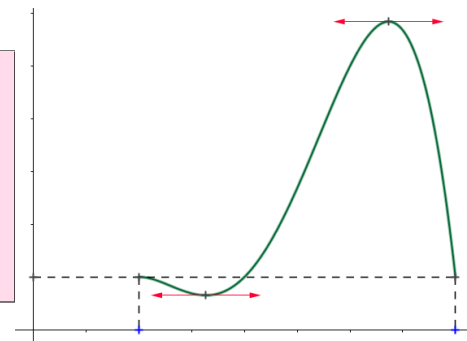
$f'(c) = 0$. ☕

Remarque. La réciproque est fausse ($x \mapsto x^3 \dots$), et l'énoncé n'est plus valide sur segment.

THÉORÈME 17.2 - (Théorème de Rolle).

Si f est dérivable sur $]a; b[$, continue sur $[a, b]$, et $f(a) = f(b)$ alors

$$\exists c \in]a; b[, f'(c) = 0$$



PREUVE. La fonction f étant continue sur le segment $[a, b]$, elle est bornée et atteint ses bornes sur $[a, b]$ (théorème des bornes atteintes). Notons m son minimum et M son maximum.

➤ Si $m = M$: alors f est constante sur $[a, b]$. Par suite : $\forall c \in]a, b[, f'(c) = 0$.

➤ Si $m < M$: distinguons alors deux sous-cas : soit $f(a) = m$, soit $f(a) \neq m$.

➢ Si $f(a) = m$: alors on a aussi $f(b) = m$ (par hypothèse), donc le maximum M de f est atteint en un réel c de $]a, b[$ (puisque $m \neq M$), et on a donc $f'(c) = 0$.

➢ Si $f(a) \neq m$: alors on a aussi $f(b) \neq m$ (par hypothèse), donc le minimum m de f est atteint en un réel c de $]a, b[$, et on a donc $f'(c) = 0$.

Conclusion. Dans tous les cas : $\exists c \in]a, b[, f'(c) = 0$.



Exemple d'application : si $f : \mathbb{R} \rightarrow \mathbb{R}$ est dérivable sur $\mathbb{R}$, et s'annule (au moins) n fois sur $\mathbb{R}$, alors f' s'annule (au moins) $n - 1$ fois sur $\mathbb{R}$.

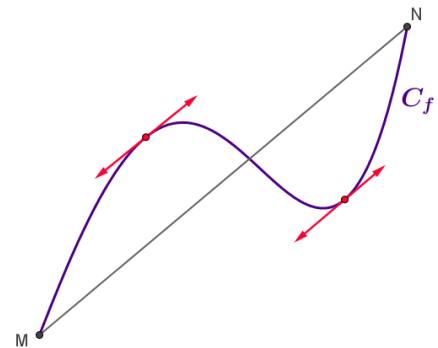
17.3 Dérivation et sens de variation — Théorème des accroissements finis

THÉORÈME 17.3 - (Théorème des accroissements finis, “TAF”). Si $f : [a, b] \rightarrow \mathbb{R}$ est dérivable sur $]a, b[$, continue sur $[a, b]$, alors

$$\exists c \in]a, b[, f'(c) = \frac{f(b) - f(a)}{b - a}.$$

Interprétation graphique : il existe au moins un réel $c \in]a, b[$ tel que la tangente à $\mathcal{C}_f$ au point d'abscisse c soit parallèle à la droite joignant les points de $\mathcal{C}_f$ d'abscisses a et b .

Interprétation génératrice de torticolis : le TAF est le “théorème de Rolle en pente”.



Interprétation cinématique : sous réserve que vous vous déplaçiez “gentiment” (= “de façon dérivable”) en voiture, il existe au moins un instant où votre vitesse instantanée aura été égale à votre vitesse moyenne sur le parcours.

PREUVE. On définit une fonction g sur $[a, b]$ en posant : $g(x) = f(x) - \frac{f(b) - f(a)}{b - a} (x - a)$.²

D'après les théorèmes généraux sur la continuité et la dérivabilité respectivement, la fonction g est continue sur $[a, b]$ et dérivable sur $]a, b[$.

En outre $g(a) = f(a) - \frac{f(b) - f(a)}{b - a} (a - a)$ d'où $g(a) = f(a)$; et $g(b) = f(b) - \frac{f(b) - f(a)}{b - a} (b - a)$ d'où : $g(b) = f(a)$.

2. Informellement, cette fonction g est destinée à “redresser la situation”, en ôtant à f ce qui l'empêche de prendre les mêmes valeurs en a et en b ; la motivation étant de se ramener à la situation du théorème de Rolle.

La fonction g vérifie les hypothèses du théorème de Rolle, dont l'application donne : $\exists c \in]a, b[, g'(c) = 0$.

Or : $\forall x \in]a, b[, g'(x) = f'(x) - \frac{f(b) - f(a)}{b - a}$.

On en déduit : $\exists c \in]a, b[, f'(c) - \frac{f(b) - f(a)}{b - a} = 0$.

D'où finalement : $\boxed{\exists c \in]a, b[, f'(c) = \frac{f(b) - f(a)}{b - a}}$. ☕

Le théorème des accroissements finis a de nombreuses applications pratiques, dont voici un échantillon.

COROLLAIRE 17.2 - Soit $f : I \rightarrow \mathbb{R}$ une fonction dérivable sur un intervalle I .

1/ f est croissante sur I SSI $f' \geq 0$ sur I

2/ f est décroissante sur I SSI $f' \leq 0$ sur I

3/ f est constante sur I SSI $f' = 0$ sur I

PREUVE. Établissons le point 1/.

► ($\Rightarrow$) — Supposons f croissante sur I .

Soit a un réel de l'intervalle I . Puisque f est dérivable sur I , elle l'est en particulier en a , et : $f'(a) = f'_d(a)$.

En d'autres termes : $f'(a) = \lim_{h \rightarrow 0} \frac{f(a+h) - f(a)}{h}$.

Or, pour tout réel $h > 0$: $f(a+h) - f(a) \geq 0$ (puisque f est supposée croissante).

Donc : $\forall h > 0, \frac{f(a+h) - f(a)}{h} \geq 0$.

Par stabilité des inégalités larges par passage à la limite, on en déduit que : $\lim_{h \rightarrow 0} \frac{f(a+h) - f(a)}{h} \geq 0$,

d'où $f'_d(a) \geq 0$ d'où $f'(a) \geq 0$.

Le réel a de I étant arbitraire dans le raisonnement ci-dessus, on peut conclure que : $\forall a \in I, f'(a) \geq 0$.

En d'autres termes, f' est positive sur I . On a donc : $\boxed{[f \text{ croissante sur } I] \Rightarrow [f' \text{ positive sur } I]}$

► ($\Leftarrow$) — Réciproquement, supposons f' positive sur I .

Soient a et b deux réels dans I , avec $a < b$. Puisque f est dérivable sur I , on peut lui appliquer le théorème des accroissements finis³ sur $[a; b]$ et affirmer que :

$$\exists c \in]a; b[, f'(c) = \frac{f(b) - f(a)}{b - a} \quad \text{d'où : } \exists c \in]a; b[, f(b) - f(a) = f'(c) \times (b - a)$$

Comme $f'(c)$ est positif par hypothèse, et que par ailleurs $b \geq a$ toujours par hypothèse, on en déduit que $f(b) - f(a) \geq 0$.

En résumé, on a établi l'implication : $[b > a] \Rightarrow [f(b) \geq f(a)]$. Donc f est croissante sur I .

On a donc : $\boxed{[f' \text{ positive sur } I] \Rightarrow [f \text{ croissante sur } I]}$

3. Puisqu'en particulier f est dérivable sur $]a, b[$ et continue sur $[a, b]$.

Conclusion. Pour f dérivable sur I , on a : $[f']$ positive sur $I] \iff [f$ croissante sur $I]$.

Les deux autres points s'en déduisent immédiatement. ☕

COROLLAIRE 17.3 - (Inégalité des accroissements finis). Si $f : I \longrightarrow \mathbb{R}$ est dérivable sur I , et s'il existe un réel K tel que $|f'| \leq K$ sur I , alors :

$$\forall (a, b) \in I^2, |f(b) - f(a)| \leq K |b - a|$$

Interprétation cinématique : si vous roulez pendant une heure en ne dépassant jamais la vitesse de 100km/h, vous parcourrez moins de 100 kilomètres !

Remarque : la condition " $|f'| \leq K$ sur I " est en particulier vérifiée lorsque I est un segment et que f est de classe $\mathcal{C}^1$.

PREUVE. Sous les hypothèses de l'énoncé, il existe un réel c compris entre a et b tel que : $f(b) - f(a) = f'(c) \times (b - a)$. Le passage aux valeurs absolues et la majoration $|f'| \leq K$ fournissent la conclusion. ☕

Application : $\forall x \in \mathbb{R}, |\sin(x)| \leq |x|$.

Une ultime conséquence du TAF :

THÉORÈME 17.4 - (Théorème de prolongement de la dérivée). Soit $f : I \longrightarrow \mathbb{R}$ une fonction, et $a \in I$. On suppose f continue sur I , dérivable sur $I \setminus \{a\}$, et qu'il existe $\ell \in \mathbb{R}$ tel que $\lim_{x \rightarrow a} f'(x) = \ell$.

Alors : $\lim_{x \rightarrow a} \frac{f(x) - f(a)}{x - a} = \ell$. En particulier, f est dérivable et $f'(a) = \ell$.

PREUVE. Soient f , a et ℓ comme dans l'énoncé. Soit x un réel de I . On peut appliquer le TAF à la fonction f entre a et x (les hypothèses sont satisfaites d'après l'énoncé) pour affirmer qu'il existe un réel c_x compris entre a et x tel que : $f'(c_x) = \frac{f(x) - f(a)}{x - a}$.

Lorsque x tend vers a , le réel c_x tend également vers a (par encadrement) d'où : $\lim_{x \rightarrow a} \frac{f(x) - f(a)}{x - a} = \lim_{x \rightarrow a} f'(c_x) = \ell$. Ce qui achève la preuve du théorème. ☕

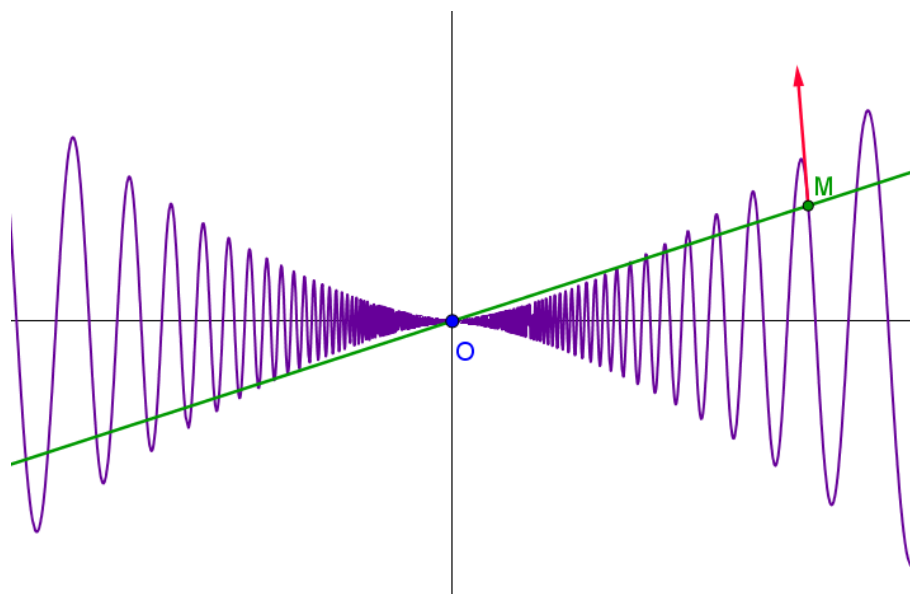
Remarque. Dans un certain sens, la "réciproque" du théorème ci-dessus ne tient pas. Il se peut que f soit dérivable en 0 sans pour autant que f' admette une limite finie en 0 ; c'est le cas par exemple pour $f : x \longmapsto |x|^\alpha \sin(1/x)$ (avec $\alpha > 1$).

La figure ci-dessous illustre cette affirmation⁴. Sur la courbe représentative de cette fonction, on choisit un point M d'abscisse a .

Lorsque M se rapproche de O (lorsque a tend vers 0), les oscillations de la droite (OM) deviennent de plus en plus faibles. De fait la pente de la droite (OM) , qui est le taux d'accroissement de f entre O et M , tend vers 0 lorsque a tend vers 0.

En revanche, toujours lorsque a tend vers 0, la tangente à la courbe représentative de f au point M se comporte de plus en plus sauvagement ; f' n'a pas de limite en 0.

4. Dont la démonstration, plutôt aisée, est un gentil petit exercice pour vérifier que l'on sait établir qu'une fonction est dérivable en un point.



17.4 Lipschitzianité

DÉFINITION 17.5 - Une fonction $f : I \longrightarrow \mathbb{R}$ est **lipschitzienne** sur I s'il existe un réel positif K tel que :

$$\forall (x, y) \in I^2, |f(y) - f(x)| \leq K |y - x|$$

On peut préciser dans cette situation que f est K -lipschitzienne.

Exemples : 1) Les fonctions cos, sin et arctan sont 1-lipschitziennes sur $\mathbb{R}$.

2) La fonction $x \longmapsto x^2$ est 2-lipschitzienne sur $[0, 1]$, et 6-lipschitzienne sur $[0, 3]$.

3) La fonction $x \longmapsto x^2$ n'est pas lipschitzienne sur $\mathbb{R}$; en effet, pour tout entier naturel n , on a $(n+1)^2 - n^2 = 2n+1 \longrightarrow +\infty$ lorsque n tend vers $+\infty$. Il ne peut donc exister aucun réel K tel que pour tout n on ait : $(n+1)^2 - n^2 \leq K$.

Etudions à présent les liens entre la notion de lipschitzianité, et les propriétés plus familières des fonctions (continuité et dérivabilité).

PROPRIÉTÉ 17.3 - Si $f : I \longrightarrow \mathbb{R}$ est lipschitzienne sur I , alors f est continue sur I .

PREUVE. Soit f une fonction lipschitzienne sur I : il existe un réel $k > 0$ tel que : $\forall (x, y) \in I^2, |f(y) - f(x)| \leq k |y - x|$.

Soit a un réel dans I , et soit $\varepsilon > 0$. On pose : $\alpha = \varepsilon/k$. On a alors : $\forall x \in I, [|x - a| < \alpha] \implies [|f(x) - f(a)| < k \times \varepsilon/k]$.

Ainsi : $\forall a \in I, \forall \varepsilon > 0, \exists \alpha > 0, \forall x \in I, [|x - a| < \alpha] \implies [|f(x) - f(a)| < \varepsilon]$. Ce qui signifie que f est continue sur I , ce qu'il fallait démontrer. ☕

On a vu plus haut que la réciproque ne tient pas (exemple de la fonction “carré” sur $\mathbb{R}_+$).

En revanche, d’après l’inégalité des accroissements finis :

PROPRIÉTÉ 17.4 - Si $f : I \rightarrow \mathbb{R}$ est dérivable sur I , et s’il existe un réel K tel que $|f'| \leq K$ sur I , alors f est K -lipschitzienne sur I .

PREUVE. Conséquence immédiate de l’inégalité des accroissements finis. ☕

Et d’ailleurs plus généralement :

PROPRIÉTÉ 17.5 - Soit $f : I \rightarrow \mathbb{R}$ une fonction dérivable sur I .

La fonction f est lipschitzienne sur I SSI f' est bornée sur I .

PREUVE. PREUVE. ► Supposons f lipschitzienne sur I . Alors il existe un réel positif k tel que : $\forall (x, y) \in I^2, |f(y) - f(x)| \leq k|y - x|$. Soit a un réel quelconque dans I . Pour tout réel h raisonnable⁵, on a :

$$\left| \frac{f(a+h) - f(a)}{h} \right| \leq \frac{k|h|}{|h|} \quad \text{d'où :} \quad \left| \frac{f(a+h) - f(a)}{h} \right| \leq k$$

Par stabilité des inégalités larges par passage à la limite, on en déduit que :

$$\lim_{h \rightarrow 0} \left| \frac{f(a+h) - f(a)}{h} \right| \leq k \quad \text{d'où :} \quad |f'(a)| \leq k$$

Le réel $a \in I$ étant arbitraire dans le raisonnement précédent, on a établi que : $\forall a \in I, |f'(a)| \leq k$. Ce qui signifie que f' est bornée sur I .

Par suite : $[f \text{ lipschitzienne sur } I] \implies [f' \text{ bornée sur } I] \quad (\spadesuit)$.

► Réciproquement, supposons f' bornée sur I . Il existe un réel M tel que : $|f'| \leq M$ sur I .

Soient x et y deux réels distincts dans I . La fonction f étant dérivable sur I , elle est continue sur le segment $[x, y]$ (ou sur le segment $[y, x]$), et dérivable sur l’intervalle $]x, y[$ (ou sur l’intervalle $]y, x[$). On peut donc appliquer le théorème des accroissements finis à la fonction f entre x et y :

il existe un réel c compris entre x et y tel que : $f(y) - f(x) = f'(c) \times (y - x)$

$\implies$ il existe un réel c compris entre x et y tel que : $|f(y) - f(x)| = |f'(c)| \times |y - x|$

$\implies$ il existe un réel c compris entre x et y tel que : $|f(y) - f(x)| \leq M \times |y - x|$

L’inégalité obtenue ci-dessus restant évidemment valide dans le cas particulier où $x = y$, on peut conclure que : $\forall (x, y) \in I^2, |f(y) - f(x)| \leq M|y - x|$. Ce qui signifie que f est M -lipschitzienne sur I .

Par suite : $[f \text{ lipschitzienne sur } I] \iff [f' \text{ bornée sur } I] \quad (\clubsuit)$.

► **Conclusion.** D’après $(\spadesuit)$ et $(\clubsuit)$: $[f \text{ lipschitzienne sur } I] \iff [f' \text{ bornée sur } I]$ ☕

En continuant sur notre lancée, on a l’ultime :

PROPRIÉTÉ 17.6 - Si $f : I \rightarrow \mathbb{R}$ est de classe $\mathcal{C}^1$ sur I , alors pour tout couple $(a, b) \in I^2$, la fonction f est lipschitzienne sur $[a, b]$ (et la constante de Lipschitz est alors : $\max_{x \in [a, b]} |f'(x)|$).

PREUVE. Conséquence de l’inégalité des accroissements finis et du TBA appliqué à f' . ☕

5. C’est à dire non nul et tel que $(a + h) \in I$.

17.5 Extension aux fonctions à valeurs complexes

Pour la définition de dérivabilité, aucun changement par rapport au cas réel.

DÉFINITION 17.6 - Soit $f : I \rightarrow \mathbb{C}$ une fonction, et soit $a \in I$. La fonction f est **dérivable en** a si la limite quand h tend vers 0 du taux d'accroissement de f entre a et $(a + h)$ existe et est finie. Lorsque c'est le cas, on note

$$f'(a) = \lim_{h \rightarrow 0} \frac{f(a+h) - f(a)}{h} \quad (\text{ou encore } f'(a) = \lim_{h \rightarrow a} \frac{f(x) - f(a)}{x - a})$$

Le point-clef pour éviter de tout reprendre depuis le début est de faire le lien entre les fonctions à valeurs réelles et celles à valeurs complexes : c'est l'objet de l'énoncé ci-dessous.

THÉORÈME 17.5 - (“Pont $\mathbb{R} \longleftrightarrow \mathbb{C}$ ” pour la dérivabilité).

Soit $f : I \rightarrow \mathbb{C}$ une fonction, et soit $a \in I$.

f est dérivable en a SSI $\operatorname{Re}(f)$ et $\operatorname{Im}(f)$ sont dérivables en a

Lorsque tel est le cas, on a : $f'(a) = (\operatorname{Re}(f))'(a) + i (\operatorname{Im}(f))'(a)$.

PREUVE. Conséquence de la définition de dérivabilité, et du théorème du “pont $\mathbb{R} \longleftrightarrow \mathbb{C}$ ” pour les limites (chapitre 14). ☕

Remarque : si f est dérivable en a , on a donc $(\operatorname{Re}(f))'(a) = \operatorname{Re}(f'(a))$ et $(\operatorname{Im}(f))'(a) = \operatorname{Im}(f'(a))$.

Par ailleurs, il résulte de ce théorème que les propriétés algébriques des dérivées (notamment d'une somme, d'un produit...) continueront d'être valables dans le cas complexe. En revanche, les propriétés faisant intervenir le sens de variation ou la notion d'extremum local ne pourront évidemment pas s'étendre, puisque cela n'a aucun sens de parler des variations d'une fonction à valeurs dans $\mathbb{C}$. On dispose néanmoins d'une variante de l'inégalité des accroissements finis, explicitement :

PROPRIÉTÉ 17.7 - Soit $f : I \rightarrow \mathbb{C}$ une fonction dérivable sur I . S'il existe un réel K tel que $|f'| \leq K$ sur I , alors : $\forall (a, b) \in I^2, |f(b) - f(a)| \leq K |b - a|$.

PREUVE. Il suffit d'appliquer l'inégalité des accroissements finis aux fonctions parties réelle et imaginaire de f . ☕

17.6 Epilogue - Nature des séries de Riemann

PROPRIÉTÉ 17.8 - (Convergence des séries de Riemann). La suite $\left(\sum_{n=1}^N \frac{1}{n^\alpha}\right)_{N \in \mathbb{N}^*}$ converge pour $\alpha > 1$, et diverge vers $+\infty$ pour $\alpha \leq 1$.

PREUVE. On suppose $\alpha > 0$ et $\alpha \neq 1$.

La fonction $f : x \in \mathbb{R}_+^* \mapsto e^{(1-\alpha)\ln x}$ est dérivable sur $\mathbb{R}_+^*$ (théorèmes généraux), et :

$$\forall x \in \mathbb{R}_+^*, f'(x) = \frac{1-\alpha}{x} e^{(1-\alpha)\ln x} = \frac{1-\alpha}{x} x^{1-\alpha} = \frac{1-\alpha}{x^\alpha}.$$

Il s'ensuit que le signe de $f'(x)$ sur $\mathbb{R}_+^*$ est exactement celui de $1-\alpha$. Par suite, si $0 < \alpha < 1$, alors f est croissante sur $\mathbb{R}_+^*$; et si $\alpha > 1$, alors f est décroissante sur $\mathbb{R}_+^*$.

► La fonction f est dérivable sur $\mathbb{R}_+^*$, donc elle l'est en particulier sur l'intervalle $[n; n+1]$ pour tout entier naturel non nul n . On peut donc lui appliquer le théorème des accroissements finis sur $[n; n+1]$, et ce faisant on obtient :

$$\exists c \in]n; n+1[, \quad \frac{1-\alpha}{c^\alpha} = \frac{(n+1)^{1-\alpha} - n^{1-\alpha}}{(n+1) - n} \iff \exists c \in]n; n+1[, \quad \frac{1}{c^\alpha} = \frac{(n+1)^{1-\alpha} - n^{1-\alpha}}{1-\alpha} \quad (\spadesuit)$$

Soit $n \in \mathbb{N}^*$. Pour tout $c \in]n; n+1[$, on a : $\frac{1}{n^\alpha} \geq \frac{1}{c^\alpha} \geq \frac{1}{(n+1)^\alpha}$ puisque la fonction puissance $x \mapsto 1/x^\alpha$ est décroissante quand α est strictement positif. En ne conservant que l'inégalité de gauche de cet encadrement, et en utilisant la relation $(\spadesuit)$, on obtient :

$$\forall \alpha \in]0; 1[, \forall n \in \mathbb{N}^*, \frac{1}{n^\alpha} \geq \frac{(n+1)^{1-\alpha} - n^{1-\alpha}}{1-\alpha}$$

D'après ce qui précède : $\forall \alpha \in]0; 1[, \forall N \in \mathbb{N}^*, \sum_{n=1}^N \frac{1}{n^\alpha} \geq \frac{1}{1-\alpha} \sum_{n=0}^N ((n+1)^{1-\alpha} - n^{1-\alpha})$, c'est à dire, la

somme de droite étant télescopique : $\forall \alpha \in]0; 1[, \forall N \in \mathbb{N}^*, \sum_{n=1}^N \frac{1}{n^\alpha} \geq \frac{1}{1-\alpha} ((N+1)^{1-\alpha} - 1)$

Si $0 < \alpha < 1$, on a : $\lim_{N \rightarrow +\infty} (N+1)^{1-\alpha} = +\infty$, d'où : $\lim_{N \rightarrow +\infty} \frac{1}{1-\alpha} ((N+1)^{1-\alpha} - 1) = +\infty$.

On déduit de cette limite, de l'inégalité précédente et du théorème de comparaison que :

$$\forall \alpha \in]0; 1[, \lim_{N \rightarrow +\infty} \sum_{n=1}^N \frac{1}{n^\alpha} = +\infty$$

► Soit $n \in \mathbb{N}^*$. Pour tout $c \in]n; n+1[$, on a : $\frac{1}{n^\alpha} \geq \frac{1}{c^\alpha} \geq \frac{1}{(n+1)^\alpha}$. On en déduit dans un premier temps que :

$$\forall \alpha \in]1; +\infty[, \forall n \in \mathbb{N}^*, \frac{1}{(n+1)^\alpha} \leq \frac{(n+1)^{1-\alpha} - n^{1-\alpha}}{1-\alpha}$$

$$\text{d'où : } \forall \alpha \in]1; +\infty[, \forall n \in \mathbb{N}^*, \frac{1}{(n+1)^\alpha} \leq \frac{n^{1-\alpha} - (n+1)^{1-\alpha}}{\alpha-1}$$

$$\text{et donc : } \forall \alpha \in]1; +\infty[, \forall N \in \mathbb{N}^*, \sum_{n=1}^N \frac{1}{(n+1)^\alpha} \leq \sum_{n=1}^N \frac{n^{1-\alpha} - (n+1)^{1-\alpha}}{\alpha-1}$$

$$\text{soit : } \forall \alpha \in]1; +\infty[, \forall N \in \mathbb{N}^*, \sum_{n=1}^N \frac{1}{(n+1)^\alpha} \leq \frac{1}{\alpha-1} (1 - (N+1)^{1-\alpha})$$

Or on a cette fois, puisque $\alpha > 1$: $\lim_{N \rightarrow +\infty} \frac{1}{\alpha-1} (1 - (N+1)^{1-\alpha}) = \frac{1}{\alpha-1}$.

On en déduit que la suite de terme général $\sum_{n=1}^N \frac{1}{(n+1)^\alpha}$ converge puisqu'elle est croissante⁶ et majorée (nous venons de l'établir).

Il ne reste plus qu'à remarquer que : $\forall N \in \mathbb{N}^*$, $u_N = \sum_{n=1}^N \frac{1}{n^\alpha} = \left[\sum_{n=1}^N \frac{1}{(n+1)^\alpha} \right] + 1 - \frac{1}{(N+1)^\alpha}$, ce qui assure que la suite de terme général u_N converge en tant que somme de suites convergentes. On peut même préciser, grâce à cette égalité que sa limite est comprise entre 1 et $1 + (\alpha-1)^{-1}$.

Conclusion : pour $\alpha > 1$, (u_N) converge et $\lim_{N \rightarrow +\infty} \sum_{n=1}^N \frac{1}{n^\alpha} = \ell$ (avec $1 \leq \ell \leq 1 + \frac{1}{\alpha-1}$). ☕

► En résumé, on a établi que :

► Si $\alpha \leq 1$: $\lim_{N \rightarrow +\infty} \left(\sum_{n=1}^N \frac{1}{n^\alpha} \right) = +\infty$

► Si $\alpha > 1$: $\lim_{N \rightarrow +\infty} \left(\sum_{n=1}^N \frac{1}{n^\alpha} \right) = \ell$ (avec $1 \leq \ell \leq 1 + \frac{1}{\alpha-1}$)

Remarque : cette propriété est très loin d'être anecdotique (encore moins que les autres). Il est en effet particulièrement utile en pratique dans l'étude des séries numériques (à la fin de cette année), ainsi que dans celle des séries de fonctions (l'an prochain).

Remarque : il existe une preuve beaucoup plus rapide et moins torturée du critère de convergence des séries de Riemann, que nous verrons au cours du chapitre sur les séries (et qui repose sur ce que nous appellerons le "théorème de comparaison série-intégrale").

6. C'est une somme de termes positifs !

Chapitre 18

Analyse asymptotique

Deux grandes parties dans ce chapitre. La première, un peu technique, consiste à introduire les relations de comparaison sur les fonctions. Essentiellement, l'intérêt de ces notions est d'apporter un "confort" de rédaction ; sur le fond peu de propriétés franchement nouvelles découleront des définitions données au cours des 3 premiers paragraphes. La seconde, consacrée aux développements limités, vous permettra en revanche de franchir un palier très important en Analyse, notamment lorsqu'il s'agira de lever des formes indéterminées dans des calculs de limites.

Préambule. Lorsque l'on crée un programme informatique, il peut être utile d'avoir une petite idée de sa durée d'exécution ; le résultat sera t-il obtenu dans la minute, ou devra t-on rester éveillé toute la nuit pour l'avoir (essayer de déterminer tous les nombres de Carmichael inférieurs à 10^6 , vous m'en direz des nouvelles) ?

Cette question est intimement liée à celle de **complexité algorithmique**, qui est essentiellement une estimation du nombre d'opérations effectuées par la machine lors de l'exécution du programme. Mais très souvent en pratique, on ne peut calculer le nombre exact d'opérations effectuées (dès qu'intervient une instruction conditionnelle, c'est cuit !), et tout au plus peut-on donner une majoration de ce nombre.

Un petit exemple valant mieux qu'un long discours, le programme ci-dessous est celui d'une fonction comptant le nombre de voyelles dans une chaîne de caractères (que l'on suppose de longueur N).

```
1  def NbVoy(mot):
2      nb = 0
3      for k in range(N):
4          if mot[k] in 'AEIOUY':
5              nb = nb + 1
6      return nb
```

Le nombre d'opérations effectuées $C(N)$ sera *au pire* égal à $4N + 1$. Cette quantité croît "comme N " vers $+\infty$, et on a pour N suffisamment grand $C(N) \leq 5N$ (par exemple, mais vous pouvez remplacer 5 par n'importe quel $k > 4$). Dans cette situation on écrira que la complexité est un $O(N)$, et on dira que la complexité est linéaire. Disposer de cette information permettra d'estimer la durée d'exécution de l'algorithme en fonction de N .

Convention : tout au long des paragraphes 1 à 3, sauf mention explicite du contraire, les fonctions considérées seront définies au voisinage de $a \in \bar{R}$ et à valeurs dans $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$.

18.1 Négligeabilité

18.1.1 Généralités

DÉFINITION 18.1 - Une fonction f est **négligeable devant g au voisinage de a** s'il existe une fonction ε définie au voisinage de a telle que : $f(x) = g(x)\varepsilon(x)$ et $\lim_{x \rightarrow a} \varepsilon(x) = 0$.

Notation. On note : $f = o_a(g)$, ou $f(x) = o_a(g(x))$, ou encore $f(x) \ll g(x)$ lorsque x tend vers a (la dernière notation étant plutôt utilisée par les Physiciens).

Notation. Dans la suite de ce chapitre, on notera $\mathbf{R}(\mathbf{a})$ l'ensemble des fonctions définies sur un voisinage raisonnable de a , la notion de "raisonnable" étant à adapter en fonction du contexte, et $\mathbf{R}_0(\mathbf{a})$ la partie de $\mathbf{R}(\mathbf{a})$ constituée des éléments ayant pour limite 0 en a .

PROPRIÉTÉ 18.1 - (Caractérisation de la négligeabilité). Sous réserve que g ne s'annule pas sur un voisinage de a , on a : $f = o_a(g)$ SSI $\lim_a \frac{f}{g} = 0$.

PREUVE. $[f = o_a(g)] \iff [\exists \varepsilon \in \mathbf{R}_0(\mathbf{a}), f(x) = g(x)\varepsilon(x)] \iff \left[\exists \varepsilon \in \mathbf{R}_0(\mathbf{a}), \frac{f(x)}{g(x)} = \varepsilon(x) \right]$
 $\iff \left[\exists \varepsilon \in \mathbf{R}_0(\mathbf{a}), \frac{f(x)}{g(x)} = \varepsilon(x) \right]$ ☕

PROPRIÉTÉ 18.2 - $f = o_a(1)$ SSI $\lim_a f = 0$.

PREUVE. $[f = o_a(1)] \iff [\exists \varepsilon \in \mathbf{R}_0(\mathbf{a}), f = \varepsilon] \iff \left[\lim_a f = 0 \right]$ ☕

PROPRIÉTÉ 18.3 - (négligeabilité et opérations algébriques).

- 1/ Si $f = o_a(g)$ alors : $\forall \lambda \in \mathbb{K}, \lambda f = o_a(g)$ (et d'ailleurs $f = o_a(\lambda g)$)
- 2/ Si $f = o_a(g)$ et $h = o_a(g)$ alors $f + h = o_a(g)$
- 3/ Si $f = o_a(g)$ alors $fh = o_a(gh)$
- 4/ Si $f = o_a(g)$ et $g = o_a(h)$ alors : $f = o_a(h)$.

Retraduction : 1) $\lambda o_a(g) = o_a(g)$ 2) $o_a(g) + o_a(g) = o_a(g)$ 3) $h o_a(g) = o_a(gh)$

PREUVE. 1/ Soient f et g deux fonctions définies au voisinage de a , telles que $f = o_a(g)$, et soit $\lambda \in \mathbb{K}$. Alors, il existe une fonction $\varepsilon \in \mathbf{R}_0(\mathbf{a})$ telle que : $f = g\varepsilon$. Il s'ensuit que : $\lambda f = g\hat{\varepsilon}$ en ayant posé $\hat{\varepsilon} = \lambda\varepsilon$. Il reste à observer puissamment que $\hat{\varepsilon} \in \mathbf{R}_0(\mathbf{a})$ pour conclure que : $\lambda f = o_a(g)$.

2/ Sous les hypothèses de l'énoncé, il existe deux fonctions ε_1 et ε_2 dans $\mathbf{R}_0(\mathbf{a})$ telles que : $f = \varepsilon_1 g$ et $h = \varepsilon_2 g$. Il s'ensuit que $f + h = \varepsilon_3 g$ en ayant noté $\varepsilon_3 = \varepsilon_1 + \varepsilon_2$. Il reste à ré-observer puissamment que $\varepsilon_3 \in \mathbf{R}_0(\mathbf{a})$ pour conclure que : $f + h = o_a(g)$.

3/ Sous les hypothèses de l'énoncé, il existe une fonction $\varepsilon \in \mathbf{R}_0(\mathbf{a})$ telle que : $f = g\varepsilon$. Alors, pour toute fonction $h \in \mathbf{R}(\mathbf{a})$, on a : $fh = gh\varepsilon$ ce qui signifie exactement que $fh = o_a(gh)$.

4/ Sous les hypothèses de l'énoncé, il existe deux fonctions ε_1 et ε_2 dans $\mathbf{R}_0(\mathbf{a})$ telles que : $f = \varepsilon_1 g$ et $g = \varepsilon_2 h$. Il s'ensuit que $f = \varepsilon_3 h$ en ayant noté $\varepsilon_3 = \varepsilon_1 \varepsilon_2$. Il reste à ré-observer puissamment que $\varepsilon_3 \in \mathbf{R}_0(\mathbf{a})$ pour conclure que : $f = o_a(h)$. ☕

18.1.2 Croissances comparées

18.1.2.1 En $+\infty$

PROPRIÉTÉ 18.4 - (CC1). Soient α et β réels, avec $\alpha < \beta$. Alors : $x^\alpha = o_{+\infty}(x^\beta)$

PREUVE. Triviale. ☕

PROPRIÉTÉ 18.5 - (CC2). Soient a et b réels, avec $0 < a < b$. Alors : $a^x = o_{+\infty}(b^x)$.

PREUVE. Triviale. ☕

PROPRIÉTÉ 18.6 - (CC3). Soient $\alpha > 0$, $\beta > 0$ et $a > 1$ trois réels. Alors :

$$(\ln x)^\beta = o(x^\alpha) \quad \text{et} \quad x^\alpha = o_{+\infty}(a^x)$$

PREUVE. Triviale. ☕

Synthèse — Echelle de négligeabilité en $+\infty$

$$e^{-x} \ll \frac{1}{x^2} \ll \frac{1}{x} \ll \frac{1}{\ln x} \ll 1 \ll x \ll x^2 \ll e^x$$

18.1.2.2 En 0^+

PROPRIÉTÉ 18.7 - (CC4). Soient α et β réels, avec $\alpha < \beta$. Alors : $x^\beta = o_{0^+}(x^\alpha)$

PREUVE. Triviale. ☕

PROPRIÉTÉ 18.8 - (CC5). Soit $\alpha > 0$ un réel. Alors : $\ln(x) = o_{0^+}\left(\frac{1}{x^\alpha}\right)$

PREUVE. Triviale. ☕

Synthèse — Echelle de négligeabilité en 0^+

$$x^2 \ll x \ll \sqrt{x} \ll \frac{1}{\ln x} \ll 1 \ll \ln x \ll \frac{1}{\sqrt{x}} \ll \frac{1}{x} \ll \frac{1}{x^2}$$

18.2 Domination (ou prépondérance)

DÉFINITION 18.2 - Une fonction f est **dominée par g au voisinage de a** s'il existe un réel $M \in \mathbb{R}_+$ tel que sur un voisinage de a on ait : $|f(x)| \leq M |g(x)|$.

Notation. On note : $f = O_a(g)$, ou $f(x) = O_a(g(x))$.

PROPRIÉTÉ 18.9 - Si $f = o_a(g)$, alors $f = O_a(g)$.

PREUVE. Si f est négligeable devant g , il existe une fonction $\varepsilon \in \mathbf{R}_0(\mathbf{a})$ telle que $f = g\varepsilon$. Puisque la fonction ε a pour limite 0 en a , alors elle est bornée au voisinage de a , d'où la conclusion. ☕

PROPRIÉTÉ 18.10 - $f = O_a(1)$ SSI f est bornée au voisinage de a .

PREUVE. Résulte de la définition de la relation de domination, et de la propriété selon laquelle f est bornée SSI sa valeur absolue est majorée. ☕

PROPRIÉTÉ 18.11 - (**caractérisation de la domination**). Sous réserve que g ne s'annule pas sur un voisinage de a , on a : $f = O_a(g)$ SSI $\frac{f}{g}$ est bornée au voisinage de a .

PREUVE. Conséquence immédiate de la propriété précédente. ☕

PROPRIÉTÉ 18.12 - (**domination et opérations algébriques**).

1/ Si $f = O_a(g)$ alors : $\forall \lambda \in \mathbb{K}, \lambda f = O_a(g)$ (et d'ailleurs $f = O_a(\lambda g)$)

2/ Si $f = O_a(g)$ et $h = O_a(g)$ alors $f + h = O_a(g)$

3/ Si $f = O_a(g)$ alors $fh = O_a(gh)$

Retraduction : 1/ $\lambda O_a(g) = O_a(g)$ 2/ $O_a(g) + O_a(g) = O_a(g)$ 3/ $h O_a(g) = O_a(gh)$

PREUVE. 1/ Immédiat. 2/ Trivial. 3/ Evident. ☕

PROPRIÉTÉ 18.13 - ("**transitivités diverses et variées**").

1/ Si $f = o_a(g)$ et $g = O_a(h)$ alors : $f = O_a(h)$.

2/ Si $f = O_a(g)$ et $g = o_a(h)$ alors : $f = o_a(h)$.

3/ Si $f = O_a(g)$ et $g = O_a(h)$ alors : $f = O_a(h)$.

PREUVE. 1/ Trivial. 2/ Clair. 3/ Banal. ☕

18.3 Equivalents

DÉFINITION 18.3 - Deux fonctions f et g sont **équivalentes au voisinage de a** s'il existe une fonction φ définie au voisinage de a telle que : $f(x) = g(x)\varphi(x)$.

Notation. On note : $f \sim_a g$, ou $f(x) \sim_{x \rightarrow a} g(x)$. On dit encore que f est un équivalent de g au voisinage de a .

PROPRIÉTÉ 18.14 - (Caractérisation de l'équivalence). Sous réserve que g ne s'annule pas sur un voisinage de a , on a :

$$f \sim_a g \text{ SSI } \lim_a \frac{f}{g} = 1$$

PREUVE. Immédiate. ☕

PROPRIÉTÉ 18.15 - (la relation d'équivalence est une relation d'équivalence, LOL).

$$1) f \sim_a f \quad 2) [f \sim_a g] \iff [g \sim_a f] \quad 3) [f \sim_a g \text{ et } g \sim_a h] \iff f \sim_a h$$

PREUVE. Triviale. ☕

PROPRIÉTÉ 18.16 - $f \sim_a g \text{ SSI } f = g + o_a(g)$.

PREUVE. On introduit $\mathbf{R}_1(\mathbf{a})$ l'ensemble des fonctions définies sur un voisinage raisonnable de a , et qui ont pour limite 1 en a .

$$[f \sim_a g] \iff [\exists \varphi \in \mathbf{R}_1(\mathbf{a}), f = g\varphi] \iff [\exists \varphi \in \mathbf{R}_1(\mathbf{a}), f = g + g(\varphi - 1)]$$

$$\iff [\exists \psi \in \mathbf{R}_0(\mathbf{a}), f = g + g\psi] \iff [f = g + o_a(g)] \quad ☕$$

PROPRIÉTÉ 18.17 - (Transitivités diverses et variées).

1/ Si $f \sim_a g$ alors : $f = O_a(g)$ (et $g = O_a(f)$)

2/ Si $f \sim_a g$ et $g = o_a(h)$ alors $f = o_a(h)$

3/ Si $f \sim_a g$ et $g = O_a(h)$ alors $f = O_a(h)$

4/ Si $f = o_a(g)$ et $g \sim h$ alors $f = o_a(h)$

5/ Si $f = O_a(g)$ et $g \sim h$ alors $f = O_a(h)$

PREUVE. Let's go!

1/ $\boxed{\text{Si } f \sim_a g \text{ alors : } f = O_a(g) \text{ (et } g = O_a(f))}$

Comme $f \sim_a g$, on a : $f = g + o_a(g)$. Or $g = O_a(g)$ et $o_a(g) = O_a(g)$. D'où $f = O_a(g) + O_a(g)$, donc : $f = O_a(g)$. La relation $g = O_a(f)$ résulte par exemple du fait que la relation d'équivalence au voisinage de a est symétrique.

2/ $\boxed{\text{Si } f \sim_a g \text{ et } g = o_a(h) \text{ alors } f = o_a(h)}$

Comme $f \sim_a g$, on a : $f = g + o_a(g)$. Or $g = o_a(h)$. D'où $f = o_a(h) + o_a(o_a(h))$, donc : $f = o_a(h) + o_a(h)$ d'où : $f = o_a(h)$.

3/ $\boxed{\text{Si } f \sim_a g \text{ et } g = O_a(h) \text{ alors } f = O_a(h)}$

Comme $f \sim_a g$, on a : $f = g + o_a(g)$. Or $g = O_a(h)$. D'où $f = O_a(h) + o_a(O_a(h))$, donc : $f = O_a(h) + o_a(h)$ d'où : $f = O_a(h)$.

4/ $\boxed{\text{Si } f = o_a(g) \text{ et } g \sim h \text{ alors } f = o_a(h)}$

Comme $g \sim h$, on a : $g = h + o_a(h)$. Or $f = o_a(g)$. D'où $f = o_a(h) + o_a(o_a(h))$, d'où : $f = o_a(h)$.

5/ $\boxed{\text{Si } f = O_a(g) \text{ et } g \sim h \text{ alors } f = O_a(h)}$

Comme $g \sim h$, on a : $g = h + o_a(h)$. Or $f = O_a(g)$. D'où $f = O_a(h) + O_a(o_a(h))$, donc $f = O_a(h) + o_a(h)$ d'où : $f = O_a(h)$. ☕

PROPRIÉTÉ 18.18 - (équivalents et opérations algébriques).

- 1) Si $f_1 \sim_a f_2$ et $g_1 \sim_a g_2$ alors : $f_1 g_1 \sim_a f_2 g_2$
- 2) Si $f_1 \sim_a f_2$ et $g_1 \sim_a g_2$ alors : $\frac{f_1}{g_1} \sim_a \frac{f_2}{g_2}$ (sous réserve...)
- 3) Si $f_1 \sim_a f_2$ alors : $\forall p \in \mathbb{Z}, f_1^p \sim_a f_2^p$ (sous réserve...)

PREUVE. Triviale³. ☕

Remarque : en revanche, on ne peut pas sommer les équivalents !

De par son extrême utilité pratique, l'énoncé suivant mérite bien le nom de théorème.

THÉORÈME 18.1 - Si $f \sim_a g$ et $\lim_{x \rightarrow a} g(x) = \ell \in \bar{\mathbb{R}}$ ou $\in \mathbb{C}$, alors $\lim_{x \rightarrow a} f(x) = \ell$

PREUVE. Triviale. ☕

18.4 Développements limités

18.4.1 Généralités

DÉFINITION 18.4 - Soit $f : I \rightarrow \mathbb{K}$ une fonction à valeurs réelles définie sur un intervalle ouvert non-vide, soit $a \in I$ et soit n un entier naturel.

La fonction f admet un **développement limité à l'ordre n au voisinage de a** s'il existe $(n+1)$ scalaires $a_0, a_1, \dots, a_n$ tels que :

$$\forall x \in I, f(x) = \left[\sum_{k=0}^n a_k (x-a)^k \right] + o((x-a)^n)$$

L'expression $\left[\sum_{k=0}^n a_k (x-a)^k \right]$ est appelée **partie régulière** du DL.

En particulier, une fonction f admet un développement limité à l'ordre n au voisinage de zéro si on peut écrire :

$$f(x) = \left[\sum_{k=0}^n a_k x^k \right] + o(x^n)$$

Exemple : $\forall n \in \mathbb{N}, \forall x \in]-1, 1[, \frac{1}{1-x} = \left[\sum_{k=0}^n x^k \right] + o(x^n)$.

PROPRIÉTÉ 18.19 - (Unicité du DL). Si f admet un développement limité à l'ordre n en a , alors il est unique.

PREUVE. Non-triviale¹. Posons pour tout entier naturel n , $P(n)$: “Si f admet un DL d'ordre n en a , il est unique”, et montrons que $P(n)$ est vraie pour tout entier naturel n par récurrence.

► Initialisation (pour $n=0$) : supposons qu'il existe deux réels a_0 et b_0 tels que, sur un voisinage de a on ait : $f(x) = a_0 + o_a(1)$ et $f(x) = b_0 + o_a(1)$. Alors² : $b_0 - a_0 = o_a(1)$. Ce qui signifie que : $\lim_{x \rightarrow a} b_0 - a_0 = 0$. Il s'ensuit que : $b_0 = a_0$, ce qui établit l'unicité dans le cas $n=0$. Donc $P(0)$ est vraie.

► Hérédité : supposons la propriété établie au rang n , pour un certain entier naturel n . Supposons également qu'il existe $2n+4$ réels $a_0, \dots, a_{n+1}, b_0, \dots, b_{n+1}$ tels que sur un voisinage de a on ait :

$$f(x) = \left[\sum_{k=0}^{n+1} a_k (x-a)^k \right] + o_a((x-a)^{n+1}) \quad \text{et} \quad f(x) = \left[\sum_{k=0}^{n+1} b_k (x-a)^k \right] + o_a((x-a)^{n+1})$$

La première (petite) astuce consiste à voir que les écritures ci-dessus fournissent en particulier deux DL à l'ordre n en a . En effet, on a³ :

$$f(x) = \left[\sum_{k=0}^n a_k (x-a)^k \right] + a_{n+1}(x-a)^{n+1} + o_a((x-a)^{n+1})$$

1. Aaah, quand même !

2. En utilisant le fait que : $o_a(1) - o_a(1) = o_a(1)$.

3. Toujours pour x raisonnablement proche de a ...

$$= \left[\sum_{k=0}^n a_k (x-a)^k \right] + (x-a)^n \underbrace{(a_{n+1}(x-a) + o_a((x-a)))}_{=\varepsilon(x)}$$

Puisqu'il est clair que : $\lim_{x \rightarrow a} \varepsilon(x) = 0$, on a donc : $f(x) = \left[\sum_{k=0}^n a_k (x-a)^k \right] + o_a((x-a)^n)$.

Naturellement, on a encore : $f(x) = \left[\sum_{k=0}^n b_k (x-a)^k \right] + o_a((x-a)^n)$.

L'hypothèse de récurrence permet alors d'affirmer que : $\forall k \in \llbracket 0, n \rrbracket, a_k = b_k$.

Par conséquent, sur un voisinage de a : $a_{n+1}(x-a)^{n+1} + o_a((x-a)^{n+1}) = b_{n+1}(x-a)^{n+1} + o_a((x-a)^{n+1})$.

Donc : $(a_{n+1} - b_{n+1})(x-a)^{n+1} = o_a((x-a)^{n+1})$; d'où $a_{n+1} - b_{n+1} = o_a(1)$; soit finalement $a_{n+1} = b_{n+1}$.

En résumé : $\forall k \in \llbracket 0, n+1 \rrbracket, a_k = b_k$. Ce qui prouve l'unicité du DL au rang $n+1$, et achève la preuve de l'hérédité.

Conclusion. Si f admet un DL à l'ordre n en a , il est unique.



18.4.2 Construction — Formule de Taylor-Young

THÉORÈME 18.2 - (FORMULE DE TAYLOR-YOUNG) : si $f \in \mathcal{C}^n(I, \mathbb{R})$ et $a \in I$, alors f admet un DL à l'ordre n en a , explicitement :

$$\forall x \in I, \quad f(x) = \left[\sum_{k=0}^n \frac{f^{(k)}(a)}{k!} (x-a)^k \right] + o((x-a)^n)$$

En particulier lorsque $a = 0$:

$$\forall x \in I, \quad f(x) = \left[\sum_{k=0}^n \frac{f^{(k)}(0)}{k!} x^k \right] + o(x^n)$$

PREUVE. Posons pour tout entier naturel n

$$\mathcal{P}(n) : \quad \text{“Si } f \in \mathcal{C}^n(I, \mathbb{R}), \text{ alors } \forall x \in I, f(x) = \left[\sum_{k=0}^n \frac{f^{(k)}(0)}{k!} x^k \right] + o((x-a)^n)”$$

et prouvons que $\mathcal{P}(n)$ est vraie pour tout entier naturel n par récurrence.

► Initialisation (pour $n = 0$) : supposons que $f \in \mathcal{C}^0(I, \mathbb{R})$, c'est-à-dire que f est continue sur I . Pour tout réel a de I , on peut judicieusement écrire :

$$\forall x \in I, \quad f(x) = f(a) + \underbrace{(f(x) - f(a))}_{=\varepsilon(x)}$$

On pose : $\varepsilon(x) = (f(x) - f(a))$. Comme f est continue en a par hypothèse, on a : $\lim_{x \rightarrow a} f(x) = f(a)$.
D'où : $\lim_{x \rightarrow a} \varepsilon(x) = 0$. On a donc obtenu, sous l'hypothèse que $f \in \mathcal{C}^0(I, \mathbb{R})$:

$$\forall x \in I, f(x) = f(a) + \varepsilon(x) \text{ avec } \lim_{x \rightarrow a} \varepsilon(x) = 0, \text{ c\`ad : } \forall x \in I, f(x) = f(a) + o_a(1)$$

Ce qui prouve que la propriété $\mathcal{P}(0)$ est vraie.

► Hérédité : supposons la propriété $\mathcal{P}(n)$ vraie pour un certain $n \in \mathbb{N}$, et montrons qu'elle est vraie au rang $(n+1)$.

Supposons que $f \in \mathcal{C}^{n+1}(I, \mathbb{R})$.

On observe (point-clef!) que sous cette hypothèse on a : $f' \in \mathcal{C}^n(I, \mathbb{R})$. On peut alors légitimement appliquer l'hypothèse de récurrence à la fonction f' pour affirmer que pour tout réel $a \in I$ on a :

$$\forall t \in I, f'(t) = \left[\sum_{k=0}^n \frac{f^{(k)}(a)}{k!} (t-a)^k \right] + o_a((t-a)^n)$$

Puis, pour un x fixé dans I , on intègre cette relation entre a et x pour obtenir (par linéarité de l'intégrale) :

$$\forall x \in I, \int_a^x f'(t) dt = \left[\sum_{k=0}^n \frac{f^{(k+1)}(a)}{k!} \int_a^x (t-a)^k dt \right] + \int_a^x o_a((t-a)^n) dt$$

Soit :

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

$$\text{D'où : } \forall x \in I, f(x) = f(a) + \left[\sum_{k=0}^n \frac{f^{(k+1)}(a)}{(k+1)!} (x-a)^{k+1} \right] + \int_a^x o_a((t-a)^n) dt$$

Et finalement :

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

Pour conclure, il "suffit" de montrer que le terme $\int_a^x o_a((t-a)^n) dt$ est négligeable devant $(x-a)^{n+1}$ au voisinage de a , c'est-à-dire qu'on peut l'écrire comme $(x-a)^{n+1} \varphi(x)$ avec $\lim_{x \rightarrow a} \varphi(x) = 0$.

On commence par noter que : $o_a((t-a)^n) = (t-a)^n \varepsilon(t)$ avec $\lim_{t \rightarrow a} \varepsilon(t) = 0$ (définition de négligeabilité).

$$\text{Puis on observe que pour tout réel } x \neq a : \int_a^x o_a((t-a)^n) dt = (x-a)^{n+1} \times \frac{\int_a^x (t-a)^n \varepsilon(t) dt}{(x-a)^{n+1}}.$$

Posons alors : $\varphi(x) = \frac{\int_a^x (t-a)^n \varepsilon(t) dt}{(x-a)^{n+1}}$, et montrons que $\varphi(x)$ tend vers 0 quand x tend vers a , ce que l'on fait en revenant à la définition de limite.

Fixons $\beta > 0$. Puisque $\lim_{t \rightarrow a} \varepsilon(t) = 0$, il existe un réel α tel que $|t-a| < \alpha \implies |\varepsilon(t)| < \beta$. Ainsi :

$$|x-a| < \alpha \implies |\varphi(x)| < \frac{1}{|x-a|^{n+1}} \times \beta \int_a^x |t-a|^n dt \quad \text{d'où} \quad |x-a| < \alpha \implies |\varphi(x)| < \frac{\beta}{n+1}.$$

En particulier : $|x-a| < \alpha \implies |\varphi(x)| < \beta$, et on a donc établi que : $\forall \beta > 0, \exists \alpha > 0, |x-a| < \alpha \implies |\varphi(x)| < \beta$.

Ce qui signifie que $\lim_{x \rightarrow a} \varphi(x) = 0$, et donc : $\int_a^x o_a((t-a)^n) dt = o_a((x-a)^{n+1})$.

Finalement, on a : $\forall x \in I, f(x) = \left[\sum_{k=0}^{n+1} \frac{f^{(k)}(a)}{k!} (x-a)^k \right] + o_a((x-a)^{n+1})$, ce qui prouve que la propriété

$\mathcal{P}(n+1)$ est vraie, montre l'hérédité de la propriété, et conclut cette preuve. ☕

RÉCIPROQUE FAUSSE ! La fonction $f : x \in \mathbb{R}^* \mapsto 1+x+x^2 \sin\left(\frac{1}{x}\right)$ prolongée par continuité en posant $f(0) = 1$ admet un développement limité à l'ordre 1 en 0, mais n'est pas de classe $\mathcal{C}^1$ en 0.

La formule de Taylor-Young permet donc d'obtenir les DL des fonctions dont on connaît les dérivées successives, notamment :

$$\triangleright \forall x \in]-1; 1[, \quad \frac{1}{1-x} = \left[\sum_{k=0}^n x^k \right] + o(x^n)$$

$$\triangleright \forall x \in]-1; 1[, \quad \frac{1}{1+x} = \left[\sum_{k=0}^n (-1)^k x^k \right] + o(x^n)$$

$$\triangleright \forall x \in \mathbb{R}, \quad e^x = \left[\sum_{k=0}^n \frac{x^k}{k!} \right] + o(x^n)$$

$$\triangleright \forall x \in \mathbb{R}, \quad \text{ch}(x) = \left[\sum_{k=0}^n \frac{x^{2k}}{(2k)!} \right] + o(x^{2n})$$

$$\triangleright \forall x \in \mathbb{R}, \quad \text{sh}(x) = \left[\sum_{k=0}^n \frac{x^{2k+1}}{(2k+1)!} \right] + o(x^{2n+1})$$

$$\triangleright \forall x \in \mathbb{R}, \quad \cos(x) = \left[\sum_{k=0}^n (-1)^k \frac{x^{2k}}{(2k)!} \right] + o(x^{2n})$$

$$\triangleright \forall x \in \mathbb{R}, \quad \sin(x) = \left[\sum_{k=0}^n (-1)^k \frac{x^{2k+1}}{(2k+1)!} \right] + o(x^{2n+1})$$

Ultime conséquence de la formule de Taylor-Young :

$$\triangleright (1+x)^\alpha = 1 + \left[\sum_{k=1}^n \frac{\prod_{i=0}^{k-1} (\alpha - i)}{k!} x^k \right] + o(x^n)$$

Dont il faut plutôt retenir la forme explicite :

$$(1+x)^\alpha = 1 + \alpha x + \frac{\alpha(\alpha-1)}{2} x^2 + \frac{\alpha(\alpha-1)(\alpha-2)}{6} x^3 + o(x^3)$$

Application : $\forall x > -1, \sqrt{1+x} = 1 + \frac{x}{2} - \frac{x^2}{8} + o(x^2)$

On peut observer que les DL des fonctions paires (*resp.* impaires) ne comportent que des puissances paires (*resp.* impaires) de x dans leurs parties régulières. La justification a posteriori de cette observation est fournie par l'énoncé ci-dessous, dont la preuve se fait par récurrence à partir de la propriété (exo) suivant laquelle la dérivée d'une fonction dérivable et paire est impaire (*resp.* ...).

PROPRIÉTÉ 18.20 - (DL et parité) : si $f \in \mathcal{C}^n(I, \mathbb{R})$ est paire (*resp.* impaire) alors la partie régulière du développement limité de f en 0 ne contient que des termes de degré pair (*resp.* impair).

PREUVE. Conséquence immédiate de la remarque faite sur la parité de f et celle de f' précédant l'énoncé de la propriété. ☕

RÉCIPROQUE FAUSSE ! La fonction $f : x \in \mathbb{R}^* \mapsto 1 + x^2 + x^3$ admet un DL à l'ordre 2 ne contenant que des termes pairs, mais n'est pas une fonction paire.

18.4.3 DL et opérations algébriques

Tout au long de cette section, on suppose que f et g sont deux fonctions définies au voisinage de 0, et y admettent un DL à l'ordre n . On note :

$$f(x) = \underbrace{a_0 + a_1x + \cdots + a_nx^n}_{P(x)} + o(x^n) \text{ et } g(x) = \underbrace{b_0 + b_1x + \cdots + b_nx^n}_{Q(x)} + o(x^n)$$

P et Q désignent donc les parties régulières de ces deux développements limités. Observons que les formules ci-dessous peuvent être adaptées au cas où f et g sont définies au voisinage de a . Il suffit de modifier ce qui doit l'être, et en particulier de remplacer partout où nécessaire les " x " par des " $(x-a)$ ".

18.4.3.1 DL d'une somme

PROPRIÉTÉ 18.21 - La fonction $f + g$ admet un DL à l'ordre n en 0, de partie régulière $\lambda P + \mu Q$.

PREUVE. ☕

18.4.3.2 DL d'une combinaison linéaire

PROPRIÉTÉ 18.22 - Soient λ et μ deux scalaires. La fonction $\lambda f + \mu g$ admet un DL à l'ordre n en 0, de partie régulière $\lambda P + \mu Q$.

PREUVE. ☕

18.4.3.3 DL d'un produit

PROPRIÉTÉ 18.23 - La fonction $f \times g$ admet un DL à l'ordre n en 0, dont la partie régulière est le produit $P \times Q$, tronqué à l'ordre n .

PREUVE. ☕

Application. DL à l'ordre 5 en 0 de $x^2 e^x$: $x^2 + x^3 + \frac{x^4}{2!} + \frac{x^5}{3!} + o(x^5)$.

18.4.3.4 DL d'un quotient

PROPRIÉTÉ 18.24 - La fonction f/g admet un DL à l'ordre n en 0, dont la partie régulière est le quotient de P et Q effectué suivant les puissances croissantes, tronqué à l'ordre n .

PREUVE. ☕

La fonction f/g admet un DL à l'ordre n en 0, dont la partie régulière est le quotient de P et Q effectué suivant les puissances croissantes, tronqué à l'ordre n .

Application 1 : $\forall x \in \left] -\frac{\pi}{2}; \frac{\pi}{2} \right[, \tan(x) = x + \frac{x^3}{3} + \frac{2x^5}{15} + \frac{17x^7}{315} + o(x^8)$

Application 2 : $\forall x \in \mathbb{R}, \operatorname{th}(x) = x - \frac{x^3}{3} + \frac{2x^5}{15} - \frac{17x^7}{315} + o(x^8)$

Remarque : si on ne veut pas faire de "HP" pour calculer le DL du quotient f/g , on essaye de se transformer cette expression pour se ramener à un produit de la forme $h \times 1/(1+u)$ (cf les deux versions données en cours pour le DL5 en 0 de la fonction tangente).

18.4.3.5 DL d'une composée

PROPRIÉTÉ 18.25 - Si $\lim_{x \rightarrow 0} g(x) = 0$, la fonction $f \circ g$ admet un DL à l'ordre n en 0, dont la partie régulière est le polynôme composé $P \circ Q$, tronqué à l'ordre n .

PREUVE. ☕

Application : $\forall x \in \mathbb{R}, e^{-x} = \left[\sum_{k=0}^n (-1)^k \frac{x^k}{k!} \right] + o(x^n)$

18.4.3.6 DL et “primitivation”

PROPRIÉTÉ 18.26 - Si F est dérivable, et si f admet un DL à l'ordre n en 0, alors F admet un DL à l'ordre $n+1$ en 0 :

$$F(x) = F(0) + a_0x + a_1\frac{x^2}{2} + \cdots + a_n\frac{x^{n+1}}{n+1} + o(x^{n+1})$$

PREUVE. 

Application 1 : $\forall x > -1, \ln(1+x) = \left[\sum_{k=1}^n (-1)^{k+1} \frac{x^k}{k} \right] + o(x^n)$

Application 2 : $\forall x \in \mathbb{R}, \arctan(x) = \left[\sum_{k=0}^n (-1)^k \frac{x^{2k+1}}{2k+1} \right] + o(x^{2n+2})$

Application 3 : $\forall x \in]-1; 1[, \arcsin(x) = x + \frac{x^3}{6} + \frac{3x^5}{40} + \frac{5x^7}{112} + o(x^8)$

Conséquence : $\forall x \in]-1; 1[, \arccos(x) = \frac{\pi}{2} - x - \frac{x^3}{6} - \frac{3x^5}{40} - \frac{5x^7}{112} + o(x^8)$

18.5 Applications des développements limités

18.5.1 Calculs de limites

On peut lever des indéterminations en effectuant un DL en 0 à un ordre judicieusement choisi.

Exemples : $\lim_{x \rightarrow 0} \frac{\sin(x)}{x} = 1$ (DL à l'ordre 1 en 0 de $\sin$) ; $\lim_{x \rightarrow 0} \frac{\sin(x) - x}{x^3} = -\frac{1}{6}$ (DL à l'ordre 3 en 0 de $\sin$) ; $\lim_{x \rightarrow 0} \frac{1}{x^2} - \frac{1}{\tan^2(x)} = \frac{2}{3}$ (DL à l'ordre 4 en 0 de $\tan^2$).

18.5.2 Tangentes et positions relatives

Soit f une fonction définie au voisinage de 0, ayant comme DL : $f(x) = a_0 + a_1x + \cdots + a_nx^n + o(x^n)$. Alors $\mathcal{C}_f$ admet une tangente T au point d'abscisse 0 d'équation $y = a_0 + a_1x$. La position relative de $\mathcal{C}_f$ et de T est donnée **au voisinage de 0** par le signe du premier terme non nul de degré ≥ 2 .

18.5.3 Recherches d'équivalents

En posant $x = 1/n$ dans les formules donnant les DL des fonctions usuelles, on peut obtenir les équivalents cités dans le formulaire.

Par exemple $\ln\left(1 + \frac{1}{n}\right) \underset{+\infty}{\sim} \frac{1}{n}$; $e^{\frac{1}{n}} - 1 \underset{+\infty}{\sim} \frac{1}{n}$; $\sqrt{1 + \frac{1}{n}} - 1 \underset{+\infty}{\sim} \frac{1}{2n} \dots$

18.5.4 Asymptotes et positions relatives

Par le biais du changement de variable “ $X = 1/x$ ” dans les DL usuels, on peut obtenir des “DL en $+\infty$ ”, appelés développements asymptotiques, permettant d’obtenir entre autres des équations d’asymptotes et des positions relatives au voisinage de $+\infty$.

Par exemple : $\forall x \in \mathbb{R}_+^*$, $\sqrt{x^2 + x + 1} = x + \frac{1}{2} - \frac{1}{8x} + o\left(\frac{1}{x}\right)$. On en déduit que la courbe représentant $x \mapsto \sqrt{x^2 + x + 1}$ admet en $+\infty$ une asymptote d’équation $y = x + 1/2$, et que $\mathcal{C}_f$ est en-dessous de celle-ci au voisinage de $+\infty$.

18.6 Appendice - DL à l’ordre n de arcsin

Problème Développement limité en 0 de la fonction arcsinus.

Partie I — DL de arcsin en 0 à l’ordre 6

0) Question gratuite : déterminer le développement limité à l’ordre 5 en 0 de la fonction $\varphi : x \mapsto \frac{\sin(x)}{\operatorname{ch}(x)}$.

Dans les quatre questions suivantes, on pourra admettre sans justification que les fonctions considérées sont définies sur $] -1, 1[$.

- 1) Rappeler l’expression donnant le développement limité à l’ordre 2 en 0 de la fonction $f_1 : x \mapsto (1+x)^\alpha$ (avec α réel non nul arbitraire).
- 2) Dédire de la question précédente le développement limité à l’ordre 2 en 0 de la fonction $f_2 : x \mapsto \frac{1}{\sqrt{1-x}}$.
- 3) Dédire de la question précédente le développement limité à l’ordre 4 en 0 de la fonction $f_3 : x \mapsto \frac{1}{\sqrt{1-x^2}}$.
- 4) Dédire de la question précédente le développement limité à l’ordre 5 en 0 de la fonction arcsin, puis son développement limité à l’ordre 6 en 0.

Partie II — Résolution d’une équation différentielle

5) On veut résoudre dans cette question l’équation différentielle

$$(E1) : \quad \cos(t) y''(t) - 2 \sin(t) y'(t) - \cos(t) y(t) = 0$$

sur l’intervalle $I = \left] -\frac{\pi}{2}, \frac{\pi}{2} \right[$.

Plus précisément, on cherche à déterminer les fonctions $y : I \longrightarrow \mathbb{R}$ de classe $\mathcal{C}^2$ sur I solutions de (E1). Or cette équation différentielle linéaire étant d’ordre 2 à coefficients non-constants, aucun résultat de votre cours ne vous permet de la résoudre directement. On propose donc de la résoudre ci-dessous par une méthode appelée “changement de fonction inconnue”.

a) Soit $y : I \rightarrow \mathbb{R}$ une fonction de classe $\mathcal{C}^2$ sur I . On pose pour tout réel $t \in I$: $z(t) = y(t) \cos(t)$.

Montrer que y est solution de (E1) si et seulement si $z'' = 0$.

b) Dédurre de la question précédente la solution générale de (E1) sur I .

6) On veut résoudre dans cette question l'équation différentielle

$$(E2) : \quad (1 - x^2) y''(x) - 3x y'(x) - y(x) = 0$$

sur l'intervalle $J =]-1, 1[$.

Plus précisément, on cherche à déterminer les fonctions $y : J \rightarrow \mathbb{R}$ de classe $\mathcal{C}^2$ sur J solutions de (E2). Comme précédemment, vous ne disposez d'aucune méthode pour la résoudre directement. On propose ici une autre méthode, qui consiste à effectuer le changement de variable $x = \sin(t)$.

a) Pour tout réel $t \in I$, on pose : $z(t) = y(\sin(t))$. Il revient au même d'écrire que : $\forall x \in J, y(x) = z(\arcsin(x))$.

Montrer que y est solution de (E2) si et seulement si z est solution de (E1).

b) Dédurre de la question précédente la solution générale de (E2) sur J .

Partie III — Une alternative à la résolution de (E2)

On considère la même équation différentielle (E2) : $(1 - x^2) y''(x) - 3x y'(x) - y(x) = 0$ que dans la partie II.

Dans cette partie⁴, on ne cherche pas à résoudre cette équation, mais à caractériser le développement limité en 0 d'une solution de (E2).

Tout au long de cette partie, f désigne une fonction de classe $\mathcal{C}^2$ solution de (E2).

7) Montrer que f est de classe $\mathcal{C}^\infty$ sur $J =]-1, 1[$.

8) Etablir que :

$$\forall n \in \mathbb{N}, \forall x \in J, \quad (1 - x^2) f^{(n+2)}(x) - (2n + 3) x f^{(n+1)}(x) - (n + 1)^2 f^{(n)}(x) = 0$$

9) Pour tout $n \in \mathbb{N}$, on pose $a_n = f^{(n)}(0)$. Etablir une relation entre a_{n+2} et a_n .

10) Soit p un entier naturel quelconque. Exprimer a_{2p+1} et a_{2p} en fonction de a_1 et a_0 respectivement et à l'aide de factorielles.

Partie IV — Conclusion

On pourra utiliser dans cette partie les résultats des parties II et III. Soit n un entier naturel.

11) Ecrire le DL à l'ordre $2n + 1$ en 0 de la fonction $g : x \mapsto \frac{\arcsin(x)}{\sqrt{1 - x^2}}$.

Indication : on pourra observer que g est la solution de (E2) telle que $g(0) = 0$ et $g'(0) = 1$.

4. Indépendante de la précédente, en dépit des apparences.

12) Ecrire le DL à l'ordre $2n$ en 0 de la fonction $h : x \mapsto \frac{1}{\sqrt{1-x^2}}$.

13) Ecrire le DL à l'ordre $2n+1$ en 0 de la fonction $\arcsin$.

Corrigé.

Partie I — DL de arcsin en 0 à l'ordre 6

1) D'après le cours : $\forall x \in]-1, 1[, (1+x)^\alpha = 1 + \alpha x + \frac{\alpha(\alpha-1)}{2} x^2 + o(x^2)$

2) D'après la question précédente : $\forall x \in]-1, 1[, \frac{1}{\sqrt{1-x}} = 1 + \frac{x}{2} + \frac{3x^2}{8} + o(x^2)$

3) D'après la question précédente, et par parité : $\forall x \in]-1, 1[, \frac{1}{\sqrt{1-x^2}} = 1 + \frac{x^2}{2} + \frac{3x^4}{8} + o(x^5)$

4) D'après la question précédente : $\forall x \in]-1, 1[, \arcsin(x) = x + \frac{x^3}{6} + \frac{3x^5}{40} + o(x^6)$

Partie II — Résolution d'une équation différentielle

5) a) La fonction z de l'énoncé est de classe $\mathcal{C}^2$ sur I puisque y est supposée $\mathcal{C}^2$ et d'après les théorèmes généraux. On peut donc la dériver deux fois sur l'intervalle I , et on ne va pas se gêner pour le faire. Deux calculs sans difficulté donnent : $\forall t \in I, z''(t) = \cos(t)y''(t) - 2\sin(t)y'(t) - \cos(t)y(t)$.

On peut donc conclure que : y est solution de (E1) si et seulement si $z'' = 0$.

b) D'après la question précédente : y est solution de (E1) si et seulement si $z'' = 0$. Or $z'' = 0$ SSI il existe deux scalaires a et b tels que : $\forall t \in I, z(t) = at + b$.

En recollant les morceaux, on en déduit que y est solution de (E1) si et seulement si il existe deux scalaires

a et b tels que : $\forall t \in I, y(t) = \frac{at+b}{\cos(t)}$.

6) a) La fonction z de l'énoncé est de classe $\mathcal{C}^2$ sur I puisque y est supposée $\mathcal{C}^2$ et d'après les théorèmes généraux. On peut donc la dériver deux fois sur l'intervalle I . On a :

$$\forall t \in I, z'(t) = \cos(t)y'(\sin(t)) \text{ puis : } z''(t) = -\sin(t)y'(\sin(t)) + \cos^2(t)y''(\sin(t)).$$

La fonction z est donc solution de (E_1) si et seulement si pour tout $t \in I$ on a :

$$-\sin(t)\cos(t)y'(\sin(t)) + \cos^3(t)y''(\sin(t)) - 2\sin(t)\cos(t)y'(\sin(t)) - \cos(t)y(\sin(t)) = 0$$

La fonction $\cos$ ne s'annulant pas sur I , on peut simplifier par $\cos(t)$ et ainsi obtenir :

$$\cos^2(t)y''(\sin(t)) - 3\sin(t)y'(\sin(t)) - y(\sin(t)) = 0$$

Via le changement de variable $x = \sin(t)$, l'assertion précédente est équivalente à la suivante :

$$\forall x \in J, (1-x^2)y''(x) - 3xy'(x) - y(x) = 0 \text{ qui est satisfaite SSI } y \text{ est solution de } (E_2).$$

Conclusion. y est solution de (E_2) si et seulement si z est solution de (E_1) .

b) D'après la question précédente et la question 5-b), on peut conclure que la solution générale de (E_2) est l'ensemble des fonctions y pour lesquelles il existe deux réels a et b tels que :

$$\forall t \in J =]-1, 1[, \quad y(t) = \frac{a \arcsin(t) + b}{\sqrt{1-t^2}}.$$

Partie III — Une alternative à la résolution de (E2)

7) D'après la question 6-b) et les théorèmes généraux, f est de classe $\mathcal{C}^\infty$ sur $J =]-1, 1[$.

8) La fonction f étant solution de l'équation (E2), on a pour tout réel x dans J : $(1-x^2)f''(x) - 3xf'(x) - f(x) = 0$. Puisqu'en outre la fonction f est de classe $\mathcal{C}^\infty$ sur J , on peut dériver n fois cette relation, et obtenir grâce à la formule de Leibniz :

$$\forall n \in \mathbb{N}, \forall x \in J, \quad (1-x^2)f^{(n+2)}(x) - (2n+3)xf^{(n+1)}(x) - (n+1)^2f^{(n)}(x) = 0.$$

9) D'après la question précédente : $\forall n \in \mathbb{N}, a_{n+2} = (n+1)^2 a_n$.

10) On peut conjecturer (même principe que dans les intégrales de Wallis) les expressions des termes pairs et impairs de la suite (a_n) , puis démontrer ces formules par récurrence.

On obtient : $\forall n \in \mathbb{N}, a_{2n} = \frac{[(2n)!]^2}{2^{2n}(n!)^2} a_0 \quad \text{et} \quad \forall n \in \mathbb{N}, a_{2n+1} = 2^{2n}(n!)^2 a_1$

Partie IV — Conclusion

11) Il "suffit" de voir que la fonction $g : x \mapsto \frac{\arcsin(x)}{\sqrt{1-x^2}}$ est la solution de (E2) telle que $g(0) = a_0 = 0$ et $g'(0) = a_1 = 1$. Par suite, d'après la question précédente et la formule de Taylor-Young :

$$\forall x \in]-1, 1[, \quad \frac{\arcsin(x)}{\sqrt{1-x^2}} = \sum_{k=0}^n \frac{2^{2k}(k!)^2}{(2k+1)!} x^{2k+1} + o(x^{2n+1})$$

12) Comme dans la question précédente, il suffit de voir que la fonction $h : x \mapsto \frac{1}{\sqrt{1-x^2}}$ est la solution de (E2) telle que $h(0) = a_0 = 1$ et $h'(0) = a_1 = 0$. Par suite, d'après la question 10 et la formule de Taylor-Young :

$$\forall x \in]-1, 1[, \quad \frac{1}{\sqrt{1-x^2}} = \sum_{k=0}^n \frac{(2k)!}{2^{2k}(k!)^2} x^{2k} + o(x^{2n})$$

13) On déduit par intégration de la formule précédente que :

$$\forall x \in]-1, 1[, \quad \arcsin(x) = \sum_{k=0}^n \frac{(2k)!}{(2k+1)2^{2k}(k!)^2} x^{2k+1} + o(x^{2n+1})$$

Chapitre 19

Polynômes et fractions rationnelles

Dans ce chapitre lignes $\mathbb{K}$ désigne $\mathbb{R}$ ou $\mathbb{C}$.

*A noter : pour faire preuve d'originalité et proposer une synthèse du chapitre avant même de l'avoir commencé, indiquons que les outils d'usage fréquent lorsqu'il s'agit de traiter une question sur les polynômes sont : **degré, racine, factorisation, coefficients, division euclidienne, décomposition en polynômes irréductibles, décomposition en éléments simples**. Cette affirmation vous paraîtra plus naturelle à la fin de ce chapitre (c'est en tout cas l'objectif).*

19.1 Généralités

DÉFINITION 19.1 - Un **polynôme (formel)** à coefficients dans un corps $\mathbb{K}$ est une suite d'éléments de $\mathbb{K}$ nulle à partir d'un certain rang.

Notations : un polynôme P à coefficients dans $\mathbb{K}$ sera en général noté $P = \sum_{k=0}^n a_k X^k$.

Explicitement, cela signifie que l'on notera $P = X^2 - 2$ plutôt que $P = (-2, 0, 1, 0, 0, \dots, 0, \dots)$.

L'ensemble des polynômes en l'indéterminée X et à coefficients dans $\mathbb{K}$ est noté $\mathbb{K}[X]$:

$$\mathbb{K}[X] = \{a_n X^n + a_{n-1} X^{n-1} + \dots + a_1 X + a_0 \quad , \quad (a_0, \dots, a_n) \in \mathbb{K}^{n+1}\}$$

Le **polynôme nul** est celui dont tous les coefficients sont nuls ; on le notera $\tilde{0}$. Plus généralement d'ailleurs, on pourra noter $\tilde{\lambda}$ le polynôme constant “égal” à λ , c'est à dire celui dont le premier coefficient est égal à λ , et tous les autres sont nuls.

DÉFINITION 19.2 - Deux polynômes sont **égaux** si tous leurs coefficients sont égaux.

► Opérations sur les polynômes.

DÉFINITION 19.3 - Soient P et Q deux polynômes dans $\mathbb{K}[X]$, et soit λ un scalaire.

On définit :

- 1/ le polynôme $P + Q$ comme le polynôme dont les coefficients sont obtenus par somme à partir de ceux de P et Q ;
- 2/ le polynôme λP comme le polynôme dont les coefficients sont obtenus en multipliant ceux de P par λ ;
- 3/ le polynôme $-P$ est le polynôme λP dans le cas particulier où $\lambda = -1$.

PROPRIÉTÉ 19.1 - (Propriétés de l'addition et de la multiplication par un scalaire dans $\mathbb{K}[X]$). Soient P , Q et R trois éléments de $\mathbb{K}[X]$, et λ et μ deux scalaires. On a :

1/ $P + Q = Q + P$	5/ $\lambda(P + Q) = \lambda P + \lambda Q$
2/ $P + \tilde{0} = \tilde{0} + P = P$	6/ $\lambda(\mu P) = (\lambda\mu)P$
3/ $P + (Q + R) = (P + Q) + R$	7/ $1P = P$
4/ $P + (-P) = \tilde{0}$	8/ $\tilde{0}P = \tilde{0}$

PREUVE. Triviale⁸. ☕

PROPRIÉTÉ 19.2 - $(\mathbb{K}[X], +)$ est un groupe abélien.

PREUVE. Il résulte des définitions et propriétés que précédentes que la loi $+$ est une loi de composition interne, associative et commutative, possédant un élément neutre (le polynôme nul $\tilde{0}$), et telle que tout élément de $\mathbb{K}[X]$ est inversible (l'inverse de P pour la loi $+$ étant son polynôme opposé $-P$). ☕

DÉFINITION 19.4 - Soient P et Q deux polynômes, que l'on note :

$$P = \sum_{k=0}^n a_k X^k \quad \text{et} \quad Q = \sum_{k=0}^m b_k X^k$$

On définit le **produit des polynômes** P et Q et on note PQ le polynôme défini en posant :

$$PQ = \sum_{k=0}^{n+m} \left(\sum_{i=0}^k a_i b_{k-i} \right) X^k$$

(avec la convention que $b_k = 0$ pour $k > m$ et $a_k = 0$ pour $k > n$).

PROPRIÉTÉ 19.3 - (Propriétés de la multiplication dans $\mathbb{K}[X]$). Soient P , Q et R trois éléments de $\mathbb{K}[X]$. On a :

$$1/ \quad PQ = QP$$

$$3/ \quad P(QR) = (PQ)R$$

$$2/ \quad P \times \tilde{0} = \tilde{0} \times P = \tilde{0}$$

$$4/ \quad P(Q + R) = PQ + PR$$

PREUVE. Triviale⁴. ☕

PROPRIÉTÉ 19.4 - $(\mathbb{K}[X], +, \times)$ est un anneau commutatif.

PREUVE. Il résulte de la propriété 19.2 que $(\mathbb{K}[X], +)$ est un groupe abélien, et des propriétés de la multiplication dans $\mathbb{K}[X]$ que $(\mathbb{K}[X], +, \times)$ est un anneau commutatif. ☕

COROLLAIRE 19.1 - (Formule de Vandermonde). Soient n et m deux entiers naturels. On a :

$$\sum_{k=0}^n \binom{n}{k} \binom{m}{n-k} = \binom{n+m}{n}$$

PREUVE. Soient n et m deux entiers naturels.

Posons $P = (1 + X)^n$ et $Q = (1 + X)^m$. On a : $PQ = \sum_{k=0}^{n+m} \sum_{i=0^k} a_i b_{k-i}$.

D'après la formule précédente, le coefficient de X^n dans PQ est donc : $\sum_{i=0}^n \binom{n}{i} \binom{m}{n-i}$.

D'autre part, puisque $PQ = (1 + X)^{n+m}$, le coefficient de X^n dans PQ est : $\binom{n+m}{n}$.

En identifiant les deux expressions, on obtient : $\binom{n+m}{n} = \sum_{i=0}^n \binom{n}{i} \binom{m}{n-i}$ ☕

19.2 Degré et coefficient dominant d'un polynôme

DÉFINITION 19.5 - Soit $P = \sum_{k=0}^n a_k X^k$ un polynôme non nul.

On appelle **degré** de P et on note $\deg(P)$ l'entier naturel :

$$\deg(P) = \max \{k \in \mathbb{N}, a_k \neq 0\}$$

Par convention : $\deg(\tilde{0}) = -\infty$.

Notons que, pour un polynôme non nul $P = \sum_{k=0}^n a_k X^k$, l'existence d'un plus grand élément dans l'ensemble $E = \{k \in \mathbb{N}, a_k \neq 0\}$ est assurée par le fait que E une partie non vide (puisque P est non nul) et majorée (puisque tous les coefficients de P sont nuls à partir d'un certain rang) de $\mathbb{N}$.

Remarque. On prendra garde au fait que l'écriture $P = \sum_{k=0}^n a_k X^k$ ne désigne pas en général un polynôme de degré n , mais un polynôme de degré au plus égal à n ; si l'on veut avoir $\deg(P) = n$, il faut de plus imposer $a_n \neq 0$.

Notation. $\mathbb{K}_n[X]$ désigne l'ensemble des polynômes de degré inférieur ou égal à n .

PROPRIÉTÉ 19.5 - (Propriétés du degré). Soient P et Q dans $\mathbb{K}[X]$. On a :

$$1/ \deg(P + Q) \leq \max(\deg(P), \deg(Q))$$

$$2/ \deg(PQ) = \deg(P) + \deg(Q)$$

$$3/ \forall \lambda \in \mathbb{K}^*, \deg(\lambda P) = \deg(P)$$

$$4/ \deg(P \circ Q) = \deg(P) \times \deg(Q)$$

PREUVE. 1/ Si P ou Q est le polynôme nul, alors la formule est immédiate (en prenant pour conventions que $-\infty \leq -\infty$, et que $-\infty \leq n$ pour tout entier n).

Supposons donc P et Q non nuls. Alors on peut noter $P = \sum_{k=0}^n a_k X^k$ et $Q = \sum_{k=0}^m b_k X^k$, où n et m désignent les degrés respectifs de P et de Q (ce qui implique que a_n et b_m sont non nuls). On distingue deux cas.

► 1er cas — Si $n \neq m$. S.N.A.L.G., on peut supposer $n > m$. On a alors : $P + Q = a_n X^n + \sum_{k=m+1}^{n-1} a_k X^k + \sum_{k=0}^m (a_k + b_k) X^k$.

Il s'ensuit, puisque $a_n \neq 0$, que : $\deg(P + Q) = n$, donc : $\deg(P + Q) = \max(\deg(P), \deg(Q))$.

► 2ème cas — Si $n = m$. On a alors : $P + Q = (a_n + b_n) X^n + \sum_{k=0}^{n-1} (a_k + b_k) X^k$.

Il ressort de cette écriture que si $a_n + b_n \neq 0$, alors $\deg(P + Q) = n$. Et si $a_n + b_n = 0$, alors $\deg(P + Q) < n$. Dans les deux cas, on a : $\deg(P + Q) \leq n$, donc : $\deg(P + Q) \leq \max(\deg(P), \deg(Q))$.

► Conclusion : $\forall (P, Q) \in \mathbb{K}[X]^2, \deg(P + Q) \leq \max(\deg(P), \deg(Q))$

2/ Si P ou Q est le polynôme nul, la formule est immédiate.

Supposons donc P et Q non nuls. On conserve les mêmes notations que dans la précédente preuve.

$$\text{On a : } PQ = \left(\sum_{k=0}^n a_k X^k \right) \left(\sum_{k=0}^m b_k X^k \right) = \sum_{k=0}^{n+m} \left(\sum_{i=0}^k a_i b_{k-i} \right) X^k = a_n b_m X^{n+m} + \sum_{k=0}^{n+m-1} \left(\sum_{i=0}^k a_i b_{k-i} \right) X^k$$

$$\text{On a clairement } \deg(a_n b_m X^{n+m}) = n + m \text{ (car } a_n b_m \neq 0) \text{ et } \deg \left(\sum_{k=0}^{n+m-1} \left(\sum_{i=0}^k a_i b_{k-i} \right) X^k \right) < n + m.$$

$$\text{Donc : } \deg(PQ) = n + m, \text{ c\`ad : } \boxed{\deg(PQ) = \deg(P) + \deg(Q)}$$

3/ On applique la formule précédente au cas particulier où $Q = \lambda$ est un polynôme constant non nul (on a alors $\deg(Q) = 0$) pour obtenir : $\boxed{\forall \lambda \in \mathbb{K}^*, \deg(\lambda P) = \deg(P)}$.

4/ Avec les notations utilisées depuis le début de la preuve, on a : $P(Q)(X) = a_n b_m^n X^{nm} + B(X)$, où $B(X)$ est un bazar de termes de degré strictement inférieur à nm .

On en déduit que : $\deg(P \circ Q) = \deg(P) \deg(Q)$. ☕

Exemple : considérons les polynômes $P = (1 + X)^n$ et $Q = (1 - X)^n$, où n désigne un entier naturel supérieur ou égal à 2.

On a : $P = X^n + nX^{n-1} + \sum_{k=0}^{n-2} X^k$ et $Q = (-1)^n X^n + (-1)^{n-1} nX^{n-1} + \sum_{k=0}^{n-2} (-1)^k X^k$.

Lorsque n est pair, on a : $P + Q = 2X^n + \sum_{k=0}^{n-1} (1 + (-1)^k) X^k$. D'où : $\deg(P + Q) = n$.

Lorsque n est impair, on a : $P + Q = 2nX^{n-1} + \sum_{k=0}^{n-2} (1 + (-1)^k) X^k$. D'où : $\deg(P + Q) = n - 1$.

Exemple d'utilisation du degré — résolution d'une équation polynomiale

Problème : résoudre dans $\mathbb{K}[X]$ l'équation $X^2 P(X) = P(X^2)$

► **Cas particulier du polynôme nul.** Clairement, le polynôme $\tilde{0}$ est solution de l'équation (♠).

► **Supposons à présent que $P \neq \tilde{0}$ est solution de l'équation.** Puisque P est non nul, il a un degré entier naturel ; notons $n = \deg(P)$.

En comparant les degrés des termes de gauche et de droite de l'équation, on obtient : $n + 2 = 2n$, ce qui entraîne $n = 2$. Il existe donc trois scalaires a, b et c , avec $a \neq 0$ tels que $P = aX^2 + bX + c$. On en déduit que P est solution de l'équation si :

$$aX^4 + bX^3 + cX^2 = aX^4 + bX^2 + c$$

et puisque deux polynômes sont égaux si et seulement si tous leurs coefficients sont deux à deux égaux, on obtient $b = 0$ et $c = 0$. Par suite : $P = aX^2$.

Réciproquement, tout polynôme P s'écrivant aX^2 (avec $a \neq 0$) est solution de l'équation. On peut donc conclure que les polynômes non nuls solutions de l'équation sont tous les aX^2 (avec $a \neq 0$) (♣).

Conclusion : d'après (♠) et (♣), les solutions de $X^2 P(X) = P(X^2)$ sont les polynômes aX^2 , avec $a \in \mathbb{K}$ arbitraire.

Dans un sens que nous ne tarderons pas à rendre plus précis, nous dirons dans cette situation que l'ensemble des solutions de l'équation est une droite vectorielle dans l'espace vectoriel des polynômes à coefficients dans $\mathbb{K}$, et plus précisément qu'il s'agit de la droite vectorielle engendrée par X^2 .

COROLLAIRE 19.2 - Soit n un entier naturel quelconque.

$\mathbb{K}_n[X]$ est stable pour la somme et la multiplication par un scalaire, c'est à dire :

$$\forall (P, Q) \in (\mathbb{K}_n[X])^2, \forall \lambda \in \mathbb{K}, P + \lambda Q \in \mathbb{K}_n[X]$$

PREUVE. Conséquence immédiate des propriétés du degré. ☕

Il revient d'ailleurs au même de dire que $\mathbb{K}_n[X]$ est **stable par combinaison linéaire**, c-à-d : si P et Q sont dans $\mathbb{K}_n[X]$, λ et μ dans $\mathbb{K}$, alors $(\lambda P + \mu Q) \in \mathbb{K}_n[X]$.

COROLLAIRE 19.3 - $(\mathbb{K}[X], +, \times)$ est un anneau commutatif **intègre**.

PREUVE. Soient P et Q deux polynômes dans $\mathbb{K}[X]$ tels que $PQ = \tilde{0}$. Alors : $\deg(P) + \deg(Q) = -\infty$, égalité qui n'est vérifiée que si au moins l'un des deux polynômes est nul.

Par suite : $[PQ = \tilde{0}] \implies [(P = \tilde{0}) \vee (Q = \tilde{0})]$. L'anneau $\mathbb{K}[X]$ est donc intègre. ☕

DÉFINITION 19.6 - Soit $P = \sum_{k=0}^{\deg(P)} a_k X^k$ un polynôme non nul.

On appelle **coefficient dominant** de P et on note $\text{cd}(P)$ le scalaire $a_{\deg(P)}$.

DÉFINITION 19.7 - Un polynôme est dit **unitaire** lorsque son coefficient dominant est égal à 1.

PROPRIÉTÉ 19.6 - Soient P et Q deux polynômes non nuls dans $\mathbb{K}[X]$. On a :

$$\text{cd}(PQ) = \text{cd}(P) \text{cd}(Q)$$

PREUVE. Puisque P et Q sont supposés non nuls, on peut noter $P = \sum_{k=0}^n a_k X^k$ et $Q = \sum_{k=0}^m b_k X^k$, où n et m désignent les degrés respectifs de P et de Q (ce qui implique que a_n et b_m sont non nuls).

$$\text{On a alors : } PQ = a_n b_m X^{n+m} + \underbrace{a_n X^n \sum_{k=0}^{m-1} b_k X^k}_{\deg \leq n+m-1} + \underbrace{b_m \sum_{k=0}^{n-1} a_k X^k}_{\deg \leq n+m-1} + \underbrace{\left(\sum_{k=0}^{n-1} a_k X^k \right) \left(\sum_{k=0}^{m-1} b_k X^k \right)}_{\deg \leq n+m-2}$$

Il s'ensuit que $\text{cd}(PQ) = a_n b_m$, c'est à dire : $\text{cd}(PQ) = \text{cd}(P) \text{cd}(Q)$. ☕

Dans certaines équations polynomiales, l'étude du degré ne suffit plus à conclure directement (comme on l'a vu un peu plus haut). C'est le cas de l'équation présentée ci-dessous.

Exemple d'utilisation du coefficient dominant — résolution d'une équation polynomiale

$$\text{Problème : résoudre dans } \mathbb{K}[X] \text{ l'équation (E)} \quad (X^2 + 1)P'' = 6P$$

Commençons par observer que le polynôme nul est trivialement solution de l'équation.

Supposons à présent que P soit un polynôme non nul solution de l'équation (E), et notons $n = \deg(P)$.

► Les cas $n = 0$ et $n = 1$ sont impossibles, puisqu'alors $P'' = \tilde{0}$.

► On suppose donc que $n \geq 2$. En notant a_n le coefficient dominant de P , on a d'une part $\text{cd}(6P) = 6a_n$ et d'autre part $\text{cd}((X^2 + 1)P'') = n(n-1)a_n$. Puisque $a_n \neq 0$ (par définition de coefficient dominant), on déduit de ces calculs que $n(n-1) = 6$. Comme cette équation du second degré a comme racines 3 et -2, on obtient finalement $n = 3$. On doit donc chercher les solutions (non nulles) de (E) parmi les polynômes de degré 3.

Posons donc : $P = aX^3 + bX^2 + cX + d$. On a : $(X^2 + 1)P'' = 6aX^3 + 2bX^2 + 6aX + 2b$, et : $6P = 6aX^3 + 6bX^2 + 6cX + 6d$.

$$\text{Par identification : } \begin{cases} 6a = 6a \\ 2b = 6b \\ 6a = 6c \\ 6d = 2b \end{cases} \iff \begin{cases} a = c \\ b = 0 \\ d = 0 \end{cases}$$

On en déduit que les polynômes de degré 3 solutions de (E) sont les polynômes $aX^3 + a = a(X^3 + X)$, avec $a \in \mathbb{K}^*$.

Puisque la formule donne encore un polynôme solution pour $a = 0$ (le polynôme $\tilde{0}$ étant solution de (E)), on peut conclure que :

L'ensemble des solutions est l'ensemble $\{a(X^3 + X) \mid a \in \mathbb{K}\}$. On peut encore noter cet ensemble : $\mathbb{K}(X^3 + X)$.

Dans un sens que nous tarderons pas à préciser, l'ensemble des solutions est l'ensemble des polynômes *colinéaires* au polynôme $(X^3 + X)$; nous dirons encore qu'il s'agit de la *droite vectorielle engendrée par le polynôme $(X^3 + X)$* .

L'utilisant des coefficients dominants permet par ailleurs de préciser la formule donnant le degré d'une somme de polynômes.

PROPRIÉTÉ 19.7 - Soient P et Q dans $\mathbb{K}[X]$. On a :

$$\deg(P + Q) = \max(\deg(P), \deg(Q))$$

$$\text{SSI } \begin{cases} \deg(P) \neq \deg(Q) \\ \text{ou} \\ [\deg(P) = \deg(Q)] \wedge [\text{cd}(P) + \text{cd}(Q) \neq 0] \end{cases}$$

PREUVE. Débarrassons-nous des cas triviaux : celui où l'un des polynômes est nul, ou celui où les degrés des polynômes sont distincts, pour lesquels la formule est évidente.

Considérons donc deux polynômes P et Q non nuls et de même degré n , de telle sorte que : $P = \sum_{k=0}^n a_k X^k$

et $Q = \sum_{k=0}^n b_k X^k$, $\text{cd}(P) = a_n$ et $\text{cd}(Q) = b_n$. Alors on a :

$$P + Q = (a_n + b_n)X^n + \underbrace{\sum_{k=0}^{n-1} (a_k + b_k) X^k}_{\deg \leq n-1}$$

On déduit de cette écriture que $\deg(P + Q) = n + m$ si et seulement si $a_n + b_n \neq 0$.

Autrement écrit : $\deg(P + Q) = \deg(P) + \deg(Q)$ si et seulement si $\text{cd}(P) + \text{cd}(Q) \neq 0$. ☕

19.3 Arithmétique dans $\mathbb{K}[X]$

19.3.1 Diviseurs et multiples, division euclidienne dans $\mathbb{K}[X]$

DÉFINITION 19.8 - Soient P et Q deux polynômes de $\mathbb{K}[X]$. On dit que P **divise** Q (ou que Q est un **multiple** de P) s'il existe un polynôme $M \in \mathbb{K}[X]$ tel que : $Q = PM$.

Exemples. Pour tout entier n non nul, le polynôme $X^n - 1$ est multiple du polynôme $\sum_{k=0}^{n-1} X^k$. Le polynôme nul ne divise que le polynôme nul, mais est multiple de tout polynôme. Un polynôme constant non nul divise tout polynôme. Le polynôme $X - j$ divise $X^2 + X + 1$. Le polynôme $X + 1$ divise $X^{2n} - 1$.

PROPRIÉTÉ 19.8 - (**Propriétés de la divisibilité dans $\mathbb{K}[X]$**). Soient A, B, C, P et Q cinq polynômes dans $\mathbb{K}[X]$.

1/ $A|A$ (**réflexivité**).

2/ Si $A|B$ et $B|C$, alors $A|C$ (**transitivité**).

3/ Si $A|B$ et $A|C$, alors $A|(BP + CQ)$.

PREUVE. 1/ Triviale. 2/ Supposons que $A|B$ et $B|C$. Il existe deux polynômes S et T tels que : $B = SA$ et $C = TB$. Par suite : $C = (ST)A$. Ce qui signifie que A divise C .

3/ Supposons que $A|B$ et $A|C$. Il existe deux polynômes S et T tels que : $B = SA$ et $C = TA$. Par suite : $BP + CQ = (PS + QT)A$. Ce qui signifie que A divise $BP + CQ$. ☕

Remarque. Il ne vous a pas échappé qu'une propriété (l'antisymétrie) manque à l'appel pour faire de la relation de divisibilité une relation d'ordre : si P et Q sont deux polynômes, les assertions $P|Q$ et $Q|P$ n'impliquent pas que $P = Q$ (considérer $P = X$ et $Q = 2X$ par exemple). Nous aurons l'occasion de revenir sur cette situation lorsque nous étudierons les polynômes *associés*.

En attendant, on peut déjà observer que la relation de divisibilité n'est pas une relation d'ordre dans $\mathbb{Z}$, mais qu'elle en est une dans $\mathbb{N}$. Le pendant de cette assertion dans $\mathbb{K}[X]$ est que la relation de divisibilité est une relation d'ordre sur l'ensemble des polynômes de $\mathbb{K}[X]$ qui sont *unitaires*.

THÉORÈME 19.1 - (**Théorème de la division euclidienne dans $\mathbb{K}[X]$**). Soient A et B dans $\mathbb{K}[X]$ avec $B \neq \tilde{0}$.

Il existe un unique couple (Q, R) dans $(\mathbb{K}[X])^2$ tel que : $A = BQ + R$ et $\deg R < \deg B$.

PREUVE. Soient A et B dans $\mathbb{K}[X]$ avec $B \neq \tilde{0}$.

Existence. ► Lorsque $\deg(A) < \deg(B)$, on peut écrire $A = B \times \tilde{0} + A$, et le couple $(\tilde{0}, A)$ convient.

► Supposons donc $\deg(A) \geq \deg(B)$: on montre l'existence d'un couple (Q, R) satisfaisant les conditions du théorème par récurrence sur le degré $n \in \mathbb{N}$ du polynôme A .

Initialisation : lorsque $\deg(A) = 0$, on a nécessairement $\deg(B) = 0$ (puisque $\deg(A) \geq \deg(B)$ et B non nul). A et B sont donc constants non nuls, et on peut donc exceptionnellement écrire $A = B \times (A/B) + \tilde{0}$. Le couple $(A/B, \tilde{0})$ convient, ce qui établit la propriété pour $n = 0$.

Hérédité : supposons la propriété établie jusqu'au rang n , et montrons qu'elle est vraie au rang $n + 1$. On considère donc un polynôme A de degré $n + 1$, et un polynôme B de degré (noté m) inférieur ou égal à $n + 1$. Il existe donc tout plein de scalaires tels que :

$$A = a_{n+1}X^{n+1} + \sum_{k=0}^n a_k X^k \quad \text{et} \quad B = b_m X^m + \sum_{k=0}^{m-1} b_k X^k, \quad \text{avec } a_{n+1} \text{ et } b_m \text{ non nuls.}$$

Posons alors : $P = A - \frac{a_{n+1}}{b_m} X^{n+1-m} B$ avec la motivation secrète (enfin pas tant que cela) de faire disparaître le terme de degré $n + 1$. On a :

$$P = \underbrace{\sum_{k=0}^n a_k X^k}_{\deg \leq n} - \underbrace{\frac{a_{n+1}}{b_m} \sum_{k=0}^{m-1} b_k X^{k+n+1-m}}_{\deg \leq n}. \quad \text{D'où : } \deg(P) \leq n.$$

On peut donc appliquer l'hypothèse de récurrence au polynôme P : il existe un couple (Q, R) de polynômes tels que $P = BQ + R$ et $\deg(R) < \deg(B)$. D'où : $A = B \left(Q + \frac{a_{n+1}}{b_m} X^{n+1-m} \right) + R$, et le couple $\left(Q + \frac{a_{n+1}}{b_m} X^{n+1-m}, R \right)$ convient.

Ce qui établit la propriété au degré $n + 1$, et l'hérédité de la propriété.

Conclusion intermédiaire. $\forall (A, B) \in \mathbb{K}[X]^2, B \neq \tilde{0}$,

$$\exists (Q, R) \in \mathbb{K}[X]^2, A = BQ + R \text{ et } \deg(R) < \deg(B).$$

Unicité. Supposons qu'il existe (Q_1, R_1) et (Q_2, R_2) deux couples de polynômes tels que $A = BQ_i + R_i$ et $\deg(R_i) < \deg(B)$. Alors : $BQ_1 + R_1 = BQ_2 + R_2$, d'où : $B(Q_1 - Q_2) = R_2 - R_1$. Le polynôme $B(Q_1 - Q_2)$, s'il n'est pas nul, est de degré supérieur ou égal à celui de B ; tandis que celui du terme de droite est strictement inférieur à celui de B : contradiction. Il s'ensuit que $B(Q_1 - Q_2) = \tilde{0}$, d'où $Q_1 - Q_2 = \tilde{0}$ (puisque $\mathbb{K}[X]$ est intègre et que B est supposé non nul).

Ainsi $Q_1 = Q_2$, et il en découle immédiatement que $R_1 = R_2$. Ce qui prouve l'unicité du couple dont l'existence a été préalablement établie.

Conclusion. $\forall (A, B) \in \mathbb{K}[X]^2, B \neq \tilde{0}, \exists! (Q, R) \in \mathbb{K}[X]^2, A = BQ + R \text{ et } \deg(R) < \deg(B).$



Remarque. Le reste peut être déterminé explicitement soit en posant et en effectuant la division euclidienne, comme dans la division de $A = X^3 + 4X^2 + 5X - 2$ par $B = X^2 - X + 1$ (exo). Il peut aussi être déterminé en révolant un système, comme dans la division de $A = X^n$ par $B = (X - 1)(X - 2)$ (exo).

Exemple. Montrons, sur l'exemple de la division de $A = 2X^5 + 2X^4 + 3X^3 + X + 1$ par $B = 3X^2 + 3X - 1$ comment l'on pose en pratique la division euclidienne.

	$2X^5 + 2X^4 + 3X^3 + X$	$3X^2 + 3X - 1$
On écrit le produit de B par $\frac{2}{3}X^3$	$- 2X^5 + 2X^4 - \frac{2}{3}X^3$	
On fait la différence	$\frac{11}{3}X^3 + X$	
On fait le produit de B par $\frac{11}{9}X$	$- \frac{11}{3}X^3 + \frac{11}{3}X^2 - \frac{11}{9}X$	$\frac{2}{3}X^3 + \frac{11}{9}X - \frac{11}{9}$
	$- \frac{11}{3}X^2 + \frac{20}{9}X$	
	$- -\frac{11}{3}X^2 - \frac{11}{3}X + \frac{11}{9}$	
Le reste a un degré inférieur à celui de B , la division est terminée	$\frac{53}{9}X - \frac{11}{9}$	

$$\text{d'où } 2X^5 + 2X^4 + 3X^3 + X = \left(\frac{2}{3}X^3 + \frac{11}{9}X - \frac{11}{9} \right) \times (3X^2 + 3X - 1) + \frac{53}{9}X - \frac{11}{9}.$$

COROLLAIRE 19.4 - Soient A et B deux polynômes de $\mathbb{K}[X]$, avec B non nul.

A est multiple de B si et seulement si le reste dans la division euclidienne de A par B est nul.

PREUVE. Supposons que le reste dans la division euclidienne de A par B soit nul. Alors il existe un (unique) polynôme Q tel que $A = BQ + \tilde{0}$, d'où A est multiple de B . La preuve de l'assertion réciproque est au moins aussi évidente. ☕

19.3.2 Polynômes associés

DÉFINITION 19.9 - Deux polynômes P et $Q \in \mathbb{K}[X]$ sont **associés** s'ils se divisent mutuellement.

Exemples. Les polynômes X^2 et $3X^2$ sont associés. Les polynômes $X - 1$ et $X^2 - 1$ ne le sont pas. Les polynômes associés au polynôme constant $\tilde{1}$ sont les polynômes constants non nuls.

PROPRIÉTÉ 19.9 - Soient P et Q deux polynômes dans $\mathbb{K}[X]$.

$$[P \text{ et } Q \text{ sont associés}] \text{ SSI } [\exists \lambda \in \mathbb{K}^*, Q = \lambda P]$$

PREUVE. Supposons P et Q associés. Alors il existe deux polynômes R et S tels que : $Q = SP$ et $P = RQ$. On en déduit que : $P = RSQ$. Si l'un des deux polynômes R ou S est nul, alors P et Q sont tous les deux nuls, et on a alors $P = 1 \times Q \dots$

Supposons à présent que R et S sont non nuls, alors on a :

$$\deg(P) = \underbrace{\deg(R) + \deg(S)}_{\in \mathbb{N}} + \deg(Q) \quad (\spadesuit)$$

Si $\deg(P) = -\infty$, alors l'égalité $(\spadesuit)$ implique que $\deg(Q) = -\infty$, ce qui signifie que P et Q sont nuls, et on a alors $P = 1 \times Q \dots$

Si $\deg(P) \in \mathbb{N}$, alors l'égalité $(\spadesuit)$ implique que $\deg(Q) \in \mathbb{N}$. En notant p, q, r et s les degrés respectifs des polynômes majuscules correspondants, les hypothèses entraînent que :

$$q = p + s \text{ et } p = r + q \quad \text{d'où} \quad q = q + r + s$$

Comme $r + s$ est un entier naturel, cette dernière égalité impose que $r + s = 0$, et donc (comme r et s sont dans $\mathbb{N}$) : $r = s = 0$.

On en déduit qu'il existe deux scalaires non nuls λ et μ tels que : $S = \lambda$ et $R = \mu$. En particulier : $Q = \lambda P$.

Conclusion intermédiaire. $[P \text{ et } Q \text{ sont associés}] \implies [\exists \lambda \in \mathbb{K}^*, Q = \lambda P]$

Réciproquement, on suppose qu'il existe $\lambda \in \mathbb{K}^*$ tel que $Q = \lambda P$. Alors P divise Q ; et puisque $P = (1/\lambda)Q$, Q divise P . Donc P et Q sont associés.

Ce qui établit l'implication réciproque de la précédente, et permet de conclure :

$$[P \text{ et } Q \text{ sont associés}] \iff [\exists \lambda \in \mathbb{K}^*, Q = \lambda P] \quad \text{☕}$$

19.3.3 Racines et factorisation dans $\mathbb{K}[X]$

PROPRIÉTÉ 19.10 - Soient α et β deux scalaires distincts, et soit P un polynôme de $\mathbb{K}[X]$.

$$\text{Si } (X - \alpha) \mid P \text{ et } (X - \beta) \mid P, \text{ alors } (X - \alpha)(X - \beta) \mid P$$

PREUVE. Voir version longue du chapitre 20. ☕

L'énoncé précédent peut encore être étendu.

COROLLAIRE 19.5 - Soient n un entier supérieur ou égal à 2, P un polynôme de $\mathbb{K}[X]$, et $\alpha_1, \dots, \alpha_n$ et β ($n+1$) scalaires deux à deux distincts.

$$\text{Si } \prod_{i=1}^n (X - \alpha_i) \mid P \text{ et } (X - \beta) \mid P, \text{ alors } \left[(X - \beta) \prod_{i=1}^n (X - \alpha_i) \right] \mid P$$

PREUVE. Voir version longue du chapitre 20. ☕

19.4 Fonctions polynomiales

19.4.1 Racines d'un polynôme

DÉFINITION 19.10 - Soit $P \in \mathbb{K}[X]$, avec : $P = \sum_{k=0}^n a_k X^k$.

On appelle **fonction polynomiale associée à P** et on note (abusivement) P la fonction définie sur $\mathbb{K}$ en posant :

$$\forall t \in \mathbb{K}, \tilde{P}(t) = \sum_{k=0}^n a_k t^k$$

Remarque. Afin de ne pas alourdir les notations, on notera encore P cette fonction polynomiale. Cela n'aura aucune fâcheuse conséquence, car nous verrons que deux polynômes distincts ne peuvent avoir même fonction polynomiale associée (en tout cas lorsque $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}^1$).

DÉFINITION 19.11 - Soit $P \in \mathbb{K}[X]$, et soit $\alpha \in \mathbb{K}$.

Le scalaire α est une **racine** de P si $P(\alpha) = 0$.

1. Ou plus généralement, dès que $\mathbb{K}$ est un corps infini.

LEMME 19.1 - Soient $P \in \mathbb{K}[X]$ et $\alpha \in \mathbb{K}$.

Le reste dans la division euclidienne du polynôme P par le polynôme $(X - \alpha)$ est le polynôme constant $\tilde{P}(\alpha)$.

PREUVE. D'après le théorème de la division euclidienne, il existe deux polynômes Q et R tels que $P = (X - \alpha)Q + R$ et $\deg(R) < 1$. Il s'ensuit que le polynôme R est constant ; il existe $\lambda \in \mathbb{K}$ tel que $R = \tilde{\lambda}$.²

Ainsi : $P = (X - \alpha)Q + \tilde{\lambda}$.

L'évaluation en α de cette relation donne alors l'égalité $P(\alpha) = \lambda$, et la conclusion. ☕

COROLLAIRE 19.6 - Soient $P \in \mathbb{K}[X]$ et $\alpha \in \mathbb{K}$.

$$[\alpha \text{ est racine de } P] \iff [(X - \alpha) \text{ divise } P]$$

PREUVE. Si α est racine de P , alors d'après le lemme précédent, le reste dans la division euclidienne de P par $(X - \alpha)$ est nul. Il revient alors au même de dire que $(X - \alpha)$ divise P .

L'implication réciproque est triviale. ☕

COROLLAIRE 19.7 - Généralisation à k racines distinctes. Soient $P \in \mathbb{K}[X]$ et $\alpha_1, \dots, \alpha_k$ des scalaires deux à deux distincts (avec $k \geq 2$). Alors :

$$[\alpha_1, \dots, \alpha_k \text{ sont racines de } P] \iff \left[\prod_{i=1}^k (X - \alpha_i) \text{ divise } P \right]$$

PREUVE. Supposons que $\alpha_1, \dots, \alpha_k$ sont racines de P .

Alors chacun des polynômes $(X - \alpha_1), \dots, (X - \alpha_k)$ divise le polynôme P , d'après le corollaire précédent.

En appliquant itérativement le corollaire 19.5 (page 475), on en déduit que : $\prod_{i=1}^k (X - \alpha_i)$ divise P .

La réciproque est immédiate. ☕

Application 1. Pour tout entier $n \geq 1$,

$$\text{le polynôme } Q = 2X^3 + 3X^2 + X \text{ divise } P = (X + 1)^{2n} - X^{2n} - 2X - 1.$$

Il est en effet assez clair que Q possède exactement trois racines réelles, qui sont 0, -1 et $-1/2$. On en déduit que $Q = 2(X + 1) \left(X + \frac{1}{2} \right) X$.

2. Il ne faudra pas m'en vouloir si j'oublie à l'avenir les " $\sim$ ", et si j'identifie λ et $\tilde{\lambda}$ lorsqu'il n'y a aucune ambiguïté.

Par ailleurs, on peut vérifier facilement que 0 , -1 et $-1/2$ sont racines de P . Puisque ces trois réels sont deux à deux distincts, on en déduit que $(X+1)\left(X+\frac{1}{2}\right)X$ divise P .

Par conséquent, Q divise P .

Application 2. Soient n , m et p trois entiers naturels. Le polynôme $1+X+X^2$ divise le polynôme $X^{3n}+X^{3m+4}+X^{3p+17}$ (exercice).

L'énoncé suivant doit sa dénomination de "théorème" à sa grande importance pratique.

THÉORÈME 19.2 - (Nombre maximal de racines d'un polynôme).

Si $P \in \mathbb{K}_n[X]$ admet $(n+1)$ racines distinctes, alors $P = \tilde{0}$.

En d'autres termes, un polynôme non nul ne peut posséder strictement plus de racines que son degré.

PREUVE. Soit P un polynôme de $\mathbb{K}_n[X]$, avec $n \in \mathbb{N}$. Supposons que P possède $(n+1)$ racines distinctes $\alpha_1, \dots, \alpha_{n+1}$. Alors P est divisible par $\prod_{i=1}^{n+1} (X - \alpha_i)$.

Il existe donc un polynôme Q tel que : $P = Q \times \prod_{i=1}^{n+1} (X - \alpha_i)$.

Si Q est non nul, alors : $\deg\left(Q \times \prod_{i=1}^{n+1} (X - \alpha_i)\right) \geq n+1$, et par suite : $\deg\left(Q \times \prod_{i=1}^{n+1} (X - \alpha_i)\right) > \deg(P)$.

Contradiction.

Il s'ensuit que Q est nul, ce qui implique que P l'est.

Conclusion. Si $P \in \mathbb{K}_n[X]$ admet $(n+1)$ racines distinctes, alors $P = \tilde{0}$.



COROLLAIRE 19.8 - (Principe du prolongement algébrique). Soient P et Q dans $\mathbb{K}_n[X]$.

S'il existe $(n+1)$ scalaires $(\alpha_i)_{1 \leq i \leq n+1}$ deux à deux distincts tels que :

$$\forall i \in \llbracket 1, n+1 \rrbracket, P(\alpha_i) = Q(\alpha_i)$$

alors $P = Q$.

PREUVE. Soient P et Q dans $\mathbb{K}_n[X]$. Supposons qu'il existe $(n+1)$ scalaires $(\alpha_i)_{1 \leq i \leq n+1}$ distincts tels que : $\forall i \in \llbracket 1, n+1 \rrbracket, P(\alpha_i) = Q(\alpha_i)$.

On pose judicieusement : $R = P - Q$. Le polynôme R est de degré au plus n (puisque P et Q le sont), et possède $(n+1)$ racines distinctes par hypothèse. D'après le théorème précédent, $R = \tilde{0}$, d'où $P = Q$.

Conclusion. Soient P et Q dans $\mathbb{K}_n[X]$. S'il existe $(n+1)$ scalaires $(\alpha_i)_{1 \leq i \leq n+1}$ deux à deux distincts tels que : $\forall i \in \llbracket 1, n+1 \rrbracket, P(\alpha_i) = Q(\alpha_i)$, alors $P = Q$.



Application - Polynômes interpolateurs de Lagrange (un cas particulier).

$$\begin{aligned} \text{L'application } F : \mathbb{R}_2[X] &\longrightarrow \mathbb{R}^3 && \text{est bijective.} \\ P &\longmapsto (P(-1), P(0), P(1)) \end{aligned}$$

Prouvons son injectivité. Soient P et Q deux polynômes de $\mathbb{R}_2[X]$ tels que $F(P) = F(Q)$. On a alors $P(-1) = Q(-1)$, $P(0) = Q(0)$ et $P(1) = Q(1)$. Puisque les polynômes P et Q sont de degré au plus 2, et qu'ils prennent les mêmes valeurs en trois scalaires distincts, le principe du prolongement algébrique permet d'affirmer que : $P = Q$.

Conclusion. $\forall (P, Q) \in (\mathbb{R}_2[X])^2$, $[F(P) = F(Q)] \implies [P = Q]$. Ainsi F est injective.

Prouvons sa surjectivité. Soit $(x, y, z) \in \mathbb{R}^3$. On construit un antécédent (qui sera unique puisque F est injective) de (x, y, z) par F par le biais des *polynômes interpolateurs de Lagrange associés aux valeurs -1 , 0 et 1* .

Explicitement, il s'agit des polynômes L_{-1} , L_0 et L_1 respectivement définis par :

$$L_{-1} = \frac{1}{2}X(X-1); \quad L_0 = -(X^2-1); \quad L_1 = \frac{1}{2}X(X+1)$$

Ces trois polynômes sont clairement des éléments de $\mathbb{R}_2[X]$ (tous trois sont de degré 2), et vérifient les propriétés suivantes :

$$\left\{ \begin{array}{l} L_{-1}(-1) = 1 \\ L_{-1}(0) = 0 \\ L_{-1}(1) = 0 \end{array} \right. ; \quad \left\{ \begin{array}{l} L_0(-1) = 0 \\ L_0(0) = 1 \\ L_0(1) = 0 \end{array} \right. ; \quad \left\{ \begin{array}{l} L_1(-1) = 0 \\ L_1(0) = 0 \\ L_1(1) = 1 \end{array} \right.$$

On pose alors : $P = xL_{-1} + yL_0 + zL_1$. Il résulte des propriétés ci-dessus que P est un polynôme de degré au plus 2, et on vérifie aisément que : $P(-1) = x$, $P(0) = y$ et $P(1) = z$. En d'autres termes $F(P) = (x, y, z)$.

Conclusion. $\forall (x, y, z) \in \mathbb{R}^3$, $\exists P \in \mathbb{R}_2[X]$, $F(P) = (x, y, z)$. Ainsi F est surjective.

Conclusion (des conclusions). L'application F est bijective.



Remarques. Il résulte de la preuve que la bijection réciproque de F est $F^{-1} : \mathbb{R}^3 \longrightarrow \mathbb{R}_2[X]$,
 $(x, y, z) \longmapsto xL_{-1} + yL_0 + zL_1$

où $L_{-1} = \frac{1}{2}X(X-1)$, $L_0 = -(X^2-1)$ et $L_1 = \frac{1}{2}X(X+1)$ désignent les polynômes interpolateurs de Lagrange associés aux valeurs -1 , 0 et 1 .

Par ailleurs, le point-clef pour établir que F est bijective est évidemment le principe du prolongement algébrique. Ce principe signifie au fond qu'un polynôme de degré $\leq n$ est un objet un peu "rigide", dans le sens où il est entièrement déterminé par la donnée de $(n+1)$ scalaires. Il est alors instructif d'observer que ces scalaires peuvent être choisis de façons diverses et variées.

Pour être plus explicite, dans le cas où $n = 2$, on peut construire par exemple 3 bijections de $\mathbb{R}_2[X]$ dans $\mathbb{R}^3$:

1) En associant à un polynôme P ses coefficients

$$F_1 : P = aX^2 + bX + c \in \mathbb{R}_2[X] \mapsto \begin{pmatrix} a \\ b \\ c \end{pmatrix} \in \mathbb{R}^3$$

$$F_2 : P \in \mathbb{R}_2[X] \mapsto \begin{pmatrix} P(\alpha) \\ P'(\alpha) \\ P''(\alpha) \end{pmatrix} \in \mathbb{R}^3$$

2) En associant à un polynôme P ses dérivées successives (en un qcque réel α ; lien avec la formule de Taylor)

3) En associant à un polynôme P ses valeurs en trois réels (par exemple $-1, 0$ et 1 ; voir plus haut)

$$F_3 : P \in \mathbb{R}_2[X] \mapsto \begin{pmatrix} P(-1) \\ P(0) \\ P(1) \end{pmatrix} \in \mathbb{R}^3$$

Ultime conséquence du corollaire 19.7 :

COROLLAIRE 19.9 - Soit $P \in \mathbb{K}_n[X]$, non nul, avec $n \geq 2$.

Si P possède n racines $\alpha_1, \dots, \alpha_n$ deux à deux distinctes, alors :

$$P = \text{cd}(P) \prod_{i=1}^n (X - \alpha_i)$$

PREUVE. Sous les hypothèses de l'énoncé, on déduit du corollaire 19.7 que $\prod_{i=1}^n (X - \alpha_i)$ divise P . Il existe donc un polynôme Q tel que :

$$Q \prod_{i=1}^n (X - \alpha_i) = P$$

Pour des raisons évidentes liées au degré, on a : $\deg Q = 0$. Il existe donc un scalaire λ tel que : $Q = \lambda$.

On a alors : $\lambda \prod_{i=1}^n (X - \alpha_i) = P$. La comparaison des coefficients dominants de chaque terme de cette égalité donne $\lambda = \text{cd}(P)$, et la conclusion. ☕

Exemple d'application. Soit $n \geq 2$ et $P = (X + 1)^n - (X - 1)^n$.

Le polynôme P est de degré $n - 1$ et de coefficient dominant égal à $2n$. Les racines du polynôme P sont les $(n - 1)$ nombres complexes $z_k = i \cotan \frac{k\pi}{n}$ pour $k = 1, \dots, n - 1$.

On en déduit que :

$$(X + 1)^n - (X - 1)^n = 2n \prod_{k=1}^{n-1} \left(X - i \cotan \frac{k\pi}{n} \right).$$

19.4.2 Dérivation (formelle) des polynômes

DÉFINITION 19.12 - Soit $P \in \mathbb{K}[X]$ un polynôme non-constant (càd de degré ≥ 1), que nous noterons : $P = \sum_{k=0}^{\deg(P)} a_k X^k$.

Le **polynôme dérivé** de P est défini en posant :

$$P' = \sum_{k=0}^{\deg(P)} k a_k X^{k-1}$$

Lorsque P est constant, on pose $P' = \tilde{0}$.

PROPRIÉTÉ 19.11 - Si $P \in \mathbb{K}[X]$ et $\deg(P) \geq 1$, alors :

$$\deg(P') = \deg(P) - 1$$

En outre, si P est non constant, alors :

$$\text{cd}(P') = \deg(P) \text{ cd}(P)$$

PREUVE. Triviale². ☕

PROPRIÉTÉ 19.12 - Les règles de la dérivation usuelle (linéarité, formule donnant la dérivée d'un produit, d'une composée) s'appliquent à la dérivation des polynômes formels.

PREUVE. RAS. ☕

DÉFINITION 19.13 - Soit n un entier naturel, et soit $P \in \mathbb{K}[X]$.

On appelle **dérivée n -ième** du polynôme P et on note $P^{(n)}$ l'élément de $\mathbb{K}[X]$ défini récursivement en posant :

$$P^{(0)} = P, \text{ et pour tout } n \in \mathbb{N}^*, P^{(n)} = (P^{(n-1)})'$$

Bien entendu, le passage à la dérivée n -ième est encore une opération linéaire :

PROPRIÉTÉ 19.13 - Si P et Q sont deux polynômes de $\mathbb{K}[X]$, λ et μ deux scalaires, et $n \in \mathbb{N}$, alors

$$(\lambda P + \mu Q)^{(n)} = \lambda P^{(n)} + \mu Q^{(n)}$$

PREUVE. RAS. ☕

De même, la formule donnant la dérivée n -ième d'un produit reste valide :

PROPRIÉTÉ 19.14 - (Formule de Leibniz dans $\mathbb{K}[X]$). Pour tout couple de polynômes (P, Q) de $(\mathbb{K}[X])^2$, et pour tout entier naturel n , on a :

$$(PQ)^{(n)} = \sum_{k=0}^n \binom{n}{k} P^{(k)} \mu Q^{(n-k)} = \sum_{k=0}^n \binom{n}{k} P^{(n-k)} \mu Q^{(k)}$$

PREUVE. Prouvons la propriété par récurrence sur n .

Posons $\mathcal{P}(n) : (PQ)^{(n)} = \sum_{k=0}^n \binom{n}{k} P^{(k)} Q^{(n-k)}$ (où P et Q désignent deux polynômes à coefficients dans $\mathbb{K}$ arbitraires).

► **Initialisation** : pour $n = 0$, on a $(PQ)^{(0)} = PQ$ et $\sum_{k=0}^0 \binom{0}{k} P^{(k)} Q^{(0-k)} = \binom{0}{0} P^{(0)} Q^{(0)} = PQ$, ce qui assure que $\mathcal{P}(0)$ est vraie.

► **Hérédité** : on suppose $\mathcal{P}(n)$ vraie pour un certain entier naturel n . On a :

$$\begin{aligned} (PQ)^{(n+1)} &= \left((PQ)^{(n)} \right)' = \sum_{k=0}^n \binom{n}{k} (P^{(k)} Q^{(n-k)})' = \sum_{k=0}^n \binom{n}{k} P^{(k+1)} Q^{(n-k)} + \sum_{k=0}^n \binom{n}{k} P^{(k)} Q^{(n+1-k)} \\ &= \sum_{k=1}^{n+1} \binom{n}{k-1} P^{(k)} Q^{(n+1-k)} + \sum_{k=0}^n \binom{n}{k} P^{(k)} Q^{(n+1-k)} = \\ &= P^{(n+1)} Q + \left[\sum_{k=1}^n \left(\binom{n}{k-1} + \binom{n}{k} \right) P^{(k)} Q^{(n+1-k)} \right] + PQ^{(n+1)} \end{aligned}$$

D'où, en appliquant la relation de Pascal³ : $(PQ)^{(n+1)} = \sum_{k=0}^{n+1} \binom{n+1}{k} P^{(k)} Q^{(n+1-k)}$

Cette relation assure que la propriété $\mathcal{P}(n+1)$ est vraie, ce qui établit l'hérédité.

Conclusion. Pour tout entier naturel n , et pour tout couple de polynômes (P, Q) de $\mathbb{K}[X]$, on a :

$$(PQ)^{(n)} = \sum_{k=0}^n \binom{n}{k} P^{(k)} Q^{(n-k)} \quad ☕$$

3. $\forall n \in \mathbb{N}^*, \forall k \in \llbracket 1, n \rrbracket, \binom{n}{k-1} + \binom{n}{k} = \binom{n+1}{k}$.

Remarque. C'est l'occasion de rappeler la classique formule :

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

Une démonstration consiste à calculer la dérivée n -ième de X^{2n} de deux façons : directement, ou en écrivant $X^{2n} = X^n \times X^n$ et en utilisant la formule de Leibniz. Pour mémoire, on utilise en cours de route la formule $(X^n)^{(k)} = \frac{(-1)^k n!}{(n-k)!} X^{n-k}$.

En continuant les analogies avec ce que l'on connaît des dérivées dans le cadre des fonctions, on peut se réjouir du fait que la formule de Taylor est vraie dans $\mathbb{K}[X]$. Mieux : puisque les dérivées n -ièmes d'un polynôme sont nulles à partir d'un certain rang, la formule est exacte (sans "o"), pour peu que l'on écrive assez de termes. Explicitement :

THÉORÈME 19.3 - (Formule de Taylor dans $\mathbb{K}[X]$, en α).

Soit $P \in \mathbb{K}[X]$ non nul, et soit $\alpha \in \mathbb{K}$. On a :

$$P = \sum_{k=0}^{\deg(P)} \frac{P^{(k)}(\alpha)}{k!} (X - \alpha)^k$$

soit encore, en notant $n = \deg(P)$:

$$P = P(\alpha) + P'(\alpha)(X - \alpha) + \frac{P''(\alpha)}{2}(X - \alpha)^2 + \cdots + \frac{P^{(n)}(\alpha)}{n!}(X - \alpha)^n$$

PREUVE. La preuve se fait par récurrence sur le degré de P , noté n . Posons donc :

$$\mathcal{P}(n) : \text{"pour } P \in \mathbb{K}[X] \text{ de degré } n, P = \sum_{k=0}^{\deg(P)} \frac{P^{(k)}(\alpha)}{k!} (X - \alpha)^k,$$

► Initialisation ($n = 0$) : pour un polynôme de degré 0, càd constant (non nul), on a $P = P(\alpha)$ pour tout scalaire α , ce qui établit l'initialisation.

► Hérédité : supposons la propriété vraie au rang n pour un certain entier naturel n . Soit P un polynôme de degré $n + 1$. Alors P' étant de degré n , on peut utiliser l'hypothèse de récurrence pour écrire :

$$P' = \sum_{k=0}^n \frac{P'^{(k)}(\alpha)}{k!} (X - \alpha)^k \text{ soit : } P' = \sum_{k=0}^n \frac{P^{(k+1)}(\alpha)}{k!} (X - \alpha)^k$$

On en déduit (par intégration formelle) :

$$\exists \lambda \in \mathbb{K}, P = \lambda + \sum_{k=0}^n \frac{P^{(k+1)}(\alpha)}{(k+1)!} (X - \alpha)^{k+1}$$

Puis, par changement d'indice dans la somme :

$$\exists \lambda \in \mathbb{K}, P = \lambda + \sum_{k=1}^{n+1} \frac{P^{(k)}(\alpha)}{k!} (X - \alpha)^k$$

Il reste à observer que $P(\alpha) = \lambda$ (en évaluant en α l'égalité précédente) pour obtenir :

$$P = P(\alpha) + \sum_{k=1}^{n+1} \frac{P^{(k)}(\alpha)}{k!} (X - \alpha)^k \quad \text{puis} \quad P = \sum_{k=0}^{n+1} \frac{P^{(k)}(\alpha)}{k!} (X - \alpha)^k$$

La dernière égalité obtenue signifie que la propriété $\mathcal{P}(n+1)$ est vraie, ce qui prouve l'hérédité et clôt cette récurrence.

Conclusion : $\forall P \in \mathbb{K}[X], \forall \alpha \in \mathbb{K}, P \neq 0, P = \sum_{k=0}^{\deg(P)} \frac{P^{(k)}(\alpha)}{k!} (X - \alpha)^k$ ☕

Même si c'est un peu redondant, on énonce ci-dessous la formule de Taylor en zéro, car elle illustre le lien existant entre les coefficients d'un polynôme et ses dérivées successives en zéro.

THÉORÈME 19.4 - (Formule de Taylor dans $\mathbb{K}[X]$, en 0).

Soit $P \in \mathbb{K}[X]$ non nul. Alors :

$$P = \sum_{k=0}^{\deg(P)} \frac{P^{(k)}(0)}{k!} X^k$$

soit encore, en notant $n = \deg(P)$:

$$P = P(0) + P'(0)X + \frac{P''(0)}{2}X^2 + \cdots + \frac{P^{(n)}(0)}{n!}X^n$$

PREUVE. Prendre $\alpha = 0$ dans l'énoncé précédent ! ☕

Remarque. “Dans l'autre sens”, la formule ci-dessus signifie que :

si $P = \sum_{k=0}^n a_k X^k,$	alors pour tout $k \in \llbracket 0, n \rrbracket, \quad a_k = k! P^{(k)}(0)$

Quelques conséquences et applications de la formule de Taylor dans $\mathbb{K}[X]$

Observation 1. L'idée clef de la formule de Taylor est que l'on peut "reconstruire" un polynôme à partir de ses dérivées successives en un point. Ou encore qu'un polynôme P de degré (au plus) n est entièrement déterminé par les $(n+1)$ valeurs $P(\alpha), P'(\alpha), \dots, P^{(n)}(\alpha)$.

Explicitement, il résulte immédiatement de la formule de Taylor dans $\mathbb{R}_2[X]$ que l'application

$$F_2 : P \in \mathbb{R}_2[X] \mapsto \begin{pmatrix} P(\alpha) \\ P'(\alpha) \\ P''(\alpha) \end{pmatrix} \in \mathbb{R}^3$$

(déjà évoquée page 478) est bijective.

Observation 2. Encore plus précisément, la formule de Taylor en 0 montre que les coefficients d'un polynôme sont (à des constantes multiplicatives près) ses dérivées successives en 0.

Exemple d'application. Problème — Déterminer $P \in \mathbb{K}[X]$ tel que $P(1) = 1$

$$\text{et } 4P = (X-1)P' + P''.$$

On peut tout d'abord chercher le degré de P . Si $\deg P = n$ avec $n \geq 0$ ($P \neq 0$ car $P(1) = 1$), $P = \sum_{k=0}^n a_k X^k$ où $a_n \neq 0$ alors

$$4P = 4a_n X^n + \dots = (X-1)(na_n X^{n-1} + \dots) + n(n-1)a_n X^{n-2} + \dots$$

d'où $4a_n = na_n$ et ainsi $n = 4$. On peut chercher P par coefficients indéterminés (exercice) ou bien penser à la formule de Taylor au point 1.

Notons (E) l'équation $4P(X) = (X-1)P'(X) + P''(X)$. Avec $X = 1$ dans (E) , on obtient $4P(1) = 4 = P''(1)$. Dérivons (E) , on obtient $(E') : 4P'(X) = P'(X) + (X-1)P''(X) + P^{(3)}(X)$ qui donne en $X = 1$, $3P'(1) = P^{(3)}(1)$. On ne change pas de méthode, on dérive (E') pour obtenir $(E'') : 3P''(X) = P''(X) + (X-1)P^{(3)}(X) + P^{(4)}(X)$ et $2P''(1) = 8 = P^{(4)}(1)$. Une dernière dérivation $(E''') : P^{(3)}(X) = (X-1)P^{(4)}(X) + P^{(5)}(X)$ qui en $X = 1$ fournit $P^{(3)}(1) = P^{(5)}(1) = 0$ car $\deg P = 4$. On dispose d'assez d'informations pour expliciter P .

Pour conclure donc, on a $P'(1) = P^{(3)}(1) = 0$, $P(1) = 1$, $P''(1) = 4$, $P^{(4)}(1) = 8$ et

$$P = 1 + \frac{4}{2!}(X-1)^2 + \frac{8}{4!}(X-1)^4 = 1 + 2(X-1)^2 + \frac{1}{3}(X-1)^4$$

19.4.3 Racines multiples

DÉFINITION 19.14 - Soient $P \in \mathbb{K}[X]$, $\alpha \in \mathbb{K}$ et $m \in \mathbb{N}^*$.

Le scalaire α est racine de P de **multiplicité au moins m** (resp. de **multiplicité m**) si $(X - \alpha)^m$ divise P (resp. $(X - \alpha)^m$ divise P et $(X - \alpha)^{m+1}$ ne divise pas P).

Exemple 1. 2 est racine de multiplicité exactement 3, 0 est racine de multiplicité exactement 4, et 1 est racine simple du polynôme $(X - 2)^3 (X - 1) X^4$.

Exemple 2. Déterminons les racines du polynôme $P = \sum_{k=0}^n \left[\binom{n}{k} 3^k (1 - X)^{3n-2k} X^k \right]$ ainsi que leur ordre de multiplicité.

$$\text{On a : } P = (1 - X)^n \sum_{k=0}^n \binom{n}{k} (3X)^k [(1 - X)^2]^{n-k} = (1 - X)^n [(1 - X^2) + 3X]^n$$

D'où : $P = (1 - X)^n (1 + X + X^2)^n$. Dans $\mathbb{R}$, P possède une unique racine, qui est 1, dont la multiplicité est égale à n .

Dans $\mathbb{C}[X]$, on peut écrire : $P = (1 - X)^n (X - j)^n (X - \bar{j})^n$. On en déduit que dans $\mathbb{C}$, P possède exactement trois racines, qui sont les racines cubiques de l'unité, toutes de multiplicité égale à n .

THÉORÈME 19.5 - Soient $P \in \mathbb{K}[X]$, $\alpha \in \mathbb{K}$ et $m \in \mathbb{N}$.

α est racine de P de multiplicité au moins m SSI $P(\alpha) = P'(\alpha) = \dots = P^{(m-1)}(\alpha) = 0$.

α est racine de P de multiplicité exactement m SSI $P(\alpha) = \dots = P^{(m-1)}(\alpha) = 0$ et $P^{(m)}(\alpha) \neq 0$.

PREUVE.

► Sens direct. Supposons que α est racine de P de multiplicité au moins m . Il existe alors un polynôme Q tel que : $P = (X - \alpha)^m Q$.

Soit alors k un entier quelconque de $\llbracket 0, m - 1 \rrbracket$. On a (via la formule de Leibniz) :

$$P^{(k)} = \sum_{i=0}^k \binom{k}{i} [(X - \alpha)^m]^{(i)} Q^{(k-i)} = \sum_{i=0}^k \binom{k}{i} \frac{m!}{(m-i)!} (X - \alpha)^{m-i} Q^{(k-i)}$$

$$\text{En évaluant en } \alpha \text{ cette relation, on obtient : } P^{(k)}(\alpha) = \sum_{i=0}^k \binom{k}{i} \frac{m!}{(m-i)!} \underbrace{0^{m-i}}_{=0 \text{ car } i < m} Q^{(k-i)}(\alpha)$$

D'où $P^{(k)}(\alpha) = 0$.

On a donc établi l'implication :

$$[\alpha \text{ est racine de } P \text{ de multiplicité au moins } m] \implies [\forall k \in \llbracket 0, m - 1 \rrbracket, P^{(k)}(\alpha) = 0].$$

En outre : P est racine de multiplicité exactement m si et seulement si $(X - \alpha)^{m+1}$ ne divise pas P . En vertu des calculs précédent, il revient au même de dire que $(X - \alpha)$ ne divise pas le polynôme Q , càd que $Q(\alpha) \neq 0$.

$$\text{Or : } P^{(m)}(\alpha) = \sum_{i=0}^m \binom{m}{i} \frac{m!}{(m-i)!} 0^{m-i} Q^{(m-i)}(\alpha) = m! Q(\alpha).$$

On a donc : $Q(\alpha) \neq 0 \iff P^{(m)}(\alpha) \neq 0$.

En résumé :

$$[\alpha \text{ est racine de } P \text{ de multiplicité exactement } m] \implies [\forall k \in \llbracket 0, m-1 \rrbracket, P^{(k)}(\alpha) = 0 \wedge P^{(m)}(\alpha) \neq 0].$$

► Réciproque. Supposons que : $\forall k \in \llbracket 0, m-1 \rrbracket, P^{(k)}(\alpha) = 0$.

D'après la formule de Taylor, on peut écrire : $P = \underbrace{\sum_{k=0}^{m-1} \frac{P^{(k)}(\alpha)}{k!} (X-\alpha)^k}_{\text{nulle par hypothèse}} + \sum_{k=m}^{\deg(P)} \frac{P^{(k)}(\alpha)}{k!} (X-\alpha)^k$.

$$\text{Par suite : } P = \sum_{k=m}^{\deg(P)} \frac{P^{(k)}(\alpha)}{k!} (X-\alpha)^k \text{ c\`ad : } P = (X-\alpha)^m \underbrace{\sum_{k=m}^{\deg(P)} \frac{P^{(k)}(\alpha)}{k!} (X-\alpha)^{\overbrace{k-m}^{\geq 0}}}_Q.$$

Il s'ensuit que $(X-\alpha)^m | P$, ce qui signifie que α est racine de multiplicité au moins égale à m de P .

Conclusion. $[\alpha \text{ est racine de } P \text{ de multiplicité au moins } m] \iff [\forall k \in \llbracket 0, m-1 \rrbracket, P^{(k)}(\alpha) = 0]$

En outre, si $P^{(m)}(\alpha) \neq 0$, alors le reste dans la division euclidienne de Q par $(X-\alpha)$ est non nul ; ce qui signifie que $(X-\alpha)$ ne divise pas Q ; ce qui signifie que $(X-\alpha)^{m+1}$ ne divise pas le polynôme P . Ainsi, α est racine de multiplicité exactement m du polynôme P .

Conclusion. $[\alpha \text{ est racine de } P \text{ de multiplicité exactement } m]$

$$\iff [\forall k \in \llbracket 0, n-1 \rrbracket, P^{(k)}(\alpha) = 0 \wedge P^{(m)}(\alpha) \neq 0] \quad \text{☕}$$

Application. L'intérêt de connaître la multiplicité d'une racine est de fournir une factorisation éventuellement peu évidente à établir par d'autres moyens. Pour illustrer cette affirmation, on peut montrer que le polynôme $P = X^{2n} - nX^{n+1} + nX^{n-1} - 1$ (avec $n \geq 3$) est divisible par $(X-1)^3$ (et pas par $(X-1)^4$), comme 1 est racine de P de multiplicité exactement 3 (exercice⁴).

En résumé, le théorème précédent fournit une méthode pratique pour calculer la multiplicité d'une racine :

CALCUL PRATIQUE DE LA MULTIPLICITÉ D'UNE RACINE

Soient $P \in \mathbb{K}[X]$ et α un scalaire.

- 1/ On vérifie que α est racine de P (ie $P(\alpha) = 0$).
- 2/ On calcule $P'(\alpha), P''(\alpha), \dots, P^{(p)}(\alpha)$ jusqu'à obtenir une dérivée non nulle.
- 3/ Le premier entier p pour lequel $P^{(p)}(\alpha) \neq 0$ est la multiplicité de α .

4. En vertu du théorème précédent, il suffit de vérifier que $P(1) = P'(1) = P''(1) = 0$ et $P^{(3)}(1) \neq 0$.

19.4.4 Polynômes irréductibles

Les polynômes irréductibles (unitaires) sont aux polynômes ce que les nombres premiers sont aux nombres entiers.

Dans ce paragraphe, lorsque l'on parlera de polynôme, on sous-entendra élément de $\mathbb{K}[X]$ (avec $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$ comme d'habitude).

DÉFINITION 19.15 - Un polynôme P non-constant (ie de degré ≥ 1) est **irréductible** si les seuls diviseurs de P sont les polynômes associés à P et à $\tilde{1}$.

Exemples. Le polynôme $X - \alpha$ est irréductible dans $\mathbb{K}[X]$, pour tout scalaire α . Le polynôme $(X - 1)^3$ n'est pas irréductible, tout comme le polynôme $X^4 + 4X^3 + 6X^2 + 4X + 1$. Le polynôme $P = X^2 + 1$ est irréductible dans $\mathbb{R}[X]$, mais pas dans $\mathbb{C}[X]$.

L'énoncé ci-dessous est la traduction dans l'anneau des polynômes du théorème de décomposition en facteurs premiers dans l'anneau des entiers.

THÉORÈME 19.6 - (Décomposition en irréductibles dans $\mathbb{K}[X]$).

Tout polynôme non nul de $\mathbb{K}[X]$ peut s'écrire comme un produit de polynômes irréductibles, de manière unique à l'ordre des facteurs et à une constante multiplicative près.

Explicitement, pour tout polynôme $P \in \mathbb{K}[X] \setminus \{\tilde{0}\}$, il existe des polynômes $P_1, \dots, P_n$ irréductibles unitaires et deux à deux distincts, et des entiers naturels non nuls $\alpha_1, \dots, \alpha_n$ tels que :

$$P = \text{cd}(P) \prod_{i=1}^n P_i^{\alpha_i}$$

Remarque. Dans le cas très particulier où P est un polynôme constant (non nul), le produit de l'énoncé ci-dessus ne contient aucun terme.

PREUVE. Prouvons l'énoncé par récurrence sur le degré de P . On pose $A(n)$ l'assertion "tout polynôme de degré n s'écrit comme un produit de polynômes irréductibles".

L'initialisation (pour $n = 0$) est immédiate.

Passons à l'hérédité, et supposons que $A(n)$ est vraie pour tout entier naturel inférieur ou égal à n .

Soit P un polynôme de degré $n + 1$. Si P est irréductible, alors on peut écrire : $P = \text{cd}(P) \times \left(\frac{1}{\text{cd}(P)} P\right)$.

L'assertion est alors vérifiée, puisque le polynôme $\left(\frac{1}{\text{cd}(P)} P\right)$ est clairement unitaire et irréductible.

Si P n'est pas irréductible. Alors P admet un diviseur Q qui est non constant, et qui n'est pas associé à P . On en déduit qu'il existe un autre polynôme D tel que : $P = QD$. Ce second polynôme D est à son tour non constant : si tel était le cas, Q serait associé à P , ce qui contredirait l'hypothèse faite sur Q .

On en déduit que : $\deg(Q)$ et $\deg(D)$ sont au plus égaux à n (puisque leur somme vaut $n + 1$, et que tous deux sont des entiers naturels non nuls).

Par hypothèse de récurrence, Q et D s'écrivent comme produits de polynômes irréductibles, et il existe donc une grande famille de polynômes unitaires et irréductibles et une tout aussi grande famille de scalaires tels que :

$$Q = \text{cd}(Q) \prod_{i=1}^n P_i^{\alpha_i} \quad \text{et} \quad D = \text{cd}(D) \prod_{i=1}^m R_i^{\beta_i}$$

$$\text{Il s'ensuit que : } P = \underbrace{\text{cd}(P)}_{=\text{cd}(Q)\text{cd}(D)} \times \prod_{i=1}^n P_i^{\alpha_i} \times \prod_{i=1}^m R_i^{\beta_i}.$$

Ainsi P s'écrit, à son coefficient dominant près, comme un produit de polynômes irréductibles et unitaires.

Ce qui prouve que $P(n+1)$ est vraie, fournit l'hérédité de la propriété, et achève donc cette preuve. ☕

Ce résultat établi, passons à la description des polynômes irréductibles dans $\mathbb{R}[X]$ et dans $\mathbb{C}[X]$; on commence par un résultat général.

PROPRIÉTÉ 19.15 - Tout polynôme de $\mathbb{K}[X]$ de degré 1 est irréductible (pour tout corps $\mathbb{K}$).

PREUVE. Soit $P \in \mathbb{K}[X]$, avec $\deg(P) = 1$, et soit D un diviseur de P . Il existe alors $Q \in \mathbb{K}[X]$ tel que $DQ = P$. En comparant les degrés des deux termes de cette égalité, on obtient $\deg(D) = 0$ ou $\deg(D) = 1$.

Si $\deg(D) = 0$, alors D est un polynôme constant non nul : $\exists \lambda \in \mathbb{K}^*$, $D = \lambda$ (et D est donc associé à 1).

Et lorsque $\deg(D) = 1$, on a $\deg(Q) = 0$, donc $\exists \lambda \in \mathbb{K}^*$, $Q = \lambda$, d'où $D = (1/\lambda)P$ (et D est donc associé à P).

On peut conclure que les diviseurs associés à P sont les polynômes associés à 1, et ceux associés à P . Donc P est irréductible dans $\mathbb{K}[X]$. ☕

Remarque. La réciproque de la propriété ci-dessus est FAUSSE. Un polynôme irréductible n'est pas nécessairement de degré 1. L'exemple-clef est celui du polynôme $X^2 + 1$, qui est irréductible dans $\mathbb{R}[X]$.

Néanmoins, la réciproque tient dans $\mathbb{C}[X]$. Plus précisément :

THÉORÈME 19.7 - (Irréductibles de $\mathbb{C}[X]$). Dans $\mathbb{C}[X]$, les polynômes irréductibles sont exactement les polynômes de degré 1.

THÉORÈME 19.8 - (Irréductibles de $\mathbb{R}[X]$). Dans $\mathbb{R}[X]$, les polynômes irréductibles sont les polynômes de degré 1 et les polynômes de degré 2 de discriminant strictement négatif.

Remarque. Le premier de ces deux énoncés est une conséquence du **théorème de D'Alembert - Gauss** qui affirme que tout polynôme non-constant de $\mathbb{C}[X]$ admet au moins une racine complexe. Le second est une conséquence du premier et du lemme bien pratique ci-dessous.



LEMME 19.2 - Soit $P \in \mathbb{R}[X]$. Si $\alpha \in \mathbb{C}$ est racine de P , alors $\bar{\alpha}$ est racine de P .

Il est temps à présent de "régler nos dettes".

PREUVE. (DU THÉORÈME 19.7). Montrons que tout polynôme irréductible à coefficients complexes est de degré 1.

Soit $P \in \mathbb{C}[X]$ un polynôme irréductible. Par définition, P est non constant, et donc $\deg(P) \geq 1$.

Raisonnons par l'absurde et supposons que $\deg(P) \geq 2$. D'après le théorème de d'Alembert-Gauss, P admet (au moins) une racine dans $\mathbb{C}$; notons-la z_0 . Alors $Q = (X - z_0)$ est un diviseur de P , qui n'est associé ni à 1 (Q étant non constant), ni à P (puisque $\deg(P) \neq 1$) : contradiction. Il s'ensuit que $\deg(P) = 1$.

Subséquentement, tout polynôme irréductible de $\mathbb{C}[X]$ est de degré 1.

D'après la propriété 19.15, il s'ensuit que les polynômes irréductibles de $\mathbb{C}[X]$ sont exactement ceux de degré 1. ☕

► **CONSÉQUENCE DU THÉORÈME 19.7.**

PROPRIÉTÉ 19.16 - (Décomposition en irréductibles dans $\mathbb{C}[X]$).

Pour tout polynôme $P \in \mathbb{C}[X] \setminus \{0\}$, il existe des complexes $z_1, \dots, z_n$ deux à deux distincts, et des entiers naturels non nuls $\alpha_1, \dots, \alpha_n$ tels que :

$$P = \text{cd}(P) \prod_{i=1}^n (X - z_i)^{\alpha_i}$$

PREUVE. Conséquence immédiate des théorèmes 19.6 et 19.7. ☕

► **APPLICATION DU THÉORÈME 19.7.**

Dans $\mathbb{C}[X]$, la décomposition en irréductibles du polynôme $X^6 - 1$ est $\prod_{\omega \in \mathbb{U}_6} (X - \omega)$, càd plus explicitement :

$$X^6 - 1 = (X - 1)(X + j)(X - j)(X + 1)(X - \bar{j})(X + \bar{j})$$

On rappelle à toute fin utile que les racines sixièmes de l'unité sont les complexes $e^{2ik\pi/6} = e^{ik\pi/3}$, avec k variant de 0 à 5. Ces six nombres complexes sont, par ordre d'apparition, 1, $-\bar{j}$, j , -1 , $\bar{j}$ et $-j$.

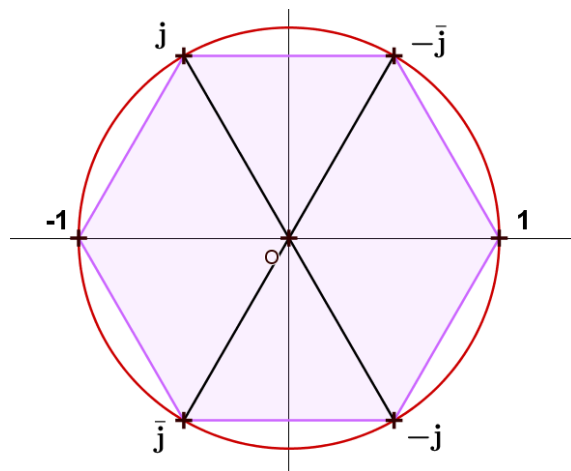


FIGURE — LES ÉLÉMENTS DE $\mathbb{U}_6$.

PREUVE. (DU THÉORÈME 19.8). Comme dans la preuve précédente, on montre dans un premier temps que les polynômes de degré 1 et les polynômes de degré 2 de discriminant strictement négatif sont irréductibles dans $\mathbb{R}[X]$.

Pour commencer, il est connu que tout polynôme de degré 1 est irréductible dans $\mathbb{K}[X]$ (quelque soit le corps $\mathbb{K}$) : c'est la propriété 19.15.

Considérons à présent un polynôme P de $\mathbb{R}[X]$ de degré 2 ayant un discriminant strictement négatif. Soit D un diviseur de P dans $\mathbb{R}[X]$. Il existe alors $Q \in \mathbb{R}[X]$ tel que $DQ = P$. En comparant les degrés des deux termes de cette égalité, on obtient $\deg(D) = 0, 1$ ou 2 .

Si $\deg(D) = 0$, alors D est un polynôme constant non nul : $\exists \lambda \in \mathbb{R}^*, D = \lambda$ (et D est donc associé à 1).

Lorsque $\deg(D) = 2$, on a $\deg(Q) = 0$, donc $\exists \lambda \in \mathbb{R}^*, Q = \lambda$, d'où $D = (1/\lambda)P$ (et D est donc associé à P).

Il reste à voir que le cas $\deg(D) = 1$ ne peut se produire ; si tel était le cas, il existerait deux réels a et b (avec $a \neq 0$) tels que $(aX + b)$ divise P . Dans ce cas, P posséderait une racine réelle $(-b/a)$, ce qui contredirait l'hypothèse suivant laquelle le discriminant de P est strictement négatif.

On peut conclure que les diviseurs associés à P sont les polynômes associés à 1, et ceux associés à P . Donc P est irréductible dans $\mathbb{R}[X]$.

Conclusion intermédiaire. Dans $\mathbb{R}[X]$, les polynômes de degré 1 et ceux de degré 2 ayant un discriminant strictement négatif sont irréductibles.

► Réciproquement à présent, montrons que tout polynôme irréductible dans $\mathbb{R}[X]$ est de l'une des deux formes précédentes. Soit $P \in \mathbb{R}[X]$ un polynôme irréductible. Par définition $\deg(P) \geq 1$.

Supposons que P soit de degré 2, avec un discriminant positif ou nul. Alors P admet (au moins) une racine dans $\mathbb{R}$; notons-la z_0 . Comme $z_0 \in \mathbb{R}$, alors $Q = (X - z_0)$ est un diviseur de P dans $\mathbb{R}[X]$, qui n'est associé ni à 1 (Q étant non constant), ni à P (puisque $\deg(P) \neq 1$) : contradiction.

Il s'ensuit que si P est de degré 2, alors son discriminant est strictement négatif.

Envisageons enfin le cas où $\deg(P) \geq 3$. D'après le théorème de d'Alembert-Gauss⁵, P admet (au moins) une racine dans $\mathbb{C}$; notons-la z_0 .

☞ Si $z_0 \in \mathbb{R}$: alors $Q = (X - z_0)$ est un diviseur de P dans $\mathbb{R}[X]$, qui n'est associé ni à 1 (Q étant non constant), ni à P (puisque $\deg(P) \neq 1$) : contradiction.

☞ Si $z_0 \in \mathbb{C} \setminus \mathbb{R}$, alors (d'après le lemme 19.2), $\overline{z_0}$ est également racine de P . Les complexes z_0 et $\overline{z_0}$ étant distincts, le polynôme $Q = (X - z_0)(X - \overline{z_0})$, (càd : $Q = X^2 - 2\operatorname{Re}(z_0)X + |z_0|^2$) divise P dans $\mathbb{C}[X]$ (a priori).

Il existe donc un polynôme $R \in \mathbb{C}[X]$ tel que : $P = QR$. On a alors : $\overline{P} = \overline{Q}\overline{R}$, et puisque P et Q sont à coefficients réels : $P = Q\overline{R}$. On en déduit que $R = \overline{R}$, d'où $R \in \mathbb{R}[X]$. Ainsi Q divise P dans $\mathbb{R}[X]$; or Q étant de degré 2, il n'est associé ni à 1, ni à P : donc P n'est pas irréductible, contradiction.

Il s'ensuit que si P est irréductible dans $\mathbb{R}[X]$, il ne peut être de degré supérieur ou égal à 3. À la lumière des résultats précédents, on peut donc conclure que dans $\mathbb{R}[X]$, les polynômes irréductibles sont les polynômes de degré 1 et les polynômes de degré 2 de discriminant strictement négatif. ☕

5. Qu'il est légitime d'appliquer au polynôme P , que l'on peut voir comme élément de $\mathbb{C}[X]$, essentiellement car $\mathbb{R} \subset \mathbb{C}$.

► **CONSÉQUENCE DU THÉORÈME 19.8.**

PROPRIÉTÉ 19.17 - (Décomposition en irréductibles dans $\mathbb{R}[X]$).

Pour tout polynôme $P \in \mathbb{R}[X] \setminus \{0\}$, il existe des réels $x_1, \dots, x_m$ deux à deux distincts, des couples de réels $(a_1, b_1), \dots, (a_p, b_p)$ deux à deux distincts et tels que $a_i^2 - 4b_i < 0$, et des entiers naturels non nuls $\alpha_1, \dots, \alpha_m, \beta_1, \dots, \beta_p$ tels que :

$$P = \text{cd}(P) \prod_{i=1}^m (X - x_i)^{\alpha_i} \prod_{i=1}^p (X^2 + a_i X + b_i)^{\beta_i}$$

PREUVE. Conséquence immédiate des théorèmes 19.6 et 19.8. ☕

► **APPLICATION DU THÉORÈME 19.8.**

On a vu un peu plus haut que la décomposition en irréductibles dans $\mathbb{C}[X]$ du polynôme $X^6 - 1$ est :

$$X^6 - 1 = (X - 1)(X + j)(X - j)(X + 1)(X - \bar{j})(X + \bar{j})$$

Pour obtenir la décomposition en irréductibles dans $\mathbb{R}[X]$ du polynôme $X^6 - 1$, on “regroupe” les termes faisant intervenir des racines conjuguées, de manière à faire apparaître des polynômes à coefficients réels, de degré 2 et de discriminant strictement négatif. A cet effet, il peut être utile de connaître la formule :

$$\forall z \in \mathbb{C}, \quad (X - z)(X - \bar{z}) = X^2 - 2\text{Re}(z)X + |z|^2$$

Dans le cas présent, on obtient :

$$X^6 - 1 = (X - 1)(X + 1) \underbrace{[(X - j)(X - \bar{j})]}_{X^2 + X + 1} \underbrace{[(X + j)(X + \bar{j})]}_{X^2 - X + 1}$$

On en déduit que la décomposition en irréductibles dans $\mathbb{R}[X]$ de $X^6 - 1$ est :

$$X^6 - 1 = (X - 1)(X + 1)(X^2 + X + 1)(X^2 - X + 1)$$

Finalement, on retiendra la :

Méthode pratique : pour décomposer un polynôme $P \in \mathbb{R}[X]$ en irréductibles dans $\mathbb{C}[X]$ (*resp.* dans $\mathbb{R}[X]$), on peut commencer par déterminer les racines complexes de P (*resp.* puis on regroupe deux à deux les termes faisant intervenir des racines conjuguées).

PREUVE. (DU LEMME 19.2, PAGE 488). Soit $P \in \mathbb{R}[X]$, et soit $\alpha \in \mathbb{C}$ une racine de P . Supposons P non nul (sinon il n'y a rien à prouver), et notons $n = \deg(P)$. Il existe $(n+1)$ réels $a_0, \dots, a_n$ tels que :

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

Puisque α est racine de P , on a : $\sum_{k=0}^n a_k \alpha^k = 0$. En prenant les conjugués des deux termes de cette égalité, on obtient :

$$\overline{\sum_{k=0}^n a_k \alpha^k} = 0 \iff \sum_{k=0}^n \overline{a_k} \overline{\alpha^k} = 0 \iff \sum_{k=0}^n a_k \overline{\alpha}^k = 0 \iff P(\overline{\alpha}) = 0$$

La première équivalence provenant de la compatibilité de la conjugaison avec l'addition, la multiplication, et l'élevation à une puissance entière ; la seconde de l'hypothèse suivant laquelle les a_k sont réels (et donc invariants par conjugaison) ; la dernière de la définition du polynôme P .

Conclusion. Soit $P \in \mathbb{R}[X]$. Si $\alpha \in \mathbb{C}$ est racine de P , alors $\bar{\alpha}$ est racine de P . ☕

► APPLICATIONS DU LEMME 19.2.

La première application du lemme est évidemment le théorème 19.8, dans la preuve duquel il joue un rôle crucial.

Une seconde application en est la factorisation d'un polynôme à coefficients réels dont on connaît une racine complexe.

Illustration : factorisation du polynôme $X^4 - X^3 + 2X^2 - 2X + 4$, sachant que $1 + i$ est racine.

► On commence par vérifier que $1 + i$ est racine de P : on a $1 + i = \sqrt{2}e^{i\pi/4}$, d'où :

$$P(1 + i) = 4e^{i\pi} - 2\sqrt{2}e^{3i\pi/4} + 4e^{i\pi/2} - 2\sqrt{2}e^{i\pi/4} + 4 = -4 + 2 - 2i + 4i - 2 - 2i + 4 = 0$$

Ainsi $1 + i$ est racine de P , et en vertu du lemme 19.2, $1 - i$ est également racine de P . Il s'ensuit que P est multiple de $(X - 1 - i)(X - 1 + i)$, c'est-à-dire de $(X^2 - 2X + 2)$.

Il existe donc un polynôme $Q \in \mathbb{R}[X]$ tel que $P = (X^2 - 2X + 2)Q$, que l'on peut déterminer en effectuant la division euclidienne de P par $(X^2 - 2X + 2)$.

$\begin{array}{r} X^4 - X^3 + 2X^2 - 2X + 4 \\ - (X^4 - 2X^3 + 2X^2) \\ \hline X^3 - 2X + 4 \\ - (X^3 - 2X^2 + 2X) \\ \hline 2X^2 - 4X + 4 \\ - (2X^2 - 4X + 4) \\ \hline 0 \end{array}$	$\begin{array}{r} X^2 - 2X + 2 \\ \hline X^2 + X + 2 \end{array}$

Par suite : $X^4 - X^3 + 2X^2 - 2X + 4 = (X^2 - 2X + 2)(X^2 + X + 2)$

► Le polynôme $X^2 + X + 2$ possède deux racines complexes conjuguées : $(-1 \pm i\sqrt{7})/2$.

Conclusion. Le polynôme possède $X^4 - X^3 + 2X^2 - 2X + 4$ quatre racines complexes : $1 \pm i$ et $\frac{-1 \pm i\sqrt{7}}{2}$.

On peut d'ailleurs déduire des calculs précédents que la décomposition en irréductibles dans $\mathbb{C}[X]$ du polynôme P est :

$$X^4 - X^3 + 2X^2 - 2X + 4 = (X - 1 - i)(X - 1 + i) \left(X - \frac{-1 + i\sqrt{7}}{2} \right) \left(X - \frac{-1 - i\sqrt{7}}{2} \right)$$

Et sa décomposition en irréductibles dans $\mathbb{R}[X]$ est :

$$X^4 - X^3 + 2X^2 - 2X + 4 = (X^2 - 2X + 2)(X^2 + X + 2)$$

19.4.5 Polynômes scindés, relations coefficients-racines

DÉFINITION 19.16 - On appelle polynôme **scindé** dans $\mathbb{K}$ tout polynôme de $\mathbb{K}[X]$ dont toutes les racines sont dans $\mathbb{K}$.

Il revient au même de dire que P est scindé ou qu'il peut s'écrire comme un produit de polynômes de degré 1 à coefficients dans $\mathbb{K}$.

Exemples. Le polynôme $X^2 + 1$ est scindé dans $\mathbb{C}$, pas dans $\mathbb{R}$.

PROPRIÉTÉ 19.18 - Tout polynôme de $\mathbb{C}[X]$ est scindé.

PREUVE. Conséquence immédiate du théorème de D'Alembert-Gauss. ☕

Si P est de degré 2 et est scindé sur $\mathbb{K}$, P peut s'écrire $P = aX^2 + bX + c = a(X - \alpha_1)(X - \alpha_2)$ où α_1, α_2 sont les racines de P . En développant le produit on a $aX^2 + bX + c = aX^2 - a(\alpha_1 + \alpha_2)X + a\alpha_1\alpha_2$ d'où

$$\alpha_1 + \alpha_2 = -\frac{b}{a} \text{ et } \alpha_1\alpha_2 = \frac{c}{a}$$

Si P est de degré 3, $P = aX^3 + bX^2 + cX + d = a(X - \alpha_1)(X - \alpha_2)(X - \alpha_3)$, en développant on a :

$$\alpha_1 + \alpha_2 + \alpha_3 = -\frac{b}{a}; \quad \alpha_1\alpha_2 + \alpha_2\alpha_3 + \alpha_3\alpha_1 = \frac{c}{a} \quad \alpha_1\alpha_2\alpha_3 = -\frac{d}{a}$$

On généralise facilement à un polynôme de degré n , $P = a_nX^n + \dots + a_0 = a_n \prod_{i=1}^n (X - \alpha_i)$ pour obtenir les **relations coefficients-racines** :

$$\begin{aligned} \sigma_1 = \alpha_1 + \dots + \alpha_n &= -\frac{a_{n-1}}{a_n} \\ \sigma_2 = \alpha_1\alpha_2 + \alpha_2\alpha_3 + \dots + \alpha_n\alpha_1 &= \frac{a_{n-2}}{a_n} \\ &\vdots \\ \sigma_n = \alpha_1\dots\alpha_n &= (-1)^n \frac{a_0}{a_n} = (-1)^n \frac{P(0)}{a_n} \end{aligned}$$

On observera l'alternance des signes. Seule les formules donnant la somme et le produit des racines sont à connaître.

On peut aussi observer que le cas $n = 2$, on retrouve le résultat classique : a et b sont racines de $X^2 - SX + P$ où $S = a + b$ et $P = ab$.

Exemple d'application. On veut résoudre l'équation $36x^3 - 12x^2 - 5x + 1 = 0$ sachant que l'une des racines est égale à la somme des deux autres.

Notons α, β, γ les trois racines et supposons par exemple que $\gamma = \alpha + \beta$. D'après les relations coefficients-racines on a : $\alpha + \beta + \gamma = -\frac{-12}{36} = 2\gamma \Rightarrow \alpha + \beta = \gamma = \frac{1}{6}$. Par ailleurs : $\alpha\beta\gamma = -\frac{1}{36}$. On en déduit que α et β sont racines de $X^2 - \left(\frac{1}{6}\right)X + \left(-\frac{1}{6}\right)$ dont les racines sont $\frac{1}{2}$ et $-\frac{1}{3}$.

19.5 Epilogue

19.5.1 Exemple d'équation polynomiale

Exercice (utilisation du degré) : résoudre dans $\mathbb{K}[X]$ l'équation $X^2P(X) = P(X^2)$

► Cas particulier du polynôme nul. Clairement, le polynôme $\tilde{0}$ est solution de l'équation (♠).

► Supposons à présent que $P \neq \tilde{0}$ est solution de l'équation. Puisque P est non nul, il a un degré entier naturel ; notons $n = \deg(P)$.

En comparant les degrés des termes de gauche et de droite de l'équation, on obtient : $n + 2 = 2n$, ce qui entraîne $n = 2$. Il existe donc trois scalaires a, b et c , avec $a \neq 0$ tels que $P = aX^2 + bX + c$. On en déduit que P est solution de l'équation si :

$$aX^4 + bX^3 + cX^2 = aX^4 + bX^2 + c$$

Et puisque deux polynômes sont égaux si et seulement si tous leurs coefficients sont deux à deux égaux, on obtient $b = 0$ et $c = 0$. Par suite : $P = aX^2$.

Réciproquement, tout polynôme P s'écrivant aX^2 (avec $a \neq 0$) est solution de l'équation. On peut donc conclure que les polynômes non nuls solutions de l'équation sont tous les aX^2 (avec $a \neq 0$) (♣).

Conclusion. D'après (♠) et (♣), les solutions de $X^2P(X) = P(X^2)$ sont les polynômes aX^2 , avec $a \in \mathbb{K}$ arbitraire.

19.5.2 Exemple de décomposition en irréductibles, dans $\mathbb{C}[X]$ et dans $\mathbb{R}[X]$

Exercice : décomposition en irréductibles de $P = X^4 + 1$ dans $\mathbb{C}[X]$, puis dans $\mathbb{R}[X]$.

► 1ère étape — Racines de P dans $\mathbb{C}$: on résout dans $\mathbb{C}$ l'équation $z^4 + 1 = 0$. Ceci revient à résoudre $z^4 = -1$, c'est-à-dire à déterminer les racines quatrièmes de -1 . En écrivant $-1 = e^{i\pi}$, on obtient déjà la première de ces racines $e^{i\pi/4}$.

Il suffit ensuite de se souvenir que “l’on passe d’une racine n -ième à la suivante en tournant de $2\pi/n$, c-à-d en multipliant par $e^{2i\pi/n}$ ”.

Dans la présente situation (puisque $n = 4$), les racines quatrièmes de -1 sont donc :

$$e^{i\pi/4}, e^{i\pi/4} \times e^{i\pi/2}, e^{i\pi/4} \times (e^{i\pi/2})^2 \text{ et } e^{i\pi/4} \times (e^{i\pi/2})^3$$

Il serait regrettable de ne pas remarquer que $e^{i\pi/2} = i$, ce qui permet de réécrire les solutions de $z^4 + 1 = 0$:

$$e^{i\pi/4}, ie^{i\pi/4}, -e^{i\pi/4}, \text{ et } -ie^{i\pi/4}$$

Rappelons aussi que les racines quatrièmes de -1 sont les affixes des sommets d’un carré inscrit dans le cercle de centre O et de rayon 1 .⁶

► **2ème étape — Décomposition en irréductibles de P dans $\mathbb{C}[X]$** : peu de travail ici, puisqu’il suffit de factoriser P à l’aide des racines que l’on vient d’obtenir.

Conclusion : la décomposition en irréductibles de P dans $\mathbb{C}[X]$ est

$$X^4 + 1 = (X - e^{i\pi/4})(X - ie^{i\pi/4})(X + e^{i\pi/4})(X + ie^{i\pi/4}) \quad \text{soit encore :} \quad X^4 + 1 = \prod_{k=0}^3 (X - i^k e^{i\pi/4})$$

► **3ème étape — Décomposition en irréductibles de P dans $\mathbb{R}[X]$** : on multiplie deux à deux les termes de la décomposition dans $\mathbb{C}[X]$ faisant intervenir des racines conjuguées.

On rappelle que pour tout $z \in \mathbb{C}$: $(X - z)(X - \bar{z}) = X^2 - 2\operatorname{Re}(z)X + |z|^2$.⁷

On obtient ainsi : $(X - e^{i\pi/4})(X + ie^{i\pi/4}) = X^2 - \sqrt{2}X + 1$ et $(X - ie^{i\pi/4})(X + e^{i\pi/4}) = X^2 + \sqrt{2}X + 1$.

Conclusion. La décomposition en irréductibles de P dans $\mathbb{R}[X]$ est :

$$X^4 + 1 = (X^2 - \sqrt{2}X + 1)(X^2 + \sqrt{2}X + 1).$$

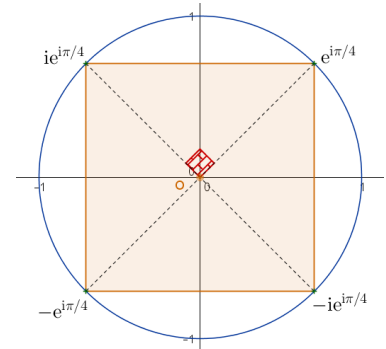
19.5.3 Division euclidienne dans $\mathbb{K}[X]$, et puissances d’une matrice

Exercice : soit $A \in M_n(\mathbb{K})$, telle que $A^2 - 4A + 3I_n = 0$. Calculer A^N pour tout $N \in \mathbb{N}$.

Observons que l’équation vérifiée par A se réécrit $P(A) = 0$, où $P = X^2 - 4X + 3$; et réalisons la division euclidienne de X^N par P (avec N entier naturel arbitraire). D’après le théorème d’Eisenstein, il existe un unique couple (Q, R) de polynômes de $\mathbb{K}[X]$ tel que : $X^N = PQ + R$ avec R de degré au plus 1.

6. Plus généralement, les racines n -ièmes d’un complexe non nul z sont les affixes des sommets d’un polygone régulier à n côtés inscrit dans le cercle de centre O et de rayon $|z|^{1/n}$.

7. Au regard de son importance pratique dans ce chapitre, cette formule n’est pas à connaître par cœur, mais doit pouvoir être retrouvée rapidement.



Autrement dit, il existe deux scalaires a et b tels que : $X^N = PQ + aX + b$ ($\spadesuit$).

Observons encore que P a comme racines 1 et 3 ; l'évaluation de la relation ($\spadesuit$) en 1 et en 3 fournit les deux équations :

$$1 = a + b \quad \text{et} \quad 3^N = 3a + b$$

On en déduit que : $a = (3^N - 1)/2$ et $b = (3 - 3^N)/2$. Par conséquent :

$$X^N = (X^2 - 4X + 3)Q + \frac{3^N - 1}{2}X + \frac{3 - 3^N}{2}$$

Revenant à notre interprétation matricielle, on obtient :

$$A^N = (A^2 - 4A + 3I_n)Q(A) + \frac{3^N - 1}{2}A + \frac{3 - 3^N}{2}I_n$$

Puisque $A^2 - 4A + 3I_n = 0$ par hypothèse, on en déduit finalement que :

$$\boxed{\forall N \in \mathbb{N}, A^N = \frac{3^N - 1}{2}A + \frac{3 - 3^N}{2}I_n}$$

19.5.4 Utilisation des multiplicités pour une factorisation

Exercice : vérifier que i est racine double de $P = X^6 + X^5 + 3X^4 + 2X^3 + 3X^2 + X + 1$.

En déduire la décomposition en irréductibles de P dans $\mathbb{C}[X]$ et dans $\mathbb{R}[X]$.

On a : $P(i) = i^6 + i^5 + 3i^4 + 2i^3 + 3i^2 + i + 1 = -1 + i + 3 - 2i - 3 + i + 1 = 0$ (ouf!).

Par ailleurs : $P' = 6X^5 + 5X^4 + 12X^3 + 6X^2 + 6X + 1$

d'où $P'(i) = 6i^5 + 5i^4 + 12i^3 + 6i^2 + 6i + 1 = 6i + 5 - 12i - 6 + 6i + 1 = 0$ (re-ouf!).

Enfin : $P'' = 30X^4 + 20X^3 + 36X^2 + 12X + 6$

d'où $P''(i) = 30X^4 + 20X^3 + 36X^2 + 12X + 6 = 30 - 20i - 36 + 12i + 6 = -8i$.

On a donc : $P(i) = P'(i) = 0$ et $P''(i) \neq 0$. On peut alors conclure que i est racine double de P , c'est à dire que i est racine de multiplicité exactement égale à 2 du polynôme P .

Il s'ensuit que P est multiple de $(X - i)^2$ (et pas de $(X - i)^3$).

Par ailleurs, P (et P' , et P'') étant à coefficients réels, on déduit de ce qui précède que $(-i)$ est également racine double de P .

Ainsi, P est divisible par $(X - i)^2(X + i)^2$, c'est à dire par $(X^2 + 1)^2$.

Il existe donc un polynôme Q de degré 2 tel que : $P = (X^2 + 1)^2 Q$.

On peut déterminer Q par division euclidienne ou par “coefficients inconnus”. Quelque soit la méthode choisie, on obtient : $Q = X^2 + X + 1$.

Par suite : $P = (X^2 + 1)^2 (X^2 + X + 1)$. Les polynômes $X^2 + 1$ et $X^2 + X + 1$ étant de discriminant strictement négatif, cette écriture est la décomposition en irréductibles de P dans $\mathbb{R}[X]$.

Conclusion. La décomposition en irréductibles de P dans $\mathbb{R}[X]$ est :

$$X^6 + X^5 + 3X^4 + 2X^3 + 3X^2 + X + 1 = (X^2 + 1)^2 (X^2 + X + 1)$$

Et la décomposition en irréductibles de P dans $\mathbb{C}[X]$ est :

$$X^6 + X^5 + 3X^4 + 2X^3 + 3X^2 + X + 1 = (X - i)^2 (X + i)^2 (X - j) (X - \bar{j})$$

19.5.5 Somme de deux binomiales indépendantes

PROPRIÉTÉ 19.19 - Si X et Y sont deux VAR indépendantes suivant la même loi binomiale $\mathcal{B}(n, p)$, alors $X + Y$ suit la loi binomiale $\mathcal{B}(2n, p)$.

PREUVE. Soient X et Y deux VAR indépendantes suivant la même loi binomiale $\mathcal{B}(n, p)$. Puisque X et Y sont à valeurs dans $\llbracket 0, n \rrbracket$, la somme $X + Y$ est à valeurs dans $\llbracket 0, 2n \rrbracket$.

Soit donc k un entier naturel dans $\llbracket 0, 2n \rrbracket$. On a :

$$P(X + Y = k) = \sum_{i=0}^k P((X = i) \cap (Y = k - i)) = \sum_{i=0}^k P(X = i) \times P(Y = k - i)$$

La première égalité provient de ce que l'évènement $(X + Y = k)$ est l'union disjointe des évènements $((X = i) \cap (Y = k - i))_{i \in \llbracket 0, k \rrbracket}$; la seconde de l'indépendance de X et Y .

$$\text{D'où : } P(X + Y = k) = \sum_{i=0}^k \binom{n}{i} p^i (1-p)^{n-i} \times \binom{n}{k-i} p^{k-i} (1-p)^{n-k+i} = \sum_{i=0}^k \binom{n}{i} \binom{n}{k-i} p^k (1-p)^{2n-k}$$

$$\text{Finalement : } P(X + Y = k) = p^k (1-p)^{2n-k} \sum_{i=0}^k \binom{n}{i} \binom{n}{k-i}.$$

$$\text{D'où (formule de Vandermonde) : } P(X + Y = k) = \binom{2n}{k} p^k (1-p)^{2n-k}$$

Conclusion : $\forall k \in \llbracket 0, 2n \rrbracket$, $P(X + Y = k) = \binom{2n}{k} p^k (1-p)^{2n-k}$. Donc $X + Y$ suit la loi binomiale

$\mathcal{B}(2n, p)$. ☕

Remarque. Ce résultat peut être généralisé de la manière suivante : si $X_1, \dots, X_N$ sont N variables aléatoires mutuellement indépendantes suivant tout la même loi binomiale de taille n et de paramètre p , alors $X_1 + \dots + X_N$ suit la loi binomiale $\mathcal{B}(Nn, p)$.

19.5.6 Polynômes de Tchebychev

Etude des polynômes de Tchebychev : degré, coefficient dominant, racines

On définit une suite de polynômes $(T_n)_{n \in \mathbb{N}}$ en posant $T_0 = 1$, $T_1 = X$, et $\forall n \in \mathbb{N}$, $T_{n+2} = 2XT_{n+1} - T_n$.

Ces polynômes sont appelés *polynômes de Tchebychev de première espèce*.

- 1) Expliciter T_2 et T_3 .
- 2) Pour tout entier naturel n , déterminer le degré de T_n et son coefficient dominant.
- 3) Etablir que : $\forall n \in \mathbb{N}$, $\forall \theta \in \mathbb{R}$, $T_n(\cos(\theta)) = \cos(n\theta)$ et $T_n(\operatorname{ch}(\theta)) = \operatorname{ch}(n\theta)$.
- 4) Soit n un entier naturel. Dédurre de la question précédente les valeurs de $T_n(1)$ et de $T'_n(1)$.
- 5) Soit à présent n un entier naturel non nul. Déterminer les racines de T_n appartenant à l'intervalle $[-1, 1]$.
Combien y en a-t-il ? Comment justifier que ces racines sont simples et qu'il n'y en a pas d'autres ?
- 6) Etudier la parité⁸ du polynôme T_n en fonction de la parité de l'entier naturel n .
- 7) Soit n un entier naturel non nul, et x un nombre réel. Montrer que : $|x| \leq 1 \iff |T_n(x)| \leq 1$.

CORRIGÉ.

1) D'après l'énoncé : $T_2 = 2XT_1 - T_0 \iff \boxed{T_2 = 2X^2 - 1}$.

Puis : $T_3 = 2XT_2 - T_1 \iff \boxed{T_3 = 4X^3 - 3X}$.

2) Montrons que : $\forall n \in \mathbb{N}^*$, $\deg(T_n) = n$ et $\operatorname{cd}(T_n) = 2^{n-1}$ par récurrence double sur n . Notons :

$$\mathcal{P}(n) : \text{“} \deg(T_n) = n \text{ et } \operatorname{cd}(T_n) = 2^{n-1} \text{”}$$

► Initialisation (pour $n = 1$ et $n = 2$) : on a $T_1 = X$ et $T_2 = 2X^2 - 1$. On observe que $\deg(T_1) = 1$ et $\deg(T_2) = 2$; et $\operatorname{cd}(T_1) = 2^0$ et $\operatorname{cd}(T_2) = 2^1$. Donc les propriétés $\mathcal{P}(1)$ et $\mathcal{P}(2)$ sont vraies.

► Hérédité : supposons la propriété vraie aux rangs n et $n + 1$ pour un certain entier naturel non nul n . On exploite alors la relation $T_{n+2} = 2XT_{n+1} - T_n$.

Dans le terme de droite de cette égalité, le degré de $2XT_{n+1}$ est par hypothèse de récurrence, égal à $1 + (n + 1) = n + 2$; tandis que celui de T_n est égal à n (toujours par hypothèse de récurrence).

Puisque $\deg(2XT_{n+1}) > \deg(T_n)$, on en déduit d'une part que $\deg(2XT_{n+1} - T_n) = \deg(2XT_{n+1}) = n + 2$, et d'autre part que : $\operatorname{cd}(2XT_{n+1} - T_n) = \operatorname{cd}(2XT_{n+1}) = 2\operatorname{cd}(T_{n+1})$.

Or, par hypothèse de récurrence, $\operatorname{cd}(T_{n+1}) = 2^n$, donc : $\operatorname{cd}(2XT_{n+1}) = 2^{n+1}$.

En résumé, on a établi que : $\deg(T_{n+2}) = n + 2$ et $\operatorname{cd}(T_{n+2}) = 2^{n+1}$, ce qui signifie que la propriété $\mathcal{P}(n + 2)$ est vraie, et prouve l'hérédité.

Conclusion : $\forall n \in \mathbb{N}^*$, $\deg(T_n) = n$ et $\operatorname{cd}(T_n) = 2^{n-1}$

3) Pour faire preuve d'originalité, prouvons la propriété $P(n)$: “ $\forall \theta \in \mathbb{R}$, $T_n(\cos(\theta)) = \cos(n\theta)$ ” par récurrence sur n .

8. Par parité du polynôme, on entend parité de la fonction polynomiale associée.

► Initialisation (pour $n = 0$ et $n = 1$) : d'une part $T_0 = 1$ d'où pour tout réel θ on a : $T_0(\cos(\theta)) = 1$. D'autre part, pour tout réel θ on a : $\cos(0 \times \theta) = 1$. Donc $P(0)$ est vraie.

D'une part $T_1 = X$ d'où pour tout réel θ on a : $T_1(\cos(\theta)) = \cos(\theta)$. D'autre part, pour tout réel θ on a : $\cos(1 \times \theta) = \cos(\theta)$. Donc $P(1)$ est vraie.

► Hérédité : supposons la propriété vraie aux rangs n et $n + 1$ pour un certain entier naturel non nul n . Alors, pour tout réel θ on a :

$$\begin{aligned} T_{n+2}(\cos(\theta)) &= 2 \cos(\theta) \underbrace{T_{n+1}(\cos(\theta))}_{=_{HR} \cos((n+1)\theta)} - \underbrace{T_n(\cos(\theta))}_{=_{HR} \cos(n\theta)} = 2 \cos(\theta) \cos((n+1)\theta) - \cos(n\theta) \\ &= \cos((n+2)\theta) + \cos(n\theta) - \cos(n\theta) \text{ d'où : } T_{n+2}(\cos(\theta)) = \cos((n+2)\theta) \end{aligned}$$

Ce qui assure que la propriété $P(n+2)$ est vraie, et prouve l'hérédité de la propriété.

Conclusion : $\forall n \in \mathbb{N}, \forall \theta \in \mathbb{R}, T_n(\cos(\theta)) = \cos(n\theta)$

4) Soit n un entier naturel. D'après la question précédente : $T_n(1) = T_n(\cos(0)) = \cos(n \times 0)$

d'où : Conclusion : $\forall n \in \mathbb{N}, T_n(1) = 1$

Considérons à présent la fonction $F : \theta \in \mathbb{R} \mapsto T_n(\cos(\theta))$. Celle-ci est dérivable sur $\mathbb{R}$ en tant que composée de fonctions qui le sont, et : $\forall \theta \in \mathbb{R}, F'(\theta) = -\sin(\theta) T'_n(\cos(\theta))$.

Comme par ailleurs on a : $\forall \theta \in \mathbb{R}, T_n(\cos(\theta)) = \cos(n\theta)$, on peut affirmer que : $\forall \theta \in \mathbb{R}, F'(\theta) = -n \sin(n\theta)$.

En identifiant les deux expressions obtenues pour F' , on a : $\sin(\theta) T'_n(\cos(\theta)) = n \sin(n\theta)$.

Alors, pour tout réel θ (non multiple de π) on a : $T'_n(\cos(\theta)) = \frac{n \sin(n\theta)}{\sin(\theta)}$.

Il reste à observer que $T'_n(1) = \lim_{x \rightarrow 1} T'_n(x)$ (puisque T_n est polynomiale donc $\mathcal{C}^\infty$ donc en particulier $\mathcal{C}^1$) pour obtenir : $T'_n(1) = \lim_{\theta \rightarrow 0} \frac{n \sin(n\theta)}{\sin(\theta)}$. Comme $\sin(\theta) \underset{0}{\sim} \theta$ et $\sin(n\theta) \underset{0}{\sim} n\theta$, on en déduit que $\frac{n \sin(n\theta)}{\sin(\theta)} \underset{0}{\sim} n^2$, ce qui entraîne $\lim_{\theta \rightarrow 0} \frac{n \sin(n\theta)}{\sin(\theta)} = n^2$ et donc : $\forall n \in \mathbb{N}, T'_n(1) = n^2$

5) Soit n un entier naturel non nul. On cherche les racines de T_n appartenant à l'intervalle $[-1, 1]$. Soit donc x un réel de cet intervalle tel que : $T_n(x) = 0$. Puisque x est compris entre -1 et 1 , il existe un unique réel θ dans $[0, \pi]$ tel que $x = \cos(\theta)$. On a alors :

$$\begin{aligned} T_n(x) = 0 &\iff T_n(\cos(\theta)) = 0 \iff \cos(n\theta) = 0 \iff n\theta \equiv \frac{\pi}{2} [\pi] \iff \theta \equiv \frac{\pi}{2n} \left[\frac{\pi}{n} \right] \\ &\iff \exists k \in \mathbb{Z}, \theta = \frac{\pi}{2n} + k \frac{\pi}{n} \iff \exists k \in \mathbb{Z}, \theta = \frac{(2k+1)\pi}{2n} \end{aligned}$$

Posons alors, pour tout $k \in [0, n-1]$, $\theta_k = \frac{(2k+1)\pi}{2n}$. Il est clair que les réels θ_k sont n réels distincts de $[0, \pi]$. Puisque la restriction de la fonction $\cos$ à l'intervalle $[0, \pi]$ est injective, les réels $(\cos(\theta_k))_{k \in [0, n-1]}$ sont n réels distincts de l'intervalle $[-1, 1]$; qui plus est, ce sont (par construction) n racines distinctes du polynôme T_n . Comme T_n est de degré n (cf question 2), on peut conclure qu'il n'y en a pas d'autres.

Conclusion : pour tout entier naturel n non nul, le polynôme T_n possède n racines distinctes (donc simples puisque T_n est de degré n) qui sont les réels : $\cos\left(\frac{(2k+1)\pi}{2n}\right)$ avec $k \in \llbracket 0, n-1 \rrbracket$.

6) Au regard des premiers termes de la suite, on peut facilement conjecturer que T_{2n} est pair tandis que T_{2n+1} est impair. On prouve par récurrence sur n la propriété $P(n)$: “ T_{2n} est pair et T_{2n+1} est impair”.

► Initialisation (pour $n = 0$) : triviale puisque $T_0 = 1$ et $T_1 = X$.

► Hérédité : soit n un entier naturel tel que $P(n)$ soit vraie, c'est-à-dire : T_{2n} est pair et T_{2n+1} est impair.

Alors : $T_{2n+2} = 2XT_{2n+1} - T_{2n}$ est la différence entre $2XT_{2n+1}$ qui est pair (en tant que produit de polynômes impairs) et T_{2n} également pair. Par suite T_{2n+2} est pair.

Puis : $T_{2n+3} = 2XT_{2n+2} - T_{2n+1}$ est la différence entre $2XT_{2n+2}$ qui est impair (en tant que produit de polynômes de parités différentes) et T_{2n+1} également impair. Par suite T_{2n+3} est impair.

La propriété est donc héréditaire.

Conclusion : pour tout entier naturel n , le polynôme T_n a la même parité que n .

7) Soient n un entier naturel non nul, et x un nombre réel.

► Supposons $|x| \leq 1$: alors existe un unique réel θ dans $[0, \pi]$ tel que $x = \cos(\theta)$. Donc $T_n(x) = T_n(\cos(\theta)) = \cos(n\theta)$. En particulier $|T_n(x)| \leq 1$ puisque $|\cos(n\theta)| \leq 1$.

On a ainsi établi que : $|x| \leq 1 \implies |T_n(x)| \leq 1$ (♠)

► Si $|x| > 1$: alors il existe un réel $\theta > 0$ tel que $x = \operatorname{ch}(\theta)$. Donc $T_n(x) = T_n(\operatorname{ch}(\theta)) = \operatorname{ch}(n\theta)$. En particulier $|T_n(x)| > 1$ puisque $|\operatorname{ch}(n\theta)| > 1$ dès lorsque n et θ sont non nuls.

On a ainsi établi que : $|x| > 1 \implies |T_n(x)| > 1$ (♣)

Conclusion : d'après (♠) et (♣), pour tout entier naturel n non nul : $|x| \leq 1 \iff |T_n(x)| \leq 1$.

19.5.7 Théorème de d'Alembert-Gauss

THÉORÈME 19.9 - (D'Alembert-Gauss).

Tout polynôme non constant de $\mathbb{C}[X]$ possède une racine dans $\mathbb{C}$.

PREUVE. NB : cette preuve est hors-programme, mais pas le théorème!!!

Soit P un polynôme de $\mathbb{C}[X]$, non constant. Notons $n \in \mathbb{N}^*$ son degré, et $P = \sum_{k=0}^n a_k X^k$.

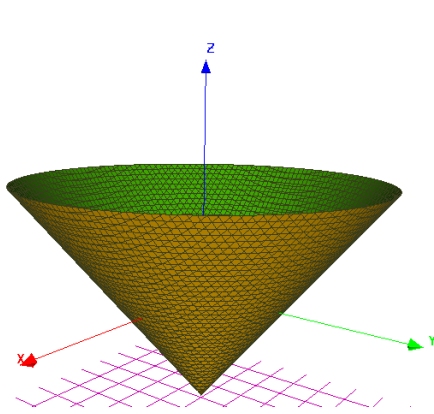
Considérons l'application $\psi : \mathbb{C} \longrightarrow \mathbb{R}_+$

$$z \longmapsto |P(z)|$$

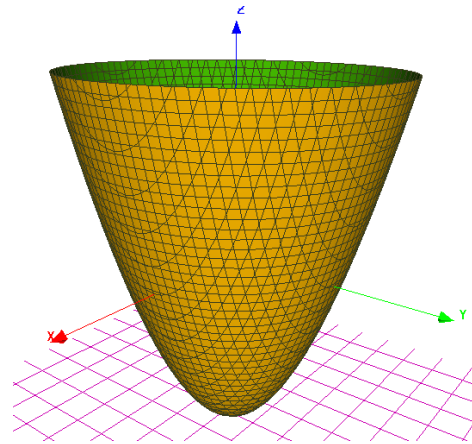
Le principe de la preuve consiste à prouver que la fonction ψ admet un minimum, puis que ce minimum est nul.

Remarque. Si vous voulez vous faire une image de ψ , il faut un peu d'imagination. Puisqu'elle est définie sur $\mathbb{C}$, et à valeurs dans $\mathbb{R}$, ψ ne sera plus représentée par une courbe, mais par une surface. Explicitement, l'image d'un complexe $z = x + iy$ par ψ est le réel positif $\psi(z)$; on se propose alors d'associer à ces données le point de $\mathbb{R}^3$ de coordonnées $(x, y, \psi(x + iy))$. L'ensemble des points obtenus de cette façon en faisant varier x et y est une surface de $\mathbb{R}^3$, que l'on pourra alors appeler surface représentative de la fonction ψ (et que l'on notera S_ψ dans les exemples ci-dessous).

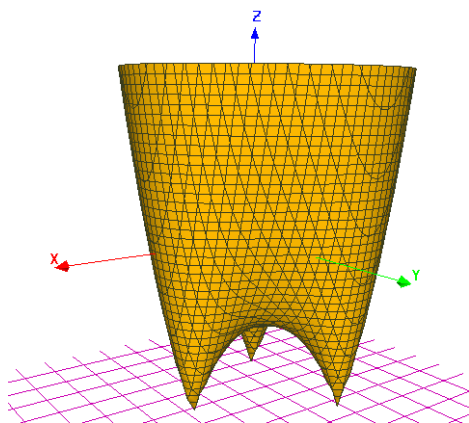
Ci-dessous, on donne quelques exemples de telles surfaces S_ψ associées à différents polynômes P .



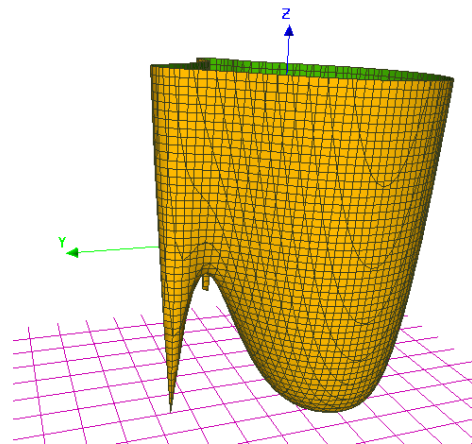
S_ψ pour $P = X$



S_ψ pour $P = X^2$



S_ψ pour $P = X^2 + X + 1$



S_ψ pour
 $P = (X - 1)(X - 1 - i)(X - 1 + i)^4$

► 1ère partie — Où l'on prouve que ψ admet un minimum.

Pour tout nombre complexe z non nul, on a : $\psi(z) = |z^n| \left| a_n + \sum_{k=0}^{n-1} \frac{a_k}{z^{n-k}} \right|$

Soit k un entier de $\llbracket 0, n-1 \rrbracket$. Lorsque $|z|$ tend vers $+\infty$, on a : $\frac{a_k}{z^{n-k}} \longrightarrow 0$.

$$\text{Donc : } \lim_{z \rightarrow +\infty} \left| a_n + \sum_{k=0}^{n-1} \frac{a_k}{z^{n-k}} \right| = |a_n|$$

Puisque $|a_n| > 0$ (a_n est non nul par hypothèse⁹), on a donc : $\lim_{z \rightarrow +\infty} \psi(z) = +\infty$.

En particulier, il existe un réel $R > 0$ tel que : $\forall z \in \mathbb{C}, (|z| > R) \implies (\psi(z) > |a_0| + 1)$ (10

Posons alors $K = D(0, R) = \{z \in \mathbb{C} / |z| \leq R\}$. K est une partie fermée et bornée de $\mathbb{C}$.¹¹ D'après le théorème des bornes atteintes¹², ψ admet un maximum et un minimum sur K .

Ne nous intéressons qu'au minimum de ψ , que l'on note m . Il existe un complexe $z_0 \in D(0, R)$ tel que : $\psi(z_0) = m$.

Alors : (1) : $\forall z \in D(0, R), \psi(z) \geq m$ puisque m est le minimum de ψ sur K .

En particulier : $|a_0| \geq m$ puisque $|a_0| = \psi(0)$. Il s'ensuit que : (2) $\forall z \in \mathbb{C}, (|z| > R) \implies (\psi(z) > m)$ d'après (

On déduit des points (1) et (2) que : $\boxed{\forall z \in \mathbb{C}, \psi(z) \geq m = \psi(z_0)}$ (

Ainsi la fonction ψ admet un minimum (global) sur $\mathbb{C}$. Fin du premier acte.

► 2ème partie — Où l'on prouve que le minimum de ψ est nul.

Commençons par une observation pas trop difficile ; la fonction ψ étant à valeurs dans $\mathbb{R}_+$, son minimum est supérieur ou égal à 0. Pour parvenir à nos fins, il “suffit” donc de prouver que m ne peut pas être strictement positif. Pour cela, rien de tel qu'un petit raisonnement par l'absurde.

Supposons que $m = \psi(z_0) > 0$. Alors la formule de Taylor appliquée en z_0 donne :

$$P(z_0 + Z) = \alpha_0 + \alpha_1 Z + \alpha_2 Z^2 + \cdots + \alpha_n Z^n$$

où les α_i désignent les complexes $\frac{P^{(i)}(z_0)}{i!}$. En particulier : $\alpha_0 = P(z_0) \neq 0$ (par hypothèse), et $\alpha_n \neq 0$ puisque $\deg(P) = n$.

Notons alors : $k = \min \{i \in \llbracket 1, n \rrbracket / \alpha_i \neq 0\}$. L'existence de ce minimum est assurée par le fait que l'ensemble considéré est une partie de $\mathbb{N}^*$ (donc minorée !) non vide (n y appartient). Avec cette définition de k , on a donc :

9. C'est le coefficient dominant de P

10. On utilise la définition de “tendre vers $+\infty$ ” pour une fonction définie sur $\mathbb{C}$ et à valeurs dans $\mathbb{R}$, que voici : soit $f : \mathbb{C} \longrightarrow \mathbb{R}$; $\lim_{z \rightarrow +\infty} f(z) = +\infty$ si pour tout réel M , il existe un réel R tel que : $\forall z \in \mathbb{C}, (|z| > R) \implies (f(z) > M)$.

La motivation pour choisir $R = |a_0| + 1$ dans le cas présent est que $\psi(0) = |a_0|$, et que l'on est donc assuré qu'il existera un voisinage de 0 sur lequel ψ sera inférieure ou égale à $|a_0| + 1$, et donc que $R \neq 0$.

11. Que K soit bornée, vous pouvez le comprendre. “Fermée”, vous pouvez en avoir l'intuition : K est un disque fermé car “on en prend le bord”. Vous verrez l'an prochain une définition vraiment mathématique de ce terme.

12. Qui affirmera que toute fonction continue $f : K \longrightarrow \mathbb{R}$ définie sur une partie fermée et bornée de $\mathbb{C}$ est elle-même bornée et atteint ses bornes. Ici, la fonction ψ est continue en tant que composée d'une fonction polynomiale, et de la fonction valeur absolue, toutes deux continues.

$$P(z_0 + Z) = \alpha_0 + \alpha_k Z^k + \alpha_{k+1} Z^{k+1} + \dots + \alpha_n Z^n$$

Notons encore ω une racine k -ième du complexe $-\frac{\alpha_0}{\alpha_k}$.

Alors pour tout réel t : $P(z_0 + \omega t) = \alpha_0 + \alpha_k \omega^k t^k + \alpha_{k+1} \omega^{k+1} t^{k+1} + \dots + \alpha_n \omega^n t^n$

d'où : $P(z_0 + \omega t) = \alpha_0 - \alpha_0 t^k + \alpha_{k+1} \omega^{k+1} t^{k+1} + \dots + \alpha_n \omega^n t^n$ soit enfin : $P(z_0 + \omega t) = \alpha_0(1 - t^k + o(t^k))$.

D'où : $\psi(z_0 + \omega t) = |\alpha_0| \times |(1 - t^k + o(t^k))|$.

Pour $t > 0$, la quantité $1 - t^k + o(t^k)$ est strictement plus petite que 1 (puisque alors l'expression $t^k - o(t^k)$ est strictement positive) sur un voisinage de 0. En particulier, il existe un réel t_0 tel que : $\psi(z_0 + \omega t_0) < |\alpha_0|$.

En posant $z_1 = z_0 + \omega t_0$ et en se rappelant que $\alpha_0 = P(z_0) = m$, on a donc établi que :

$$\exists z_1 \in \mathbb{C}, \psi(z_1) < \psi(z_0)$$

ce qui contredit violemment le fait que $\psi(z_0)$ est le minimum global de ψ .

Il s'ensuit que $m = \psi(z_0) = 0$, ce qui prouve que ψ s'annule au moins une fois dans $\mathbb{C}$, et donc que P s'annule au moins une fois dans $\mathbb{C}$.

Conclusion. Tout polynôme non constant de $\mathbb{C}[X]$ possède une racine dans $\mathbb{C}$. ☕

Chapitre 20

Probabilités et variables aléatoires usuelles

*Tout au long de ce chapitre, on se place dans la situation où l'univers de l'expérience aléatoire est **fini**.*

20.1 Expérience aléatoire et univers

On donne dans ce premier paragraphe quelques définitions “à la volée”.¹

En premier lieu, on appelle **expérience aléatoire** une expérience dont on ne peut prévoir le résultat (de façon certaine). Une **issue** (ou **éventualité**) est un résultat possible lors d'une expérience aléatoire. Exemples d'expériences aléatoires : lancer de dé, de pièce, pas de nain parceque c'est interdit, jeux divers et variés (cartes, boules, urnes...), note lors d'une évaluation...

L'**univers** (souvent noté Ω) d'une expérience aléatoire est l'ensemble de ses issues.

Un **évènement** est une partie de l'univers. L'**évènement certain** (*resp.* **impossible**) est Ω (*resp.* $\emptyset$). Un évènement élémentaire est un singleton de l'univers (càd un évènement constitué d'une unique éventualité).

20.2 Opérations sur les évènements

On commence ce paragraphe de la même façon que le second, en rappelant des définitions bien connues.

DÉFINITION 20.1 - Soient A et $B \in \mathcal{P}(\Omega)$ (il revient au même d'écrire : “soient A et B deux évènements lors d'une même expérience aléatoire”).

- 1/ L'évènement $A \cap B$ est l'évènement qui est réalisé SSI A et B le sont ;
- 2/ L'évènement $A \cup B$ est l'évènement qui est réalisé SSI A ou B l'est.
- 3/ L'évènement $\overline{A}$ est l'évènement qui est réalisé SSI A ne l'est pas.

1. Essentiellement, il s'agit de définitions qui vous sont familières depuis le Lycée.

Exemples. Lors d'un lancer de dé à 6 faces, on considère les évènements A : “obtenir un résultat pair” et B : “obtenir au moins 5”.

Dans cette situation on a : $A = \{2, 4, 6\}$; $B = \{5, 6\}$; $A \cap B = \{6\}$; $A \cup B = \{2, 4, 5, 6\}$; $\bar{A} = \{1, 3, 5\}$; $\bar{B} = \{1, 2, 3, 4\}$.

Remarque. Ces considérations sur les évènements ne sont finalement, au vocabulaire près, qu'une retranscription des notions introduites lors du chapitre 1 dans le cadre de la théorie des ensembles. On dispose donc déjà d'un nombre impressionnant de propriétés sur les opérations sur les ensembles (distributivités et lois de Morgan notamment). Rappelons-les explicitement.

PROPRIÉTÉ 20.1 - Soient A , B et C trois éléments de $\mathcal{P}(\Omega)$. On a :

$$1/ A \cap \Omega = A; \quad A \cup \bar{A} = \Omega; \quad A \cap \bar{A} = \emptyset$$

$$2/ A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$$

$$3/ A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$$

$$4/ \overline{(A \cap B)} = \bar{A} \cup \bar{B}$$

$$5/ \overline{(A \cup B)} = \bar{A} \cap \bar{B}$$

PREUVE. Voir chapitre 1.



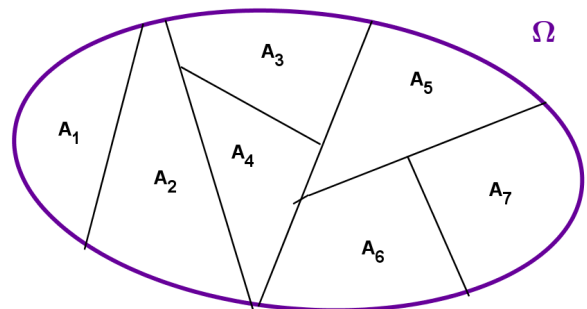
DÉFINITION 20.2 - Deux évènements A et B sont **incompatibles** (ou **disjoints**) si $A \cap B = \emptyset$.

Exemples. Lors d'un lancer de dé, les évènements “obtenir un résultat pair” et “obtenir 3 ou 5” sont incompatibles.

Dans un autre registre, les évènements “habiter Dunkerque” et “ne pas marcher dans le vent” sont incompatibles.

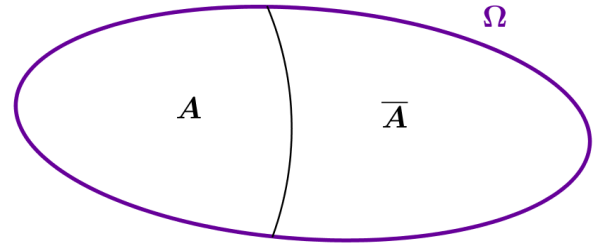
DÉFINITION 20.3 - Soit Ω l'univers d'une expérience aléatoire. Un **système complet d'évènements** (en abrégé SCE ; on dit aussi **partition**) est une famille $(A_i)_{i=1, \dots, n}$ d'évènements deux à deux disjoints tels que $\bigcup_{i=1, \dots, n} A_i = \Omega$.

Illustration. Ci-contre, les évènements $A_1, \dots, A_7$ constituent un système complet d'évènements de l'univers Ω .



Exemple. Lorsque l'expérience aléatoire est un lancer de dé, les évènements $A_1 = \{1, 2, 3\}$, $A_2 = \{4, 5\}$ et $A_3 = \{6\}$ constituent un système complet d'évènements.

Exemple fondamental. Pour tout évènement A , la famille $(A, \bar{A})$ est un système complet d'évènements, puisque $A \cup \bar{A} = \Omega$ et $A \cap \bar{A} = \emptyset$ (voir illustration ci-contre).



20.3 Probabilités

20.3.1 Généralités

DÉFINITION 20.4 - Une **probabilité** $\mathbb{P}$ sur Ω est une application

$$\mathbb{P} : \mathcal{P}(\Omega) \longrightarrow [0; 1]$$

telle que :

1/ $\mathbb{P}(\Omega) = 1$

2/ Si A et B sont deux parties disjointes de Ω , alors $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$

Exemple. Dans le contexte d'un lancer de pièce, on définit une probabilité sur Ω en posant : $\mathbb{P}(\text{pile}) = 1/5$ et $\mathbb{P}(\text{face}) = 4/5$.

PROPRIÉTÉ 20.2 - (Propriétés des probabilités) Soit Ω l'univers d'une expérience aléatoire, et soit $\mathbb{P}$ une probabilité sur Ω .

1/ Pour tout évènement $A \in \mathcal{P}(\Omega)$, $0 \leq \mathbb{P}(A) \leq 1$

2/ $\mathbb{P}(\emptyset) = 0$

3/ Pour tout évènement $A \in \mathcal{P}(\Omega)$, $\mathbb{P}(\bar{A}) = 1 - \mathbb{P}(A)$

4/ Pour tout couple (A, B) d'évènements de $\mathcal{P}(\Omega)$, $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$

PREUVE. Le point 1 résulte de la définition de probabilité : puisque $\mathbb{P}$ est une probabilité sur Ω , elle est à valeurs dans $[0, 1]$.

Le point 2 est une nouvelle conséquence de la définition. En effet, on a d'une part $\mathbb{P}(\Omega) = 1$; et d'autre part les évènements Ω et $\emptyset$ sont disjoints. On en déduit que :

$$\mathbb{P}(\Omega \cup \emptyset) = \mathbb{P}(\Omega) + \mathbb{P}(\emptyset) \quad \text{d'où : } \mathbb{P}(\emptyset) = 1 - \mathbb{P}(\Omega) \quad \text{soit } \mathbb{P}(\emptyset) = 0.$$

Prouvons le point 3 : soit $A \in \mathcal{P}(\Omega)$. Les évènements A et $\bar{A}$ étant disjoints, on a : $\mathbb{P}(A \cup \bar{A}) = \mathbb{P}(A) + \mathbb{P}(\bar{A})$. Par ailleurs, puisque $A \cup \bar{A} = \Omega$, on a : $\mathbb{P}(A \cup \bar{A}) = 1$. On déduit de ces deux relations que : $\mathbb{P}(\bar{A}) = 1 - \mathbb{P}(A)$.

Prouvons le point 4 : soient A et B deux évènements dans Ω . On observe judicieusement que l'évènement $A \cup B$ est la réunion disjointe² de $A \setminus B$ et de B .

2. Dire qu'une partie P est la *réunion disjointe* de deux parties P_1 et P_2 signifie que $P = P_1 \cup P_2$, et que $P_1 \cap P_2 = \emptyset$.

Il s'ensuit que :

$$\mathbb{P}(A \cup B) = \mathbb{P}(A \setminus B) + \mathbb{P}(B) \quad (\spadesuit)$$

Par ailleurs, l'évènement A est aussi la réunion disjointe de $A \setminus B$ et de $A \cap B$. D'où :

$$\mathbb{P}(A) = \mathbb{P}(A \setminus B) + \mathbb{P}(A \cap B) \quad (\clubsuit)$$

On déduit de $(\spadesuit)$ et de $(\clubsuit)$ que : $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$. ☕

Remarque. Les deux affirmations concernant les réunions disjointes dans la preuve du point 4 peuvent être retrouvées rapidement à l'aide d'un dessin ; et on peut naturellement les prouver en utilisant les définitions des ensembles intervenant dans ces affirmations (c'est un exercice facile de théorie des ensembles).

PROPRIÉTÉ 20.3 - (Croissance) Pour toute probabilité $\mathbb{P}$ sur $\mathcal{P}(\Omega)$, et pour tout couple (A, B) d'évènements dans Ω , on a :

$$[A \subset B] \implies [\mathbb{P}(A) \leq \mathbb{P}(B)]$$

PREUVE. Soient A et B deux évènements dans Ω . Si $A \subset B$, alors : $B = A \cup (B \setminus A)$. Or cette union est disjointe, donc : $\mathbb{P}(B) = \mathbb{P}(A) + \mathbb{P}(B \setminus A)$. Une probabilité étant toujours positive, on en déduit que : $\mathbb{P}(B) \geq \mathbb{P}(A)$. ☕

Remarque. Une autre observation, qui concerne l'exemple précédant les propriétés cette fois-ci. L'exemple de probabilité donné dans le contexte de l'expérience de pile ou face peut être généralisé à un univers fini. C'est l'objet de la propriété suivante.

PROPRIÉTÉ 20.4 - Soit Ω un univers fini, que l'on se propose de noter $\Omega = \{x_1, \dots, x_n\}$, où n désigne un entier naturel non nul (et les x_i les issues de l'expérience aléatoire).

Soit $(p_1, \dots, p_n)$ un n -uplets de réels tels que :

$$1/ \forall i \in \llbracket 1, n \rrbracket, p_i \in [0, 1]; \quad 2/ \sum_{i=1}^n p_i = 1$$

Alors il existe une unique probabilité $\mathbb{P}$ sur $\mathcal{P}(\Omega)$ telle que :

$$\forall i \in \llbracket 1, n \rrbracket, \mathbb{P}(\{x_i\}) = p_i$$

Traduction informelle. On définit de manière unique une probabilité sur un univers fini en choisissant les probabilités de chacun des évènements élémentaires, de telle sorte que la somme de toutes ces probabilités soit égale à 1.

PREUVE. Avec les notations de l'énoncé, il s'agit de vérifier que l'application $\mathbb{P}$ est correctement définie, à valeurs dans $[0, 1]$, telle que $\mathbb{P}(\Omega) = 1$, et telle que $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$ dès lors que A et B sont deux parties disjointes de Ω .

Soit A une partie de Ω . Alors : $A = \{x_{i_1}, \dots, x_{i_m}\}$ (en ayant noté m le cardinal de A). On peut observer que A est la réunion disjointe de ses singletons, ce que l'on peut noter :

$$A = \bigsqcup_{j=1}^m \{x_{i_j}\}$$

Il s'ensuit que : $\mathbb{P}(A) = \sum_{j=1}^m \mathbb{P}(\{x_{i_j}\}) = \sum_{j=1}^m p_{i_j}$.

L'application $\mathbb{P}$ est donc correctement définie ; il reste à vérifier qu'elle est à valeurs dans $[0, 1]$.

Or, en conservant les notations précédemment introduites, on a :

$$0 \leq \mathbb{P}(A) = \sum_{j=1}^m p_{i_j} \leq \sum_{i=1}^n p_i = 1 = 1$$

puisque les réels p_i sont tous positifs ou nuls et de somme égale à 1 par hypothèse. Il s'ensuit donc que $\mathbb{P}$ est définie sur $\mathcal{P}(\Omega)$, et à valeurs dans $[0, 1]$.

Enfin, il résulte de la construction de $\mathbb{P}$ que $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$ dès lors que A et B sont deux parties disjointes de Ω , puisque la probabilité d'un évènement est la somme des probas des évènements élémentaires qui le composent.

Pour conclure, cette probabilité $\mathbb{P}$ est unique, car s'il existait une autre probabilité $\mathbb{P}'$ telle que : $\forall i \in \llbracket 1, n \rrbracket, \mathbb{P}'(\{x_i\}) = p_i$, alors il en résulterait (voir calculs faits plus haut) que $\mathbb{P}'(A) = \mathbb{P}(A)$ pour tout évènement A de $\mathcal{P}(\Omega)$. ☕

Exemple-clef — Situation d'équiprobabilité. Notons $\Omega = \{\omega_1, \dots, \omega_n\}$, et notons pour tout $i \in \llbracket 1, n \rrbracket, p_i = \mathbb{P}(\{\omega_i\})$.

D'après la propriété précédente, on définit une probabilité $\mathbb{P}$ sur $\mathcal{P}(\Omega)$ en posant :

$$\forall i \in \llbracket 1, n \rrbracket, \quad \mathbb{P}(\{\omega_i\}) = \frac{1}{n}$$

On parle alors d'**équiprobabilité**.

Dans cette situation, on a pour tout évènement A dans Ω :

$$\mathbb{P}(A) = \frac{\text{card } A}{\text{card } \Omega} = \frac{\text{card } A}{n}$$

Pour faire le lien avec une affirmation que vous avez peut-être déjà entendue, cette formule correspond au quotient :

$$\frac{\text{nombre de cas favorables}}{\text{nombre de cas possibles}}$$

Exemple. Quand on lance un dé non truqué (ou équilibré), on est dans la situation d'équiprobabilité, dans le sens où la probabilité d'obtenir chacune des faces est égal à $1/6$.

3. Où $\text{card } A$ désigne le **cardinal** de A , c'est-à-dire son nombre d'éléments.

20.3.2 Formule des probabilités totales

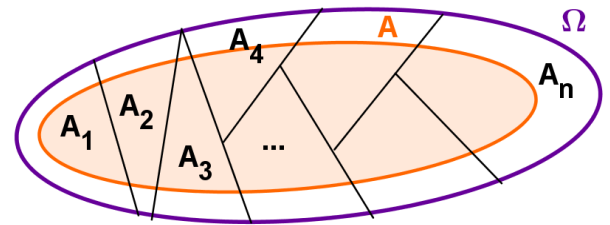
Place maintenant à l'une des propriétés essentielles de ce chapitre de probabilités.

THÉORÈME 20.1 - (Formule des probabilités totales) Soit $(A_i)_{i=1,\dots,n}$ un système complet d'événements de Ω .

On a :

$$\forall A \in \mathcal{P}(\Omega), \quad \mathbb{P}(A) = \sum_{i=1}^n \mathbb{P}(A \cap A_i)$$

Illustration. Sur la figure ci-contre, on a fait apparaître le fait que la probabilité de l'événement A est la somme des probabilités des “petits morceaux” qui composent A , c'est-à-dire la somme des probabilités des événements $A \cap A_1, \dots, A \cap A_n$.



PREUVE. Soit $(A_i)_{i=1,\dots,n}$ un système complet d'événements, et soit A un événement quelconque de Ω .

Le point-clef consiste ici à observer que les événements $(A \cap A_i)$ sont deux à deux disjoints, et recouvrent A . En clair, il s'agit d'établir les deux faits suivants :

$$1) \forall (i, j) \in \llbracket 1, n \rrbracket^2, i \neq j \implies (A \cap A_i) \cap (A \cap A_j) = \emptyset \quad \text{et} \quad 2) \bigcup_{i=1}^n (A \cap A_i) = A$$

► Pour le point 1 : il suffit d'observer que pour tout couple (i, j) d'entiers (compris entre 1 et n), on a ⁴ : $(A \cap A_i) \cap (A \cap A_j) = A \cap (A_i \cap A_j)$. Or si $i \neq j$, alors $A_i \cap A_j = \emptyset$ puisque les A_i constituent un SCE par hypothèse. On en déduit le point 1).

► Pour le point 2 : comme $(A_i)_{i=1,\dots,n}$ est un système complet d'événements, on peut déjà affirmer que :

$$\Omega = \bigcup_{i=1}^n A_i. \text{ On en déduit que : } A \cap \Omega = A \cap \left(\bigcup_{i=1}^n A_i \right). \text{ En utilisant encore une fois la distributivité de } \cap$$

par rapport à $\cup$, on obtient : $A = \bigcup_{i=1}^n (A \cap A_i)$, ce qui établit le point 2.

► Conclusion : d'après ce qui précède $\mathbb{P}(A) = \mathbb{P}\left(\bigcup_{i=1}^n (A \cap A_i)\right)$. Et puisque d'après le 1), les événements $(A \cap A_i)$ sont deux à deux disjoints, la probabilité de leur réunion est la somme de leurs probabilités individuelles, soit :

$$\mathbb{P}\left(\bigcup_{i=1}^n (A \cap A_i)\right) = \sum_{i=1}^n \mathbb{P}(A \cap A_i) \quad \text{d'où} \quad \boxed{\mathbb{P}(A) = \sum_{i=1}^n \mathbb{P}(A \cap A_i)} \quad \text{☕}$$

4. En utilisant l'associativité de l'intersection d'une part, et la propriété décoiffante selon laquelle $A \cap A = A$.

COROLLAIRE 20.1 - Soit A un évènement quelconque. On a :

$$\forall B \in \mathcal{P}(\Omega), \quad \mathbb{P}(B) = \mathbb{P}(B \cap A) + \mathbb{P}(B \cap \bar{A})$$

PREUVE. Il suffit d'appliquer le théorème 20.1 au système complet d'évènements $(A, \bar{A})$. ☕

On renvoie à la prochaine section pour des applications de la FPT.

20.4 Probabilités conditionnelles et évènements indépendants

DÉFINITION 20.5 - Soient A et B deux évènements, avec $\mathbb{P}(A) \neq 0$. La **probabilité conditionnelle de B sachant A** (càd sachant que A est réalisé) est :

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

Exemples. Lors du lancer d'un dé équilibré, on considère les évènements :

A : "obtenir un résultat pair" B : "obtenir au plus 3" et C : "obtenir au moins 3"

$$\text{On a : } \mathbb{P}_A(B) = \frac{1/6}{1/2} = \frac{1}{3}; \quad \mathbb{P}_A(C) = \frac{2/6}{1/2} = \frac{2}{3}; \quad \mathbb{P}_C(A) = \frac{2/6}{2/3} = \frac{1}{2}; \dots$$

Remarque. Il résulte de la définition que : $\mathbb{P}(B \cap A) = \mathbb{P}_A(B) \times \mathbb{P}(A)$.

PROPRIÉTÉ 20.5 - (Propriétés des probabilités conditionnelles) Soit A un évènement tel que $\mathbb{P}(A) \neq 0$. On a :

1/ Pour tout évènement B : $\mathbb{P}_A(\bar{B}) = 1 - \mathbb{P}_A(B)$.

2/ Pour tout couple (B_1, B_2) d'évènements : $\mathbb{P}_A(B_1 \cup B_2) = \mathbb{P}_A(B_1) + \mathbb{P}_A(B_2) - \mathbb{P}_A(B_1 \cap B_2)$.

PREUVE. Soit A un évènement tel que $\mathbb{P}(A) \neq 0$.

Alors pour tout évènement B , on a : $\Omega = B \cup \bar{B}$. D'où : $A \cap \Omega = A \cap (B \cup \bar{B})$.

Or $A \cap \Omega = A$ (puisque $A \subset \Omega$); et $A \cap (B \cup \bar{B}) = (A \cap B) \cup (A \cap \bar{B})$ (distributivité de $\cap$ par rapport à $\cup$).

On en déduit que $\mathbb{P}(A) = \mathbb{P}((A \cap B) \cup (A \cap \bar{B}))$.

Or l'union $(A \cap B) \cup (A \cap \bar{B})$ est disjointe⁵, donc (par définition de probabilité) :

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

En résumé, on a établi que : $\mathbb{P}(A) = \mathbb{P}(A \cap B) + \mathbb{P}(A \cap \bar{B})$, d'où en divisant tous les termes de cette égalité par $\mathbb{P}(A)$ (qui est non nulle) : $1 = \mathbb{P}_A(B) + \mathbb{P}_A(\bar{B})$, soit finalement : $\mathbb{P}_A(\bar{B}) = 1 - \mathbb{P}_A(B)$.

5. C'est-à-dire : $(A \cap B) \cap (A \cap \bar{B}) = \emptyset$. La preuve de ce fait est triviale, puisque si un élément appartient aux deux termes de l'intersection, alors il est à la fois dans B , et dans son complémentaire $\bar{B}$.

► Pour le second point, on considère B_1 et B_2 deux évènements arbitraires dans Ω . On a alors :

$$\begin{aligned}\mathbb{P}_A(B_1 \cup B_2) &= \frac{\mathbb{P}((B_1 \cup B_2) \cap A)}{\mathbb{P}(A)} = \frac{\mathbb{P}((B_1 \cap A) \cup (B_2 \cap A))}{\mathbb{P}(A)} \\ &= \frac{\mathbb{P}(B_1 \cap A) + \mathbb{P}(B_2 \cap A) - \mathbb{P}((B_1 \cap B_2) \cap A)}{\mathbb{P}(A)} = \mathbb{P}_A(B_1) + \mathbb{P}_A(B_2) - \mathbb{P}_A(B_1 \cap B_2)\end{aligned}$$



On aurait pu faire l'économie de cette preuve en établissant que $\mathbb{P}_A$ définit une probabilité sur $\mathcal{P}(\Omega)$ (à ce titre, elle vérifie donc les propriétés des probas).

PROPRIÉTÉ 20.6 - (Propriétés des probabilités conditionnelles) Soit A un évènement tel que $\mathbb{P}(A) \neq 0$.

L'application $\mathbb{P}_A$ est une probabilité sur $\mathcal{P}(\Omega)$.

PREUVE. Soit A un évènement tel que $\mathbb{P}(A) \neq 0$. Il s'agit de vérifier que $\mathbb{P}_A$ est une application de $\mathcal{P}(\Omega)$ à valeurs dans $[0, 1]$, et que $\mathbb{P}_A(B \cup C) = \mathbb{P}_A(B) + \mathbb{P}_A(C)$ pour tout couple (B, C) d'évènements disjoints dans Ω .

Soit $B \in \mathcal{P}(\Omega)$ un évènement quelconque. On a : $(B \cap A) \subset A$. D'où (en vertu de la propriété 20.3 page 508) : $0 \leq \mathbb{P}(B \cap A) \leq \mathbb{P}(A)$.

On en déduit que : $0 \leq \mathbb{P}_A(B) \leq 1$.

En résumé : $\forall B \in \mathcal{P}(\Omega), \mathbb{P}_A(B) \in [0, 1]$. Autrement dit, $\mathbb{P}_A$ est définie sur $\mathcal{P}(\Omega)$ et à valeurs dans $[0, 1]$.

Considérons à présents B et C deux évènements disjoints dans Ω . On a, d'après le point 2 de la propriété 20.5 :

$$\mathbb{P}_A(B \cup C) = \mathbb{P}_A(B) + \mathbb{P}_A(C) - \mathbb{P}_A(B \cap C)$$

Or B et C étant disjoints (ou incompatibles), on a : $\mathbb{P}_A(B \cap C) = 0$. Il s'ensuit que :

$$\mathbb{P}_A(B \cup C) = \mathbb{P}_A(B) + \mathbb{P}_A(C)$$

On peut alors conclure que $\mathbb{P}_A$ est une probabilité sur $\mathcal{P}(\Omega)$. 

DÉFINITION 20.6 - Deux évènements A et B sont **indépendants** lorsque :

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

Quand tel est le cas, on a : $\mathbb{P}_A(B) = \mathbb{P}(B)$.

Exemples. Lors du lancer d'un dé équilibré, on considère les évènements :

A : "obtenir un résultat pair" B : "obtenir au plus 3" et C : "obtenir au moins 3"

On a : $\mathbb{P}(A \cap C) = 1/3$, $\mathbb{P}(A) = 1/2$ et $\mathbb{P}(C) = 2/3$. D'où : $\mathbb{P}(A \cap C) = \mathbb{P}(A) \times \mathbb{P}(C)$. Les événements A et C **sont indépendants**.

On a : $\mathbb{P}(A \cap B) = 1/6$, $\mathbb{P}(A) = 1/2$ et $\mathbb{P}(B) = 1/2$. D'où : $\mathbb{P}(A \cap B) \neq \mathbb{P}(A) \times \mathbb{P}(B)$. Les événements A et B **ne sont pas indépendants**.

Remarque. Attention !!! **Indépendant $\neq$ Incompatible**

Dans l'exemple précédent, les événements A et C sont indépendants mais ne sont pas incompatibles. D'une manière générale d'ailleurs, lorsque deux événements A et B sont incompatibles (càd si $\mathbb{P}(A \cap B) = 0$), ils ne sont pas indépendants dès lors que $\mathbb{P}(A)$ et $\mathbb{P}(B)$ sont non nulles.

La notion de probabilité conditionnelle permet de donner une autre traduction de la formule des probabilités totales (théorème 20.1, page 510).

THÉORÈME 20.2 - (Formule des probabilités totales, revisitée) Soit $(A_i)_{i=1,\dots,n}$ un système complet d'événements de Ω .

On suppose que : $\forall i \in \llbracket 1, n \rrbracket, \mathbb{P}(A_i) \neq 0$.

Alors :

$$\forall A \in \mathcal{P}(\Omega), \quad \mathbb{P}(A) = \sum_{i=1}^n \mathbb{P}_{A_i}(A) \times \mathbb{P}(A_i)$$

PREUVE. Sous les hypothèses de l'énoncé, on a : $\forall i \in \llbracket 1, n \rrbracket, \mathbb{P}(A \cap A_i) = \mathbb{P}_{A_i}(A) \times \mathbb{P}(A_i)$. Cette remarque faite, on déduit directement l'égalité de celle établie dans le théorème 20.1. ☕

Remarque. Cette formule justifie la méthode consistant à calculer des probabilités à l'aide d'un arbre pondéré, où l'on "ajoute les probas correspondant aux différents chemins, et où l'on multiplie les probas correspondant aux branches successives".

Exemple d'application 1. On dispose d'un dé cubique équilibré dont une face porte le numéro 1, deux faces portent le numéro 2 et trois faces portent le numéro 3.

On dispose également d'une urne contenant dix boules indiscernables au toucher, portant les lettres L, O, G, A, R, I, T, H, M, E.

Un joueur fait une partie en deux étapes.

Première étape : il jette le dé et note le numéro obtenu.

Deuxième étape.

- Si le dé indique 1, il tire au hasard une boule de l'urne. Il gagne la partie si cette boule porte une voyelle et il perd dans le cas contraire ;
- Si le dé indique 2, il tire au hasard et simultanément deux boules de l'urne. Il gagne la partie si chacune de ces deux boules porte une voyelle et il perd dans le cas contraire ;
- Si le dé indique 3, il tire au hasard et simultanément trois boules de l'urne. Il gagne la partie si chacune de ces trois boules porte une voyelle et il perd dans le cas contraire.

On définit les événements suivants :

- D_1 : “le dé indique 1” ;
- D_2 : “le dé indique 2” ;
- D_3 : “le dé indique 3” ;
- G : “la partie est gagnée”

1/ Calculer les probabilités conditionnelles $P_{D_1}(G)$, $P_{D_2}(G)$ et $P_{D_3}(G)$.

2/ Calculer la probabilité que le joueur gagne la partie.

3/ Calculer la probabilité que le joueur gagne la partie s’il obtient un 1 lors du lancer de dé.

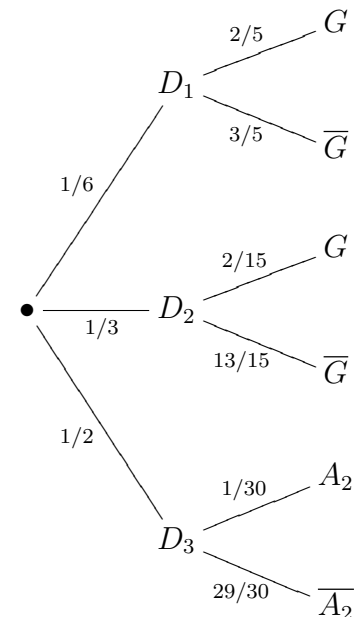
Corrigé. Il peut être utile de représenter l’expérience décrite dans l’énoncé à l’aide d’un arbre pondéré (même si ce n’est pas explicitement demandé dans l’énoncé).

$$1) P_{D_1}(G) = \frac{4}{10} \text{ soit } \boxed{P_{D_1}(G) = \frac{2}{5}}; P_{D_2}(G) = \frac{\binom{4}{2}}{\binom{10}{2}} = \frac{6}{45} \text{ soit}$$

$$\boxed{P_{D_2}(G) = \frac{2}{15}}; \text{ enfin } P_{D_3}(G) = \frac{\binom{4}{3}}{\binom{10}{3}} = \frac{4}{120} \text{ d'où } \boxed{P_{D_3}(G) = \frac{1}{30}}.$$

Les événements D_1 , D_2 et D_3 constituent une partition de l’univers de cette expérience. La formule des probabilités totales donne alors : $P(G) = P(G \cap D_1) + P(G \cap D_2) + P(G \cap D_3) = P_{D_1}(G) \times P(D_1) + P_{D_2}(G) \times P(D_2) + P_{D_3}(G) \times P(D_3)$. Par conséquent :

$$P(G) = \frac{2}{5} \times \frac{1}{6} + \frac{1}{3} \times \frac{2}{15} + \frac{1}{2} \times \frac{1}{30} = \frac{2}{30} + \frac{2}{45} + \frac{1}{60} = \frac{12 + 8 + 3}{180} \text{ d'où } \boxed{P(G) = \frac{23}{180}}$$



$$3) \text{ La probabilité demandée est } P_G(D_1). \text{ Or : } P_G(D_1) = \frac{P(G \cap D_1)}{P(G)} = \frac{\left(\frac{2}{30}\right)}{\left(\frac{23}{180}\right)} \text{ soit } \boxed{P_G(D_1) = \frac{12}{23}}.$$

Exemple d’application 2. Dans un stand de tir, un tireur effectue des tirs successifs pour atteindre plusieurs cibles. La probabilité que la première cible soit atteinte est $\frac{1}{2}$. Lorsqu’une cible est atteinte, la probabilité que la suivante le soit est $\frac{3}{4}$. Lorsqu’une cible n’est pas atteinte, la probabilité que la suivante soit atteinte est $\frac{1}{2}$.

On note, pour tout entier naturel n non nul :

• A_n l'évènement : « la n-ième cible est atteinte ». • $\bar{A}_n$ l'évènement : « la n-ième cible n'est pas atteinte ». • a_n la probabilité de l'évènement A_n • b_n la probabilité de l'évènement $\bar{A}_n$.	2) Montrer que, pour tout $n \in \mathbb{N}$, $n \geq 1$: $a_{n+1} = \frac{3}{4}a_n + \frac{1}{2}b_n$, puis : $a_{n+1} = \frac{1}{4}a_n + \frac{1}{2}$ 3) Dédurre de la question précédente la limite de la suite (a_n).
- 1) Donner a_1 et b_1 . Calculer a_2 et b_2 .

Corrigé. 1) D'après l'énoncé : $a_1 = 1/2$ et $b_1 = 1/2$.

Au second tir, on peut appliquer la formule des probabilités totales au système complet d'évènements $(A_1, \bar{A}_1)$:

$$a_2 = p(A_2) = p_{A_1}(A_2) \times p(A_1) + p_{\bar{A}_1}(A_2) \times p(\bar{A}_1) = \frac{3}{4} \times \frac{1}{2} + \frac{1}{2} \times \frac{1}{2} \text{ soit : } a_2 = \frac{5}{8}.$$

Et puisque $b_2 = p(\bar{A}_2)$, on a donc $b_2 = \frac{5}{8}$.

2) Soit n un entier naturel non nul. On peut appliquer la formule des probabilités totales au système complet d'évènements $(A_n, \bar{A}_n)$:

$$a_{n+1} = p(A_{n+1}) = p_{A_n}(A_{n+1}) \times p(A_n) + p_{\bar{A}_n}(A_{n+1}) \times p(\bar{A}_n) = \frac{3}{4}a_n + \frac{1}{2}b_n \text{ soit : } \forall n \in \mathbb{N}^*, a_{n+1} = \frac{3}{4}a_n + \frac{1}{2}b_n.$$

Comme en outre $b_n = 1 - a_n$, on a aussi : $\forall n \in \mathbb{N}^*, a_{n+1} = \frac{1}{4}a_n + \frac{1}{2}$

3) D'après la question précédente, la suite (a_n) est arithmético-géométrique ; elle est convergente puisque le coefficient $1/4$ est en module strictement inférieur à 1. Et sa limite ℓ vérifie : $\ell = \frac{1}{4}\ell + \frac{1}{2}$ d'où : $\ell = \frac{2}{3}$.

Conclusion : la suite (a_n) converge et $\lim_{n \rightarrow +\infty} a_n = \frac{2}{3}$.

Pour finir, on dispose également d'un corollaire "bis" de la formule des probabilités totales, que voici.

COROLLAIRE 20.2 - Soient B un évènement dans Ω , tel que $\mathbb{P}(B) \neq 0$ et $\mathbb{P}(B) \neq 1$.

Alors :

$$\forall A \in \mathcal{P}(\Omega), \quad \mathbb{P}(A) = \mathbb{P}_B(A) \times \mathbb{P}(B) + \mathbb{P}_{\bar{B}}(A) \times \mathbb{P}(\bar{B})$$

PREUVE. Il suffit d'appliquer le théorème 20.2 au système complet d'évènements $(A, \bar{A})$. 

On peut déduire de la formule des probabilités totales l'énoncé suivant.

PROPRIÉTÉ 20.7 - (l'une des formules de Bayes) Soit $(A_j)_{j=1,\dots,n}$ un système complet d'événements.

On suppose que : $\forall j \in \llbracket 1, n \rrbracket, \mathbb{P}(A_j) \neq 0$.

Alors pour tout événement $B \in \mathcal{P}(\Omega)$ tel que $\mathbb{P}(B) \neq 0$ on a :

$$\forall i \in \llbracket 1, n \rrbracket, \mathbb{P}_B(A_i) = \frac{\mathbb{P}_{A_i}(B) \times \mathbb{P}(A_i)}{\sum_{j=1}^n \mathbb{P}_{A_j}(B) \times \mathbb{P}(A_j)}$$

PREUVE. Soit i un entier compris entre 1 et n . Par définition de probabilité conditionnelle, on a : $\mathbb{P}_B(A_i) = \frac{\mathbb{P}(B \cap A_i)}{\mathbb{P}(B)}$, ce que l'on peut encore écrire : $\mathbb{P}_B(A_i) = \frac{\mathbb{P}_{A_i}(B) \times \mathbb{P}(A_i)}{\mathbb{P}(B)}$.

Or, d'après la FPT : $\mathbb{P}(B) = \sum_{j=1}^n \mathbb{P}_{A_j}(B) \times \mathbb{P}(A_j)$, d'où la conclusion. ☕

20.5 Variables aléatoires réelles

20.5.1 Généralités

DÉFINITION 20.7 - Un **espace probabilisé** est un couple $(\Omega, \mathbb{P})$, où Ω est un univers et $\mathbb{P}$ une probabilité sur Ω .

Exemple. Un univers fini $\Omega = \{\omega_1, \dots, \omega_n\}$ muni de l'équiprobabilité est un exemple d'espace probabilisé.

DÉFINITION 20.8 - Une **variable aléatoire réelle** (en abrégé **VAR**) X sur un espace probabilisé $(\Omega, \mathbb{P})$ est une application $X : \Omega \longrightarrow \mathbb{R}$.

Exemples.

1/ On note X_1 le nombre de bonnes réponses obtenues par un étudiant répondant au hasard aux 10 questions d'un QCM. Dans ce cadre, X_1 est une VAR.

2/ On note X_2 le nombre de QCM nécessaires à l'étudiant précédemment évoqué pour qu'il fasse un "sans faute". X_2 est un autre exemple de VAR.

Remarque. Lorsque X est une VAR, $X(\Omega)$ désigne l'ensemble des valeurs prises par l'application X .

Cette notation n'est pas nouvelle, puisque l'on a défini en général la notation $f(E)$ pour désigner l'image d'une application $f : E \rightarrow F$ (dans le chapitre sur les applications).

Exemples. En reprenant le contexte des exemples précédents, on a :

$$X_1(\Omega) = \llbracket 0, 10 \rrbracket \quad \text{et} \quad X_2(\Omega) = \mathbb{N}^*$$

20.5.2 Espérance et variance d'une VAR

DÉFINITION 20.9 - Soit X une VAR. L'**espérance mathématique** de X , notée $E(X)$ est le nombre réel :

$$E(X) = \sum_{\omega \in \Omega} X(\omega) \mathbb{P}(\{\omega\}) \text{ (ou } E(X) = \sum_{\omega \in \Omega} X(\omega) \mathbb{P}(X = \omega))$$

Remarque. Dans le cas où Ω est fini (càd dans le seul cas que nous étudierons cette année), la somme de la définition ci-dessus comporte un nombre fini de termes et elle est donc bien définie (ouf!).

En outre, on peut noter $x_1, \dots, x_n$ les différentes valeurs prises par X , et $p_1, \dots, p_n$ les probabilités correspondantes. Avec ces notations, l'espérance de X est le réel :

$$E(X) = \sum_{i=1}^n x_i p_i$$

Exemple. Un joueur lance une pièce truquée de telle sorte que la probabilité d'obtenir pile soit égale à p (avec $p \in [0, 1]$). Le joueur gagne 10 euros s'il obtient pile, et perd 3 euros s'il obtient face. On note X la variable aléatoire égale au gain du joueur.⁶

L'espérance de X est alors : $E(X) = 10 \times p + (-3) \times (1 - p)$ soit $E(X) = -3 + 13p$.

Au passage, on dit qu'un jeu est équitable lorsque l'espérance du gain est nulle. Sur cet exemple, le jeu est équitable lorsque $p = 3/13$. Il est défavorable au joueur lorsque $p < 3/13$, et favorable au joueur lorsque $p > 3/13$.

Remarque. On dit souvent espérance (tout court) pour désigner l'espérance mathématique de X . De plus, l'espérance est également appelée **valeur moyenne** de X .

DÉFINITION 20.10 - Soit X une VAR. La **variance de X** est le réel noté $V(X)$ et défini en posant :

$$V(X) = \sum_{\omega \in \Omega} \mathbb{P}(\{\omega\}) (X(\omega) - E(X))^2 \text{ (ou } V(X) = \sum_{i=1}^n [x_i - E(X)]^2 p_i)$$

Remarque. Dans le cas où Ω est fini (càd...), la somme de la définition ci-dessus comporte un nombre fini de termes et elle est donc bien définie (re-ouf!).

Exemple. On reprend l'exemple donné précédemment : un joueur lance une pièce truquée de telle sorte que la probabilité d'obtenir pile soit égale à p (avec $p \in [0, 1]$). Le joueur gagne 10 euros s'il obtient pile, et perd 3 euros s'il obtient face. On note X la variable aléatoire égale au gain du joueur. On a déjà calculé l'espérance de X : $E(X) = -3 + 13p$.

6. Dans ce contexte, on a donc : $X(\Omega) = \{-3, 10\}$.

La variance de X est :

$$\begin{aligned} V(X) &= [10 - (-3 + 13p)]^2 \times p + [-3 - (-3 + 13p)]^2 \times (1 - p) = p(13 - 13p)^2 + (1 - p)(13p)^2 \\ &= 169p - 338p^2 + 169p^2 + 169p^2 - 169p^3 = 169p(1 - p^2) \end{aligned}$$

20.5.3 Propriétés de l'espérance et la variance

PROPRIÉTÉ 20.8 - Soit X une VAR sur un espace probabilisé $(\Omega, \mathbb{P})$. Alors :

$$\forall (a, b) \in \mathbb{R}^2, E(aX + b) = aE(X) + b$$

PREUVE. Notons $\Omega = \{\omega_1, \dots, \omega_n\}$, et pour tout entier i compris entre 1 et n notons : $x_i = X(\omega_i)$ et $p_i = P(X = x_i)$. Soient a et b deux réels arbitraires.

La variable aléatoire $aX + b$ prend les valeurs : $ax_1 + b, \dots, ax_n + b$ avec les probabilités $p_1, \dots, p_n$. Il s'ensuit que :

$$E(aX + b) = \sum_{i=1}^n (ax_i + b) p_i = \sum_{i=1}^n (ax_i p_i + b p_i) = a \sum_{i=1}^n x_i p_i + b \underbrace{\sum_{i=1}^n p_i}_{=1} = aE(X) + b \quad \text{☕}$$

Soit X une VAR sur un espace probabilisé $(\Omega, \mathbb{P})$, et soit f une fonction définie sur $\mathbb{R}$ et à valeurs dans $\mathbb{R}$. L'application composée $f \circ X$ est définie sur Ω , et à valeurs réelles. A ce titre, $f \circ X$ est une VAR, que l'on note parfois simplement $f(X)$.

Exemple. On reprend l'exemple donné plus haut : un joueur lance une pièce truquée de telle sorte que la probabilité d'obtenir pile soit égale à p (avec $p \in [0, 1]$)...

Considérons par exemple pour f la fonction carrée. Dans ce cas : $f(X) = X^2$ est la VAR qui peut prendre comme valeurs 9 et 100, avec les probabilités respectives $(1 - p)$ et p .

THÉORÈME 20.3 - (Formule de transfert). Soit X une VAR sur un espace probabilisé $(\Omega, \mathbb{P})$, et soit f une fonction définie sur $\mathbb{R}$ et à valeurs réelles.

On a :

$$E(f(X)) = \sum_{i=1}^n f(x_i) \mathbb{P}(X = x_i)$$

PREUVE. Notons comme d'habitude $\Omega = \{\omega_1, \dots, \omega_n\}$, et pour tout entier i compris entre 1 et n notons : $x_i = X(\omega_i)$ et $p_i = P(X = x_i)$. Soit $f \in \mathbb{R}^{\mathbb{R}}$.

La variable aléatoire $f(X) = f \circ X$ prend les valeurs : $f(x_1), \dots, f(x_n)$ avec les probabilités $p_1, \dots, p_n$. Il s'ensuit que :

$$E(f(X)) = \sum_{i=1}^n f(x_i) p_i = \sum_{i=1}^n f(x_i) \mathbb{P}(X = x_i) \quad \text{☕}$$

PROPRIÉTÉ 20.9 - Soit X une VAR sur un espace probabilisé $(\Omega, \mathbb{P})$. Alors :

$$\forall (a, b) \in \mathbb{R}^2, V(aX + b) = a^2 V(X)$$

PREUVE. Notons $\Omega = \{\omega_1, \dots, \omega_n\}$, et pour tout entier i compris entre 1 et n notons : $x_i = X(\omega_i)$ et $p_i = P(X = x_i)$. Soient a et b deux réels arbitraires.

La variable aléatoire $aX + b$ prend les valeurs : $ax_1 + b, \dots, ax_n + b$ avec les probabilités $p_1, \dots, p_n$. Il s'ensuit que :

$$\begin{aligned} V(aX + b) &= \sum_{i=1}^n (ax_i + b - E(aX + b))^2 p_i = \sum_{i=1}^n (ax_i + b - aE(X) - b)^2 p_i \\ &= \sum_{i=1}^n a^2 (x_i - E(X))^2 p_i = a^2 \sum_{i=1}^n (x_i - E(X))^2 p_i = a^2 V(X) \end{aligned} \quad \text{☕}$$

THÉORÈME 20.4 - (Formule de Koenig-Huygens). Avec les mêmes notations que depuis le début de ce paragraphe :

$$V(X) = E(X^2) - [E(X)]^2$$

PREUVE. Avec les notations usuelles :

$$V(X) = \sum_{i=1}^n (x_i - E(X))^2 p_i = \underbrace{\sum_{i=1}^n x_i^2 p_i}_{=E(X^2)} - 2E(X) \underbrace{\sum_{i=1}^n x_i p_i}_{=E(X)} + E(X)^2 \underbrace{\sum_{i=1}^n p_i}_{=1}$$

D'où : $V(X) = E(X^2) - 2E(X)^2 + E(X)^2$. Ainsi : $V(X) = E(X^2) - E(X)^2$. ☕

PROPRIÉTÉ 20.10 - (Inégalité de Markov (Марков)). Soit X une VAR positive. On a :

$$\forall a \in \mathbb{R}_+, a\mathbb{P}(X \geq a) \leq E(X)$$

PREUVE. Soit X une VAR positive sur un espace probabilisé $(\Omega, \mathbb{P})$, et soit a un réel positif. On a :

$$E(X) = \sum_{\omega \in \Omega} X(\omega) \mathbb{P}(\{\omega\}) \geq \sum_{\omega \in \Omega, X(\omega) \geq a} X(\omega) \mathbb{P}(\{\omega\}) \geq a \times \sum_{\omega \in \Omega, X(\omega) \geq a} \mathbb{P}(\{\omega\}) = a\mathbb{P}(X \geq a)$$

Soit : $\forall a \in \mathbb{R}_+, a\mathbb{P}(X \geq a) \leq E(X)$ ☕

THÉORÈME 20.5 - (Inégalité de Bienaymé-Tchebychev (Чебышев)) Soit X une VAR. On a :

$$\forall \varepsilon > 0, \mathbb{P}(|X - E(X)| \geq \varepsilon) \leq \frac{V(X)}{\varepsilon^2}$$

PREUVE. Soient X une VAR, et ε un réel > 0 . On a : $[|X - E(X)| \geq \varepsilon] \iff [(X - E(X))^2 \geq \varepsilon^2]$.

En appliquant l'inégalité de Markov à la VAR (positive) $(X - E(X))^2$ (et en prenant $a = \varepsilon^2$), on obtient :

$$E((X - E(X))^2) \geq \varepsilon^2 \mathbb{P}((X - E(X))^2 \geq \varepsilon^2)$$

$$\text{D'où : } \mathbb{P}((X - E(X))^2 \geq \varepsilon^2) = \mathbb{P}(|X - E(X)| \geq \varepsilon) \leq \frac{E((X - E(X))^2)}{\varepsilon^2} \quad (\spadesuit)$$

Il reste à voir que : $E((X - E(X))^2) = V(X)$ (

En effet :

$$E((X - E(X))^2) = E(X^2 - 2XE(X) + E(X)^2) = E(X^2) - 2E(X)^2 + \underbrace{E(E(X))}_{=E(X)} = E(X^2) - E(X)^2$$

Ce qui justifie (

Conclusion. D'après () et () : $\forall \varepsilon > 0, \mathbb{P}(|X - E(X)| \geq \varepsilon) \leq \frac{V(X)}{\varepsilon^2}$



Remarques. Ce dernier résultat s'applique en général à toute VAR X **admettant une variance**. Dans le cadre dans lequel nous avons travaillé (celui d'un univers fini), cette condition est automatiquement vérifiée. En dehors de ce cadre, il convient de rajouter l'hypothèse supplémentaire que $E(X^2)$ existe, c'est-à-dire que $\sum_{\omega \in \Omega} (X(\omega))^2 \mathbb{P}(\{\omega\})$ existe.

La seconde remarque consiste à prendre un peu de recul sur les deux derniers résultats énoncés. Finalement, l'inégalité de Марков ne “sert essentiellement qu'à” démontrer celle de Bienaymé-Чебышев.⁷ Cette dernière a une signification à la fois profonde et très intuitive : la probabilité qu'une variable aléatoire prenne des valeurs éloignées de sa moyenne est d'autant plus faible que cet éloignement est grand.

⁷. Ce qui n'est pas rien!!!

20.6 Variables aléatoires réelles : lois usuelles

Remarque préliminaire. Soit X une VAR sur un espace probabilisé $(\Omega, \mathbb{P})$.

Déterminer la **loi (de probabilité)** de X , c'est déterminer :

- $X(\Omega)$: l'ensemble des valeurs prises par la variable aléatoire X
- et pour chaque $x \in X(\Omega)$, la probabilité $\mathbb{P}(X = x)$

Dans le cas où $X(\Omega)$ est fini (càd encore une fois le seul cas étudié cette année), on peut noter comme d'habitude $X(\Omega) = \{x_1, \dots, x_n\}$. Déterminer la loi de X revient donc essentiellement à remplir un tableau comme ci-dessous :

x_i	x_1	x_2	$\dots$	x_n
$p_i = \mathbb{P}(X = x_i)$	p_1	p_2	$\dots$	p_n

En pratique, on ne présentera néanmoins la loi d'une VAR sous cette forme que lorsque le nombre de valeurs prises par X est connu et "petit". Dans les autres cas, on explicitera la loi en décrivant donc $X(\Omega)$, et en précisant :

$$\forall x \in X(\Omega), \quad \mathbb{P}(X = x) = \dots$$

20.6.1 Loi uniforme

DÉFINITION 20.11 - Soient m et n deux entiers naturels, avec $n < m$.

Une VAR X suit la **loi uniforme sur** $\llbracket n, m \rrbracket$ si X peut prendre toutes les valeurs de $\llbracket n, m \rrbracket$ avec la probabilité $1/(m - n + 1)$.

Pour insister sur ce que l'on remarqué en préliminaire, une VAR X définie sur un univers Ω suit la loi uniforme sur $\llbracket n, m \rrbracket$ lorsque :

$$\text{➤ } X(\Omega) = \llbracket n, m \rrbracket$$

$$\text{➤ } \forall k \in \llbracket n, m \rrbracket, \quad P(X = k) = \frac{1}{m - n + 1}$$

Exemple. Lors d'un lancer de dé bien équilibré, la VAR X égale au numéro de la face obtenue suit une loi uniforme sur $\llbracket 1, 6 \rrbracket$.

PROPRIÉTÉ 20.11 - Si X suit la loi uniforme sur $\llbracket n, m \rrbracket$, alors :

$$E(X) = \frac{n+m}{2}$$

PREUVE. Sous les hypothèses de l'énoncé, on a :

$$E(X) = \sum_{k=n}^m k \times \frac{1}{m-n+1} = \frac{1}{m-n+1} \sum_{k=n}^m k = \frac{1}{m-n+1} \times (m-n+1) \times \frac{n+m}{2}$$

D'où : $E(X) = \frac{n+m}{2}$ ☕

20.6.2 Loi de Bernoulli

DÉFINITION 20.12 - Une VAR X suit une **loi de Bernoulli de paramètre p** si X peut prendre les valeurs 1 et 0 avec les probabilités respectives p et $1-p$.

Pour insister une nouvelle fois sur ce que l'on remarqué en préliminaire, une VAR X définie sur un univers Ω suit la loi de Bernoulli de paramètre p lorsque :

➤ $X(\Omega) = \{0, 1\}$

➤ $\mathbb{P}(X=0) = 1-p$ et $\mathbb{P}(X=1) = p$

x_i	0	1
p_i	$1-p$	p

Exemple. Lors d'un lancer d'une pièce de monnaie truquée de telle sorte que l'on obtienne pile avec une probabilité égale à p , on note X la VAR qui vaut 1 si l'obtient pile (succès) et face sinon (échec). La VAR X suit alors la loi de Bernoulli de paramètre p .

PROPRIÉTÉ 20.12 - Si X suit la loi de Bernoulli de paramètre p , alors :

$$E(X) = p \quad \text{et} \quad V(X) = p(1-p)$$

PREUVE. Sous les hypothèses de l'énoncé : $E(X) = p \times 1 + (1-p) \times 0$ d'où : $E(X) = p$.

Par ailleurs : $V(X) = (1-p)^2 \times p + (-p)^2 \times (1-p) = p - 2p^2 + p^3 + p^2 - p^3 = p - p^2 = p(1-p)$. ☕

20.6.3 Loi binomiale

DÉFINITION 20.13 - Soient n un entier naturel et p un réel de $[0; 1]$.

Une VAR X suit la **loi binomiale de taille n et de paramètre p** (X suit la loi $\mathcal{B}(n, p)$) si X est la somme de n VAR indépendantes suivant toutes la même loi de Bernoulli de paramètre p .

Il revient alors au même de dire que X **compte** le nombre de succès au cours de n expériences **indépendantes** ne pouvant conduire qu'à **deux issues** (on dit aussi n **épreuves de Bernoulli**) : "succès" avec la probabilité p , ou "échec" avec la probabilité $(1-p)$.

Si X suit la loi binomiale $\mathcal{B}(n, p)$, alors par définition : $X(\Omega) = \llbracket 0; n \rrbracket$ (puisque le nombre de succès au cours de n épreuves de Bernoulli indépendantes est un entier compris entre 0 et n). Pour décrire la loi de X , il reste à préciser les probabilités $\mathbb{P}(X = k)$ pour tout entier k compris entre 0 et n . C'est l'objet de l'énoncé ci-dessous.

THÉORÈME 20.6 - Si X suit la loi $\mathcal{B}(n, p)$ alors

$$\forall k \in \llbracket 0; n \rrbracket, \quad \mathbb{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

Observation rassurante. Dans ce contexte, on obtient $\sum_{k=0}^n \mathbb{P}(X = k) = 1$ comme conséquence du binôme de Newton (vérifiez que vous savez le justifier!).

PREUVE. Supposons que X suit la loi binomiale de taille n et de paramètre p . Alors X est la somme de n VAR indépendantes suivant toutes la même loi de Bernoulli de paramètre p .

Soit k un entier compris entre 0 et n . La probabilité que l'on obtienne exactement k succès suivis de $(n-k)$ échecs est, puisque les expériences sont supposés indépendantes : $p^k (1-p)^{n-k}$.

Mais cette configuration (k succès, puis $(n-k)$ échecs) n'est pas la seule pour obtenir $X = k$. Il s'agit de choisir les positions des succès (ou des échecs) pour obtenir le nombre de configurations conduisant à ce que l'évènement $X = k$ soit réalisé : il y a exactement $\binom{n}{k}$ (ou ce qui revient au même⁸ $\binom{n}{n-k}$) de le faire.

Il s'ensuit que : $\mathbb{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$. ☕

En résumé :

Si X suit la loi $\mathcal{B}(n, p)$ alors

$$\triangleright X(\Omega) = \llbracket 0, n \rrbracket$$

$$\triangleright \forall k \in \llbracket 0; n \rrbracket, \quad \mathbb{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

THÉORÈME 20.7 - (**Espérance et variance de la loi binomiale**) Si X suit la loi binomiale de taille n et de paramètre p , alors :

$$E(X) = np \quad \text{et} \quad V(X) = np(1-p) = npq \quad (\text{avec } q = 1-p).$$

PREUVE. $\triangleright$ On a : $E(X) = \sum_{k=0}^n k \binom{n}{k} p^k (1-p)^{n-k}$, d'où : $E(X) = \sum_{k=1}^n k \binom{n}{k} p^k (1-p)^{n-k}$.

Or pour tout $k \in \mathbb{N}^*$: $k \binom{n}{k} = k \frac{n!}{k! (n-k)!} = \frac{n!}{(k-1)! (n-k)!} = n \frac{(n-1)!}{(k-1)! (n-k)!} = n \binom{n-1}{k-1}$

8. D'après la propriété de symétrie des coefficients binomiaux.

$$\begin{aligned} \text{Donc : } E(X) &= \sum_{k=1}^n n \binom{n-1}{k-1} p^k (1-p)^{n-k} = n \sum_{k=1}^n \binom{n-1}{k-1} p^k (1-p)^{n-k} = n \sum_{k=0}^{n-1} \binom{n-1}{k} p^{k+1} (1-p)^{n-1-k} \\ &= np \sum_{k=0}^{n-1} \binom{n-1}{k} p^k (1-p)^{n-1-k} = np \underbrace{(p + (1-p))^{n-1}}_{=1} \text{ donc } \boxed{E(X) = np} \end{aligned}$$

$$\blacktriangleright \text{ On a }^9 : E(X(X-1)) = \sum_{k=0}^n k(k-1) \binom{n}{k} p^k (1-p)^{n-k} = \sum_{k=2}^n k(k-1) \binom{n}{k} p^k (1-p)^{n-k}.$$

Or $\forall k \in \mathbb{N} \setminus \{0, 1\}$ on a :

$$k(k-1) \binom{n}{k} = k(k-1) \frac{n!}{k!(n-k)!} = \frac{n!}{(k-2)!(n-k)!} = n(n-1) \frac{(n-2)!}{(k-2)!(n-k)!} = n(n-1) \binom{n-2}{k-2}$$

Donc :

$$\begin{aligned} E(X(X-1)) &= \sum_{k=2}^n n(n-1) \binom{n-2}{k-2} p^k (1-p)^{n-k} = n(n-1) \sum_{k=1}^n \binom{n-2}{k-2} p^k (1-p)^{n-k} = \\ &= n(n-1) \sum_{k=0}^{n-2} \binom{n-2}{k} p^{k+2} (1-p)^{n-2-k} = n(n-1) p^2 \sum_{k=0}^{n-2} \binom{n-2}{k} p^k (1-p)^{n-2-k} \end{aligned}$$

D'où $\boxed{E(X(X-1)) = n(n-1)p^2}$. Or : $E(X(X-1)) = E(X^2) - E(X)$.

Par suite : $E(X^2) = n(n-1)p^2 + np$, donc : $V(X) = n(n-1)p^2 + np - (np)^2 = np - np^2$ soit :

$$\boxed{V(X) = np(1-p)}.$$



Exemples de question amusantes sur la loi binomiale.

EXERCICE 20.1 - Une variable aléatoire réelle X suit la loi binomiale de taille n et de paramètre p . Quelle est la loi suivie par la variable aléatoire $Y = n - X$?

EXERCICE 20.2 - Une variable aléatoire réelle X suit la loi binomiale de taille n et de paramètre $p \in]0; 1[$. Pour quelle valeur de k , la probabilité $p_k = P(X = k)$ est-elle maximale ?

EXERCICE 20.3 - Soit X une variable aléatoire réelle suivant la loi binomiale de taille n et de paramètre $p \in]0; 1[$. On note $b(k, n, p) = P(X = k)$.

1) Pour quelle valeur m de k , le coefficient $b(k, n, p)$ est-il maximal ?

2) Etudier la monotonie de la fonction $f : x \mapsto x^m(1-x)^{n-m}$ sur $[0; 1]$.

3) Vérifier que si $m \in [np; (n+1)p]$ alors : $b\left(m, n, \frac{m}{n+1}\right) \leq b(m, n, p) \leq b\left(m, n, \frac{m}{n}\right)$

4) Proposer un encadrement analogue pour $m \in [(n+1)p - 1; np]$

9. D'après la formule de transfert.

EXERCICE 20.4 - Un étudiant résout un QCM constitué de n questions offrant chacune quatre réponses possibles. Pour chaque question, et indépendamment les unes des autres, il a la probabilité p de savoir résoudre celle-ci. Dans ce cas il produit la bonne réponse. Si en revanche, il ne sait pas résoudre la question, il choisit arbitrairement l'une des quatre réponses possibles. On note X la variable aléatoire déterminant le nombre de questions qu'il savait résoudre et Y le nombre de questions qu'il a correctement résolues parmi celles où il a répondu "au hasard".

- 1) Montrer que pour j et k deux entiers de $\llbracket 0, n \rrbracket$ on a :
$$\binom{n-j}{k-j} \binom{n}{j} = \binom{k}{j} \binom{n}{k}$$
- 2) Décrire la loi de $Z = X + Y$.
- 3) Calculer espérance et variance de Z .

EXERCICE 20.5 - Un archer tire sur n cibles. A chaque tir, il a la probabilité p de toucher la cible et les tirs sont supposés indépendants. Il tire une première fois sur chaque cible et on note X le nombre de cibles atteintes lors de ce premier jet. L'archer tire ensuite une seconde fois sur les cibles restantes et l'on note Y le nombre de cibles touchées lors de cette tentative. Déterminer la loi de la variable $Z = X + Y$.

20.7 Epilogue

20.7.1 Problème — Sur la loi géométrique et les fonctions génératrices

Présentation du problème : cet énoncé est une compilation de questions extraites de divers concours sur les probabilités. Si cet énoncé a un objectif principal (le calcul de la probabilité de l'avant-dernière question du problème¹⁰), d'autres thèmes seront abordés avant d'y parvenir.

*Explicitement, la courte introduction (**partie I**) est constituée d'une question de cours, et d'une seconde question qui ne sera réutilisée que pour la question 17-d de la partie V.*

*La **partie II** est l'étude d'une nouvelle loi de probabilité (ici, nouvelle = non étudiée en cours); elle est complètement indépendante du reste du problème, et peut être considérée comme un exercice à part.*

*Les **parties III et IV** sont fortement liées; mais les résultats de la partie III utiles pour la partie IV sont indiqués dans l'énoncé.*

*Pour conclure, la **partie V** comporte une partie de pur calcul matriciel, et s'achève en apothéose par le calcul d'amusantes probabilités que je vous laisse découvrir, histoire tout de même de laisser un peu de place au suspense.*

Partie I — Une formule générale

On se place dans le cadre général d'un espace probabilisé $(\Omega, \mathbb{P})$.

- 1) Soient A et B deux évènements quelconques de Ω . Rappeler la formule donnant $\mathbb{P}(A \cup B)$.
- 2) Soient A , B et C trois évènements quelconques de Ω . Montrer que :

$$\mathbb{P}(A \cup B \cup C) = \mathbb{P}(A) + \mathbb{P}(B) + \mathbb{P}(C) - \mathbb{P}(A \cap B) - \mathbb{P}(A \cap C) - \mathbb{P}(B \cap C) + \mathbb{P}(A \cap B \cap C)$$

Partie II — Loi géométrique

Soit p un réel tel que : $0 < p < 1$.

Un joueur répète de manière successive et indépendante une expérience au cours de laquelle il peut connaître un succès (avec la probabilité p) ou un échec (avec la probabilité $q = 1 - p$).

Le jeu s'arrête au premier succès.

On note X la variable aléatoire égale au rang du premier succès.

Au cours de cette partie II, on pourra au besoin noter S_k l'évènement "le joueur a obtenu un succès au cours de la k -ième expérience", et E_k l'évènement $\overline{S}_k$ (pour tout entier naturel non nul k).

- 3) Quelles sont les valeurs prises par la variable aléatoire X ?
- 4) Justifier soigneusement que $\mathbb{P}(X = 2) = qp$. Que vaut $\mathbb{P}(X = 3)$?
- 5) Pour tout entier naturel non nul k , calculer $\mathbb{P}(X = k)$.
- 6) Pour tout entier naturel non nul N , calculer $\mathbb{P}(X \leq N)$.

¹⁰. Attendez un peu avant de la regarder.

7) Etablir que : $\lim_{N \rightarrow +\infty} \left[\sum_{k=1}^N \mathbb{P}(X = k) \right] = 1.$

Remarque : ce résultat est rassurant, dans le sens où il signifie que la formule obtenue dans la question 5 définit effectivement une loi de probabilité. Il est alors raisonnable de se demander quelle est la valeur moyenne de X ; c'est l'objet de la question suivante.

8) **Calcul de l'espérance.**

a) Pour tout entier naturel N non nul, calculer $\sum_{k=1}^N k q^{k-1}.$

b) En déduire que l'espérance de X est égale à $1/p$, c'est-à-dire établir que¹¹ :

$$E(X) = \lim_{N \rightarrow +\infty} \left[\sum_{k=1}^N k \mathbb{P}(X = k) \right] = \frac{1}{p}$$

9) **Calcul de la variance.**

a) Pour tout entier naturel N non nul, calculer $\sum_{k=1}^N k(k-1) q^{k-2}.$

b) A l'aide de la question précédente, montrer que :

$$E(X(X-1)) = \lim_{N \rightarrow +\infty} \left[\sum_{k=1}^N k(k-1) \mathbb{P}(X = k) \right] = \frac{2q}{p^2}$$

c) Déterminer la variance de X .

Partie III — Dérivées successives de fonctions polynomiales

Dans toute cette partie, on note Q_k (pour tout entier naturel k) la fonction polynomiale définie sur $\mathbb{R}$ en posant :

$$\forall x \in \mathbb{R}, \quad Q_k(x) = x^k$$

10) Soit k un entier naturel quelconque. Montrer que :
$$\begin{cases} \forall i \in [0, k], & Q_k^{(i)} = \frac{k!}{(k-i)!} Q_{k-i} \\ \forall i > k, & Q_k^{(i)} = 0 \end{cases}$$

11) Dédurre de la question précédente que : $\forall (k, i) \in \mathbb{N}^2, Q_k^{(i)}(0) = \delta_{ki} \times k!$

12) Soit n un entier naturel et P la fonction polynomiale définie sur $\mathbb{R}$ par

$$\forall x \in \mathbb{R}, \quad P(x) = \sum_{k=0}^n a_k x^k \quad ("P(x) = a_n x^n + \dots + a_1 x + a_0")$$

où $(a_0, \dots, a_n)$ est un $(n+1)$ -uplet de réels quelconques.

Montrer que : $\boxed{\forall k \in [0, n], P^{(k)}(0) = k! a_k}$

11. Pour qu'il n'y ait pas de confusion possible, il s'agit ici d'établir "seulement" la seconde égalité, c'est-à-dire que la limite de la somme vaut effectivement $1/p$. La même remarque s'appliquera à la question 9-b.

Partie IV — Fonctions génératrices

Soit $n \in \mathbb{N}$, et soit X une variable aléatoire à valeurs dans $\llbracket 0, n \rrbracket$. On appelle **fonction génératrice de X** la fonction G_X définie par

$$\forall s \in \mathbb{R}, G_X(s) = E(s^X)$$

13) Deux exemples pour commencer.

- a) On suppose dans cette question que X est une variable aléatoire suivant la loi uniforme sur $\llbracket 0, n \rrbracket$. Vérifier que :

$$\forall s \in \mathbb{R} \setminus \{1\}, G_X(s) = \frac{1}{n+1} \times \frac{1-s^{n+1}}{1-s}$$

- b) On suppose dans cette question que X suit la loi binomiale de taille n et de paramètre p (avec $p \in [0, 1]$). Donner l'expression de $G_X(s)$ (le résultat final ne devra plus comporter de “ $\sum$ ”).

14) On revient au cas général, où X est une variable aléatoire quelconque à valeurs dans $\llbracket 0, n \rrbracket$.

- a) Justifier que G_X est une fonction polynomiale. Calculer $G_X(1)$.
- b) Exprimer les probabilités $P(X = k)$ en fonction des dérivées successives de G_X en 0 pour tout entier $k \in \llbracket 0, n \rrbracket$ (on pourra utiliser la question 12).
- c) En déduire que deux variables aléatoires X et Y à valeurs dans $\llbracket 0, n \rrbracket$ ont même loi si et seulement si $G_X = G_Y$.
- d) Exprimer l'espérance et la variance de X en fonction de $G'_X(1)$ et $G''_X(1)$.

Partie V — Calcul matriciel et probabilités

On considère la matrice

$$M = \begin{pmatrix} X & X & X \\ Y & Y & Y \\ Z & Z & Z \end{pmatrix}$$

où X, Y et Z désignent trois réels quelconques.

15) La matrice M est-elle inversible ?

16) a) Calculer M^2 .

- b) Une matrice carrée N est dite **idempotente** lorsqu'elle vérifie $N^2 = N$.

Démontrer que M est une matrice idempotente si et seulement si $X = Y = Z = 0$ ou $S = 1$.

17) On suppose à présent que X, Y et Z désignent trois variables aléatoires réelles indépendantes, suivant la même loi binomiale $\mathcal{B}(n, p)$ (n désignant un entier naturel, et p un réel dans $[0; 1]$).

- a) Donner sans calcul l'espérance et la variance de X .

b) Montrer que $X + Y$ suit la loi binomiale $\mathcal{B}(2n, p)$.

On pourra à cette fin utiliser la **formule de Vandermonde** :

$$\sum_{k=0}^q \binom{n}{k} \binom{m}{q-k} = \binom{n+m}{q} \quad \text{si } 0 \leq q \leq n+m$$

Par ailleurs, dans la suite du problème, on pourra admettre que $X + Y + Z$ suit la loi binomiale $\mathcal{B}(3n, p)$.

c) Quelle est la probabilité pour que la matrice M soit idempotente ?

d) On suppose que $p = \frac{1}{2}$. Quelle est la probabilité pour qu'au moins une ligne de M soit égale à la somme des deux autres ?

CORRIGÉ.

Partie I — Une formule générale

1) Pur cours¹² : $\boxed{\forall (A, B) \in \mathcal{P}(\Omega)^2, \mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)}$.

2) Soient A, B et C trois évènements quelconques de Ω . On utilise la question précédente en écrivant par exemple :

$$\mathbb{P}(A \cup B \cup C) = \mathbb{P}((A \cup B) \cup C) = \mathbb{P}(A \cup B) + \mathbb{P}(C) - \mathbb{P}((A \cup B) \cap C)$$

$$\text{D'où : } \mathbb{P}(A \cup B \cup C) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B) + \mathbb{P}(C) - \mathbb{P}((A \cup B) \cap C) \quad (\spadesuit)$$

Reste à calculer $\mathbb{P}((A \cup B) \cap C)$. En avant :

$$\mathbb{P}((A \cup B) \cap C) = \mathbb{P}((A \cap C) \cup (B \cap C)) = \mathbb{P}(A \cap C) + \mathbb{P}(B \cap C) - \mathbb{P}((A \cap C) \cap (B \cap C))$$

$$\text{d'où : } \mathbb{P}((A \cup B) \cap C) = \mathbb{P}(A \cap C) + \mathbb{P}(B \cap C) - \mathbb{P}(A \cap B \cap C) \quad (\clubsuit)$$

Grâce à $(\spadesuit)$ et $(\clubsuit)$:

$$\boxed{\mathbb{P}(A \cup B \cup C) = \mathbb{P}(A) + \mathbb{P}(B) + \mathbb{P}(C) - \mathbb{P}(A \cap B) - \mathbb{P}(A \cap C) - \mathbb{P}(B \cap C) + \mathbb{P}(A \cap B \cap C)}$$

Partie II — Loi géométrique

3) $\boxed{X(\Omega) = \mathbb{N}^*}$.

4) $X = 2$ si et seulement si le joueur a connu un échec lors de la première expérience, et un succès lors de la seconde. En d'autres termes, $X = 2$ si et seulement si l'évènement $E_1 \cap S_2$ est réalisé. Or, puisque les expériences sont supposées indépendantes, on peut affirmer que : $\mathbb{P}(E_1 \cap S_2) = \mathbb{P}(E_1) \times \mathbb{P}(S_2)$, c'est-à-dire : $\boxed{\mathbb{P}(X = 2) = qp}$.

Sur le même modèle : $\boxed{\mathbb{P}(X = 3) = q^2p}$.

12. De Seconde.

- 5) On généralise le raisonnement de la question précédente. Soit k un entier naturel non nul. $X = k$ si et seulement si le joueur a connu un échec lors des $(k - 1)$ premières expériences, et un succès lors de la k -ième. En d'autres termes, $X = k$ si et seulement si l'évènement $\left(\bigcap_{i=1}^{k-1} E_i\right) \cap S_k$ est réalisé. Once again, puisque les expériences sont supposées indépendantes, on peut affirmer que :

$$\mathbb{P}\left(\left(\bigcap_{i=1}^{k-1} E_i\right) \cap S_k\right) = \mathbb{P}\left(\bigcap_{i=1}^{k-1} E_i\right) \times \mathbb{P}(S_k) = \left[\prod_{i=1}^{k-1} \mathbb{P}(E_i)\right] \times \mathbb{P}(S_k), \text{ c'est-à-dire : } \boxed{\mathbb{P}(X = k) = q^{k-1}p}.$$

- 6) Soit N un entier naturel non nul N . On a :

$$\mathbb{P}(X \leq N) = \sum_{k=1}^N \mathbb{P}(X = k) = \sum_{k=1}^N q^{k-1}p = p \sum_{k=1}^N q^{k-1} = p \sum_{k=0}^{N-1} q^k = p \frac{1 - q^N}{1 - q} \quad 13$$

De plus : $p = 1 - q$. Finalement : $\boxed{\mathbb{P}(X \leq N) = 1 - q^N}$

- 7) Puisque d'après l'énoncé $0 < q < 1$ on a : $\lim_{N \rightarrow +\infty} q^N = 0$. On déduit de cette remarquable observation

et de la question précédente que : $\boxed{\lim_{N \rightarrow +\infty} \left[\sum_{k=1}^N \mathbb{P}(X = k) \right] = 1}$

8) Calcul de l'espérance.

- a) Soit N un entier naturel non nul. Posons, pour tout $q \in]0; 1[$, $\varphi(q) = \sum_{k=1}^N q^k$. La fonction φ est dérivable sur $]0; 1[$ (elle est polynomiale), et par linéarité de la dérivation : $\forall q \in]0; 1[$, $\varphi'(q) = \sum_{k=1}^N kq^{k-1}$ (

► D'autre part, on a : $\forall q \in]0; 1[$, $\varphi(q) = q \frac{1 - q^N}{1 - q} = \frac{q - q^{N+1}}{1 - q}$.

Par suite, pour tout réel $q \in]0; 1[$, on a : $\varphi'(q) = \frac{(1 - (N + 1)q^N)(1 - q) + q - q^{N+1}}{(1 - q)^2}$

D'où : $\forall q \in]0; 1[$, $\varphi'(q) = \frac{1 - (N + 1)q^N + Nq^{N+1}}{(1 - q)^2}$ (

On déduit de (\boxed{\forall N \in \mathbb{N}^*, \sum_{k=1}^N kq^{k-1} = \frac{1 - (N + 1)q^N + Nq^{N+1}}{(1 - q)^2}}

- b) Soit N un entier naturel non nul. On a : $\sum_{k=1}^N k\mathbb{P}(X = k) = \sum_{k=1}^N kq^{k-1}p = p \sum_{k=1}^N kq^{k-1}$.

D'où, d'après la question précédente : $\sum_{k=1}^N k\mathbb{P}(X = k) = p \times \frac{1 - (N + 1)q^N + Nq^{N+1}}{(1 - q)^2}$

13. Légitime car $q \neq 1$ d'après l'énoncé.

Puisque d'après l'énoncé $0 < q < 1$, et que $1 - q = p$, on obtient en passant à la limite lorsque N tend vers $+\infty$ dans la relation précédente :¹⁴

$$\lim_{N \rightarrow +\infty} \left[\sum_{k=1}^N k \mathbb{P}(X = k) \right] = p \times \frac{1}{p^2} \quad \text{soit :} \quad \boxed{\lim_{N \rightarrow +\infty} \left[\sum_{k=1}^N k \mathbb{P}(X = k) \right] = \frac{1}{p}}$$

9) Calcul de la variance.

a) Soit N un entier naturel non nul. En reprenant les notations de la question 8-a, la fonction φ est deux fois dérivable sur $]0; 1[$ (car elle est toujours polynomiale), et donc : $\forall q \in]0; 1[, \varphi''(q) =$

$$\sum_{k=2}^N k(k-1) q^{k-2} \quad (\spadesuit).$$

► D'autre part, on a : $\forall q \in]0; 1[, \varphi'(q) = \frac{1 - (N+1)q^N + Nq^{N+1}}{(1-q)^2}$.

Par suite, pour tout réel $q \in]0; 1[$, on a :

$$\varphi''(q) = \frac{(-N(N+1)q^{N-1} + N(N+1)q^N)(1-q)^2 + 2(1-q)(1 - (N+1)q^N + Nq^{N+1})}{(1-q)^4}$$

$$\iff \varphi''(q) = \frac{(-N(N+1)q^{N-1} + N(N+1)q^N)(1-q) + 2(1 - (N+1)q^N + Nq^{N+1})}{(1-q)^3}$$

$$\iff \varphi''(q) = \frac{2 - N(N+1)q^{N-1} - 2(N+1)q^N - (N^2 - N)q^{N+1}}{(1-q)^3} \quad (\clubsuit)$$

On déduit de $(\spadesuit)$ et $(\clubsuit)$ que :

$$\boxed{\forall N \in \mathbb{N}^*, \sum_{k=2}^N k(k-1)q^{k-2} = \frac{2 - N(N+1)q^{N-1} - 2(N+1)q^N - (N^2 - N)q^{N+1}}{(1-q)^3}}$$

b) Soit $N \in \mathbb{N}^*$. On a : $\sum_{k=1}^N k(k-1)\mathbb{P}(X = k) = \sum_{k=2}^N k(k-1)pq^{k-1} = pq \sum_{k=2}^N k(k-1)q^{k-2}$.

D'après de la question précédente (et puisque $1 - q = p$) :

$$\sum_{k=1}^N k(k-1)\mathbb{P}(X = k) = pq \times \frac{2 - N(N+1)q^{N-1} - 2(N+1)q^N - (N^2 - N)q^{N+1}}{p^3}$$

Subséquentement : $\lim_{N \rightarrow +\infty} \left[\sum_{k=1}^N k(k-1)\mathbb{P}(X = k) \right] = pq \frac{2}{p^3}$

soit : $\boxed{\lim_{N \rightarrow +\infty} \left[\sum_{k=1}^N k(k-1)\mathbb{P}(X = k) \right] = \frac{2q}{p^2}}$

c) D'après la formule de Koenig-Huygens : $V(X) = E(X^2) - E(X)^2 = E(X(X-1)) + E(X) - E(X)^2$.

D'après les deux questions précédentes, on a donc :

$$V(X) = \frac{2q}{p^2} + \frac{1}{p} - \frac{1}{p^2} = \frac{2q + p - 1}{p^2} \quad \text{d'où :} \quad \boxed{V(X) = \frac{q}{p^2}}$$

14. Il faut aussi noter que : $\lim_{N \rightarrow +\infty} Nq^N$ par croissances comparées.

Partie III — Dérivées successives de fonctions polynomiales

10) Démontrons la propriété par récurrence sur k , en notant

$$\mathcal{P}(k) \quad \text{la propriété :} \quad \begin{cases} \forall i \in \llbracket 0, k \rrbracket, & Q_k^{(i)} = \frac{k!}{(k-i)!} Q_{k-i} \\ \forall i > k, & Q_k^{(i)} = 0 \end{cases}$$

► Initialisation ($k=0$) : on a $Q^{(0)} = 1 = \frac{0!}{0!} Q_0$, et pour tout $i > 0$, $Q_0^{(i)} = 0$. La propriété est initialisée.

► Hérédité : supposons la propriété vraie pour un certain entier naturel k , et montrons qu'elle est vraie au rang $k+1$.

• Soit $i \in \llbracket 1, k \rrbracket$: on peut observer que $Q_{k+1} = Q_k Q_1$ et utiliser la formule de Leibniz pour obtenir :

$$Q_{k+1}^{(i)} = \sum_{j=0}^i \binom{i}{j} Q_1^{(j)} Q_k^{(i-j)} = Q_1 Q_k^{(i)} + i Q_k^{(i-1)}$$

Il s'ensuit que pour tout réel x :

$$Q_{k+1}^{(i)}(x) = x Q_k^{(i)}(x) + i Q_k^{(i-1)}(x)$$

D'où, par hypothèse de récurrence :

$$\begin{aligned} Q_{k+1}^{(i)}(x) &= \frac{k!}{(k-i)!} x^{k-i+1} + i \frac{k!}{(k-i+1)!} x^{k-i+1} = \frac{k! [(k-i+1) + i]}{(k-i+1)!} x^{k-i+1} \\ &= \frac{(k+1)!}{(k+1-i)!} x^{k+1-i} \end{aligned}$$

Autrement écrit : $\forall i \in \llbracket 0, k \rrbracket, \quad Q_{k+1}^{(i)} = \frac{(k+1)!}{(k+1-i)!} Q_{k+1-i}$

• Pour $i = k+1$: on peut observer que $Q_{k+1}^{(k+1)} = \left[Q_{k+1}^{(k)} \right]'$ et utiliser la question précédente pour obtenir :

$$Q_{k+1}^{(k+1)} = (k+1)! Q_1' \quad \text{d'où :} \quad Q_{k+1}^{(k+1)} = (k+1)! Q_0$$

A la lumière de ce résultat et du précédent, on a donc : $\forall i \in \llbracket 0, k \rrbracket, \quad Q_{k+1}^{(i)} = \frac{(k+1)!}{(k+1-i)!} Q_{k+1-i} \quad (\spadesuit)$

• Pour $i > k+1$: puisque la fonction $Q_{k+1}^{(k+1)}$ est constante, toutes ses dérivées sont nulles. On en déduit donc que : $\forall i > k+1, \quad Q_{k+1}^{(i)} = 0 \quad (\clubsuit)$

• En résumé : $\begin{cases} \forall i \in \llbracket 0, k+1 \rrbracket, & Q_{k+1}^{(i)} = \frac{(k+1)!}{(k+1-i)!} Q_{k+1-i} \\ \forall i > k+1, & Q_{k+1}^{(i)} = 0 \end{cases}$

Ce qui prouve que la propriété $\mathcal{P}(n+1)$ est vraie, établit l'hérédité, et achève cette récurrence.

<u>Conclusion</u> : $\forall k \in \mathbb{N},$ $\begin{cases} \forall i \in \llbracket 0, k \rrbracket, & Q_k^{(i)} = \frac{k!}{(k-i)!} Q_{k-i} \\ \forall i > k, & Q_k^{(i)} = 0 \end{cases}$

11) Soient k et i deux entiers naturels.

- Pour $i > k$: alors $Q_k^{(i)} = 0$ d'où en particulier : $Q_k^{(i)}(0) = 0$.
- Pour $i \leq k$: alors $Q_k^{(i)} = \frac{k!}{(k-i)!} Q_{k-i}$ et en particulier $Q_k^{(i)}(0) = \frac{k!}{(k-i)!} 0^{k-i}$.

Or : $0^{k-i} = 0$ dès que $k \neq i$, et $0^{k-i} = 1$ pour $i = k$.

- En résumé : $\forall (k, i) \in \mathbb{N}^2, Q_k^{(i)}(0) = \begin{cases} k! & \text{si } i = k \\ 0 & \text{sinon} \end{cases}$

On en déduit que : $\forall (k, i) \in \mathbb{N}^2, Q_k^{(i)}(0) = \delta_{ki} \times k!$

12) Considérons P la fonction polynomiale définie sur $\mathbb{R}$ par $\forall x \in \mathbb{R}, P(x) = \sum_{m=0}^n a_m x^m$. Il revient au

même d'écrire : $\forall x \in \mathbb{R}, P(x) = \sum_{m=0}^n a_m Q_m$. La fonction P est de classe $\mathcal{C}^\infty$ sur $\mathbb{R}$ (car polynomiale) et :

$$\forall k \in \llbracket 0, n \rrbracket, P^{(k)}(0) = \sum_{m=0}^n a_m Q_m^{(k)}(0) = \sum_{m=0}^n a_m \delta_{km} m!; \quad \text{d'où : } \boxed{\forall k \in \llbracket 0, n \rrbracket, P^{(k)}(0) = (k!) a_k}.$$

Partie IV — Fonctions génératrices

13) Deux exemples pour commencer.

a) Dans le cas où X suit la loi uniforme $\mathcal{U}(\llbracket 0, n \rrbracket)$, on a : $\forall k \in \llbracket 0, n \rrbracket, \mathbb{P}(X = k) = \frac{1}{n+1}$.

<u>Conclusion</u> : si X suit une loi uniforme $\mathcal{U}(\llbracket 0, n \rrbracket)$, alors : $\forall s \in \mathbb{R} \setminus \{1\}, G_X(s) =$ $\frac{1}{n+1} \sum_{k=0}^n s^k$ Ce que l'on peut réécrire : $\forall s \in \mathbb{R} \setminus \{1\}, G_X(s) = \frac{1}{1-s^{n+1}} \times \frac{1-s^{n+1}}{1-s}$

b) Si X suit une loi binomiale de taille n et de paramètre p , alors X est à valeurs dans $\llbracket 0, n \rrbracket$ et :

$$\forall k \in \llbracket 0, n \rrbracket, \mathbb{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

Par suite, pour tout réel s on a :

$$G_X(s) = \sum_{k=0}^n s^k \binom{n}{k} p^k (1-p)^{n-k} = \sum_{k=0}^n \binom{n}{k} (ps)^k (1-p)^{n-k} \text{ d'où :}$$

$\forall s \in \mathbb{R}, G_X(s) = [ps + (1-p)]^n$

14) a) Par définition de l'espérance : $\forall s \in \mathbb{R}, G_X(s) = \sum_{k=0}^n \mathbb{P}(X = k) s^k$ donc G_X est polynomiale en s .

On a : $G_X(s) = \sum_{k=0}^n \mathbb{P}(X = k) 1^k = \sum_{k=0}^n \mathbb{P}(X = k)$ d'où : $G_X(1) = 1$.

b) Pour tout polynôme Q de degré inférieur ou égal à n : $Q = \sum_{k=0}^n \frac{Q^{(k)}(0)}{k!} X^k$.

Lorsque G_X désigne la fonction génératrice d'une variable aléatoire X à valeurs dans $\llbracket 0; n \rrbracket$, on a d'une part $G_X = \sum_{k=0}^n \mathbb{P}(X = k) s^k$ et d'autre part $G_X = \sum_{k=0}^n \frac{G_X^{(k)}(0)}{k!} s^k$. Par identification on en déduit que :

$$\forall k \in \llbracket 0; n \rrbracket, \quad \mathbb{P}(X = k) = \frac{G_X^{(k)}(0)}{k!}$$

c) Soient X et Y deux variables aléatoires à valeurs dans $\llbracket 0; n \rrbracket$. Si $G_X = G_Y$, alors :

$\forall k \in \llbracket 0; n \rrbracket, G_X^{(k)}(0) = G_Y^{(k)}(0)$ d'où (grâce à 2-a) : $\forall k \in \llbracket 0; n \rrbracket, \mathbb{P}(X = k) = \mathbb{P}(Y = k)$.

Donc X et Y ont même loi. Réciproquement, supposons que X et Y ont même loi, alors :

$\forall k \in \llbracket 0; n \rrbracket, \mathbb{P}(X = k) = \mathbb{P}(Y = k)$, d'où $\forall k \in \llbracket 0; n \rrbracket, G_X^{(k)}(0) = G_Y^{(k)}(0)$.

Ce qui entraîne que les polynômes G_X et G_Y sont égaux, puisqu'ils ont les mêmes coefficients.

Conclusion : X et Y ont même loi si et seulement si $G_X = G_Y$.

d) Par définition de la fonction génératrice de X : $\forall s \in \mathbb{R}, G_X(s) = \sum_{k=0}^n \mathbb{P}(X = k) s^k$. La fonction G_X est dérivable (car polynomiale donc de classe $\mathcal{C}^\infty$) sur $\mathbb{R}$ et : $\forall s \in \mathbb{R}, G'_X(s) = \sum_{k=1}^n k \mathbb{P}(X = k) s^{k-1}$.

En particulier : $G'_X(1) = \sum_{k=1}^n k \mathbb{P}(X = k)$. Conclusion : $G'_X(1) = E(X)$.

En continuant sur notre lancée :

$\forall s \in \mathbb{R}, G''_X(s) = \sum_{k=2}^n k(k-1) \mathbb{P}(X = k) s^{k-2}$ d'où : $G''_X(1) = \sum_{k=2}^n k(k-1) \mathbb{P}(X = k)$

En développant :

$$G''_X(1) = \left[\sum_{k=2}^n k^2 \mathbb{P}(X = k) \right] - \left[\sum_{k=2}^n k \mathbb{P}(X = k) \right]$$

$$\iff G''_X(1) = \left[\sum_{k=1}^n k^2 \mathbb{P}(X = k) - \mathbb{P}(X = 1) \right] - \left[\sum_{k=1}^n k \mathbb{P}(X = k) - \mathbb{P}(X = 1) \right]$$

$$\Longleftrightarrow G''_X(1) = \left[\sum_{k=1}^n k^2 \mathbb{P}(X=k) \right] - \left[\sum_{k=1}^n k \mathbb{P}(X=k) \right]$$

$$\Longleftrightarrow G''_X(1) = E(X^2) - E(X)$$

Or la variance $V(X)$ de X est donnée par : $V(X) = E(X^2) - [E(X)]^2$.

D'où : $V(X) = G''_X(1) + G'_X(1) - (G'_X(1))^2$.

Partie V — Calcul matriciel et probabilités

15) Soient X , Y et Z trois réels quelconques. La matrice M est inversible si et seulement si

$$\forall \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} \in \mathbb{R}^3, \quad \exists! \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{R}^3, \quad M \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$$

$$\text{Or : } M \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} \Longleftrightarrow \begin{cases} Xx_1 + Xx_2 + Xx_3 = b_1 & (L_1) \\ Yx_1 + Yx_2 + Yx_3 = b_2 & (L_2) \\ Zx_1 + Zx_2 + Zx_3 = b_3 & (L_3) \end{cases}$$

$$\Longleftrightarrow \begin{cases} Xx_1 + Xx_2 + Xx_3 = b_1 & (L_1) \\ Yx_1 + Yx_2 + Yx_3 = b_2 & (L_2) \\ 0 = Yb_2 - Zb_3 & Y(L_2) - Z(L_3) \end{cases}$$

Il s'ensuit que le système ne peut admettre de solution si $Yb_2 - Zb_3 \neq 0$. Par suite : M n'est pas inversible.

$$16) \text{ a) } M^2 = \begin{pmatrix} X & X & X \\ Y & Y & Y \\ Z & Z & Z \end{pmatrix} \begin{pmatrix} X & X & X \\ Y & Y & Y \\ Z & Z & Z \end{pmatrix}$$

$$= \begin{pmatrix} X^2 + XY + XZ & X^2 + XY + XZ & X^2 + XY + XZ \\ XY + Y^2 + YZ & XY + Y^2 + YZ & XY + Y^2 + YZ \\ XZ + YZ + Z^2 & XZ + YZ + Z^2 & XZ + YZ + Z^2 \end{pmatrix}$$

$$\text{soit : } M^2 = \begin{pmatrix} X(X+Y+Z) & X(X+Y+Z) & X(X+Y+Z) \\ Y(X+Y+Z) & Y(X+Y+Z) & Y(X+Y+Z) \\ Z(X+Y+Z) & Z(X+Y+Z) & Z(X+Y+Z) \end{pmatrix} \text{ ou encore : } M^2 = (X+Y+Z)M$$

b) D'après la question précédente, $M^2 = M$ si et seulement si $M = 0_{M_3(\mathbb{R})}$ ou $X+Y+Z = 1$. Or $M = 0_{M_3(\mathbb{R})}$ si et seulement si $X=Y=Z=0$.

Par suite :

$$M \text{ est idempotente si et seulement si } X=Y=Z=0 \text{ ou } S=1 \text{ (en ayant noté } S=X+Y+Z \text{)}.$$

$$17) \text{ a) Si } X \text{ suit la loi binomiale } \mathcal{B}(n, p), \text{ alors } E(X) = np \text{ et } V(X) = np(1-p).$$

b) Puisque X et Y sont à valeurs dans $\llbracket 0, n \rrbracket$, la variable aléatoire $X + Y$ est à valeurs dans $\llbracket 0, 2n \rrbracket$. Soit donc k un entier naturel compris entre 0 et $2n$. On a :

$$\mathbb{P}(X + Y = k) = \sum_{j=0}^k \mathbb{P}([X = j] \cap [Y = k - j])$$

Puisque X et Y sont indépendantes : $\mathbb{P}(X + Y = k) = \sum_{j=0}^k \mathbb{P}(X = j) \times \mathbb{P}(Y = k - j)$

$$\text{D'où : } \mathbb{P}(X + Y = k) = \sum_{j=0}^k \binom{n}{j} p^j (1-p)^{n-j} \binom{n}{k-j} p^{k-j} (1-p)^{n-k+j}$$

$$\text{D'où : } \mathbb{P}(X + Y = k) = \sum_{j=0}^k \binom{n}{j} \binom{n}{k-j} p^k (1-p)^{2n-k}$$

$$\text{D'où : } \mathbb{P}(X + Y = k) = p^k (1-p)^{2n-k} \sum_{j=0}^k \binom{n}{j} \binom{n}{k-j}$$

$$\text{D'après la formule de Vandermonde : } \mathbb{P}(X + Y = k) = \binom{2n}{k} p^k (1-p)^{2n-k}$$

$$\text{En résumé : } \forall k \in \llbracket 0, 2n \rrbracket, \quad \mathbb{P}(X + Y = k) = \binom{2n}{k} p^k (1-p)^{2n-k}.$$

Conclusion : $X + Y$ suit la loi binomiale $\mathcal{B}(2n, p)$.

c) Notons $\mathbb{P}_1$ la probabilité que M soit idempotente. D'après la question 16-b, M est idempotente si et seulement si $X = Y = Z = 0$ ou $X + Y + Z = 1$.

Puisque ces deux événements sont disjoints, on a : $\mathbb{P}_1 = \mathbb{P}(X = Y = Z = 0) + \mathbb{P}(X + Y + Z = 1)$.

Les VAR X , Y et Z étant indépendantes, on a : $\mathbb{P}(X = Y = Z = 0) = \mathbb{P}(X = 0) \mathbb{P}(Y = 0) \mathbb{P}(Z = 0) = (1-p)^{3n}$.

Par ailleurs, d'après l'indication de l'énoncé ¹⁵, on a : $\mathbb{P}(X + Y + Z = 1) = 3np(1-p)^{3n-1}$.

Il s'ensuit que : $\mathbb{P}_1 = (1-p)^{3n} + 3np(1-p)^{3n-1} = (1-p)^{3n-1} (1-p + 3np)$.

Conclusion : la probabilité que M soit idempotente est $(1-p)^{3n-1} ((3n-1)p + 1)$.

d) Notons $\mathbb{P}_2$ la probabilité qu'au moins une ligne soit égale à la somme des deux autres.

Alors : $\mathbb{P}_2 = \mathbb{P}((X = Y + Z) \cup (Y = X + Z) \cup (Z = X + Y)) = \mathbb{P}(E_X \cup E_Y \cup E_Z)$ en notant respectivement E_X , E_Y et E_Z les événements $(X = Y + Z)$ et $(Y = X + Z)$ et $(Z = X + Y)$.

En utilisant la formule bien connue donnant la probabilité d'une union de trois événements, on en déduit que :

$$\mathbb{P}_2 = \mathbb{P}(E_X) + \mathbb{P}(E_Y) + \mathbb{P}(E_Z) - \mathbb{P}(E_X \cap E_Y) - \mathbb{P}(E_X \cap E_Z) - \mathbb{P}(E_Y \cap E_Z) + \mathbb{P}(E_X \cap E_Y \cap E_Z)$$

Pour des raisons de symétrie évidentes, on a : $\mathbb{P}(E_X) = \mathbb{P}(E_Y) = \mathbb{P}(E_Z)$ et

$$\mathbb{P}(E_X \cap E_Y) = \mathbb{P}(E_X \cap E_Z) = \mathbb{P}(E_Y \cap E_Z).$$

Ainsi : $\boxed{\mathbb{P}_2 = 3\mathbb{P}(E_X) - 3\mathbb{P}(E_X \cap E_Y) + \mathbb{P}(E_X \cap E_Y \cap E_Z)} \quad (\spadesuit)$

¹⁵. $X + Y + Z$ suit la loi binomiale $\mathcal{B}(3n, p)$.

► Calcul de $\mathbb{P}(E_X)$. Les évènements $(X = k)_{k \in \llbracket 0; n \rrbracket}$ constituent une partition de l'univers, et on peut donc appliquer la formule des probabilités totales pour obtenir :

$$\mathbb{P}(E_X) = \mathbb{P}(X = Y + Z) = \sum_{k=0}^n \mathbb{P}((X = k) \cap (Y + Z = k))$$

Et puisque X et $Y + Z$ sont des variables aléatoires indépendantes :

$$\begin{aligned} \mathbb{P}(E_X) &= \mathbb{P}(X = Y + Z) = \sum_{k=0}^n \mathbb{P}(X = k) \mathbb{P}(Y + Z = k) \\ \iff \mathbb{P}(E_X) &= \sum_{k=0}^n \mathbb{P}(X = k) \sum_{j=0}^k \mathbb{P}((Y = j) \cap (Z = k - j)) \end{aligned}$$

Comme Y et Z sont des variables aléatoires indépendantes on a :

$$\begin{aligned} \mathbb{P}(E_X) &= \sum_{k=0}^n \mathbb{P}(X = k) \sum_{j=0}^k \mathbb{P}(Y = j) \mathbb{P}(Z = k - j) \\ \iff \mathbb{P}(E_X) &= \sum_{k=0}^n \binom{n}{k} \frac{1}{2^n} \sum_{j=0}^k \binom{n}{j} \frac{1}{2^n} \binom{n}{k-j} \frac{1}{2^n} \\ \iff \mathbb{P}(E_X) &= \frac{1}{2^{3n}} \sum_{k=0}^n \binom{n}{k} \sum_{j=0}^k \binom{n}{j} \binom{n}{k-j} \end{aligned}$$

Or, d'après la formule de Vandermonde : $\sum_{j=0}^k \binom{n}{j} \binom{n}{k-j} = \binom{2n}{k}$. D'où : $\mathbb{P}(E_X) = \frac{1}{2^{3n}} \sum_{k=0}^n \binom{n}{k} \binom{2n}{k}$.

En utilisant le fait que : $\forall k \in \llbracket 0; n \rrbracket$, $\binom{n}{k} = \binom{n}{n-k}$ ("symétrie" des coefficients binomiaux), on en déduit que :

$$\mathbb{P}(E_X) = \frac{1}{2^{3n}} \sum_{k=0}^n \binom{n}{n-k} \binom{2n}{k}$$

En appliquant une fois encore la formule de Vandermonde, on obtient : $\boxed{\mathbb{P}(E_X) = \frac{1}{2^{3n}} \binom{3n}{n} \quad (\heartsuit)}.$

► Calcul de $\mathbb{P}(E_X \cap E_Y)$. L'évènement $E_X \cap E_Y$ est réalisé si et seulement si $X = Y + Z$ et $Y = X + Z$, ce qui entraîne que $X = Y$ et $Z = 0$. Par conséquent :

$$\mathbb{P}(E_X \cap E_Y) = \sum_{k=0}^n \mathbb{P}((X = k) \cap (Y = k) \cap (Z = 0))$$

Comme X , Y et Z sont des variables aléatoires indépendantes on a :

$$\begin{aligned} \mathbb{P}(E_X \cap E_Y) &= \sum_{k=0}^n \mathbb{P}(X = k) \mathbb{P}(Y = k) \mathbb{P}(Z = 0) \\ \iff \mathbb{P}(E_X \cap E_Y) &= \sum_{k=0}^n \binom{n}{k} \frac{1}{2^n} \times \binom{n}{k} \frac{1}{2^n} \times \frac{1}{2^n}. \end{aligned}$$

Donc : $\mathbb{P}(E_X \cap E_Y) = \frac{1}{2^{3n}} \sum_{k=0}^n \left[\binom{n}{k} \right]^2$ c'est-à-dire : $\mathbb{P}(E_X \cap E_Y) = \frac{1}{2^{3n}} \binom{2n}{n} \quad (\diamond)$

► Calcul de $\mathbb{P}(E_X \cap E_Y \cap E_Z)$. L'évènement $E_X \cap E_Y \cap E_Z$ est réalisé si et seulement si $X = Y + Z$, $Y = X + Z$ et $Z = X + Y$, ce qui entraîne que $X = Y = Z = 0$. D'où :

$$\mathbb{P}(E_X \cap E_Y \cap E_Z) = \mathbb{P}((X = 0) \cap (Y = 0) \cap (Z = 0))$$

Comme X , Y et Z sont des variables aléatoires indépendantes on a :

$$\mathbb{P}(E_X \cap E_Y \cap E_Z) = \mathbb{P}(X = 0) \mathbb{P}(Y = 0) \mathbb{P}(Z = 0)$$

Donc : $\mathbb{P}(E_X \cap E_Y \cap E_Z) = \frac{1}{2^{3n}} \quad (\clubsuit)$

Ainsi, d'après $(\spadesuit)$, $(\heartsuit)$, $(\diamond)$ et $(\clubsuit)$ on a : $\mathbb{P}_2 = 3 \frac{1}{2^{3n}} \binom{3n}{n} - 3 \frac{1}{2^{3n}} \binom{2n}{n} + \frac{1}{2^{3n}}$

Conclusion : si X , Y et Z sont des variables aléatoires mutuellement indépendantes, chacune suivant une loi binomiale $\mathcal{B}(n, 1/2)$ (avec $n \in \mathbb{N}^*$), la probabilité $\mathbb{P}_2$ qu'au moins une ligne de la matrice $M = \begin{pmatrix} X & X & X \\ Y & Y & Y \\ Z & Z & Z \end{pmatrix}$ soit égale à la somme des deux autres est : $\mathbb{P}_2 = \frac{1}{2^{3n}} \left[3 \binom{3n}{n} - 3 \binom{2n}{n} + 1 \right]$

20.7.2 Problème — Un processus de Markov à trois positions

Présentation du problème : cet énoncé reprend le cadre de l'exercice portant sur le processus de Markov d'un DM. Mais si la trame est identique, le contenu à proprement parler est très différent.

Explicitement, les données du problème sont les mêmes, mais on ne fera ici que l'étude de deux des trois suites $(a_n)_n$, $(b_n)_n$ et $(c_n)_n$. Cette étude sera faite par le biais de matrices carrées de taille 2, ce qui simplifiera spectaculairement les calculs¹⁶, notamment lorsqu'il s'agira d'élever la matrice du problème à la puissance n (**partie IV**) ; puis d'en déduire les expressions en fonction de n de a_n , b_n et c_n (**partie V**).

Sans trop de surprise, il y a un prix à payer pour cette simplification. Au lieu d'étudier une suite de vecteurs $(V_n)_n$ "géométrique" (caractérisée par la relation de récurrence $V_{n+1} = MV_n$), c'est une suite de vecteurs "arithmético-géométrique" qui interviendra dans le présent problème. Les propriétés de cette suite de vecteurs seront étudiées dans les **parties II et III**.

Au passage, prenez garde au fait que les notions de suites géométrique et arithmético-géométrique ne sont définies que pour les nombres complexes, et pas pour les suites de vecteurs ; par conséquent, aucune propriété connue sur ce type de suite ne peut être admise dans ce problème, tout résultat les concernant doit être démontré (cela sera de toute façon précisé dans l'énoncé).

Enfin, quitte à radoter, vous pouvez admettre des résultats intermédiaires donnés dans l'énoncé pour traiter les questions qui suivent.

Une impulsion se propage toutes les secondes entre trois positions A , B et C .

Notons A_n (resp. B_n , C_n) l'évènement "être en A (resp. en B , en C) à l'instant n ", et notons encore :

$$a_n = P(A_n), \quad b_n = P(B_n) \quad \text{et} \quad c_n = P(C_n).$$

¹⁶ Ouf !

On sait que si à la seconde n

- ⇒ l'impulsion est arrivée en A , la probabilité pour qu'elle se transmette en B est de $3/4$, et pour qu'elle se transmette en C est de $1/4$;
- ⇒ l'impulsion est arrivée en B , la probabilité pour qu'elle se transmette en A est de $3/4$, et pour qu'elle se transmette en C est de $1/4$;
- ⇒ l'impulsion est arrivée en C , elle se transmet en B

à la seconde $n + 1$.

Rappelons qu'il a été établi en DM que :

$$(F1) \quad \forall n \in \mathbb{N}, \quad \begin{cases} a_{n+1} = \frac{3}{4} b_n \\ b_{n+1} = \frac{3}{4} a_n + c_n \\ c_{n+1} = \frac{1}{4} a_n + \frac{1}{4} b_n \end{cases}$$

Ces relations pourront être utilisées sans être redémontrées. Pour terminer, on suppose qu'initialement (à la seconde $n = 0$), l'impulsion est en A .

Partie I — Préambule

- 1) Justifier brièvement que : $(F2) \quad \forall n \in \mathbb{N}, \quad c_n = 1 - a_n - b_n$
- 2) Etablir à l'aide de la question précédente qu'il existe une matrice $M \in M_2(\mathbb{R})$ et un vecteur $B \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2$ que l'on déterminera tels que :

$$(F3) \quad \forall n \in \mathbb{N}, \quad \begin{pmatrix} a_{n+1} \\ b_{n+1} \end{pmatrix} = M \begin{pmatrix} a_n \\ b_n \end{pmatrix} + B$$

Notation : dans la suite du problème, on notera $V_n = \begin{pmatrix} a_n \\ b_n \end{pmatrix}$. La formule ci-dessus pourra donc être écrite

$$\forall n \in \mathbb{N}, \quad V_{n+1} = MV_n + B$$

L'objectif principal du problème est d'étudier la suite de vecteurs (V_n) , et en particulier sa convergence.

Partie II — Recherche d'une éventuelle limite.

- 3) Calculer $(I_2 - M)$ (on rappelle que I_2 désigne la matrice identité de $M_2(\mathbb{R})$). Sans calculer son inverse, justifier que $(I_2 - M)$ est inversible.
- 4) On pose $(F4) \quad L = (I_2 - M)^{-1} B$.

Sans calculer explicitement L , établir que : $(F5) \quad L = ML + B$

5) Calculer explicitement L , et vérifier que¹⁷ :

$$(F6) \quad L = \begin{pmatrix} \frac{12}{35} \\ \frac{16}{35} \end{pmatrix}$$

Partie III — Etude d'une suite auxiliaire.

On pose : $(F7) \quad \forall n \in \mathbb{N}, \quad W_n = V_n - L$

6) Montrer que : $(F8) \quad \forall n \in \mathbb{N}, \quad W_{n+1} = MW_n$

7) Etablir par récurrence que : $(F9) \quad \forall n \in \mathbb{N}, \quad W_n = M^n W_0$

Partie IV — Calcul de M^n .

8) On pose : $P = \begin{pmatrix} -3 & 1 \\ 1 & -1 \end{pmatrix}$. Justifier que P est inversible, puis calculer P^{-1} .

9) Calculer $P^{-1}MP$, et s'assurer que l'on obtient une matrice diagonale, que l'on notera D .

10) Calculer D^n pour tout entier naturel n .

11) En déduire l'expression de M^n en fonction de n .

 **Notation (conseil)** : dans la suite du problème, les coefficients de la matrice M^n pourront être notés $\alpha_n, \beta_n, \gamma_n$ et δ_n . On pourra donc écrire $M^n = \begin{pmatrix} \alpha_n & \beta_n \\ \gamma_n & \delta_n \end{pmatrix}$ et n'utiliser les expressions de $\alpha_n, \dots$ qu'en cas de nécessité absolue, c'est-à-dire pas avant la seconde partie de la question 13.

Partie V — Calcul des différentes limites.

12) Déduire des questions précédentes l'expression de W_n en fonction de n .

13) En déduire l'expression de V_n en fonction de n , puis les expressions de a_n et b_n en fonction de n .

14) En déduire $\lim_{n \rightarrow \infty} a_n$ et $\lim_{n \rightarrow \infty} b_n$; puis $\lim_{n \rightarrow \infty} c_n$.

CORRIGÉ.

PARTIE I – Préambule. 1) A chaque seconde n , les événements A_n, B_n et C_n constituent une partition de l'univers. Il s'ensuit que $\mathbb{P}(A_n) + \mathbb{P}(B_n) + \mathbb{P}(C_n) = 1$, soit $a_n + b_n + c_n = 1$. Donc : $\forall n \in \mathbb{N}, c_n = 1 - a_n - b_n$.

2) D'après la question précédente, et les rappels faits dans l'énoncé (formule (F1)), on a :

$$\forall n \in \mathbb{N}, \quad \begin{cases} a_{n+1} = \frac{3}{4}b_n \\ b_{n+1} = -\frac{1}{4}a_n - b_n + 1 \end{cases} \quad \text{d'où : } \forall n \in \mathbb{N}, \quad \begin{pmatrix} a_{n+1} \\ b_{n+1} \end{pmatrix} = \underbrace{\begin{pmatrix} 0 & 3/4 \\ -1/4 & -1 \end{pmatrix}}_{=M} \begin{pmatrix} a_n \\ b_n \end{pmatrix} + \underbrace{\begin{pmatrix} 0 \\ 1 \end{pmatrix}}_{=B}$$

17. Le but de cette question est de faire une vérification intermédiaire; si "votre L est juste", il y a de bonnes chances que vos calculs de M, B et $(I_2 - M)^{-1}$ le soient aussi.

PARTIE II – Recherche d’une éventuelle limite.

3) D’après la question précédente : $I_2 - M = \begin{pmatrix} 1 & -3/4 \\ 1/4 & 2 \end{pmatrix}$. Son déterminant vaut $35/16$; puisqu’il est non nul, la matrice $I_2 - M$ est inversible.

4) En multipliant à gauche par $I_2 - M$ les deux termes de l’égalité (F4), on obtient : $(I_2 - M)L = B$, puis en développant le premier terme : $I_2L - ML = B$. D’où finalement : $L = ML + B$.

5) On peut exploiter (F4) ; il suffit alors de calculer l’inverse de la matrice $(I_2 - M)$.

$$\text{On a : } (I_2 - M)^{-1} = \frac{16}{35} \begin{pmatrix} 2 & 3/4 \\ -1/4 & 1 \end{pmatrix}, \text{ d'où : } L = \frac{16}{35} \begin{pmatrix} 2 & 3/4 \\ -1/4 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \frac{16}{35} \begin{pmatrix} 3/4 \\ 1 \end{pmatrix} = \begin{pmatrix} 12/35 \\ 16/35 \end{pmatrix}$$

PARTIE III – Etude d’une suite auxiliaire.

6) Soit n un entier naturel arbitraire. Par définition, on a : $W_{n+1} = V_{n+1} - L$ et en utilisant les relations (F3) et (F5), on obtient :

$$W_{n+1} = MV_n + B - (ML + B) \text{ soit : } W_{n+1} = MV_n - ML = M(V_n - L). \text{ Par suite : } \forall n \in \mathbb{N}, W_{n+1} = MW_n.$$

7) Notons $\mathcal{P}(n)$ la propriété “ $W_n = M^n W_0$ ”. L’initialisation (pour $n = 0$) est immédiate puisque $M^0 = I_2$. Reste à établir l’hérédité.

Supposons que $\mathcal{P}(n)$ soit vraie pour un certain entier naturel n . En utilisant la question précédente, on peut écrire que $W_{n+1} = MW_n$, puis en utilisant l’hypothèse de récurrence : $W_{n+1} = M(M^n W_0)$. Par suite : $W_{n+1} = M^{n+1} W_0$, donc $\mathcal{P}(n+1)$ est vraie, ce qui prouve l’hérédité.

$$\text{Conclusion : } \forall n \in \mathbb{N}, W_n = M^n W_0.$$

PARTIE IV – Calcul de M^n .

$$8) \text{ On a } \det(P) = 2, \text{ donc } P \text{ est inversible. En outre : } P^{-1} = \frac{1}{2} \begin{pmatrix} -1 & -1 \\ -1 & -3 \end{pmatrix} = -\frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & 3 \end{pmatrix}.$$

$$9) P^{-1}MP = -\frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} 0 & 3/4 \\ -1/4 & -1 \end{pmatrix} \begin{pmatrix} -3 & 1 \\ 1 & -1 \end{pmatrix} = -\frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} 3/4 & -3/4 \\ -1/4 & 3/4 \end{pmatrix} \\ = -\frac{1}{2} \begin{pmatrix} 1/2 & 0 \\ 0 & 3/2 \end{pmatrix} \text{ d'où : } P^{-1}MP = \begin{pmatrix} -1/4 & 0 \\ 0 & -3/4 \end{pmatrix}$$

$$10) \text{ D'après le cours : } \forall n \in \mathbb{N}, D^n = \begin{pmatrix} (-1/4)^n & 0 \\ 0 & (-3/4)^n \end{pmatrix}.$$

11) C’est une question classique ! Notons $\mathcal{P}(n)$ la propriété “ $M^n = PD^n P^{-1}$ ”. L’initialisation (pour $n = 0$) est immédiate puisque $M^0 = I_2$ et $PD^0 P^{-1} = PI_2 P^{-1} = PP^{-1} = I_2$. Reste à établir l’hérédité.

Supposons que $\mathcal{P}(n)$ soit vraie pour un certain entier naturel n . Alors $M^{n+1} = M^n \times M$, et en utilisant la définition de D (à savoir $D = P^{-1}MP$), on peut écrire : $M = PDP^{-1}$. De plus, d’après l’hypothèse de récurrence : $M^n = PD^n P^{-1}$.

Par suite : $M^{n+1} = PD^n P^{-1} PDP^{-1} = PD^n I_2 DP^{-1} = PD^n DP^{-1} = PD^{n+1} P^{-1}$, et cette dernière écriture permet d’affirmer que $\mathcal{P}(n+1)$ est vraie, ce qui prouve l’hérédité.

$$\text{Conclusion : } \forall n \in \mathbb{N}, M^n = PD^n P^{-1}.$$

Soit n un entier naturel arbitraire. Il suffit à présent d'utiliser le résultat ci-dessus pour obtenir l'expression de M^n pour tout entier naturel n . Signalons simplement qu'en prenant quelques précautions, on peut rendre ce produit un peu plus "confortable" à effectuer, et minimiser le nombre de fractions intervenant dans les calculs. Explicitement :

$$\begin{aligned}
 M^n &= \begin{pmatrix} -3 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} (-1/4)^n & 0 \\ 0 & (-3/4)^n \end{pmatrix} \begin{pmatrix} -\frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & 3 \end{pmatrix} \end{pmatrix} \\
 \Leftrightarrow M^n &= -\frac{1}{2} \begin{pmatrix} -3 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} (-1/4)^n & 0 \\ 0 & (-3/4)^n \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 3 \end{pmatrix} \\
 \Leftrightarrow M^n &= -\frac{1}{2} \begin{pmatrix} -3 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} (-1/4)^n & (-1/4)^n \\ (-3/4)^n & 3(-3/4)^n \end{pmatrix} \\
 \Leftrightarrow M^n &= -\frac{1}{2} \begin{pmatrix} -3 \left(-\frac{1}{4}\right)^n + \left(-\frac{3}{4}\right)^n & -3 \left(-\frac{1}{4}\right)^n + 3 \left(-\frac{3}{4}\right)^n \\ \left(-\frac{1}{4}\right)^n - \left(-\frac{3}{4}\right)^n & \left(-\frac{1}{4}\right)^n - 3 \left(-\frac{3}{4}\right)^n \end{pmatrix}
 \end{aligned}$$

Conclusion : $\forall n \in \mathbb{N}, M^n = \begin{pmatrix} \frac{3}{2} \left(-\frac{1}{4}\right)^n - \frac{1}{2} \left(-\frac{3}{4}\right)^n & \frac{3}{2} \left(-\frac{1}{4}\right)^n - \frac{3}{2} \left(-\frac{3}{4}\right)^n \\ -\frac{1}{2} \left(-\frac{1}{4}\right)^n + \frac{1}{2} \left(-\frac{3}{4}\right)^n & -\frac{1}{2} \left(-\frac{1}{4}\right)^n + \frac{3}{2} \left(-\frac{3}{4}\right)^n \end{pmatrix}$

PARTIE V – Calculs des différentes limites.

12) D'après la question 7, on a : $\forall n \in \mathbb{N}, W_n = M^n W_0$. En outre, puisque d'après l'énoncé l'impulsion est en A à la seconde 0, on a : $a_0 = 1$ et $b_0 = 0$, soit : $W_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$. Par conséquent, en suivant l'indication de l'énoncé qui suggérerait de noter $\alpha_n, \beta_n, \gamma_n$ et δ_n les coefficients de M^n , on a :

$$\forall n \in \mathbb{N}, W_n = \begin{pmatrix} \alpha_n & \beta_n \\ \gamma_n & \delta_n \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{soit :} \quad \boxed{\forall n \in \mathbb{N}, W_n = \begin{pmatrix} \alpha_n \\ \gamma_n \end{pmatrix}}.$$

13) Par définition de la suite $(W_n)_n$ (formule (F7)), on a : $\forall n \in \mathbb{N}, V_n = W_n + L$. On en déduit, grâce à la question précédente et la question 5 que :

$$\forall n \in \mathbb{N}, V_n = \begin{pmatrix} \alpha_n \\ \gamma_n \end{pmatrix} + \begin{pmatrix} 12/35 \\ 16/35 \end{pmatrix} \quad \text{soit :} \quad \boxed{\forall n \in \mathbb{N}, V_n = \begin{pmatrix} \alpha_n + \frac{12}{35} \\ \gamma_n + \frac{16}{35} \end{pmatrix}}$$

On en déduit donc que pour tout $n \in \mathbb{N}$ on a : $a_n = \alpha_n + \frac{12}{35}$ et $b_n = \gamma_n + \frac{16}{35}$. Et puisqu'il faut bien finir par être explicite :

$$\forall n \in \mathbb{N}, \begin{cases} a_n = \frac{12}{35} + \frac{3}{2} \left(-\frac{1}{4}\right)^n - \frac{1}{2} \left(-\frac{3}{4}\right)^n \\ b_n = \frac{16}{35} - \frac{1}{2} \left(-\frac{1}{4}\right)^n + \frac{1}{2} \left(-\frac{3}{4}\right)^n \end{cases}$$

14) On observe que $\lim_{n \rightarrow \infty} \left[\frac{3}{2} \left(-\frac{1}{4} \right)^n - \frac{1}{2} \left(-\frac{3}{4} \right)^n \right] = 0$ et $\lim_{n \rightarrow \infty} \left[-\frac{1}{2} \left(-\frac{1}{4} \right)^n + \frac{1}{2} \left(-\frac{3}{4} \right)^n \right] = 0$.

Il s'ensuit que : $\boxed{\lim_{n \rightarrow \infty} a_n = \frac{12}{35}}$ et $\boxed{\lim_{n \rightarrow \infty} b_n = \frac{16}{35}}$.

En passant alors à la limite dans la relation $c_n = 1 - a_n - b_n$, on en déduit que $\boxed{\lim_{n \rightarrow \infty} c_n = \frac{7}{35} = \frac{1}{5}}$.

Chapitre 21

Espaces vectoriels et applications linéaires

Tout au long de ce chapitre, $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$.

21.1 Espaces vectoriels

DÉFINITION 21.1 - Un $\mathbb{K}$ -espace vectoriel $(E, +, \cdot)$ est un ensemble E muni d'une loi de composition interne (notée "+"), et d'une loi externe

$$\begin{aligned}\mathbb{K} \times E &\longrightarrow E \\ (\lambda, \vec{v}) &\longmapsto \lambda \cdot \vec{v}\end{aligned}$$

satisfaisant les conditions suivantes :

1/ $(E, +)$ est un groupe abélien ; le neutre est noté $\vec{0}_E$ et appelé **vecteur nul**.

2/ $\forall \vec{v} \in E, 1 \cdot \vec{v} = \vec{v}$

3/ $\forall (\lambda, \mu) \in \mathbb{K}^2, \forall \vec{v} \in E, \lambda \cdot (\mu \cdot \vec{v}) = (\lambda\mu) \cdot \vec{v}$

4/ $\forall (\lambda, \mu) \in \mathbb{K}^2, \forall \vec{v} \in E, (\lambda + \mu) \cdot \vec{v} = \lambda \cdot \vec{v} + \mu \cdot \vec{v}$

5/ $\forall \lambda \in \mathbb{K}, \forall (\vec{v}, \vec{w}) \in E^2, \lambda \cdot (\vec{v} + \vec{w}) = \lambda \cdot \vec{v} + \lambda \cdot \vec{w}$

Définition d'espace vectoriel (bis) : au-delà de la très formelle définition ci-dessus, ce qu'il faut retenir est qu'un $\mathbb{K}$ -espace vectoriel (ou $\mathbb{K}$ -ev) est un ensemble muni d'une "**addition**" ("loi +") et d'une "**multiplication par un scalaire**" ("loi ."), qui vérifient les propriétés que vous leur connaissez dans le cadre des vecteurs du plan (ou de l'espace).

Notation : on note $(E, +, \cdot)$ un $\mathbb{K}$ -ev lorsque l'on veut insister sur le fait que E est muni d'une loi "+" et d'une loi ".". En pratique cependant, lorsque les lois sont l'addition et la multiplication par un scalaire usuelles, on notera simplement E un espace vectoriel.

La première motivation pour introduire ce nouveau concept est que tous les ensembles ayant servi de cadre aux précédents chapitres sont naturellement des espaces vectoriels. Plus précisément, on dispose d'une très impressionnante série de

“THÉORÈMES GÉNÉRAUX” SUR LES ESPACES VECTORIELS

THÉORÈME 21.1 - $(\mathbb{R}^2, +, \cdot)$ est un $\mathbb{R}$ -espace vectoriel (l'espace vectoriel des vecteurs usuels du plan).

THÉORÈME 21.2 - $(\mathbb{R}^3, +, \cdot)$ est un $\mathbb{R}$ -espace vectoriel (l'espace vectoriel des vecteurs usuels de l'espace usuel).

Et plus généralement :

THÉORÈME 21.3 - Pour tout $n \in \mathbb{N}^*$, $(\mathbb{K}^n, +, \cdot)$ est un $\mathbb{K}$ -espace vectoriel (l'espace vectoriel des n -uplets d'éléments de $\mathbb{K}$).

THÉORÈME 21.4 - $(\mathbb{R}^{\mathbb{N}}, +, \cdot)$ est un $\mathbb{R}$ -espace vectoriel (l'espace vectoriel des suites réelles).

Pour le même prix :

THÉORÈME 21.5 - $(\mathbb{C}^{\mathbb{N}}, +, \cdot)$ est un $\mathbb{C}$ -espace vectoriel, et $(\mathbb{K}^{\mathbb{N}}, +, \cdot)$ est un $\mathbb{K}$ -espace vectoriel (les espaces vectoriels des suites complexes et à valeurs dans $\mathbb{K}$ respectivement).

THÉORÈME 21.6 - $(\mathbb{K}[X], +, \cdot)$ est un $\mathbb{K}$ -espace vectoriel (l'espace vectoriel des polynômes à coefficients dans $\mathbb{K}$).

THÉORÈME 21.7 - $(\mathcal{F}(I, \mathbb{K}), +, \cdot)$ est un $\mathbb{K}$ -espace vectoriel (l'espace vectoriel des fonctions à valeurs dans $\mathbb{K}$; avec I un intervalle de $\mathbb{R}$).

THÉORÈME 21.8 - $(M_{n,p}(\mathbb{K}), +, \cdot)$ est un $\mathbb{K}$ -espace vectoriel (l'espace vectoriel des matrices à n lignes et p colonnes à coefficients dans $\mathbb{K}$).

Remarque : les démonstrations de ces théorèmes ont été (ou seront) brièvement évoquées en classe au cours du premier cours d'algèbre linéaire. Elles reposent toutes sur le même principe : vérifier laborieusement que dans chaque cas, on dispose d'une lci et d'une lce satisfaisant les assertions énoncées dans la définition de $\mathbb{K}$ -ev. Plus tard, on disposera heureusement d'autres outils pour montrer que “quelquechose” est un espace vectoriel. En attendant, il est impératif d'avoir en tête les énoncés listés plus haut : ce sont les “briques élémentaires” pour la suite des événements.

DÉFINITION 21.2 - Un **vecteur** est un élément d'un espace vectoriel.

Dans le sens où ce sont des éléments d'espaces vectoriels, on peut ainsi affirmer que la matrice I_3 , la suite de Fibonacci, le 23ème polynôme de Tchebychev et encore la fonction sinus sont des vecteurs (même si cela ne fait rire que l'auteur de ces lignes).

21.2 Sous-espaces vectoriels

DÉFINITION 21.3 - Soit E un $\mathbb{K}$ -espace vectoriel.

Un **sous-espace vectoriel** de E est une partie F de E , qui est un $\mathbb{K}$ -espace vectoriel.

Remarque. Seule, cette définition est d'une remarquable inutilité pratique. Et ne devient intéressante qu'une fois couplée avec la caractérisation suivante des sev :

PROPRIÉTÉ 21.1 - (Caractérisation des SEV). Soit E un $\mathbb{K}$ -espace vectoriel. Un ensemble F est un sous-espace vectoriel de E si les conditions suivantes sont vérifiées :

- (SEV1) $F \subset E$ (*inclusion*)
- (SEV2) $\vec{0}_E \in F$ (F “*passe par l'origine*”)
- (SEV3) $\forall (\vec{v}, \vec{w}) \in F^2, \vec{v} + \vec{w} \in F$ (*stabilité pour la somme*)
- (SEV4) $\forall \vec{v} \in F, \forall \lambda \in \mathbb{K}, \lambda \cdot \vec{v} \in F$ (*stabilité pour la multiplication par un scalaire*)

PREUVE. Blabla. ☕

En pratique, on peut gagner du temps en observant que les deux dernières conditions, traduisant respectivement la stabilité de F pour la loi interne de E (“l'addition”) et la loi externe de E (la “multiplication par un scalaire”) peuvent être vérifiées simultanément. Explicitement, on a la :

PROPRIÉTÉ 21.2 - (Caractérisation des SEV 2.0). soit E un $\mathbb{K}$ -espace vectoriel. Un ensemble F est un sous-espace vectoriel de E si les conditions suivantes sont vérifiées :

- (SEV1) $F \subset E$ (*inclusion*)
- (SEV2) $\vec{0}_E \in F$ (F “*passe par l'origine*”)
- (SEV3') $\forall (\vec{v}, \vec{w}) \in F^2, \forall (\lambda, \mu) \in \mathbb{K}^2, \lambda \cdot \vec{v} + \mu \cdot \vec{w} \in F$ (*stabilité par combinaison linéaire*)

PREUVE. Blabla. ☕

Remarque technique : l'assertion (SEV3') $\forall (\vec{v}, \vec{w}) \in F^2, \forall (\lambda, \mu) \in \mathbb{K}^2, \lambda \cdot \vec{v} + \mu \cdot \vec{w} \in F$ est encore équivalente à la suivante :

$$(SEV3'') \forall (\vec{v}, \vec{w}) \in F^2, \forall \mu \in \mathbb{K}, \vec{v} + \mu \cdot \vec{w} \in F.$$

EXEMPLES DE SOUS-ESPACES VECTORIELS

Exemple 1 : pour tout espace vectoriel E , E est un sev de E ; à l'opposé, $\{\vec{0}_E\}$ est un sev de E .

Exemple 2 : $\mathcal{C}^0(\mathbb{R}, \mathbb{R})$ est un sev $\mathcal{F}(\mathbb{R}, \mathbb{R})$; $\mathcal{C}^1(\mathbb{R}, \mathbb{R})$ est un sev de $\mathcal{C}^0(\mathbb{R}, \mathbb{R})$; et $\mathcal{C}^\infty(\mathbb{R}, \mathbb{R})$ est un sev de $\mathcal{C}^1(\mathbb{R}, \mathbb{R})$ (donc de $\mathcal{C}^0(\mathbb{R}, \mathbb{R})$ donc de...).

En revanche, l'ensemble des bijections continues sur $\mathbb{R}$ à valeurs dans $\mathbb{R}$ n'est pas un sev de $\mathcal{C}^0(\mathbb{R}, \mathbb{R})$ (pourquoi?).

Exemple 3 : si $\vec{v}$ est un vecteur non nul d'un $\mathbb{K}$ -ev E , alors $\mathbb{K}\vec{v} = \{\lambda \cdot \vec{v} / \lambda \in \mathbb{K}\}$ est un sev de E appelé **droite vectorielle** engendrée par $\vec{v}$.

Exemple 4 : l'ensemble des solutions de l'E.D. $(1+x^2)y' + 2xy = 0$ est un sev de $\mathcal{C}^1(\mathbb{R}, \mathbb{R})$; explicitement, c'est la droite vectorielle engendrée par la fonction $x \in \mathbb{R} \mapsto (1+x^2)^{-1}$. Mais l'ensemble des solutions de l'E.D. $(1+x^2)y' + 2xy = 1$ n'est pas un sev de $\mathcal{C}^1(\mathbb{R}, \mathbb{R})$ (pourquoi?).

Exemple 5 : pour tout $n \in \mathbb{N}$, $\mathbb{K}_n[X]$ est un sev de $\mathbb{K}[X]$; mais l'ensemble des polynômes de degré égal à n n'est pas un sev de $\mathbb{K}[X]$ (pourquoi?).

Exemple 6 : l'ensemble des matrices symétriques est un sev de $M_n(\mathbb{K})$; l'ensemble $GL_n(\mathbb{K})$ des matrices inversibles n'est pas un sev de $M_n(\mathbb{K})$ (pourquoi?).

Exemple 7 : l'ensemble des suites réelles convergentes est un sev de $\mathbb{R}^{\mathbb{N}}$; l'ensemble des suites réelles croissantes n'est pas un sev de $\mathbb{R}^{\mathbb{N}}$ (pourquoi?).

Remarque : puisqu'un sev contient toujours au moins le vecteur nul $\vec{0}_E$, un sev n'est jamais vide.

PROPRIÉTÉ 21.3 - Soient F_1 et F_2 deux sous-espaces vectoriels d'un même $\mathbb{K}$ -ev E .

Alors $F_1 \cap F_2$ est un sev de E .

PREUVE. Soient F_1 et F_2 deux sev de E .

- Alors $F_1 \cap F_2 \subset E$, puisque F_1 et F_2 sont deux parties de E .
- On peut affirmer que $\vec{0}_E \in F_1 \cap F_2$, car $\vec{0}_E \in F_1$ et $\vec{0}_E \in F_2$.
- Soient enfin $\vec{u}$ et $\vec{v}$ deux vecteurs dans $F_1 \cap F_2$, et λ et μ deux scalaires. Alors : $\lambda \vec{u} + \mu \vec{v}$ appartient à F_1 (*resp.* F_2) car $\vec{u}$ et $\vec{v}$ appartiennent en particulier à F_1 (*resp.* à F_2) et F_1 (*resp.* F_2) est un sev de E (donc stable par combinaison linéaire). Il s'ensuit que : $\forall (\vec{u}, \vec{v}) \in (F_1 \cap F_2)^2, \forall (\lambda, \mu) \in \mathbb{K}^2, \lambda \vec{u} + \mu \vec{v} \in F_1 \cap F_2$.

En résumé, $F_1 \cap F_2$ est une partie du $\mathbb{K}$ -ev E , contenant le vecteur nul de E , et stable par combinaison linéaire. En d'autres termes, $F_1 \cap F_2$ est un sev de E . ☕

Exemples : dans $\mathbb{R}^2$, l'intersection des droites vectorielles $\mathbb{R} \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ et $\mathbb{R} \cdot \begin{pmatrix} 2 \\ -3 \end{pmatrix}$ (ie d'équations respectives $y = x$ et $3x + 2y = 0$) est le singleton $\{\vec{0}_{\mathbb{R}^2}\}$. Dans $\mathbb{R}^3$, l'intersection des plans d'équations

respectives $x = 0$ et $y - z = 0$ est la droite vectorielle engendrée par $\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$. Dans $M_2(\mathbb{C})$, l'intersection du sev des matrices symétriques et du sev des matrices triangulaires supérieures est le sev des matrices diagonales ; et l'intersection du sev des matrices symétriques et du sev des matrices antisymétriques est réduit à la matrice nulle.

Remarque. Il résulte de la propriété précédente et de la remarque faite juste avant son énoncé que **l'intersection de deux sous-espaces vectoriels n'est à son tour JAMAIS vide** puisqu'elle contient au moins le vecteur nul.

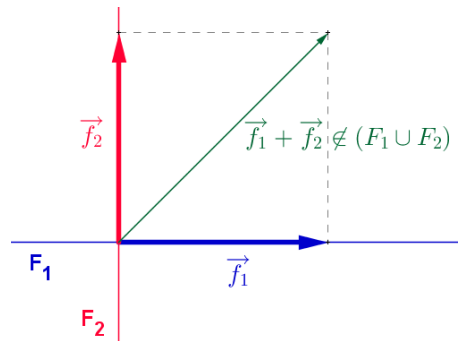
Remarque. En général, l'union de deux sev n'est pas un sev. Par exemple, dans $\mathbb{R}^2$ l'union des droites d'équations $x = 0$ et $y = 0$ n'est pas un sev (cette union n'étant pas stable pour l'addition). Même conclusion dans $\mathbb{R}^3$ pour l'union des plans d'équations $x + y = 0$ et $2x - y + 3z = 0$. D'ailleurs :

PROPRIÉTÉ 21.4 - soient E un $\mathbb{K}$ -ev, F_1 et F_2 deux sev de E .

Alors $(F_1 \cup F_2)$ est un sev de E si et seulement si $F_1 \subset F_2$ ou $F_2 \subset F_1$.

► Illustration de la propriété 21.4.

L'union de deux sev n'est en général pas stable pour la somme.



PREUVE. Soient F_1 et F_2 deux sev d'un même $\mathbb{K}$ -ev E .

Sens direct : supposons que $F_1 \cup F_2$ est un sev de E .

Raisonnons par l'absurde, en supposant que : $(F_1 \not\subset F_2) \wedge (F_2 \not\subset F_1)$ ($\spadesuit$).

Alors il existe un vecteur $\vec{f}_1 \in F_1$, $\vec{f}_1 \notin F_2$; et il existe un vecteur $\vec{f}_2 \in F_2$, $\vec{f}_2 \notin F_1$.

Considérons le vecteur $\vec{V} = \vec{f}_1 + \vec{f}_2$. Comme $\vec{f}_1$ et $\vec{f}_2$ appartiennent tous deux à $F_1 \cup F_2$, et que l'on a supposé que $F_1 \cup F_2$ est un sev, on peut affirmer que $\vec{V} \in F_1 \cup F_2$.

Par suite $\vec{V} \in F_1$, ou $\vec{V} \in F_2$. Or si $\vec{V} \in F_1$, alors $\vec{V} - \vec{f}_1 \in F_1$; mais $\vec{V} - \vec{f}_1 = \vec{f}_2 \notin F_1$ par hypothèse, d'où une contradiction. On aboutit de même à une contradiction si l'on suppose $\vec{V} \in F_2$.

Il s'ensuit que l'hypothèse de départ ($\spadesuit$) est erronée. Donc : $F_1 \subset F_2$ ou $F_2 \subset F_1$.

En résumé, on a établi que : $(F_1 \cup F_2 \text{ est un sev de } E) \implies (F_1 \subset F_2) \vee (F_2 \subset F_1)$

Réciproquement : supposons que $F_1 \subset F_2$ ou $F_2 \subset F_1$.

Si $F_1 \subset F_2$, alors $F_1 \cup F_2 = F_2$, qui est un sev de E . Raisonement analogue en permutant "1" et "2".

Par suite : $(F_1 \subset F_2) \vee (F_2 \subset F_1) \implies (F_1 \cup F_2 \text{ est un sev de } E)$.

Conclusion : $(F_1 \subset F_2) \vee (F_2 \subset F_1) \iff (F_1 \cup F_2 \text{ est un sev de } E)$.



21.3 Combinaisons linéaires, familles génératrices

DÉFINITION 21.4 - Soit E un $\mathbb{K}$ -espace vectoriel, et soit $\mathcal{F} = (\vec{v}_i)_{i \in \llbracket 1, n \rrbracket}$ une famille finie de vecteurs de E .

On note $\text{Vect}(\mathcal{F})$ l'ensemble des combinaisons linéaires de vecteurs de $\mathcal{F}$, c'est-à-dire l'ensemble des vecteurs de E s'écrivant : $\sum_{i=1}^n \alpha_i \vec{v}_i$, avec $(\alpha_i)_{i \in \llbracket 1, n \rrbracket} \in \mathbb{K}^n$.

PROPRIÉTÉ 21.5 - Avec les notations précédentes : $\text{Vect}(\mathcal{F})$ est un sous-espace vectoriel de E .

PREUVE. Banale. ☕

MÉTHODE — La propriété précédente fournit une nouvelle méthode pour prouver qu'une partie F d'un ev E est un sous-espace vectoriel. Il "suffit" en effet de montrer que tout élément de F est combinaison linéaire d'éléments de E que l'on détermine par le calcul ; voir exemples un peu plus bas (et ceux traités en classe).

DÉFINITION 21.5 - Soit F un sev de E . On appelle **famille génératrice** de F une famille $(\vec{v}_i)_{i \in \llbracket 1, n \rrbracket}$ de vecteurs de E telle que :

$$F = \text{Vect} \left((\vec{v}_i)_{i \in \llbracket 1, n \rrbracket} \right)$$

Exemple 1 : on a $\mathbb{R}^2 = \text{Vect}((1, 0); (0, 1))$.

Exemple 2 : dans l'espace vectoriel $M_2(\mathbb{R})$, le sev $S_2(\mathbb{R})$ des matrices symétriques peut s'écrire :

$$\text{Vect}(E_{11}, E_{22}, E_{12} + E_{21})$$

Et le sev $A_2(\mathbb{R})$ des matrices antisymétriques s'écrit : $\text{Vect}(E_{12} - E_{21})$; il s'agit donc d'une droite vectorielle dans l'espace des matrices carrées de taille 2.

Exemple 3 : notons F la partie de $\mathbb{R}^{\mathbb{N}}$ constituée des suites $(x_n)_n$ telles que :

$$\forall n \in \mathbb{N}, x_{n+2} = x_{n+1} + x_n. \text{ On a : } F = \text{Vect} \left(\left(\left(\frac{1 + \sqrt{5}}{2} \right)^n \right)_n, \left(\left(\frac{1 - \sqrt{5}}{2} \right)^n \right)_n \right).$$

F est en particulier un sev de $\mathbb{R}^{\mathbb{N}}$ et en anticipant un peu, on peut affirmer que c'est un plan de $\mathbb{R}^{\mathbb{N}}$, puisqu'il est engendré par deux vecteurs non colinéaires (et la suite de Fibonacci est un élément de ce plan).

Attention : il n'y a pas unicité de la famille génératrice d'un sev. Autrement dit, deux familles distinctes de vecteurs peuvent engendrer le même sev. Explicitement :

Exemple 4 : on a $\mathbb{R}_2[X] = \text{Vect}(1, X, X^2)$ (par définition de polynôme), $\mathbb{R}_2[X] = \text{Vect}(1, (X - \alpha), (X - \alpha)^2)$ (formule de Taylor pour les polynômes), $\mathbb{R}_2[X] = \text{Vect}(L_{-1}, L_0, L_1)$ (où les L_i désignent les polynômes interpolateurs de Lagrange associés aux scalaires $-1, 0$ et 1).

Plus généralement :

PROPRIÉTÉ 21.6 - Soient E un $\mathbb{K}$ -espace vectoriel, $\mathcal{F} = (\vec{v}_i)_{i \in [1, n]}$ et $\mathcal{F}' = (\vec{w}_j)_{j \in [1, p]}$ deux familles de vecteurs de E . On a :

$$\text{Vect}((\vec{v}_i)_{i \in [1, n]}) = \text{Vect}((\vec{w}_j)_{j \in [1, p]}) \text{ SSI}$$

$$\forall i \in [1, n], v_i \in \text{Vect}(\mathcal{F}') \text{ et } \forall j \in [1, p], w_j \in \text{Vect}(\mathcal{F})$$

PREUVE. Raisonnons par double implication.

► Sens direct ($\implies$). Supposons que $\text{Vect}(\mathcal{F}) = \text{Vect}(\mathcal{F}')$.

Soit $i \in [1, n]$. On observe judicieusement que $\vec{v}_i \in \text{Vect}(\mathcal{F})$ ¹. D'où par hypothèse : $v_i \in \text{Vect}(\mathcal{F}')$.

Ainsi : $\forall i \in [1, n], v_i \in \text{Vect}(\mathcal{F}')$.

Pour des raisons de symétrie évidentes, on a aussi : $\forall j \in [1, p], w_j \in \text{Vect}(\mathcal{F})$.

Ce qui prouve la première implication.

► Réciproquement ($\impliedby$). Supposons que $\forall i \in [1, n], v_i \in \text{Vect}(\mathcal{F}')$ et $\forall j \in [1, p], w_j \in \text{Vect}(\mathcal{F})$; et montrons que $\text{Vect}(\mathcal{F}) = \text{Vect}(\mathcal{F}')$ par double inclusion.

Soit $\vec{V} \in \text{Vect}(\mathcal{F})$. Alors il existe n scalaires $\alpha_1, \dots, \alpha_n$ tels que : $\vec{V} = \sum_{i=1}^n \alpha_i \vec{v}_i$ (♠).

En outre, par hypothèse, chacun des vecteurs $\vec{v}_i$ est combinaison linéaire des vecteurs de la famille $\mathcal{F}'$.

Pour un entier i donné dans $[1, n]$, il existe p scalaires $x_{i,1}, \dots, x_{i,p}$ tels que : $\vec{v}_i = \sum_{j=1}^p x_{i,j} \vec{w}_j$ (♣).

On déduit de (♠) et de (♣) que : $\vec{V} = \sum_{i=1}^n \alpha_i \sum_{j=1}^p x_{i,j} \vec{w}_j = \sum_{i=1}^n \sum_{j=1}^p \alpha_i x_{i,j} \vec{w}_j = \sum_{j=1}^p \left(\sum_{i=1}^n \alpha_i x_{i,j} \right) \vec{w}_j$

D'où finalement : $\vec{V} = \sum_{j=1}^p \beta_j \vec{w}_j$ en ayant posé : $\beta_j = \sum_{i=1}^n \alpha_i x_{i,j}$. En particulier : $\vec{V} \in \text{Vect}(\mathcal{F}')$.

Ainsi : $\left[\vec{V} \in \text{Vect}(\mathcal{F}) \right] \implies \left[\vec{V} \in \text{Vect}(\mathcal{F}') \right]$. D'où l'inclusion : $\text{Vect}(\mathcal{F}) \subset \text{Vect}(\mathcal{F}')$.

L'inclusion réciproque s'en déduit par symétrie, et on peut conclure que : $\text{Vect}(\mathcal{F}) = \text{Vect}(\mathcal{F}')$, ce qui prouve la réciproque de la propriété.

Conclusion. ($\text{Vect}(\mathcal{F}) = \text{Vect}(\mathcal{F}')$)

$$\text{SSI } (\forall i \in [1, n], v_i \in \text{Vect}(\mathcal{F}') \text{ et } \forall j \in [1, p], w_j \in \text{Vect}(\mathcal{F})). \quad \text{☕}$$

1. Informellement car : $\vec{v}_i = 0 \times \vec{v}_1 + 0 \times \vec{v}_2 + \dots + 1 \times \vec{v}_i + \dots + 0 \times \vec{v}_n$, et "hyperformellement" car : $\vec{v}_i = \sum_{j=1}^n \delta_{ij} \vec{v}_j$.

Pour finir, observons que la notion de famille génératrice peut être étendue à celle d'une famille quelconque de vecteurs (et pas seulement d'une famille finie). Explicitement :

DÉFINITION 21.6 - Soient E un $\mathbb{K}$ -espace vectoriel, I un ensemble, et $\mathcal{F} = (\vec{v}_i)_{i \in I}$ une famille de vecteurs de E .

On note $\text{Vect}(\mathcal{F})$ l'ensemble des combinaisons linéaires finies de vecteurs de $\mathcal{F}$, c'est-à-dire l'ensemble des vecteurs de E s'écrivant : $\sum_{j=1}^p \alpha_{i_j} \vec{v}_{i_j}$, avec $(\alpha_{i_j})_{j \in [1,p]} \in \mathbb{K}^p$.

PROPRIÉTÉ 21.7 - Avec les notations précédentes : $\text{Vect}(\mathcal{F})$ est un sous-espace vectoriel de E .

PREUVE. Banale. ☕

Exemple. On peut ainsi affirmer que $\mathbb{K}[X] = \text{Vect}\left((X^k)_{k \in \mathbb{N}}\right)$, ou encore que pour tout scalaire α on a $\mathbb{K}[X] = \text{Vect}\left(\left((X - \alpha)^k\right)_{k \in \mathbb{N}}\right)$ (d'après la formule de Taylor pour les polynômes).

Avec cette définition, on peut également écrire que $\mathbb{R}^{\mathbb{N}} = \text{Vect}\left((e_i)_{i \in \mathbb{N}}\right)$, où e_i désigne la suite dont le i -ème terme vaut 1 et dont tous les autres sont nuls.

21.4 Applications linéaires

DÉFINITION 21.7 - Une application $f : E \longrightarrow F$ entre deux $\mathbb{K}$ -ev est **linéaire** si :

$$\forall (\lambda, \mu) \in \mathbb{K}^2, \forall (\vec{u}, \vec{v}) \in E^2, f(\lambda \vec{u} + \mu \vec{v}) = \lambda f(\vec{u}) + \mu f(\vec{v})$$

Un **endomorphisme** de E est une application linéaire de E dans lui-même (" $F = E$ ").

Remarque 1. Une application linéaire est donc une application entre espaces vectoriels, compatible avec les lois de ces espaces dans le sens où pour une telle application, *"l'image d'une combinaison linéaire est la combinaison linéaire des images"*.

Remarque 2. Trois propriétés peuvent être obtenues comme conséquences immédiates de la définition. Si $f : E \longrightarrow F$ est linéaire, alors :

$$\blacktriangleright \forall (\vec{u}, \vec{v}) \in E^2, f(\vec{u} + \vec{v}) = f(\vec{u}) + f(\vec{v})$$

("l'image d'une somme est la somme des images")

$$\blacktriangleright \forall (\lambda, \vec{v}) \in \mathbb{K} \times E, f(\lambda \vec{v}) = \lambda f(\vec{v})$$

("f est compatible avec la loi externe")

$$\blacktriangleright f(\vec{0}_E) = \vec{0}_F \text{ ("f envoie le vecteur nul sur le vecteur nul")}$$

Notations : $\mathcal{L}(E, F)$ désigne l'ensemble des applications linéaires de E dans F , $\mathcal{L}(E)$ celui des endomorphismes de E .

Exemple 1. L'application

$$\begin{aligned}\mathbb{R}^2 &\longrightarrow \mathbb{R}^2 \\ (x, y) &\longmapsto (x + y, 0)\end{aligned}$$

est linéaire.

Exemple 2. L'application

$$\begin{aligned}\mathbb{R}^{\mathbb{N}} &\longrightarrow \mathbb{R}^2 \\ (u_n) &\longmapsto (u_0, u_1)\end{aligned}$$

est linéaire.

Exemple 3. L'application

$$\begin{aligned}\mathbb{K}_n[X] &\longrightarrow \mathbb{K}_{n-1}[X] \\ P &\longmapsto P'\end{aligned}$$

est linéaire.

Exemple 4. L'application

$$\begin{aligned}\mathbb{R}[X] &\longrightarrow \mathbb{R} \\ P &\longmapsto \int_0^\pi P(t) \sin(t) \, dt\end{aligned}$$

est linéaire.

Exemple 5. L'application

$$\begin{aligned}\mathrm{M}_n(\mathbb{K}) &\longrightarrow \mathrm{M}_n(\mathbb{K}) \\ A &\longmapsto {}^t A\end{aligned}$$

est linéaire.

Exemple 6. L'application

$$\begin{aligned}\mathcal{C}^1(\mathbb{R}, \mathbb{C}) &\longrightarrow \mathcal{C}^0(\mathbb{R}, \mathbb{C}) \\ f &\longmapsto f' - 2f\end{aligned}$$

est linéaire.

Exemple 7. L'application

$$\begin{aligned}\mathrm{M}_n(\mathbb{K}) &\longrightarrow \mathbb{K} \\ A &\longmapsto \mathrm{tr}(A)\end{aligned}$$

est linéaire.

PROPRIÉTÉ 21.8 - La somme de deux applications linéaires est linéaire ; plus généralement une combinaison linéaire d'applications linéaires est linéaire ; la composée de deux applications linéaires est linéaire. Plus précisément :

$$1/ \forall (f, g) \in \mathcal{L}(E, F)^2, (f + g) \in \mathcal{L}(E, F)$$

$$2/ \forall (f, g) \in \mathcal{L}(E, F)^2, \forall (\lambda, \mu) \in \mathbb{K}^2, (\lambda f + \mu g) \in \mathcal{L}(E, F)$$

$$3/ \forall (f, g) \in \mathcal{L}(E, F) \times \mathcal{L}(F, G), g \circ f \in \mathcal{L}(E, G)$$

PREUVE. Triviale. ☕

21.5 Noyau et image d'une application linéaire

21.5.1 $\ker$ et Im

DÉFINITION 21.8 - Soit $f \in \mathcal{L}(E, F)$. On appelle **noyau de f** et on note $\ker f$ le sous-ensemble de E suivant :

$$\ker f = \left\{ \vec{v} \in E \mid f(\vec{v}) = \vec{0}_F \right\}$$

On appelle **image de f** et on note $\text{Im } f$ le sous-ensemble de F suivant :

$$\text{Im } f = \{ \vec{w} \in F \mid \exists \vec{v} \in E, f(\vec{v}) = \vec{w} \}$$

Remarque. En fait vous connaissiez déjà ce type d'ensemble, depuis le chapitre portant sur les applications. Le noyau de f est en effet l'image réciproque par f du singleton $\{ \vec{0}_F \}$, et l'image de f est l'image directe de E par f .

Exemple. L'application qui à un polynôme P associe son polynôme dérivé induit un endomorphisme de $\mathbb{K}_n[X]$, dont le noyau est le sev des polynômes constants, et l'image est le sev des polynômes de degré au plus 1.

PROPRIÉTÉ 21.9 - Si $f \in \mathcal{L}(E, F)$, alors $\ker f$ est un sev de E et $\text{Im } f$ est un sev de F .

PREUVE. Soit $f : E \longrightarrow F$ une application linéaire.

Montrons que $\ker f$ est un sev de E . Par définition, $\ker f$ est une partie de E . Le vecteur nul y appartient, puisque $f(\vec{0}_E) = \vec{0}_F$ pour toute application linéaire f entre E et F .

Il reste à établir que $\ker f$ est stable par combinaison linéaire. Soient $\vec{u}$ et $\vec{v}$ dans $\ker f$, et λ et μ deux scalaires.

Alors : $f(\lambda \vec{u} + \mu \vec{v}) = \lambda f(\vec{u}) + \mu f(\vec{v}) = \lambda \vec{0}_F + \mu \vec{0}_F = \vec{0}_F$; la première égalité provenant de la définition d'application linéaire, et la seconde du fait que $\vec{u}$ et $\vec{v}$ sont dans le noyau de f par hypothèse. Par conséquent : $\lambda \vec{u} + \mu \vec{v} \in \ker f$, et $\ker f$ est donc stable par combinaison linéaire.

Conclusion : $\ker f$ est une partie de E , contenant $\vec{0}_E$, et stable par combinaison linéaire. Par suite, $\ker f$ est un sev de E .

Montrons que $\text{Im} f$ est un sev de F . Par définition, $\text{Im} f$ est une partie de F . Le vecteur nul (de F) y appartient, puisque $\vec{0}_F = f(\vec{0}_E)$ pour toute application linéaire f entre E et F .

Il reste à établir que $\text{Im} f$ est stable par combinaison linéaire. Soient $\vec{u}$ et $\vec{v}$ dans $\text{Im} f$, et λ et μ deux scalaires.

Par hypothèse, il existe deux vecteurs $\vec{a}$ et $\vec{b}$ tels que $f(\vec{a}) = \vec{u}$ et $f(\vec{b}) = \vec{v}$. Par conséquent :

$f(\lambda \vec{a} + \mu \vec{b}) = \lambda f(\vec{a}) + \mu f(\vec{b}) = \lambda \vec{u} + \mu \vec{v}$; la première égalité provenant de la définition d'application linéaire, et la seconde des définitions respectives des vecteurs $\vec{u}$ et $\vec{v}$.

Il s'ensuit que $\lambda \vec{u} + \mu \vec{v} \in \text{Im} f$ (puisque'il admet pour antécédent $\lambda \vec{a} + \mu \vec{b}$). Ainsi $\text{Im} f$ est stable par combinaison linéaire.

Conclusion : $\text{Im} f$ est une partie de F , contenant $\vec{0}_F$, et stable par combinaison linéaire. Par suite, $\text{Im} f$ est un sev de F . ☕

Remarques pratiques. Déterminer le noyau d'une application linéaire revient à résoudre une équation (" $f(\vec{v}) = 0$ "). Il faut alors avoir présent à l'esprit que suivant l'espace vectoriel dans lequel on travaille, cette équation peut être une équation polynomiale, une équation différentielle, un système linéaire, une équation matricielle...

Déterminer l'image d'une application linéaire peut s'avérer une opération plus délicate. Dans le cas où l'on dispose d'une famille génératrice de l'espace de départ (par exemple si $E = \mathbb{R}^3$ ou $\mathbb{K}_n[X]$ ou $M_n(\mathbb{K})$), la propriété ci-dessous donne une méthode pratique pour y parvenir.

PROPRIÉTÉ 21.10 - Soient E et F deux $\mathbb{K}$ -ev, et $f \in \mathcal{L}(E, F)$.

On suppose qu'il existe une famille $(\vec{v}_i)_{i \in [1, n]}$ génératrice de E , c'est à dire une famille de vecteurs de E tels que $E = \text{Vect}((\vec{v}_i)_{i \in [1, n]})$. Alors :

$$\text{Im} f = \text{Vect}((f(\vec{v}_i))_{i \in [1, n]})$$

PREUVE. ► Montrons l'inclusion : $\text{Im} f \subset \text{Vect}((f(\vec{v}_i))_{i \in [1, n]})$. Soit $\vec{V} \in \text{Im} f$. Il existe un vecteur $\vec{u}$ de E tel que : $f(\vec{u}) = \vec{V}$.

Puisque la famille $(\vec{v}_i)_{i \in [1, n]}$ est génératrice de E : $\exists (\lambda_1, \dots, \lambda_n) \in \mathbb{K}^n$, $\vec{u} = \sum_{i=1}^n \lambda_i \vec{v}_i$.

Ainsi : $\vec{V} = f\left(\sum_{i=1}^n \lambda_i \vec{v}_i\right)$ et par linéarité de f : $\vec{V} = \sum_{i=1}^n \lambda_i f(\vec{v}_i)$, d'où : $\vec{V} \in \text{Vect}((f(\vec{v}_i))_{i \in [1, n]})$.

Conclusion : $\text{Im} f \subset \text{Vect}((f(\vec{v}_i))_{i \in [1, n]})$

► Montrons l'inclusion : $\text{Vect}((f(\vec{v}_i))_{i \in [1, n]}) \subset \text{Im} f$. Soit $\vec{V} \in \text{Vect}((f(\vec{v}_i))_{i \in [1, n]})$.

Alors : $\exists (\lambda_1, \dots, \lambda_n) \in \mathbb{K}^n$, $\vec{V} = \sum_{i=1}^n \lambda_i f(\vec{v}_i)$.

Par linéarité de f , on a : $\vec{V} = f\left(\sum_{i=1}^n \lambda_i \vec{v}_i\right)$, d'où $\vec{V} \in \text{Im} f$.

Conclusion : $\text{Vect}\left((f(\vec{v}_i))_{i \in \llbracket 1, n \rrbracket}\right) \subset \text{Im} f$

Finalement : $\boxed{\text{Im} f = \text{Vect}\left((f(\vec{v}_i))_{i \in \llbracket 1, n \rrbracket}\right)}$ ☕

Exemple 1. Considérons l'application linéaire $f : M_n(\mathbb{K}) \longrightarrow \mathbb{K}$.

$$A \longmapsto \text{tr}(A)$$

Le noyau de f est le sev de $M_n(\mathbb{K})$ constitué des matrices de trace nulle.

Son image est : $\text{Im} f = \text{Vect}\left((f(E_{ij}))_{(i,j) \in \llbracket 1, n \rrbracket^2}\right)$.

En particulier : $\text{Vect}(f(E_{11})) \text{Im} f \subset \mathbb{K}$. D'où : $\text{Vect}(1) \subset \text{Im} f \subset \mathbb{K}$.

Or : $\text{Vect}(1) = \mathbb{K}$. On en déduit donc que $\text{Im} f = \mathbb{K}$, càd que f est surjective.

Nous aurons l'occasion de revenir en détails sur ce type de situation lorsque nous étudierons les *formes linéaires*.

Exemple 2. Considérons l'application linéaire $f : \mathbb{R}_2[X] \longrightarrow \mathbb{R}_2[X]$.

$$P \longmapsto P - XP'$$

Le noyau de f est le sev des solutions dans $\mathbb{R}_2[X]$ de l'équation polynomiale $P - XP' = 0$. On peut obtenir facilement que : $\ker f = \text{Vect}(X)$.

Son image est : $\text{Im} f = \text{Vect}(f(1), f(X), f(X^2))$.

Or : $f(1) = 1$, $f(X) = 0$ et $f(X^2) = -X^2$. On en déduit que : $\text{Im} f = \text{Vect}(1, 0, -X^2)$, càd que : $\text{Im} f = \text{Vect}(1, X^2)$.²

Remarque “pour l’avenir”. Dans l'exemple ci-dessus, le noyau de f pourra être appelé *droite vectorielle* engendrée par X , et l'image de f *plan vectoriel* engendré par 1 et X^2 .

THÉORÈME 21.9 - Soit $f \in \mathcal{L}(E, F)$. Alors :

1/ f est injective SSI $\ker f = \{\vec{0}_E\}$

2/ f est surjective SSI $\text{Im} f = F$.

PREUVE. On raisonne par double implication, pour faire preuve d'originalité.

► Supposons f injective. On a déjà $\{\vec{0}_E\} \subset \ker f$ puisque $\ker f$ est un sev de E . Réciproquement, soit $\vec{v}$ un vecteur de $\ker f$. Alors $f(\vec{v}) = \vec{0}_F$. Comme par ailleurs f est linéaire, on a aussi : $f(\vec{0}_E) = \vec{0}_F$. Donc $f(\vec{v}) = f(\vec{0}_E)$, et on en déduit $\vec{v} = \vec{0}_E$ par injectivité de f . Ce qui assure que $\ker f = \{\vec{0}_E\}$.

En résumé : $[f \text{ injective}] \implies [\ker f = \{\vec{0}_E\}]$.

2. En vertu de la formule de Taylor appliquée dans $\mathbb{R}_2[X]$ en 0, il s'agit donc du sev de $\mathbb{R}_2[X]$ constitué des polynômes P tels que $P'(0) = 0$.

► Réciproquement, supposons $\ker f = \{\vec{0}_E\}$. Soient $\vec{v}$ et $\vec{w}$ deux vecteurs de E tels que $f(\vec{v}) = f(\vec{w})$. Alors : $f(\vec{v}) - f(\vec{w}) = \vec{0}_F$, d'où par linéarité de f : $f(\vec{v} - \vec{w}) = \vec{0}_F$. Par suite, $\vec{v} - \vec{w} \in \ker f$, et puisque le noyau de f est réduit au vecteur nul par hypothèse, on en déduit que $\vec{v} - \vec{w} = \vec{0}_E$, d'où : $\vec{v} = \vec{w}$.

On vient donc d'établir que $[f(\vec{v}) = f(\vec{w})] \implies [\vec{v} = \vec{w}]$, ce qui prouve l'injectivité de f .

Par suite : $[\ker f = \{\vec{0}_E\}] \implies [f \text{ injective}]$.

On déduit de cette implication et de celle établie plus haut l'équivalence :

$$[\ker f = \{\vec{0}_E\}] \iff [f \text{ injective}]$$

L'équivalence concernant la surjectivité est une conséquence immédiate des définitions de surjectivité et d'image. ☕

21.5.2 Isomorphismes et automorphismes

DÉFINITION 21.9 - Soient E et F deux $\mathbb{K}$ -ev. On appelle **isomorphisme** entre E et F (*resp.* **automorphisme** de E) une application linéaire $f : E \longrightarrow F$ bijective (*resp.* un endomorphisme de E bijectif).

Les ev E et F sont **isomorphes** s'il existe un isomorphisme entre E et F .

Exemples : l'application qui à tout polynôme P de $\mathbb{R}_2[X]$ associe le triplet de ses coefficients est un isomorphisme de $\mathbb{R}_2[X]$ dans $\mathbb{R}^3$ (les ev $\mathbb{R}_2[X]$ et $\mathbb{R}^3$ sont donc isomorphes). L'application qui à toute matrice carrée de $M_n(\mathbb{K})$ associe sa transposée induit un automorphisme de $M_n(\mathbb{K})$. Voir cours (passé et à venir d'ici la fin de l'année) pour d'autres exemples.

Notation : $GL(E)$ est l'ensemble des automorphismes de E .

PROPRIÉTÉ 21.11 - La composée de deux iso(auto)morphismes est un iso(auto)morphisme.

Si f est un isomorphisme entre E et F , alors f^{-1} est un isomorphisme entre F et E .

Si $f \in \mathcal{L}(E, F)$ et $g \in \mathcal{L}(F, G)$ sont deux isomorphismes, alors $g \circ f$ est un isomorphisme de E dans G , et $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$.

PREUVE. Soient f et g deux isomorphismes, tels que $g \circ f$ soit défini. Alors l'application $g \circ f$ est bijective (car la composée de deux bijections en est une), et $g \circ f$ est linéaire (c'est une propriété immédiate de ce chapitre). Ainsi $g \circ f$ est un isomorphisme, ce qui prouve la première assertion.

Passons à la seconde : considérons f un isomorphisme entre E et F . Puisque f est en particulier une application bijective, sa réciproque $f^{-1} \in E^F$ existe. Soient alors $\vec{U}$ et $\vec{V}$ deux vecteurs de F , et soient λ et μ deux scalaires.

Puisque f est bijective (c'est un isomorphisme), il existe un (unique) vecteur $\vec{u} \in E$ et un (unique) vecteur $\vec{v} \in E$ tels que : $\vec{U} = f(\vec{u})$ et $\vec{V} = f(\vec{v})$ (càd : $\vec{u} = f^{-1}(\vec{U})$ et $\vec{v} = f^{-1}(\vec{V})$). On a alors :

$$f^{-1}(\lambda\vec{U} + \mu\vec{V}) = f^{-1}(\lambda f(\vec{u}) + \mu f(\vec{v})) = f^{-1}(f(\lambda\vec{u} + \mu\vec{v})) = \lambda\vec{u} + \mu\vec{v} = \lambda f^{-1}(\vec{U}) + \mu f^{-1}(\vec{V})$$

En résumé : $\forall (\vec{U}, \vec{V}) \in F^2, \forall (\lambda, \mu) \in \mathbb{K}^2, f^{-1}(\lambda\vec{U} + \mu\vec{V}) = \lambda f^{-1}(\vec{U}) + \mu f^{-1}(\vec{V})$. Ce qui signifie que f^{-1} est linéaire. Puisqu'en outre f^{-1} est bijective, c'est un isomorphisme entre F et E , ce qui prouve la seconde assertion. ☕

COROLLAIRE 21.1 - $(\text{GL}(E), \circ)$ est un groupe, appelé **groupe linéaire de E** .

PREUVE. Conséquence immédiate de la propriété précédente. ☕

21.6 Sommes de sev et sev supplémentaires

21.6.1 Sommes de sev

DÉFINITION 21.10 - Soient F et G deux sev d'un même $\mathbb{K}$ -ev E . La **somme de F et G** , notée $F + G$ est l'ensemble des vecteurs de E s'écrivant comme somme d'un vecteur de F et d'un vecteur de G , soit :

$$F + G = \left\{ \vec{v} \in E / \exists (\vec{f}, \vec{g}) \in F \times G, \vec{v} = \vec{f} + \vec{g} \right\}$$

PROPRIÉTÉ 21.12 - Avec les notations précédentes, $F + G$ est un sev de E .

PREUVE. Immédiate. ☕

Remarque. En outre, $F + G$ est le plus petit (pour l'inclusion) sev de E contenant F et G : si F' est un autre sev de E contenant F et G , alors F' contient $F + G$ (exercice).

PROPRIÉTÉ 21.13 - Soient $\vec{v}_1, \dots, \vec{v}_n, \vec{w}_1, \dots, \vec{w}_p$ ($n + p$) vecteurs de E . On a :

$$\text{Vect}(\vec{v}_1, \dots, \vec{v}_n) + \text{Vect}(\vec{w}_1, \dots, \vec{w}_p) = \text{Vect}(\vec{v}_1, \dots, \vec{v}_n, \vec{w}_1, \dots, \vec{w}_p)$$

PREUVE. Longue, mais triviale. ☕

Remarque. La définition de somme de sev peut-être généralisée à une somme à n termes.

DÉFINITION 21.11 - Soient p un entier naturel supérieur ou égal à 2, et $F_1, \dots, F_p$ p sev d'un même $\mathbb{K}$ -ev E .

La **somme** $F_1 + \dots + F_p$ est l'ensemble :

$$F_1 + \dots + F_p = \left\{ \vec{v} \in E / \exists (\vec{f}_1, \dots, \vec{f}_p) \in F_1 \times \dots \times F_p, \vec{v} = \sum_{i=1}^p \vec{f}_i \right\}$$

Exemples. $\mathbb{R}_2[X] = \text{Vect}(1) + \text{Vect}(X) + \text{Vect}(X^2)$; $M_n(\mathbb{K}) = T_n^+(\mathbb{K}) + T_n^-(\mathbb{K}) + D_n(\mathbb{K}) \dots$

PROPRIÉTÉ 21.14 - Avec les notations de la définition 21.11, $F_1 + \dots + F_p$ est un sev de E .

PREUVE. Récurrence immédiate à partir de la propriété 21.12, après avoir observé que l'addition est associative : $(F_1 + F_2) + F_3 = F_1 + (F_2 + F_3)$. ☕

Remarque. En outre, $F_1 + \dots + F_p$ est le plus petit (pour l'inclusion) sev de E contenant $F_1, \dots, F_p$: si F' est un autre sev de E contenant $F_1, \dots, F_p$, alors F' contient $F_1 + \dots + F_p$ (exercice).

21.6.2 Sev supplémentaires

DÉFINITION 21.12 - Deux sev F et G d'un même $\mathbb{K}$ -ev E sont dits **supplémentaires (dans E)** lorsque tout vecteur de E s'écrit de manière unique (à l'ordre près) comme somme d'un vecteur de F et d'un vecteur de G , c'est à dire :

$$\forall \vec{v} \in E, \quad \exists! (\vec{f}, \vec{g}) \in F \times G, \quad \vec{v} = \vec{f} + \vec{g}$$

Dans ce cas, on note : $E = F \oplus G$, et on dit que E est la **somme directe** de F et de G .

Exemples. Avec les notations usuelles pour les vecteurs que vous connaissez bien, on a :

$$\mathbb{R}^2 = \text{Vect} \left(\begin{pmatrix} 1 \\ i \end{pmatrix} \right) \oplus \text{Vect} \left(\begin{pmatrix} 1 \\ j \end{pmatrix} \right) \quad \text{et} \quad \mathbb{R}^3 = \text{Vect} \left(\begin{pmatrix} 1 \\ i \\ j \end{pmatrix} \right) \oplus \text{Vect} \left(\begin{pmatrix} 1 \\ k \end{pmatrix} \right)$$

Dans un autre registre, on a encore : $\mathbb{R}_3[X] = \text{Vect}(X, X^2, X^3) \oplus \text{Vect}(1), \dots$

En revanche $F = \text{Vect}(X, X^2)$ et $G = \text{Vect}(X^3)$ ne sont pas supplémentaires dans $\mathbb{R}_3[X]$ (car $F + G \subsetneq E$) ; et $F' = \text{Vect}(1, X, X^2)$ et $G' = \text{Vect}(X^2, X^3)$ ne sont pas supplémentaires non plus dans $\mathbb{R}_3[X]$ ("pas d'unicité de la décomposition").

Autres exemples (les "inratables"). Dans $M_n(\mathbb{K})$, les sev des matrices symétriques $S_n(\mathbb{K})$ et $A_n(\mathbb{K})$ sont supplémentaires (toute matrice carrée s'écrit de manière unique comme somme d'une matrice symétrique et d'une antisymétrique). On a donc : $M_n(\mathbb{K}) = S_n(\mathbb{K}) \oplus A_n(\mathbb{K})$.

Enfin si I désigne un intervalle de $\mathbb{R}$ symétrique par rapport à 0, l'espace vectoriel $\mathcal{F}(I, \mathbb{K})$ des fonctions définies sur I à valeurs dans $\mathbb{K}$ est la somme directe des sev $\mathcal{P}(I, \mathbb{K})$ et $\mathcal{I}(I, \mathbb{K})$ des fonctions paires et impaires respectivement : $\mathcal{F}(I, \mathbb{K}) = \mathcal{P}(I, \mathbb{K}) \oplus \mathcal{I}(I, \mathbb{K})$ (toute fonction définie sur une partie de $\mathbb{R}$ symétrique par rapport à 0 s'écrit de manière unique comme somme d'une fonction paire et d'une fonction impaire).

PROPRIÉTÉ 21.15 - (Caractérisation des sev supplémentaires). Soient F et G deux sev d'un même $\mathbb{K}$ -ev E . Alors :

$$[F \text{ et } G \text{ sont supplémentaires dans } E] \iff [F \cap G = \{ \vec{0} \} \text{ et } E = F + G]$$

Remarque informelle. " $E = F + G$ donne l'existence de la décomposition, et la condition $F \cap G = \{ \vec{0} \}$ entraîne l'unicité".

PREUVE. Soient donc F et G deux sev d'un même $\mathbb{K}$ -ev E . Raisonnons par double implication.

► Sens direct : supposons que F et G soient supplémentaires dans E (càd $E = F \oplus G$). En particulier, on a déjà $E = F + G$.

Montrons que $F \cap G = \{\vec{0}\}$. Soit $\vec{v} \in F \cap G$. Alors le vecteur $\vec{v}$ peut s'écrire $\vec{v} = \underbrace{\vec{v}}_{\in F} + \underbrace{\vec{0}}_{\in G}$ et $\vec{v} = \underbrace{\vec{0}}_{\in F} + \underbrace{\vec{v}}_{\in G}$. Les sev F et G étant supplémentaires dans E , on en déduit que : $\vec{v} = \vec{0}$. Par conséquent : $F \cap G = \{\vec{0}\}$.

Conclusion : $[E = F \oplus G] \implies [F \cap G = \{\vec{0}\} \text{ et } E = F + G] \text{ (♠)}.$

► Réciproquement : supposons que $F \cap G = \{\vec{0}\}$ et $E = F + G$, et considérons un vecteur arbitraire $\vec{v}$ de E . La seconde hypothèse permet d'affirmer qu'il existe $\vec{f} \in F$ et $\vec{g} \in G$ tels que : $\vec{v} = \vec{f} + \vec{g}$.

Il reste à établir l'unicité de cette écriture. Supposons qu'il existe $\vec{f}$ et $\vec{f}'$ dans F , et $\vec{g}$ et $\vec{g}'$ dans G tels que : $\vec{v} = \vec{f} + \vec{g}$ et $\vec{v} = \vec{f}' + \vec{g}'$. Alors : $\vec{f} - \vec{f}' = \vec{g}' - \vec{g}$. On en déduit que $\vec{f} - \vec{f}' \in F \cap G$. Par hypothèse, ceci implique que : $\vec{f} - \vec{f}' = \vec{0}$. D'où $\vec{f} = \vec{f}'$ et donc $\vec{g} = \vec{g}'$.

Par suite : $\forall \vec{v} \in E, \exists! (\vec{f}, \vec{g}) \in F \times G, \vec{v} = \vec{f} + \vec{g}$. Autrement dit : $E = F \oplus G$.

Conclusion : $[F \cap G = \{\vec{0}\} \text{ et } E = F + G] \implies [E = F \oplus G] \text{ (♣)}.$

Synthèse : d'après (♠) et (♣) : $[F \cap G = \{\vec{0}\} \text{ et } E = F + G] \iff [E = F \oplus G]$.



Application. Soit $E = \mathcal{C}([0; \pi], \mathbb{R})$. On considère les deux sev de E suivants :

$$F = \left\{ f \in E / f(0) = f\left(\frac{\pi}{2}\right) = 0 \right\} \quad \text{et} \quad G = \text{Vect}(\cos, \sin).$$

Montrons que F et G sont supplémentaires dans E .

► Montrons que $F \cap G = \{\mathbf{0}\}$.³

Soit $\varphi \in F \cap G$. Alors : $\exists (a, b) \in \mathbb{R}^2, \varphi = a \cos + b \sin$ (puisque $f \in G$).

$$\text{Comme par ailleurs } \varphi \in F, \text{ on a : } \begin{cases} \varphi(0) = 0 \iff a = 0 \\ \varphi\left(\frac{\pi}{2}\right) = 0 \iff b = 0 \end{cases}$$

On en déduit que $\varphi = \mathbf{0}$. D'où : $F \cap G = \{\mathbf{0}\}$.

► Montrons que $E = F + G$.

Soit $\varphi \in E$. Pour tout couple de réels (a, b) on a : $\varphi = (\varphi - a \cos - b \sin) + (a \cos + b \sin)$.

3. On a noté $\mathbf{0}$ le vecteur nul de E , c'est-à-dire la fonction identiquement nulle sur $[0, \pi]$.

Puisqu'il est clair que $(a \cos + b \sin)$ appartient à G , il suffit de choisir a et b tels que $(\varphi - a \cos - b \sin)$ soit dans F . On cherche donc a et b tels que :

$$\begin{cases} (\varphi - a \cos - b \sin)(0) = 0 \\ (\varphi - a \cos - b \sin)\left(\frac{\pi}{2}\right) = 0 \end{cases} \iff \begin{cases} a = \varphi(0) \\ b = \varphi\left(\frac{\pi}{2}\right) \end{cases}$$

En résumé : $\forall \varphi \in E, \exists (f, g) \in F \times G, \varphi = f + g$ avec
$$\begin{cases} f = \varphi - \varphi(0) \cos - \varphi\left(\frac{\pi}{2}\right) \sin \\ \text{et} \\ g = \varphi(0) \cos + \varphi\left(\frac{\pi}{2}\right) \sin \end{cases}$$

On en déduit que : $E = F + G$.

► On a établi que $E = F + G$ et $F \cap G = \{\mathbf{0}\}$. D'après la caractérisation des sous-espaces supplémentaires, on peut conclure que : $E = F \oplus G$.

21.6.3 Projections et symétries

DÉFINITION 21.13 - Soient E un $\mathbb{K}$ -ev E ; F et G deux sev supplémentaires de E (càd tels que $E = F \oplus G$), de telle sorte que :

$$\forall \vec{v} \in E, \exists! (\vec{f}, \vec{g}) \in F \times G, \vec{v} = \vec{f} + \vec{g}$$

Ces hypothèses et notations étant posées, on appelle :

- **projection sur F parallèlement à G** l'application p_F de E dans E définie par $p_F(\vec{v}) = \vec{f}$;
- **symétrie par rapport à F parallèlement à G** l'application s_F de E dans E définie par $s_F(\vec{v}) = \vec{f} - \vec{g}$

Exemple naïf. L'application $p : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ définie en posant $p(x, y) = (x, 0)$ est la projection sur $\text{Vect}((1, 0))$ parallèlement à $\text{Vect}((0, 1))$.

L'application $s : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ définie en posant $p(x, y) = (x, -y)$ est la symétrie par rapport à $\text{Vect}((1, 0))$ parallèlement à $\text{Vect}((0, 1))$.

PROPRIÉTÉ 21.16 - (Propriétés des projections et symétries). Avec les hypothèses et notations de la définition précédente :

1/ p_F est un endomorphisme de E , et $p_F^2 = p_F$. En outre : $\ker(p_F) = G$ et $\text{Im}(p_F) = F$.

2/ s_F est un endomorphisme de E , et $s_F^2 = \text{id}_E$ (s_F est une involution). En particulier, s_F est un automorphisme de E ($s_F \in \text{GL}(E)$).

PREUVE. Soient F et G deux sev supplémentaires de E (càd tels que $E = F \oplus G$). Par définition (de sev supplémentaires) : $\forall \vec{v} \in E, \exists! (\vec{f}, \vec{g}) \in F \times G, \vec{v} = \vec{f} + \vec{g}$. On rappelle alors que la projection sur F parallèlement à G est l'application p_F de E dans E définie en posant $p_F(\vec{v}) = \vec{f}$.

► Montrons la linéarité de p_F . Soient $\vec{v}$ et $\vec{w}$ deux vecteurs de E , et soient λ et μ deux scalaires. Puisque $E = F \oplus G$, il existe deux vecteurs $\vec{f}$ et $\vec{f}'$ dans F , et deux vecteurs $\vec{g}$ et $\vec{g}'$ dans G uniques tels que : $\vec{v} = \vec{f} + \vec{g}$ et $\vec{w} = \vec{f}' + \vec{g}'$. Notons déjà que : $\vec{f} = p_F(\vec{v})$ et $\vec{f}' = p_F(\vec{w})$ (§).

En outre : $\lambda \vec{v} + \mu \vec{w} = \underbrace{\lambda \vec{f} + \mu \vec{f}'}_{\in F} + \underbrace{\lambda \vec{g} + \mu \vec{g}'}_{\in G}$. D'où : $p_F(\lambda \vec{v} + \mu \vec{w}) = \lambda \vec{f} + \mu \vec{f}'$ (b).

On déduit de (b) et de (§) que : $p_F(\lambda \vec{v} + \mu \vec{w}) = \lambda p_F(\vec{v}) + \mu p_F(\vec{w})$. Ce qui prouve que p_F est linéaire.
D'où $p_F \in \mathcal{L}(E)$.

► Montrons que $p_F^2 = p_F$. Avec les notations précédemment introduites : $p_F(\vec{v}) = \vec{f}$. Et puisque $\vec{f} = \underbrace{\vec{f}}_{\in F} + \underbrace{\vec{0}}_{\in G}$, on a : $p_F^2(\vec{v}) = p_F(\vec{f}) = \vec{f}$. Il s'ensuit que : $\forall \vec{v} \in E, p_F^2(\vec{v}) = p_F(\vec{v})$. Donc : $p_F^2 = p_F$.

► Montrons que $\ker(p_F) = G$. Par double inclusion : si $\vec{g} \in G$, alors $\vec{g} = \underbrace{\vec{0}}_{\in F} + \underbrace{\vec{g}}_{\in G}$, d'où $p_F(\vec{g}) = \vec{0}$,
 càd : $\vec{g} \in \ker(p_F)$. Ainsi : $G \subset \ker(p_F)$.

Réciproquement, si $\vec{v} = \vec{f} + \vec{g}$ (toujours avec les mêmes notations) est dans $\ker(p_F)$, alors $\vec{f} = \vec{0}$, d'où $\vec{v} = \vec{g}$. Par suite : $\vec{v} \in G$, ce qui prouve l'inclusion : $\ker(p_F) \subset G$.
Finalement : $\ker(p_F) = G$.

► Montrons que $\text{Im}(p_F) = F$. Il résulte de la définition que $\text{Im}(p_F) \subset F$. Établissons l'inclusion réciproque. Pour tout $\vec{f} \in F$, on a : $\vec{f} = \vec{f} + \vec{0}$, d'où : $\vec{f} = p_F(\vec{f})$. Par conséquent : $\vec{f} \in \text{Im}(p_F)$, ce qui prouve que : $F \subset \text{Im}(p_F)$.

Finalement : $\text{Im}(p_F) = F$.

Ce travail fait, les assertions concernant s_F sont immédiates. ☕

Remarque. Avec les notations et hypothèses de la propriété précédente, on a :

$$\vec{v} \in \text{Im}(p_F) \iff \vec{v} \in F \iff p_F(\vec{v}) = \vec{v} \iff \vec{v} \in \ker(\text{id}_E - p_F)$$

D'où : $F = \text{Im}(p_F) = \ker(\text{id}_E - p_F)$.

Exemples rigolos. La fonction cosinus hyperbolique est l'image de la fonction exponentielle par la projection sur le sev des fonctions paires parallèlement au sev des fonctions impaires. Et pour une matrice $A \in M_n(\mathbb{K})$, la transposée de A est l'image de A par la symétrie par rapport au sev des matrices symétriques parallèlement au sev des matrices antisymétriques.

21.7 Projecteurs

DÉFINITION 21.14 - Soit E un $\mathbb{K}$ -ev. Un **projecteur** de E est un endomorphisme p de E tel que $p^2 = p$.

Par exemple, une projection (au sens défini dans le paragraphe précédent) est un cas particulier de projecteur. L'objet de ce paragraphe est de prouver qu'il n'y a pas d'autre situation, càd que tout projecteur est une projection ; ceci se fait en trois étapes.

PROPRIÉTÉ 21.17 - (Propriétés des projecteurs). Soit $p \in \mathcal{L}(E)$.

Si p est un projecteur, alors :

1/ $\text{id}_E - p$ est un projecteur

2/ $\text{Im}(p) = \ker(\text{id}_E - p)$

3/ $\ker(p) = \text{Im}(\text{id}_E - p)$

PREUVE. 1/ Soit p un projecteur de E . On a : $(\text{id}_E - p) \circ (\text{id}_E - p) = \text{id}_E - p - p + \underbrace{p^2}_{=p} = \text{id}_E - p$. Ce

qui prouve que $\boxed{\text{id}_E - p \text{ est un projecteur de } E}$.

2/ Soit $\vec{v} \in E$. Supposons que $\vec{v} \in \text{Im}(p)$.

Alors : $\exists \vec{u} \in E, \vec{v} = p(\vec{u})$.

D'où : $p(\vec{v}) = p^2(\vec{u})$ et puisque p est un projecteur : $p(\vec{v}) = p(\vec{u})$.

Il s'ensuit que : $p(\vec{v}) = \vec{v}$, càd que : $\vec{v} \in \ker(\text{id}_E - p)$. Ce qui prouve que : $\text{Im}(p) \subset \ker(\text{id}_E - p)$.

Etablissons l'inclusion réciproque : soit $\vec{v} \in \ker(\text{id}_E - p)$.

Alors : $(\text{id}_E - p)(\vec{v}) = \vec{0}$, d'où : $\text{id}_E(\vec{v}) - p(\vec{v}) = \vec{0}$, donc : $\vec{v} = p(\vec{v})$. En particulier : $\vec{v} \in \text{Im}(p)$.

Ce qui prouve que : $\ker(\text{id}_E - p) \subset \text{Im}(p)$.

On déduit de ce qui précède que : $\boxed{\ker(\text{id}_E - p) = \text{Im}(p)}$.

3/ D'après le point 1, l'endomorphisme $\text{id}_E - p$ est un projecteur de E .

On peut donc lui appliquer le résultat du 2 pour affirmer que : $\ker(\text{id}_E - (\text{id}_E - p)) = \text{Im}(\text{id}_E - p)$, càd que : $\boxed{\ker(p) = \text{Im}(\text{id}_E - p)}$. ☕

Remarque. Pour donner un peu de sens aux propriétés ci-dessus, on pourra observer que dans la situation du paragraphe précédent (où $E = F \oplus G$), la projection p_F est un projecteur, et $\text{id}_E - p_F = p_G$ (qui est effectivement un autre projecteur). En continuant de considérer $p = p_F$, on a alors $\text{Im}(p) = F = \ker(p_G)$; et $\ker(p) = G = \text{Im}(p_G)$.

PROPRIÉTÉ 21.18 - (Caractérisation des projecteurs). Soit $p \in \mathcal{L}(E)$.

LASSE :

1/ p est un projecteur

2/ $\text{id}_E - p$ est un projecteur

3/ $E = \ker(p) \oplus \ker(\text{id}_E - p)$

PREUVE. Montrons que $(1) \Leftrightarrow (2)$ et que $(1) \Leftrightarrow (3)$.

► D'après la propriété précédente, si p est un projecteur, alors $\text{id}_E - p$ est un projecteur; ce qui prouve l'implication $(1) \Rightarrow (2)$. On obtient la réciproque en remplaçant p par $\text{id}_E - p$ dans ce raisonnement. D'où :

$\boxed{(1) \Leftrightarrow (2)}$.

► Montrons que (1) $\Rightarrow$ (3). Supposons que p soit un projecteur. Soit $\vec{v}$ un élément de E .

On écrit judicieusement : $\vec{v} = (\vec{v} - p(\vec{v})) + p(\vec{v})$ (♠).

Notons alors que : $p(\vec{v} - p(\vec{v})) = p(\vec{v}) - p^2(\vec{v}) = p(\vec{v}) - p(\vec{v}) = \vec{0}$.⁴ D'où : $\vec{v} - p(\vec{v}) \in \ker(p)$ (♡).

Par ailleurs : $p(p(\vec{v})) = p(\vec{v})$ (puisque p est un projecteur). Il revient au même d'écrire : $p(p(\vec{v})) = \text{id}_E(p(\vec{v}))$, d'où : $(p - \text{id}_E)(p(\vec{v})) = \vec{0}$. Par suite : $p(\vec{v}) \in \ker(\text{id}_E - p)$ (♣).

On déduit de (♠), (♡) et (♣) que : $E = \ker(p) + \ker(\text{id}_E - p)$ (◇).

En outre, supposons que $\vec{w} \in \ker(p) \cap \ker(\text{id}_E - p)$, alors $p(\vec{w}) = \vec{0}$ (puisque $\vec{w} \in \ker(p)$) et $p(\vec{w}) = \vec{w}$ (puisque $\vec{w} \in \ker(\text{id}_E - p)$). D'où $\vec{w} = \vec{0}$. Il s'ensuit que : $\ker(p) \cap \ker(\text{id}_E - p) = \{\vec{0}\}$ (b).

D'après (◇), (b) et la caractérisation des sev supplémentaires, on peut conclure : $E = \ker(p) \oplus \ker(\text{id}_E - p)$.

► Montrons réciproquement que (3) $\Rightarrow$ (1). Supposons que $E = \ker(p) \oplus \ker(\text{id}_E - p)$, et montrons que p est un projecteur.

Soit $\vec{v}$ un élément de E . Par hypothèse, il existe $\vec{f} \in \ker(p)$ et $\vec{g} \in \ker(\text{id}_E - p)$ uniques tels que : $\vec{v} = \vec{f} + \vec{g}$. On a alors : $p(\vec{v}) = p(\vec{f}) + p(\vec{g})$, puis : $p(\vec{v}) = \vec{g}$.

Il s'ensuit que : $p^2(\vec{v}) = p(\vec{g}) = \vec{g}$, càd : $p^2(\vec{v}) = \vec{g}$.

Par conséquent : $\forall \vec{v} \in E, p^2(\vec{v}) = p(\vec{v})$. Donc : $p^2 = p$, ce qui signifie que p est un projecteur, et achève la preuve l'implication (3) $\Rightarrow$ (1), donc de l'équivalence (3) $\Leftrightarrow$ (1), et de la propriété. ☕

THÉORÈME 21.10 - Soient E un $\mathbb{K}$ -ev, et $p \in \mathcal{L}(E)$.

Si p est un projecteur de E , alors p est la projection sur $\text{Im}(p)$ parallèlement à $\ker(p)$.

PREUVE. Soit p un projecteur de E . D'après la question de cours 6, on a : $E = \ker(p) \oplus \ker(\text{id}_E - p)$, et d'après la propriété 21.18 : $\ker(\text{id}_E - p) = \text{Im}(p)$.

Soit $\vec{v} \in E$. Alors il existe $\vec{f} \in \ker(p)$ et $\vec{g} \in \ker(\text{id}_E - p) = \text{Im}(p)$ uniques tels que : $\vec{v} = \vec{f} + \vec{g}$ (♠).

On a alors : $p(\vec{v}) = p(\vec{f}) + p(\vec{g})$, puis : $p(\vec{v}) = \vec{g}$.

Par ailleurs, notons π la projection sur $\ker(\text{id}_E - p) = \text{Im}(p)$ parallèlement à $\ker(p)$ (cette projection est bien définie puisque $\ker(p)$ et $\ker(\text{id}_E - p)$ sont supplémentaires dans E). Par définition (de projection), on a alors d'après (♠) : $\pi(\vec{v}) = \vec{g}$.

Il s'ensuit que : $\forall \vec{v} \in E, p(\vec{v}) = \pi(\vec{v})$, ce qui prouve que $p = \pi$. Ainsi le projecteur p est la projection sur $\text{Im}(p)$ parallèlement à $\ker(p)$, ce qu'il fallait démontrer. ☕

4. La première égalité provenant de la linéarité de p , et la seconde de l'hypothèse selon laquelle p est un projecteur.

21.8 Epilogue — Etude des symétries dans un espace vectoriel

Au cours du chapitre, nous avons étudié en cours les propriétés des projections dans un espace vectoriel. Cette étude s'est faite en deux temps. En premier lieu, on a introduit la notion de projection sur un sev parallèlement à un autre, et étudié les propriétés d'une telle projection. Dans un second temps, nous avons cherché à caractériser les projections (par le biais de la notion de projecteur).

L'objectif majeur de ce problème est de faire une étude analogue dans le cadre des symétries.

Rappel de cours. Soit E un $\mathbb{K}$ -ev E ; soient F et G deux sev supplémentaires de E (càd tels que $E = F \oplus G$). Ainsi : $\forall \vec{v} \in E, \exists! (\vec{f}, \vec{g}) \in F \times G, \vec{v} = \vec{f} + \vec{g}$. Ces hypothèses et notations étant posées, on appelle :

- **projection sur F parallèlement à G** l'application p_F de E dans E définie en posant $p_F(\vec{v}) = \vec{f}$;
- **symétrie par rapport à F parallèlement à G** l'application s_F de E dans E définie en posant $s_F(\vec{v}) = \vec{f} - \vec{g}$.

Partie I — Deux exemples

- 1) Dans cette question, E désigne le $\mathbb{R}$ -espace vectoriel $M_2(\mathbb{R})$ des matrices carrées de taille 2 à coefficients réels.

On pose $F = \text{Vect}(E_{11})$, et G désigne l'ensemble des matrices de E de trace nulle.

- a) Montrer que G est un sev de E , et en déterminer une famille génératrice.
- b) Montrer que F et G sont supplémentaires dans E .
- c) Donner l'expression de la projection p_F sur F parallèlement à G , puis celle de la projection p_G sur G parallèlement à F .

Dans les deux cas, il s'agit de calculer l'image d'une matrice $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ par l'application étudiée.

- d) Que vaut $p_F \circ p_G$?
- e) Donner l'expression de la symétrie s_F par rapport à F parallèlement à G .
- f) Donner l'expression de l'endomorphisme de E : $\varphi = s_F - \text{id}_E$. Quel lien y a-t-il entre φ et p_G ?

- 2) Dans cette question, E désigne le $\mathbb{R}$ -espace vectoriel $\mathbb{R}_2[X]$ des polynômes à coefficients réels et de degré au plus 2.

On pose $F = \text{Vect}(X + 1)$, et G désigne l'ensemble des polynômes P de E tels que : $\int_{-1}^1 P(t) dt = 0$.

- a) Montrer que G est un sev de E , et en déterminer une famille génératrice.
- b) Montrer que F et G sont supplémentaires dans E .
- c) Donner l'expression de la projection p_F sur F parallèlement à G , puis celle de la symétrie s_F par rapport à F parallèlement à G .
- d) Expliciter l'image du polynôme X par p_F .

Partie II — Généralités sur les projections et symétries

Tout au long de cette partie, E désigne un $\mathbb{K}$ -ev, et F et G deux sev supplémentaires dans E (on suppose donc $E = F \oplus G$).

- 1) Montrer que l'application s_F est un endomorphisme de E .
- 2) Etablir une relation entre les endomorphismes $s_F - \text{id}_E$ et p_G .
- 3) De même, établir une relation entre les endomorphismes $s_F + \text{id}_E$ et p_F .
- 4) Déterminer $\ker(s_F - \text{id}_E)$ et $\ker(s_F + \text{id}_E)$.

Partie III — Réflexions dans un espace vectoriel

Tout au long de cette partie, E désigne un $\mathbb{K}$ -ev. On introduit alors une nouvelle définition : un endomorphisme s de E est appelée une **réflexion** si $s^2 = \text{id}_E$.⁵

Soit s un endomorphisme de E .

- 1) Montrer que si s est une réflexion, alors $E = \ker(s - \text{id}_E) \oplus \ker(s + \text{id}_E)$.
- 2) Réciproquement, montrer que si $E = \ker(s - \text{id}_E) \oplus \ker(s + \text{id}_E)$, alors l'endomorphisme s est une réflexion.
- 3) On pose $p = \frac{1}{2}(s + \text{id}_E)$. Montrer que p est un projecteur si et seulement si s est une réflexion.

Partie IV — Epilogue : toute réflexion est une symétrie

Dans cette partie, E désigne toujours un $\mathbb{K}$ -ev, et s un endomorphisme de E .

Montrer que si s est une réflexion, alors s est une symétrie par rapport à F parallèlement à G , où F et G sont deux sev de E que l'on précisera.

CORRIGÉ.

Partie I — Deux exemples

- 1) a) Soit $M \in M_2(\mathbb{R})$. Alors :

$$[M \in G] \iff \left[\exists (a, b, c) \in \mathbb{R}^3, M = \begin{pmatrix} a & b \\ c & -a \end{pmatrix} \right] \iff [\exists (a, b, c) \in \mathbb{R}^3, M = a(E_{11} - E_{22}) + bE_{12} + cE_{21}]$$

On en déduit que : $G = \text{Vect}((E_{11} - E_{22}), E_{12}, E_{21})$ (en particulier G est un sev de $M_2(\mathbb{R})$).

- b) Soit $M \in M_2(\mathbb{R})$. Il existe quatre réels a, b, c et d tels que : $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. On cherche à présent quatre réels x, y, z et t tels que : $M = xE_{11} + y(E_{11} - E_{22}) + zE_{12} + tE_{21}$. Cette égalité entre matrices donne lieu au système :

$$\begin{cases} x + y = a \\ z = b \\ t = c \\ -y = d \end{cases} \quad \text{dont la résolution aisée donne : } \begin{cases} x = a + d \\ z = b \\ t = c \\ y = -d \end{cases}$$

5. Les symétries sont des exemples de réflexion.

Ainsi, la matrice M peut s'écrire comme somme d'une matrice dans F et d'une matrice dans G , explicitement :

$$M = \underbrace{\begin{pmatrix} a+d & 0 \\ 0 & 0 \end{pmatrix}}_{\in F} + \underbrace{\begin{pmatrix} -d & b \\ c & d \end{pmatrix}}_{\in G}$$

En outre, cette écriture est unique (unicité de la solution du système précédent). Par suite : $\boxed{M_2(\mathbb{R}) = F \oplus G}$.

c) D'après la question précédente, les projections p_F et p_G sont respectivement données par :

$$\forall M = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_2(\mathbb{R}), \quad p_F(M) = \begin{pmatrix} a+d & 0 \\ 0 & 0 \end{pmatrix} \wedge p_G(M) = \begin{pmatrix} -d & b \\ c & d \end{pmatrix}.$$

d) D'après les propriétés des projections, on a : $\text{Im}(p_G) = G$ et $\ker(p_F) = G$. Il s'ensuit que : $\boxed{p_F \circ p_G = 0_{\mathcal{L}(E)}}$.

e) D'après la question b) :

$$\forall M = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_2(\mathbb{R}), \quad s_F(M) = \begin{pmatrix} a+d & 0 \\ 0 & 0 \end{pmatrix} - \begin{pmatrix} -d & b \\ c & d \end{pmatrix} \text{ soit : } \boxed{s_F(M) = \begin{pmatrix} a+2d & -b \\ -c & -d \end{pmatrix}}$$

f) D'après la question précédente, pour toute matrice $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_2(\mathbb{R})$ on a :

$$\varphi(M) = (s_F - \text{id}_E)(M) = s_F(M) - M = \begin{pmatrix} 2d & -2b \\ -2c & -2d \end{pmatrix}. \text{ Par suite : } \boxed{\varphi = s_F - \text{id}_E = -2p_G}.$$

2) a) Soit $P \in \mathbb{R}_2[X]$. Il existe trois réels a, b et c tels que : $P = aX^2 + bX + c$. Alors :

$$[P \in G] \iff \left[\int_{-1}^1 at^2 + bt + c dt = 0 \right] \iff \left[\int_0^1 at^2 + c dt = 0 \right] \iff [a = -3c]$$

On en déduit que : $[P \in G] \iff [\exists (b, c) \in \mathbb{R}^2, P = c(3X^2 - 1) + bX]$. Par suite : $\boxed{G = \text{Vect}(X, 3X^2 - 1)}$.

b) Soit $P \in \mathbb{R}_2[X]$. Il existe trois réels a, b et c tels que : $P = aX^2 + bX + c$. On cherche alors trois réels x, y et z tels que : $P = x(X+1) + yX + z(3X^2 - 1)$, c'est-à-dire tels que : $P = 3zX^2 + (x+y)X + x - z$. L'identification entre les deux écritures de P fournit alors le système :

$$\begin{cases} 3z = a \\ x + y = b \\ x - z = c \end{cases} \quad \text{dont la résolution donne : } \begin{cases} z = a/3 \\ y = b - c - (a/3) \\ x = c + (a/3) \end{cases}$$

Il s'ensuit que P peut s'écrire comme somme d'un polynôme de F et d'un polynôme de G explicitement :

$$P = \underbrace{\left(c + \frac{a}{3}\right)(X+1)}_{\in F} + \underbrace{\left(b - c - \frac{a}{3}\right)X + \frac{a}{3}(3X^2 - 1)}_{\in G}$$

Comme en outre cette écriture est unique (d'après les calculs précédents), on peut conclure : $\boxed{\mathbb{R}_2[X] = F \oplus G}$.

c) D'après la question précédente, et d'après les définitions respectives de p_F et de s_F , on a :

$$\boxed{\begin{array}{l} p_F : \mathbb{R}_2[X] \longrightarrow \mathbb{R}_2[X] \\ aX^2 + bX + c \longmapsto \left(c + \frac{a}{3}\right)(X+1) \end{array}}$$

$$\text{et } \boxed{\begin{array}{l} s_F : \mathbb{R}_2[X] \longrightarrow \mathbb{R}_2[X] \\ aX^2 + bX + c \longmapsto \left(c + \frac{a}{3}\right)(X+1) - \left(b - c - \frac{a}{3}\right)X + \frac{a}{3}(3X^2 - 1) \end{array}}$$

d) La projection sur F parallèlement au sev G a pour noyau le sev G . Puisque $X \in G$, on en déduit : $p_F(X) = 0$.

Partie II — Généralités sur les projections et symétries

Tout au long de cette partie, E désigne un $\mathbb{K}$ -ev, et F et G deux sev supplémentaires dans E (on suppose donc $E = F \oplus G$).

1) Soient $\vec{u}$ et $\vec{v} \in E$, et soient λ et μ deux scalaires. Par hypothèse, il existe $\vec{f}$ et $\vec{f}'$ dans F , $\vec{g}$ et $\vec{g}'$ dans G uniques tels que : $\vec{u} = \vec{f} + \vec{g}$ et $\vec{v} = \vec{f}' + \vec{g}'$. Alors on a : $\lambda \vec{u} + \mu \vec{v} = \underbrace{(\lambda \vec{f} + \mu \vec{f}')}_{\in F} + \underbrace{(\lambda \vec{g} + \mu \vec{g}')}_{\in G}$.

D'où :

$$s_F(\lambda \vec{u} + \mu \vec{v}) = (\lambda \vec{f} + \mu \vec{f}') - (\lambda \vec{g} + \mu \vec{g}') = \lambda (\vec{f} - \vec{g}) + \mu (\vec{f}' - \vec{g}') = \lambda s_F(\vec{u}) + \mu s_F(\vec{v})$$

Conclusion. $s_F \in \mathcal{L}(E)$.

2) Soit $\vec{v} \in E$. Par hypothèse : $\exists! (\vec{f}, \vec{g}) \in F \times G$, $\vec{v} = \vec{f} + \vec{g}$. On a alors :

$(s_F - \text{id}_E)(\vec{v}) = s_F(\vec{v}) - \vec{v} = \vec{f} - \vec{g} - \vec{f} - \vec{g} = -2\vec{g}$. Par ailleurs : $p_G(\vec{v}) = \vec{g}$. On en déduit que : $s_F - \text{id}_E = -2p_G$.

3) Soit $\vec{v} \in E$. Par hypothèse : $\exists! (\vec{f}, \vec{g}) \in F \times G$, $\vec{v} = \vec{f} + \vec{g}$. On a alors :

$(s_F + \text{id}_E)(\vec{v}) = s_F(\vec{v}) + \vec{v} = \vec{f} - \vec{g} + \vec{f} + \vec{g} = 2\vec{f}$. Par ailleurs : $p_F(\vec{v}) = \vec{f}$. On en déduit que : $s_F + \text{id}_E = 2p_F$.

4) D'après les deux questions précédentes : $\ker(s_F - \text{id}_E) = F$ et $\ker(s_F + \text{id}_E) = G$ (on peut alors observer que : $E = \ker(s_F - \text{id}_E) \oplus \ker(s_F + \text{id}_E)$).

Partie III — Réflexions dans un espace vectoriel

Tout au long de cette partie, E désigne un $\mathbb{K}$ -ev. On introduit alors une nouvelle définition : un endomorphisme s de E est appelée une **réflexion** si $s^2 = \text{id}_E$.⁶

1) Supposons que s soit une réflexion ; on a donc $s \in \mathcal{L}(E)$ et $s^2 = \text{id}_E$. Soit $\vec{v}$ un vecteur de E . Procédons par analyse-synthèse pour établir que $\ker(s - \text{id}_E)$ et $\ker(s + \text{id}_E)$ sont supplémentaires dans E .

► Analyse : supposons qu'il existe $\vec{f} \in \ker(s - \text{id}_E)$ et $\vec{g} \in \ker(s + \text{id}_E)$ tels que $\vec{v} = \vec{f} + \vec{g}$. Alors : $s(\vec{v}) = s(\vec{f}) + s(\vec{g})$ d'où $s(\vec{v}) = \vec{f} - \vec{g}$. On a donc :

$$\begin{cases} \vec{v} = \vec{f} + \vec{g} \\ s(\vec{v}) = \vec{f} - \vec{g} \end{cases} \quad \text{d'où :} \quad \begin{cases} \vec{f} = \frac{1}{2}(\vec{v} + s(\vec{v})) \\ \vec{g} = \frac{1}{2}(\vec{v} - s(\vec{v})) \end{cases}$$

► Synthèse : on écrit donc $\vec{v} = \underbrace{\frac{1}{2}(\vec{v} + s(\vec{v}))}_{=\vec{v}_1} + \underbrace{\frac{1}{2}(\vec{v} - s(\vec{v}))}_{=\vec{v}_2}$.

$$\text{On a : } s(\vec{v}_1) = s\left(\frac{1}{2}(\vec{v} + s(\vec{v}))\right) = \frac{1}{2}\left(s(\vec{v}) + \underbrace{s^2(\vec{v})}_{=\vec{v}}\right) = \frac{1}{2}(\vec{v} + s(\vec{v})) = \vec{v}_1.$$

6. Les symétries sont des exemples de réflexion, et comme dans le cours sur les projecteurs, l'objectif de la fin du problème est d'établir qu'il n'existe pas d'autres réflexions que les symétries (c'est précisément l'objet de l'unique question de la partie IV).

D'où : $\boxed{\vec{v}_1 \in \ker(s - \text{id}_E) \text{ (♠)}}$

$$\text{Et : } s(\vec{v}_2) = s\left(\frac{1}{2}(\vec{v} - s(\vec{v}))\right) = \frac{1}{2}\left(s(\vec{v}) - \underbrace{s^2(\vec{v})}_{=\vec{v}}\right) = -\frac{1}{2}(\vec{v} - s(\vec{v})) = -\vec{v}_2.$$

D'où : $\boxed{\vec{v}_2 \in \ker(s + \text{id}_E) \text{ (♣)}}$

On déduit alors de (♠) et de (♣) que : $\boxed{E = \ker(s - \text{id}_E) + \ker(s + \text{id}_E) \text{ (♡)}}$.

Enfin, soit $\vec{v} \in \ker(s - \text{id}_E) \cap \ker(s + \text{id}_E)$. Alors : $s(\vec{v}) = \vec{v}$ et $s(\vec{v}) = -\vec{v}$ d'où $\vec{v} = -\vec{v}$ et par suite : $\vec{v} = \vec{0}$. Donc : $\boxed{\ker(s - \text{id}_E) \cap \ker(s + \text{id}_E) = \{\vec{0}\} \text{ (◇)}}$.

On déduit alors de (♡), (◇) et de la caractérisation des sous-espaces vectoriels supplémentaires que :

$$E = \ker(s - \text{id}_E) \oplus \ker(s + \text{id}_E).$$

Conclusion. Soit $s \in \mathcal{L}(E)$. Alors : $\boxed{[s^2 = \text{id}_E] \implies [E = \ker(s - \text{id}_E) \oplus \ker(s + \text{id}_E)]}$.

2) Réciproquement, soit s un endomorphisme de E tel que : $E = \ker(s - \text{id}_E) \oplus \ker(s + \text{id}_E)$. Soit $\vec{v} \in E$. Par hypothèse il existe alors $\vec{f} \in \ker(s - \text{id}_E)$ et $\vec{g} \in \ker(s + \text{id}_E)$ uniques tels que : $\vec{v} = \vec{f} + \vec{g}$.

$$\text{On a alors : } s(\vec{v}) = s(\vec{f} + \vec{g}) = s(\vec{f}) + s(\vec{g}) = \vec{f} - \vec{g}.$$

$$\text{Puis : } s^2(\vec{v}) = s(\vec{f} - \vec{g}) = s(\vec{f}) - s(\vec{g}) = \vec{f} + \vec{g} = \vec{v}.$$

En résumé : $\forall \vec{v} \in E, s^2(\vec{v}) = \vec{v}$. D'où $s^2 = \text{id}_E$; s est donc une réflexion.

Conclusion. Soit $s \in \mathcal{L}(E)$. Alors : $\boxed{[E = \ker(s - \text{id}_E) \oplus \ker(s + \text{id}_E)] \implies [s^2 = \text{id}_E]}$.

3) Soit s un endomorphisme de E , et posons $p = \frac{1}{2}(s + \text{id}_E)$.

$$[p \text{ est un projecteur}] \iff [p^2 = p] \iff \left[\frac{1}{4}(s + \text{id}_E)^2 = \frac{1}{2}(s + \text{id}_E)\right] \iff \left[\frac{1}{4}(s^2 + 2s + \text{id}_E) = \frac{1}{2}(s + \text{id}_E)\right]$$

$$\iff [s^2 + \text{id}_E = 2\text{id}_E] \iff [s^2 = \text{id}_E] \iff [s \text{ est une réflexion}]$$

Conclusion. Pour $s \in \mathcal{L}(E)$, posons $p = \frac{1}{2}(s + \text{id}_E)$. On a : $[p \text{ est un projecteur}] \iff [s \text{ est une réflexion}]$.

En résumé, on a établi dans cette partie une propriété caractérisant les réflexions (c'est le pendant de la propriété caractérisant les projecteurs démontrée en cours), explicitement :

PROPRIÉTÉ (caractérisation des réflexions). Soit $s \in \mathcal{L}(E)$. LASSE :

- 1) s est une réflexion (càd $s^2 = \text{id}_E$)
- 2) $E = \ker(s - \text{id}_E) \oplus \ker(s + \text{id}_E)$
- 3) $p = \frac{1}{2}(s + \text{id}_E)$ est un projecteur de E (càd $p^2 = p$)

Partie IV — Epilogue : toute réflexion est une symétrie

Dans cette question, on suppose que s est une réflexion. D'après la question 1 de la partie précédente, on a alors : $E = \ker(s - \text{id}_E) \oplus \ker(s + \text{id}_E)$.

Soit $\vec{v} \in E$. Il existe alors $\vec{f} \in \ker(s - \text{id}_E)$ et $\vec{g} \in \ker(s + \text{id}_E)$ uniques tels que : $\vec{v} = \vec{f} + \vec{g}$. D'où : $s(\vec{v}) = \vec{f} - \vec{g}$.

Par ailleurs, en notant σ la symétrie par rapport à $F = \ker(s - \text{id}_E)$ parallèlement à $G = \ker(s + \text{id}_E)$ (qui est bien définie puisque $E = F \oplus G$), on a : $\sigma(\vec{v}) = \vec{f} - \vec{g}$.

D'où : $\forall \vec{v} \in E, s(\vec{v}) = \sigma(\vec{v})$. Donc : $s = \sigma$.

Conclusion. Soit $s \in \mathcal{L}(E)$. Si s est une réflexion, alors s est la symétrie par rapport à $\ker(s - \text{id}_E)$ parallèlement à $\ker(s + \text{id}_E)$.

Chapitre 22

Espaces vectoriels de dimension finie

Dans ce chapitre, E désigne un $\mathbb{K}$ -espace vectoriel, avec $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$. Dorénavant, on pourra noter v au lieu de $\vec{v}$ un élément de E .

22.1 Familles libres et bases d'un espace vectoriel

DÉFINITION 22.1 - Deux vecteurs u et v d'un $\mathbb{K}$ -espace vectoriel E sont **colinéaires** s'il existe un scalaire λ tel que $u = \lambda v$, ou s'il existe un scalaire λ tel que $v = \lambda u$.

PROPRIÉTÉ 22.1 - Deux vecteurs u et v non nuls d'un $\mathbb{K}$ -espace vectoriel E sont colinéaires SSI il existe $(\alpha, \beta) \in \mathbb{K}^2$, avec $(\alpha, \beta) \neq (0, 0)$ tels que $\alpha u + \beta v = 0_E$.

PREUVE. Si u et v sont colinéaires, on peut supposer (quitte à renommer u et v) supposer qu'il existe un scalaire λ tel que $u = \lambda v$. Alors : $1u - \lambda v = 0$, ce qui entraîne la seconde assertion (le couple $(1, -\lambda)$ étant non-nul).

Réciproquement, supposons qu'il existe $(\alpha, \beta) \in \mathbb{K}^2$, avec $(\alpha, \beta) \neq (0, 0)$ tels que $\alpha u + \beta v = 0_E$. Si $\alpha \neq 0$, alors $u = \lambda v$ avec $\lambda = -\beta/\alpha$. Et si $\alpha = 0$, alors $\beta \neq 0$, et on déduit alors que $v = \lambda u$ avec $\lambda = -\alpha/\beta$. Dans les deux cas, on a établi que u et v sont colinéaires, ce qui achève la preuve de cette réciproque, et de la propriété. ☕

Remarque. A l'opposé, u et v sont non colinéaires si et seulement si la relation $\alpha u + \beta v = 0_E$ implique $\alpha = 0$ et $\beta = 0$.

DÉFINITION 22.2 - Soit n un entier naturel non nul.

n vecteurs $v_1, \dots, v_n$ d'un $\mathbb{K}$ -ev E sont **linéairement indépendants** si :

$$\left[\sum_{i=1}^n \lambda_i v_i = 0_E \right] \implies [\forall i \in [1; n], \lambda_i = 0] \quad (\text{avec } (\lambda_1, \dots, \lambda_n) \in \mathbb{K}^n)$$

Remarque. La définition fournit une méthode pour prouver que n vecteurs $v_1, \dots, v_n$ d'un $\mathbb{K}$ -ev E sont linéairement indépendants. Explicitement, on suppose qu'il existe n scalaires $\lambda_1, \dots, \lambda_n$ tels que $\lambda_1 v_1 + \dots + \lambda_n v_n = 0_E$, et on établit que cette égalité n'est satisfaite que lorsque tous les λ_i sont nuls.

Exemples. Deux vecteurs sont linéairement indépendants SSI ils sont non colinéaires.

Si n (avec $n \geq 3$) vecteurs sont linéairement indépendants alors ils sont deux à deux non colinéaires. ATTENTION : la réciproque de cette dernière affirmation est FAUSSE. Dans $\mathbb{R}^2$, les trois vecteurs $u(1, 0)$, $v(0, 1)$ et $w(1, 1)$ sont deux à deux non colinéaires ; mais u , v et w ne sont pas linéairement indépendants puisque $w = u + v$.

DÉFINITION 22.3 - Soient $v_1, \dots, v_n$ n vecteurs d'un $\mathbb{K}$ -ev E .

La famille $\mathcal{F} = (v_i)_{i \in [1; n]}$ est **libre** si les vecteurs $v_1, \dots, v_n$ sont linéairement indépendants ; dans le cas contraire on dit que $\mathcal{F}$ est **liée**.

PROPRIÉTÉ 22.2 - Soient $v_1, \dots, v_n$ n vecteurs d'un $\mathbb{K}$ -ev E .

La famille $\mathcal{F} = (v_i)_{i \in [1; n]}$ est **liée** SSI il existe un entier $i \in [1, n]$ tel que : $v_i \in \text{Vect}(v_1, \dots, \widehat{v_i}, \dots, v_n)$.

PREUVE. Raisonnement par double implication.

► Sens direct. Supposons que la famille $\mathcal{F} = (v_i)_{i \in [1; n]}$ est liée. Alors il existe $(\alpha_1, \dots, \alpha_n) \in \mathbb{K}^n \setminus \{0_{\mathbb{K}^n}\}$, $\sum_{i=1}^n \alpha_i v_i = 0$. Puisque le n -uplet est non nul, il existe $i \in [1, n]$ tel que $\alpha_i \neq 0$. On peut alors écrire :

$$\alpha_i v_i = - \sum_{k=1, k \neq i}^n \alpha_k v_k \quad \text{d'où} \quad v_i = \sum_{k=1, k \neq i}^n \beta_k v_k \quad \text{en ayant posé : } \beta_k = -\frac{\alpha_k}{\alpha_i}.$$

Par suite : $v_i \in \text{Vect}(v_1, \dots, \widehat{v_i}, \dots, v_n)$ ce qui prouve la première implication.

► Réciproque. Supposons qu'il existe $i \in [1, n]$ tel que $v_i \in \text{Vect}(v_1, \dots, \widehat{v_i}, \dots, v_n)$. Alors il existe $(n-1)$ scalaires $\alpha_1, \dots, \alpha_i, \dots, \alpha_n$ tels que : $v_i = \sum_{k=1, k \neq i}^n \alpha_k v_k$. Il existe donc n scalaires $\beta_1, \dots, \beta_n$ non tous nuls tels que $\sum_{k=1}^n \alpha_k v_k = 0$ (en ayant posé $\beta_k = \alpha_k$ pour tout $k \neq i$, et $\alpha_i = -1$). D'où la famille $\mathcal{F}$ est liée, ce qui achève la preuve de la seconde implication.

Conclusion. $\mathcal{F} = (v_i)_{i \in [1; n]}$ est liée

SSI il existe un entier $i \in [1, n]$ tel que : $v_i \in \text{Vect}(v_1, \dots, \widehat{v_i}, \dots, v_n)$ ☕

Exemples :

- soit v un vecteur de E . La famille $\mathcal{F} = (v)$ est libre SSI $v \neq 0_E$.
- dans $\mathbb{R}^2$, la famille $\mathcal{F} = (v_1, v_2)$ est libre SSI $\det(v_1, v_2) \neq 0$.
- dans $\mathbb{R}^3$: la famille $\mathcal{F} = (u(1, 2, 3), v(4, 5, 6), w(7, 8, 9))$ est liée puisque $w = 2v - u$. Et la famille $\mathcal{G} = (v_1, v_2)$ est libre SSI $v_1 \wedge v_2 \neq \vec{0}$
- dans $\mathbb{K}_2[X]$, la famille $\mathcal{F} = (1, X, X^2)$ est (très clairement) libre.

La famille $\mathcal{G} = ((X-1)(X-2), X(X-2), X(X-1))$ est (un peu moins clairement) libre, puisque les 3 polynômes qui la constituent sont (à des constantes multiplicatives près) les polynômes d'interpolation de Lagrange de degré 2 associés aux valeurs 0, 1 et 2.

- dans $\mathbb{K}[X]$, on considère n polynômes $P_1, \dots, P_n$ non nuls.

On suppose que : $\deg(P_1) < \deg(P_2) < \dots < \deg(P_n)$. Alors la famille $\mathcal{F} = (P_1, \dots, P_n)$ est libre (toute famille échelonnée de polynômes est libre).

- dans $M_2(\mathbb{K})$, la famille $\mathcal{F} = \left(\underbrace{\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}}_{E_{11}}, \underbrace{\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}}_{E_{12}}, \underbrace{\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}}_{E_{21}}, \underbrace{\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}}_{E_{22}} \right)$ est libre.

- dans $\mathcal{C}^0(\mathbb{R}, \mathbb{R})$, la famille $\mathcal{F} = (\text{ch}, \text{sh}, \exp)$ est liée puisque $\exp = \text{ch} + \text{sh}$; mais la famille $\mathcal{G} = (\text{ch}, \text{sh})$ est libre.

DÉFINITION 22.4 - Une **base** d'un espace vectoriel E est une famille génératrice et libre de E .

Exemples :

- la famille $\mathcal{F} = (e_1(1, 0); e_2(0, 1))$ est une base de $\mathbb{R}^2$, appelée **base canonique de $\mathbb{R}^2$** . Définitions analogues pour les bases canoniques de $\mathbb{R}^3$, $\mathbb{R}^n$ et $\mathbb{K}^n$.
- la famille $\mathcal{F} = (E_{11}, E_{12}, E_{21}, E_{22})$ est la **base canonique de $M_2(\mathbb{K})$** .
- la famille $\mathcal{F} = (1, X, X^2)$ est une base de $\mathbb{K}_2[X]$, appelée **base canonique de $\mathbb{K}_2[X]$** ; plus généralement, la famille $(1, X, \dots, X^n)$ est la **base canonique de $\mathbb{K}_n[X]$** .
- La famille $\mathcal{G} = ((X-1)(X-2), X(X-2), X(X-1))$ est une autre base de $\mathbb{K}_2[X]$.

22.2 Espaces vectoriels de dimension finie

22.2.1 Dimension

DÉFINITION 22.5 - On appelle **$\mathbb{K}$ -espace vectoriel de dimension finie** un $\mathbb{K}$ -espace vectoriel E ayant une famille génératrice finie.

Exemples : $\mathbb{R}$, $\mathbb{R}^2$, $\mathbb{R}^n$, $\mathbb{K}^n$, $M_{np}(\mathbb{K})$, $\mathbb{K}_n[X]$ sont des espaces vectoriels de dimension finie.

En revanche, $\mathbb{K}[X]$ et $\mathcal{C}^0(\mathbb{R}, \mathbb{R})$ (entre autres exemples) ne sont pas des espaces vectoriels de dimension finie.

La notion de dimension repose essentiellement sur les deux lemmes suivants.

LEMME 22.1 - Soit E un $\mathbb{K}$ -ev.

- 1/ Toute famille de vecteurs de E contenue dans une famille libre est libre.
- 2/ Toute famille de vecteurs de E contenant une famille génératrice est génératrice.

PREUVE. Triviale. ☕

LEMME 22.2 - Dans un espace vectoriel E engendré par n vecteurs, le cardinal d'une famille libre est majoré par n .

PREUVE. Une formulation équivalente de ce lemme est : dans un espace vectoriel E engendré par n vecteurs, toute famille de cardinal strictement plus grand que n est liée.

Pour l'établir, il suffit de prouver que dans un espace vectoriel E engendré par n vecteurs, toute famille de cardinal $(n + 1)$ est liée. En effet, si cette propriété est valide, alors toute famille de cardinal strictement plus grand que n contiendra une sous famille de cardinal $(n + 1)$ liée, et sera donc elle-même liée.

Posons donc $\mathcal{P}(n)$ l'assertion : “dans un espace vectoriel E engendré par n vecteurs, toute famille de cardinal $(n + 1)$ est liée”, et démontrons $\mathcal{P}(n)$ par récurrence sur n .

Initialisation ($n = 0$). Soit E un espace vectoriel engendré par 0 vecteur... Alors $E = \{0\}$, et toute famille de cardinal 1 est liée, puisqu'une telle famille ne peut contenir que le vecteur nul. Cette brillante démonstration permet d'affirmer que $\mathcal{P}(0)$ est vraie.

Hérédité. Supposons que $\mathcal{P}(n)$ est vraie pour un certain entier naturel n , et démontrons qu'elle est vraie au rang $(n + 1)$.

Soit E un espace vectoriel, $\mathcal{G} = (v_1, \dots, v_{n+1})$ une famille génératrice de E , et $\mathcal{F} = (u_1, \dots, u_{n+2})$ une famille de vecteurs de E de cardinal $(n + 2)$. Il s'agit de prouver que $\mathcal{F}$ est liée.

Introduisons le sev F de E engendré par les n premiers vecteurs de $\mathcal{G}$: $F = \text{Vect}(v_1, \dots, v_n)$.

Soit i un entier compris entre 1 et $n + 2$. Puisque $\mathcal{G}$ est une famille génératrice de E , il existe $(n + 1)$ scalaires $\alpha_{i1}, \dots, \alpha_{in+1}$ tels que :

$$u_i = \sum_{k=1}^{n+1} \alpha_{ik} v_k = \underbrace{\left(\sum_{k=1}^n \alpha_{ik} v_k \right)}_{\in F} + \alpha_{in+1} v_{n+1}$$

En renommant $w_i = \left(\sum_{k=1}^n \alpha_{ik} v_k \right)$ et $\alpha_i = \alpha_{in+1}$, on peut conclure que :

$$\forall i \in \llbracket 1, n + 2 \rrbracket, \exists w_i \in F, \exists \alpha_i \in \mathbb{K}, u_i = w_i + \alpha_i v_{n+1} \quad (\spadesuit)$$

Deux cas peuvent alors se présenter :

➤ Premier cas : tous les α_i sont nuls. Alors : $\forall i \in \llbracket 1, n + 2 \rrbracket, \exists w_i \in F, u_i = w_i$. Ainsi la famille $(w_1, \dots, w_{n+2})$ est une famille de cardinal $(n + 2)$ dans F , qui est engendré par n vecteurs. Par hypothèse de récurrence, elle est liée.¹

➤ Second cas : les α_i ne sont pas tous nuls. Alors il existe $i \in \llbracket 1, n + 2 \rrbracket$ tel que $\alpha_i \neq 0$. Qui à renuméroter les vecteurs, on peut supposer que $\alpha_{n+2} \neq 0$.

On introduit alors la nouvelle famille de vecteurs $(U_1, \dots, U_{n+1})$ en posant :

$$\forall i \in \llbracket 1, n + 1 \rrbracket, U_i = u_i - \frac{\alpha_i}{\alpha_{n+2}} u_{n+2}$$

On a :

$$\forall i \in \llbracket 1, n + 1 \rrbracket, U_i = w_i + \alpha_i v_{n+1} - \frac{\alpha_i}{\alpha_{n+2}} (w_{n+2} + \alpha_{n+2} v_{n+1})$$

1. Cette affirmation s'obtient “en deux coups” ; l'hypothèse de récurrence permet seulement d'affirmer que la famille $(w_1, \dots, w_{n+1})$ est liée. A fortiori, la famille $(w_1, \dots, w_{n+2})$, qui la contient, est liée.

D'où : $\forall i \in \llbracket 1, n+1 \rrbracket$, $U_i = w_i - \frac{\alpha_i}{\alpha_{n+2}} w_{n+2}$. En particulier : $\forall i \in \llbracket 1, n+1 \rrbracket$, $U_i \in F$.

La famille $(U_1, \dots, U_{n+1})$ est ainsi une famille de $(n+1)$ vecteurs de F . Puisque F est engendré par n vecteurs (par construction), l'hypothèse de récurrence permet d'affirmer que cette famille est liée. Il existe donc $(n+1)$ scalaires $\beta_1, \dots, \beta_{n+1}$ non tous nuls tels que :

$$\sum_{i=1}^{n+1} \beta_i U_i = 0 \text{ d'où : } \sum_{i=1}^{n+1} \beta_i \left(u_i - \frac{\alpha_i}{\alpha_{n+2}} u_{n+2} \right) = 0 \text{ d'où : } \sum_{i=1}^{n+1} \beta_i u_i - \left(\sum_{i=1}^{n+1} \frac{\beta_i \alpha_i}{\alpha_{n+2}} \right) u_{n+2} = 0$$

En posant : $\forall i \in \llbracket 1, n+1 \rrbracket$, $\gamma_i = \beta_i$ et $\gamma_{n+2} = - \sum_{i=1}^{n+1} \frac{\beta_i \alpha_i}{\alpha_{n+2}}$,

on peut conclure qu'il existe $(n+2)$ scalaires $\gamma_1, \dots, \gamma_{n+2}$ non tous nuls² tels que : $\sum_{i=1}^{n+2} \gamma_i u_i = 0$. Il s'ensuit que la famille $\mathcal{F} = (u_1, \dots, u_{n+2})$ est liée.

➤ Dans les deux cas, la famille $\mathcal{F}$ est liée, ce qui prouve que la propriété $\mathcal{P}(n+1)$ est vraie, fournit l'hérédité, et achève donc la preuve du lemme.

Conclusion. Dans un espace vectoriel E engendré par n vecteurs, le cardinal d'une famille libre est majoré par n . ☕

THÉORÈME 22.1 - Dans un espace vectoriel (non nul) de dimension finie, toutes les bases ont le même cardinal.

PREUVE. Soit E un $\mathbb{K}$ -ev de dimension finie, et soient $\mathcal{B}$ et $\mathcal{B}'$ deux bases de E . Notons n et p les cardinaux respectifs de $\mathcal{B}$ et $\mathcal{B}'$.

Puisque $\mathcal{B}$ est une base, c'est en particulier une famille libre de E , qui est engendré par p vecteurs (puisque $\mathcal{B}'$ est génératrice). D'après le lemme, on en déduit que : $n \leq p$.

En permutant les rôles de $\mathcal{B}$ et $\mathcal{B}'$ dans le court raisonnement ci-dessus, on obtient : $p \leq n$.

Finalement $n = p$, ce qui permet de conclure :

Deux bases d'un même $\mathbb{K}$ -ev de dimension finie ont même cardinal.



Le théorème rend légitime la définition suivante :

DÉFINITION 22.6 - Soit E un $\mathbb{K}$ -espace vectoriel (non nul) de dimension finie.

On appelle **dimension** de E le cardinal d'une quelconque de ses bases (par convention, la dimension de l'espace vectoriel nul est zéro).

Notation. La dimension du $\mathbb{K}$ -ev est notée $\dim_{\mathbb{K}} E$, ou $\dim E$ s'il n'y a pas d'ambiguïté sur $\mathbb{K}$.

Exemples :

➤ $\dim \mathbb{R}^3 = 3$; $\dim_{\mathbb{K}} \mathbb{K}^n = n$; $\dim \mathbb{R}_n[X] = n+1$; $\dim M_{n,p}(\mathbb{R}) = n \times p$; ...

2. Puisque les β_i sont non tous nuls.

- $\dim M_n(\mathbb{R}) = n^2$; $\dim \mathcal{S}_n(\mathbb{R}) = \frac{n(n+1)}{2}$; $\dim \mathcal{A}_n(\mathbb{R}) = \frac{n(n-1)}{2}$
- Un exemple “plus torturé” : $\dim_{\mathbb{C}} \mathbb{C} = 1$ mais $\dim_{\mathbb{R}} \mathbb{C} = 2$. En effet, une base du $\mathbb{C}$ -ev $\mathbb{C}$ est (1) , mais une base du $\mathbb{R}$ -ev $\mathbb{C}$ est $(1, i)$. “Dans le même style” : $\dim_{\mathbb{C}} \mathbb{C}_3[X] = 4$ mais $\dim_{\mathbb{R}} \mathbb{C}_3[X] = 8$.

22.2.2 Théorème de la base extraite

THÉORÈME 22.2 - (“De la base extraite”). De toute famille génératrice on peut extraire une base. Plus explicitement : si $(v_1, \dots, v_n)$ est une famille génératrice d'un $\mathbb{K}$ -espace vectoriel non nul E , alors il existe un entier $p \in \llbracket 1; n \rrbracket$ tel que $(v_1, \dots, v_p)$ soit une base de E (quitte à renuméroter les vecteurs).

PREUVE. Soient E un $\mathbb{K}$ -ev non nul de dimension finie, et $\mathcal{F} = (v_1, \dots, v_n)$ une famille génératrice de E . Deux cas peuvent se présenter :

- Soit $\mathcal{F}$ est libre : auquel cas $\mathcal{F}$ est alors une base de E , et on a gagné.
- Soit $\mathcal{F}$ est liée : auquel cas il existe un vecteur v_i qui s'exprime comme combinaison linéaire des $n-1$ autres vecteurs de la famille $\mathcal{F}$. Sans nuire à la généralité, on peut supposer que : $v_n \in \text{Vect}(v_1, \dots, v_{n-1})$. Il s'ensuit que la famille $\mathcal{F}' = \text{Vect}(v_1, \dots, v_{n-1})$ est génératrice de E .

On recommence la discussion avec la famille $\mathcal{F}'$: soit elle est libre et c'est une base, soit elle est liée et on peut alors exprimer v_{n-1} comme combinaison linéaire des $n-2$ autres vecteurs de $\mathcal{F}'$...

Ce processus itératif se termine nécessairement (en un nombre fini d'étapes) car le cardinal de $\mathcal{F}$ est fini (et au pire, on obtiendra une base constituée d'un seul vecteur non nul, puisque E est supposé non nul).

Conclusion. De toute famille génératrice d'un ev de dimension finie, on peut extraire une base. ☕

COROLLAIRE 22.1 - Tout espace vectoriel (non nul) de dimension finie admet une base.

PREUVE. Par définition, un espace vectoriel de dimension finie admet une famille génératrice finie ; d'après le théorème précédent, on peut extraire de cette famille une base. ☕

COROLLAIRE 22.2 - Soit E un $\mathbb{K}$ -ev non nul de dimension finie n .

- 1/ Toute famille génératrice de E possède au moins n éléments.
- 2/ Toute famille génératrice de E de cardinal n est une base.

PREUVE. Soit E un $\mathbb{K}$ -ev non nul de dimension finie n . Soit $\mathcal{F}$ une famille génératrice de E ; puisque l'on peut en extraire une base, $\mathcal{F}$ est nécessairement de cardinal $\geq n$. Par ailleurs, si $\mathcal{F}$ est de cardinal n , alors toute sous-famille de $\mathcal{F}$ de cardinal $< n$ ne peut plus être plus être génératrice, car $\dim E = n$. Il s'ensuit que $\mathcal{F}$ est déjà une base de E . ☕

22.2.3 Théorème de la base incomplète

THÉORÈME 22.3 - (De la base incomplète). Soit E un $\mathbb{K}$ -ev non nul de dimension finie n . Soient encore :

- $\mathcal{F} = (v_1, \dots, v_p)$ (avec $p \leq n$) une famille libre de E ;
- et $\mathcal{B} = (w_1, \dots, w_n)$ une base de E .

Alors il existe $(n - p)$ vecteurs $w_1, \dots, w_{n-p}$ (quitte à renuméroter les vecteurs) de $\mathcal{B}$ tels que la famille $\mathcal{F} = (v_1, \dots, v_p, w_1, \dots, w_{n-p})$ est une base de E .

Interprétation. C'est le théorème de la base extraite "à l'envers" ; l'énoncé ci-dessus signifie que toute famille libre $\mathcal{F}$ peut-être complétée en une base de E . Encore plus précisément, les vecteurs que l'on adjoint à $\mathcal{F}$ pour obtenir une base peuvent être pris (avec précaution, cf exemple suivant la preuve du théorème) dans une base choisie à l'avance.

PREUVE. Soit E un $\mathbb{K}$ -ev non nul de dimension finie n , $\mathcal{F} = (v_1, \dots, v_p)$ une famille libre de E , et $\mathcal{B} = (w_1, \dots, w_n)$ une base de E . Deux cas peuvent se présenter :

- Soit $\mathcal{F}$ est génératrice : auquel cas elle est de cardinal n , et il n'y a aucun vecteur à lui adjoindre pour obtenir une base de E (c'en est déjà une!).
- Soit $\mathcal{F}$ n'est pas génératrice : auquel cas il existe un vecteur $w_i \notin \text{Vect}(\mathcal{F})$. Sans nuire à la généralité, on peut supposer qu'il s'agit de w_1 . On considère alors la nouvelle famille $\mathcal{F}' = (v_1, \dots, v_p, w_1)$; celle-ci est libre (essentiellement car $w_1 \notin \text{Vect}(\mathcal{F})$, et que $\mathcal{F}$ est libre).

Deux cas se présentent pour la famille $\mathcal{F}'$: soit elle est génératrice, soit elle ne l'est pas. Dans le second cas, on peut supposer que $w_2 \notin \text{Vect}(\mathcal{F}')$, et on considère la nouvelle famille libre $\mathcal{F}'' = (v_1, \dots, v_p, w_1, w_2) \dots$

Comme dans la preuve du théorème de la base extraite, ce processus itératif ne comporte qu'un nombre fini d'étapes, puisque la famille $\mathcal{B}$ est de cardinal fini. Mieux, puisque E est de dimension n , on sait que ce processus comporte exactement $(n - p)$ étapes, pour obtenir une base $\mathcal{B}' = (v_1, \dots, v_p, w_1, \dots, w_{n-p})$ de E .

Conclusion. Toute famille libre d'un ev E ($\neq 0$) de dimension finie est contenue dans une base de E . ☕

Exemples. Dans $E = \mathbb{R}^3$, la famille $\mathcal{F} = (u(1, 1, 1), v(1, 2, 3))$ est libre. On peut obtenir une base de $\mathbb{R}^3$ en rajoutant à $\mathcal{F}$ un vecteur de la base canonique. Dans ce cas particulier, la famille $\mathcal{B} = (u, v, w)$ est une base avec $w = e_1$ ou e_2 ou e_3 .

MAIS pour la famille $\mathcal{F}' = (u'(1, 1, 0), v'(1, 2, 0))$ (qui est libre aussi), la famille $\mathcal{B}' = (u', v', w')$ n'est une base que pour $w' = e_3$.

COROLLAIRE 22.3 - Soit E un $\mathbb{K}$ -ev non nul de dimension finie n .

- 1) Toute famille libre de E possède au plus n éléments.
- 2) Toute famille libre de E de cardinal n est une base.

PREUVE. Conséquence directe du lemme 22.2 et du théorème 22.3. 

Application (tellement classique). Soit $n \in \mathbb{N}^*$, et $(P_1, \dots, P_n)$ une famille échelonnée de polynômes de $\mathbb{K}[X]$ non nuls, càd une famille de polynômes non nuls tels que : $\deg(P_1) < \deg(P_2) \dots < \deg(P_n)$. Alors la famille $(P_1, \dots, P_n)$ est libre.

Autrement dit : toute famille échelonnée de polynômes non nuls est libre.

Preuve “rapide” : soit $n \in \mathbb{N}^*$, et $(P_1, \dots, P_n)$ une famille de polynômes de $\mathbb{K}[X]$ non nuls tels que : $\deg(P_1) < \deg(P_2) \dots < \deg(P_n)$. Supposons qu’il existe n scalaires $\alpha_1, \dots, \alpha_n$ tels que :

$$\underbrace{\alpha_1 P_1 + \dots + \alpha_n P_n}_P = 0$$

Si $\alpha_n \neq 0$, alors $\deg(P) = \deg(P_n)$.³ Ceci est absurde, puisque $P = 0$.

Il s’ensuit que $\alpha_n = 0$. On a donc :

$$\underbrace{\alpha_1 P_1 + \dots + \alpha_{n-1} P_{n-1}}_P = 0$$

En reprenant le raisonnement précédent, on obtient $\alpha_{n-1} = 0$, puis $\alpha_{n-2} = 0 \dots$ et enfin $\alpha_1 = 0$.

On peut donc conclure que la famille $(P_1, \dots, P_n)$ est libre.

Conclusion. Toute famille échelonnée de polynômes non nuls est libre. 

Remarque. La preuve “propre” (sans petits points) se fait grâce au même raisonnement, mais par le biais d’une récurrence sur le nombre de polynômes n de la famille.

Application. Soit $\alpha \in \mathbb{R}$ arbitraire. Dans $\mathbb{R}_n[X]$, la famille $\mathcal{F} = \left((X - \alpha)^k \right)_{k \in [0; n]}$ est libre (car échelonnée) et de cardinal $(n + 1) = \dim \mathbb{R}_n[X]$. C’est donc une base de $\mathbb{R}_n[X]$.

PROPRIÉTÉ 22.3 - Soit E un $\mathbb{K}$ -ev non nul de dimension finie n , et F un sev de E . Alors :

- 1) $\dim F \leq n$
- 2) $\dim F = n$ SSI $F = E$

PREUVE. 1/ Soit F un sev de E . Puisque F est de dimension finie (que l’on peut supposer non nulle⁴), F admet une base $\mathcal{F}$, de cardinal égal à $\dim(F)$ (par définition de dimension). Cette base $\mathcal{F}$ est en particulier une famille libre de vecteurs de E ; à ce titre, son cardinal est majoré par $\dim(E)$.⁵ En d’autres termes : $\dim(F) \leq \dim(E)$.

3. En effet, les P_i étant de degrés deux à deux distincts, on a $\deg(\alpha_1 P_1 + \dots + \alpha_n P_n) = \max(\deg(\alpha_1 P_1), \dots, \deg(\alpha_n P_n))$.

4. Si $F = \{0\}$, alors $\dim(F) = 0$.

5. Puisque dans un espace vectoriel de dimension n , une famille libre possède au plus n éléments.

2/ Supposons que $\dim F = n$. Alors F admet une base $\mathcal{F}$ de cardinal n . Cette base $\mathcal{F}$ est en particulier une famille libre de vecteurs de E . On en déduit que $\mathcal{F}$ est une base de E , puisqu'une famille de E est une base dès qu'elle est libre et que son cardinal est égal à la dimension de E . Par conséquent : $E = \text{Vect}(\mathcal{F}) = F$.

On a ainsi montré l'implication : $[\dim(F) = n] \implies [F = E]$. La réciproque est triviale. ☕

Exemples d'illustration. Les sev de $\mathbb{R}^2$ sont $\{\vec{0}\}$ (dimension 0), les droites vectorielles $\text{Vect}(v)$ (avec $v \in \mathbb{R}^2$ non nul ; dimension 1) et $\mathbb{R}^2$ (dimension 2).

Dans $\mathbb{R}^3$, les sev sont $\{\vec{0}\}$ (dimension 0), les droites vectorielles $\text{Vect}(v)$ (avec $v \in \mathbb{R}^3$ non nul ; dimension 1), les plans vectoriels $\text{Vect}(u, v)$ (avec u et $v \in \mathbb{R}^3$ non colinéaires ; dimension 2) et $\mathbb{R}^3$ (dimension 3).

Exemple d'application. soit n un entier naturel. L'application linéaire $\varphi : P \in \mathbb{K}_n[X] \mapsto (P - P') \in \mathbb{K}_n[X]$ est un automorphisme de $\mathbb{K}_n[X]$ (autrement écrit, $\varphi \in \text{GL}(\mathbb{K}_n[X])$).

Prouvons ce dernier point. L'application φ est bien définie car si P est un polynôme de degré au plus n , alors il en est de même du polynôme $P - P'$. En outre, un calcul aisé permet d'établir que φ est linéaire, c'est à dire que : $\forall (\lambda, \mu) \in \mathbb{K}^2, \varphi(\lambda P + \mu Q) = \lambda \varphi(P) + \mu \varphi(Q)$. Par suite : $\varphi \in \mathcal{L}(\mathbb{K}_n[X])$.

L'espace vectoriel $\mathbb{K}_n[X]$ a pour base canonique $\mathcal{B} = (X^i)_{i \in [0, n]}$. Il s'ensuit que : $\text{Im} \varphi = \text{Vect}((\varphi(X^i))_{i \in [0, n]})$.

Or : $(\varphi(X^i))_{i \in [0, n]} = (X^i - iX^{i-1})_{i \in [0, n]}$. La famille des polynômes $X^i - iX^{i-1}$ est libre car échelonnée, et de cardinal $(n+1)$ égal à la dimension de $\mathbb{K}_n[X]$. C'est donc une base de $\mathbb{K}_n[X]$, et on en déduit que φ est un isomorphisme.

Conclusion. $\varphi \in \text{GL}(\mathbb{K}_n[X])$.

22.3 Dimension d'une somme de sev

PROPRIÉTÉ 22.4 - Soient E un $\mathbb{K}$ -ev de dimension finie, et F et G deux sev supplémentaires dans E (ie $E = F \oplus G$). On a : $\dim E = \dim F + \dim G$.

PREUVE. La preuve consiste à prouver que $\mathcal{B}$ et $\mathcal{B}'$ étant deux bases de F et G respectivement, la famille $\mathcal{B} \cup \mathcal{B}'$ est génératrice et libre dans E sous l'hypothèse que $E = F \oplus G$; ce qui est à la fois aisé et long à écrire. ☕

L'énoncé ci-dessous est une conséquence du théorème de la base incomplète.

THÉORÈME 22.4 - (Existence d'un supplémentaire en dimension finie). Soient E un $\mathbb{K}$ -ev de dimension finie, et F un sev de E .

Il existe un sev G de E tel que : $E = F \oplus G$.

PREUVE. Notons $n = \dim E$, et considérons F un sev de E . F est de dimension finie (notons $p = \dim F$) et on a $0 \leq p \leq n$. Lorsque $p = 0$ ou $p = n$, la propriété est immédiate ($E = \{0\} \oplus E = E \oplus \{0\}$).

Par la suite, on peut donc supposer que : $1 \leq p \leq n - 1$.

Puisque F est un espace vectoriel non nul, il admet une base : $\mathcal{B}_0 = (v_1, \dots, v_p)$. Cette base est en particulier une famille libre de E ; d'après le théorème de la base incomplète, il existe $(n - p)$ vecteurs $w_1, \dots, w_{n-p}$ de E tels que la famille $\mathcal{B} = (v_1, \dots, v_p, w_1, \dots, w_{n-p})$ est une base de E .

Montrons que : $E = F \oplus G$ avec $G = \text{Vect}(w_1, \dots, w_{n-p})$.

► Primo, observons que : $F+G = \text{Vect}(v_1, \dots, v_p) + \text{Vect}(w_1, \dots, w_{n-p})$ d'où : $F+G = \text{Vect}(v_1, \dots, v_p, w_1, \dots, w_{n-p})$. Or la famille $\mathcal{B}$ étant une base de E , elle est en particulier génératrice de E . On en déduit que :

$$\boxed{F + G = E \quad (\spadesuit)}.$$

► Soit $V \in F \cap G$. Alors il existe n scalaires $\alpha_1, \dots, \alpha_p, \beta_1, \dots, \beta_{n-p}$ tels que : $V = \sum_{i=1}^p \alpha_i v_i = \sum_{j=1}^{n-p} \beta_j w_j$. On

en déduit que : $\sum_{i=1}^p \alpha_i v_i - \sum_{j=1}^{n-p} \beta_j w_j = 0$. Puisque $\mathcal{B}$ est en particulier une famille libre de E , on en déduit

que tous les α_i et tous les β_j sont nuls. Par suite : $V = 0$. D'où : $\boxed{F \cap G = \{0\} \quad (\clubsuit)}$.

► On déduit alors de $(\spadesuit)$, $(\clubsuit)$ et de la caractérisation des sev supplémentaires que $E = F \oplus G$.

Conclusion. Tout sev d'un $\mathbb{K}$ -ev de dimension finie E admet un supplémentaire dans E .



THÉORÈME 22.5 - (Des 4 dimensions ou relation de Grassmann). Soient E un $\mathbb{K}$ -ev de dimension finie, et F et G deux sev de E . Alors :

$$\dim(F + G) = \dim F + \dim G - \dim(F \cap G)$$

PREUVE. ► Puisque F et G sont deux sev de E , et que E est de dimension finie, F et G sont également de dimension finie. L'intersection $F \cap G$ est donc elle aussi de dimension finie. Par suite $F \cap G$ admet un supplémentaire dans F que nous noterons H ; on a donc $\boxed{(F \cap G) \oplus H = F \quad (\spadesuit)}$.

► Montrons que H et G sont supplémentaires dans $F + G$

En premier lieu : $H + (F \cap G) = F$, d'où $H + (F \cap G) + G = F + G$. Or puisque $(F \cap G) \subset G$, on a : $H + (F \cap G) + G = H + G$. Par suite : $F + G = H + G$.

En second lieu : $H \cap G = (H \cap F) \cap G$ car H est inclus dans F , ce qui implique : $(H \cap F) = H$. On en déduit, par associativité de l'intersection que : $H \cap G = H \cap (F \cap G)$. Or $H \cap (F \cap G) = \{0_E\}$ car H et $(F \cap G)$ sont supplémentaires dans F par hypothèse.

Donc : $\boxed{F + G = H \oplus G \quad (\clubsuit)}$.

On déduit de $(\spadesuit)^6$ que : $\dim H = \dim F - \dim(F \cap G)$. Et on déduit de $(\clubsuit)$ que : $\dim H = \dim(F + G) - \dim G$. En identifiant les deux expressions obtenues pour $\dim H$, on obtient : $\dim F - \dim(F \cap G) = \dim(F + G) - \dim G$ puis la conclusion : $\boxed{\dim(F + G) = \dim F + \dim G - \dim(F \cap G)}$. 

Exemple d'application. Dans $E = \mathbb{R}^3$, on considère les sous-espaces vectoriels

$F = \{(x, y, z) \in \mathbb{R}^3 / x + y + z = 0\}$ et $G = \text{Vect}((1, 1, 1))$. Il est aisé de prouver que $\dim F = 2$, et il est immédiat que $\dim G = 1$. D'après le théorème ci-dessus, il suffit alors de prouver que $F \cap G = \{\vec{0}\}$ pour conclure que $\mathbb{R}^3 = F \oplus G$ (gain de temps important).

6. Et de la formule donnant la dimension d'une somme directe : $\dim(F_1 \oplus F_2) = \dim F_1 + \dim F_2$.

DÉFINITION 22.7 - Soit E un $\mathbb{K}$ -ev de dimension n . On appelle **hyperplan** de E un sev de E de dimension $n - 1$.

PROPRIÉTÉ 22.5 - Soient E un $\mathbb{K}$ -ev de dimension n , H un hyperplan de E et v un vecteur de E tel que $v \notin H$. Alors :

$$E = H \oplus \text{Vect}(v)$$

PREUVE. Commençons par observer que $H \subset H + \text{Vect}(v) \subset E$. Il s'ensuit que : $n - 1 \leq \dim(H + \text{Vect}(v)) \leq n$.

Supposons que : $\dim(H + \text{Vect}(v)) = n - 1$. Alors : $H + \text{Vect}(v) = H$ puisque H est contenu dans $(H + \text{Vect}(v))$ et que ces deux espaces auraient alors la même dimension. Cette égalité impliquerait $v \in H$, ce qui est en contradiction avec l'hypothèse.

Par suite : $\dim(H + \text{Vect}(v)) = n$, ce qui entraîne : $H + \text{Vect}(v) = E$ (♣).

Par ailleurs : $\{\vec{0}_E\} \subset (H \cap \text{Vect}(v)) \subset \text{Vect}(v)$. Il s'ensuit que : $0 \leq \dim(H \cap \text{Vect}(v)) \leq 1$.

Supposons que : $\dim(H \cap \text{Vect}(v)) = 1$. Alors : $H \cap \text{Vect}(v) = \text{Vect}(v)$ puisque $H \cap \text{Vect}(v)$ est contenu dans $\text{Vect}(v)$ et que ces deux espaces auraient alors la même dimension. De nouveau, cette égalité impliquerait $v \in H$, ce qui est en contradiction avec l'hypothèse.

Par suite : $\dim(H \cap \text{Vect}(v)) = 0$, ce qui entraîne : $H \cap \text{Vect}(v) = \{\vec{0}_E\}$ (♠).

D'après (♣) et (♠), on peut conclure que : $E = H \oplus \text{Vect}(v)$, ce qu'il fallait démontrer. ☕

Exemples d'application.

1/ Dans $\mathbb{K}^n$ (avec $n \geq 2$), on considère $H = \{(x_1, \dots, x_n), x_1 + \dots + x_n = 0\}$ et $D = \text{Vect}((1, \dots, 1))$. On a : $\mathbb{K}^n = H \oplus D$.

2/ Dans $M_2(\mathbb{R})$, on considère $H = \{A \in M_2(\mathbb{R}) / \text{tr}(A) = 0\}$. H est l'hyperplan des matrices de trace nulle, et on a : $M_2(\mathbb{R}) = H \oplus \text{Vect}(M)$ pour toute matrice $M \in M_2(\mathbb{R})$ de trace non nulle.

3/ Dans un $\mathbb{K}$ -ev E de dimension $n \geq 2$, l'intersection de deux hyperplans distincts est de dimension $n - 2$.

En effet, d'après la formule des quatre dimensions : $\dim(H_1 + H_2) = \dim H_1 + \dim H_2 - \dim(H_1 \cap H_2)$.

On a : $H_1 \subset (H_1 + H_2) \subset E$. D'où : $n - 1 \leq \dim(H_1 + H_2) \leq n$.

Supposons que $\dim(H_1 + H_2) = n - 1$. Alors : $H_1 = H_1 + H_2$ (puisque alors $H_1 \subset (H_1 + H_2)$ et $\dim H_1 = \dim(H_1 + H_2)$). Or par hypothèse, H_1 et H_2 sont distincts, donc il existe $h_2 \in H_2 \setminus H_1$, et donc $H_1 \neq H_1 + H_2$.

Il s'ensuit que : $\dim(H_1 + H_2) = n$. Par conséquent : $n = 2(n - 1) + \dim(H_1 \cap H_2)$, d'où : $\dim(H_1 \cap H_2) = n - 2$.

Conclusion. Dans un $\mathbb{K}$ -ev de dimension $n \geq 2$, l'intersection de deux hyperplans distincts est dimension $n - 2$.

22.4 Coordonnées d'un vecteur dans une base, matrice de passage

22.4.1 Coordonnées

THÉORÈME 22.6 - Soit E un $\mathbb{K}$ -ev de dimension n , et soit $\mathcal{B} = (v_1, \dots, v_n)$ une base de E . Alors :

$$\forall V \in E, \exists! (\alpha_1, \dots, \alpha_n) \in \mathbb{K}^n, V = \sum_{i=1}^n \alpha_i v_i$$

DÉFINITION 22.8 - Les scalaires $(\alpha_1, \dots, \alpha_n)$ du théorème ci-dessus sont appelés **coordonnées du vecteur V dans la base $\mathcal{B}$** .

Remarque : le point crucial est de remarquer que les coordonnées d'un vecteur sont relatives à la base choisie.

PREUVE. (du théorème 22.6). Sous les hypothèses de l'énoncé, pour un vecteur v de E , il existe un n -uplet $(\alpha_1, \dots, \alpha_n)$ de scalaires tel que : $v = \sum_{i=1}^n \alpha_i v_i$, puisque la famille $\mathcal{B}$ est en particulier génératrice.

L'unicité du n -uplet est une conséquence directe du fait que la famille $\mathcal{B}$ est libre. ☕

Exemple. Dans $\mathbb{R}_2[X]$, les coordonnées du polynôme $P = aX^2 + bX + c$ sont (c, b, a) dans la base canonique $\mathcal{B} = (1, X, X^2)$. Les coordonnées de P sont $(P(0), P(1), P(2))$ dans la base $\mathcal{B}' = (L_0, L_1, L_2)$ des polynômes de Lagrange de degré 2 associés aux valeurs 0, 1 et 2. Enfin, d'après la formule de Taylor pour les polynômes, les coordonnées de P sont $(P(\alpha), P'(\alpha), P''(\alpha)/2)$ dans la base $\mathcal{B}'' = (1, X - \alpha, (X - \alpha)^2)$ (pour tout réel α).

22.4.2 Changement de base et matrice de passage

DÉFINITION 22.9 - Soit E un $\mathbb{K}$ -ev de dimension n ($n \neq 0$), soient $\mathcal{B} = (v_1, \dots, v_n)$ et $\mathcal{B}' = (w_1, \dots, w_n)$ deux bases de E .

La **matrice de passage de la base $\mathcal{B}$ à la base $\mathcal{B}'$** notée $P_{\mathcal{B}\mathcal{B}'}$ (ou $P_{\mathcal{B}}^{\mathcal{B}'}$) est la matrice de $M_n(\mathbb{K})$

$$P_{\mathcal{B}\mathcal{B}'} = (\alpha_{ij})_{1 \leq i \leq n, 1 \leq j \leq n}$$

les scalaires α_{ij} étant caractérisés par :

$$\forall j \in \llbracket 1, n \rrbracket, w_j = \sum_{i=1}^n \alpha_{ij} v_i$$

Traduction : $P_{\mathcal{B}\mathcal{B}'}$ est la matrice obtenue en mettant en colonnes les coordonnées des vecteurs de $\mathcal{B}'$ dans la base $\mathcal{B}$ (attention à l'ordre!).

PROPRIÉTÉ 22.6 - (Effet d'un changement de base sur les coordonnées). Soit E un $\mathbb{K}$ -ev de dimension n ($n \neq 0$), soient $\mathcal{B}$ et $\mathcal{B}'$ deux bases de E .

Pour tout vecteur V de E , notons $X_{\mathcal{B}}$ et $X_{\mathcal{B}'}$ les n -uplets de ses coordonnées dans les bases $\mathcal{B}$ et $\mathcal{B}'$ respectivement. On a :

$$X_{\mathcal{B}} = P_{\mathcal{B}\mathcal{B}'} X_{\mathcal{B}'}$$

PREUVE. ► Notations et remarques préliminaires. Soient $\mathcal{B} = (v_i)_{i \in [1, n]}$ et $\mathcal{B}' = (w_i)_{i \in [1, n]}$ deux bases de E .

On note $P_{\mathcal{B}\mathcal{B}'} = (a_{ij})$ la matrice de passage de la base $\mathcal{B}$ à la base $\mathcal{B}'$, où les scalaires a_{ij} sont caractérisés par : $\forall j \in [1, n], w_j = \sum_{i=1}^n a_{ij} v_i$.

Soit v un vecteur de E . Il existe $2n$ scalaires $(x_1, \dots, x_n, y_1, \dots, y_n)$ tels que : $v = \sum_{i=1}^n x_i v_i = \sum_{j=1}^n y_j w_j$.

Ainsi le n -uplet des coordonnées de v dans la base $\mathcal{B}$ (resp. $\mathcal{B}'$) est : $X_{\mathcal{B}} = (x_1, \dots, x_n)$ (resp. $X_{\mathcal{B}'} = (y_1, \dots, y_n)$).

Ces notations étant posées, observons à présent que l'égalité de l'énoncé peut être traduite par :

$$[X_{\mathcal{B}} = P_{\mathcal{B}\mathcal{B}'} X_{\mathcal{B}'}] \iff \left[\forall i \in [1, n], x_i = \sum_{j=1}^n a_{ij} y_j \right] \quad (\spadesuit)$$

► On a : $v = \sum_{j=1}^n y_j w_j = \sum_{j=1}^n y_j \sum_{i=1}^n a_{ij} v_i = \sum_{j=1}^n \sum_{i=1}^n a_{ij} y_j v_i = \sum_{i=1}^n \left(\sum_{j=1}^n a_{ij} y_j \right) v_i$ et $v = \sum_{i=1}^n x_i v_i$.

On déduit de ces deux égalités et de l'unicité des coordonnées de v dans la base $\mathcal{B}$ que : $\forall i \in [1, n], x_i = \sum_{j=1}^n a_{ij} y_j$, ce qui à la lumière de $(\spadesuit)$ permet de conclure : $X_{\mathcal{B}} = P_{\mathcal{B}\mathcal{B}'} X_{\mathcal{B}'}$. ☕

Exemple d'application. Soit $E = \mathbb{R}_2[X]$. On considère la famille $\mathcal{B}' = (X^2 - 1, X + 1, X)$.

1/ Montrer que $\mathcal{B}'$ est une base de E .

2/ A l'aide d'une matrice de passage, déterminer les coordonnées de $P = aX^2 + bX + c$ dans la base $\mathcal{B}'$.

CORRIGÉ.

1/ Dans cette situation, la famille $\mathcal{B}'$ n'est pas une famille échelonnée de polynômes, et on doit donc utiliser la définition de famille libre pour établir que $\mathcal{B}'$ l'est.

Supposons donc qu'il existe trois réels a, b et c tels que $a(X^2 - 1) + b(X + 1) + cX = \tilde{0}$. L'évaluation de cette relation en (-1) donne $c = 0$. Dans un second temps, on évalue $a(X^2 - 1) + b(X + 1) = \tilde{0}$ en 1 , ce qui donne $b = 0$. On en déduit finalement $a = 0$. Par suite, la famille $\mathcal{B}'$ est libre, et puisque son cardinal est égal à la dimension de $\mathbb{R}_2[X]$, on peut conclure : $\mathcal{B}'$ est une base de $\mathbb{R}_2[X]$.

2/ La matrice de passage de la base canonique $\mathcal{B} = (1, X, X^2)$ à la base $\mathcal{B}' = (X^2 - 1, X + 1, X)$ est :

$$P_{\mathcal{B}\mathcal{B}'} = \begin{pmatrix} -1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \end{pmatrix}$$

Cette matrice est inversible (comme toute matrice de passage, voir un peu plus loin) et :

$$P_{\mathcal{B}\mathcal{B}'} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} \iff \begin{cases} -x_1 + x_2 = b_1 \\ x_2 + x_3 = b_2 \\ x_1 = b_3 \end{cases} \iff \begin{cases} x_2 = b_1 + b_3 \\ x_3 = b_2 - b_1 - b_3 \\ x_1 = b_3 \end{cases} \Rightarrow$$

$$P_{\mathcal{B}\mathcal{B}'}^{-1} = P_{\mathcal{B}'\mathcal{B}} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 1 \\ -1 & 1 & -1 \end{pmatrix}$$

Soit $P = aX^2 + bX + c$ un polynôme de $\mathbb{R}_2[X]$. Notons $X_{\mathcal{B}}$ (resp. $X_{\mathcal{B}'}$) le triplet des coordonnées de P dans la base $\mathcal{B}$ (resp. $\mathcal{B}'$). On a :

$$X_{\mathcal{B}} = \begin{pmatrix} c \\ b \\ a \end{pmatrix} \text{ et } X_{\mathcal{B}'} = P_{\mathcal{B}'\mathcal{B}} X_{\mathcal{B}} \text{ d'où : } X_{\mathcal{B}'} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 1 \\ -1 & 1 & -1 \end{pmatrix} \begin{pmatrix} c \\ b \\ a \end{pmatrix} \text{ soit : } X_{\mathcal{B}'} = \begin{pmatrix} a \\ a + c \\ b - (a + c) \end{pmatrix}$$

► **Conclusion.** $\forall (a, b, c) \in \mathbb{R}^3, aX^2 + bX + c = a(X^2 - 1) + (a + c)(X + 1) + (a - (b + c))X$

L'énoncé suivant a une importance plus technique que pratique (on vous demandera souvent de faire un changement de base dans un problème, rarement deux) :

PROPRIÉTÉ 22.7 - Soit E un $\mathbb{K}$ -ev de dimension n ($n \neq 0$), soient $\mathcal{B}$, $\mathcal{B}'$ et $\mathcal{B}''$ trois bases de E . On a :

$$P_{\mathcal{B}\mathcal{B}''} = P_{\mathcal{B}\mathcal{B}'} P_{\mathcal{B}'\mathcal{B}''}$$

PREUVE. ► Notations et remarques préliminaires. Soient $\mathcal{B} = (v_i)_{i \in [1, n]}$, $\mathcal{B}' = (w_i)_{i \in [1, n]}$ et $\mathcal{B}'' = (z_i)_{i \in [1, n]}$ trois bases de E .

On note :

— $P_{\mathcal{B}\mathcal{B}'} = (a_{ij})$ la matrice de passage de la base $\mathcal{B}$ à la base $\mathcal{B}'$, où les scalaires a_{ij} sont caractérisés par :

$$\forall j \in [1, n], \quad w_j = \sum_{i=1}^n a_{ij} v_i.$$

— $P_{\mathcal{B}'\mathcal{B}''} = (b_{ij})$ la matrice de passage de la base $\mathcal{B}'$ à la base $\mathcal{B}''$, où les scalaires b_{ij} sont caractérisés par :

$$\forall j \in [1, n], \quad z_j = \sum_{i=1}^n b_{ij} w_i.$$

— $P_{\mathcal{B}\mathcal{B}''} = (c_{ij})$ la matrice de passage de la base $\mathcal{B}$ à la base $\mathcal{B}''$, où les scalaires c_{ij} sont caractérisés par :

$$\forall j \in [1, n], \quad z_j = \sum_{i=1}^n c_{ij} v_i.$$

Avec ces notations, on peut reformuler la propriété à démontrer :

$$[P_{\mathcal{B}\mathcal{B}''} = P_{\mathcal{B}\mathcal{B}'} P_{\mathcal{B}'\mathcal{B}''}] \iff \left[\forall (i, j) \in [1, n]^2, \quad c_{ij} = \sum_{k=1}^n a_{ik} b_{kj} \right]$$

► Soit j un entier compris entre 1 et n . On a d'une part :
$$z_j = \sum_{i=1}^n c_{ij} v_i \quad (\spadesuit).$$

Et d'autre part : $z_j = \sum_{k=1}^n b_{kj} w_k = \sum_{k=1}^n b_{kj} \left(\sum_{i=1}^n a_{ik} v_i \right)$ donc :
$$z_j = \sum_{i=1}^n \left(\sum_{k=1}^n b_{kj} a_{ik} \right) v_i \quad (\clubsuit).$$

On déduit alors de $(\spadesuit)$, $(\clubsuit)$ et de l'unicité des coordonnées de z_j dans la base $\mathcal{B}$ que : $\forall i \in \llbracket 1, n \rrbracket, c_{ij} = \sum_{k=1}^n b_{kj} a_{ik}$. L'entier j étant arbitraire dans ce raisonnement, on peut conclure que : $\forall i \in \llbracket 1, n \rrbracket, \forall j \in \llbracket 1, n \rrbracket, c_{ij} = \sum_{k=1}^n b_{kj} a_{ik}$.

Par suite :
$$P_{\mathcal{B}\mathcal{B}''} = P_{\mathcal{B}\mathcal{B}'} P_{\mathcal{B}'\mathcal{B}''}.$$
 ☕

La conséquence spectaculaire de cet énoncé étant la propriété (plus pratique que technique) ci-dessous :

PROPRIÉTÉ 22.8 - Soit E un $\mathbb{K}$ -ev de dimension n ($n \neq 0$), soient $\mathcal{B}$ et $\mathcal{B}'$ deux bases de E . Alors $P_{\mathcal{B}\mathcal{B}'} \in \text{GL}_n(\mathbb{K})$, et $[P_{\mathcal{B}\mathcal{B}'}]^{-1} = P_{\mathcal{B}'\mathcal{B}}$.

PREUVE. Triviale. ☕

Pour enfoncer le clou : **toute matrice de passage est inversible**

22.5 Classification des ev de dimension finie

PROPRIÉTÉ 22.9 - Soient E et F deux $\mathbb{K}$ -ev, avec E de dimension $n \geq 1$; $\mathcal{B} = (\vec{v}_i)_{i \in \llbracket 1, n \rrbracket}$ une base de E ; et $f \in \mathcal{L}(E, F)$.

Enfin on note $\mathcal{F} = (f(\vec{v}_i))_{i \in \llbracket 1, n \rrbracket}$. Alors :

- 1/ $[f \text{ est injective}] \iff [\mathcal{F} \text{ est libre}]$
- 2/ $[f \text{ est surjective}] \iff [\mathcal{F} \text{ est génératrice de } F]$
- 3/ $[f \text{ est un isomorphisme}] \iff [\mathcal{F} \text{ est une base de } F]$

PREUVE. ► Montrons le point 1 : raisonnons par double implication.

► Montrons que $[f \text{ est injective}] \implies [\mathcal{F} \text{ est libre}]$. On suppose f injective.

Supposons qu'il existe n scalaires $\alpha_1, \dots, \alpha_n$ tels que :
$$\sum_{i=1}^n \alpha_i f(v_i) = 0_F$$

Alors, par linéarité de f on a : $f\left(\sum_{i=1}^n \alpha_i v_i\right) = 0_F$

Puisque f est injective, on en déduit que : $\sum_{i=1}^n \alpha_i v_i = 0_E$. Et comme les vecteurs v_i sont linéairement indépendants, on a : $\forall i \in \llbracket 1, n \rrbracket, \alpha_i = 0$. Donc $\mathcal{F}$ est libre.

Conclusion : $[f \text{ est injective}] \implies [\mathcal{F} \text{ est libre}]$.

➤ Montrons que réciproquement $[\mathcal{F} \text{ est libre}] \implies [f \text{ est injective}]$. On suppose que la famille $\mathcal{F}$ est libre.

Soit $V \in \ker f$. Puisque les vecteurs v_i constituent une base de $E : \exists! (\alpha_1, \dots, \alpha_n) \in \mathbb{K}^n, V = \sum_{i=1}^n \alpha_i v_i$.

$$\begin{aligned} \text{Ainsi : } [V \in \ker f] &\iff [f(V) = 0_F] \iff \left[f \left(\sum_{i=1}^n \alpha_i v_i \right) = 0_F \right] \iff \left[\sum_{i=1}^n \alpha_i f(v_i) = 0_F \right] \\ &\iff [\forall i \in \llbracket 1, n \rrbracket, \alpha_i = 0] \end{aligned}$$

l'avant dernière équivalence provenant de la linéarité de f , et la dernière de l'hypothèse de liberté faite sur $\mathcal{F}$. Il s'ensuit que $V = 0_E$. Par suite $\ker f = \{0_E\}$.

Donc f est injective. **Conclusion** : $[\mathcal{F} \text{ est libre}] \implies [f \text{ est injective}]$.

Finalement : $[f \text{ est injective}] \iff [\mathcal{F} \text{ est libre}]$.

➤ Montrons le point 2 :⁷

$$\begin{aligned} [f \text{ est surjective}] &\iff [\text{Im } f = F] \iff [\text{Vect}((f(v_i))_{i \in \llbracket 1, n \rrbracket}) = F] \iff [\text{Vect}(\mathcal{F}) = F] \\ &\iff [\mathcal{F} \text{ est une famille génératrice de } F] \end{aligned}$$

➤ Montrons le point 3 : $[f \text{ est un isomorphisme}] \iff [f \text{ est injective et surjective}]$

$$\iff [\mathcal{F} \text{ est libre et génératrice de } F] \iff [\mathcal{F} \text{ est une base de } F] \quad \text{☕}$$

COROLLAIRE 22.4 - Soient E et F deux $\mathbb{K}$ -ev de dimension finie.

Si E et F sont isomorphes, alors $\dim E = \dim F$

PREUVE. Conséquence immédiate de la propriété précédente. ☕

La réciproque est vraie, dans le sens rendu plus précis par l'énoncé ci-dessous :

THÉORÈME 22.7 - (Classification des ev de dimension finie).

Tout $\mathbb{K}$ -espace vectoriel de dimension n est isomorphe à $\mathbb{K}^n$.

PREUVE. Soit E un $\mathbb{K}$ -ev de dimension n non nulle.⁸ E admet une base $\mathcal{B} = (v_1, \dots, v_n)$. On définit alors les applications

$$\begin{aligned} \varphi : \mathbb{K}^n &\longrightarrow E & \text{et} & & \psi : E &\longrightarrow \mathbb{K}^n \\ (\alpha_1, \dots, \alpha_n) &\longmapsto \sum_{i=1}^n \alpha_i v_i & & & V &\longmapsto (x_1, \dots, x_n) \end{aligned}$$

7. La preuve de ce point repose en grande partie sur la célèbre **propriété** : soient E et F deux $\mathbb{K}$ -ev, et $f \in \mathcal{L}(E, F)$. On suppose qu'il existe une famille $(v_i)_{i \in \llbracket 1, n \rrbracket}$ génératrice de E , c'est-à-dire une famille de vecteurs de E tels que $E = \text{Vect}((v_i)_{i \in \llbracket 1, n \rrbracket})$. Alors : $\text{Im } f = \text{Vect}((f(v_i))_{i \in \llbracket 1, n \rrbracket})$.

8. Sinon $E = \{0\}$ et il n'y a pas grand chose à faire.

où l'on a noté $(x_1, \dots, x_n)$ le n -uplet des coordonnées du vecteur V dans la base $\mathcal{B}$.

Il est immédiat que φ et ψ sont linéaires, et que φ et ψ sont réciproques l'une de l'autre. On en déduit que φ et ψ sont des isomorphismes, d'où la conclusion :

tout $\mathbb{K}$ -espace vectoriel de dimension n est isomorphe à $\mathbb{K}^n$.



Interprétation. A isomorphisme près, $\mathbb{K}^n$ est le seul $\mathbb{K}$ -ev de dimension n .

Interprétation libre... A un costume près, il existe un unique Manneken-Pis.



De la même manière :

- l'espace vectoriel des polynômes à coefficients réels de degré au plus un : $\mathbb{R}_1[X] = \text{Vect}(1, X)$;
- l'espace vectoriel F des solutions (à valeurs réelles) de l'équation différentielle linéaire $y'' - 3y' + 2y = 0$, càd : $F = \text{Vect}(x \mapsto e^x, x \mapsto e^{2x})$;

- l'espace vectoriel D des matrices diagonales dans $M_2(\mathbb{R})$: $D = \text{Vect}(E_{11}, E_{22})$;
- l'espace vectoriel E des suites récurrentes réelles linéaires d'ordre 2 satisfaisant la relation de récurrence $\forall n \in \mathbb{N}, u_{n+2} = u_{n+1} + u_n$, càd :

$$E = \text{Vect} \left(\left(\left(\frac{1 + \sqrt{5}}{2} \right)^n \right)_{n \in \mathbb{N}}, \left(\left(\frac{1 - \sqrt{5}}{2} \right)^n \right)_{n \in \mathbb{N}} \right)$$

sont quatre costumes différents pour le même $\mathbb{R}$ -espace vectoriel $\mathbb{R}^2$.

Remarque. Dans le théorème de classification, il est surtout important de retenir l'idée de la preuve. Si E est un ev de dimension n , il admet une base $\mathcal{B} = (v_1, \dots, v_n)$. Un isomorphisme $\varphi : \mathbb{K}^n \rightarrow E$ peut alors être défini en posant : $\varphi(\alpha_1, \dots, \alpha_n) = \sum_{i=1}^n \alpha_i v_i$; l'isomorphisme réciproque φ^{-1} associe à tout vecteur v de E le n -uplet de ses coordonnées dans la base $\mathcal{B}$.

COROLLAIRE 22.5 - Deux $\mathbb{K}$ -ev de même dimension sont isomorphes.

PREUVE. Conséquence immédiate du théorème de classification.

Exemple. $S_3(\mathbb{R})$ (le sev des matrices symétriques de $M_3(\mathbb{R})$), $M_{3,2}(\mathbb{R})$ et $\mathbb{R}_5[X]$ sont des $\mathbb{R}$ -espaces vectoriels deux à deux isomorphes, tous trois étant isomorphes à $\mathbb{R}^6$.

22.6 Epilogue

22.6.1 Exemples de sev supplémentaires

Dans cet exercice, on considère le $\mathbb{R}$ -espace vectoriel $E = \mathbb{R}_{26}[X]$ des polynômes à coefficients réels de degré au plus 26.

On note $G = \text{Vect}(X^{19} + 2X^4 - 5X)$ le sous-espace vectoriel de E engendré par le polynôme $Q = X^{19} + 2X^4 - 5X$.

Et on note F la partie de E suivante :

$$F = \{P \in E / P(0) = P(1)\}$$

- 1) Montrer que F est un sous-espace vectoriel de E . Préciser la dimension de F .
- 2) Etablir que $E = F \oplus G$.

CORRIGÉ.

1) L'application $\varphi : P \in \mathbb{R}_{26}[X] \mapsto P(0) - P(1)$ est clairement une forme linéaire sur $\mathbb{R}_{26}[X]$, non nulle (car $\varphi(X) \neq 0$). Puisque $F = \ker(\varphi)$, on en déduit que F est un hyperplan de $\mathbb{R}_{26}[X]$.

En d'autres termes F est un sev de $\mathbb{R}_{26}[X]$ de dimension 26.

2) Comme F est un hyperplan de $\mathbb{R}_{26}[X]$ et que $(X^{19} + 2X^4 - 5X) \notin F$, on a :

$$\mathbb{R}_{26}[X] = F \oplus \text{Vect}(X^{19} + 2X^4 - 5X).$$

22.6.2 Un exemple de changement de base

Dans cet exercice, on considère l'espace vectoriel $E = M_2(\mathbb{K})$ des matrices carrées d'ordre 2 à coefficients dans $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$.

On rappelle que la **base canonique** de E est la base $\mathcal{B} = (E_{11}, E_{12}, E_{21}, E_{22})$ avec :

$$E_{11} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}; \quad E_{12} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}; \quad E_{21} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}; \quad E_{22} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

1) On considère la famille $\mathcal{B}' = (M_1, M_2, M_3, M_4)$ de $M_2(\mathbb{K})$, où l'on a posé :

$$M_1 = \begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix}; \quad M_2 = \begin{pmatrix} -1 & 1 \\ -1 & 0 \end{pmatrix}; \quad M_3 = \begin{pmatrix} 0 & -1 \\ 1 & -1 \end{pmatrix}; \quad M_4 = \begin{pmatrix} 0 & 0 \\ -1 & 1 \end{pmatrix}$$

Montrer que $\mathcal{B}'$ est une base de $M_2(\mathbb{K})$.

2) Ecrire la matrice de passage $P_{\mathcal{B}\mathcal{B}'}$ de la base $\mathcal{B}$ à la base $\mathcal{B}'$.

Dans la suite de l'exercice, on notera simplement P la matrice $P_{\mathcal{B}\mathcal{B}'}$.

3) Justifier brièvement que P est inversible, et calculer P^{-1} .

4) On note A la matrice $\begin{pmatrix} 2 & -3 \\ 1 & 4 \end{pmatrix}$. Déterminer les coordonnées de A dans la base $\mathcal{B}'$.

CORRIGÉ.

1) Montrons que la famille $\mathcal{B}'$ est libre.

Soient a, b, c et d quatre réels tels que $aM_1 + bM_2 + cM_3 + dM_4 = 0_E$, alors :

$$\begin{pmatrix} a-b & b-a-c \\ c-b-d & d-c \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \iff \begin{cases} b=a \\ b=a+c \\ c=b+d \\ c=d \end{cases} \iff \begin{cases} a=0 \\ b=0 \\ c=0 \\ d=0 \end{cases}$$

Il s'ensuit que la famille $\mathcal{B}'$ est libre ; et puisque son cardinal (4) est égal à la dimension de E , on peut conclure : $\boxed{\mathcal{B}' \text{ est une base de } E}$.

2) On a : $\boxed{P_{\mathcal{B}\mathcal{B}'} = \begin{pmatrix} 1 & -1 & 0 & 0 \\ -1 & 1 & -1 & 0 \\ 0 & -1 & 1 & -1 \\ 0 & 0 & -1 & 1 \end{pmatrix}}$.

3) La matrice P est inversible car c'est une matrice de passage. Calculons son inverse en résolvant un système linéaire. On fixe donc B un vecteur arbitraire de $\mathbb{K}^4$, et on cherche $X \in \mathbb{K}^4$ tel que $PX = B$.

$$PX = B \iff \begin{cases} x_1 - x_2 = b_1 \\ -x_1 + x_2 - x_3 = b_2 \\ -x_2 + x_3 - x_4 = b_3 \\ -x_3 + x_4 = b_4 \end{cases} \iff \begin{cases} x_2 = -b_3 - b_4 \\ x_1 = b_1 - b_3 - b_4 \\ x_3 = -b_1 - b_2 \\ x_4 = -b_1 - b_2 + b_4 \end{cases}$$

Conclusion : P est inversible et $P^{-1} = P_{\mathcal{B}'\mathcal{B}} = \begin{pmatrix} 1 & 0 & -1 & -1 \\ 0 & 0 & -1 & -1 \\ -1 & -1 & 0 & 0 \\ -1 & -1 & 0 & 1 \end{pmatrix}$.

4) Notons $X_{\mathcal{B}}$ et $X_{\mathcal{B}'}$ les quadruplets de coordonnées de A dans les bases $\mathcal{B}$ et $\mathcal{B}'$ respectivement. On a : $X_{\mathcal{B}} = (2, -3, 1, 4)$ et $X_{\mathcal{B}'} = P_{\mathcal{B}'\mathcal{B}}X_{\mathcal{B}}$.

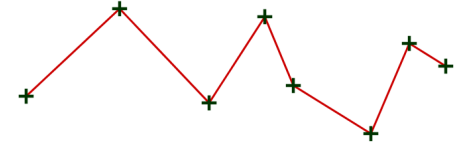
Ainsi : $X_{\mathcal{B}'} = \begin{pmatrix} 1 & 0 & -1 & -1 \\ 0 & 0 & -1 & -1 \\ -1 & -1 & 0 & 0 \\ -1 & -1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 2 \\ -3 \\ 1 \\ 4 \end{pmatrix} \iff \boxed{X_{\mathcal{B}'} = \begin{pmatrix} -3 \\ -5 \\ 1 \\ 5 \end{pmatrix}}$.

22.6.3 Interpolation polynomiale et B-splines

ENONCÉ.

► **Motivation.** *Etant donnés un nombre fini de points du plan, on cherche à décrire une trajectoire (c'est-à-dire une courbe plane) passant par ces n points.*

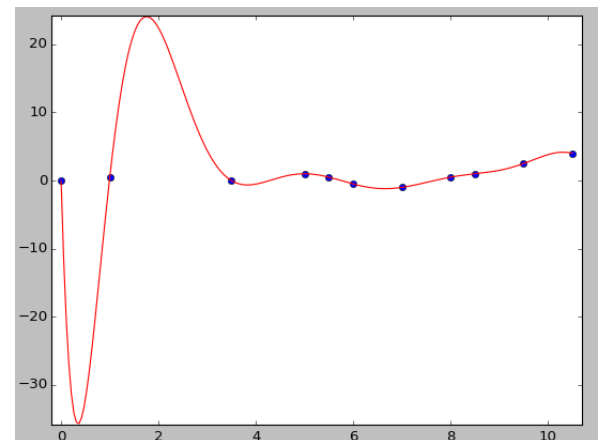
Une première idée pourrait être de joindre les points par une ligne polygonale (voir figure ci-contre).



Cette approche est un peu trop “brutale” si l'on souhaite par exemple que la trajectoire obtenue puisse être celle d'un véhicule incapable de tourner sur lui-même... On souhaite donc préciser le problème, et obtenir une trajectoire “lisse” passant par les n points.

Une stratégie pour y parvenir est d'interpoler les n points par des polynômes de degré au plus $(n + 1)$, par le biais notamment des polynômes de Lagrange ; c'est l'objet de la partie A.

Cette stratégie a cependant un fâcheux inconvénient : si la trajectoire obtenue passe effectivement par les points, elle peut se comporter “sauvagement” en-dehors des points d'interpolation (voir illustration ci-contre).



Une seconde stratégie consiste alors à contrôler non seulement les points d'interpolation, mais également les tangentes à la trajectoire en ces points ; c'est l'objet des parties B et C.

► **Notations et conventions.** Tout au long de ce problème, n désigne un entier naturel non nul.

On suppose que $\alpha_0, \dots, \alpha_n$ sont $(n + 1)$ réels 2 à 2 distincts tels que : $\alpha_0 < \alpha_1 < \dots < \alpha_n$.

Enfin, on pose :
$$U = \prod_{i=0}^n (X - \alpha_i) \quad \text{c'est-à-dire}^9 : U = (X - \alpha_0)(X - \alpha_1) \cdots (X - \alpha_n).$$

PARTIE A - Interpolation polynomiale

Pour tout entier naturel p non-nul, on définit l'application :

$$\begin{aligned} \varphi_p : \mathbb{R}_p[X] &\longrightarrow \mathbb{R}^{n+1} \\ P &\longmapsto (P(\alpha_0), P(\alpha_1), \dots, P(\alpha_n)) \end{aligned}$$

On admet que l'application φ_p est linéaire.

1/ Montrer que si $p \leq n$, alors l'application φ_p est injective.

2/ Montrer que si $p \geq n + 1$, alors $\ker \varphi_p$ est un sev de $\mathbb{R}_p[X]$ dont on précisera une base et la dimension.

9. Plus explicitement mais moins rigoureusement.

3/ Dans cette question, on suppose que $p = n$.

On note $\mathcal{F} = ((L_k)_{k=0,\dots,n})$ la famille des polynômes de Lagrange associés aux réels $\alpha_0, \dots, \alpha_n$. Ces polynômes sont définis comme suit :

$$\forall k \in \llbracket 0, n \rrbracket, \quad L_k(X) = \prod_{j=0, j \neq k}^n \frac{X - \alpha_j}{\alpha_k - \alpha_j}$$

Chacun des polynômes L_k vérifie les conditions :

$$(*) \quad L_k(\alpha_k) = 1 \text{ et pour tout entier } i \neq k \text{ dans } \llbracket 0, n \rrbracket, L_k(\alpha_i) = 0$$

a/ Etablir que la famille $\mathcal{F} = (L_0, L_1, \dots, L_n)$ est une base de $\mathbb{R}_n[X]$.

b/ Déterminer $\text{Im}(\varphi_n)$. Puis établir que φ_n est un isomorphisme de $\mathbb{R}_n[X]$ dans $\mathbb{R}^{n+1}$.

PARTIE B - Interpolation polynomiale avec contraintes

Soit k un entier quelconque dans $\llbracket 0, n-1 \rrbracket$. On considère l'application (dont on admet la linéarité)

$$\begin{aligned} \psi_k : \mathbb{R}_3[X] &\longrightarrow \mathbb{R}^4 \\ P &\longmapsto (P(\alpha_k), P(\alpha_{k+1}), P'(\alpha_k), P'(\alpha_{k+1})) \end{aligned}$$

On note $\mathcal{B} = \{e_1, e_2, e_3, e_4\}$ la base canonique de $\mathbb{R}^4$.

4/ Etablir que ψ_k est injective.

5/ Déterminer explicitement un polynôme $P_{1,k}$ de $\mathbb{R}_3[X]$ tel que : $\psi_k(P_{1,k}) = e_1$.

6/ Déterminer explicitement un polynôme $P_{3,k}$ de $\mathbb{R}_3[X]$ tel que : $\psi_k(P_{3,k}) = e_3$.

7/ En déduire les expressions des polynômes $P_{2,k}$ et $P_{4,k}$ de $\mathbb{R}_3[X]$ tels que : $\psi_k(P_{2,k}) = e_2$ et $\psi_k(P_{4,k}) = e_4$.

8/ Etablir que l'application ψ_k est un isomorphisme de $\mathbb{R}_3[X]$ dans $\mathbb{R}^4$.

9/ Soit $(y_0, y_1, z_0, z_1) \in \mathbb{R}^4$. Exprimer en fonction de $P_{1,k}, \dots, P_{4,k}$ l'unique polynôme Q de $\mathbb{R}_3[X]$ tel que :

$$Q(\alpha_k) = y_0; \quad Q(\alpha_{k+1}) = y_1; \quad Q'(\alpha_k) = z_0 \quad \text{et} \quad Q'(\alpha_{k+1}) = z_1$$

PARTIE C - Interpolation polynomiale avec contraintes étendue

On définit les polynômes $P_0, \dots, P_n$ de $\mathbb{R}_{2n+1}[X]$ en posant :

$$\forall k \in \llbracket 0, n \rrbracket, \quad P_k = \prod_{i=0, i \neq k}^n (X - \alpha_i)^2$$

On définit également les polynômes $Q_0, \dots, Q_n$ de $\mathbb{R}_{2n+1}[X]$ en posant :

$$\forall k \in \llbracket 0, n \rrbracket, \quad Q_k = (X - \alpha_k) \prod_{i=0, i \neq k}^n (X - \alpha_i)^2$$

10/ Soit k un entier naturel. Préciser le degré de P_k , et celui de Q_k .

- 11/ Soit k un entier naturel. Etablir que $P_k(\alpha_k) \neq 0$, ainsi que $Q_k(\alpha_k) = 0$ et $Q'_k(\alpha_k) \neq 0$.
- 12/ Montrer que la famille $B = (P_0, \dots, P_n, Q_0, \dots, Q_n)$ est une famille libre ; puis justifier que c'est une base de $\mathbb{R}_{2n+1}[X]$.
- 13/ On considère l'application (dont on admet la linéarité)

$$\begin{array}{ccc} \rho : \mathbb{R}_{2n+1}[X] & \longrightarrow & \mathbb{R}^{2n+2} \\ P & \longmapsto & (P(\alpha_0), \dots, P(\alpha_n), P'(\alpha_0), \dots, P'(\alpha_n)) \end{array}$$

- a/ Etablir que ρ est injective.
- b/ Etablir que la famille $\mathcal{F}' = \{\rho(P_0), \dots, \rho(P_n), \rho(Q_0), \dots, \rho(Q_n)\}$ est une famille libre de $\mathbb{R}^{2n+2}$.
- c/ Justifier que ρ est un isomorphisme de $\mathbb{R}_{2n+1}[X]$ dans $\mathbb{R}^{2n+2}$.
- 14/ Soient $y_0, \dots, y_n, z_0, \dots, z_n$ ($2n+2$) nombres réels. Etablir qu'il existe un unique polynôme R de $\mathbb{R}_{2n+1}[X]$ tel que :

$$\forall k \in \llbracket 0, n \rrbracket, [R(\alpha_k) = y_k \quad \wedge \quad R'(\alpha_k) = z_k]$$

Remarque. Ce résultat assure l'existence et l'unicité d'un polynôme de degré au plus $2n+1$ passant par $(n+1)$ points du plan donnés avec des tangentes imposées en ces points. L'inconvénient majeur de cette méthode est que plus le nombre de points augmente, plus le degré du polynôme d'interpolation augmente ; et plus ce degré augmente, plus la machine (qui sera chargée de faire les calculs) sera susceptible de commettre des erreurs importantes.

Une "juste milieu" consiste en fait à recoller les polynômes d'interpolation de degré au plus 3 obtenus sur chaque intervalle $[\alpha_k, \alpha_{k+1}]$ pour obtenir ce que l'on appelle une courbe polynomiale par morceaux, ou encore une *B-spline*.

PARTIE D - Interpolation par des B-splines

- 15/ Ecrire en langage Python une fonction `Poly4(a,b)` qui reçoit comme paramètres deux flottants a et b , et qui retourne les 4 polynômes P_1, P_2, P_3 et P_4 tels que :

$$\begin{array}{ll} P_1(a) = 1 \text{ et } P_1(b) = P'_1(a) = P'_1(b) = 0 ; & P_3(b) = 1 \text{ et } P_3(a) = P'_3(a) = P'_3(b) = 0 ; \\ P'_2(a) = 1 \text{ et } P_2(a) = P_2(b) = P'_2(b) = 0 ; & P'_4(b) = 1 \text{ et } P_4(a) = P_4(b) = P'_4(b) = 0. \end{array}$$

Indication : en Python, le polynôme $(X-a)(X-b)(X-c)$ peut être défini par l'instruction `poly1d([a,b,c])`.

- 16/ Ecrire en langage Python une fonction `Interp(LX,LY,LP)` qui reçoit comme paramètres trois listes de longueur 2, $LX = [a, b]$, $LY = [y_a, y_b]$, et $LP = [p_a, p_b]$, et qui retourne le polynôme Q de $\mathbb{R}_3[X]$ tel que :

$$Q(a) = y_a, Q(b) = y_b, Q'(a) = p_a \text{ et } Q'(b) = p_b$$

- 17/ Ecrire en langage Python une fonction `InterpGLUE(LABS,LORD,LPEN)` qui reçoit comme paramètres trois listes de longueur $(n+1)$, $LABS = [x_0, \dots, x_n]$, $LORD = [y_0, \dots, y_n]$, et $LPEN = [p_0, \dots, p_n]$, et qui retourne la liste des polynômes $[Q_0, \dots, Q_{n-1}]$ tels que pour tout entier $k \in \llbracket 0, n-1 \rrbracket$, Q_k est l'unique polynôme de $\mathbb{R}_3[X]$ vérifiant

$$Q_k(x_k) = y_k, Q_k(x_{k+1}) = y_{k+1}, Q'_k(x_k) = p_k \text{ et } Q'_k(x_{k+1}) = p_{k+1}$$

CORRIGÉ.

PARTIE A - Interpolation polynomiale

1/ Soit P un polynôme de $\mathbb{R}_p[X]$ tel que $\varphi_p(P) = 0$. Par définition de l'application φ_p , P a au moins $(n+1)$ racines $(\alpha_0, \dots, \alpha_n)$. Or P est de degré inférieur ou égal à p (donc à n) par hypothèse, et il est donc nécessairement nul (principe du prolongement algébrique).

Par conséquent : $\ker \varphi_p = \{0\}$, ce qui prouve l'injectivité de φ_p .

Remarque. Une ultime remarque sur ce thème : vous pouvez noter $\tilde{0}$ ou $0_{\mathbb{K}_3[X]}$ le polynôme nul, et il existe peut-être des cas ambigus (je n'en connais cependant aucun) où cette notation est indispensable. On ne vous en voudra pas si vous le notez simplement 0. L'important est que vous sachiez ce dont vous parlez, et de ne pas écrire qu'un polynôme est égal à un vecteur de $\mathbb{R}^n$ ou à une matrice.

2/ En revanche, si $p \geq n+1$, alors P est dans le noyau de φ_p si et seulement si il admet pour racines $\alpha_0, \dots, \alpha_n$, c'est-à-dire si et seulement si il est divisible par le polynôme $U = \prod_{i=0}^n (X - \alpha_i)$. Un polynôme P de $\mathbb{R}_p[X]$ est donc dans le noyau de φ_p si et seulement si il existe un polynôme Q tel que : $P = QU$. Dans cette égalité, le terme de gauche (P) est de degré inférieur ou égal à p ; et celui de droite est de degré : $n+1 + \deg Q$. On en déduit que : $n+1 + \deg Q \leq p$, c'est-à-dire : $\deg Q \leq p - n - 1$, d'où : $Q \in \mathbb{R}_{p-n-1}[X]$.

En conclusion, on a montré que : $\ker \varphi_p = \{P \in \mathbb{R}_p[X] \mid \exists Q \in \mathbb{R}_{p-n-1}[X], P = QU\}$.

On en déduit qu'une base de $\ker \varphi_p$ est la famille $\mathcal{B} = (U, XU, \dots, X^{p-n-1}U)$, et que $\dim(\ker \varphi_p) = p - n$.

3/ a/ Supposons qu'il existe $(n+1)$ scalaires $\beta_0, \dots, \beta_n$ tels que : $\sum_{k=0}^n \beta_k L_k = 0$.

On observe alors que pour tout $i \in \llbracket 0, n \rrbracket$ on a : $\sum_{k=0}^n \beta_k L_k(\alpha_i) = \beta_i$ (par construction des polynômes de Lagrange).

Par conséquent : $\forall i \in \llbracket 0, n \rrbracket, \beta_i = 0$. Ce qui prouve que la famille $\mathcal{F}$ est libre.

Puisqu'en outre le cardinal de $\mathcal{F}$ est égal à la dimension de $\mathbb{R}_n[X]$, $\mathcal{F}$ est une base de $\mathbb{R}_n[X]$.

Remarque. La maladresse ultime ici consistait à écrire (sans le justifier) que la famille $\mathcal{F} = \{L_0, \dots, L_n\}$ est génératrice, puis à établir (plus ou moins rigoureusement) qu'elle était libre, et conclure. Puisque $\mathcal{F}$ est de cardinal égal à la dimension de $\mathbb{R}_n[X]$, il est équivalent d'établir que $\mathcal{F}$ est génératrice ou libre.

Donc, pour être explicite, en avançant sans le justifier que $\mathcal{F}$ est génératrice, vous admettez la réponse à la question.

b/ D'après la question précédente, on a : $\text{Im}(\varphi_n) = \text{Vect}(\varphi_n(L_0), \dots, \varphi_n(L_n))$.

Remarque. Il est totalement inapproprié d'utiliser ici la base canonique de $\mathbb{R}_n[X]$, choix qui conduit à des calculs inexploitable. Une fois encore, le choix de la base $\{L_0, \dots, L_n\}$ était commandé par la règle pratique : "je cherche la réponse à la question n en m'inspirant de la question $n-1$ "¹⁰.

Or il résulte de la définition de φ_n d'une part et de la construction des polynômes de Lagrange d'autre part que pour tout entier k entre 0 et n on a : $\varphi_n(L_k) = e_{k+1}$, où e_{k+1} désigne le $(k+1)$ -ième vecteur de la base canonique de $\mathbb{R}^{n+1}$.

10. Pour $n \in \mathbb{N}$, $n \geq 2$, bien entendu.

Par suite : $\boxed{\text{Im}(\varphi_n) = \mathbb{R}^{n+1}}$. D'où φ_n est surjective. Puisque φ_n est également injective (cf question 1), on peut conclure que $\boxed{\varphi_n \text{ est un isomorphisme de } \mathbb{R}_n[X] \text{ dans } \mathbb{R}^{n+1}}$.

PARTIE B - Interpolation polynomiale avec contraintes

Soit k un entier quelconque dans $\llbracket 0, n-1 \rrbracket$. On considère l'application (dont on admet la linéarité)

$$\begin{aligned} \psi_k : \mathbb{R}_3[X] &\longrightarrow \mathbb{R}^4 \\ P &\longmapsto (P(\alpha_k), P(\alpha_{k+1}), P'(\alpha_k), P'(\alpha_{k+1})) \end{aligned}$$

4/ Soit $P \in \mathbb{R}_3[X]$. Si $\psi_k(P) = 0$, alors P admet deux racines doubles distinctes (α_k et α_{k+1}), c'est-à-dire quatre racines (comptées avec multiplicité). Puisque P est de degré au plus 3, on en déduit qu'il est nul. Ce qui assure que : $\ker(\psi_k) = \{0\}$.

Conclusion. L'application ψ_k est injective.

Remarque. Dans certaines copies, il était ici fait référence à la question 1. Même avec beaucoup d'imagination, le raisonnement est difficile à suivre, étant donné que l'application de la question 1 est définie différemment de celle-ci. Il revenait donc exactement au même d'écrire : " ψ_k est injective car la fonction exponentielle est strictement croissante sur $\mathbb{R}$ "; les deux assertions sont vraies, mais ne sont aucunement liées.

5/ On cherche un polynôme $P_{1,k}$ de $\mathbb{R}_3[X]$ tel que $\psi_k(P_{1,k}) = e_1$, c'est-à-dire tel que :

$$P_{1,k}(\alpha_k) = 1 \quad \text{et} \quad P_{1,k}(\alpha_{k+1}) = P'_{1,k}(\alpha_k) = P'_{1,k}(\alpha_{k+1}) = 0$$

Un tel polynôme admet α_{k+1} pour racine double. Puisqu'en outre son degré est inférieur ou égal à 3, il existe deux réels a et b tels que :

$$P_{1,k} = (aX + b)(X - \alpha_{k+1})^2$$

$$\text{On a alors : } \begin{cases} P_{1,k}(\alpha_k) = 1 \\ P'_{1,k}(\alpha_k) = 0 \end{cases} \iff \begin{cases} (a\alpha_k + b)(\alpha_k - \alpha_{k+1})^2 = 1 \\ a(\alpha_k - \alpha_{k+1})^2 + 2(a\alpha_k + b)(\alpha_k - \alpha_{k+1}) = 0 \end{cases}$$

$$\iff \begin{cases} a\alpha_k + b = \frac{1}{(\alpha_k - \alpha_{k+1})^2} \\ a(\alpha_k - \alpha_{k+1}) + 2a\alpha_k + 2b = 0 \end{cases} \iff \begin{cases} a\alpha_k + b = \frac{1}{(\alpha_k - \alpha_{k+1})^2} \\ a(3\alpha_k - \alpha_{k+1}) + 2b = 0 \end{cases} \iff \begin{cases} a = \frac{2}{(\alpha_{k+1} - \alpha_k)^3} \\ b = \frac{3\alpha_k - \alpha_{k+1}}{(\alpha_k - \alpha_{k+1})^3} \end{cases}$$

$$\iff \begin{cases} a = \frac{2}{(\alpha_{k+1} - \alpha_k)^3} \\ b = \frac{3\alpha_k - \alpha_{k+1}}{(\alpha_k - \alpha_{k+1})^3} \end{cases}$$

Conclusion. Il existe un unique polynôme $P_{1,k}$ tel que $\psi_k(P_{1,k}) = e_1$.

$$\text{Explicitement : } P_{1,k} = \frac{1}{(\alpha_{k+1} - \alpha_k)^3} (2X + \alpha_{k+1} - 3\alpha_k)(X - \alpha_{k+1})^2$$

6/ On cherche un polynôme $P_{3,k}$ de $\mathbb{R}_3[X]$ tel que $\psi_k(P_{3,k}) = e_3$, c'est-à-dire tel que :

$$P'_{3,k}(\alpha_k) = 1 \quad \text{et} \quad P_{3,k}(\alpha_k) = P_{3,k}(\alpha_{k+1}) = P'_{3,k}(\alpha_{k+1}) = 0$$

Un tel polynôme admet α_{k+1} pour racine double et α_k pour racine simple. Puisqu'en outre son degré est inférieur ou égal à 3, il existe un réel a tel que :

$$P_{3,k} = a(X - \alpha_k)(X - \alpha_{k+1})^2$$

$$\text{On a alors : } P'_{3,k}(\alpha_k) = 1 \iff a(\alpha_k - \alpha_{k+1})^2 = 1 \iff a = \frac{1}{(\alpha_k - \alpha_{k+1})^2}$$

Conclusion. Il existe un unique polynôme $P_{3,k}$ tel que $\psi_k(P_{3,k}) = e_3$.

$$\text{Explicitement : } P_{3,k} = \frac{1}{(\alpha_k - \alpha_{k+1})^2} (X - \alpha_k)(X - \alpha_{k+1})^2$$

7/ On cherche à présent un polynôme $P_{2,k}$ de $\mathbb{R}_3[X]$ tel que $\psi_k(P_{2,k}) = e_2$, c'est à dire tel que :

$$P_{2,k}(\alpha_{k+1}) = 1 \quad \text{et} \quad P_{2,k}(\alpha_k) = P'_{2,k}(\alpha_k) = P'_{2,k}(\alpha_{k+1}) = 0$$

Un tel polynôme admet α_k pour racine double. Puisqu'en outre son degré est inférieur ou égal à 3, il existe deux réels a et b tels que :

$$P_{2,k} = (aX + b)(X - \alpha_k)^2$$

Il suffit de permuter les rôles de α_k et α_{k+1} dans la question 5 pour conclure.

Conclusion. Il existe un unique polynôme $P_{2,k}$ tel que $\psi_k(P_{2,k}) = e_2$.

$$\text{Explicitement : } P_{2,k} = \frac{1}{(\alpha_k - \alpha_{k+1})^3} (2X + \alpha_k - 3\alpha_{k+1})(X - \alpha_k)^2$$

De même, on obtient par symétrie à partir de la question 6 que :

Il existe un unique polynôme $P_{4,k}$ tel que $\psi_k(P_{4,k}) = e_4$.

$$\text{Explicitement : } P_{4,k} = \frac{1}{(\alpha_k - \alpha_{k+1})^2} (X - \alpha_{k+1})(X - \alpha_k)^2$$

8/ Puisque les polynômes $P_{1,k}$, $P_{2,k}$, $P_{3,k}$ et $P_{4,k}$ appartiennent à $\mathbb{R}_3[X]$, on a :

$$\text{Vect}(P_{1,k}, P_{2,k}, P_{3,k}, P_{4,k}) \subset \mathbb{R}_3[X], \text{ d'où : } \text{Vect}(\psi_k(P_{1,k}), \psi_k(P_{2,k}), \psi_k(P_{3,k}), \psi_k(P_{4,k})) \subset \text{Im}(\psi_k) \quad (\spadesuit).$$

Or, d'après les questions 5 à 7 : $\text{Vect}(\psi_k(P_{1,k}), \psi_k(P_{2,k}), \psi_k(P_{3,k}), \psi_k(P_{4,k})) = \text{Vect}(e_1, e_2, e_3, e_4) = \mathbb{R}^4 \quad (\clubsuit)$.

On déduit de $(\spadesuit)$ et $(\clubsuit)$ que l'application ψ_k est surjective. Puisque l'on a déjà établi que ψ_k est injective (question 4), on peut conclure : ψ_k est un isomorphisme de $\mathbb{R}_3[X]$ dans $\mathbb{R}^4$.

9/ D'après les questions 5 à 8, le polynôme $Q = y_0 P_{1,k} + y_1 P_{3,k} + z_0 P_{2,k} + z_1 P_{4,k}$ est l'unique antécédent de (y_0, z_0, y_1, z_1) par l'application ψ_k . C'est donc l'unique polynôme de $\mathbb{R}_3[X]$ tel que :

$$Q(\alpha_k) = y_0; \quad Q(\alpha_{k+1}) = y_1; \quad Q'(\alpha_k) = z_0 \quad \text{et} \quad Q'(\alpha_{k+1}) = z_1$$

PARTIE C - Interpolation polynomiale avec contraintes étendue

On définit les polynômes $P_0, \dots, P_n$ de $\mathbb{R}_{2n+1}[X]$ en posant :

$$\forall k \in \llbracket 0, n \rrbracket, \quad P_k = \prod_{i=0, i \neq k}^n (X - \alpha_i)^2$$

On définit également les polynômes $Q_0, \dots, Q_n$ de $\mathbb{R}_{2n+1}[X]$ en posant :

$$\forall k \in \llbracket 0, n \rrbracket, \quad Q_k = (X - \alpha_k) \prod_{i=0, i \neq k}^n (X - \alpha_i)^2$$

10/ Soit k un entier naturel. On a : $\deg(P_k) = 2n$ et $\deg(Q_k) = 2n + 1$.

Remarque. Attention : l'intervalle d'entiers $\llbracket 0, n \rrbracket$ contient $(n+1)$ valeurs. En outre, pour deux polynômes P et Q , on a : $\deg(PQ) = \deg(P) + \deg(Q)$ (et non $\deg(P) \times \deg(Q)$).

11/ Soit k un entier naturel. Il résulte de la définition de P_k que $P_k(\alpha_k) \neq 0$ puisque les racines de P_k sont connues, et que α_k n'est pas l'une d'entre elles.

Il résulte par ailleurs de la définition de Q_k que α_k est racine simple de Q_k , d'où : $Q_k(\alpha_k) = 0$ et $Q'_k(\alpha_k) \neq 0$.

Remarque. Observez que la définition de l'énoncé ($P_k = \prod_{i=0, i \neq k}^n (X - \alpha_i)^2$) donne la décomposition en irréductibles dans $\mathbb{R}[X]$ de P_k (la même remarque vaut pour Q_k). Les racines de P_k sont donc les scalaires α_i avec $i \neq k$, et toutes sont de multiplicité exactement égale à 2.

12/ Soient $\beta_0, \dots, \beta_n, \gamma_0, \dots, \gamma_n$ tels que : $\sum_{k=0}^n \beta_k P_k + \sum_{k=0}^n \gamma_k Q_k = 0$ ($\spadesuit$)

En évaluant la relation ($\spadesuit$) en α_i , avec $i \in \llbracket 0, n \rrbracket$, on obtient : $\beta_i P_i(\alpha_i) = 0$. Puisque $P_i(\alpha_i) \neq 0$ (question précédente), on en déduit que $\beta_i = 0$.

Par suite : $\sum_{k=0}^n \gamma_k Q_k = 0$. En dérivant cette relation, et en l'évaluant en α_i , avec $i \in \llbracket 0, n \rrbracket$, on obtient : $\gamma_i Q'_i(\alpha_i) = 0$. Puisque $Q'_i(\alpha_i) \neq 0$ (question précédente), on en déduit que $\gamma_i = 0$.

Finalement : $\left[\sum_{k=0}^n \beta_k P_k + \sum_{k=0}^n \gamma_k Q_k = 0 \right] \implies [\forall k \in \llbracket 0, n \rrbracket, \beta_k = 0 \wedge \gamma_k = 0]$.

La famille $B = (P_0, \dots, P_n, Q_0, \dots, Q_n)$ est donc une famille libre de polynômes de $\mathbb{R}_{2n+1}[X]$. Puisque son cardinal $(2n+2)$ est égal à la dimension de $\mathbb{R}_{2n+1}[X]$, c'est une base de $\mathbb{R}_{2n+1}[X]$.

13/ a/ Soit $P \in \mathbb{R}_{2n+1}[X]$. Si $\rho(P) = 0$, alors P admet $(n+1)$ racines doubles distinctes (tous les α_k), c'est-à-dire $(2n+2)$ racines (comptées avec multiplicité). Puisque P est de degré au plus $2n+1$, on en déduit qu'il est nul. Ce qui assure que : $\ker(\rho) = \{0\}$.

Conclusion. L'application ρ est injective.

b/ D'après les questions précédentes, l'application ρ est injective, et la famille $(P_0, \dots, P_n, Q_0, \dots, Q_n)$ est une base de $\mathbb{R}_{2n+1}[X]$. Puisque l'image d'une base par une application injective est une famille libre de l'espace d'arrivée, on peut conclure que :

la famille $\mathcal{F}' = \{\rho(P_0), \dots, \rho(P_n), \rho(Q_0), \dots, \rho(Q_n)\}$ est une famille libre de $\mathbb{R}^{2n+2}$.

c/ D'après le cours, on a : $\text{Im}(\rho) = \text{Vect}(\mathcal{F}')$. Or d'après la question précédente, $\mathcal{F}'$ est une base de $\mathbb{R}^{2n+2}$, puisque c'est une famille libre de cardinal $(2n+2)$ dans un espace vectoriel de dimension $(2n+2)$. Par conséquent : $\text{Vect}(\mathcal{F}') = \mathbb{R}^{2n+2}$. Ce qui implique que ρ est surjective.

Puisqu'en outre ρ est injective (question 13-a), on peut conclure.

Conclusion. L'application ρ est un isomorphisme de $\mathbb{R}_{2n+1}[X]$ dans $\mathbb{R}^{2n+2}$.

14/ Soient $y_0, \dots, y_n, z_0, \dots, z_n$ ($2n+2$) nombres réels. D'après la question précédente, il existe un unique polynôme R dans $\mathbb{R}_{2n+1}[X]$ tel que $\rho(R) = (y_0, \dots, y_n, z_0, \dots, z_n)$ (puisque ρ est un isomorphisme de $\mathbb{R}_{2n+1}[X]$ dans $\mathbb{R}^{2n+2}$).

En d'autres termes, il existe un unique polynôme R de $\mathbb{R}_{2n+1}[X]$ (et $R = \rho^{-1}(y_0, \dots, y_n, z_0, \dots, z_n)$) tel que : $\forall k \in \llbracket 0, n \rrbracket, [R(\alpha_k) = y_k \quad \wedge \quad R'(\alpha_k) = z_k]$.

PARTIE D - Interpolation par des B-splines

15/ La fonction `Poly4(a,b)` ci-dessous reçoit comme paramètres deux flottants a et b , et retourne les 4 polynômes P_1, P_2, P_3 et P_4 tels que :

$$\begin{aligned} P_1(a) &= 1 \text{ et } P_1(b) = P'_1(a) = P'_1(b) = 0; & P_3(b) &= 1 \text{ et } P_3(a) = P'_3(a) = P'_3(b) = 0; \\ P'_2(a) &= 1 \text{ et } P_2(a) = P_2(b) = P'_2(b) = 0; & P'_4(b) &= 1 \text{ et } P_4(a) = P_4(b) = P'_4(b) = 0. \end{aligned}$$

```

1 def Poly4(a,b):
2     P1 = (((b-a)**(-3))*2)*np.poly1d([(3*a-b)/2,b,b], True)
3     P2 = ((b-a)**(-2))*np.poly1d([a,b,b], True)
4     P3 = (((a-b)**(-3))*2)*np.poly1d([(3*b-a)/2,a,a], True)
5     P4 = ((b-a)**(-2))*np.poly1d([a,a,b], True)
6     return P1, P2, P3, P4

```

16/ La fonction `Interp(LX,LY,LP)` ci-dessous reçoit comme paramètres trois listes de longueur 2, $\text{LX} = [a, b]$, $\text{LY} = [y_a, y_b]$, et $\text{LP} = [p_a, p_b]$, et retourne le polynôme Q de $\mathbb{R}_3[X]$ tel que :

$$Q(a) = y_a, Q(b) = y_b, Q'(a) = p_a \text{ et } Q'(b) = p_b$$

```

1 def Interp(LX,LY,LP):
2     LISPOL = Poly4(LX[0],LX[1])
3     Q = LY[0]*LISPOL[0] + LP[0]*LISPOL[1] + LY[1]*LISPOL[2] + LP[1]*LISPOL[3]
4     return Q

```

17/ La fonction `InterpGLUE(LABS,LORD,LPEN)` reçoit comme paramètres trois listes de longueur $(n+1)$, $\text{LABS} = [x_0, \dots, x_n]$, $\text{LORD} = [y_0, \dots, y_n]$, et $\text{LPEN} = [p_0, \dots, p_n]$, et retourne la liste des polynômes $[Q_0, \dots, Q_{n-1}]$ tels que pour tout entier $k \in \llbracket 0, n-1 \rrbracket$, Q_k est l'unique polynôme de $\mathbb{R}_3[X]$ vérifiant

$$Q_k(x_k) = y_k, Q_k(x_{k+1}) = y_{k+1}, Q'_k(x_k) = p_k \text{ et } Q'_k(x_{k+1}) = p_{k+1}$$

```

1 def InterpGLUE(LABS,LORD,LPEN):
2     ListPOL = []
3     N = len(LABS)
4     for i in range(N-1):
5         Q = Interp([LABS[i],LABS[i+1]], [LORD[i],LORD[i+1]], [LPEN[i],LPEN[i+1]])
6         ListPOL = ListPOL + [Q]
7     return ListPOL

```

COMPLÉMENT

Finalelement, l'interpolation, comment ça marche et à quoi ça sert ?

1/ Comment ça marche ?

On cherche à interpoler les points du plan de coordonnées suivantes :

$$(0, 2); \quad (1, 1); \quad (1.5, -2); \quad (3, 1); \quad (4, -1); \quad (7, 1)$$

c'est-à-dire à construire une courbe passant par ces points. Si l'on souhaite que cette courbe soit lisse, on peut chercher une fonction polynomiale répondant au problème.

L'idée qui motive la construction des polynômes de Lagrange est la suivante. Pour obtenir facilement l'équation d'une courbe passant par ces 6 points, on cherche un polynôme de degré assez grand (au plus 5 dans le cas présent) en commençant par renommer les abscisses des points :

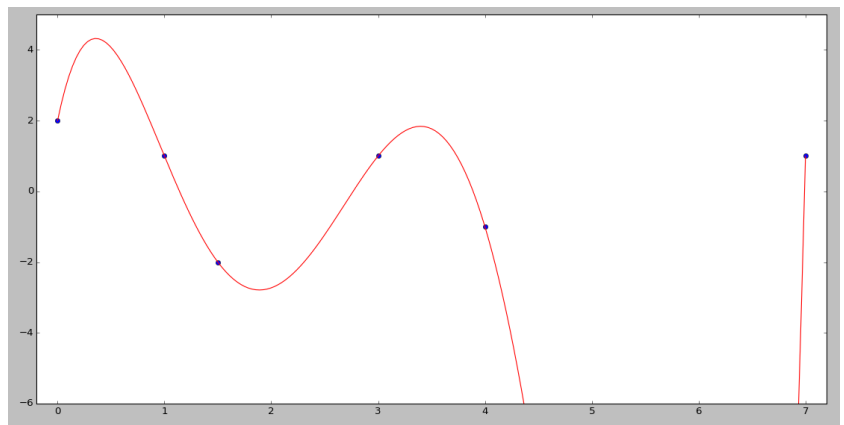
$$\alpha_0 = 0; \quad \alpha_1 = 1; \quad \alpha_2 = 1.5; \quad \alpha_3 = 3; \quad \alpha_4 = 4; \quad \alpha_5 = 7$$

Puis on construit les polynômes $L_0, \dots, L_5$ qui prennent la valeur 1 en l'un des α_i et la valeur 0 en tous les autres α_i . Avec cette merveilleuse famille, il suffit de poser :

$$Q = 2 \times L_0 + 1 \times L_1 - 2 \times L_2 + 1 \times L_3 - 1 \times L_4 + 1 \times L_5$$

pour obtenir un polynôme Q dont la courbe représentative passe effectivement par les 6 points.

Le problème est que l'approximation alors obtenue est de très mauvaise qualité en-dehors des points d'interpolation, comme l'illustre la figure ci-contre (où j'ai découpé une partie de la courbe représentative de Q , qui “descend jusque dans les -40 ”... Ce qui commence à faire un rude détour!).



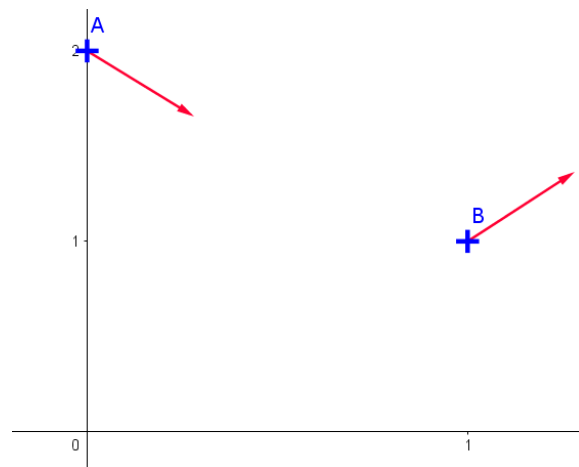
D'où l'idée d'attaquer le problème un peu plus finement, en fixant non seulement les points d'interpolation, mais également les tangentes en ces points (c'est-à-dire les valeurs et les valeurs des dérivées).

La première méthode pour parvenir à ce but est de construire des polynômes de “Super-Lagrange” (les P_k et Q_k de la partie C); elle est toutefois numériquement risquée, dans le sens où les polynômes d'interpolation sont alors de degré “très vite très grands”, et que les erreurs de calcul risquent donc d'exploser assez vite.

La seconde méthode consiste à “casser le problème en petits morceaux” (c’est l’analogue de la célèbre règle en Informatique : “diviser pour régner”), et à chercher une solution sur un seul sous-intervalle de la forme $[\alpha_k, \alpha_{k+1}]$.

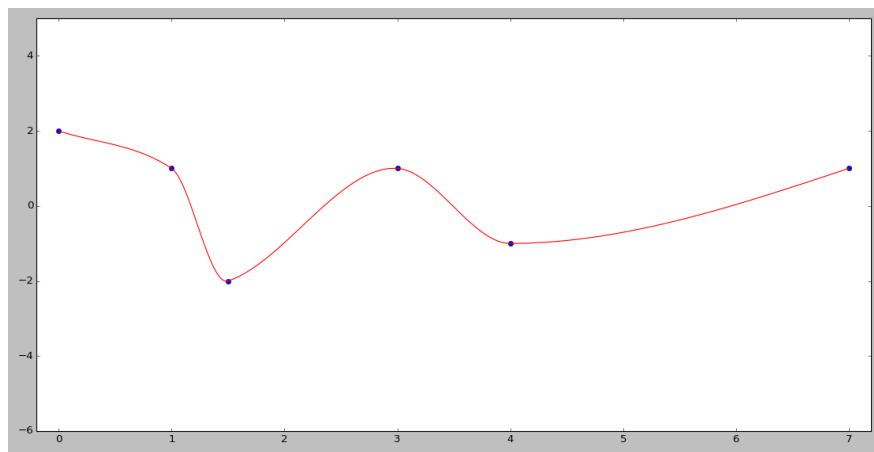
Pour cela, il “suffit” de construire non plus deux polynômes prenant les valeurs 0 et 1 en α_k et α_{k+1} , mais quatre polynômes : deux pour contrôler les valeurs en α_k et α_{k+1} , et deux pour contrôler les valeurs des dérivées en α_k et α_{k+1} .

Le détail de cette construction faisait l’objet de la partie B.



Une fois obtenu le polynôme désiré sur un sous-intervalle, on réitère la construction sur chaque sous-intervalle de $[\alpha_0, \alpha_n]$ pour obtenir une **courbe polynomiale par morceaux**, également appelée *B-spline* (ce processus de “recollement” faisait l’objet de la partie D).

Le fait de contrôler également les valeurs des dérivées permet de construire une approximation bien meilleure qu’avec les polynômes de Lagrange. A titre d’illustration, la courbe ci-dessous est celle de la *B-spline* passant par les 6 points donnés au début de cette discussion. La comparaison avec la figure de la page précédente vaut la peine !...



2/ A quoi ça sert ?

A titre culturel, les courbes B-splines sont utilisées dans les logiciels de CAO (par exemple en conception automobile¹¹), en Physique moléculaire¹², en Statistiques pour obtenir des fonctions de lissage (autres que les droites de régression par exemple) et par suite dans l’obtention de modèles correspondant à des mesures physico-chimiques, ainsi que pour la conception de lentilles ophtalmiques, ce qui n’est déjà pas rien.

11. D’ailleurs, la notion de courbe B-spline est très liée à celle de courbe de Bézier, cette dernière devant son nom à Pierre Bézier, qui était ingénieur chez Renault, et qui a utilisé ces (ses!) courbes pour le dessin et la conception de pièces.

12. J’ai trouvé sur le web quelques articles sur ce thème, auquel je confesse ne comprendre aucun mot en dehors de “B-spline”.

Chapitre 23

Intégrale sur un segment

Tout au long de ce chapitre, a et b désignent deux réels tels que $a < b$.

En Terminale, l'intégrale d'une fonction continue sur le segment $[a, b]$ a été définie comme "l'aire sous la courbe...". Or cette définition, assez intuitive géométriquement, doit être précisée dans le sens où l'aire d'un domaine n'est connue et définie que dans des situations très particulières.

L'idée maîtresse pour donner un sens à $\int_a^b f$ est qu'en dernière analyse, la seule aire que l'on sache vraiment bien calculer est celle d'un rectangle... On peut donc définir l'intégrale d'une fonction constante (c'est l'aire d'un rectangle); puis l'intégrale d'une fonction en escalier, c'est-à-dire constante par morceaux (comme une somme d'aires de rectangles). En "passant à la limite", on pourra alors donner un sens rigoureux à l'intégrale d'une fonction continue (et même continue par morceaux) sur un segment.

23.1 Continuité uniforme

DÉFINITION 23.1 - Soit $f : I \longrightarrow \mathbb{K}$ une fonction à valeurs dans $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$. f est **uniformément continue** sur I si :

$$\forall \varepsilon > 0, \exists \alpha > 0, \forall (x, y) \in I^2, |y - x| < \alpha \implies |f(y) - f(x)| < \varepsilon$$

Exemple. La fonction carrée $x \in [0, 1] \longmapsto x^2$ est uniformément continue sur $[0, 1]$.

Contre-exemple. La fonction carrée $x \in [0, +\infty[\longmapsto x^2$ n'est pas uniformément continue sur $[0, +\infty[$.

La définition de continuité uniforme peut vous faire penser à celle de fonction lipschitzienne (introduite dans le chapitre portant sur la dérivabilité).¹ De fait les deux notions sont liées par l'énoncé ci-dessous.

PROPRIÉTÉ 23.1 - Si f est lipschitzienne sur I , alors f est uniformément continue sur I .

PREUVE. En exercice. Idée : il suffit d'utiliser les définitions de fonction lipschitzienne et de continuité uniforme, de fixer $\varepsilon > 0$ et de prendre $\alpha = \varepsilon/k$. ☕

1. Rappelons qu'une fonction $f : I \longrightarrow \mathbb{K}$ est lipschitzienne s'il existe un réel positif k tel que : $\forall (x, y) \in I^2, |f(y) - f(x)| \leq k \times |y - x|$.

L'implication réciproque est FAUSSE, comme l'illustre l'exemple suivant.

Contre-exemple. La fonction racine carrée $x \in [0, 1] \mapsto \sqrt{x}$ est uniformément continue sur $[0, 1]$, mais n'est pas lipschitzienne sur $[0, 1]$.

En particulier :

PROPRIÉTÉ 23.2 - Si f est de classe $\mathcal{C}^1$ et à dérivée bornée sur I , alors f est uniformément continue sur I .

PREUVE. En exercice. Idée : c'est une révision d'une propriété déjà vue cette année. Si f est de classe $\mathcal{C}^1$ et à dérivée bornée sur I , alors f est lipschitzienne sur I (conséquence du TAF). D'où la conclusion grâce à la propriété précédente. ☕

PROPRIÉTÉ 23.3 - Si f est uniformément continue sur I , alors f est continue sur I .

PREUVE. Triviale. ☕

Le point clef de ce paragraphe, voire de ce chapitre, est que la réciproque de la propriété précédente est vraie, sous réserve que l'on se restreigne à un segment.

THÉORÈME 23.1 - (Heine). Soit $f : [a, b] \rightarrow \mathbb{R}$.

$$[f \text{ est uniformément continue sur } [a, b]] \quad \text{SSI} \quad [f \text{ est continue sur } [a, b]]$$

PREUVE. Raisonnons par double implication.

"De la droite vers la gauche" : supposons f continue sur $[a, b]$. Raisonnons par l'absurde et supposons que f ne soit pas uniformément continue sur $[a, b]$. Alors :

$$\exists \varepsilon > 0, \forall \alpha > 0, \exists (x, y) \in [a, b]^2, [|y - x| < \alpha] \wedge [|f(y) - f(x)| \geq \varepsilon]$$

Pour un tel $\varepsilon > 0$, on pose $\alpha_n = \frac{1}{n+1}$. Il existe un couple (x_n, y_n) de réels de $[a, b]$ tels que :

$$|y_n - x_n| < \frac{1}{n+1} \text{ et } |f(y_n) - f(x_n)| > \varepsilon$$

En faisant varier n , on obtient deux suites (x_n) et (y_n) d'éléments de $[a, b]$.

En particulier, la suite (x_n) est bornée. D'après la propriété de Bolzano-Weierstrass², on peut en extraire une sous-suite convergente, que nous noterons $(x_{\varphi(n)})$.³ Posons : $c = \lim_{n \rightarrow +\infty} x_{\varphi(n)}$.

Notons alors que pour tout entier naturel n on a :

$$|y_{\varphi(n)} - c| \leq |y_{\varphi(n)} - x_{\varphi(n)}| + |x_{\varphi(n)} - c| \leq \frac{1}{\varphi(n) + 1} + |x_{\varphi(n)} - c|$$

Or $\lim_{n \rightarrow +\infty} \frac{1}{\varphi(n) + 1} = 0$ (puisque $\varphi(n) \geq n$ pour tout n) et $\lim_{n \rightarrow +\infty} |x_{\varphi(n)} - c| = 0$ (puisque la suite $(x_{\varphi(n)})$ converge vers c). Il s'ensuit que : $\lim_{n \rightarrow +\infty} |y_{\varphi(n)} - c| = 0$. Donc la suite $(y_{\varphi(n)})$ converge et a pour limite c .

2. C'est donc à cet endroit précis qu'intervient le fait que l'on travaille sur un segment.

3. Où $\varphi : \mathbb{N} \rightarrow \mathbb{N}$ est strictement croissante, et vérifie donc : $\forall n \in \mathbb{N}, \varphi(n) \geq n$.

La fonction f étant continue en c (par hypothèse), on a alors (par continuité séquentielle) :

$$\lim_{n \rightarrow +\infty} f(x_{\varphi(n)}) = f(c) \text{ et } \lim_{n \rightarrow +\infty} f(y_{\varphi(n)}) = f(c)$$

Or ceci contredit très clairement le fait que : $\forall n \in \mathbb{N}, |f(y_{\varphi(n)}) - f(x_{\varphi(n)})| \geq \varepsilon > 0$.

On en déduit donc que f est uniformément continue sur $[a, b]$, ce qui établit l'implication :

$$[f \text{ est continue sur } [a, b]] \implies [f \text{ est uniformément continue sur } [a, b]].$$

Réciproquement : supposons f uniformément continue sur $[a, b]$. Fixons $x \in [a, b]$, et $\varepsilon > 0$. Puisque f est uniformément continue sur $[a, b]$, il existe $\alpha > 0$ tel que pour tout réel y dans $[a, b]$ on ait : $[|y - x| < \alpha] \implies [f(y) - f(x)] < \varepsilon$, ce qui signifie que la limite de $f(y)$ lorsque y tend vers x existe et vaut $f(x)$, ce qui signifie encore que f est continue en x . Le réel x étant arbitraire dans $[a, b]$, on peut donc affirmer que f est continue sur $[a, b]$.

Ce qui prouve l'implication : $[f \text{ est uniformément continue sur } [a, b]] \implies [f \text{ est continue sur } [a, b]]$.

Conclusion. f est uniformément continue sur $[a, b]$ SSI f est continue sur $[a, b]$.



23.2 Subdivisions

DÉFINITION 23.2 - Une **subdivision** $\sigma = (x_0, \dots, x_n)$ du segment $[a, b]$ est la donnée de $n + 1$ réels $x_0, \dots, x_n$ tels que : $x_0 = a < x_1 < \dots < x_n = b$. On appelle **pas** de la subdivision σ le réel : $p = \max_{i \in \llbracket 0, n \rrbracket} (x_{i+1} - x_i)$.

Une subdivision est dite **régulière** (ou **à pas constant**) lorsque : $\forall i \in \llbracket 0, n \rrbracket, x_{i+1} - x_i = \frac{b - a}{n}$.

Illustration. Ci-dessous, une subdivision (non régulière) $\sigma = (x_0, \dots, x_n)$ d'un intervalle $[a, b]$.

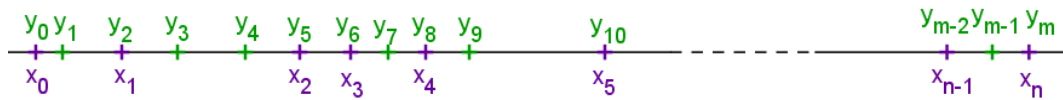


Remarque. Les subdivisions régulières vous sont particulièrement familières, puisque ce sont celles que l'on utilise dans la méthode d'Euler, ainsi que dans les méthodes numériques de calcul intégral.

DÉFINITION 23.3 - Le **support** de la subdivision $\sigma = (x_0, \dots, x_n)$ de $[a, b]$ est l'ensemble noté $\text{Supp}(\sigma) = \{x_0, \dots, x_n\}$.

Une subdivision σ' de $[a, b]$ est **plus fine** que σ si $\text{Supp}(\sigma) \subset \text{Supp}(\sigma')$.

Illustration. Ci-dessous, la subdivision $\sigma' = (y_0, \dots, y_m)$ est plus fine que $\sigma = (x_0, \dots, x_n)$.



Remarque. La relation binaire “être plus fine que” est une relation d'ordre (qui n'est pas totale) sur l'ensemble des subdivisions du segment $[a, b]$ (exercice).

DÉFINITION 23.4 - Soient σ et σ' deux subdivisions de $[a, b]$. La **réunion** des subdivisions σ et σ' , notée $\sigma \vee \sigma'$ est la subdivision de $[a, b]$ dont le support est $\text{Supp}(\sigma) \cup \text{Supp}(\sigma')$.

Lemme. Avec les notations introduites plus haut, $\sigma \vee \sigma'$ est plus fine que σ et σ' .

PREUVE. En exercice (conséquence immédiate des définitions). ☕

23.3 Fonctions en escalier

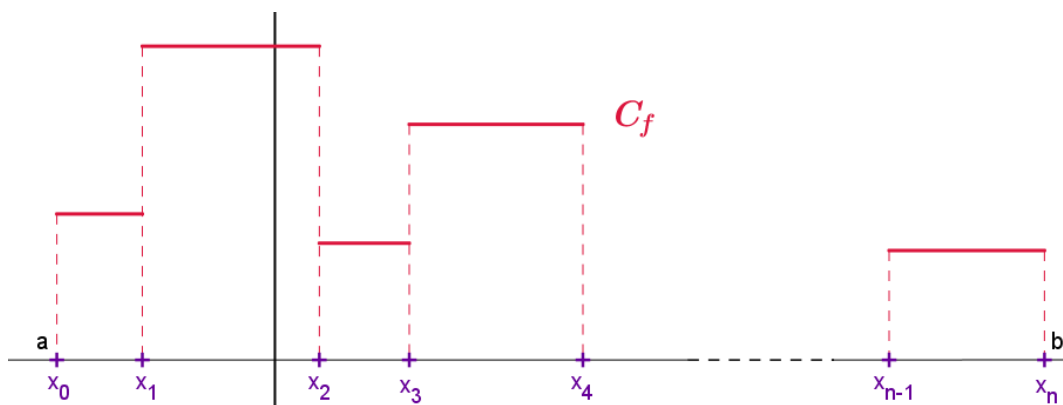
DÉFINITION 23.5 - Une fonction $f : [a, b] \rightarrow \mathbb{R}$ est une **fonction en escalier** (sur $[a, b]$) s'il existe une subdivision $\sigma = (x_0, \dots, x_n)$ de $[a, b]$ et n réels $y_0, \dots, y_{n-1}$ tels que :

$$\forall i \in \llbracket 0, n-1 \rrbracket, \forall x \in]x_i, x_{i+1}[, f(x) = y_i$$

Dans cette situation, on dit que σ est **adaptée** à la fonction f .

Informellement, une fonction en escalier est une fonction “constante par paliers” ; par ailleurs, les valeurs d'une fonction f aux différents points de la subdivision n'ont qu'une importance très relative.

Illustration. Ci-dessous, un exemple de courbe représentation d'une fonction en escalier.



Exemples. Une fonction constante sur un segment est un cas particulier de fonction en escalier. La restriction de la fonction de Heaviside à un segment est une fonction en escalier. Enfin, la restriction de la fonction partie entière à un segment est un autre exemple de fonction en escalier.

Remarque. Observons que si σ est une subdivision adaptée à la fonction f , alors toute subdivision plus fine que σ l'est aussi.

Notation. Par la suite, nous noterons $\mathcal{E}([a, b], \mathbb{R})$ l'ensemble des fonctions en escalier sur $[a, b]$.

PROPRIÉTÉ 23.4 - En notant $\bullet$ la multiplication par un scalaire : $(\mathcal{E}([a, b], \mathbb{R}), +, \bullet)$ est un $\mathbb{R}$ -espace vectoriel.

En outre, $(\mathcal{E}([a, b], \mathbb{R}), +, \times)$ est un anneau commutatif (non intègre).

PREUVE. Essentiellement, les propriétés ci-dessus signifient que les fonctions en escalier sont stables par somme, produit, multiplication par un scalaire. Techniquement, ceci découle du fait que pour deux fonctions en escalier sur $[a, b]$, il existe une subdivision adaptée à f et à g : la réunion d'une subdivision adaptée à f et d'une subdivision adaptée à g . ☕

23.4 Fonctions continues par morceaux

La notion de fonction continue par morceaux généralise celle de fonction continue.

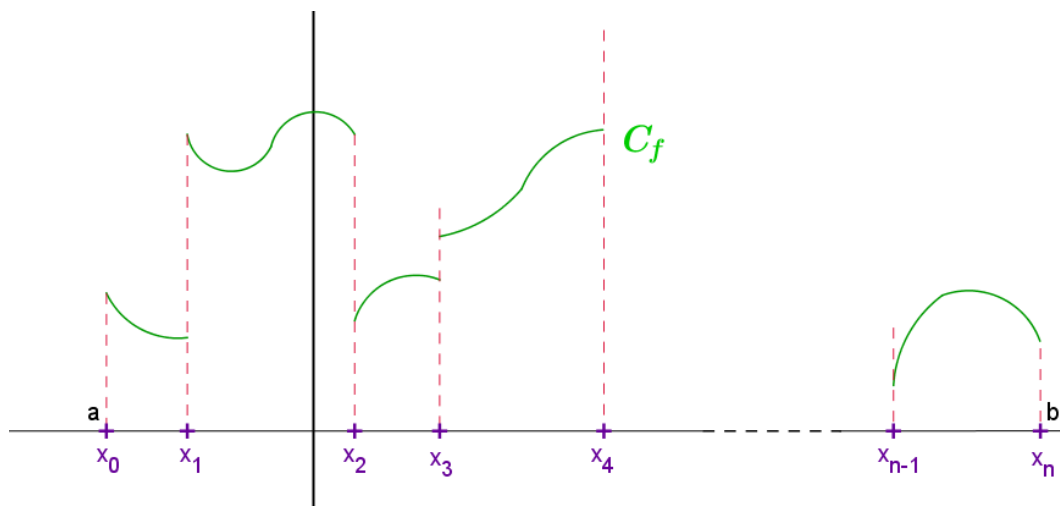
DÉFINITION 23.6 - Une fonction $f : [a, b] \rightarrow \mathbb{R}$ est **continue par morceaux** s'il existe une subdivision

$$\sigma = (x_0, \dots, x_n) \text{ de } [a, b] \text{ telle que :}$$

- 1) $\forall i \in \llbracket 0, n-1 \rrbracket, f|_{]x_i, x_{i+1}[} \in \mathcal{C}^0(]x_i, x_{i+1}[, \mathbb{R})$
- 2) pour tout entier i compris entre 1 et $n-1$, les limites de f à gauche et à droite de x_i existent et sont finies ; la limite de f à droite (*resp.* à gauche) de $x_0 = a$ (*resp.* de $x_n = b$) existe et est finie.

Dans cette situation, on dit encore que σ est **adaptée** à la fonction f .

Illustration. Ci-dessous, un exemple de courbe représentation d'une fonction continue par morceaux.



Remarque. Observons que comme dans le cas des fonctions en escalier, si σ est une subdivision adaptée à la fonction f , alors toute subdivision plus fine que σ l'est aussi.

Notation. On note $\mathcal{C}_{\text{pm}}^0([a, b], \mathbb{R})$ l'ensemble des fonctions continues par morceaux sur $[a, b]$.

Puisqu'il est clair que toute fonction continue (*resp.* en escalier) sur $[a, b]$ est en particulier continue par morceaux sur $[a, b]$, on a l'inclusion :

$$\mathcal{C}^0([a, b], \mathbb{R}) \subset \mathcal{C}_{\text{pm}}^0([a, b], \mathbb{R}) \text{ (resp. } \mathcal{E}([a, b], \mathbb{R}) \subset \mathcal{C}_{\text{pm}}^0([a, b], \mathbb{R}))$$

PROPRIÉTÉ 23.5 - Toute fonction continue par morceaux sur $[a, b]$ est bornée $[a, b]$.

PREUVE. En exo. Idée : pour une fonction continue par morceaux, on peut utiliser le TBA sur chaque morceau⁴ ; en outre, les valeurs de f aux points d'une subdivision adaptée sont en nombre fini. ☕

PROPRIÉTÉ 23.6 - En notant $\bullet$ la multiplication par un scalaire : $(\mathcal{C}_{\text{pm}}^0([a, b], \mathbb{R}), +, \bullet)$ est un $\mathbb{R}$ -espace vectoriel.

En outre $(\mathcal{C}_{\text{pm}}^0([a, b], \mathbb{R}), +, \times)$ est un anneau commutatif (non intègre).

PREUVE. En exercice. Il s'agit quand même d'une vérification un peu rébarbative. Le seul point nécessitant un peu d'attention est que pour f et g deux fonctions continues par morceaux, on peut choisir une subdivision adaptée à la fois à f et g (en prenant la réunion d'une subdivision adaptée à f et d'une subdivision adaptée à g). ☕

De nouveau, les propriétés ci-dessus signifient en particulier que les fonctions continues par morceaux sont stables par somme, produit, multiplication par un scalaire.

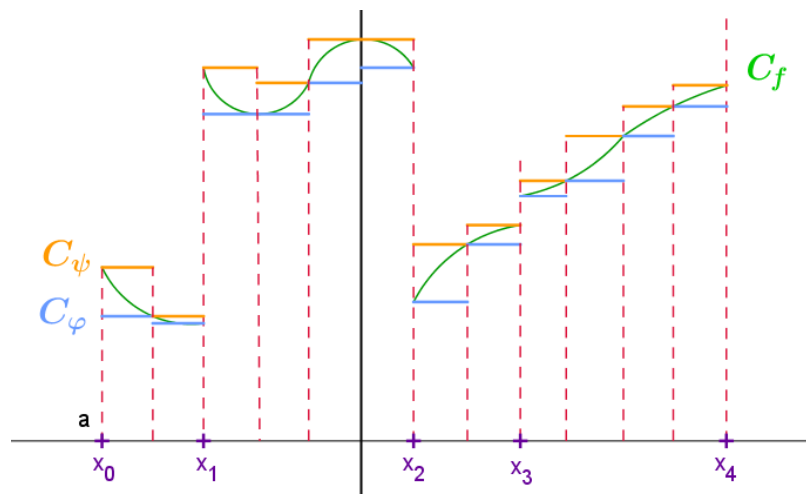
L'énoncé ci-dessous joue un rôle fondamental dans la construction de l'intégrale d'une fonction continue (par morceaux ou non) sur un segment, qui constitue l'objet du paragraphe suivant.

THÉORÈME 23.2 - (Approximation uniforme par des fonctions en escalier).

Toute fonction continue par morceaux sur $[a, b]$ "peut être approchée arbitrairement près et de manière uniforme par une fonction en escalier" dans le sens plus précis suivant :

$$\forall f \in \mathcal{C}_{\text{pm}}^0([a, b], \mathbb{R}), \forall \varepsilon > 0, \exists (\varphi, \psi) \in \mathcal{E}([a, b], \mathbb{R})^2,$$

$$\forall x \in [a, b], [\varphi(x) \leq f(x) \leq \psi(x)] \wedge [0 \leq \psi(x) - \varphi(x) < \varepsilon]$$



4. Il faut quand même faire un peu de "chirurgie"...

PREUVE. ► Commençons par établir le théorème dans le cas particulier où $f \in \mathcal{C}^0([a, b], \mathbb{R})$.

Soit $\varepsilon > 0$.

Puisque f est continue sur $[a, b]$, alors f est uniformément continue sur $[a, b]$ (d'après le théorème de Heine). Il existe donc un réel $\alpha > 0$ tel que : $\forall (x, y) \in [a, b]^2, [|y - x| < \alpha] \implies [|f(y) - f(x)| < \varepsilon]$.

Choisissons alors un entier $n \in \mathbb{N}^*$ tel que : $\frac{b-a}{n} < \alpha$,⁵ et notons σ la subdivision régulière de $[a, b]$ de pas $1/n$. On a : $\sigma = (x_0, \dots, x_n)$ avec $x_i = a + i(b-a)/n$.

En exploitant de nouveau la continuité de f sur $[x_i, x_{i+1}]$ (f étant continue sur $[a, b]$), le théorème des bornes atteintes permet d'affirmer que f est bornée sur chacun de ces intervalles. Explicitement :

$$\forall i \in \llbracket 0, n-1 \rrbracket, \exists (m_i, M_i) \in \mathbb{R}^2, m_i = \min_{[x_i, x_{i+1}]} f, M_i = \max_{[x_i, x_{i+1}]} f$$

En outre : $0 \leq M_i - m_i < \varepsilon$ puisque pour tout couple de réels (x, y) dans $[x_i, x_{i+1}]$, on a : $|f(y) - f(x)| < \varepsilon$ (par construction des x_i).

On construit alors deux fonctions en escalier φ et ψ en posant :

$$\forall x \in [a, b], \varphi(x) = \begin{cases} m_i & \text{si } x \in [x_i, x_{i+1}[\\ f(b) & \text{si } x = b \end{cases} \quad \text{et} \quad \forall x \in [a, b], \psi(x) = \begin{cases} M_i & \text{si } x \in [x_i, x_{i+1}[\\ f(b) & \text{si } x = b \end{cases}$$

Il reste à vérifier que les fonctions φ et ψ satisfont les assertions du théorème, ce qui est une formalité. Soit x un réel de $[a, b]$.

Deux cas se présentent : soit il existe i tel que $x \in [x_i, x_{i+1}[$ (ce qui équivaut à dire que $x \in [a, b[$), soit $x = b$. Dans le premier cas, on a : $0 \leq \psi(x) - \varphi(x) = M_i - m_i$ (par construction de φ et ψ) et $M_i - m_i < \varepsilon$ (par construction de la subdivision σ).

D'où : $\forall x \in [a, b[, [\varphi(x) \leq f(x) \leq \psi(x)] \wedge [0 \leq \psi(x) - \varphi(x) < \varepsilon]$.

Dans le second cas : on a $[\varphi(b) \leq f(b) \leq \psi(b)] \wedge [0 \leq \psi(b) - \varphi(b) < \varepsilon]$ puisque $\varphi(b) = f(b) = \psi(b) \dots$

En résumé, on a établi que :

$$\boxed{\forall f \in \mathcal{C}^0([a, b], \mathbb{R}), \forall \varepsilon > 0, \exists (\varphi, \psi) \in \mathcal{E}([a, b], \mathbb{R})^2, \forall x \in [a, b], \\ [\varphi(x) \leq f(x) \leq \psi(x)] \wedge [0 \leq \psi(x) - \varphi(x) < \varepsilon]}$$

► “Il n'y a plus qu'à” étendre le résultat aux fonctions continues par morceaux.

Considérons donc à présent $f \in \mathcal{C}_{\text{pm}}^0([a, b], \mathbb{R})$, et fixons $\varepsilon > 0$.

La fonction f étant continue par morceaux sur $[a, b]$, il existe une subdivision $\sigma = (x_0, \dots, x_n)$ de l'intervalle $[a, b]$ telle que sur chacun des intervalles $]x_i, x_{i+1}[$, la fonction f est continue.

Soit $i \in \llbracket 0, n-1 \rrbracket$. Sur $]x_i, x_{i+1}[$, la fonction f est continue et en outre : $\lim_{x \rightarrow x_i^+} f(x) = \ell_i^+$ et $\lim_{x \rightarrow x_{i+1}^-} f(x) = \ell_{i+1}^-$

où ℓ_i^+ et ℓ_{i+1}^- désignent deux réels. On prolonge alors $f|_{]x_i, x_{i+1}[}$ par continuité (en x_i et en x_{i+1}) en posant :

$$\forall x \in [x_i, x_{i+1}], \widehat{f}_i(x) = \begin{cases} f(x) & \text{si } x \in]x_i, x_{i+1}[\\ \ell_i^+ & \text{si } x = x_i \\ \ell_{i+1}^- & \text{si } x = x_{i+1} \end{cases}$$

5. Un tel entier n existe puisque $\mathbb{R}$ est archimédien.

La fonction $\widehat{f}_i$ ainsi construite est continue sur $[x_i, x_{i+1}]$, et on peut donc lui appliquer le résultat précédemment établi. Il existe deux fonctions φ_i et $\psi_i \in \mathcal{C}_{\text{pm}}^0([a, b], \mathbb{R})$ telles que :

$$\forall x \in [x_i, x_{i+1}], \left[\varphi_i(x) \leq \widehat{f}_i(x) \leq \psi_i(x) \right] \wedge [0 \leq \psi_i(x) - \varphi_i(x) < \varepsilon]$$

En particulier : $\forall x \in]x_i, x_{i+1}[$, $\left[\varphi_i(x) \leq \widehat{f}_i(x) \leq \psi_i(x) \right] \wedge [0 \leq \psi_i(x) - \varphi_i(x) < \varepsilon]$

D'où par définition de la fonction $\widehat{f}_i$ on a : $\forall x \in]x_i, x_{i+1}[$, $\varphi_i(x) \leq f(x) \leq \psi_i(x)$

Puis on recolle les fonctions φ_i et ψ_i en définissant deux fonctions Φ et Ψ telles que :

$$\forall x \in [a, b], \Phi(x) = \begin{cases} \varphi_i(x) & \text{si } \exists i \in \llbracket 0, n-1 \rrbracket, x \in]x_i, x_{i+1}[\\ f(x_i) & \text{si } \exists i \in \llbracket 0, n \rrbracket, x = x_i \end{cases}$$

$$\text{et } \forall x \in [a, b], \Psi(x) = \begin{cases} \psi_i(x) & \text{si } \exists i \in \llbracket 0, n-1 \rrbracket, x \in]x_i, x_{i+1}[\\ f(x_i) & \text{si } \exists i \in \llbracket 0, n \rrbracket, x = x_i \end{cases}$$

Les fonctions Φ et Ψ ainsi construites sont en escalier sur $[a, b]$; de plus on a $0 \leq \Psi(x) - \Phi(x) < \varepsilon$ sur chacun des intervalles $]x_i, x_{i+1}[$ par définition de ces fonctions, et $0 \leq \Psi(x_i) - \Phi(x_i) < \varepsilon$ puisqu'on a imposé à Φ et Ψ de coïncider avec f en chacun des réels x_i .

En résumé : $\exists (\varphi, \psi) \in \mathcal{E}([a, b], \mathbb{R})^2$, $\forall x \in [a, b]$, $[\varphi(x) \leq f(x) \leq \psi(x)] \wedge [0 \leq \psi(x) - \varphi(x) < \varepsilon]$

Comme le réel $\varepsilon > 0$ et la fonction $f \in \mathcal{C}_{\text{pm}}^0([a, b], \mathbb{R})$ sont arbitraires, on peut conclure :

$$\forall f \in \mathcal{C}_{\text{pm}}^0([a, b], \mathbb{R}), \forall \varepsilon > 0, \exists (\varphi, \psi) \in \mathcal{E}([a, b], \mathbb{R})^2, \forall x \in [a, b],$$

$$[\varphi(x) \leq f(x) \leq \psi(x)] \wedge [0 \leq \psi(x) - \varphi(x) < \varepsilon]$$



23.5 Intégrale d'une fonction continue par morceaux

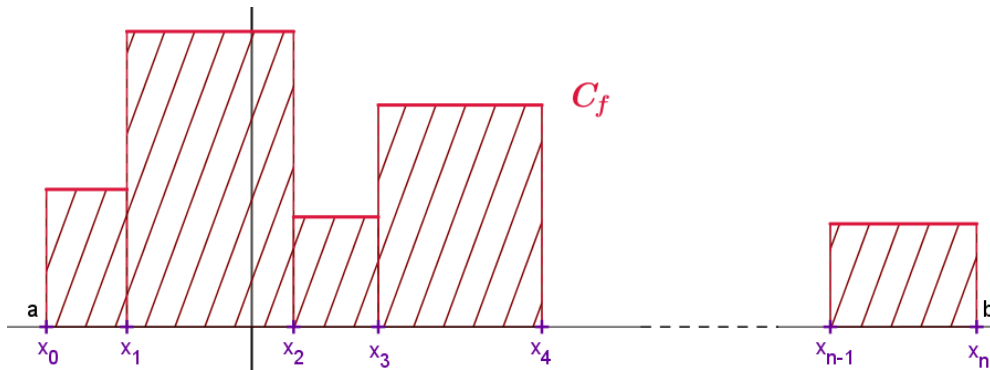
23.5.1 Intégrale d'une fonction en escalier

DÉFINITION 23.7 - Soit $f \in \mathcal{E}([a, b], \mathbb{R})$, et soit $\sigma = (x_0, \dots, x_n)$ une subdivision de $[a, b]$ adaptée à f .

On appelle **intégrale de f sur $[a, b]$ relativement à σ** le réel :

$$I_{[a, b], \sigma}(f) = \sum_{i=0}^{n-1} y_i (x_{i+1} - x_i)$$

où les y_i désignent les valeurs de la fonction f sur chacun des intervalles $]x_i, x_{i+1}[$.



Remarque. Soit f une fonction continue sur $[a, b]$. On se place dans le cas particulier où σ est une subdivision régulière de $[a, b]$, de pas $h = \frac{b-a}{n}$. On note φ la fonction en escalier égale à $f(x_i)$ sur chacun des intervalles $[x_i, x_{i+1}]$. On retrouve alors la formule $I_{[a,b],\sigma}(\varphi) = \frac{b-a}{n} \sum_{i=0}^{n-1} f\left(a + i \frac{b-a}{n}\right)$ intervenant dans la méthode des rectangles (à gauche) pour le calcul d'une valeur approchée de $\int_a^b f$.

PROPRIÉTÉ 23.7 - Mêmes notations que dans la définition précédente. La valeur du réel $I_{[a,b],\sigma}(f)$ est indépendante de la subdivision σ adaptée à f .

En d'autres termes, si σ et σ' sont deux subdivisions de $[a, b]$ adaptées à f , alors : $I_{[a,b],\sigma}(f) = I_{[a,b],\sigma'}(f)$.

PREUVE. En exercice. Idée : il suffit d'utiliser les définitions de $I_{[a,b],\sigma}(f)$, de subdivision plus fine qu'une autre, et de tenir compte du fait que les valeurs de f aux points de la subdivision σ ne jouent aucun rôle dans le calcul de $I_{[a,b],\sigma}(f)$. Pour ce faire, on montre que le fait de rajouter un point à une subdivision ne modifie pas la valeur de l'intégrale ; puis on prouve que $I_{[a,b],\sigma}(f) = I_{[a,b],\sigma'}(f)$ en utilisant le fait que ces deux intégrales sont égales à une même troisième : $I_{[a,b],\sigma \vee \sigma'}(f)$. ☕

DÉFINITION 23.8 - Mêmes notations que dans la définition et la propriété précédentes. La valeur commune à toutes les $I_{[a,b],\sigma}(f)$ est appelée **intégrale de la fonction en escalier f sur l'intervalle $[a, b]$** . On la note : $I_{[a,b]}(f)$.

23.5.2 Propriétés de l'intégrale des fonctions en escalier

PROPRIÉTÉ 23.8 - (Positivité) $\forall f \in \mathcal{E}([a, b], \mathbb{R}), [f \geq 0] \implies [I_{[a,b]}(f) \geq 0]$

PREUVE. Conséquence immédiate de la définition de $I_{[a,b],\sigma}(f)$. ☕

PROPRIÉTÉ 23.9 - (Croissance) $\forall (f, g) \in \mathcal{E}([a, b], \mathbb{R})^2, [g \geq f] \implies [I_{[a,b]}(g) \geq I_{[a,b]}(f)]$

PREUVE. Conséquence immédiate de la propriété précédente. ☕

PROPRIÉTÉ 23.10 - (Linéarité). L'application $f \in \mathcal{E}([a, b], \mathbb{R}), \mapsto I_{[a, b]}(f)$ est une forme linéaire sur $\mathcal{E}([a, b], \mathbb{R})$.

PREUVE. Conséquence immédiate (quoique pénible à écrire) de la définition de $I_{[a, b], \sigma}(f)$. ☕

PROPRIÉTÉ 23.11 - (Relation de Chasles).

$$\forall c \in]a, b[, \forall f \in \mathcal{E}([a, b], \mathbb{R}), I_{[a, b]}(f) = I_{[a, c]}(f) + I_{[c, b]}(f)$$

PREUVE. Nouvelle conséquence immédiate de la définition de $I_{[a, b], \sigma}(f)$. ☕

23.5.3 Intégrale d'une fonction continue par morceaux

En avant pour quelques notations. Pour toute fonction $f \in \mathcal{C}_{\text{pm}}^0([a, b], \mathbb{R})$, on introduit les ensembles :

$$\mathcal{E}^+(f) = \{\varphi \in \mathcal{E}([a, b], \mathbb{R}), \varphi \geq f\} \text{ et } \mathcal{E}_-(f) = \{\varphi \in \mathcal{E}([a, b], \mathbb{R}), \varphi \leq f\}$$

des fonctions en escalier sur $[a, b]$ respectivement minorées et majorées par f . On introduit également les ensembles des valeurs des intégrales correspondantes :

$$\mathcal{J}^+(f) = \{I_{[a, b]}(\varphi), \varphi \in \mathcal{E}^+(f)\} \text{ et } \mathcal{J}_-(f) = \{I_{[a, b]}(\varphi), \varphi \in \mathcal{E}_-(f)\}$$

THÉORÈME 23.3 - Soit $f \in \mathcal{C}_{\text{pm}}^0([a, b], \mathbb{R})$. L'ensemble $\mathcal{J}^+(f)$ admet une borne inférieure, l'ensemble $\mathcal{J}_-(f)$ admet une borne supérieure, et :

$$\inf \mathcal{J}^+(f) = \sup \mathcal{J}_-(f)$$

PREUVE. Soit f une fonction de $\mathcal{C}_{\text{pm}}^0([a, b], \mathbb{R})$. Alors f est en particulier bornée sur $[a, b]$ (puisque toute fonction continue par morceaux sur un segment l'est). Il existe donc deux réels m et M tels que : $m \leq f \leq M$.

La fonction constante égale à m (*resp.* M) sur $[a, b]$ est en escalier sur $[a, b]$ et appartient à $\mathcal{E}_-(f)$ (*resp.* $\mathcal{E}^+(f)$).

Donc le réel $m(b-a)$ (intégrale de la première fonction évoquée ci-dessus) appartient à $\mathcal{J}_-(f)$; et $M(b-a)$ appartient à $\mathcal{J}^+(f)$. Il s'ensuit que les ensembles $\mathcal{J}_-(f)$ et $\mathcal{J}^+(f)$ sont non vides.

En outre, pour toute fonction $\varphi \in \mathcal{E}_-(f)$, on a : $I_{[a, b]}(\varphi) \leq M(b-a)$, puisque dès lors que φ appartient à $\mathcal{E}_-(f)$, on a en particulier $\varphi \leq f$ d'où $\varphi \leq M$. L'ensemble $\mathcal{J}_-(f)$ est donc majoré (par $M(b-a)$). De manière symétrique, l'ensemble $\mathcal{J}^+(f)$ est minoré (par $m(b-a)$).

Comme l'ensemble $\mathcal{J}_-(f)$ (*resp.* $\mathcal{J}^+(f)$) est non vide et majoré (*resp.* minoré), il admet une borne supérieure (*resp.* une borne inférieure) que nous noterons A (*resp.* B).

On a donc : $A = \sup \{I_{[a, b]}(\varphi), \varphi \in \mathcal{E}_-(f)\}$ et $B = \inf \{I_{[a, b]}(\varphi), \varphi \in \mathcal{E}^+(f)\}$.

Prouvons que $A = B$. Soit $\varphi \in \mathcal{E}_-(f)$ et $\psi \in \mathcal{E}^+(f)$; on a $\varphi \leq f \leq \psi$. Donc : $I_{[a,b]}(\varphi) \leq I_{[a,b]}(\psi)$ (croissance de l'intégrale pour les fonctions en escalier). Cette inégalité étant valable pour toute fonction $\psi \in \mathcal{E}^+(f)$, on en déduit que $I_{[a,b]}(\varphi)$ est un minorant de $\mathcal{J}^+(f)$. D'où : $I_{[a,b]}(\varphi) \leq B$ (puisque B est le plus grand des minorants de $\mathcal{J}^+(f)$).

Il s'ensuit que B est un majorant de $\mathcal{J}_-(f)$, donc : $A \leq B$ ($\spadesuit$) (puisque A est le plus petit des majorants de $\mathcal{J}_-(f)$).

Par ailleurs, d'après le théorème 23.2, il existe pour tout réel $\varepsilon > 0$ deux fonctions $\varphi \in \mathcal{E}_-(f)$ et $\psi \in \mathcal{E}^+(f)$ telles que : $0 \leq \psi - \varphi < \varepsilon$. Ainsi : $\varphi \leq f \leq \psi$ d'où $I_{[a,b]}(\varphi) \leq A$ et $I_{[a,b]}(\psi) \geq B$.

De ces observations et de l'encadrement $0 \leq \psi - \varphi < \varepsilon$, on déduit alors que : $B \leq A + \varepsilon(b - a)$. Cette dernière inégalité étant valable pour un $\varepsilon > 0$ arbitraire, on a finalement : $B \leq A$ ($\clubsuit$).

Au bout du compte, ($\spadesuit$) et ($\clubsuit$) impliquent : $A = B$, c'est-à-dire :

$$\sup \{I_{[a,b]}(\varphi), \varphi \in \mathcal{E}_-(f)\} = \inf \{I_{[a,b]}(\varphi), \varphi \in \mathcal{E}^+(f)\}$$



DÉFINITION 23.9 - Soit $f \in \mathcal{C}_{\text{pm}}^0([a, b], \mathbb{R})$. On appelle **intégrale de f sur $[a, b]$** le réel

$$\int_a^b f = \inf \mathcal{J}^+(f) \quad (= \sup \mathcal{J}_-(f))$$

Remarque. Pour les fonctions en escalier, cette définition coïncide (ouf!) avec celle d'intégrale donnée dans le paragraphe précédent. Explicitement : $\forall f \in \mathcal{E}([a, b], \mathbb{R}), I_{[a,b]}(f) = \int_a^b f$.

23.5.4 Propriétés de l'intégrale des fonctions continues par morceaux

Les preuves des énoncés ci-dessous peuvent s'interpréter comme des passages à la limite des propriétés analogues pour les intégrales des fonctions en escalier.

PROPRIÉTÉ 23.12 - (Positivité).

$$\forall f \in \mathcal{C}_{\text{pm}}^0([a, b], \mathbb{R}), [f \geq 0] \implies \left[\int_a^b f \geq 0 \right]$$

PREUVE. Soit $f \in \mathcal{C}_{\text{pm}}^0([a, b], \mathbb{R})$. Supposons que f soit positive sur $[a, b]$.

Alors la fonction constante égale à 0 appartient à $\mathcal{E}_-(f) = \{\varphi \in \mathcal{E}([a, b], \mathbb{R}), \varphi \leq f\}$. En particulier : $0 \in \mathcal{J}_-(f)$. Il s'ensuit que $\sup \mathcal{J}_-(f) \geq 0$, ce qui signifie que : $\int_a^b f \geq 0$.

PROPRIÉTÉ 23.13 - (Linéarité).

L'application $f \in \mathcal{C}_{\text{pm}}^0([a, b], \mathbb{R}), \mapsto \int_a^b f$ est une forme linéaire sur $\mathcal{C}_{\text{pm}}^0([a, b], \mathbb{R})$.

PREUVE. Aisé mais pénible.

PROPRIÉTÉ 23.14 - (Croissance).

$$\forall (f, g) \in \mathcal{C}_{\text{pm}}^0([a, b], \mathbb{R})^2, \quad [g \geq f] \implies \left[\int_a^b g \geq \int_a^b f \right]$$

PREUVE. Il suffit d'appliquer la propriété de positivité à la fonction $g - f$, puis d'utiliser celle de linéarité pour conclure. ☕

PROPRIÉTÉ 23.15 - (Relation de Chasles).

$$\forall c \in]a, b[, \quad \forall f \in \mathcal{C}_{\text{pm}}^0([a, b], \mathbb{R}), \quad \int_a^b f = \int_a^c f + \int_c^b f$$

PREUVE. Immédiate. ☕

Les deux énoncés ci-dessous sont des conséquences de la positivité de l'intégrale.

PROPRIÉTÉ 23.16 - (Majoration)

$$\forall f \in \mathcal{C}_{\text{pm}}^0([a, b], \mathbb{R}), \quad \left| \int_a^b f \right| \leq \int_a^b |f|$$

PREUVE. Immédiate. ☕

THÉORÈME 23.4 - (Inégalité de Cauchy-Schwarz-Буняковский)

$$\forall (f, g) \in \mathcal{C}_{\text{pm}}^0([a, b], \mathbb{R})^2,$$

$$\left| \int_a^b fg \right| \leq \left(\int_a^b f^2 \right)^{1/2} \left(\int_a^b g^2 \right)^{1/2}$$

PREUVE. Pour tout réel t , posons : $P(t) = \int_a^b (f + tg)^2$; cette intégrale est bien définie, puisque sous les hypothèses de l'énoncé, la fonction $(f + tg)^2$ est continue sur $[a, b]$. Clairement : $\forall t \in \mathbb{R}, P(t) \geq 0$.

D'autre part, pour tout réel t : $P(t) = \left(\int_a^b g^2 \right) t^2 + 2 \left(\int_a^b fg \right) t + \int_a^b f^2$. La fonction P est donc polynomiale de degré au plus 2. On distingue alors deux cas.

► Premier cas : si $\int_a^b g^2 = 0$. Puisque la fonction g^2 est continue sur $[a, b]$ (par hypothèse : $g \in \mathcal{C}^0([a, b], \mathbb{R})$) et positive sur $[a, b]$, on en déduit que g^2 est identiquement nulle sur $[a, b]$ (propriété faisant l'objet de la question de cours 2) ; donc g est identiquement nulle sur $[a, b]$.

En particulier l'inégalité $\left| \int_a^b fg \right| \leq \left(\int_a^b f^2 \right)^{1/2} \left(\int_a^b g^2 \right)^{1/2}$ est vérifiée car les deux termes de celle-ci sont nuls.

► Second cas : si $\int_a^b g^2 \neq 0$. Alors la fonction P est polynomiale de degré 2, et puisque son signe est constant, son discriminant Δ est négatif ou nul. Or :

$$[\Delta \leq 0] \iff \left[4 \left(\int_a^b fg \right)^2 - 4 \left(\int_a^b f^2 \right) \left(\int_a^b g^2 \right) \leq 0 \right] \iff \left[\left(\int_a^b fg \right)^2 \leq \left(\int_a^b f^2 \right) \left(\int_a^b g^2 \right) \right]$$

$$\iff \left[\left| \int_a^b fg \right| \leq \left(\int_a^b f^2 \right)^{1/2} \left(\int_a^b g^2 \right)^{1/2} \right]$$

► **Conclusion.** $\forall (f, g) \in \mathcal{C}^0([a, b], \mathbb{R})^2, \quad \left| \int_a^b fg \right| \leq \left(\int_a^b f^2 \right)^{1/2} \left(\int_a^b g^2 \right)^{1/2}$ ☕

23.6 Propriétés de l'intégrale d'une fonction continue

PROPRIÉTÉ 23.17 - Si f est une fonction continue et positive sur $[a, b]$, alors :

$$\left[\int_a^b f = 0 \right] \iff [f = 0]$$

PREUVE. Raisonnement par double implication. Soit f une fonction continue et positive sur $[a, b]$.

► Sens direct. Supposons que $\int_a^b f = 0$ (♠).

Raisonnons par l'absurde et supposons que $f \neq 0$ (f n'est pas identiquement nulle sur $[a, b]$).

Alors : $\exists c \in [a, b], f(c) > 0$. On distingue deux cas (suivant que c est ou n'est pas une borne de l'intervalle $[a, b]$).

► Premier cas : si $c \in]a, b[$. Puisque la fonction f est continue en c , il existe un réel strictement positif $\alpha > 0$ tel que : $\forall x \in]c - \alpha, c + \alpha[, f(x) \geq \frac{f(c)}{2}$. En intégrant terme à terme cette inégalité on obtient :

$$\int_{c-\alpha}^{c+\alpha} f \geq \alpha f(c).$$

D'où : $\int_a^b f = \underbrace{\int_a^{c-\alpha} f}_{\geq 0} + \int_{c-\alpha}^{c+\alpha} f + \underbrace{\int_{c+\alpha}^b f}_{\geq 0} \geq \alpha f(c)$. Comme $\alpha f(c) > 0$, on en déduit $\int_a^b f > 0$; contradiction

(avec (♠)).

► Second cas : si $c = a$ (ou b). Supposons que $c = a$. Puisque la fonction f est continue en a , il existe un réel strictement positif $\alpha > 0$ tel que : $\forall x \in [a, a + \alpha[, f(x) \geq \frac{f(a)}{2}$. En intégrant terme à terme on

obtient : $\int_a^{a+\alpha} f \geq \alpha f(a)/2$.

D'où : $\int_a^b f = \int_a^{a+\alpha} f + \underbrace{\int_{a+\alpha}^b f}_{\geq 0} \geq \alpha f(a)/2$. Comme $\alpha f(a)/2 > 0$, on en déduit que $\int_a^b f > 0$; contradiction

(avec (♠)).

Le cas $c = b$ se traite de manière analogue.

Conclusion. $\left[\int_a^b f = 0 \right] \implies [f = 0]$

► Réciproque. Giga-triviale.

Conclusion. $\left[\int_a^b f = 0 \right] \iff [f = 0]$ ☕

DÉFINITION 23.10 - Pour toute $f \in \mathcal{C}_{\text{pm}}^0([a, b], \mathbb{R})$, on appelle **valeur moyenne** de f sur $[a, b]$ le réel : $\frac{1}{b-a} \int_a^b f$.

On obtient comme conséquence du théorème des bornes atteintes et du théorème des valeurs intermédiaires l'énoncé suivant.

PROPRIÉTÉ 23.18 - (**Formule de la moyenne**) Si f est une fonction continue sur $[a, b]$, alors :

$$\exists c \in [a, b], f(c) = \frac{1}{b-a} \int_a^b f.$$

PREUVE. Voir preuve de la propriété 14.2 page 374. ☕

La fin de ce paragraphe est consacrée aux liens avec le cours d'informatique portant sur les méthodes numériques de calcul intégral.

DÉFINITION 23.11 - Pour toute fonction $f \in \mathcal{C}^0([a, b], \mathbb{R})$, et pour tout entier n non nul, on appelle (n -ième) **somme de Riemann** de f sur $[a, b]$ le réel : $\frac{b-a}{n} \sum_{k=0}^{n-1} f\left(a + k \frac{b-a}{n}\right)$.

Il résulte de la construction de l'intégrale que la suite des sommes de Riemann d'une fonction continue converge vers $\int_a^b f$; en d'autres termes, la méthode des rectangles (à gauche) converge. En outre, dans le cas des fonctions de classe $\mathcal{C}^1$, on dispose d'un énoncé plus fin donnant une majoration de la distance entre la n -ième somme de Riemann et l'intégrale de f sur $[a, b]$, que voici :

THÉORÈME 23.5 - Pour toute fonction f de classe $\mathcal{C}^1$ sur $[a, b]$ on a :

$$\left| \int_a^b f(x) dx - S_n(f) \right| \leq \frac{(b-a)^2}{2n} M_1 \quad \text{avec} \quad M_1 = \max_{x \in [a, b]} |f'(x)|$$

PREUVE. Soit f une fonction de classe $\mathcal{C}^1$ sur $[a, b]$, et soit n un entier naturel non nul. On pose :

$$\forall k \in \llbracket 0, n \rrbracket, x_k = a + k \left(\frac{b-a}{n} \right)$$

de telle sorte que $\sigma = (x_0, \dots, x_n)$ est la subdivision régulière de $[a, b]$ de pas $h = \frac{b-a}{n}$.

► Observons que f' étant continue sur $[a, b]$ par hypothèse, le théorème des bornes atteintes assure qu'elle est bornée et atteint ses bornes sur $[a, b]$. Donc : $\max_{x \in [a, b]} |f'(x)|$ existe. Par la suite, on notera M_1 ce maximum.

► Soit k un entier dans $\llbracket 0, n-1 \rrbracket$, et soit $x \in [x_k, x_{k+1}]$. Sur le segment $[x_k, x]$, la fonction f vérifie les hypothèses du théorème des accroissements finis (f est continue sur $[x_k, x]$ et dérivable sur $]x_k, x[$), que l'on ne se prive donc pas d'appliquer :

$$\exists c_k \in]x_k, x[, f'(c_k) = \frac{f(x) - f(x_k)}{x - x_k}. \text{ D'où : } \exists c_k \in]x_k, x[, f(x) - f(x_k) = f'(c_k)(x - x_k)$$

Comme en outre : $|f'(c_k)| \leq M_1$, on déduit de ce qui précède : $-M_1(x - x_k) \leq f(x) - f(x_k) \leq M_1(x - x_k)$

Cette dernière inégalité étant valide pour un réel x quelconque dans le segment $[x_k, x_{k+1}]$, on peut l'intégrer terme à terme pour obtenir :

$$-M_1 \int_{x_k}^{x_{k+1}} (x - x_k) dx \leq \int_{x_k}^{x_{k+1}} [f(x) - f(x_k)] dx \leq M_1 \int_{x_k}^{x_{k+1}} (x - x_k) dx$$

$$\text{D'où : } -M_1 \frac{(x_{k+1} - x_k)^2}{2} \leq \left(\int_{x_k}^{x_{k+1}} f(x) dx \right) - (x_{k+1} - x_k) f(x_k) \leq M_1 \frac{(x_{k+1} - x_k)^2}{2}$$

$$\text{Donc : } -\frac{M_1 h^2}{2} \leq \left(\int_{x_k}^{x_{k+1}} f(x) dx \right) - h f(x_k) \leq \frac{M_1 h^2}{2}$$

Pour conclure, on somme l'encadrement ci-dessus en faisant varier k de 0 à $n-1$:

$$-\sum_{k=0}^{n-1} \frac{M_1 h^2}{2} \leq \sum_{k=0}^{n-1} \int_{x_k}^{x_{k+1}} f(x) dx - h \sum_{k=0}^{n-1} f(x_k) \leq \sum_{k=0}^{n-1} \frac{M_1 h^2}{2}$$

$$\text{Soit finalement : } -\frac{n M_1 h^2}{2} \leq \int_a^b f(x) dx - h \sum_{k=0}^{n-1} f(x_k) \leq \frac{n M_1 h^2}{2}$$

$$\text{càd : } \left| \int_a^b f(x) dx - \frac{b-a}{n} \sum_{k=0}^{n-1} f\left(a + k \frac{b-a}{n}\right) \right| \leq \frac{M_1 (b-a)^2}{2n}$$



23.7 Brève extension aux fonctions à valeurs complexes

DÉFINITION 23.12 - Pour toute fonction $f \in \mathcal{C}_{\text{pm}}^0([a, b], \mathbb{C})$, on appelle **intégrale de f sur $[a, b]$** le nombre complexe : $\int_a^b f = \left(\int_a^b \text{Re}(f) \right) + i \left(\int_a^b \text{Im}(f) \right)$.

Les preuves des trois propriétés ci-dessous sont des conséquences des propriétés analogues pour les fonctions à valeurs réelles et de l'inégalité triangulaire.

PROPRIÉTÉ 23.19 - (**Linéarité**). L'application $f \in \mathcal{C}_{\text{pm}}^0([a, b], \mathbb{C}), \mapsto \int_a^b f$ est une forme linéaire sur $\mathcal{C}_{\text{pm}}^0([a, b], \mathbb{C})$.

PROPRIÉTÉ 23.20 - (**Majoration**).

$$\forall f \in \mathcal{C}_{\text{pm}}^0([a, b], \mathbb{C}), \left| \int_a^b f \right| \leq \int_a^b |f|$$

PROPRIÉTÉ 23.21 - (Relation de Chasles).

$$\forall c \in]a, b[, \forall f \in \mathcal{C}_{\text{pm}}^0([a, b], \mathbb{C}), \quad \int_a^b f = \int_a^c f + \int_c^b f$$

23.8 Primitives et intégrales

THÉORÈME 23.6 - (Théorème fondamental de l'intégration). Soit f une fonction continue sur un intervalle I , et $a \in I$.

La fonction $F : x \in I \mapsto \int_a^x f(t) dt$ est la primitive de f sur I s'annulant en a .

PREUVE. Soit h un réel non nul et raisonnable. On a : $\frac{F(a+h) - F(a)}{h} = \frac{1}{h} \int_a^{a+h} f(t) dt$.

La fonction f étant continue en a , on a : $f(t) = f(a) + o_a(1)$. Il s'ensuit que :

$$\frac{F(a+h) - F(a)}{h} = f(a) + \frac{1}{h} \int_a^{a+h} o_a(1) dt$$

Il reste à voir que $\lim_{h \rightarrow 0} \frac{1}{h} \int_a^{a+h} o_a(1) dt = 0$ pour conclure que :

$$\lim_{h \rightarrow 0} \frac{F(a+h) - F(a)}{h} = f(a)$$

Ainsi F est dérivable en a , et $F' = f$.

L'unicité de F provient de ce que deux primitives d'une même fonction diffèrent d'une constante sur un intervalle, et que cette constante est fixée par la condition $F(a) = 0$. ☕

COROLLAIRE 23.1 - Soit f une fonction continue sur un intervalle I , et soit G une primitive de f sur I . Alors :

$$\forall (a, b) \in I^2, \quad \int_a^b f(t) dt = [G(t)]_a^b$$

PREUVE. Triviale. ☕

23.9 Formule de Taylor avec reste intégrale et inégalité de Taylor-Lagrange

THÉORÈME 23.7 - (Formule de Taylor avec reste intégrale). Soit $f : [a; b] \longrightarrow \mathbb{R}$ une fonction de classe $\mathcal{C}^{n+1}$, alors :

$$f(b) = f(a) + \frac{b-a}{1!} f'(a) + \cdots + \frac{(b-a)^n}{n!} f^{(n)}(a) + \frac{1}{n!} \int_a^b (b-t)^n f^{(n+1)}(t) dt$$

PREUVE. Montrons par récurrence sur $n \in \mathbb{N}$ la propriété $P(n)$: “ $\forall f \in \mathcal{C}^{n+1}$,

$$f(b) = \sum_{k=0}^n \frac{f^{(k)}(a)}{k!} (b-a)^k + \frac{1}{n!} \int_a^b (b-t)^n f^{(n+1)}(t) dt$$

► **Initialisation.** Soit f une fonction de classe $\mathcal{C}^1$ sur $[a, b]$. On a alors : $f(b) - f(a) = \int_a^b f'(t) dt$, ce qui établit la propriété pour $n = 0$.

► **Hérédité.** Supposons $P(n)$ vraie pour un certain entier naturel n . Soit f une fonction de classe $\mathcal{C}^{n+2}$ sur $[a, b]$. Alors f est en particulier de classe $\mathcal{C}^{n+1}$ sur $[a, b]$, et on peut donc lui appliquer l’hypothèse de récurrence. Ce faisant on obtient :

$$f(b) = \sum_{k=0}^n \frac{f^{(k)}(a)}{k!} (b-a)^k + \frac{1}{n!} \int_a^b (b-t)^n f^{(n+1)}(t) dt$$

Les fonctions $t \longmapsto -\frac{(b-t)^{n+1}}{n+1}$ et $t \longmapsto f^{(n+1)}(t)$ étant de classe $\mathcal{C}^1$ sur $[a, b]$, on peut utiliser une intégration par parties pour obtenir :

$$f(b) = \sum_{k=0}^n \frac{f^{(k)}(a)}{k!} (b-a)^k + \frac{1}{n!} \left[-\frac{(b-t)^{n+1}}{n+1} f^{(n+1)}(t) \right]_a^b + \frac{1}{n! \times (n+1)} \int_a^b (b-t)^{n+1} f^{(n+2)}(t) dt$$

$$\text{D'où : } f(b) = \sum_{k=0}^n \frac{f^{(k)}(a)}{k!} (b-a)^k + \frac{(b-a)^{n+1}}{(n+1)!} f^{(n+1)}(a) + \frac{1}{(n+1)!} \int_a^b (b-t)^{n+1} f^{(n+2)}(t) dt$$

Soit finalement : $f(b) = \sum_{k=0}^{n+1} \frac{f^{(k)}(a)}{k!} (b-a)^k + \frac{1}{(n+1)!} \int_a^b (b-t)^{n+1} f^{(n+2)}(t) dt$, ce qui assure que $P(n+1)$ est vraie et établit l’hérédité de la propriété.

Conclusion. Soit $f : [a; b] \longrightarrow \mathbb{R}$ une fonction de classe $\mathcal{C}^{n+1}$, alors :

$$f(b) = \sum_{k=0}^n \frac{f^{(k)}(a)}{k!} (b-a)^k + \frac{1}{n!} \int_a^b (b-t)^n f^{(n+1)}(t) dt \quad \text{☕}$$

La propriété ci-dessous est hors-programme ; cela ne signifie pas qu'elle est trop compliquée pour que vous puissiez la comprendre, mais juste que vous ne pouvez pas l'utiliser sans démonstration. . .

PROPRIÉTÉ 23.22 - (Inégalité de Taylor-Lagrange). Soit $f : [a; b] \longrightarrow \mathbb{R}$ une fonction de classe $\mathcal{C}^{n+1}$, alors :

$$\left| f(b) - \sum_{k=0}^n \frac{(b-a)^k}{k!} f^{(k)}(a) \right| \leq \frac{(b-a)^{n+1}}{(n+1)!} M_{n+1} \quad \text{avec} \quad M_{n+1} = \sup_{x \in [a; b]} |f^{(n+1)}(x)|$$

PREUVE. Soit n un entier naturel, et soit f une fonction de classe $\mathcal{C}^{n+1}$ sur $[a, b]$. D'après la formule de Taylor-Lagrange :

$$f(b) = \left[\sum_{k=0}^n \frac{f^{(k)}(a)}{k!} (b-a)^k \right] + \frac{1}{n!} \int_a^b (b-t)^n f^{(n+1)}(t) dt$$

$$\text{D'où : } \left| f(b) - \left[\sum_{k=0}^n \frac{f^{(k)}(a)}{k!} (b-a)^k \right] \right| = \frac{1}{n!} \left| \int_a^b (b-t)^n f^{(n+1)}(t) dt \right|$$

$$\text{Par suite : } \left| f(b) - \left[\sum_{k=0}^n \frac{f^{(k)}(a)}{k!} (b-a)^k \right] \right| \leq \frac{1}{n!} \int_a^b |(b-t)^n f^{(n+1)}(t)| dt$$

$$\text{Donc : } \left| f(b) - \left[\sum_{k=0}^n \frac{f^{(k)}(a)}{k!} (b-a)^k \right] \right| \leq \frac{1}{n!} \int_a^b (b-t)^n |f^{(n+1)}(t)| dt$$

Or, $f^{(n+1)}$ étant continue sur le segment $[a, b]$ par hypothèse, le théorème des bornes atteintes assure qu'elle est bornée et atteint ses bornes sur $[a, b]$. Donc : $\max_{x \in [a, b]} |f^{(n+1)}(x)|$ existe. Par la suite, on notera M_{n+1} ce maximum.

$$\text{On a donc : } \left| f(b) - \left[\sum_{k=0}^n \frac{f^{(k)}(a)}{k!} (b-a)^k \right] \right| \leq \frac{M_{n+1}}{n!} \int_a^b (b-t)^n dt$$

$$\text{Or : } \int_a^b (b-t)^n dt = \left[-\frac{(b-t)^{n+1}}{n+1} \right]_a^b = \frac{(b-a)^{n+1}}{n+1}.$$

$$\text{D'où finalement : } \left| f(b) - \left[\sum_{k=0}^n \frac{f^{(k)}(a)}{k!} (b-a)^k \right] \right| \leq \frac{M_{n+1} (b-a)^{n+1}}{(n+1)!} \quad \text{☕}$$

Conséquences de l'inégalité de Taylor-Lagrange.

► **Application 1.** La série de terme général $\frac{1}{n!}$ converge et a pour somme e. En raccourci : $\sum_{n \in \mathbb{N}} \frac{1}{n!} = e$.

► **Application 2.** Plus généralement : $\forall x \in \mathbb{R}, \sum_{n \in \mathbb{N}} \frac{x^n}{n!} = e^x$.

► **Application 3.** $\sum_{n \in \mathbb{N}} \frac{(-1)^n}{n+1} = \ln 2$

23.10 Epilogue

23.10.1 Vers les séries numériques

Les exemples d'application de la formule de Taylor-Lagrange donnés en fin de dernier paragraphe préparent la transition vers la notion de série, déjà évoquée à de multiples reprises depuis le début d'année.

Pour être plus explicite, on commence par détailler le résultat de l'application 3. L'assertion $\sum_{n \in \mathbb{N}} \frac{(-1)^n}{n+1} = \ln 2$ signifie que :

$$\lim_{N \rightarrow +\infty} \left[\sum_{n=0}^N \frac{(-1)^n}{n+1} \right] = \ln 2$$

Pour établir ce résultat, on applique la formule de Taylor-Lagrange à la fonction $f : x \mapsto \ln(1+x)$ entre 0 et 1, ce qui est légitime puisque cette fonction est de classe $\mathcal{C}^\infty$ sur $] -1, +\infty[$ selon les théorèmes généraux. On a alors pour tout entier naturel N :

$$\ln(2) = \left[\sum_{k=0}^N \frac{f^{(k)}(0)}{k!} \right] + \frac{1}{N!} \int_0^1 (1-t)^N f^{(N+1)}(t) dt$$

En observant que $f^{(0)}(0) = f(0) = 0$, et que pour tout entier k non nul et tout réel $x > -1$, $f^{(k)}(x) = \frac{(-1)^{k-1} (k-1)!}{(1+x)^k}$, on en déduit que :

$$\ln(2) = \left[\sum_{k=1}^N \frac{(-1)^{k-1}}{k} \right] + \frac{1}{N!} \int_0^1 (1-t)^N \frac{(-1)^N N!}{(1+t)^{N+1}} dt$$

Il s'ensuit que :

$$\left| \ln(2) - \sum_{k=0}^{N-1} \frac{(-1)^k}{k+1} \right| \leq \int_0^1 (1-t)^N dt = \frac{1}{N+1}$$

D'où : $\lim_{N \rightarrow +\infty} \left| \ln(2) - \sum_{k=0}^{N-1} \frac{(-1)^k}{k+1} \right| = 0.$

On en déduit que : $\lim_{N \rightarrow +\infty} \sum_{k=0}^{N-1} \frac{(-1)^k}{k+1} = \ln(2)$, cqfd. Dans ce contexte, on dit que *la série de terme général $(-1)^k/(k+1)$ est convergente, et a pour somme $\ln(2)$.*

23.10.2 Méthodes d'intégration numérique

Notations. Ci-dessous la liste des conventions et notations utilisées tout au long du problème.

► a et b sont deux réels, avec $a < b$; et n désigne un entier naturel non nul.

► On note $(x_0, x_1, \dots, x_n)$ la subdivision régulière de $[a, b]$ de pas $h = \frac{b-a}{n}$, c'est à dire que :

$$\forall k \in \llbracket 0, n \rrbracket, x_k = a + kh$$

► f désigne une fonction continue sur $[a, b]$. On note : $I = \int_a^b f(x) dx$; et pour tout entier k compris entre 0 et $(n-1)$ on note : $I_k = \int_{x_k}^{x_{k+1}} f(x) dx$.

► Lorsque φ désigne une fonction définie sur un segment $[c, d]$, on dira que φ est de classe $\mathcal{C}^m$ sur $[c, d]$ si : 1) φ est de classe $\mathcal{C}^m$ sur $]c, d[$; 2) φ est m fois dérivable à droite en c (*resp.* à gauche en d); 3) $\varphi^{(m)}$ est continue à droite en c (*resp.* à gauche en d).

► Si $P = \sum_{k=0}^p a_k X^k$ désigne un polynôme à coefficients réels, on identifiera celui-ci avec la fonction

polynomiale qui lui est associée, c-à-d la fonction définie sur $\mathbb{R}$ en posant : $\forall t \in \mathbb{R}, P(t) = \sum_{k=0}^p a_k t^k$

Problématique. *Il existe différentes méthodes pour calculer une valeur approchée de l'intégrale d'une fonction continue sur un segment $[a, b]$. Le principe des méthodes intervenant dans ce problème est le même : on construit une subdivision régulière $(x_0, \dots, x_n)$ de l'intervalle $[a, b]$, et sur chacun des sous-intervalles $[x_k, x_{k+1}]$ on approche f par une fonction polynomiale P (de telle sorte que l'intégrale de P soit "simple" à calculer). Cette fonction polynomiale peut être de degré au plus 0 (méthode des rectangles médians, partie I), de degré au plus 1 (méthode des trapèzes, partie II) ou de degré au plus 2 (méthode de Simpson, partie III). L'objectif principal de ce problème est de comparer ces différentes méthodes.*

1. On rappelle ci-dessous un énoncé rencontré cette année.

THÉORÈME DE ROLLE. Soient c et d deux réels tels que $c < d$, et φ une fonction continue sur $[c, d]$ et dérivable sur $]c, d[$, telle que $\varphi(c) = \varphi(d)$. Sous ces hypothèses :

$$\exists \alpha \in]c, d[, \varphi'(\alpha) = 0$$

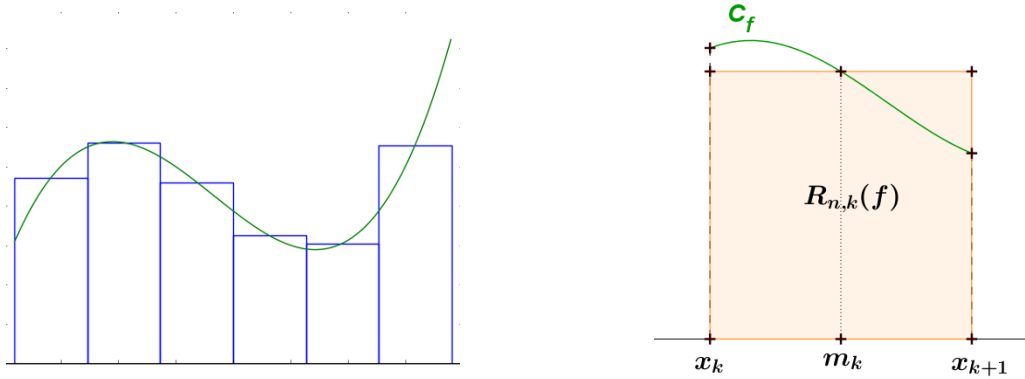
Redémontrer le théorème de Rolle (*indication : on pourra distinguer deux cas suivant que la fonction φ est ou n'est pas constante sur $[a, b]$*).

Partie I — Méthode des rectangles médians

Tout au long de la partie I, on suppose que f est de classe $\mathcal{C}^2$ sur $[a, b]$.

Par ailleurs dans cette partie, on approche f sur chacun des segments $[x_k, x_{k+1}]$ par la fonction constante égale $f(m_k)$, où $m_k = \frac{x_k + x_{k+1}}{2}$ désigne le milieu du segment $[x_k, x_{k+1}]$.

On note $R_{n,k}(f)$ l'aire algébrique du rectangle construit sur le segment $[x_k, x_{k+1}]$, et de "hauteur" $f(m_k)$; et on note $R_n(f) = \sum_{k=0}^{n-1} R_{n,k}(f)$. Les figures ci-dessous illustrent cette construction.



2. Justifier que :
$$I - R_n(f) = \sum_{k=0}^{n-1} \int_{x_k}^{x_{k+1}} [f(x) - f(m_k)] dx$$

3. **Estimation de l'erreur.** Soient c et d deux réels quelconques de $[a, b]$, avec $c < d$.

a. Soit A un nombre réel. On définit une fonction Φ sur $[c, d]$ en posant :

$$\forall x \in [c, d], \Phi(x) = f(d) - f(x) - (d-x)f'(x) - A \frac{(d-x)^2}{2}$$

Etablir qu'il existe un unique réel A , que l'on explicitera⁶, tel que : $\Phi(c) = \Phi(d) = 0$.

b. On suppose à présent A est l'unique réel déterminé à la question précédente. Etablir que :

$$\exists \alpha \in]c, d[, A = f''(\alpha)$$

c. En déduire que :
$$f(d) = f(c) + (d-c)f'(c) + \frac{(d-c)^2}{2} f''(\alpha)$$

d. Soit k un entier quelconque compris entre 0 et $(n-1)$. Etablir que :

$$\forall x \in [x_k, x_{k+1}], |f(x) - f(m_k) - (x - m_k)f'(m_k)| \leq M_2 \frac{(x - (m_k))^2}{2}$$

e. En déduire que :

$$\forall k \in \llbracket 0, n-1 \rrbracket, |I_k - R_{n,k}(f)| \leq M_2 \frac{(b-a)^3}{24n^3}$$

puis que :
$$|I - R_n(f)| \leq M_2 \frac{(b-a)^3}{24n^2}$$

6. On exprimera A en fonction de $c, d, f(c), f(d)$ et $f'(c)$.

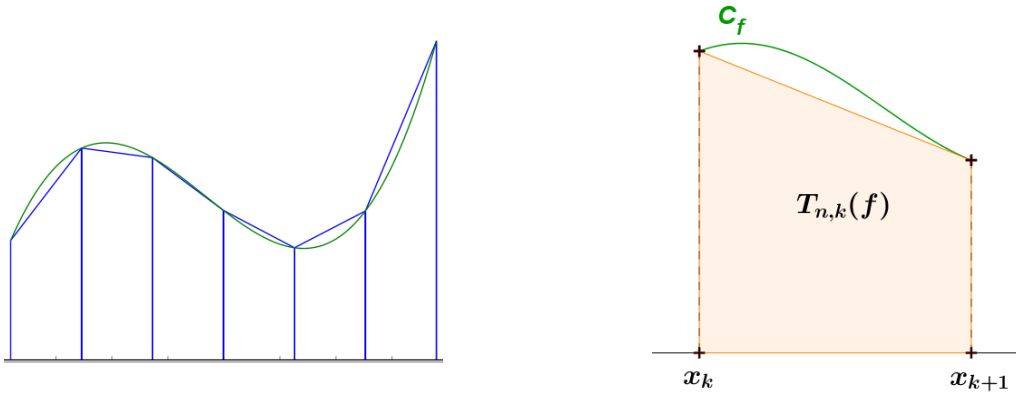
Partie II — Méthode des trapèzes

Tout au long de la partie II, on suppose que f est de classe $\mathcal{C}^2$ sur $[a, b]$.

Par ailleurs dans cette partie, on approche f sur chacun des segments $[x_k, x_{k+1}]$ par la fonction affine coïncidant avec la fonction f en x_k et en x_{k+1} , càd la fonction g_k définie en posant :

$$\forall x \in [x_k, x_{k+1}], \quad g_k(x) = \frac{f(x_{k+1}) - f(x_k)}{h} (x - x_k) + f(x_k)$$

On note $T_{n,k}(f)$ l'aire algébrique du trapèze construit sur le segment $[x_k, x_{k+1}]$, et de "hauteurs" $f(x_k)$ et $f(x_{k+1})$; et on note $T_n(f) = \sum_{k=0}^{n-1} T_{n,k}(f)$. Les figures ci-dessous illustrent cette construction.



Muni de ces notations, on peut écrire : $T_n(f) = \sum_{k=0}^{n-1} \int_{x_k}^{x_{k+1}} g_k(x) dx$.

4. Justifier que :
$$T_n(f) = \frac{b-a}{2n} \sum_{k=0}^{n-1} [f(x_k) + f(x_{k+1})]$$

5. Etablir que :
$$\left| \int_a^b f(x) dx - T_n(f) \right| \leq \sum_{k=0}^{n-1} \left| \int_{x_k}^{x_{k+1}} f(x) - g_k(x) dx \right|$$

6. **Estimation de l'erreur.** Soient c et d deux réels quelconques de $[a, b]$, avec $c < d$.

a. Pour tout réel x dans $[x_k, x_{k+1}]$, on pose : $\delta_k(x) = f(x) - g_k(x)$. A l'aide de deux intégrations par parties, établir que :

$$\int_{x_k}^{x_{k+1}} \delta_k(x) dx = \frac{1}{2} \int_{x_k}^{x_{k+1}} \delta_k''(x) (x - x_k) (x - x_{k+1}) dx$$

b. En déduire que :

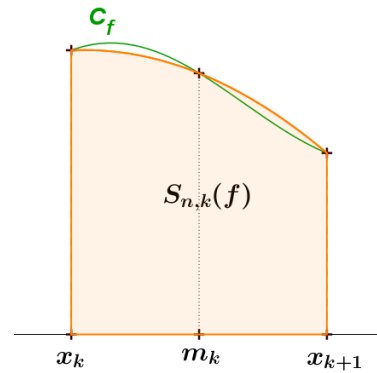
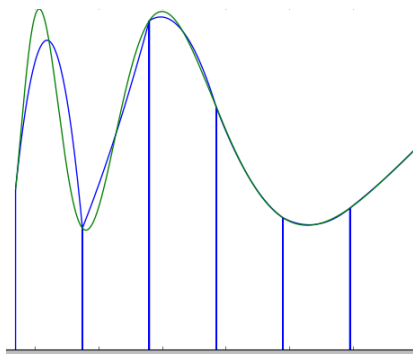
$$\forall k \in \llbracket 0, n-1 \rrbracket, \quad |I_k - T_{n,k}(f)| \leq M_2 \frac{(b-a)^3}{12n^3}; \quad \text{puis que : } |I - T_n(f)| \leq M_2 \frac{(b-a)^3}{12n^2}$$

Partie III — Méthode de Simpson

Tout au long de la partie III, on suppose que f est de classe $\mathcal{C}^4$ sur $[a, b]$.

Par ailleurs dans cette partie, on approche f sur chacun des segments $[x_k, x_{k+1}]$ par la fonction polynomiale P_k de degré au plus 2 coïncidant avec la fonction f en x_k , x_{k+1} et en m_k (avec $m_k = (x_k + x_{k+1})/2$).

On note $S_{n,k}(f)$ l'intégrale de P_k sur le segment $[x_k, x_{k+1}]$; et on note $S_n(f) = \sum_{k=0}^{n-1} S_{n,k}(f)$. Les figures ci-dessous illustrent cette construction.



7. Soient c et d deux réels quelconques de $[a, b]$, avec $c < d$; et $m = (c + d)/2$ le milieu du segment $[c, d]$. On considère le polynôme P de $\mathbb{R}_2[X]$ défini par :

$$P(X) = \frac{f(c)}{(c-m)(c-d)}(X-m)(X-d) + \frac{f(m)}{(m-c)(m-d)}(X-c)(X-d) + \frac{f(d)}{(d-m)(d-c)}(X-c)(X-m)$$

On notera encore P la fonction définie sur $\mathbb{R}$ associée à ce polynôme.

Montrer que P est l'unique polynôme de $\mathbb{R}_2[X]$ tel que : $P(c) = f(c)$, $P(d) = f(d)$ et $P(m) = f(m)$.

Remarque : ceci établit l'existence et l'unicité du polynôme $P_k \in \mathbb{R}_2[X]$ évoqué dans le préambule de cette partie.

8. On conserve tout au long de cette question les notations de la précédente. On souhaite calculer l'intégrale de P sur $[c, d]$.⁷

a. A l'aide d'un changement de variable, établir que :

$$\int_c^d P(x) dx = (d-c) \int_0^1 P((d-c)u + c) du$$

b. On note L_0 , L_1 et $L_{1/2}$ les polynômes interpolateurs de Lagrange associés aux réels 0, 1 et 1/2. Explicitement :

$$L_0(X) = 2X^2 - 3X + 1; \quad L_1(X) = 2X^2 - X; \quad L_{1/2}(X) = 4X - 4X^2$$

7. En évitant de très lourds calculs.

On continuera de noter L_0 , L_1 et $L_{1/2}$ les fonctions polynomiales associées à ces polynômes.⁸

Etablir que :

$$\forall u \in [0, 1], P((d-c)u + c) = f(c)L_0(u) + f(d)L_1(u) + f(m)L_{1/2}(u)$$

c. Calculer les intégrales $\int_0^1 L_0(u) du$, $\int_0^1 L_1(u) du$ et $\int_0^1 L_{1/2}(u) du$.

d. Dédurre des questions précédentes que :

$$\int_c^d P(x) dx = \frac{d-c}{6} (f(c) + 4f(m) + f(d))$$

A la lumière des calculs précédents, la méthode de Simpson consiste à approcher $\int_a^b f(x) dx$ par :

$$S_n(f) = \frac{b-a}{6n} \sum_{k=0}^{n-1} (f(x_k) + 4f(m_k) + f(x_{k+1})) \quad \text{où on a noté : } m_k = \frac{x_k + x_{k+1}}{2} \quad (\star)$$

Comme dans les deux premières parties, on cherche à majorer l'erreur $|I - S_n(f)|$.

9. Estimation de l'erreur. Jusqu'à la fin de cette question, c et d continuent de désigner deux réels quelconques de $[a, b]$, avec $c < d$. On rappelle par ailleurs que dans cette partie, la fonction f est supposée de classe $\mathcal{C}^4$ sur $[a, b]$ et à valeurs réelles.

On définit une fonction g sur $\left[0, \frac{d-c}{2}\right]$ en posant :

$$\forall t \in \left[0, \frac{d-c}{2}\right], g(t) = \int_{m-t}^{m+t} f(x) dx - \frac{t}{3} (f(m-t) + 4f(m) + f(m+t)) \quad \text{où } m = \frac{c+d}{2}$$

a. Après avoir brièvement justifié que la fonction g est de classe $\mathcal{C}^4$, calculer ses dérivées successives g' , g'' et $g^{(3)}$. On vérifiera en particulier que :

$$\forall t \in \left[0, \frac{d-c}{2}\right], g^{(3)}(t) = \frac{t}{3} (f^{(3)}(m-t) - f^{(3)}(m+t))$$

b. Soit t un réel non nul de $\left[0, \frac{d-c}{2}\right]$. Etablir que : $\exists \alpha \in]m-t, m+t[, g^{(3)}(t) = -\frac{2t^2}{3} f^{(4)}(\alpha)$

c. D'après la question précédente, on a : $\forall t \in \left[0, \frac{d-c}{2}\right], |g^{(3)}(t)| \leq \frac{2M_4}{3} t^2$ où $M_4 = \sup_{x \in [a, b]} |f^{(4)}(x)|$.

En déduire successivement les majorations suivantes, valides pour tout réel $t \in \left[0, \frac{d-c}{2}\right]$:

$$\Leftrightarrow |g^{(2)}(t)| \leq \frac{2M_4}{9} t^3 \quad \Leftrightarrow |g'(t)| \leq \frac{M_4}{18} t^4 \quad \Leftrightarrow |g(t)| \leq \frac{M_4}{90} t^5$$

8. On pourra observer que $L_0(0) = 1$, et que L_0 s'annule en $1/2$ et 1 . En adaptant ce qui doit l'être, les mêmes propriétés sont satisfaites par les polynômes L_1 et $L_{1/2}$.

d. En déduire que : $|I - S_n(f)| \leq \frac{M_4(b-a)^5}{2880n^4}$

Partie IV — Vers une généralisation : approximation polynomiale de degré arbitraire

L'objectif de cette partie est de donner une piste pour généraliser les méthodes précédentes, en montrant en particulier que toute fonction de classe $\mathcal{C}^{n+1}$ sur un segment peut être "correctement approchée" par un polynôme de degré n dans un sens rendu plus précis par l'ultime question de ce problème.

Soit n un entier naturel non nul, et $\alpha_0, \dots, \alpha_n$ ($n+1$) réels distincts du segment $[a, b]$.

On pose pour tout $k \in \llbracket 0, n \rrbracket$: $L_k = \prod_{j=0, j \neq k}^n \frac{X - \alpha_j}{\alpha_k - \alpha_j}$.

10. Soit k un entier naturel arbitraire. Quel est le degré de L_k ?

11. Etablir que pour tout $\forall k \in \llbracket 0, n \rrbracket, \forall i \in \llbracket 0, n \rrbracket, L_k(\alpha_i) = \delta_{ik}$.

12. Soit f une fonction définie sur $[a, b]$ et à valeurs réelles. On pose : $P = \sum_{k=0}^n f(\alpha_k) L_k$.

Montrer que P est l'unique polynôme de $\mathbb{R}_n[X]$ tel que : $\forall i \in \llbracket 0, n \rrbracket, P(\alpha_i) = f(\alpha_i)$.

13. On désire maintenant mesurer la qualité de l'approximation réalisée lorsque l'on approche la fonction f par le polynôme P défini ci-dessus. Pour cela, on suppose (jusqu'à la fin du problème) que f est une fonction de classe $\mathcal{C}^{n+1}$ sur $[a, b]$, et on pose par ailleurs :

$$\Pi_{n+1} = \prod_{i=0}^n (X - \alpha_i)$$

On veut prouver qu'il existe un réel $\beta \in [a, b]$ tel que : $f(x) - P(x) = \frac{\Pi_{n+1}(x)}{(n+1)!} f^{(n+1)}(\beta)$ (*)

a. Soit $x \in \{\alpha_0, \dots, \alpha_n\}$. Etablir le résultat (*).

b. Soit à présent $x \in [a, b], x \notin \{\alpha_0, \dots, \alpha_n\}$. On introduit la fonction F définie par :

$$\forall t \in [a, b], F(t) = f(t) - P(t) - K \Pi_{n+1}(t)$$

avec K constante réelle choisie de sorte que : $F(x) = 0$.

Justifier l'existence de la constante K et observer que F possède au moins $(n+2)$ valeurs d'annulation distinctes. En déduire l'existence d'un réel $\beta \in [a, b]$ tel que : $F^{(n+1)}(\beta) = 0$. Conclure.

c. En déduire que : $\|f - P\| \leq \frac{\|\Pi_{n+1}\|}{(n+1)!} \|f^{(n+1)}\|$ (où l'on a noté : $\|g\| = \sup_{x \in [a, b]} |g(x)|$)

CORRIGÉ.

1) **Preuve du théorème de Rolle.**

THÉORÈME DE ROLLE. Soient c et d deux réels tels que $c < d$, et φ une fonction continue sur $[c, d]$ et dérivable sur $]c, d[$. On suppose que $\varphi(c) = \varphi(d)$. Sous ces hypothèses :

$$\exists \alpha \in]c, d[, \varphi'(\alpha) = 0$$

La fonction φ étant continue sur le segment $[c, d]$, elle est bornée et atteint ses bornes sur $[c, d]$ (théorème des bornes atteintes). Notons m son minimum et M son maximum.

► Si $m = M$: alors φ est constante sur $[c, d]$. Par suite : $\forall \alpha \in]c, d[, \varphi'(\alpha) = 0$.

► Si $m < M$: distinguons alors deux sous-cas : soit $\varphi(c) = m$, soit $\varphi(c) \neq m$.

➤ Si $\varphi(c) = m$: alors on a aussi $\varphi(d) = m$ (par hypothèse), donc le maximum M de φ est atteint en un réel α de $]c, d[$ (puisque $m \neq M$), et on a donc $\varphi'(\alpha) = 0$.

➤ Si $\varphi(c) \neq m$: alors on a aussi $\varphi(d) \neq m$ (par hypothèse), donc le minimum m de φ est atteint en un réel α de $]c, d[$, et on a donc $\varphi'(\alpha) = 0$.

Conclusion. Dans tous les cas : $\exists \alpha \in]c, d[, \varphi'(\alpha) = 0$.

Partie I — Méthode des rectangles médians

2) D'après la relation de Chasles pour les intégrales, on a : $I = \sum_{k=0}^{n-1} \int_{x_k}^{x_{k+1}} f(x) dx$.

Il s'ensuit déjà que : $I - R_n(f) = \sum_{k=0}^{n-1} \left[\int_{x_k}^{x_{k+1}} f(x) dx - R_{n,k}(f) \right]$.

Par ailleurs, pour tout entier k compris entre 0 et $n-1$, on a : $R_{n,k}(f) = h \times f(m_k) = \int_{x_k}^{x_{k+1}} f(m_k) dx$.

On en déduit grâce à (♠) et par linéarité de l'intégrale que : $I - R_n(f) = \sum_{k=0}^{n-1} \int_{x_k}^{x_{k+1}} [f(x) - f(m_k)] dx$.

3) **Estimation de l'erreur.**

a) Avec la définition de l'énoncé, on a : $\Phi(c) = f(d) - f(c) - (d-c)f'(c) - A \frac{(d-c)^2}{2}$ et $\Phi(d) = 0$.

Par conséquent : $[\Phi(c) = \Phi(d) = 0] \iff \left[f(d) - f(c) - (d-c)f'(c) - A \frac{(d-c)^2}{2} = 0 \right] \iff$

$$\left[A = \frac{2}{(d-c)^2} (f(d) - f(c) - (d-c)f'(c)) \right].$$

Ce qui prouve l'existence et l'unicité du réel A tel que $\Phi(c) = \Phi(d) = 0$.

b) D'après les théorèmes généraux et l'hypothèse faite sur f , la fonction Φ est dérivable sur $]c, d[$ et continue sur $[c, d]$; on peut lui appliquer le théorème de Rolle pour affirmer : $\exists \alpha \in]c, d[, \Phi'(\alpha) = 0$.

Or pour tout réel $x \in]c, d[$ on a :

$$\Phi'(x) = -f'(x) - (d-x)f''(x) + f'(x) + A(d-x) = -(d-x)f''(x) + A(d-x)$$

Par suite : $\exists \alpha \in]c, d[, -(d-\alpha)f''(\alpha) + A(d-\alpha) = 0 \iff (d-\alpha)f''(\alpha) = A(d-\alpha) \iff A = f''(\alpha)$.⁹

Conclusion. $\exists \alpha \in]c, d[, A = f''(\alpha)$.

c) On déduit des deux questions précédentes qu'il existe un réel $\alpha \in]c, d[$ tel que :

$$f''(\alpha) = \frac{2}{(d-c)^2} (f(d) - f(c) - (d-c)f'(c)) \iff f(d) - f(c) - (d-c)f'(c) = \frac{(d-c)^2}{2} f''(\alpha)$$

$$\iff f(d) = f(c) + (d-c)f'(c) + \frac{(d-c)^2}{2} f''(\alpha)$$

Conclusion. $\exists \alpha \in]c, d[, f(d) = f(c) + (d-c)f'(c) + \frac{(d-c)^2}{2} f''(\alpha)$.

d) La fonction f étant supposée de classe $\mathcal{C}^2$ sur $[a, b]$, sa dérivée seconde est continue sur $[a, b]$; à ce titre f'' est bornée sur $[a, b]$ d'après le théorème des bornes atteintes.

Par suite : $\exists M_2 \in \mathbb{R}_+, \forall x \in [a, b], |f''(x)| \leq M_2$.

e) Soit k un entier quelconque compris entre 0 et $(n-1)$, et soit x un réel du segment $[x_k, x_{k+1}]$. D'après la question c), il existe un réel α compris entre m_k et x tel que :

$$f(x) = f(m_k) + (x - m_k)f'(m_k) + \frac{(x - m_k)^2}{2} f''(\alpha)$$

$$\text{d'où : } |f(x) - f(m_k) - (x - m_k)f'(m_k)| = \frac{(x - m_k)^2}{2} |f''(\alpha)|$$

On en déduit grâce à d) que : $|f(x) - f(m_k) - (x - m_k)f'(m_k)| \leq M_2 \frac{(x - m_k)^2}{2}$.

L'entier k et le réel x étant arbitraires dans ce qui précède, on peut conclure :

$$\forall k \in \llbracket 0, n-1 \rrbracket, \forall x \in [m_k, m_{k+1}], |f(x) - f(m_k) - (x - m_k)f'(m_k)| \leq M_2 \frac{(x - m_k)^2}{2}$$

f) Soit k un entier quelconque compris entre 0 et $(n-1)$, et soit x un réel du segment $[x_k, x_{k+1}]$. En appliquant la formule de Taylor avec reste intégrale entre m_k et x , on obtient :

$$f(x) = f(m_k) + (x - m_k)f'(m_k) + \int_{m_k}^x (x-t) f''(t) dt \quad (\spadesuit)$$

Puisque la fonction f'' est continue entre m_k et x , elle y est bornée et atteint ses bornes : il existe donc deux valeurs m et M de f'' telles que : $m \leq f \leq M$.

Dans le cas où $x \geq m_k$, on en déduit que : $m \int_{m_k}^x (x-t) dt \leq \int_{m_k}^x (x-t) f''(t) dt \leq M \int_{m_k}^x (x-t) dt$

9. La dernière équivalence provenant de ce que $(d-\alpha) \neq 0$, puisque $\alpha \in]c, d[$.

$$\text{D'où : } m \frac{(x - m_k)^2}{2} \leq \int_{m_k}^x (x - t) f''(t) dt \leq M \frac{(x - m_k)^2}{2}$$

$$\text{D'où : } m \leq \frac{2}{(x - m_k)^2} \int_{m_k}^x (x - t) f''(t) dt \leq M$$

Ainsi, le terme du milieu de cet encadrement apparaît comme une valeur intermédiaire de la fonction f'' , et on peut donc affirmer qu'il existe un réel α compris entre m_k et x tel que :

$$f''(\alpha) = \frac{2}{(x - m_k)^2} \int_{m_k}^x (x - t) f''(t) dt \text{ d'où : } \int_{m_k}^x (x - t) f''(t) dt = \frac{(x - m_k)^2}{2} f''(\alpha) \quad (\clubsuit)$$

$$\text{On déduit de } (\spadesuit) \text{ et de } (\clubsuit) \text{ que : } \boxed{f(x) = f(m_k) + (x - m_k)f'(m_k) + \frac{(x - m_k)^2}{2} f''(\alpha)}.$$

Dans le cas où $x < m_k$, le même raisonnement donne la même conclusion (il suffit de changer les signes des inégalités du premier encadrement de cette page).

g) Soit k un entier quelconque compris entre 0 et $(n - 1)$.

► Par linéarité de l'intégrale :

$$\int_{x_k}^{x_{k+1}} f(x) - f(m_k) - (x - m_k)f'(m_k) dx = \int_{x_k}^{x_{k+1}} f(x) - f(m_k) dx - \int_{x_k}^{x_{k+1}} (x - m_k)f'(m_k) dx$$

$$\text{Or }^{10} : \int_{x_k}^{x_{k+1}} (x - m_k)f'(m_k) dx = 0.$$

$$\text{Par suite : } \int_{x_k}^{x_{k+1}} f(x) - f(m_k) - (x - m_k)f'(m_k) dx = \int_{x_k}^{x_{k+1}} f(x) - f(m_k) dx \quad (\spadesuit)$$

$$\text{► Par ailleurs : } \left| \int_{x_k}^{x_{k+1}} f(x) - f(m_k) - (x - m_k)f'(m_k) dx \right| \leq \int_{x_k}^{x_{k+1}} |f(x) - f(m_k) - (x - m_k)f'(m_k)| dx$$

D'après la question e) on a alors :

$$\left| \int_{x_k}^{x_{k+1}} f(x) - f(m_k) - (x - m_k)f'(m_k) dx \right| \leq \int_{x_k}^{x_{k+1}} M_2 \frac{(x - m_k)^2}{2} dx$$

Il reste à observer que :

$$\int_{x_k}^{x_{k+1}} M_2 \frac{(x - m_k)^2}{2} dx = \frac{M_2}{6} [(x - m_k)^3]_{x_k}^{x_{k+1}} = \frac{M_2}{6} \times 2 \times \left(\frac{h}{2}\right)^3 = \frac{M_2}{24} h^3$$

$$\text{pour obtenir : } \left| \int_{x_k}^{x_{k+1}} f(x) - f(m_k) - (x - m_k)f'(m_k) dx \right| \leq M_2 \frac{(b - a)^3}{24n^3} \quad (\clubsuit).$$

$$\text{On déduit alors de } (\spadesuit) \text{ et de } (\clubsuit) \text{ que : } \boxed{\forall k \in \llbracket 0, n - 1 \rrbracket, \left| \int_{x_k}^{x_{k+1}} f(x) - f(m_k) dx \right| \leq M_2 \frac{(b - a)^3}{24n^3}}.$$

10. Par un calcul sans difficulté ou par un argument géométrique évident.

D'après la question 2 : $I - R_n(f) = \sum_{k=0}^{n-1} \int_{x_k}^{x_{k+1}} [f(x) - f(m_k)] dx$. Grâce au résultat que l'on vient d'établir

et à l'inégalité triangulaire, on en déduit : $|I - R_n(f)| \leq \sum_{k=0}^{n-1} M_2 \frac{(b-a)^3}{24n^3}$.

D'où finalement : $|I - R_n(f)| \leq M_2 \frac{(b-a)^3}{24n^2}$.

4) a) La fonction `rectangles(f, a, b, n)` ci-dessous reçoit comme paramètres deux réels a et b , une fonction f et un entier naturel non nul n , et retourne la valeur de $R_n(f)$.

```

1  def rectangles(f, a, b, n):
2      S = 0
3      h = (b - a) / n
4      for k in range(n):
5          S = S + h * f(a + (h / 2) + k * h)
6      return S

```

b) La fonction `rectangles_v2(f, a, b, M_2, eps)` ci-dessous reçoit comme paramètres deux réels a et b , une fonction f , un réel M_2 (un majorant de $|f''|$ sur $[a, b]$), et un réel eps , et retourne une valeur de $R_n(f)$ approchant $\int_a^b f$ avec une erreur inférieure ou égale à eps .

```

1  def rectangles_v2(f, a, b, M_2, eps):
2      S = 0
3      n = 1
4      while (M_2 * (b - a)**3 / (24 * n**2)) > eps:
5          n = n + 1
6      return rectangles(f, a, b, n)

```

Partie II — Méthode des trapèzes

5) La réponse à cette question repose essentiellement sur la formule donnant l'aire d'un trapèze... Pour tout entier k compris entre 0 et $n-1$ on a : $T_{n,k}(f) = h \times \frac{f(x_k) + f(x_{k+1})}{2}$.

$$\text{D'où : } T_n(f) = \sum_{k=0}^{n-1} h \times \frac{f(x_k) + f(x_{k+1})}{2} = \frac{h}{2} \sum_{k=0}^{n-1} [f(x_k) + f(x_{k+1})]$$

$$\text{C'est à dire : } T_n(f) = \frac{b-a}{2n} \sum_{k=0}^{n-1} [f(x_k) + f(x_{k+1})].$$

6) En utilisant la relation de Chasles pour les intégrales, on peut déjà écrire :

$$\int_a^b f(x) dx - T_n(f) = \sum_{k=0}^{n-1} \int_{x_k}^{x_{k+1}} f(x) dx - \sum_{k=0}^{n-1} T_{n,k}(f).$$

Or par définition de g_k , on a : $\forall k \in \llbracket 0, n-1 \rrbracket$, $T_{n,k}(f) = \int_{x_k}^{x_{k+1}} g_k(x) dx$.

Grâce à ces observations et par linéarité de l'intégrale : $\int_a^b f(x) dx - T_n(f) = \sum_{k=0}^{n-1} \int_{x_k}^{x_{k+1}} f(x) - g_k(x) dx$.

7) **Estimation de l'erreur.** Soit k un entier compris entre 0 et $(n-1)$.

a) La fonction δ'_k et la fonction qui à tout réel x du segment $[x_k, x_{k+1}]$ associe $(x - x_k)(x - x_{k+1})$ sont de classe $\mathcal{C}^1$ sur $[x_k, x_{k+1}]$, ce qui rend légitime l'intégration par parties suivante :

$$\int_{x_k}^{x_{k+1}} \delta''_k(x) (x - x_k)(x - x_{k+1}) dx = \underbrace{[\delta'_k(x) (x - x_k)(x - x_{k+1})]_{x_k}^{x_{k+1}}}_{=0} - \int_{x_k}^{x_{k+1}} \delta'_k(x) (2x - x_k - x_{k+1}) dx$$

La fonction δ_k et la fonction qui à tout réel x du segment $[x_k, x_{k+1}]$ associe $2x - x_k - x_{k+1}$ sont de classe $\mathcal{C}^1$ sur $[x_k, x_{k+1}]$, ce qui rend légitime la seconde intégration par parties suivante :

$$\int_{x_k}^{x_{k+1}} \delta''_k(x) (x - x_k)(x - x_{k+1}) dx = \underbrace{[\delta_k(x) (2x - x_k - x_{k+1})]_{x_k}^{x_{k+1}}}_{=0} + 2 \int_{x_k}^{x_{k+1}} \delta_k(x) dx$$

Conclusion. $\int_{x_k}^{x_{k+1}} \delta_k(x) dx = \frac{1}{2} \int_{x_k}^{x_{k+1}} \delta''_k(x) (x - x_k)(x - x_{k+1}) dx$.

b) Soit k un entier quelconque compris entre 0 et $(n-1)$. D'après la question précédente, et par définition de δ_k on a :

$$\int_{x_k}^{x_{k+1}} \delta_k(x) dx = \frac{1}{2} \int_{x_k}^{x_{k+1}} \delta''_k(x) (x - x_k)(x - x_{k+1}) dx$$

Observons alors que pour tout réel $x \in [x_k, x_{k+1}]$, on a : $\delta''_k(x) = f''(x)$, puisque g_k est affine (et qu'elle a à ce titre une dérivée seconde identiquement nulle). Par suite :

$$\left| \int_{x_k}^{x_{k+1}} \delta_k(x) dx \right| \leq \frac{M_2}{2} \int_{x_k}^{x_{k+1}} |(x - x_k)(x - x_{k+1})| dx = -\frac{M_2}{2} \int_{x_k}^{x_{k+1}} (x - x_k)(x - x_{k+1}) dx \quad (\spadesuit)$$

$$\begin{aligned} \text{Or : } \int_{x_k}^{x_{k+1}} (x - x_k)(x - x_{k+1}) dx &= \left[\frac{x^3}{3} - \frac{x_k + x_{k+1}}{2} x^2 + x_k x_{k+1} x \right]_{x_k}^{x_{k+1}} \\ &= \frac{1}{6} (2x_{k+1}^3 - 3x_k x_{k+1}^2 - 3x_{k+1}^3 + 6x_k x_{k+1}^2 - 2x_k^3 + 3x_k^3 + 3x_k^2 x_{k+1} - 6x_k^2 x_{k+1}) \\ &= \frac{1}{6} (x_k^3 - 3x_k^2 x_{k+1} + 3x_k x_{k+1}^2 - x_{k+1}^3) = \frac{1}{6} (x_k - x_{k+1})^3 = -\frac{h^3}{6} \quad (\clubsuit) \end{aligned}$$

D'après $(\spadesuit)$ et $(\clubsuit)$: $\left| \int_{x_k}^{x_{k+1}} \delta_k(x) dx \right| \leq M_2 \frac{h^3}{12}$

Autrement écrit : $\forall k \in \llbracket 0, n-1 \rrbracket, \left| \int_{x_k}^{x_{k+1}} [f(x) - g_k(x)] dx \right| \leq M_2 \frac{(b-a)^3}{12n^3}$

On déduit de ce résultat, de la question 6 et de l'inégalité triangulaire que :

$$\left| \int_a^b f(x) dx - T_n(f) \right| \leq \sum_{k=0}^{n-1} M_2 \frac{(b-a)^3}{12n^3} \quad \text{d'où : } \left| I - T_n(f) \right| \leq M_2 \frac{(b-a)^3}{12n^2}.$$

Partie III — Méthode de Simpson

8) Le polynôme P est de degré au plus 2 (trivial), et il coïncide avec f en c , m et d (c'est une vérification immédiate). Puisque les réels c , m et d sont distincts, le principe du prolongement algébrique assure l'unicité d'un polynôme P vérifiant ces conditions.

Conclusion. $\exists ! P \in \mathbb{R}_2[X]$, $P(c) = f(c)$, $P(d) = f(d)$, $P(m) = f(m)$.

9) a) Le très mystérieux changement de variable suggéré dans l'énoncé consiste à poser : $x = (d - c)u + c$. Ainsi : $dx = (d - c) du$ et il reste à observer que u prend la valeur 0 (*resp.* 1) lorsque x prend la valeur c (*resp.* d) pour obtenir :

$$\int_c^d P(x) dx = (d - c) \int_0^1 P((d - c)u + c) du.$$

b) Notons $Q = f(c)L_0(u) + f(d)L_1(u) + f(m)L_{1/2}(u)$. Le polynôme Q est de degré au plus 2, en tant que combinaison linéaire des polynômes de Lagrange (qui sont dans $\mathbb{R}_2[X]$). De plus : $Q(0) = f(c)$, $Q(1/2) = f(m)$ et $Q(1) = f(d)$.

D'autre part, la fonction $u \mapsto P((d - c)u + c)$ est elle aussi polynomiale de degré au plus 2, et vérifie : $P(c) = f(c)$, $P(m) = f(m)$ et $P(d) = f(d)$.

On en déduit que P et Q sont deux fonctions polynomiales de degré au plus 2, telles que : $P(0) = Q(0)$, $P(1/2) = Q(1/2)$ et $P(1) = Q(1)$. D'après le principe du prolongement algébrique, elles sont égales.

Conclusion. $\forall u \in [0, 1]$, $P((d - c)u + c) = f(c)L_0(u) + f(d)L_1(u) + f(m)L_{1/2}(u)$.

c) On a : $\int_0^1 L_0(u) du = \int_0^1 2u^2 - 3u + 1 = \frac{1}{6} [4x^3 - 9x^2 + 6x]_0^1$ d'où : $\int_0^1 L_0(u) du = \frac{1}{6}$

On a : $\int_0^1 L_1(u) du = \int_0^1 2u^2 - u = \frac{1}{6} [4x^3 - 3x^2]_0^1$ d'où : $\int_0^1 L_1(u) du = \frac{1}{6}$

Enfin : $\int_0^1 L_{1/2}(u) du = \int_0^1 4u - 4u^2 = \frac{1}{6} [12x^2 - 8x^3]_0^1$ d'où : $\int_0^1 L_{1/2}(u) du = \frac{4}{6}$

d) D'après la question 9-a : $\int_c^d P(x) dx = (d - c) \int_0^1 P((d - c)u + c) du$

On en déduit grâce à la question 9-b : $\int_c^d P(x) dx = (d - c) \int_0^1 f(c)L_0(u) + f(d)L_1(u) + f(m)L_{1/2}(u) du$

D'où, par linéarité de l'intégrale :

$$\int_c^d P(x) dx = (d - c) \left[f(c) \int_0^1 L_0(u) du + f(d) \int_0^1 L_1(u) du + f(m) \int_0^1 L_{1/2}(u) du \right]$$

Enfin, d'après la question précédente : $\int_c^d P(x) dx = (d - c) \left[f(c) \times \frac{1}{6} + f(d) \times \frac{1}{6} + f(m) \times \frac{4}{6} \right]$

C'est à dire : $\int_c^d P(x) dx = \frac{d - c}{6} (f(c) + 4f(m) + f(d))$.

10) **Estimation de l'erreur.** a) Notons F une primitive de f sur $[c, d]$ (l'existence de F est assurée par la continuité de f sur $[c, d]$). On a :

$$\forall t \in \left[0, \frac{d-c}{2}\right], \int_{m-t}^{m+t} f(t) dt = F(m+t) - F(m-t)$$

En particulier, la fonction $t \in \left[0, \frac{d-c}{2}\right] \mapsto \int_{m-t}^{m+t} f(t) dt$ est de classe $\mathcal{C}^4$ (et même $\mathcal{C}^5$) sur $\left[0, \frac{d-c}{2}\right]$.

Cette observation faite, les théorèmes généraux et l'hypothèse selon laquelle f est de classe $\mathcal{C}^4$ sur $[a, b]$ impliquent que la fonction g est de classe $\mathcal{C}^4$ sur $\left[0, \frac{d-c}{2}\right]$. Il est alors légitime de calculer ses dérivées successives pour obtenir :

$$\begin{aligned} \forall t \in \left[0, \frac{d-c}{2}\right], g'(t) &= f(m+t) + f(m-t) - \frac{1}{3} (f(m-t) + 4f(m) + f(m+t)) \\ &\quad - \frac{t}{3} (-f'(m-t) + f'(m+t)) \end{aligned}$$

$$\iff \forall t \in \left[0, \frac{d-c}{2}\right], g'(t) = \frac{2}{3}f(m+t) + \frac{2}{3}f(m-t) - \frac{4}{3}f(m) - \frac{t}{3} (-f'(m-t) + f'(m+t))$$

$$\begin{aligned} \implies \forall t \in \left[0, \frac{d-c}{2}\right], g''(t) &= \frac{2}{3}f'(m+t) - \frac{2}{3}f'(m-t) - \frac{t}{3} (f''(m-t) + f''(m+t)) \\ &\quad - \frac{1}{3} (-f'(m-t) + f'(m+t)) \end{aligned}$$

$$\iff \forall t \in \left[0, \frac{d-c}{2}\right], g''(t) = \frac{1}{3}f'(m+t) - \frac{1}{3}f'(m-t) - \frac{t}{3} (f''(m-t) + f''(m+t))$$

$$\begin{aligned} \implies \forall t \in \left[0, \frac{d-c}{2}\right], g^{(3)}(t) &= \frac{1}{3}f''(m+t) + \frac{1}{3}f''(m-t) - \frac{t}{3} (-f^{(3)}(m-t) - f^{(3)}(m+t)) \\ &\quad - \frac{1}{3}f''(m-t) - \frac{1}{3}f''(m+t) \end{aligned}$$

$$\text{Par suite : } \boxed{\forall t \in \left[0, \frac{d-c}{2}\right], g^{(3)}(t) = \frac{t}{3} (f^{(3)}(m-t) - f^{(3)}(m+t))}$$

b) Soit t un réel strictement positif dans $\left[0, \frac{d-c}{2}\right]$. La fonction $f^{(3)}$ est de classe $\mathcal{C}^1$ sur $[m-t, m+t]$, puisque $[m-t, m+t] \subset [c, d] \subset [a, b]$ et que f est supposée de classe $\mathcal{C}^4$ sur $[a, b]$.

On peut donc appliquer le théorème des accroissements finis à la fonction $f^{(3)}$ sur $[m-t, m+t]$ pour affirmer :

$$\exists \alpha \in]m-t, m+t[, f^{(3)}(m+t) - f^{(3)}(m-t) = 2tf^{(4)}(\alpha)$$

On déduit de cette relation et de celle établie dans la question précédente que :

$$\exists \alpha \in]m-t, m+t[, g^{(3)}(t) = \frac{t}{3} \times (-2tf^{(4)}(\alpha)) = -\frac{2t^2}{3} f^{(4)}(\alpha)$$

Conclusion. $\exists \alpha \in]m - t, m + t[$, $g^{(3)}(t) = -\frac{2t^2}{3} f^{(4)}(\alpha)$

c) Tout au long de cette question, t désigne un réel de $\left[0, \frac{d-c}{2}\right]$.

On a : $g^{(2)}(t) = \int_0^t g^{(3)}(u) du + g^{(2)}(0)$

Or $g^{(2)}(0) = 0$ (d'après 10-a). D'où : $|g^{(2)}(t)| = \left| \int_0^t g^{(3)}(u) du \right| \leq \int_0^t |g^{(3)}(u)| du$

D'après la question précédente : $|g^{(2)}(t)| \leq \frac{2M_4}{3} \int_0^t u^2 du$ c'est à dire : $|g^{(2)}(t)| \leq \frac{2M_4}{9} t^3$

On réitère le processus avec g' . On a : $g'(t) = \int_0^t g^{(2)}(u) du + g'(0)$

Or $g'(0) = 0$ (d'après 10-a). D'où : $|g'(t)| = \left| \int_0^t g^{(2)}(u) du \right| \leq \int_0^t |g^{(2)}(u)| du$

D'après ce qui précède : $|g'(t)| \leq \frac{2M_4}{9} \int_0^t u^3 du$ c'est à dire : $|g'(t)| \leq \frac{M_4}{18} t^4$

On ne change pas de méthode, et on réitère le processus avec g . On a : $g(t) = \int_0^t g'(u) du + g(0)$

Or $g(0) = 0$ (par définition même). D'où : $|g(t)| = \left| \int_0^t g'(u) du \right| \leq \int_0^t |g'(u)| du$

D'après ce qui précède : $|g(t)| \leq \frac{M_4}{18} \int_0^t u^4 du$ c'est à dire : $|g(t)| \leq \frac{M_4}{90} t^5$

d) Soit k un entier compris entre 0 et $n-1$. On commence par observer que d'après la question précédente¹¹ :

$$\left| \int_{x_k}^{x_{k+1}} f(x) dx - \frac{h}{6} (f(x_k) + 4f(m_k) + f(x_{k+1})) \right| \leq \frac{M_4}{90} \left(\frac{h}{2}\right)^5$$

D'où : $\left| \int_{x_k}^{x_{k+1}} f(x) dx - \frac{b-a}{6n} (f(x_k) + 4f(m_k) + f(x_{k+1})) \right| \leq \frac{M_4 (b-a)^5}{2880n^5}$

On en déduit, grâce à la définition de $S_n(f)$ et par le biais d'une ultime application de l'inégalité triangulaire que :

$$|I - S_n(f)| \leq \frac{M_4 (b-a)^5}{2880n^4}$$

23.10.3 Буняковски, ou l'histoire d'un mathématicien oublié

Les résultats que l'on utilise de nos jours en Mathématiques, et en particulier ceux du cours de cette année, ont connu de nombreuses évolutions. Leurs énoncés sont souvent le fruit du travail de plusieurs Mathématiciens, contemporains ou non. Il arrive alors que le nom de l'un d'entre eux soit oublié... L'inégalité de Cauchy-Schwarz-Buniakovski (Буняковски) en fait partie¹².

11. On applique le résultat de la question 10.c, en prenant $c = x_k$ et $d = x_{k+1}$ (de telle sorte que $d - c = h$) et $t = h/2$.

12. Au même titre d'ailleurs que, paraît-il, le théorème de d'Alembert-Gauss ou le théorème de Rolle.

Dans le cas de l'inégalité évoquée dans ce chapitre, il me semble¹³ particulièrement injuste¹⁴ que le nom de Буняковски soit souvent omis.

Pour achever cette parenthèse plus légère, le fait qu'une découverte mathématique ne porte pas toujours le nom de son auteur est un cas particulier de la *loi d'éponymie de Stigler*, énoncée sous ce nom par le statisticien et humoriste¹⁵ Stephen Stigler en 1999, qui affirme pour faire court :

“Une découverte scientifique ne porte jamais le nom de son auteur.”

Selon Stigler, cette loi ne serait pas de lui, mais de R. Merton (sociologue américain). Il ajoute, à propos des noms propres donnés aux découvertes scientifiques (éponymies), qu'ils “ne sont que rarement proposés et jamais universellement reconnus, sauf si celui qui nomme est très distant dans le temps ou dans l'espace du scientifique ainsi honoré”.

13. Et c'est extrêmement subjectif.

14. Il ne faut bien entendu y voir aucune sympathie pro-slave.

15. Je ne sais pas si Stigler est vraiment un humoriste, mais je crois qu'il faut une énorme dose d'auto-dérision et d'amour du second degré pour énoncer une loi comme celle dont il est question dans ce paragraphe.

Chapitre 24

Applications linéaires en dimension finie

24.1 Rang d'une famille, rang d'une application linéaire

DÉFINITION 24.1 - Soient E un $\mathbb{K}$ -ev et $\mathcal{F}$ une famille de vecteurs de E .

On appelle **rang de la famille** $\mathcal{F}$ la dimension du sous-espace vectoriel de E engendré par $\mathcal{F}$.

Notation. On note $\text{rg}(\mathcal{F}) = \dim_{\mathbb{K}} \text{Vect}(\mathcal{F})$ le rang de $\mathcal{F}$.

Exemples. Dans $\mathbb{R}^2$, le rang de la famille $\mathcal{F} = \{(1, 2), (1, -1)\}$ est 2 ; dans $\mathbb{R}^3$, le rang de la famille $\mathcal{F} = \{(1, 2, 3), (4, 5, 6), (7, 8, 9)\}$ est 2 ; dans $\mathbb{R}[X]$, le rang de la famille $\mathcal{F} = \{1, X + 1, X^3 - 2, X^7\}$ est 4 ; dans $\mathcal{C}^0(\mathbb{R}, \mathbb{R})$, le rang de la famille $\mathcal{F} = \{\exp, \text{ch}, \text{sh}\}$ est 2 ; dans un $\mathbb{K}$ -ev de dimension n , une famille $\mathcal{F}$ est génératrice si et seulement si son rang est égal à n .

DÉFINITION 24.2 - Soit $f \in \mathcal{L}(E, F)$. Le **rang de f** est la dimension de l'image de f . On le note : $\text{rg}(f) = \dim_{\mathbb{K}} \text{Im } f$ (ou simplement $\text{rg}(f) = \dim \text{Im } f$ s'il n'y a pas d'ambiguïté sur $\mathbb{K}$).

PROPRIÉTÉ 24.1 - Soit $f \in \mathcal{L}(E, F)$, où E et F sont deux $\mathbb{K}$ -ev de dimension finie.

1/ $\text{rg}(f) \leq \min(\dim E, \dim F)$

2/ $[\text{rg}(f) = \dim F] \iff [f \text{ est surjective}]$

3/ $[\text{rg}(f) = \dim E] \iff [f \text{ est injective}]$

PREUVE. 1/ D'une part, $f(E)$ est un sev de F , donc : $\dim(f(E)) \leq \dim F$. Par ailleurs, si $\mathcal{B} = (v_1, \dots, v_n)$ désigne une base de E (en ayant noté $n = \dim E$), on a : $f(E) = \text{Vect}(f(v_1), \dots, f(v_n))$. Le sev $f(E)$ étant engendré par n vecteurs, il est au plus de dimension n . En clair : $\dim(f(E)) \leq \dim E$.

On a donc $\dim(f(E)) \leq \dim E$ et $\dim(f(E)) \leq \dim F$, d'où : $\text{rg}(f) \leq \min(\dim E, \dim F)$.

2/ Si $\text{rg}(f) = \dim F$, alors $f(E)$ est un sev de F de dimension égale à $\dim F$; par suite $f(E) = F$. Ce qui signifie que f est surjective, et prouve l'implication de la gauche vers la droite. La réciproque est triviale.

3/ On a déjà établi que f est injective si et seulement si l'image d'une famille libre par f est encore libre.

Le point 3/ s'en déduit. ☕

24.2 Le théorème du rang et ses conséquences

Droit au but !

THÉORÈME 24.1 - (Théorème du rang). Soient E et F deux $\mathbb{K}$ -ev de dimension finie, et $f \in \mathcal{L}(E, F)$. On a :

$$\dim E = \dim \ker f + \operatorname{rg} f$$

PREUVE. ➤ Puisque $\ker f$ est un sev de E , et que E est de dimension finie, $\ker f$ admet un supplémentaire S dans E . En clair, il existe un sev S de E tel que : $E = \ker f \oplus S$. On construit alors une application :

$$\begin{aligned} \tilde{f} : S &\longrightarrow \operatorname{Im} f \\ \vec{v} &\longmapsto f(\vec{v}) \end{aligned}$$

En premier lieu, observons que f est bien définie (c'est-à-dire à valeurs dans $\operatorname{Im} f$), linéaire (car f l'est), et injective (car f l'est).

➤ Montrons que $\tilde{f}$ est surjective : soit $\vec{w} \in \operatorname{Im} f$. Il existe $\vec{v} \in E$ tel que : $f(\vec{v}) = \vec{w}$ (♠).

Par ailleurs, comme $E = \ker f \oplus S : \exists ! (\vec{u}, \vec{s}) \in \ker f \times S$, $\vec{v} = \vec{u} + \vec{s}$ (♡)

De (♠) et (♡), on déduit que : $\vec{w} = f(\vec{u} + \vec{s}) = \underbrace{f(\vec{u})}_{=\vec{0}_F} + f(\vec{s})$ d'où $\vec{w} = f(\vec{s})$, et puisque $s \in S$,

$$\vec{w} = \tilde{f}(\vec{s}).$$

En résumé : $\forall \vec{w} \in \operatorname{Im} f, \exists \vec{s} \in S, \vec{w} = \tilde{f}(\vec{s})$. Ce qui signifie que $\tilde{f}$ est surjective.

➤ Il s'ensuit que $\tilde{f}$ est un isomorphisme entre S et $\operatorname{Im} f$. D'où en particulier : $\dim S = \dim \operatorname{Im} f$.

➤ D'autre part, on déduit de $E = \ker f \oplus S$ la relation : $\dim E = \dim \ker f + \dim S$.

De ces deux égalités entre dimensions, on déduit enfin : $\dim E = \dim \ker f + \dim \operatorname{Im} f$, soit encore :

$$\dim E = \dim \ker f + \operatorname{rg} f. \quad \text{☕}$$

Remarques. Il suffit en fait de supposer que seul E est de dimension finie dans cet énoncé. En outre, le point clef de la preuve est que f induit un isomorphisme entre un supplémentaire de $\ker f$ dans E et l'image de f (ce résultat est parfois appelé théorème “noyau/image”).

COROLLAIRE 24.1 - Soient E et F deux $\mathbb{K}$ -ev de dimension finie, et $f \in \mathcal{L}(E, F)$.

1/ Si $\dim E > \dim F$, alors f n'est pas injective.

2/ Si $\dim E < \dim F$, alors f n'est pas surjective.

PREUVE. Triviale. ☕

COROLLAIRE 24.2 - Soient E et F deux $\mathbb{K}$ -ev de même dimension finie ($\dim E = \dim F$), et $f \in \mathcal{L}(E, F)$. LASSE :

- 1/ f est injective.
- 2/ f est surjective.
- 3/ f est un isomorphisme.

PREUVE. Notons $n = \dim E = \dim F$, et observons que d'après le théorème du rang :

$$\dim E = \dim \ker f + \dim \operatorname{Im} f \quad (\spadesuit)$$

- 1) $\implies$ 2) : si f est injectif, alors $\ker f = \{\vec{0}_E\}$ donc $\dim \ker f = 0$ d'où grâce à $(\spadesuit)$: $\dim \operatorname{Im} f = n$. Donc $\operatorname{Im} f = F$ et f est donc surjectif.
- 2) $\implies$ 3) : si f est surjectif, alors $\operatorname{Im} f = F$ donc $\dim \operatorname{Im} f = n$ d'où grâce à $(\spadesuit)$: $\dim \ker f = 0$. Donc $\ker f = \{\vec{0}_E\}$ et f est injectif, d'où f est un isomorphisme.
- 3) $\implies$ 1) : trivial.
- Conclusion : $1) \implies 2) \implies 3) \implies 1)$ donc : $\boxed{1) \iff 2) \iff 3)}$. ☕

En particulier :

COROLLAIRE 24.3 - Soient E un $\mathbb{K}$ -ev de dimension finie, et $f \in \mathcal{L}(E)$ (un endomorphisme de E). LASSE :

- 1/ f est injective.
- 2/ f est surjective.
- 3/ f est un automorphisme de E .

PREUVE. Prendre $E = F$ dans le corollaire précédent ! ☕

Application. L'application linéaire $\varphi : P \in \mathbb{K}_n[X] \mapsto (P - P') \in \mathbb{K}_n[X]$ est un automorphisme de $\mathbb{K}_n[X]$ (ie : $\varphi \in \operatorname{GL}(\mathbb{K}_n[X])$).

En effet :

- L'application φ est bien définie car si P est un polynôme de degré au plus n , alors il en est de même du polynôme $P - P'$. En outre, un calcul aisé permet d'établir que φ est linéaire, c'est à dire que : $\forall (\lambda, \mu) \in \mathbb{K}^2, \varphi(\lambda P + \mu Q) = \lambda \varphi(P) + \mu \varphi(Q)$. Par suite : $\boxed{\varphi \in \mathcal{L}(\mathbb{K}_n[X])}$.
- Soit $P \in \mathbb{K}_n[X]$. Alors : $P \in \ker \varphi \iff P = P'$. Or seul le polynôme nul est égal à son polynôme dérivé. Donc $\ker \varphi = \{\vec{0}\}$. Donc φ est injectif.
- Puisque φ est un endomorphisme injectif de $\mathbb{K}_n[X]$ (qui est de dimension finie), on déduit du corollaire ci-dessus que $\boxed{\varphi \text{ est un automorphisme de } \mathbb{K}_n[X]}$.

ATTENTION : l'énoncé précédent n'est valide qu'en dimension finie, comme l'illustre l'exemple ci-dessous.

Contre-exemple en dimension infinie. L'application $\Delta : P \in \mathbb{K}[X] \mapsto P' \in \mathbb{K}[X]$ est un endomorphisme surjectif de $\mathbb{K}[X]$. Cependant, Δ n'est pas un automorphisme de $\mathbb{K}[X]$.

L'application $\Delta : P \in \mathbb{K}[X] \mapsto P' \in \mathbb{K}[X]$ est un endomorphisme de $\mathbb{K}[X]$, par linéarité de la dérivation formelle.

Montrons sa surjectivité : soit $P \in \mathbb{K}[X]$. Il existe un entier naturel n (n'importe quel majorant du degré de P) et $n+1$ scalaires $\alpha_0, \dots, \alpha_n$ tels que : $P = \sum_{k=0}^n \alpha_k X^k$. On pose judicieusement : $Q = \sum_{k=0}^n \frac{\alpha_k}{k+1} X^{k+1}$, de telle sorte que $\Delta(Q) = P$.

On a ainsi établi que tout $P \in \mathbb{K}[X]$ appartient à l'image de Δ , ce qui prouve la surjectivité de Δ .

En revanche, Δ n'est pas injectif (par exemple car $\Delta(X) = \Delta(X+3)$), et n'est donc pas un automorphisme de $\mathbb{K}[X]$.

Conclusion. Δ est un endomorphisme injectif et non surjectif de $\mathbb{K}[X]$.

On achève ce paragraphe par une propriété qui prendra toute sa signification dans la suite des événements.

PROPRIÉTÉ 24.2 - (Invariance du rang par composition avec un isomorphisme). Soient E , F et G trois $\mathbb{K}$ -ev de dimension finie, $f \in \mathcal{L}(E, F)$ et $g \in \mathcal{L}(F, G)$.

1/ $[g \text{ est un isomorphisme}] \implies [\text{rg}(g \circ f) = \text{rg}(f)]$

2/ $[f \text{ est un isomorphisme}] \implies [\text{rg}(g \circ f) = \text{rg}(g)]$

PREUVE. Pseudo-triviale. ☕

Remarque : pour le 2), il suffit en fait de supposer que f est surjective.

24.3 Formes linéaires et hyperplans

24.3.1 Formes linéaires

DÉFINITION 24.3 - Soit E un $\mathbb{K}$ -ev.

Une **forme linéaire sur E** est une application linéaire de E dans $\mathbb{K}$ (un élément de $\mathcal{L}(E, \mathbb{K})$).

Notation. On pourra noter E^* l'espace vectoriel des formes linéaires sur E .

Exemples : 1) L'application : $(x, y) \in \mathbb{R}^2 \mapsto y \in \mathbb{R}$ est une forme linéaire sur $\mathbb{R}^2$.

2) L'application : $(x, y, z) \in \mathbb{R}^3 \mapsto x + y - 2z \in \mathbb{R}$ est une forme linéaire sur $\mathbb{R}^3$.

3) L'application (d'évaluation en 0) : $P \in \mathbb{K}[X] \mapsto P(0) \in \mathbb{K}$ est une forme linéaire sur $\mathbb{K}[X]$.

4) L'application : $f \in \mathcal{C}^0(\mathbb{R}, \mathbb{K}) \mapsto \int_a^b f(t)dt \in \mathbb{R}$ est une forme linéaire sur $\mathcal{C}^0(\mathbb{R}, \mathbb{K})$ (pour tous réels distincts a et b).

5) L'application $\text{tr} : M \in M_n(\mathbb{K}) \mapsto \text{tr}(M) = \sum_{i=1}^n m_{ii} \in \mathbb{K}$ qui à une matrice carrée de $M_n(\mathbb{K})$ associe sa trace est une forme linéaire sur $M_n(\mathbb{K})$.

6) L'application $(u_n)_n \in \mathbb{K}^{\mathbb{N}} \mapsto u_0 \in \mathbb{K}$ qui à une suite à valeurs dans $\mathbb{K}$ associe son premier terme est une forme linéaire sur $\mathbb{K}^{\mathbb{N}}$.

7) Dans $\mathbb{R}^2$ muni du produit scalaire usuel, on fixe un vecteur $\vec{v}$. L'application $\vec{u} \in \mathbb{R}^2 \mapsto \vec{u} \cdot \vec{v} \in \mathbb{R}$ qui à un vecteur $\vec{u}$ associe son produit scalaire avec $\vec{v}$ est une forme linéaire sur $\mathbb{R}^2$.

8) Soit $n \in \mathbb{N}$, $n \geq 2$. L'application $\Delta : M \in M_n(\mathbb{K}) \mapsto \sum_{i=1}^n (m_{i1} - m_{i2}) \in \mathbb{K}$ qui à une matrice carrée de $M_n(\mathbb{K})$ associe la différence des sommes des coefficients de ses première et seconde colonne est une forme linéaire sur $M_n(\mathbb{K})$.

PROPRIÉTÉ 24.3 - Soient E un $\mathbb{K}$ -ev, et f une forme linéaire sur E . Si f n'est pas nulle, alors f est surjective.

Autrement dit, toute forme linéaire non nulle est surjective.

PREUVE. $f(E)$ est un sev de $\mathbb{K}$, qui est de dimension 1. Si f est non nulle, alors $f(E)$ est de dimension égale à 1, donc $f(E) = \mathbb{K}$ d'où la conclusion. ☕

PROPRIÉTÉ 24.4 - Soit E un $\mathbb{K}$ -ev de dimension finie.

Si f une forme linéaire non nulle sur E , alors $\ker f$ est un hyperplan de E .

PREUVE. Conséquence immédiate de la propriété précédente et du théorème du rang. ☕

Exemples. Dans chacun des exemples 1) à 8) évoqués précédemment, la forme f est surjective. Dans l'exemple 1), le noyau de f est la droite d'équation $y = 0$; dans l'exemple 2), le noyau de f est le plan d'équation $x + y - 2z = 0$; dans l'exemple 5), le noyau de f est l'hyperplan de $M_n(\mathbb{K})$ constitué des matrices de trace nulle; dans le 7), sous réserve que $\vec{v} \neq \vec{0}$, le noyau de f est la droite vectorielle de vecteur normal $\vec{v}$; dans le 8), le noyau de f est l'hyperplan de $M_n(\mathbb{K})$ constitué des matrices dont les sommes des coefficients des deux premières colonnes sont égales. Tous ces noyaux sont des hyperplans dans leurs espaces vectoriels ambiants respectifs.

Dans les autres exemples, les espaces vectoriels considérés n'étant pas de dimension finie, les noyaux des différentes formes linéaires ne sont pas des hyperplans. Ce sont néanmoins des sev, et explicitement : dans l'exemple 3), le noyau de f est le sev de $\mathbb{K}[X]$ des polynômes s'annulant en 0; dans l'exemple 4), le noyau de f est le sev de $\mathcal{C}^0(\mathbb{R}, \mathbb{K})$ des fonctions continues et d'intégrale nulle sur $[a, b]$; dans l'exemple 6), le noyau de f est le sev de $\mathbb{K}^{\mathbb{N}}$ des suites à valeurs dans $\mathbb{K}$ de premier terme nul.

24.3.2 Correspondance entre formes linéaires et hyperplans

THÉORÈME 24.2 - Soit E un $\mathbb{K}$ -ev de dimension $n \geq 1$, et soit H un sev de E . LASSE :

- 1/ H est un hyperplan de E ;
- 2/ il existe un vecteur $\vec{v} \in E$, $\vec{v} \neq \vec{0}_E$ tel que $E = H \oplus \text{Vect}(\vec{v})$;
- 3/ il existe $f \in E^*$, $f \neq 0_{E^*}$ (une forme linéaire sur E non nulle) telle que $H = \ker f$.

PREUVE. $1/ \implies 3/$. Soit H un hyperplan de E . Puisque H est un $\mathbb{K}$ -ev de dimension finie (égale à $n-1$), il admet une base $\mathcal{B} = (\vec{v}_i)_{i \in [1, n-1]}$, qui est en particulier une famille libre de vecteurs de E .

En vertu du théorème de la base incomplète, cette base peut être complétée en une base de E . Explicitement, il existe un vecteur $\vec{w} \in E$ (et évidemment $\vec{w} \notin H$) tel que $\mathcal{B}' = (\vec{v}_1, \dots, \vec{v}_{n-1}, \vec{w})$ est une base de E .

Pour tout vecteur $\vec{v}$ de E , il existe alors un unique n -uplet $(\alpha_1, \dots, \alpha_{n-1}, \beta)$ de scalaires tel que :

$$\vec{v} = \left(\sum_{i=1}^{n-1} \alpha_i \vec{v}_i \right) + \beta \vec{w}.$$

On considère la forme coordonnée suivante :

$$\begin{array}{ccc} w^* : E & \longrightarrow & \mathbb{K} \\ \vec{v} & \longmapsto & \beta \end{array}$$

Il est alors clair que : $H = \ker w^*$.

Conclusion : $[H \text{ est un hyperplan de } E] \implies [\text{il existe } f \in E^*, f \neq 0_{E^*} \text{ telle que } H = \ker f]$

► $3/ \iff 1/$. Supposons qu'il existe $f \in E^*$, $f \neq 0_{E^*}$ telle que $H = \ker f$. Puisque f est une forme linéaire non nulle, elle est surjective, donc de rang 1. En vertu du théorème du rang, on a donc : $\dim \ker f = n-1$. Donc $H = \ker f$ est un hyperplan de E .

Conclusion : $[\text{il existe } f \in E^*, f \neq 0_{E^*} \text{ telle que } H = \ker f] \implies [H \text{ est un hyperplan de } E]$

► Ainsi : **Conclusion** : $[\text{il existe } f \in E^*, f \neq 0_{E^*} \text{ telle que } H = \ker f] \iff [H \text{ est un hyperplan de } E]$.

► La preuve de $1/ \iff 2/$ est contenue dans le raisonnement précédent. ☕

Le théorème ci-dessus affirme en particulier que tout hyperplan est le noyau d'une forme linéaire ; mais pour un hyperplan fixé, on n'a pas unicité de cette forme, et l'énoncé suivant précise ce fait.

THÉORÈME 24.3 - Soit E un $\mathbb{K}$ -ev de dimension $n \geq 1$.

Deux formes linéaires sur E ont même noyau SSI elles sont proportionnelles.

En d'autres termes, pour f et g dans E^* , on a :

$$[\ker f = \ker g] \iff [\exists \lambda \in \mathbb{K}^*, g = \lambda f]$$

PREUVE. Raisonnement par double implication. Supposons $\ker f = \ker g$. Il existe alors $(v_1, \dots, v_{n-1}, w)$ une base de E telle que $(v_1, \dots, v_{n-1})$ soit une base de $\ker f$, et $w \notin \ker f$.

Posons alors : $\lambda = \frac{g(w)}{f(w)}$. Cette définition est légitime puisque $f(w)$ et $g(w)$ sont deux scalaires, et $g(w)$ puisque $w \notin \ker g$ par hypothèse.

Soit alors V un vecteur quelconque de E . Il existe n scalaires tels que : $V = \sum_{k=1}^n \alpha_k v_k + \beta w$.

Par hypothèse et par linéarité de f (resp. de g), on a : $f(V) = \beta f(w)$ et $g(V) = \beta g(w)$.

Il s'ensuit que : $g(V) = \lambda f(V)$.

Le vecteur V étant arbitraire dans le raisonnement précédent, on a établi que : $g = \lambda f$. Ce qui prouve la première implication.

La réciproque est triviale. ☕

Pour un $\mathbb{K}$ -ev de dimension finie, on a ainsi une **correspondance bijective** entre les hyperplans de E et les formes linéaires non nulles définies à une constante multiplicative près.

THÉORÈME 24.4 - Soit E un $\mathbb{K}$ -ev de dimension $n \geq 1$, et soit F un sev de dimension r .

Alors F est l'intersection de $n - r$ hyperplans.

Plus explicitement, il existe $n - r$ formes linéaires non nulles $(f_i)_{i \in [1, n-r]}$ sur E telles que :

$$F = \bigcap_{i=1}^{n-r} \ker f_i$$

PREUVE. Soit F un sev de dimension r d'un $\mathbb{K}$ -espace vectoriel E de dimension $n \geq 1$.

Puisque F est de dimension finie, il admet une base $(\vec{v}_1, \dots, \vec{v}_r)$, que l'on peut compléter (grâce au théorème de la base incomplète) pour obtenir une base : $\mathcal{B} = (\vec{v}_1, \dots, \vec{v}_r, \vec{w}_1, \dots, \vec{w}_{n-r})$ de E .

Notons alors $w_1^*, \dots, w_{n-r}^*$ les formes coordonnées associées aux vecteurs $\vec{w}_i$, c'est-à-dire pour mémoire les applications :

$$\begin{array}{ccc} w_i^* : E & \longrightarrow & \mathbb{K} \\ \vec{v} = \sum_{i=1}^r \alpha_i \vec{v}_i + \sum_{i=1}^{n-r} \beta_i \vec{w}_i & \longmapsto & \beta_i \end{array}$$

Les applications w_i^* sont des formes linéaires non nulles sur E (qui est un $\mathbb{K}$ -ev de dimension finie), donc leurs noyaux sont des hyperplans de E (conséquence du théorème du rang).

Prouvons alors l'égalité : $F = \bigcap_{i=1}^{n-r} \ker w_i^*$ par double inclusion.

► Montrons $F \subset \bigcap_{i=1}^{n-r} \ker w_i^*$. Soit $\vec{v} \in F$. Alors il existe r scalaires $\alpha_1, \dots, \alpha_r$ tels que : $\vec{v} = \sum_{i=1}^r \alpha_i \vec{v}_i$ (tous les β_i sont donc nuls). On en déduit que : $\forall i \in \llbracket 1, n-r \rrbracket, w_i^*(\vec{v}) = 0$. Il revient au même de dire que : $\forall i \in \llbracket 1, n-r \rrbracket, \vec{v} \in \ker w_i^*$.

Par suite : $\vec{v} \in \bigcap_{i=1}^{n-r} \ker w_i^*$. D'où : $F \subset \bigcap_{i=1}^{n-r} \ker w_i^* \quad (\spadesuit)$

► Montrons $\bigcap_{i=1}^{n-r} \ker w_i^* \subset F$. Soit $\vec{v} \in \bigcap_{i=1}^{n-r} \ker w_i^*$. Alors il existe n scalaires $\alpha_1, \dots, \alpha_r, \beta_1, \dots, \beta_{n-r}$ tels que : $\vec{v} = \sum_{i=1}^r \alpha_i \vec{v}_i + \sum_{i=1}^{n-r} \beta_i \vec{w}_i$. Lorsque $\vec{v} \in \bigcap_{i=1}^{n-r} \ker w_i^*$, on a : $\forall i \in \llbracket 1, n-r \rrbracket, w_i^*(\vec{v}) = 0$. Ainsi tous les β_i

sont nuls et on en déduit que $\vec{v} = \sum_{i=1}^r \alpha_i \vec{v}_i$, et donc : $\vec{v} \in F$. D'où : $\bigcap_{i=1}^{n-r} \ker w_i^* \subset F \quad (\heartsuit)$ On déduit

de $(\spadesuit)$ et de $(\heartsuit)$ que : $\bigcap_{i=1}^{n-r} \ker w_i^* = F$. Puisque les sev $\ker w_i^*$ sont des hyperplans de E , on a établi que :

Si $\dim E = n$, tout sev de dimension r de E est l'intersection de $(n-r)$ hyperplans de E



Précisions concernant la preuve ci-dessus. Dans cette démonstration, on a construit deux bases. Une pour F (celle des vecteurs $\vec{v}_i$) et une pour un supplémentaire de F dans E (celle des vecteurs $\vec{w}_i$). Pour que ces bases existent¹, il est indispensable de supposer que $r > 0$ et $n-r > 0$. Il convient donc de traiter à part dans la preuve les cas $r = 0$ et $r = n$. On observe simplement que si $r = n$, alors $F = E$ et F est bien l'intersection de... zéro hyperplan de E ; et dans le cas où $r = 0$, on a $F = \{\vec{0}_E\}$, et F est l'intersection des n hyperplans $\ker v_i^*$, où les v_i^* sont les formes coordonnées associées à une quelconque base $(\vec{v}_i)_{i \in \llbracket 1, n \rrbracket}$ de E . Fin de la parenthèse.

D'une certaine manière, le théorème précédent généralise ce que vous connaissez déjà de la géométrie dans l'espace de Terminale, où vous avez vu qu'une droite peut être définie comme l'intersection de deux plans (et donc par deux équations).

Exemple d'application (du théorème du rang) et d'illustration (du théorème ci-dessus). Dans $M_{10}(\mathbb{R})$, l'ensemble F des matrices dont les sommes des coefficients de chaque colonne sont égales est un sous-espace vectoriel. Plus précisément, F est de dimension 91, et c'est donc l'intersection de 9 hyperplans.

1. Pour qu'elles contiennent au moins un vecteur.

24.4 Parenthèse : inverse à gauche, inverse à droite

Une ultime application du théorème du rang :

PROPRIÉTÉ 24.5 - Soit E un $\mathbb{K}$ -ev de dimension finie, et soit $f \in \mathcal{L}(E)$. LASSE :

1/ $f \in \text{GL}(E)$

2/ $\exists g \in \mathcal{L}(E), g \circ f = \text{id}_E$ (et alors $g = f^{-1}$)

3/ $\exists g \in \mathcal{L}(E), f \circ g = \text{id}_E$ (et alors $g = f^{-1}$)

PREUVE. Supposons que : $\exists g \in \mathcal{L}(E), g \circ f = \text{id}_E$. Alors f est injective, et donc bijective puisque c'est un endomorphisme d'un ev de dimension finie. D'où f est un automorphisme de E . Ce qui montre l'implication : 2/ $\implies$ 1/.

L'implication : 3/ $\implies$ 1/ se démontre de façon analogue, en remplaçant dans le discours précédent “injective” par “surjective”.

Les implications 1/ $\implies$ 2/ et 1/ $\implies$ 3/ sont immédiates. ☕

En clair, un endomorphisme de E est bijectif si et seulement si il admet un inverse à gauche **ou** à droite.

La traduction matricielle de cette propriété est la suivante :

PROPRIÉTÉ 24.6 - Soit $n \in \mathbb{N}^*$, et soit $A \in M_n(\mathbb{K})$. LASSE :

1) $A \in \text{GL}_n(\mathbb{K})$

2) $\exists B \in M_n(\mathbb{K}), BA = I_n$ (et alors $B = A^{-1}$)

3) $\exists B \in M_n(\mathbb{K}), AB = I_n$ (et alors $B = A^{-1}$)

PREUVE. Comme indiqué plus haut, c'est l'interprétation en termes de matrices de la propriété précédente (voir paragraphes suivants pour la correspondance entre applications linéaires et matrices). ☕

Observons pour conclure que l'énoncé est complètement **faux** dans un espace vectoriel de dimension **infinie**. Les endomorphismes de $\mathbb{R}[X]$ (qui est un $\mathbb{R}$ -ev de dimension infinie)

$$\begin{array}{ccc} \Delta : \mathbb{R}[X] \longmapsto \mathbb{R}[X] & \text{et} & \mathcal{J} : \mathbb{R}[X] \longmapsto \mathbb{R}[X] \\ P \longmapsto P' & & P = \sum_{k=0}^n a_k X^k \longmapsto \sum_{k=0}^n \frac{a_k}{k+1} X^{k+1} \end{array}$$

vérifient $\Delta \circ \mathcal{J} = \text{id}_{\mathbb{R}[X]}$. Donc Δ est inversible à droite, et $\mathcal{J}$ est inversible à gauche. Mais ni l'un ni l'autre ne sont bijectifs, puisque Δ (*resp.* $\mathcal{J}$) n'est clairement pas injectif (*resp.* pas surjectif).

24.5 Matrice d'une application linéaire

24.5.1 Définition et généralités

DÉFINITION 24.4 - Soient E et F deux $\mathbb{K}$ -ev, de dimensions respectives p et n ; $f \in \mathcal{L}(E, F)$; $\mathcal{B} = (\vec{v}_1, \dots, \vec{v}_p)$ une base de E ; et $\mathcal{B}' = (\vec{w}_1, \dots, \vec{w}_n)$ une base de F .

La **matrice de f dans les bases $\mathcal{B}$ et $\mathcal{B}'$** , notée $\text{Mat}_{\mathcal{B}, \mathcal{B}'}(f)$ est la matrice de $M_{np}(\mathbb{K})$ suivante :

$$\text{Mat}_{\mathcal{B}, \mathcal{B}'}(f) = (\alpha_{ij})_{1 \leq i \leq n, 1 \leq j \leq p}$$

où les scalaires α_{ij} sont caractérisés par :

$$\forall j \in \llbracket 1, p \rrbracket, f(\vec{v}_j) = \sum_{i=1}^n \alpha_{ij} \vec{w}_i$$

Traduction. La matrice de f est obtenue en écrivant en colonnes les coordonnées dans la base $\mathcal{B}'$ des images par f des vecteurs de la base $\mathcal{B}$ (par $\heartsuit$).

Notation. Lorsque f est un endomorphisme de E , et que $\mathcal{B}$ est une base de E , on note simplement $\text{Mat}_{\mathcal{B}}(f)$ la matrice $\text{Mat}_{\mathcal{B}, \mathcal{B}}(f)$.

ILLUSTRATION : on reprend les notations de la définition.

- $f : E \longrightarrow F$ est une application linéaire ;
- $\mathcal{B} = (v_1, \dots, v_n)$ est une base de E ;
- $\mathcal{B}' = (w_1, \dots, w_m)$ est une base de F ;

$$M_{\mathcal{B}, \mathcal{B}'}(f) = \begin{matrix} f(v_1) \cdots f(v_n) \\ w_1 \\ w_2 \\ \vdots \\ w_m \end{matrix} \left(\begin{array}{c} \\ \\ \\ \end{array} \right)$$

EXEMPLES

Exemple 1. On considère l'application linéaire $f : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ définie par $f(x, y, z) = (x + 2y, 3y - 4z)$. On appelle $\mathcal{B}$ et $\mathcal{B}'$ les bases canoniques respectives de $\mathbb{R}^3$ et de $\mathbb{R}^2$. Plus précisément, on note :

$$\mathcal{B} = (e_1, e_2, e_3) \quad \text{et} \quad \mathcal{B}' = (e'_1, e'_2)$$

On a :

$$f(e_1) = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = 1 e'_1 + 0 e'_2; \quad f(e_2) = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = 2 e'_1 + 3 e'_2; \quad f(e_3) = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = 0 e'_1 + -4 e'_2;$$

D'où :

$$M_{\mathcal{B}, \mathcal{B}'}(f) = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 3 & -4 \end{pmatrix} \quad \text{Remarque : } [M_{\mathcal{B}, \mathcal{B}'}(f)] \begin{pmatrix} x \\ y \\ z \end{pmatrix} = (x + 2y, 3y - 4z)$$

Remarque : on pose $\mathcal{B}'' = (e'_2, e'_1)$. On a alors :

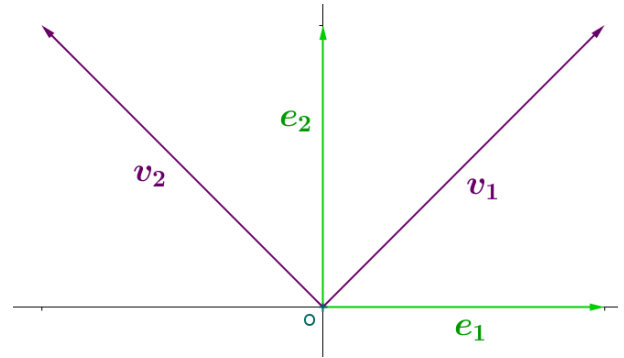
$$M_{\mathcal{B}, \mathcal{B}''}(f) = \begin{pmatrix} 0 & 3 & -4 \\ 1 & 2 & 0 \end{pmatrix}$$

Exemple 2. Dans cet exemple on considère $E = \mathbb{R}^2$. On considère la projection orthogonale p sur (Ox) , qui est comme chacun sait une application linéaire $p : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ (et donc un endomorphisme).

1) Dans un premier temps, on prend $\mathcal{B} = \mathcal{B}' = (e_1, e_2)$ (base canonique de $\mathbb{R}^2$). Dans ce cas, on parle de matrice de l'endomorphisme p dans la base $\mathcal{B}$ (et pas “dans les bases $\mathcal{B}$ et $\mathcal{B}'$ ”); et cette matrice est notée $M_{\mathcal{B}}(p)$ (et pas $M_{\mathcal{B}, \mathcal{B}}(p)$).

Comme : $p(e_1) = 1 e_1 + 0 e_2$ et $p(e_2) = 0 e_1 + 0 e_2$,
la matrice de p dans la base $\mathcal{B}$ est :

$$M_{\mathcal{B}}(p) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$



2) Dans un second temps, on prend $\mathcal{B} = (v_1, v_2)$ et $\mathcal{B}' = (e_1, e_2)$ (voir figure).

Comme : $p(v_1) = 1 e_1 + 0 e_2$ et $p(v_2) = -e_1 + 0 e_2$,
la matrice de p dans les bases $\mathcal{B}$ et $\mathcal{B}'$ est :

$$M_{\mathcal{B}, \mathcal{B}'}(p) = \begin{pmatrix} 1 & 0 \\ -1 & 0 \end{pmatrix}$$

3) Dans un dernier temps, on prend $\mathcal{B} = \mathcal{B}' = (v_1, v_2)$.

Comme : $p(v_1) = (1/2) v_1 - (1/2) v_2$ et $p(v_2) = -(1/2) v_1 + (1/2) v_2$,
la matrice de p dans la base $\mathcal{B}'$ est :

$$M_{\mathcal{B}'}(p) = \frac{1}{2} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$$

Remarque. Il ressort de ces premiers exemples qu'une application linéaire ne possède pas une unique matrice. On ne peut parler de la matrice d'une application linéaire que relativement à une base de l'év de départ et une base de l'év d'arrivée. En outre, une fois ces bases choisies, la matrice associée à une application linéaire peut être très “simple” ou plus “compliquée”; nous verrons que ces différentes matrices possèdent néanmoins des propriétés communes.

Exemple 3. L'application nulle de E dans F a pour matrice la matrice nulle, dans toutes bases $\mathcal{B}$ de E et $\mathcal{B}'$ de F .

Exemple 4. L'application linéaire $\varphi : P \in \mathbb{R}_2[X] \mapsto (P(0), P(1), P(2)) \in \mathbb{R}^3$ a pour matrice

$$\text{Mat}_{\mathcal{B}, \mathcal{B}'}(\varphi) = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \end{pmatrix} \text{ où } \mathcal{B} \text{ et } \mathcal{B}' \text{ désignent les bases canoniques de } \mathbb{R}_2[X] \text{ et de } \mathbb{R}^3 \text{ respective-}$$

ment. En revanche, on a :

$$\text{Mat}_{\mathcal{B}'', \mathcal{B}'}(\varphi) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \text{ si on choisit pour } \mathcal{B}'' \text{ la base de } \mathbb{R}_2[X] \text{ constituée des polynômes de}$$

Lagrange associés aux valeurs 0, 1 et 2.

Exemple 5. L'endomorphisme $f : P \in \mathbb{R}_3[X] \mapsto P(X+1) \in \mathbb{R}_3[X]$ a pour matrice

$$\text{Mat}_{\mathcal{B}}(f) = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 1 \end{pmatrix} \text{ où } \mathcal{B} \text{ désigne la base canonique de } \mathbb{R}_3[X].$$

Exemple 6. L'application linéaire $f : (x, y, z, t) \in \mathbb{R}^4 \mapsto (x+2y+3t, 4x-5z+6t) \in \mathbb{R}^2$ a pour matrice

$$\text{Mat}_{\mathcal{B}, \mathcal{B}'}(f) = \begin{pmatrix} 1 & 2 & 3 & 0 \\ 4 & 0 & -5 & 6 \end{pmatrix} \text{ où } \mathcal{B} \text{ et } \mathcal{B}' \text{ désignent les bases canoniques de } \mathbb{R}^4 \text{ et } \mathbb{R}^2 \text{ respectivement.}$$

Exemple 7. L'application identité $\text{id} : \mathbb{R}_2[X] \rightarrow \mathbb{R}_2[X]$ a pour matrice

$$\text{Mat}_{\mathcal{B}}(\text{id}) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \text{I}_3 \text{ où } \mathcal{B} \text{ désigne une base quelconque de } \mathbb{R}_2[X]. \text{ Mais on a :}$$

$$\text{Mat}_{\mathcal{B}', \mathcal{B}}(\text{id}) = \begin{pmatrix} 1 & -1 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix} = P_{\mathcal{B}\mathcal{B}'} \text{ où } \mathcal{B} \text{ désigne la base canonique de } \mathbb{R}_2[X], \text{ et } \mathcal{B}' \text{ la base}$$

$$(1, (X-1), (X-1)^2).$$

Exemple 8. Plus généralement, si E est un $\mathbb{K}$ -ev de dimension $n \geq 1$, on a

$$\text{Mat}_{\mathcal{B}}(\text{id}_E) = \text{I}_n \text{ pour toute base } \mathcal{B} \text{ de } E, \text{ et :}$$

$$\text{Mat}_{\mathcal{B}', \mathcal{B}}(\text{id}_E) = P_{\mathcal{B}\mathcal{B}'} \text{ pour toutes bases } \mathcal{B} \text{ et } \mathcal{B}' \text{ de } E.$$

L'énoncé ci-dessous précise le lien entre une application linéaire f et une matrice la représentant.

PROPRIÉTÉ 24.7 - Soient E et F deux $\mathbb{K}$ -ev, de dimensions respectives p et n ; $f \in \mathcal{L}(E, F)$; $\mathcal{B}$ une base de E , et $\mathcal{B}'$ une base de F . Soit $\vec{v}$ un vecteur de E , dont on note $X_{\mathcal{B}}$ le p -uplet des coordonnées de $\vec{v}$ dans la base $\mathcal{B}$. On note encore $Y_{\mathcal{B}'}$ le n -uplet des coordonnées de $f(\vec{v})$ dans la base $\mathcal{B}'$. Alors :

$$Y_{\mathcal{B}'} = \text{Mat}_{\mathcal{B}, \mathcal{B}'}(f) X_{\mathcal{B}}$$

PREUVE. Notons :

- $\mathcal{B} = (\vec{v}_1, \dots, \vec{v}_p)$ une base de E ;
- $\mathcal{B}' = (\vec{w}_1, \dots, \vec{w}_n)$ une base de F ;
- $\text{Mat}_{\mathcal{B}, \mathcal{B}'}(f) = (a_{ij})_{1 \leq i \leq n, 1 \leq j \leq p} \in M_{np}(\mathbb{K})$ la matrice de f dans les bases $\mathcal{B}$ et $\mathcal{B}'$, de telle sorte que :

$$\forall j \in \llbracket 1, p \rrbracket, f(\vec{v}_j) = \sum_{i=1}^n a_{ij} \vec{w}_i \quad (E1)$$

- $\vec{v}$ un vecteur de E , de coordonnées $X_{\mathcal{B}} = (x_1, \dots, x_p)$ dans la base $\mathcal{B}$, de telle sorte que :

$$\vec{v} = \sum_{i=1}^p x_i \vec{v}_i \quad (E2) ;$$

- $Y_{\mathcal{B}'} = (y_1, \dots, y_n)$ le n -uplet des coordonnées de $f(\vec{v})$ dans la base $\mathcal{B}'$, de telle sorte que :

$$f(\vec{v}) = \sum_{i=1}^n y_i \vec{w}_i \quad (E3) ;$$

- Notons enfin qu'avec ces notations on a : $[Y_{\mathcal{B}'} = \text{Mat}_{\mathcal{B}, \mathcal{B}'}(f) X_{\mathcal{B}}] \iff \left[\forall i \in \llbracket 1, n \rrbracket, y_i = \sum_{j=1}^p a_{ij} x_j \right] \quad (\spadesuit)$

En avant pour la preuve ! D'une part : $f(\vec{v}) = \sum_{i=1}^n y_i \vec{w}_i \quad (E3)$

$$\begin{aligned} \text{D'autre part : } f(\vec{v}) &= f\left(\sum_{j=1}^p x_j \vec{v}_j\right) = \sum_{j=1}^p x_j f(\vec{v}_j) = \sum_{j=1}^p x_j \sum_{i=1}^n a_{ij} \vec{w}_i = \sum_{j=1}^p \sum_{i=1}^n a_{ij} x_j \vec{w}_i \\ &= \sum_{i=1}^n \left(\sum_{j=1}^p a_{ij} x_j\right) \vec{w}_i \quad (E4) \end{aligned}$$

On déduit alors de (E3), (E4) et de l'unicité des coordonnées de $f(\vec{v})$ dans la base $\mathcal{B}'$ que :

$$\forall i \in \llbracket 1, n \rrbracket, y_i = \sum_{j=1}^p a_{ij} x_j$$

En vertu de $(\spadesuit)$, ceci équivaut à : $\boxed{Y_{\mathcal{B}'} = \text{Mat}_{\mathcal{B}, \mathcal{B}'}(f) X_{\mathcal{B}}}$.



THÉORÈME 24.5 - Mêmes notations et hypothèses que dans la propriété ci-dessus. L'application

$$\begin{aligned}\psi : \mathcal{L}(E, F) &\longmapsto M_{np}(\mathbb{K}) \\ f &\longmapsto \text{Mat}_{\mathcal{B}, \mathcal{B}'}(f)\end{aligned}$$

est une bijection.

PREUVE. Il s'agit ici de vérifier que, une base $\mathcal{B}$ de E et une base $\mathcal{B}'$ de F étant fixées, l'application qui à une application linéaire de E dans F associe sa matrice dans les bases $\mathcal{B}$ et $\mathcal{B}'$ est bijective (deux applications linéaires distinctes sont représentées par deux matrices distinctes (injectivité), et toute matrice de $M_{np}(\mathbb{K})$ est celle d'une application linéaire de E dans F (surjectivité)). La surjectivité est claire, et l'injectivité provient de ce qu'une application linéaire f est uniquement déterminée par l'image par f d'une base de E . ☕

24.5.2 Compatibilité avec les opérations matricielles

PROPRIÉTÉ 24.8 - (Linéarité) Soient E et F deux $\mathbb{K}$ -ev, de dimensions respectives p et n ; $\mathcal{B}$ une base de E , et $\mathcal{B}'$ une base de F .

Soient f et $g \in \mathcal{L}(E, F)$, soient λ et μ deux scalaires :

$$\text{Mat}_{\mathcal{B}, \mathcal{B}'}(\lambda f + \mu g) = \lambda \text{Mat}_{\mathcal{B}, \mathcal{B}'}(f) + \mu \text{Mat}_{\mathcal{B}, \mathcal{B}'}(g)$$

PREUVE. Longue mais facile. ☕

COROLLAIRE 24.4 - L'application $\psi : f \in \mathcal{L}(E, F) \mapsto \text{Mat}_{\mathcal{B}, \mathcal{B}'}(f) \in M_{np}(\mathbb{K})$ est un isomorphisme.

PREUVE. L'application ψ est bien entendu celle du théorème 24.5.2, et le corollaire une conséquence immédiate de la propriété précédente. ☕

COROLLAIRE 24.5 - Si E et F sont deux $\mathbb{K}$ -ev de dimension finie, alors $\mathcal{L}(E, F)$ est un $\mathbb{K}$ -ev de dimension finie et :

$$\dim_{\mathbb{K}} \mathcal{L}(E, F) = \dim_{\mathbb{K}} E \times \dim_{\mathbb{K}} F$$

PREUVE. Conséquence du corollaire précédent, et du fait que deux espaces vectoriels isomorphes ont la même dimension. ☕

PROPRIÉTÉ 24.9 - soient E , F et G trois $\mathbb{K}$ -ev, de bases respectives $\mathcal{B}$, $\mathcal{B}'$ et $\mathcal{B}''$; soient $f \in \mathcal{L}(E, F)$ et $g \in \mathcal{L}(F, G)$. Alors :

$$\text{Mat}_{\mathcal{B}, \mathcal{B}''}(g \circ f) = \text{Mat}_{\mathcal{B}', \mathcal{B}''}(g) \times \text{Mat}_{\mathcal{B}, \mathcal{B}'}(f)$$

PREUVE. Copier, coller et adapter la preuve de la propriété 24.7. ☕

PROPRIÉTÉ 24.10 - Soient E et F deux $\mathbb{K}$ -ev de même dimension, de bases respectives $\mathcal{B}$ et $\mathcal{B}'$; et soit $f \in \mathcal{L}(E, F)$.

f est un isomorphisme de E de F SSI sa matrice dans toutes bases $\mathcal{B}$ et $\mathcal{B}'$ (de E et F respectivement) est inversible.

PREUVE. Si f est un isomorphisme de E dans F , alors d'après la propriété précédente :

$$\text{Mat}_{\mathcal{B}', \mathcal{B}}(f^{-1}) \times \text{Mat}_{\mathcal{B}, \mathcal{B}'}(f) = \text{Mat}_{\mathcal{B}, \mathcal{B}}(f^{-1} \circ f)$$

pour toutes bases $\mathcal{B}$ et $\mathcal{B}'$ (de E et F respectivement). Or : $\text{Mat}_{\mathcal{B}, \mathcal{B}}(f^{-1} \circ f) = \text{Mat}_{\mathcal{B}, \mathcal{B}}(\text{id}_E) = I_n$.

On en déduit que $\text{Mat}_{\mathcal{B}, \mathcal{B}'}(f)$ est inversible, ce qui prouve la première implication.

Réciproquement, si $\text{Mat}_{\mathcal{B}, \mathcal{B}'}(f)$ est inversible, alors il existe une matrice N telle que : $N \times \text{Mat}_{\mathcal{B}, \mathcal{B}'}(f) = I_n$. En vertu du théorème, cette matrice N correspond est celle d'une unique application g dans les bases $\mathcal{B}'$ et $\mathcal{B}$. L'égalité matricielle ci-dessus signifie alors que $g \circ f = \text{id}_E$ (et N étant l'inverse de $\text{Mat}_{\mathcal{B}, \mathcal{B}'}(f)$, on a également $f \circ g = \text{id}_F$). On en déduit que f est un isomorphisme, ce qui complète la preuve. ☕

En particulier :

PROPRIÉTÉ 24.11 - Soient E un $\mathbb{K}$ -ev de dimension n , et $f \in \mathcal{L}(E)$ (un endomorphisme de E). Alors :

$$[f \in \text{GL}(E)] \iff [\text{Mat}_{\mathcal{B}}(f) \in \text{GL}_n(\mathbb{K})] \quad (\mathcal{B} \text{ désignant une base quelconque de } E).$$

En outre, si $f \in \text{GL}(E)$, on a : $\text{Mat}_{\mathcal{B}}(f^{-1}) = (\text{Mat}_{\mathcal{B}}(f))^{-1}$.

PREUVE. Conséquence immédiate de la propriété précédente. ☕

Exemple. L'endomorphisme $f : P \in \mathbb{R}_3[X] \mapsto P - P' \in \mathbb{R}_3[X]$ a pour matrice

$$A = \text{Mat}_{\mathcal{B}}(f) = \begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & -2 & 0 \\ 0 & 0 & 1 & -3 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \text{où } \mathcal{B} \text{ désigne la base canonique de } \mathbb{R}_3[X].$$

A est clairement inversible et il s'ensuit que f est un automorphisme de $\mathbb{R}_3[X]$. Son automorphisme réciproque a pour matrice dans la base $\mathcal{B}$:

$$A^{-1} = \text{Mat}_{\mathcal{B}}(f^{-1}) = \begin{pmatrix} 1 & 1 & 2 & 6 \\ 0 & 1 & 2 & 6 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

et explicitement on a donc :

$$\forall (a, b, c, d) \in \mathbb{R}^4, \quad f^{-1}(aX^3 + bX^2 + cX + d)$$

$$= aX^3 + (b + 3a)X^2 + (c + 2b + 6a)X + 6a + 2b + c + d$$

24.6 Rang d'une matrice

DÉFINITION 24.5 - Soit $A \in M_{np}(\mathbb{K})$.

On appelle **application linéaire canoniquement associée à la matrice** A l'application linéaire $f_A \in \mathcal{L}(\mathbb{K}^p, \mathbb{K}^n)$ définie en posant :

$$\forall X \in \mathbb{K}^p, f_A(X) = AX$$

Exemple. Posons $A = \begin{pmatrix} 1 & 1 & 1 & 2 \\ 1 & 2 & 0 & -1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$. L'application linéaire f_A canoniquement associée à A est l'application $f_A \in \mathcal{L}(\mathbb{R}^4, \mathbb{R}^3)$ définie en posant :

$$\forall (x, y, z, t) \in \mathbb{R}^4, f_A(x, y, z, t) = (x + y + z + 2t, x + 2y - t, t) \in \mathbb{R}^3$$

DÉFINITION 24.6 - Soit $A \in M_{np}(\mathbb{K})$. On appelle :

- 1/ **noyau de** A et on note $\ker A$ le noyau de f_A ;
- 2/ **image de** A et on note $\text{Im } A$ l'image de f_A ;
- 3/ **rang de** A et on note $\text{rg } A$ le rang de f_A .

Exemple. En reprenant la matrice A de l'exemple ci-dessus, l'image de A est $\mathbb{R}^3$, donc le rang de A est 3. Le noyau de A est la droite vectorielle de $\mathbb{R}^4$ engendrée par le vecteur $\vec{v}(2, -1, -1, 0)$.

Les propriétés ci-dessous sont obtenues en traduisant matriciellement celles déjà établies pour le rang d'une application linéaire.

PROPRIÉTÉ 24.12 - Soit $A \in M_{np}(\mathbb{K})$. On a :

- 1/ $\text{rg}(A) \leq \min(n, p)$;
- 2/ $[\text{rg}(A) = n] \iff [f_A \text{ est surjective}]$;
- 3/ $[\text{rg}(A) = p] \iff [f_A \text{ est injective}]$;
- 4/ *Théorème du rang* : $p = \dim \ker A + \text{rg}(A)$.

PREUVE. Voir preuves des propriétés 24.1 page 635, et 24.1 page 636.



En particulier :

PROPRIÉTÉ 24.13 - Soit $A \in M_n(\mathbb{K})$. On a :

$$[A \in GL_n(\mathbb{K})] \iff [\text{rg}(A) = n] \iff [\ker A = \{\vec{0}_{\mathbb{K}^n}\}] \iff [f_A \in GL(\mathbb{K}^n)]$$

PREUVE. Conséquences immédiates des propriétés précédentes. ☕

24.7 Changement de base

THÉORÈME 24.6 - (Formule du changement de base) Soit E un $\mathbb{K}$ -ev de dimension finie, $\mathcal{B}$ et $\mathcal{B}'$ deux bases de E , et $f \in \mathcal{L}(E)$. Alors :

$$M_{\mathcal{B}'}(f) = P_{\mathcal{B},\mathcal{B}'}^{-1} M_{\mathcal{B}}(f) P_{\mathcal{B},\mathcal{B}'}$$

PREUVE. Soient E un $\mathbb{K}$ -ev de dimension n ; $f \in \mathcal{L}(E)$; $\mathcal{B}$ et $\mathcal{B}'$ deux bases base de E .

Notons : $M = M_{\mathcal{B}}(f)$ (resp. $M' = M_{\mathcal{B}'}(f)$) la matrice de f dans la base $\mathcal{B}$ (resp. $\mathcal{B}'$).

Notons : $P = P_{\mathcal{B},\mathcal{B}'}$ la matrice de passage de la base $\mathcal{B}$ à la base $\mathcal{B}'$, de telle sorte que $P^{-1} = P_{\mathcal{B}',\mathcal{B}}$ est la matrice de passage de la base $\mathcal{B}'$ à la base $\mathcal{B}$.

Soit V un vecteur de E , dont on note $X_{\mathcal{B}}$ (resp. $X_{\mathcal{B}'}$) le n -uplet des coordonnées dans la base $\mathcal{B}$ (resp. $\mathcal{B}'$).

On introduit encore $Y_{\mathcal{B}}$ (resp. $Y_{\mathcal{B}'}$) le n -uplet des coordonnées de $f(V)$ dans la base $\mathcal{B}$ (resp. $\mathcal{B}'$).

On a alors : $Y_{\mathcal{B}'} = M' X_{\mathcal{B}'}$ (♠) (et $Y_{\mathcal{B}} = M X_{\mathcal{B}}$) (propriété de la matrice d'une application linéaire dans une base donnée).

Par ailleurs : $Y_{\mathcal{B}'} = P^{-1} Y_{\mathcal{B}}$ (propriété de la matrice de passage). D'où : $Y_{\mathcal{B}'} = P^{-1} M X_{\mathcal{B}}$ et donc : $Y_{\mathcal{B}'} = P^{-1} M P X_{\mathcal{B}'}$ (♣)

D'après (♠) et (♣) : $M' X_{\mathcal{B}'} = P^{-1} M P X_{\mathcal{B}'}$. Le vecteur V (et par suite le n -uplet $X_{\mathcal{B}'}$) étant arbitraire dans ce raisonnement, on a ainsi établi que :

$\forall X \in \mathbb{K}^n, M' X = P^{-1} M P X$ On en déduit que : $M' = P^{-1} M P$ en vertu du lemme suivant la conclusion.

Conclusion. $M_{\mathcal{B}'}(f) = P_{\mathcal{B},\mathcal{B}'}^{-1} M_{\mathcal{B}}(f) P_{\mathcal{B},\mathcal{B}'}$. ☕

LEMME 24.1 - $\forall (A, B) \in M_n(\mathbb{K})^2, [A = B] \iff [\forall X \in \mathbb{K}^n, AX = BX]$

PREUVE. Notons $A = (a_{ij})$ et $B = (b_{ij})$.

Notons aussi, pour tout entier $j \in \llbracket 1, n \rrbracket$, $e_j = (0, \dots, 0, \underbrace{1}_j, 0, \dots, 0)$ l'élément de $\mathbb{K}^n$ dont la j -ème coordonnée vaut 1 et toutes les autres sont nulles.

Supposons que $\forall X \in \mathbb{K}^n, AX = BX$. Alors en particulier : $\forall j \in \llbracket 1, n \rrbracket, Ae_j = Be_j$.

Or : $Ae_j = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{ij} \\ \vdots \\ a_{nj} \end{pmatrix}$ et $Be_j = \begin{pmatrix} b_{1j} \\ \vdots \\ b_{ij} \\ \vdots \\ b_{nj} \end{pmatrix}$ (le produit de A par e_j est égal à la j -ème colonne de la matrice A).

On en déduit donc que : $\forall i \in \llbracket 1, n \rrbracket, a_{ij} = b_{ij}$.

Puisqu'en outre l'entier j est arbitraire dans les calculs précédents, on a établi que : $\forall (i, j) \in \llbracket 1, n \rrbracket^2, a_{ij} = b_{ij}$.

Ce qui signifie que $A = B$, et prouve donc l'implication "de la droite vers la gauche".

Réciproque triviale. ☕

Remarque concernant la formule du changement de base. Avec les notations classiquement utilisées en exercice, on peut retenir cette propriété sous la forme $A' = P^{-1}AP$. Le tout étant de se souvenir qui sont A, A' et P dans cette formule. Un autre moyen mnémotechnique : souvent, la matrice M est celle d'un endomorphisme f donné dans l'énoncé, et on cherche à la *diagonaliser*, c'est-à-dire à exhiber une base $\mathcal{B}'$ de E telle que la matrice de f dans la base $\mathcal{B}'$ soit diagonale. Lorsque cela est possible, on obtient alors la formule $D = P^{-1}MP$, que l'on peut réécrire $MP = PD$; on trouve ainsi la formule d'algèbre linéaire préférée des PSI. Bref, tous les moyens sont bons, l'important étant que vous sachiez retrouver ce résultat en strictement moins de deux essais.

THÉORÈME 24.7 - (Formule du changement de base, généralisation). Soient E et F deux $\mathbb{K}$ -ev de dimension finie, $\mathcal{B}$ et $\mathcal{B}'$ deux bases de E , $\mathcal{C}$ et $\mathcal{C}'$ deux bases de F , et $f \in \mathcal{L}(E, F)$. Alors :

$$\text{Mat}_{\mathcal{B}'\mathcal{C}'}(f) = P_{\mathcal{C}',\mathcal{C}} \text{Mat}_{\mathcal{B},\mathcal{C}}(f) P_{\mathcal{B},\mathcal{B}'}$$

PREUVE. Analogue à celle du théorème 24.6. ☕

On peut alors observer que les notions de rang (d'une matrice d'une part, d'une application linéaire d'autre part) coïncident et que le rang de f est celui d'une quelconque de ses matrices (indépendance du choix des bases). Explicitement :

PROPRIÉTÉ 24.14 - Mêmes hypothèses et notations que dans le théorème ci-dessus. On a :

$$\text{rg}(\text{Mat}_{\mathcal{B}'\mathcal{C}'}(f)) = \text{rg}(\text{Mat}_{\mathcal{B},\mathcal{C}}(f))$$

PREUVE. La preuve de cette propriété repose essentiellement sur le théorème précédent et sur le fait que le rang d'une application linéaire est invariant par composition (à gauche ou à droite) avec un isomorphisme (voir propriété 24.2 page 638). ☕

PREUVE. Soit $A \in M_{np}(\mathbb{K})$. Montrons l'équivalence $[\text{rg}(A) = r] \iff [A \equiv J_r]$ par double implication.

Supposons que A soit de rang r . Il revient au même de dire que l'application linéaire canoniquement associée $f_A \in \mathcal{L}(\mathbb{K}^p, \mathbb{K}^n)$ est de rang r . Le noyau de f_A est donc de dimension $p - r$ (d'après le théorème du rang), et admet donc un supplémentaire dans E , que l'on note S et qui est de dimension r (d'après le théorème assurant l'existence d'un supplémentaire).

On construit alors une base $\mathcal{B} = (u_1, \dots, u_r, v_1, v_{p-r})$ de $\mathbb{K}^p$, de telle sorte que $(u_1, \dots, u_r)$ est une base de S , et $(v_1, \dots, v_{p-r})$ est une base de $\ker f_A$.

Puisque $f|_S$ est injective (cf preuve du théorème du rang), la famille $(f(u_1), \dots, f(u_r))$ est libre. A son tour, elle peut être complétée en une base $\mathcal{B}' = (f(u_1), \dots, f(u_r), w_1, \dots, w_{n-r})$ de $\mathbb{K}^n$ (d'après le théorème de la base incomplète).

On a alors par construction : $\text{Mat}_{\mathcal{B}, \mathcal{B}'}(f_A) = J_r$.

Par ailleurs, en notant P (resp. P') la matrice de passage de la base canonique à la base $\mathcal{B}$ de $\mathbb{K}^p$ (resp. à la base $\mathcal{B}'$ de $\mathbb{K}^n$), la formule du changement de base assure que :

$$J_r = \text{Mat}_{\mathcal{B}, \mathcal{B}'}(f_A) = P'AP$$

Puisque P et P' sont inversibles (car toute matrice de passage l'est), cette dernière égalité signifie que A et J_r sont équivalentes. Ce qui prouve la première implication. La réciproque est triviale.

La seconde équivalence est une conséquence immédiate de la première et du théorème du rang. ☕

COROLLAIRE 24.6 - $\forall A \in M_{np}(\mathbb{K}) \quad \text{rg}({}^tA) = \text{rg}(A)$

PREUVE. Supposons que $A \in M_{np}(\mathbb{K})$ soit de rang r . Alors il existe deux matrices $P \in GL_n(\mathbb{K})$ et $Q \in GL_p(\mathbb{K})$ telles que : $PAQ = J_r$. On en déduit : ${}^tQ {}^tA {}^tP = {}^tJ_r$. Or ${}^tQ$ et ${}^tP$ sont inversibles, et la matrice ${}^tJ_r$ est la matrice " J_r " de $M_{pn}(\mathbb{K})$.

Il s'ensuit que ${}^tA$ est équivalente à J_r , et que $\text{rg}({}^tA) = r$ d'après le théorème précédent. On a ainsi établi que si A est de rang r , alors sa transposée est de rang r .

L'implication réciproque provient de la précédente appliquée à la transposée ${}^tA$ (en lieu et place de A), puisque ${}^t({}^tA) = A$. ☕

24.8.2 Matrices semblables

Nous rappelons ici la définition de la relation de similitude sur les matrices, déjà rencontrée dans le chapitre portant sur le calcul matriciel.

DÉFINITION 24.8 - Deux matrices A et B de $M_n(\mathbb{K})$ sont **semblables** s'il existe une matrice $P \in GL_n(\mathbb{K})$ telle que : $B = P^{-1}AP$.

Notation. Dans ces lignes, nous noterons $A \sim B$ lorsque A est semblable à la matrice B .

Remarque. Une différence de taille par rapport au paragraphe précédent. Si deux matrices peuvent être équivalentes sans être carrées, on ne pourra en revanche parler de matrices semblables que si l'on considère deux matrices carrées et de même dimension.

PROPRIÉTÉ 24.16 - La relation de similitude est une relation d'équivalence sur $M_n(\mathbb{K})$, càd :

1/ **Réflexivité** : $\forall A \in M_n(\mathbb{K}), A \simeq A$

2/ **Symétrie** : $\forall (A, B) \in M_n(\mathbb{K})^2, [A \simeq B] \iff [B \simeq A]$

3/ **Transitivité** : $\forall (A, B, C) \in M_n(\mathbb{K})^3, [(A \simeq B) \wedge (B \simeq C)] \implies [A \simeq C]$

PREUVE. Voir propriété 13.22, page 335 ☕

Désormais, nous pourrons donc dire que deux matrices A et B de $M_n(\mathbb{K})$ sont semblables (l'ordre n'ayant pas d'importance).

Exemple. Soit f un endomorphisme d'un espace vectoriel E de dimension finie, et soient deux bases $\mathcal{B}$ et $\mathcal{B}'$ deux bases de E . Alors : $\text{Mat}_{\mathcal{B}'}(f) \sim \text{Mat}_{\mathcal{B}}(f)$ puisqu'il existe une matrice P telle que $\text{Mat}_{\mathcal{B}}(f) = P \text{Mat}_{\mathcal{B}'}(f) P^{-1}$ (en l'occurrence $P = P_{\mathcal{B}\mathcal{B}'}$). Réciproquement, deux matrices de $M_n(\mathbb{K})$ sont semblables si elles représentent le même endomorphisme de $\mathbb{K}^n$ (puisque toute matrice P inversible peut être interprétée comme une matrice de passage).

PROPRIÉTÉ 24.17 - Deux matrices semblables sont équivalentes.

$$\forall (A, B) \in M_n(\mathbb{K})^2, [A \sim B] \implies [A \equiv B]$$

PREUVE. Immédiate. ☕

Attention : réciproque fausse ! Les matrices $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ et $B = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ sont équivalentes (il suffit d'observer qu'elles sont de même rang), mais ne sont pas semblables (elles n'ont pas même trace).

PROPRIÉTÉ 24.18 - (Invariants de similitude)

1/ $\forall (A, B) \in M_n(\mathbb{K})^2, [A \sim B] \implies [\text{rg}(A) = \text{rg}(B)]$. Deux matrices semblables ont même rang, ou encore le rang est un **invariant de similitude**.

2/ $\forall (A, B) \in M_n(\mathbb{K})^2, [A \sim B] \implies [\text{tr}(A) = \text{tr}(B)]$. Deux matrices semblables ont même trace, ou encore la trace est un **invariant de similitude**.

3/ $\forall (A, B) \in M_n(\mathbb{K})^2, [A \sim B] \implies [\det(A) = \det(B)]$. Deux matrices semblables ont même déterminant, ou encore le déterminant est un **invariant de similitude**.

PREUVE. 1/ Si A et B sont semblables, alors il existe une matrice inversible P telle que : $B = P^{-1}AP$. Puisque le rang est invariant par composition à gauche ou à droite avec un isomorphisme, on en déduit que $\text{rg}(A) = \text{rg}(B)$. Le rang est donc un invariant de similitude.

2/ Voir preuve de la propriété 13.23, page 336.

3/ Voir chapitre "Déterminants". ☕

24.8.3 Trace d'un endomorphisme

DÉFINITION 24.9 - Soit E un $\mathbb{K}$ -ev de dimension finie, et soit $f \in \mathcal{L}(E)$.

On appelle **trace de l'endomorphisme** f et on note $\text{tr}(f)$ la trace de $\text{Mat}_{\mathcal{B}}(f)$ où $\mathcal{B}$ désigne une quelconque base de E .

Remarque. Il résulte du paragraphe précédent que cette définition est sans ambiguïté, puisque la trace est un invariant de similitude.

PROPRIÉTÉ 24.19 - Soit E un $\mathbb{K}$ -ev de dimension finie. On a :

$$1/ \forall (f, g) \in (\mathcal{L}(E))^2, \forall (\lambda, \mu) \in \mathbb{K}^2, \text{tr}(\lambda f + \mu g) = \lambda \text{tr}(f) + \mu \text{tr}(g) \quad (\text{linéarité de la trace})$$

$$2/ \forall (f, g) \in (\mathcal{L}(E))^2, \text{tr}(g \circ f) = \text{tr}(f \circ g)$$

PREUVE. Conséquence des propriétés déjà établies dans le cadre des traces des matrices. ☕

Un cas particulier important :

PROPRIÉTÉ 24.20 - (**Trace d'un projecteur**) Soit $p \in \mathcal{L}(E)$ un $\mathbb{K}$ -ev de dimension finie, tel que $p^2 = p$.

On a : $\text{tr}(p) = \text{rg}(p)$

PREUVE. Soient E un $\mathbb{K}$ -ev de dimension finie n , et $p \in \mathcal{L}(E)$ un projecteur. On peut affirmer que $E = \ker(p) \oplus \text{Im}(p)$ (cf propriétés des projecteurs justement !). On choisit alors une base $\mathcal{B}' = (v_1, \dots, v_r)$ de $\text{Im}(p)$ (l'entier r désignant le rang de p) et une base $\mathcal{B}'' = (w_1, \dots, w_{n-r})$ de $\ker(p)$, de telle sorte que $\mathcal{B} = \mathcal{B}' \cup \mathcal{B}''$ est une base de E .

Par construction des vecteurs v_j , on a : $\forall j \in \llbracket 1, r \rrbracket, p(v_j) = v_j$, car $v_j \in \text{Im}(p)$, et $\text{Im}(p) = \ker(\text{id}_E - p)$ (re-cf propriétés des projecteurs).

Et par construction des vecteurs w_j , on a : $\forall j \in \llbracket 1, n-r \rrbracket, p(w_j) = 0$, car $w_j \in \ker(p)$.

Il s'ensuit que la matrice de p dans la base $\mathcal{B}$ est : $M_{\mathcal{B}}(p) = J_r =$

$$\begin{pmatrix} 1 & 0 & \dots & \dots & \dots & 0 & \dots & 0 \\ 0 & 1 & \dots & \dots & \vdots & \vdots & \dots & \vdots \\ \vdots & & \ddots & & \vdots & \vdots & \dots & \vdots \\ \vdots & & & \ddots & \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & \dots & 1 & 0 & \dots & 0 \\ 0 & \dots & \dots & \dots & 0 & 0 & \dots & 0 \\ \vdots & & & & \vdots & \vdots & & \vdots \\ 0 & \dots & \dots & \dots & 0 & 0 & \dots & 0 \end{pmatrix}$$

On en déduit que $\text{tr}(M_{\mathcal{B}}(p)) = \text{tr}(J_r) = r$. Par suite : $\text{tr}(p) = r$. Ainsi : $\text{tr}(p) = \text{rg}(p)$. ☕

24.9 Calcul pratique du rang d'une matrice

Le but de ce paragraphe est de formaliser matriciellement les techniques mises en place plus tôt dans l'année notamment pour la méthode du pivot de Gauss. Explicitement, nous appellerons **opérations élémentaires sur les lignes** d'une matrice $A \in M_{np}(\mathbb{K})$:

- 1/ les **permutations** de deux lignes L_i et L_j ;
- 2/ les **dilatations** de rapport λ , qui consistent à remplacer une ligne L_i par la ligne λL_i , avec $\lambda \in \mathbb{K}^*$
- 3/ les **transvections**, qui consistent à remplacer une ligne L_j par $L_j - \lambda L_i$, avec $\lambda \in \mathbb{K}^*$.

Ces opérations élémentaires sur les lignes de $A \in M_n(\mathbb{K})$ peuvent être interprétées matriciellement. Plus précisément :

- 1/ on appelle **matrice de permutation** la matrice $P_{i,j} = I_n - E_{ii} - E_{jj} + E_{ij} + E_{ji} \in M_n(\mathbb{K})$ (c'est la matrice identité dont on a permuté la i -ème et la j -ème lignes). La matrice $P_{i,j}A$ est la matrice A dont on a permuté les lignes i et j ;
- 2/ on appelle **matrice de transvection** la matrice $T_{i,j,\lambda} = I_n - \lambda E_{ji} \in M_n(\mathbb{K})$ (c'est la matrice identité dont on a permuté la i -ème et la j -ème lignes). La matrice $T_{i,j,\lambda}A$ est la matrice A dans laquelle la ligne L_j a été remplacée par $L_j - \lambda L_i$;
- 3/ Enfin, notons que la multiplication de A à gauche par $H_{i,\lambda} = I_n + (\lambda - 1) E_{ii}$ permet de remplacer dans A la ligne L_i par λL_i .

Notons que :

$$\forall (i, j) \in \llbracket 1, n \rrbracket^2, \forall \lambda \in \mathbb{K}^*, P_{i,j} \in GL_n(\mathbb{K}), T_{i,j,\lambda} \in GL_n(\mathbb{K}) \text{ et } H_{i,\lambda} \in GL_n(\mathbb{K})$$

Le point-clef est qu'il existe alors une suite d'opérations élémentaires sur les lignes de A permettant de faire apparaître des zéros dans la première colonne de A (en pratique, on applique l'algorithme du pivot de Gauss). Dans le détail, il existe un couple d'entiers $(i, j) \in \llbracket 1, n \rrbracket^2$, et $(n-1)$ scalaires $\lambda_2, \dots, \lambda_n$ tels que :

$$T_{1,n,\lambda_n} \cdots T_{1,2,\lambda_2} P_{ij} A = \left(\begin{array}{c|ccc} p_1 & * & \cdots & * \\ \hline 0 & & & \\ \vdots & & A' & \\ 0 & & & \end{array} \right)$$

Deux cas peuvent alors se présenter :

- si $p_1 \neq 0$: alors $\text{rg}(A) = 1 + \text{rg}(A')$;
- si $p_1 = 0$: alors $\text{rg}(A) = \text{rg}(A')$.

La justification de ces égalités provient essentiellement du fait que $\text{rg}(T_{1,n,\lambda_n} \cdots T_{1,2,\lambda_2} P_{ij} A) = \text{rg}(A)$, puisque les matrices de transvection et de permutation sont inversibles, et que le rang d'une matrice est invariant par multiplication (à gauche ou à droite) par une matrice inversible.

Dans les deux cas, on est ramené au calcul d'une matrice de $M_{n-1,p-1}(\mathbb{K})$. En répétant le procédé autant de fois que nécessaire, on peut déterminer facilement le rang de A .

Exemple d'application 1 : rang de $\begin{pmatrix} 1 & 1 & \alpha \\ 1 & \alpha & 1 \\ \alpha & 1 & 1 \end{pmatrix}$ en fonction de $\alpha \in \mathbb{K}$:

$$\begin{aligned} \text{rg} \begin{pmatrix} 1 & 1 & \alpha \\ 1 & \alpha & 1 \\ \alpha & 1 & 1 \end{pmatrix} &= \text{rg} \begin{pmatrix} 1 & 1 & \alpha \\ 0 & \alpha - 1 & 1 - \alpha \\ 0 & 1 - \alpha & 1 - \alpha^2 \end{pmatrix} = 1 + \text{rg} \begin{pmatrix} \alpha - 1 & 1 - \alpha \\ 1 - \alpha & 1 - \alpha^2 \end{pmatrix} = 1 + \text{rg} \begin{pmatrix} \alpha - 1 & 1 - \alpha \\ 0 & 2 - \alpha - \alpha^2 \end{pmatrix} = \\ &= 1 + \begin{cases} 0 & \text{si } \alpha = 1 \\ 1 & \text{si } \alpha = 2 \\ 2 & \text{si } \alpha \in \mathbb{K} \setminus \{1, 2\} \end{cases} \end{aligned}$$

Exemple d'application 2 : calcul du rang de $\begin{pmatrix} 1 & 1 & 1 \\ b+c & a+c & a+b \\ bc & ac & ab \end{pmatrix}$ en fonction de a, b et $c \in \mathbb{K}$ (voir cours).

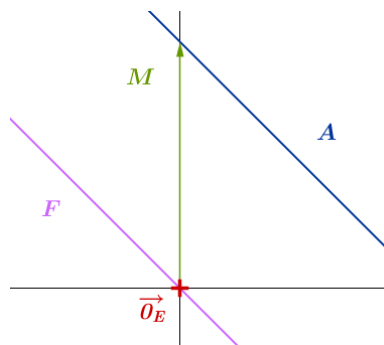
Remarque : on peut définir de manière analogue à ce qui précède les opérations élémentaires sur les colonnes de A , et définir les matrices de permutation et de transvection correspondantes. Mais ce n'est pas indispensable, puisque les colonnes de A sont les lignes de sa transposée ${}^t A$, et qu'une matrice et sa transposée ont même rang.

24.10 Sous-espaces affines d'un espace vectoriel

DÉFINITION 24.10 - Soit E un $\mathbb{K}$ -ev. Une partie A de E est un **sous-espace affine** de E s'il existe un élément M de E et un sev F de E tel que :

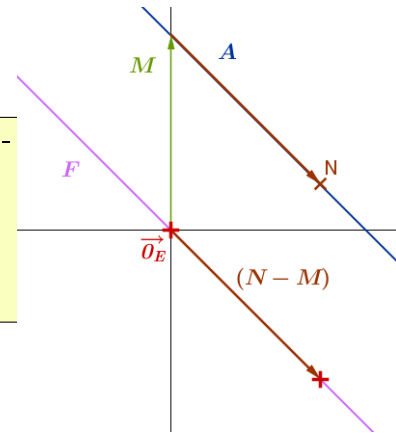
$$A = M + F = \{m + \vec{f} / \vec{f} \in F\}$$

Exemple et illustration. $A = \{(x, y) \in \mathbb{R}^2 / y = 1 - x\}$ est un sous-espace affine de $\mathbb{R}^2$, avec $M = (0, 1)$ et $F = \text{Vect}(1, -1)$.



PROPRIÉTÉ 24.21 - Soient E un $\mathbb{K}$ -ev, et $A = M + F$ un sous-espace affine de E . Soit $N \in E$. Alors :

$$[N \in A] \iff [(N - M) \in F]$$



PREUVE. Avec les notations de la propriété :

$$[N \in A] \iff [\exists f \in F, N = M + f] \iff [\exists f \in F, N - M = f] \iff [(N - M) \in F] \quad \text{☕}$$

Conséquence : si A est un sous-espace affine, alors le point M de la définition n'est pas unique, dans un sens rendu plus précis par l'énoncé suivant.

PROPRIÉTÉ 24.22 - Soient E un $\mathbb{K}$ -ev, et $A = M + F$ un sous-espace affine de E .

Alors : $\forall N \in A, A = N + F$

PREUVE. Soit $N \in A$. Alors $M - N \in F$ d'après la propriété précédente.

Soit alors $P \in M + F$; il existe un élément f de F tel que : $P = M + f$. D'où $P = N + \underbrace{(M - N) + f}_{\in F}$.

Par suite : $P \in N + F$. Ce qui prouve l'inclusion : $M + F \subset N + F$.

Réciproquement, si $P \in N + F$; il existe un élément f de F tel que : $P = N + f$. D'où $P = M + \underbrace{(N - M) + f}_{\in F}$. Par suite : $P \in M + F$. Ce qui prouve l'inclusion : $M + F \supset N + F$.

Finalement : $M + F = N + F$, cette égalité étant valide pour tout point N de A . ☕

En revanche, le sev F de la définition est défini de manière unique.

PROPRIÉTÉ 24.23 - Soient E un $\mathbb{K}$ -ev, et A un sous-espace affine de E .

S'il existe deux éléments M et M' de E , et F et F' deux sev de E tels que : $A = M + F = M' + F'$, alors $F = F'$.

PREUVE. Supposons qu'il existe deux éléments M et M' de E , et F et F' deux sev de E tels que : $A = M + F = M' + F'$, alors $F = F'$.

Il résulte de ces hypothèses que M et M' appartiennent à A . On en déduit que : $M - M' \in F$ (d'après la propriété 24.21), et que : $M' - M \in F'$ (même propriété). Ainsi : $M - M' \in F \cap F'$.

Soit à présent f un élément de F . On a : $M + f \in A$, donc : $M + f = M' + M - M' + f \in A$; donc $M - M' + f \in F'$ (puisque $A = M' + F'$).

On déduit de cette égalité et du fait que $M - M'$ appartient à F' que : $f \in F'$.

Ce qui établit l'inclusion $F \subset F'$.

L'autre inclusion est analogue. On peut donc conclure que : $F = F'$. ☕

La propriété précédente permet d'introduire la notion de *direction* d'un sous-espace affine.

DÉFINITION 24.11 - Soient E un $\mathbb{K}$ -ev, et A un sous-espace affine de E .

On appelle **direction** de A l'unique sev F de E tel que : $A = M + F$ où M désigne un quelconque élément de A .

Exemple : le sous-espace affine de $A = \{(x, y) \in \mathbb{R}^2 / y = 1 - x\}$ de $\mathbb{R}^2$ a pour direction le sev $F = \text{Vect}(1, -1)$.

DÉFINITION 24.12 - Soient E un $\mathbb{K}$ -ev, et A un sous-espace affine de E .

On appelle **dimension de** A la dimension de sa direction.

En particulier, on dit que A est une **droite** (*resp.* un **plan**, *resp.* un **hyperplan**) **affine** si la direction de A est de dimension 1 (*resp.* 2, *resp.* $\dim(E) - 1$).

Exemple 1. $A = \{(x, y) \in \mathbb{R}^2 / y = 1 - x\}$ est une droite affine dans $\mathbb{R}^2$.

Exemple 2. Notons $E = \mathcal{C}^\infty(\mathbb{R}, \mathbb{R})$. L'ensemble $A = \{y \in E / y'' - 5y' + 6y = e^t\}$ est un plan affine dans E , de direction $F = \{y \in E / y'' - 5y' + 6y = 0\}$, c'est à dire : $F = \text{Vect}(f, g)$ où $f : t \in \mathbb{R} \mapsto e^{2t}$ et $g : t \in \mathbb{R} \mapsto e^{3t}$.

En effet, tout élément φ de A s'écrit $\underbrace{f_p}_{= "M''"} + \underbrace{\lambda f + \mu g}_{\in "F''"}$ avec λ et μ réels, et où $f_p : t \in \mathbb{R} \mapsto e^t/2$ désigne une solution particulière de l'EDL $y'' - 5y' + 6y = e^t$.

Exemple 3. Soit $n \in \mathbb{N}^*$. L'ensemble $A = \{M \in M_n(\mathbb{K}) / \text{tr}(M) = 1\}$ est un hyperplan affine de $M_n(\mathbb{K})$, de direction le sev des matrices de $M_n(\mathbb{K})$ ayant une trace nulle. Dans cette situation, on a en effet : $A = E_{11} + \ker(\text{tr})$, où tr désigne la forme linéaire qui à une matrice de $M_n(\mathbb{K})$ associe sa trace.

PROPRIÉTÉ 24.24 - (Intersection de 2 sea). Soient E un $\mathbb{K}$ -ev, A et A' deux sous-espaces affines de E . On suppose qu'il existe deux éléments M et M' de E , et F et F' deux sev de E tels que : $A = M + F$ et $A' = M' + F'$.

L'intersection $A \cap A'$ est soit l'ensemble vide, soit un sous-espace affine de direction $F \cap F'$.

PREUVE. Supposons que $A \cap A'$ soit non vide. Il existe alors $P \in A \cap A'$. Dans ce contexte, établissons que $A \cap A' = P + (F \cap F')$.

Soit $M \in A \cap A'$. Alors $M - P \in F$ et $M - P \in F'$ (d'après la propriété 24.21). Il s'ensuit que : $M - P \in F \cap F'$. Or : $M = P + (M - P)$. On en déduit que : $M \in P + (F \cap F')$.

Ce qui prouve l'inclusion : $A \cap A' \subset P + (F \cap F')$.

L'inclusion réciproque est immédiate. ☕

24.11 Equations linéaires

DÉFINITION 24.13 - On appelle **équation linéaire** une équation de la forme $f(\vec{v}) = \vec{b}$ où $f \in \mathcal{L}(E, F)$ (avec E et F deux $\mathbb{K}$ -ev), $\vec{v} \in E$ et $\vec{b} \in F$.

Terminologie. Notons (E) l'équation $f(\vec{v}) = \vec{b}$. L'équation $f(\vec{v}) = \vec{0}_F$ est appelée **équation homogène associée à (E)**; et $\vec{b}$ est appelé le **second membre** de (E).

Exemples.

- Le système $\begin{cases} ax + by + cz = b_1 \\ dx + ey + fz = b_2 \end{cases}$ d'inconnue $(x, y, z) \in \mathbb{R}^3$ est une équation linéaire.
- L'équation différentielle $a(t)y' + b(t)y = c(t)$ d'inconnue $y \in \mathcal{C}^\infty(\mathbb{R}, \mathbb{R})$ est une équation linéaire.
- L'équation différentielle $ay'' + by'(t) + cy(t) = d(t)$ d'inconnue $y \in \mathcal{C}^\infty(\mathbb{R}, \mathbb{R})$ est une équation linéaire.
- L'équation $P - P' + P'' = X^n$ d'inconnue $P \in \mathbb{K}_n[X]$ est une équation linéaire.
- L'équation $\text{tr}(M) = 1$ d'inconnue $M \in M_n(\mathbb{K})$ est une équation linéaire.

THÉORÈME 24.9 - Notons (E) l'équation linéaire $f(\vec{v}) = \vec{b}$ (mêmes hypothèses et notation que dans la définition ci-dessus), et notons S l'ensemble des solutions de (E).

Si S n'est pas vide, alors S est un sous-espace affine de E de direction $\ker f$.

Plus précisément, si $S \neq \emptyset$, alors : $\exists \vec{v}_p \in E, S = \vec{v}_p + \ker f$.

PREUVE. Supposons que $S \neq \emptyset$. Alors il existe $\vec{v}_p \in E, f(\vec{v}_p) = \vec{b}$. Établissons l'égalité $S = \vec{v}_p + \ker f$.

$$[\vec{w} \in S] \iff [f(\vec{w}) = f(\vec{v}_p)] \iff [\vec{w} - \vec{v}_p \in \ker f] \iff [\exists \vec{u} \in \ker f, \vec{w} - \vec{v}_p = \vec{u}]$$

$$\iff [\exists \vec{u} \in \ker f, \vec{w} = \vec{v}_p + \vec{u}] \iff [\vec{w} \in \vec{v}_p + \ker f]$$

La comparaison des assertions situées aux deux extrémités de cette chaîne d'équivalences donne l'égalité : $S = \vec{v}_p + \ker f$.

Conclusion. L'ensemble des solutions de l'équation linéaire $f(\vec{v}) = \vec{b}$, s'il n'est pas vide, est un sous-espace affine de E . ☕

Applications immédiates. L'ensemble des solutions du système $\begin{cases} ax + by + cz = b_1 \\ dx + ey + fz = b_2 \end{cases}$ est un sous-espace affine de $\mathbb{R}^3$; l'ensemble des solutions de l'équation différentielle $a(t)y' + b(t)y = c(t)$ (ou de $ay'' + by'(t) + cy(t) = d(t)$) est un sous-espace affine de $\mathcal{C}^\infty(\mathbb{R}, \mathbb{R})$; l'ensemble des solutions de l'équation $P - P' + P'' = X^n$ est un sous-espace affine de $\mathbb{K}_n[X]$; l'ensemble des solutions de l'équation $\text{tr}(M) = 2016$ d'inconnue $M \in M_n(\mathbb{K})$ est un sous-espace affine de $M_n(\mathbb{K})$; l'ensemble des matrices de $M_n(\mathbb{K})$ dont la somme des coefficients de chaque colonne est égale à 2 est un autre sous-espace affine de $M_n(\mathbb{K})$; etc...

Application détaillée. Soient $\lambda \in \mathbb{K}$, et $n \in \mathbb{N}^*$. L'ensemble des matrices de trace égale à λ est un hyperplan affine de $M_n(\mathbb{K})$.

En effet, on a : $\forall (A, B) \in M_n(\mathbb{K})^2, \forall (\lambda, \mu) \in \mathbb{K}^2, \text{tr}(\lambda A + \mu B) = \lambda \text{tr}(A) + \mu \text{tr}(B)$. Ce qui signifie encore que la trace donne lieu à une forme linéaire sur $M_n(\mathbb{K})$; puisqu'elle est non nulle, son noyau (qui est le sev de $M_n(\mathbb{K})$ constitué des matrices de trace nulle) est un hyperplan de $M_n(\mathbb{K})$.

Cette remarque faite l'ensemble L des matrices de trace égale à λ est l'ensemble des solutions de l'équation $\text{tr}(A) = \lambda$. Puisque d'une part cette équation est linéaire, et que d'autre part elle admet une solution (par exemple $A_p = \lambda E_{11}$), le théorème précédent permet d'affirmer que L est un sous-espace affine de $M_n(\mathbb{K})$, de direction le sev F des matrices de $M_n(\mathbb{K})$ ayant une trace nulle. Plus précisément : $L = A_p + F$.

Un exemple d'application un peu moins immédiate. Soient n et p deux entiers naturels non nuls, et soient $\alpha_0, \dots, \alpha_n, b_0, \dots, b_n$ ($2n+2$) scalaires. On cherche à déterminer les polynômes P de degré au plus p tels que

$$\forall i \in \llbracket 0, n \rrbracket, P(\alpha_i) = b_i \quad ((\spadesuit))$$

A cette fin, on définit les polynômes interpolateurs de Lagrange $(L_k)_{k \in \llbracket 0, n \rrbracket}$ associés aux scalaires α_k en posant :

$$\forall k \in \llbracket 0, n \rrbracket, L_k = \prod_{i=0, i \neq k}^n \frac{X - \alpha_i}{\alpha_k - \alpha_i}$$

Puis on introduit le polynôme : $Q_p = \sum_{k=0}^n b_k L_k \in \mathbb{K}_n[X]$.

Q_p vérifie la propriété : $\forall i \in \llbracket 0, n \rrbracket, Q_p(\alpha_i) = b_i$.

Enfin, on note :

$$\begin{aligned} \varphi_p : \mathbb{K}_p[X] &\longrightarrow \mathbb{K}^{n+1} \\ P &\longmapsto (P(\alpha_0), P(\alpha_1), \dots, P(\alpha_n)) \end{aligned}$$

Plusieurs cas se présentent alors :

- si $p > n$: alors l'ensemble des polynômes solution de $(\spadesuit)$ est un sous-espace affine de $\mathbb{K}_p[X]$, contenant Q_p et de direction $\ker \varphi_p$. C'est un sous-espace affine de dimension $p - n$.
- si $p = n$: alors Q_p est l'unique polynôme solution de $(\spadesuit)$. Dans ce cas particulier, l'ensemble des solutions est un sous-espace affine de $\mathbb{K}_p[X]$ de dimension 0.
- si $p < n$: alors soit on a de la chance, soit on n'en a pas ! Plus sérieusement, soit l'ensemble des solutions de $(\spadesuit)$ est vide, soit il contient un unique élément (qui n'est pas Q_p , puisque $Q_p \notin \mathbb{K}_p[X]$).

Interprétation géométrique. Par $n+1$ points du plan d'abscisses distinctes passe une unique courbe représentative d'un polynôme de degré au plus n . En particulier, par 3 points non alignés du plan passe une unique parabole.

24.12 Epilogue

24.12.1 Suite des noyaux itérés

Exercice : (Suites des noyaux itérés et des images itérées)

Soit E un $\mathbb{K}$ -espace vectoriel de dimension finie, et soit f un endomorphisme de E .

Pour tout entier naturel k , on pose : $I_k = \text{Im } f^k$ et $N_k = \ker f^k$.

- 1) Montrer que la suite $(I_k)_k$ est décroissante, et que la suite $(N_k)_k$ est croissante.
- 2) Montrer qu'il existe un entier $s \in \mathbb{N}$ tel que $I_{s+1} = I_s$ et $N_{s+1} = N_s$.
- 3) On note s_0 le plus petit des entiers s vérifiant la propriété ci-dessus. On admet que : $\forall s \geq s_0, I_s = I_{s_0}$ et $N_s = N_{s_0}$. Montrer que $E = I_{s_0} \oplus N_{s_0}$.

Soient E un $\mathbb{K}$ -espace vectoriel de dimension finie, et $f \in \mathcal{L}(E)$.

- 1) Décroissance de la suite des images itérées $(I_k)_k$. Soit k un entier naturel quelconque. Si $\vec{v} \in I_{k+1}$, alors il existe un vecteur $\vec{w} \in E$ tel que : $\vec{v} = f^{k+1}(\vec{w})$. D'où : $\vec{v} = f^k(f(\vec{w}))$, et $\vec{v} \in \text{Im } f^k = I_k$.

Conclusion. La suite des images itérées d'un endomorphisme est décroissante : $\forall k \in \mathbb{N}, \text{Im } f^{k+1} \subset \text{Im } f^k$.

- Croissance de la suite des noyaux itérés $(N_k)_k$. Soit k un entier naturel quelconque. Si $\vec{v} \in N_k$, alors $f^k(\vec{v}) = \vec{0}_E$. Puisque f est linéaire, on a : $f(f^k(\vec{v})) = f^{k+1}(\vec{v}) = \vec{0}_E$, d'où : $\vec{v} \in \ker f^{k+1} = N_{k+1}$.

Conclusion. La suite des noyaux itérés d'un endomorphisme est croissante : $\forall k \in \mathbb{N}, \ker f^k \subset \ker f^{k+1}$.

- 2) Pour tout entier k , posons : $d_k = \dim N_k$. D'après la question précédente, la suite (d_k) est croissante. De plus, puisque E est de dimension finie, la suite (d_k) est majorée (par $\dim E$). Croissante et majorée, la suite (d_k) converge donc. Or une suite convergente d'entiers naturels est stationnaire². Il existe donc un entier s tel que : $\forall n \geq s, d_n = d_s$.

En particulier : $d_{s+1} = d_s$. Ainsi N_s est un sev de N_{s+1} ; mais puisque ces deux sev sont de même dimension, on en déduit que $N_s = N_{s+1}$.

Conclusion. $\exists s \in \mathbb{N}, \underbrace{\ker f^s}_{N_s} = \underbrace{\ker f^{s+1}}_{N_{s+1}}$.

Une double application du théorème du rang (aux endomorphismes f^s et f^{s+1}) donne alors : $\dim \text{Im } f^s = \dim \text{Im } f^{s+1}$. Puisque $\text{Im } f^{s+1}$ est un sev de $\text{Im } f^s$, et que tous deux sont de même dimension, on en déduit que : $\text{Im } f^s = \text{Im } f^{s+1}$.

Conclusion. $\exists s \in \mathbb{N}, \ker f^s = \ker f^{s+1}$ et $\text{Im } f^s = \text{Im } f^{s+1}$.

2. Preuve de ce point, qui pourra être admis en colle : soit (u_n) une suite d'entiers naturel convergente. Notons ℓ sa limite. Commençons par observer que $\ell \in \mathbb{N}$. En effet, si ℓ n'était pas dans $\mathbb{N}$, alors il existerait un rang n_0 tel que : $[n \geq n_0] \implies [|u_n - \ell| < \varepsilon = \min(\ell - [\ell], [\ell] + 1 - \ell)]$. Ce qui est absurde puisque l'intervalle $] \ell - \varepsilon, \ell + \varepsilon [$ ne contient aucun entier.

Donc $\ell \in \mathbb{N}$. En utilisant une nouvelle fois la définition de limite, on peut affirmer que : $\exists n_1 \in \mathbb{N}, \forall n \in \mathbb{N}, [n \geq n_1] \implies [|u_n - \ell| < 1/2]$. Or l'intervalle $] \ell - 1/2, \ell + 1/2 [$ ne contient que l'entier ℓ : la condition $|u_n - \ell| < 1/2$ est alors équivalente à $u_n = \ell$. En résumé, on a établi que : $\exists n_1 \in \mathbb{N}, \forall n \in \mathbb{N}, [n \geq n_1] \implies [u_n = \ell]$; ce qui signifie que la suite (u_n) est stationnaire.

3) Dans la présente situation, $\ker f^{s_0}$ et $\operatorname{Im} f^{s_0}$ sont deux sev de E (puisque $f^{s_0} \in \mathcal{L}(E)$).

Montrons que $\ker f^{s_0} \cap \operatorname{Im} f^{s_0} = \{\vec{0}_E\}$. Soit $\vec{v} \in (\ker f^{s_0} \cap \operatorname{Im} f^{s_0})$. Alors $\vec{v}$ appartient en particulier à $\operatorname{Im} f^{s_0}$, donc : $\exists \vec{w} \in E, \vec{v} = f^{s_0}(\vec{w})$.

Puisque par ailleurs $\vec{v} \in \ker f^{s_0}$, on a : $f^{2s_0}(\vec{w}) = \vec{0}_E$, donc $\vec{w} \in \ker f^{2s_0}$. Or par hypothèse $\ker f^{2s_0} = \ker f^{s_0}$, donc $\vec{w} \in \ker f^{s_0}$, d'où : $\vec{v} = f^{s_0}(\vec{w}) = \vec{0}_E$. Ainsi : $\ker f^{s_0} \cap \operatorname{Im} f^{s_0} = \{\vec{0}_E\}$.

Comme $\ker f^{s_0} \cap \operatorname{Im} f^{s_0} = \{\vec{0}_E\}$ et que $\dim \ker f^{s_0} + \dim \operatorname{Im} f^{s_0} = E$ (théorème du rang), on a :

$$E = \ker f^{s_0} \oplus \operatorname{Im} f^{s_0}.$$

Remarque. Si l'on ne veut pas admettre ce qui l'est dans l'énoncé...

Pour tout entier naturel k , on pose : $I_k = \operatorname{Im} f^k$ et $N_k = \ker f^k$.

On note s_0 le plus petit des entiers s vérifiant la propriété ci-dessus.

Etablir que : $\forall s \geq s_0, I_s = I_{s_0}$ et $N_s = N_{s_0}$.

Prouvons par récurrence sur l'entier naturel k la propriété $\mathcal{P}(k)$: " $I_{s_0+k} = I_{s_0}$ ".

L'initialisation provient de la définition de l'entier s_0 .

Etablissons l'hérédité : supposons que $\mathcal{P}(k)$ est vraie pour un certain entier naturel non nul k .

On a $I_{s_0+k+1} \subset I_{s_0}$ (on a déjà établi que la suite des images itérées est décroissante).

Réciproquement : soit $V \in I_{s_0}$. Alors : $V \in I_{s_0+k}$ (hypothèse de récurrence).

Par suite : $\exists U \in E, V = f^{s_0+k}(U) = f^k(f^{s_0}(U))$.

Puisque $f^{s_0}(U) \in I_{s_0}$, et que $I_{s_0} = I_{s_0+1}$, on a : $\exists W \in E, f^{s_0}(U) = f^{s_0+1}(W)$.

Ainsi : $V = f^k(f^{s_0+1}(W))$ c-à-d : $V = f^{s_0+k+1}(W)$ d'où : $V \in I_{s_0+k+1}$. On en déduit que $I_{s_0} \subset I_{s_0+k+1}$ et par double inclusion : $I_{s_0} = I_{s_0+k+1}$. Récurrence établie.

Conclusion. $\forall k \in \mathbb{N}^*, I_{s_0+k} = I_{s_0}$.

La croissance de la suite des noyaux et une double application du théorème du rang (aux endomorphismes f^{s_0} et f^{s_0+k}) permettent d'affirmer : $\forall k \in \mathbb{N}^*, N_{s_0+k} = N_{s_0}$.

24.12.2 Espace vectoriel dual

On rappelle que le dual E^* d'un espace vectoriel E est l'espace vectoriel des formes linéaires sur E .

PROPRIÉTÉ 24.25 - Tout espace vectoriel de dimension fini est isomorphe à son dual.

PREUVE. Soit E un espace vectoriel de dimension $n \in \mathbb{N}^*$. Notons $\mathcal{B} = (e_1, \dots, e_n)$ une base de E , et notons $\mathcal{B}^* = (e_1^*, \dots, e_n^*)$ la famille des formes coordonnées associée.

Montrons que $\mathcal{B}^*$ est une base de E^* en établissant qu'elle est génératrice est libre.

Soit $(\alpha_1, \dots, \alpha_n)$ une famille de n scalaires telle que : $\sum_{i=1}^n \alpha_i e_i^* = 0_{E^*}$. Alors :

$$\left[\forall j \in \llbracket 1, n \rrbracket, \left(\sum_{i=1}^n \alpha_i e_i^* \right) (e_j) = 0 \right] \implies [\forall j \in \llbracket 1, n \rrbracket, \alpha_j = 0]$$

Ce qui assure que la famille $\mathcal{B}^*$ est libre.

Considérons à présent $f \in E^*$. Montrons que : $f = \sum_{i=1}^n f(e_i) e_i^*$.

Pour cela, il suffit d'observer que :

$$\forall j \in \llbracket 1, n \rrbracket, \left(\sum_{i=1}^n f(e_i) e_i^* \right) (e_j) = f(e_j)$$

Il s'ensuit que les formes linéaires f et $\sum_{i=1}^n f(e_i) e_i^*$ coïncident sur une base de E . A ce titre, elles sont égales. Ce qui assure que la famille $\mathcal{B}^*$ engendre E^* .

On peut donc conclure que $\mathcal{B}^*$ est une base de E^* . Il s'ensuit que $\dim E^* = \dim E$, ce qui implique que E et E^* sont isomorphes.³ ☕

24.12.3 Un nouvel invariant de similitude : le polynôme caractéristique

Dans le paragraphe consacré aux matrices semblables, nous avons rencontré trois invariants de similitude : le rang, la trace et le déterminant. Il en existe encore un autre, relevant du programme de Spé, et que vous utiliserez donc l'an prochain : le polynôme caractéristique. Sans vouloir empiéter sur le programme de mes collègues de Spé, voici deux mots sur cette notion.⁴

DÉFINITION 24.14 - Soit $A \in M_n(\mathbb{K})$ une matrice. Le **polynôme caractéristique** de la matrice A est le polynôme $\chi_A(\lambda) \in \mathbb{K}[\lambda]$ défini en posant :

$$\chi_A(\lambda) = \det(\lambda I_n - A)$$

PROPRIÉTÉ 24.26 - Le polynôme caractéristique est un invariant de similitude.

Explicitement, si A et B sont deux matrices de $M_n(\mathbb{K})$ semblables, alors $\chi_A = \chi_B$.

PREUVE. Si A et B sont semblables, alors il existe une matrice $P \in GL_n(\mathbb{K})$ telle que : $B = P^{-1}AP$.

On a alors :

$$\chi_B(\lambda) = \det(\lambda I_n - P^{-1}AP) = \det(P^{-1}(\lambda I_n - A)P) = \underbrace{\det(P^{-1}) \det(P)}_{=1} \det(\lambda I_n - A) = \chi_A(\lambda)$$

Ainsi : $[A \text{ et } B \text{ semblables}] \implies [\chi_A = \chi_B]$. ☕

3. En vertu du théorème de classification des espaces vectoriels de dimension finie.

4. En fait, après avoir parlé des précédents invariants de similitude, l'occasion était trop belle de présenter les polynômes caractéristiques ; d'autant plus que la définition de ce polynôme d'une part, et la preuve du fait que c'est un invariant de similitude d'autre part, ne font intervenir que des notions relevant purement du programme de Sup. En revanche, l'utilisation et les autres propriétés des polynômes caractéristiques demandent des outils spécifiques que vous ne verrez que l'an prochain.

24.12.4 Une propriété des endomorphismes nilpotents

Exercice classique. Soit E un $\mathbb{K}$ -ev de dimension n , et f un endomorphisme de E nilpotent et non nul. On note p le plus petit entier tel que $f^p = 0_{\mathcal{L}(E)}$.

- 1) Montrer qu'il existe un vecteur V de E tel que la famille : $(V, f(V), f^2(V), \dots, f^{p-1}(V))$ est libre.
- 2) En déduire que $f^n = 0_{\mathcal{L}(E)}$.

Corrigé.

1) Puisque p est le plus petit entier tel que $f^p = 0_{\mathcal{L}(E)}$, on a : $f^{p-1} \neq 0_{\mathcal{L}(E)}$. Il existe donc un vecteur V tel que $f^{p-1}(V) \neq 0_E$.

Supposons alors qu'il existe p scalaires $\alpha_0, \dots, \alpha_{p-1}$ tels que : $\alpha_0 V + \alpha_1 f(V) + \dots + \alpha_{p-1} f^{p-1}(V) = 0_E$.

En appliquant f^{p-1} à cette relation, on obtient : $\alpha_0 f^{p-1}(V) + \alpha_1 f^p(V) + \dots + \alpha_{p-1} f^{2p-2}(V) = 0_E$.

Or puisque $f^p = 0_{\mathcal{L}(E)}$, cette dernière égalité implique : $\alpha_0 f^{p-1}(V) = 0_E$.

Comme $f^{p-1}(V) \neq 0_E$, on en déduit $\alpha_0 = 0$.

La relation initiale se réécrit alors : $\alpha_1 f(V) + \dots + \alpha_{p-1} f^{p-1}(V) = 0_E$.

En appliquant f^{p-2} à cette égalité, un raisonnement analogue fournit $\alpha_1 = 0$.

En répétant le procédé autant de fois que nécessaire, on obtient finalement : $\alpha_0 = \alpha_1 = \dots = \alpha_p = 0$.

Il s'ensuit que la famille $(V, f(V), f^2(V), \dots, f^{p-1}(V))$ est libre.

Conclusion. Soit f un endomorphisme de E (de dimension finie) nilpotent et non nul. On note p le plus petit entier tel que $f^p = 0_{\mathcal{L}(E)}$. Il existe un vecteur V de E tel que la famille : $(V, f(V), f^2(V), \dots, f^{p-1}(V))$ est libre.

2) Soit f un endomorphisme de E nilpotent et non nul.

Raisonnons par l'absurde et supposons que $f^n \neq 0_{\mathcal{L}(E)}$.

Alors d'après la propriété précédente, il existe un vecteur V de E tel que la famille : $\mathcal{F} = (V, f(V), \dots, f^n(V))$ est libre. Or ceci est absurde, puisque la famille $\mathcal{F}$ est de cardinal $n+1$, et que E est de dimension n (or la cardinal d'une famille libre d'un ev de dimension finie est majoré par la dimension de cet ev).

Donc : $f^n = 0_{\mathcal{L}(E)}$.

Conclusion. Soit f un endomorphisme de E (de dimension finie) nilpotent et non nul. On a : $f^n = 0_{\mathcal{L}(E)}$.

Remarque : le résultat établi dans cet exercice admet évidemment une traduction matricielle.

Conclusion. Soit $A \in M_n(\mathbb{K})$ une matrice nilpotente. On a : $A^n = 0_{M_n(\mathbb{K})}$.

24.12.5 A propos des endomorphismes cycliques

L'énoncé ci-dessous est celui d'un problème ayant pour but de présenter la notion d'endomorphisme cyclique, et d'établir quelques propriétés générales de ces endomorphismes.

Soit E un $\mathbb{R}$ -espace vectoriel. On note id l'identité de E .

Pour tout endomorphisme f de E , on note $f^0 = \text{id}$, et pour tout entier naturel k , $f^{k+1} = f \circ f^k$.

On dit qu'un endomorphisme f de E est **cyclique d'ordre p** s'il existe un élément $\vec{a}$ de E vérifiant les trois conditions suivantes :

$$(C1) \quad f^p(\vec{a}) = \vec{a};$$

$$(C2) \quad \text{la famille } (\vec{a}, f(\vec{a}), \dots, f^{p-1}(\vec{a})) \text{ est génératrice de } E;$$

$$(C3) \quad \text{la famille } (\vec{a}, f(\vec{a}), \dots, f^{p-1}(\vec{a})) \text{ est constituée d'éléments deux à deux distincts.}$$

Lorsque ces conditions sont remplies, la famille $(\vec{a}, f(\vec{a}), \dots, f^{p-1}(\vec{a}))$ est alors appelée un **cycle** de E .

Partie 1 — Exemples

1) Dans cette question, on pose $E = \mathbb{R}^2$. On considère l'endomorphisme f de $\mathbb{R}^2$ défini en posant : $\forall (x, y) \in \mathbb{R}^2, f(x, y) = (-y, x)$.

En considérant $\vec{a} = (1, 0)$, montrer que f est cyclique d'ordre p , l'entier p étant à préciser.

2) Dans cette question, on pose $E = \text{Vect}(\cos, \sin)$ (sous-espace vectoriel de $\mathbb{R}^{\mathbb{R}}$ engendré par les fonctions $\cos$ et $\sin$).

a) Déterminer la dimension de E .

b) Soit $p \in \mathbb{N} \setminus \{0, 1, 2\}$. Pour $f \in E$, on note $\tau_p(f)$ la fonction définie sur $\mathbb{R}$ par :

$$\forall x \in \mathbb{R}, (\tau_p(f))(x) = f\left(x + \frac{2\pi}{p}\right)$$

Montrer que $\tau_p(f) \in E$.

c) Montrer que l'application $\tau_p : f \mapsto \tau_p(f)$ est un endomorphisme de E .

d) On pose $f = \sin$. Exprimer, pour tout $k \in \mathbb{N}$, $\tau_p^k(f)$.

e) Montrer que : $\forall (k, m) \in \mathbb{N}^2, \tau_p^k(f) = \tau_p^m(f) \implies k \equiv m \pmod{p}$

f) Montrer que τ_p est cyclique d'ordre p .

Partie 2 — Cas général

Dans cette partie E désigne un $\mathbb{R}$ -espace vectoriel de dimension $n \in \mathbb{N}^*$.

On considère un endomorphisme f de E cyclique d'ordre p , et on note : $(\vec{a}, f(\vec{a}), \dots, f^{p-1}(\vec{a}))$ un cycle de f .

3) Justifier que $p \geq n$.

4) a) Observer que : $\forall k \in \mathbb{N}, f^p(f^k(\vec{a})) = f^k(\vec{a})$.

b) En déduire que $f^p = \text{id}$. L'endomorphisme f est-il bijectif?

- c) Justifier brièvement que $\ker(f - \text{id})$ et $\ker(\text{id} + f + \dots + f^{p-1})$ sont des sous-espaces vectoriels de E . Etablir qu'ils sont supplémentaires dans E .
- 5) On note m le plus grand des entiers k tels que la famille $(\vec{a}, f(\vec{a}), \dots, f^{k-1}(\vec{a}))$ soit libre.
- 6) a) Montrer que $f^m(\vec{a})$ est combinaison linéaire des m vecteurs $\vec{a}, f(\vec{a}), \dots, f^{m-1}(\vec{a})$.
- b) Montrer par récurrence que pour tout entier $k \geq m$, le vecteur $f^k(\vec{a})$ est combinaison linéaire des vecteurs $\vec{a}, f(\vec{a}), \dots, f^{m-1}(\vec{a})$.
- c) En déduire que $m = n$, et que la famille $(\vec{a}, f(\vec{a}), \dots, f^{n-1}(\vec{a}))$ est une base de E .
- 7) Soit g un endomorphisme de E commutant avec f .

On note $\alpha_0, \dots, \alpha_{n-1}$ les n scalaires tels que : $g(\vec{a}) = \sum_{k=0}^{n-1} \alpha_k f^k(\vec{a})$.

On considère alors h l'endomorphisme de E défini par : $h = \sum_{k=0}^{n-1} \alpha_k f^k$.

- a) Montrer que f et h commutent.
- b) Montrer que : $\forall k \in \mathbb{N}, g(f^k(\vec{a})) = h(f^k(\vec{a}))$.
- c) En déduire que $g = h$.
- d) Quels sont les endomorphismes de E commutant avec f ?

CORRIGÉ

Partie 1 — Exemples

1) $f(\vec{a}) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$; $f^2(\vec{a}) = \begin{pmatrix} -1 \\ 0 \end{pmatrix}$; $f^3(\vec{a}) = \begin{pmatrix} 0 \\ -1 \end{pmatrix}$; et $f^4(\vec{a}) = \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \vec{a}$.

Conclusion. L'endomorphisme f est cyclique d'ordre 4.

2) a) Puisque les fonctions $\cos$ et $\sin$ ne sont pas colinéaires, $\dim \text{Vect}(\cos, \sin) = 2$.

b) Soit $p \in \mathbb{N} \setminus \{0, 1, 2\}$, et soit $f = \alpha \cos + \beta \sin \in E$. On a :

$$\begin{aligned} \forall x \in \mathbb{R}, (\tau_p(f))(x) &= \alpha \cos\left(x + \frac{2\pi}{p}\right) + \beta \sin\left(x + \frac{2\pi}{p}\right) = \alpha \cos(x) \cos\left(\frac{2\pi}{p}\right) - \alpha \sin(x) \sin\left(\frac{2\pi}{p}\right) \\ &\quad + \beta \sin(x) \cos\left(\frac{2\pi}{p}\right) + \beta \cos(x) \sin\left(\frac{2\pi}{p}\right) \end{aligned}$$

$$\text{D'où : } (\tau_p(f)) = \gamma \cos + \delta \sin \text{ avec : } \begin{cases} \gamma = \alpha \cos\left(\frac{2\pi}{p}\right) + \beta \sin\left(\frac{2\pi}{p}\right) \\ \text{et} \\ \delta = \beta \cos\left(\frac{2\pi}{p}\right) - \alpha \sin\left(\frac{2\pi}{p}\right) \end{cases}$$

Conclusion. $\forall f \in E, \tau_p(f) \in E$.

$$\begin{aligned}
 \text{c) } \forall (f, g) \in E^2, \forall (\lambda, \mu) \in \mathbb{R}^2, \forall x \in \mathbb{R}, [\tau_p(\lambda f + \mu g)](x) &= (\lambda f + \mu g)\left(x + \frac{2\pi}{p}\right) \\
 &= \lambda f\left(x + \frac{2\pi}{p}\right) + \mu g\left(x + \frac{2\pi}{p}\right) = \lambda(\tau_p(f))(x) + \mu(\tau_p(g)(x))
 \end{aligned}$$

Conclusion. $\forall (f, g) \in E^2, \forall (\lambda, \mu) \in \mathbb{R}^2, \tau_p(\lambda f + \mu g) = \lambda \tau_p(f) + \mu \tau_p(g)$. Donc : $\tau_p \in \mathcal{L}(E)$.

d) Soit $k \in \mathbb{N}$ et $f = \sin$.

$$\text{On a : } \forall x \in \mathbb{R}, (\tau_p^k(f))(x) = \sin\left(x + \frac{2k\pi}{p}\right) = \sin\left(\frac{2k\pi}{p}\right) \cos(x) + \cos\left(\frac{2k\pi}{p}\right) \sin(x)$$

e) Soient k et m deux entiers tels que : $\tau_p^k(f) = \tau_p^m(f)$. Alors d'après la question précédente on a :

$$\begin{aligned}
 \sin\left(\frac{2k\pi}{p}\right) \cos + \cos\left(\frac{2k\pi}{p}\right) \sin &= \sin\left(\frac{2m\pi}{p}\right) \cos + \cos\left(\frac{2m\pi}{p}\right) \sin \\
 \text{La famille } (\cos, \sin) \text{ étant libre, on déduit de cette égalité que : } &\begin{cases} \sin\left(\frac{2k\pi}{p}\right) = \sin\left(\frac{2m\pi}{p}\right) \\ \cos\left(\frac{2k\pi}{p}\right) = \cos\left(\frac{2m\pi}{p}\right) \end{cases}
 \end{aligned}$$

$$\text{D'où : } \frac{2k\pi}{p} = \frac{2m\pi}{p} \pmod{2\pi} \text{ donc : } k \equiv m \pmod{p}.$$

$$\text{Ainsi : } \forall (k, m) \in \mathbb{N}^2, \tau_p^k(f) = \tau_p^m(f) \implies k \equiv m \pmod{p}.$$

f) D'après la question d, on a : $\tau_p^p(f) = f$. De plus, la famille $(f, \tau_p(f))$ est génératrice de E puisqu'elle est libre (f et $\tau_p(f)$ ne sont pas colinéaires) dans un espace vectoriel de dimension 2 ; a fortiori donc, la famille $(f, \tau_p(f), \dots, \tau_p^{p-1}(f))$ est génératrice de E . Enfin, les éléments de cette famille sont deux à deux distincts, toujours d'après la question d et car la famille $(\cos, \sin)$ est libre.

Conclusion. τ_p est un endomorphisme de E cyclique d'ordre p .

Partie 2 — Cas général

3) Puisque la famille $(\vec{a}, f(\vec{a}), \dots, f^{p-1}(\vec{a}))$ est génératrice de E , son cardinal (p) est supérieur ou égal à la dimension de E (n). D'où : $p \geq n$.

$$\text{4) a) } \forall k \in \mathbb{N}, f^p(f^k(\vec{a})) = f^{p+k}(\vec{a}) = f^k(f^p(\vec{a})) = f^k(\vec{a}). \text{ Ainsi : } \forall k \in \mathbb{N}, f^p(f^k(\vec{a})) = f^k(\vec{a}).$$

b) Soit $\vec{v} \in E$. Puisque la famille $(\vec{a}, f(\vec{a}), \dots, f^{p-1}(\vec{a}))$ est génératrice de E :

$$\exists (\alpha_0, \dots, \alpha_{p-1}) \in \mathbb{R}^p, \vec{v} = \sum_{k=0}^{p-1} \alpha_k f^k(\vec{a})$$

$$\text{Par suite : } f^p(\vec{v}) = f^p\left(\sum_{k=0}^{p-1} \alpha_k f^k(\vec{a})\right) = \sum_{k=0}^{p-1} \alpha_k f^{p+k}(\vec{a}) = \sum_{k=0}^{p-1} \alpha_k f^k(\vec{a}) = \vec{v}.$$

$$\text{Ainsi : } \forall \vec{v} \in E, f^p(\vec{v}) = \vec{v}. \quad [f^p = \text{id}_E]. \text{ Il s'ensuit que : } [f \in \text{GL}(E), \text{ d'automorphisme réciproque } f^{p-1}].$$

c) Les parties $\ker(f - \text{id})$ et $\ker(\text{id} + f + \dots + f^{p-1})$ sont des noyaux d'endomorphismes de E , à ce titre, ce sont des sous-espaces vectoriels de E .

Soit $\vec{v} \in (\ker(f - \text{id}) \cap \ker(\text{id} + f + \dots + f^{p-1}))$. Alors $f(\vec{v}) = \vec{v}$. D'où : $(\text{id} + f + \dots + f^{p-1})(\vec{v}) = p\vec{v}$. Comme $\vec{v} \in \ker(\text{id} + f + \dots + f^{p-1})$, on a donc $\vec{v} = \vec{0}$.

D'où : $\ker(f - \text{id}) \cap \ker(\text{id} + f + \dots + f^{p-1}) = \{\vec{0}\}$.

Soit à présent : $\vec{v} \in E$. Posons : $\vec{u} = \frac{1}{p} (\vec{v} + f(\vec{v}) + \dots + f^{p-1}(\vec{v}))$ et $\vec{w} = \vec{v} - \vec{u}$, de telle sorte que : $\vec{v} = \vec{u} + \vec{w}$.

Alors : $f(\vec{u}) = \frac{1}{p} \left(f(\vec{v}) + f^2(\vec{v}) + \dots + \underbrace{f^p(\vec{v})}_{=\vec{v}} \right) = \frac{1}{p} (\vec{v} + f(\vec{v}) + \dots + f^{p-1}(\vec{v})) = \vec{u}$.

D'où : $\vec{u} \in \text{Ker}(f - \text{id})$.

D'autre part : $(\text{id} + f + \dots + f^{p-1})(\vec{w}) = (\text{id} + f + \dots + f^{p-1})(\vec{v}) - (\text{id} + f + \dots + f^{p-1})(\vec{u}) = p\vec{u} - p\vec{u} = \vec{0}$. D'où $\vec{w} \in \text{Ker}(\text{id} + f + \dots + f^{p-1})$.

En résumé : $E = \text{Ker}(f - \text{id}) + \text{Ker}(\text{id} + f + \dots + f^{p-1})$. De plus : $\ker(f - \text{id}) \cap \ker(\text{id} + f + \dots + f^{p-1}) = \{\vec{0}\}$. On peut conclure : $E = \text{Ker}(f - \text{id}) \oplus \text{Ker}(\text{id} + f + \dots + f^{p-1})$.

6) a) Par définition de l'entier m , la famille $(\vec{a}, f(\vec{a}), \dots, f^m(\vec{a}))$ est liée. Subséquemment il existe $(m+1)$ scalaires $\alpha_0, \dots, \alpha_m$ non tous nuls tels que :

$$\sum_{k=0}^m \alpha_k f^k(\vec{a}) = \vec{0} \text{ d'où : } \alpha_m f^m(\vec{a}) = - \sum_{k=0}^{m-1} \alpha_k f^k(\vec{a}).$$

Si α_m était nul, alors la relation ci-dessus établirait l'existence de m scalaires $\alpha_0, \dots, \alpha_{m-1}$ non tous nuls tels que : $\sum_{k=0}^{m-1} \alpha_k f^k(\vec{a}) = \vec{0}$, ce qui est absurde puisque la famille $(\vec{a}, f(\vec{a}), \dots, f^{m-1}(\vec{a}))$ est libre. Il

s'ensuit que $\alpha_m \neq 0$, d'où : $f^m(\vec{a}) = \sum_{k=0}^{m-1} \left(-\frac{\alpha_k}{\alpha_m} \right) f^k(\vec{a})$.

Conclusion. $f^m(\vec{a}) \in \text{Vect}(\vec{a}, f(\vec{a}), \dots, f^{m-1}(\vec{a}))$.

b) Montrons par récurrence que la propriété $P(K)$: " $f^{m+K}(\vec{a}) \in \text{Vect}(\vec{a}, f(\vec{a}), \dots, f^{m-1}(\vec{a}))$ " est vraie pour tout $K \in \mathbb{N}$.

La propriété est vraie pour $K = 0$ d'après la question précédente, d'où l'initialisation.

Passons à l'hérédité : supposons $P(K)$ vraie pour un certain entier naturel K . Alors : $f^{m+K+1}(\vec{a}) = f(f^{m+K}(\vec{a}))$. Par hypothèse de récurrence, il existe m scalaires $\alpha_0, \dots, \alpha_{m-1}$ tels que :

$$f^{m+K}(\vec{a}) = \sum_{k=0}^{m-1} \alpha_k f^k(\vec{a})$$

D'où : $f^{m+K+1}(\vec{a}) = \sum_{k=0}^{m-1} \alpha_k f^{k+1}(\vec{a}) = \underbrace{\sum_{k=0}^{m-2} \alpha_k f^{k+1}(\vec{a})}_{\in \text{Vect}(\vec{a}, f(\vec{a}), \dots, f^{m-1}(\vec{a}))} + \underbrace{\alpha_{m-1} f^m(\vec{a})}_{\in \text{Vect}(\vec{a}, f(\vec{a}), \dots, f^{m-1}(\vec{a}))}$

Par suite : $f^{m+K+1}(\vec{a}) \in \text{Vect}(\vec{a}, f(\vec{a}), \dots, f^{m-1}(\vec{a}))$. Ce qui établit l'hérédité.

Conclusion. $\forall K \in \mathbb{N}, f^{m+K}(\vec{a}) \in \text{Vect}(\vec{a}, f(\vec{a}), \dots, f^{m-1}(\vec{a}))$

$$\iff \forall k \in \mathbb{N}, [k \geq m] \implies [f^k(\vec{a}) \in \text{Vect}(\vec{a}, f(\vec{a}), \dots, f^{m-1}(\vec{a}))]$$

c) D'après la question précédente, on a : $\text{Vect}(\vec{a}, f(\vec{a}), \dots, f^{p-1}(\vec{a})) = \text{Vect}(\vec{a}, f(\vec{a}), \dots, f^{m-1}(\vec{a}))$.

La famille $\text{Vect}(\vec{a}, f(\vec{a}), \dots, f^{m-1}(\vec{a}))$ étant génératrice et libre, c'est une base de E .

7) a) D'après l'énoncé : $h = \sum_{k=0}^{n-1} \alpha_k f^k$. Comme il est clair que f commute avec toute itérée de f , et que la partie de $\mathcal{L}(E)$ des endomorphismes commutant avec f est un sous-espace vectoriel, on peut affirmer que : $f \circ h = h \circ f$.

b) Soit $k \in \mathbb{N}$. Puisque g commute avec f par hypothèse, alors g commute avec toute itérée de f (récurrence immédiate). On a donc : $g(f^k(\vec{a})) = f^k(g(\vec{a})) = f^k\left(\sum_{i=0}^{n-1} \alpha_i f^i(\vec{a})\right) = \sum_{i=0}^{n-1} \alpha_i f^{i+k}(\vec{a})$

$$= \sum_{i=0}^{n-1} \alpha_i f^i(f^k(\vec{a})) = \left(\sum_{i=0}^{n-1} \alpha_i f^i\right)(f^k(\vec{a})) = h(f^k(\vec{a}))$$

Conclusion. $\forall k \in \mathbb{N}, g(f^k(\vec{a})) = h(f^k(\vec{a}))$.

c) D'après la question précédente, g et h coïncident sur la famille $(\vec{a}, f(\vec{a}), \dots, f^{p-1}(\vec{a}))$. Par linéarité, elles sont donc égales sur $\text{Vect}(\vec{a}, f(\vec{a}), \dots, f^{p-1}(\vec{a}))$, c'est à dire sur E puisque la famille $(\vec{a}, f(\vec{a}), \dots, f^{p-1}(\vec{a}))$ est supposée génératrice.

Conclusion. $h = g$.

d) Nous avons déjà observé que le sous-espace vectoriel (notons-le $\text{COM}(f)$) des endomorphismes de E commutant avec f contient les itérées de f , et par suite, toutes les combinaisons linéaires de ses itérées. En particulier : $\text{Vect}(\text{id}, f, \dots, f^{p-1}) \subset \text{COM}(f)$.

D'après c, l'inclusion réciproque est vraie, puisque tout endomorphisme commutant avec f est combinaison linéaire des f^k .

Conclusion. $\text{COM}(f) = \{g \in \mathcal{L}(E) / g \circ f = f \circ g\} = \text{Vect}(\text{id}, f, \dots, f^{p-1})$.

24.12.6 Hyperplans dans $M_n(\mathbb{K})$

L'objectif principal de ce problème est de prouver que dans $M_n(\mathbb{K})$, les matrices inversibles sont "partout". Sans précision, cette affirmation n'aurait pas grand sens... Pour lui donner davantage de consistance, notez que vous avez déjà entendu une phrase du même type au cours de cette année, lorsque nous avons établi que les nombres rationnels sont "partout" dans l'ensemble des nombres réels, c'est à dire que $\mathbb{Q}$ est dense dans $\mathbb{R}$.

De façon analogue, on peut montrer que $GL_n(\mathbb{K})$ est dense dans $M_n(\mathbb{K})$; ce résultat figure d'ailleurs au programme de Spé, mais dans l'état actuel de nos connaissances, il nous manque de précieux outils (notamment les notions de distance et de voisinage dans $M_n(\mathbb{K})$), primo pour lui donner un sens, et secundo pour le prouver.

Plus raisonnablement donc, l'objectif du présent problème est d'établir un résultat un peu plus modeste mais néanmoins spectaculaire : tout hyperplan de l'espace des matrices carrées contient une matrice inversible.

Tout au long de ce problème, on note $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$; $E = M_n(\mathbb{K})$ le $\mathbb{K}$ -ev des matrices carrées de taille n à coefficients dans $\mathbb{K}$; et n désignera un entier supérieur ou égal à 2.

Un élément de E sera noté $M = (m_{ij})$, et on notera encore E_{ij} les matrices élémentaires de E . On notera E^* l'espace vectoriel dual de E , c'est à dire l'espace vectoriel $\mathcal{L}(E, \mathbb{K})$ des formes linéaires sur E .

On rappelle que la trace de $M \in E$ est le scalaire noté $\text{tr}(M)$, et défini par : $\text{tr}(M) = \sum_{i=1}^n m_{ii}$.

Enfin, à toute matrice U de E , on associe l'application
$$\left[\begin{array}{ccc} T_U : E & \longrightarrow & \mathbb{K} \\ M & \longmapsto & \text{tr}(UM) \end{array} \right].$$

1) Montrer que pour tout $U \in E$, l'application T_U est une forme linéaire sur E .

2) Soit U un élément de E . On pose : $H_U = \text{Ker}(T_U)$.

a) Si U est la matrice nulle, déterminer H_U .

b) Supposons à présent que U est une matrice non nulle dans E .

Montrer qu'il existe un couple d'entiers (i_0, j_0) tel que $T_U(E_{i_0, j_0}) \neq 0$. Décrire alors $\text{Im}(T_U)$ et préciser la dimension de H_U .

3) Pour $(i, j) \in \llbracket 1, n \rrbracket^2$, on note $T_{i,j} = T_{E_{i,j}}$.

a) Soient k et l deux entiers dans $\llbracket 1, n \rrbracket$. Calculer : $T_{i,j}(E_{k,l})$.

b) En déduire que les n^2 éléments $T_{i,j}$ de E^* constituent une base de E^* .

c) Montrer que l'application
$$\left[\begin{array}{ccc} \varphi : E & \longrightarrow & E^* \\ U & \longmapsto & T_U \end{array} \right]$$
 est un isomorphisme.

4) Soit H un hyperplan de E .

- a) Soit $A \in E$. A quelle condition sur A les sous-espaces vectoriels H et $\text{Vect}(A)$ sont-ils supplémentaires dans E ? (on pourra rappeler le résultat sans démonstration).
- b) Etablir qu'il existe un élément ℓ non nul de E^* tel que $H = \text{Ker}(\ell)$.
- c) Montrer qu'il existe un élément U de E tel que : $H = H_U$.

5) Pour tout entier $r \in \llbracket 1, n \rrbracket$, on note $J_r = \sum_{i=1}^r E_{ii}$,⁵ et A la matrice définie par :

$$A = \begin{pmatrix} 0 & 0 & \cdots & 0 & 1 \\ 1 & 0 & \ddots & \vdots & 0 \\ 0 & \ddots & \ddots & 0 & \vdots \\ \vdots & \ddots & \ddots & 0 & \vdots \\ 0 & \cdots & 0 & 1 & 0 \end{pmatrix}$$

- a) Prouver que A est inversible.
- b) Prouver que A appartient à l'hyperplan H_{J_r} .
- c) Conclure que chaque hyperplan H de E contient au moins une matrice inversible.⁶

CORRIGÉ

1) Soit $U \in E$. Alors : $\forall (M, N) \in E^2, \forall (\lambda, \mu) \in \mathbb{K}^2, T_U(\lambda M + \mu N) = \text{tr}(U(\lambda M + \mu N))$

$$= \text{tr}(\lambda UM + \mu UN) = \lambda \text{tr}(UM) + \mu \text{tr}(UN) = \lambda T_U(M) + \mu T_U(N)$$

L'application U est linéaire, et à valeurs dans $\mathbb{K}$. On peut donc conclure que T_U est une forme linéaire sur E .

Conclusion. $\forall U \in E, T_U \in E^*$.

2) a) Si U est la matrice nulle, alors la forme linéaire T_U est identiquement nulle, et son noyau est donc E tout entier. **Conclusion.** $[U = 0_E] \implies [\text{Ker}(T_U) = E]$.

b) Soit U une matrice non nulle de E . En notant $U = (u_{ij}) \in E$, on peut donc affirmer qu'il existe un couple (i_1, j_1) d'entiers naturels tel que : $u_{i_1 j_1} \neq 0$. Considérons alors la matrice produit : $C = U E_{j_1 i_1}$, et calculons ses coefficients diagonaux. Soit i un entier compris entre 1 et n . On a :

$$C_{ii} = \sum_{k=1}^n u_{ik} (E_{j_1 i_1})_{ki} = \sum_{k=1}^n u_{i_1 k} \delta_{i i_1} \delta_{j_1 k} = u_{i_1 j_1} \delta_{ii_1}$$

Il s'ensuit que les coefficients diagonaux de la matrice C sont nuls, à l'exception d'un seul : $C_{i_1 i_1} = u_{i_1 j_1}$. Par suite : $\text{tr}(C) = u_{i_1 j_1}$. En posant : $i_0 = j_1$ et $j_0 = i_1$, on a donc : $\text{tr}(U E_{i_0 j_0}) \neq 0$, c'est à dire : $T_U(E_{i_0 j_0}) \neq 0$.

5. Il s'agit ici de la même matrice J_r que celle rencontrée dans le cours d'algèbre linéaire, c'est à dire la matrice de $M_n(\mathbb{K})$ qui est diagonale, dont les r premiers coefficients diagonaux sont égaux à 1 et dont les $(n-r)$ derniers sont nuls.

6. Indication : on rappelle que si U est une matrice de E de rang r , alors il existe deux matrices inversibles P et Q dans E telles que $PUQ = J_r$ (les matrices U et J_r sont équivalentes).

Conclusion. $\forall U \in E, U \neq 0_E, \exists (i_0, j_0) \in \llbracket 1, n \rrbracket^2, T_U(E_{i_0 j_0}) \neq 0$.

Dans ces conditions, T_U étant une forme linéaire non nulle, elle est surjective ($\text{Im}(T_U) = \mathbb{K}$) et son noyau est un hyperplan de E , c'est à dire que H_U est un sev de E de dimension $n^2 - 1$.

3) Soit $(i, j) \in \llbracket 1, n \rrbracket^2$ fixé.

a) Soient k et l deux entiers dans $\llbracket 1, n \rrbracket$. En essayant de garder son calme, on obtient : $E_{ij}E_{kl} = \delta_{il}\delta_{jk}E_{il}$. La trace de cette matrice est nulle dès que $i \neq l$, égale à 1 si $i = l$ (et $j = k$). D'où :

Conclusion. $\forall (i, j, k, l) \in \llbracket 1, n \rrbracket^4, T_{ij}(E_{kl}) = \delta_{il}\delta_{jk}$.

b) Montrons que la famille $(T_{ij})_{1 \leq i, j \leq n}$ est libre. Supposons qu'il existe n^2 scalaires α_{ij} tq : $\sum_{i=1}^n \sum_{j=1}^n \alpha_{ij} T_{ij} = 0_{E^*}$.

D'après la question précédente, pour tout couple d'entiers (k, l) entre 1 et n on a :

$$\sum_{i=1}^n \sum_{j=1}^n \alpha_{ij} T_{ij}(E_{kl}) = 0 \quad \text{d'où} \quad \alpha_{lk} T_{lk}(E_{kl}) = 0 \quad \text{d'où} \quad \alpha_{lk} = 0.$$

Donc : $\forall (k, l) \in \llbracket 1, n \rrbracket^2, \alpha_{lk} = 0$. Il s'ensuit que la famille $(T_{ij})_{1 \leq i, j \leq n}$ est libre. Puisque son cardinal (n^2) est égal à la dimension de E^* , on peut conclure que : la famille $(T_{ij})_{1 \leq i, j \leq n}$ est une base de E^* .

c) D'après la question précédente, l'image de la base canonique $(E_{ij})_{1 \leq i, j \leq n}$ de E par l'application φ est la famille $(T_{ij})_{1 \leq i, j \leq n}$, qui est une base de E^* . Il s'ensuit que φ est un isomorphisme entre E et son dual E^* .

4) a) Pur cours : soit H un hyperplan de E . Alors : $\forall A \in E, [E = H \oplus \text{Vect}(A)] \iff [A \notin H]$.

b) Pur cours (adapté au cas particulier de $E = M_n(\mathbb{K})$) de nouveau. Puisque H est un hyperplan de E , H admet une base de cardinal $n^2 - 1$: $(M_1, \dots, M_{n^2-1})$. Choisissons alors une matrice $A \in (E \setminus H)$. La famille $\mathcal{B} = (M_1, \dots, M_{n^2-1}, A)$ est une base de E . On définit alors une forme linéaire :

$$\begin{aligned} \ell : E &\longrightarrow \mathbb{K} & \text{où } \beta \text{ est tel que : } & M = \left(\sum_{i=1}^{n^2-1} \alpha_i M_i \right) + \beta A \\ M &\longmapsto \beta \end{aligned}$$

Et il est alors clair que : $H = \text{Ker}(\ell)$. Pour tout hyperplan H de E , il existe $\ell \in E^*$ telle que : $H = \text{Ker}(\ell)$.⁷

c) D'après la question précédente, il existe un élément ℓ de E^* tel que $H = \text{Ker}(\ell)$. On utilise alors l'isomorphisme de la question 3-c pour affirmer qu'il existe une (unique) matrice $U \in E$ telle que $\ell = T_U$.

Conclusion. Pour tout hyperplan H de E , il existe $U \in E$ telle que : $H = \text{Ker}(T_U)$.

5) a) ► La matrice A est inversible car ses colonnes constituent une base de $\mathbb{K}^n$.

► La matrice A est inversible car ses lignes constituent une base de $\mathbb{K}^n$.

► Pour tout couple d'entiers (i, j) entre 1 et n , on note $P_{i,j} \in E$ la matrice de permutation définie en posant : $P_{i,j} = I_n - E_{ii} - E_{jj} + E_{ij} + E_{ji}$. Rappelons que $P_{i,j} \in \text{GL}_n(\mathbb{K})$.

Alors : $\left(\prod_{i=0}^{n-2} P_{n-i-1, n-i} \right) A = I_n$.⁸ On a donc : $\text{rg}(A) = \text{rg}(I_n) = n$ puisque le rang d'une matrice est

7. Au passage, c'est un résultat prouvé dans un cadre plus général dans le cours, et vous pouviez l'utiliser directement, sans le redémontrer.

8. A est égale à I_n modulo une permutation circulaire de ses lignes.

invariant par multiplication par une matrice inversible. Il s'ensuit que A est inversible, puisqu'une matrice de $M_n(\mathbb{K})$ est inversible si et seulement si elle est de rang maximal.

► On a : $A \left(\prod_{i=1}^{n-1} P_{i,i+1} \right) = I_n$.⁹ On a donc : $\text{rg}(A) = \text{rg}(I_n) = n$ puisque le rang d'une matrice est invariant par multiplication par une matrice inversible. Il s'ensuit que A est inversible, puisqu'une matrice de $M_n(\mathbb{K})$ est inversible si et seulement si elle est de rang maximal.

Bref, quelle que soit la justification choisie : $A \in \text{GL}_n(\mathbb{K})$.

b) Notons $C = (c_{ij})$ la matrice produit $J_r A$. Pour déterminer la trace de C , il suffit de connaître ses éléments diagonaux. A cette fin, fixons un entier i entre 1 et n . On a : $c_{ii} = \sum_{k=1}^n (J_r)_{ik} A_{ki}$.

Or par construction de la matrice J_r , le coefficient $(J_r)_{ik}$ est nul si $i \neq k$ (et éventuellement non nul si $i = k$). D'où : $c_{ii} = (J_r)_{ii} A_{ii}$. La matrice A étant de diagonale nulle, on a $A_{ii} = 0$. Il s'ensuit que $c_{ii} = 0$.

Ce raisonnement étant valable pour un i arbitraire, on peut affirmer : $\forall i \in \llbracket 1, n \rrbracket, c_{ii} = 0$. Par suite : $\text{tr}(C) = 0$. Autrement dit : $\text{tr}(J_r A) = 0$, soit encore : $A \in H_{J_r}$.

c) Soit H un hyperplan de E . D'après la question 4-c, il existe une matrice $U \in M_n(\mathbb{K})$ telle que $H = H_U$. On peut de plus observer que U est non nulle (sinon H_U n'est pas un hyperplan, cf question 2-a). Le rang de U est donc un entier compris entre 1 et n ; notons le r . On sait que U est équivalente à J_r , et il existe donc deux matrices P et Q dans $\text{GL}_n(\mathbb{K})$ telles que : $PUQ = J_r$ (d'où $U = P^{-1}J_rQ^{-1}$).

D'après la question précédente, on a : $A \in H_{J_r}$, c'est à dire : $\text{tr}(J_r A) = 0$. Posons alors $A' = QAP$. La matrice A' est inversible puisque par hypothèse les matrices Q , A et P le sont, et que le produit est la loi interne de $\text{GL}_n(\mathbb{K})$. On a :

$$\begin{aligned} \text{tr}(UA') &= \text{tr}(P^{-1}J_rQ^{-1}QAP) = \text{tr}(P^{-1}J_rAP) \\ &= \text{tr}((P^{-1}J_rA)P) = \text{tr}(P(P^{-1}J_rA)) = \text{tr}(PP^{-1}J_rA) = \text{tr}(J_rA) = 0 \end{aligned}$$

Par suite, $A' \in \text{Ker}(T_U)$, soit $A' \in H_U$ c'est à dire $A' \in H$. L'hyperplan H contient donc une matrice inversible.

Conclusion. Tout hyperplan de $M_n(\mathbb{K})$ contient une matrice inversible.

9. A est égale à I_n modulo une permutation circulaire de ses colonnes.

24.12.7 Transformée de Fourier discrète

Dans la suite de ces lignes, n désigne un entier supérieur ou égal à 2.

DÉFINITION 3 — Soit $X = (x_0, \dots, x_{n-1})$ un n -uplet de scalaires. On appelle **transformée de Fourier discrète de X** , et on note $\hat{X}$ le n -uplet $(y_0, \dots, y_{n-1})$ où :

$$\forall k \in \llbracket 0, n-1 \rrbracket, \quad y_k = \sum_{l=0}^{n-1} x_l e^{-i \frac{2lk\pi}{n}} \quad \text{ou} \quad y_k = \sum_{l=0}^{n-1} x_l \omega^{-kl}$$

On peut traduire matriciellement cette définition. Le procédé qui y est décrit donne lieu à une application (notons la TFD) de $\mathbb{C}^n$ dans $\mathbb{C}^n$. Cette application est linéaire et admet une matrice dans la base canonique de $\mathbb{C}^n$. Explicitons cette matrice ; en notant à présent (et seulement dans cette matrice) $\omega = e^{-i \frac{2\pi}{n}}$, la matrice de TFD est

$$M_n = \begin{pmatrix} 1 & 1 & \cdots & 1 & 1 \\ 1 & \omega & \cdots & \omega^{n-2} & \omega^{n-1} \\ \vdots & & & & \vdots \\ 1 & \omega^{n-2} & \cdots & \omega^{(n-2)^2} & \omega^{(n-2)(n-1)} \\ 1 & \omega^{n-1} & \cdots & \omega^{(n-1)(n-2)} & \omega^{(n-1)^2} \end{pmatrix}$$

La transformée de Fourier discrète du n -uplet X est donc le n -uplet $\hat{X} = M_n X$. On peut donc réinterpréter la définition précédente comme suit :

DÉFINITION 3-BIS — La transformée de Fourier discrète est l'application

$$\begin{aligned} TFD : \mathbb{C}^n &\longrightarrow \mathbb{C}^n \\ X &\longmapsto M_n X \end{aligned}$$

Propriétés de la TFD

PROPRIÉTÉ 5 (LINÉARITÉ DE LA TFD) — La TFD est une application linéaire de $\mathbb{C}^n$ dans $\mathbb{C}^n$.

La matrice de TFD (M_n) est un cas particulier de **matrice de Vandermonde**. Ces matrices possèdent des propriétés remarquables, et sont en particulier inversibles. Plus précisément :

PROPRIÉTÉ 6 (INVERSE DE LA TFD) — Mêmes notations que précédemment.

$$\text{La matrice } M_n \text{ est inversible et } M_n^{-1} = \frac{1}{n} \overline{M_n}.$$

La preuve de la propriété précédente consiste à calculer explicitement le coefficient du produit des matrices M_n et $\overline{M_n}$, et à observer que les coefficients diagonaux de ce produit sont égaux à n , ceux situés en dehors de la diagonale étant nuls (on montre ainsi que $M_n \overline{M_n} = nI_n$).

Par suite, l'application TFD est bijective, et sa bijection réciproque est encore une application linéaire de $\mathbb{C}^n$ dans $\mathbb{C}^n$.

PROPRIÉTÉ 6-BIS (RE-INVERSE DE LA TFD)— Mêmes notations que précédemment. L'application qui à un n -uplet de scalaires $Y = (y_0, \dots, y_{n-1})$ associe le n -uplet de scalaires $X = (x_0, \dots, x_{n-1})$ où

$$\forall k \in \llbracket 0, n-1 \rrbracket, \quad x_k = \frac{1}{n} \sum_{l=0}^{n-1} y_l e^{+i \frac{2lk\pi}{n}} \quad \text{ou} \quad x_k = \frac{1}{n} \sum_{l=0}^{n-1} y_l \omega^{kl}$$

est l'automorphisme réciproque de la TFD (et sera noté dans la suite TFD^{-1}).

Une autre conséquence de la formule $M_n \overline{M_n} = nI_n$ est la propriété appelée **effet d'une conjugaison** dans le programme.

PROPRIÉTÉ 7 (EFFET D'UNE CONJUGAISON)— Mêmes notations que précédemment.

Pour tout n -uplet de scalaires $(x_0, \dots, x_{n-1})$:

$$\overline{\text{TFD}(\overline{x_0}, \dots, \overline{x_{n-1}})} = n \text{TFD}^{-1}(x_0, \dots, x_{n-1})$$

24.12.8 Origines de la diagonalisation

Dans ce problème, E désigne un $\mathbb{K}$ -espace vectoriel de dimension finie non nulle, et u un endomorphisme de E .

On dit que $\vec{v} \in E$ est un **vecteur propre de u** s'il est non nul et s'il existe un scalaire λ tel que $u(\vec{v}) = \lambda \vec{v}$. Le scalaire λ est alors appelé **valeur propre de u** .

"Réciproquement", un scalaire λ est une valeur propre de u s'il existe un vecteur $\vec{v} \in E$ non nul tel que $u(\vec{v}) = \lambda \vec{v}$.

En résumé, lorsqu'il existe $\vec{v}$ et λ satisfaisant les conditions ci-dessus, on dira que $\vec{v}$ est un vecteur propre associé à la valeur propre λ .

On note $\text{Sp}(u)$ l'ensemble des valeurs propres de u . Ce sous-ensemble de $\mathbb{K}$ est appelé **spectre** de u .

Partie 1 — Polynômes annulateurs d'un endomorphisme

On rappelle que si u est un endomorphisme de E , on convient que $u^0 = \text{id}$, et pour tout entier naturel non nul n on pose $u^n = u \circ u^{n-1}$.¹⁰

Par ailleurs à tout polynôme $P = \sum_{k=0}^n a_k X^k$ de $\mathbb{K}[X]$, on peut associer un endomorphisme de E noté $P(u)$,

et défini en posant : $P(u) = \sum_{k=0}^n a_k u^k$.¹¹

Dans ce contexte, un polynôme P est dit **annulateur de u** si $P(u) = 0_{\mathcal{L}(E)}$.

10. Plus explicitement mais moins rigoureusement : $u^n = \underbrace{u \circ u \circ \dots \circ u}_{n \text{ fois}}$

11. Par exemple, si $P = X^3 - 4X + 2$, alors $P(u) = u^3 - 4u + 2\text{id}$.

1) Exemples de polynômes annulateurs.

- a) On considère l'endomorphisme u_1 de $\mathbb{R}^2$ canoniquement associé à la matrice $A_1 = \begin{pmatrix} 0 & 2 \\ 2 & 0 \end{pmatrix}$. Montrer que $P_1 = X^2 - 4$ est un polynôme annulateur de u_1 .
- b) En recherchant concision et précision dans vos justifications, expliciter un polynôme annulateur (P_{u_i}) pour chacun des endomorphismes (u_i) suivants :

$$\text{☞ } u_2 = \text{id}$$

$$\text{☞ } u_3 = 0_{\mathcal{L}(E)}$$

$$\text{☞ } u_4 : M_n(\mathbb{R}) \longrightarrow M_n(\mathbb{R})$$

$$A \longmapsto {}^t A$$

$$\text{☞ } u_5 \text{ un endomorphisme nilpotent de } E \text{ (on pourra noter au besoin } n = \dim E)$$

$$\text{☞ } u_6 \text{ un projecteur de } E$$

$$\text{☞ l'endomorphisme } u_7 \text{ de } \mathbb{R}^4 \text{ canoniquement associé à la matrice } A_7 = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}$$

- 2) **Existence d'un polynôme annulateur.** On rappelle que E est de dimension finie (notée n), et que u est un endomorphisme de E . Etablir que la famille $(\text{id}, u, u^2, \dots, u^{n^2})$ est liée.

En déduire l'existence d'un polynôme annulateur de u .

- 3) **Structure de l'ensemble des polynômes annulateurs.** Soit encore u un endomorphisme de E . On note I_u l'ensemble des polynômes annulateurs de u , soit :

$$I_u = \{P \in \mathbb{K}[X] \mid P(u) = 0_{\mathcal{L}(E)}\}$$

On peut observer que I_u contient au moins le polynôme nul $\widetilde{0}$.

- a) Montrer que si P et Q appartiennent à I_u , alors $(P + Q)$ appartient à I_u .
- b) Montrer que si P appartient à I_u , alors pour tout polynôme $R \in \mathbb{K}[X]$, (RP) appartient à I_u .

Ces propriétés permettent d'affirmer que I_u est un *idéal* de $\mathbb{K}[X]$. Comme il est bien connu qu'un idéal de $\mathbb{K}[X]$ est monogène, on peut affirmer qu'il existe un polynôme $P_u \in \mathbb{K}[X]$ tel que $I_u = P_u \mathbb{K}[X]$. Mieux, il existe un unique polynôme unitaire¹² annulateur de u noté π_u tel que :

$$I_u = \pi_u \mathbb{K}[X] \quad \text{soit encore} \quad I_u = \{\pi_u Q \mid Q \in \mathbb{K}[X]\}$$

Ce polynôme π_u , qui divise donc tout autre polynôme annulateur de u , est appelé **polynôme minimal** de l'endomorphisme u .

12. De coefficient dominant égal à 1.

Partie 2 — Valeurs propres et sous-espaces propres

- 4) Soit $\lambda \in \mathbb{K}$. Montrer que λ est une valeur propre de u si et seulement si $u - \lambda \text{id}$ n'est pas un automorphisme¹³.
- 5) Soit $\lambda \in \mathbb{K}$. On suppose que $u - \lambda \text{id}$ n'est pas un automorphisme. Montrer que les vecteurs propres associés à λ sont les vecteurs non nuls de $\text{Ker}(u - \lambda \text{id})$.

Notation : par la suite, si λ est une valeur propre de u , on notera $E_\lambda(u) = \text{Ker}(u - \lambda \text{id})$. Ce sev de E sera appelé **sous-espace propre** de u associé à la valeur propre λ .

- 6) Soit λ une valeur propre de u . Montrer que le sev $E_\lambda(u)$ est stable par u .
- 7) Soit $k \in \mathbb{N}^*$. Soient $\lambda_1, \dots, \lambda_k$ des valeurs propres de u , et $F_1, \dots, F_k$ des sev de $E_{\lambda_1}(u), \dots, E_{\lambda_k}(u)$ respectivement. Montrer que $F_1 + \dots + F_k$ est stable par u .

Partie 3 — Valeurs propres et polynômes annulateurs

- 8) Soient P un polynôme (dans $\mathbb{K}[X]$), et $\vec{v}$ un vecteur propre de u , associé à une valeur propre λ . Montrer que $[P(u)](\vec{v}) = P(\lambda) \cdot \vec{v}$.
- 9) Soit $P \in \mathbb{K}[X]$ un polynôme annulateur de u .¹⁴ Dédurre de la question précédente que $\text{Sp}(u) \subset \text{rac}(P)$, où $\text{rac}(P)$ désigne l'ensemble des racines de P .
- 10) On note π_u le polynôme minimal de u (défini à la fin de la partie I). Soit λ une racine de π_u . On écrit alors $\pi_u = (X - \lambda)Q$, avec $Q \in \mathbb{K}[X]$. En supposant que λ n'est pas valeur propre de u , contredire la minimalité de π_u .
- 11) En déduire que $\text{Sp}(u) = \text{rac}(\pi_u)$.
- 12) En déduire que u admet un nombre fini de valeurs propres.

Partie 4 — Diagonalisabilité

- 13) Soient P et Q deux polynômes premiers entre eux. En utilisant le théorème de Bezout, montrer que :

$$\text{Ker}((PQ)(u)) = \text{Ker}(P(u)) \oplus \text{Ker}(Q(u))$$

Plus généralement, on peut montrer (mais on ne demande pas de la faire ici) que si $P_1, \dots, P_k$ sont des polynômes deux à deux premiers entre eux, alors :

$$\text{Ker}((P_1 \cdots P_k)(u)) = \text{Ker}(P_1(u)) \oplus \cdots \oplus \text{Ker}(P_k(u))$$

- 14) On dit que u est **diagonalisable** s'il existe une base $\mathcal{B}$ de E telle que $\text{Mat}_{\mathcal{B}}(u)$ soit diagonale.
- a) Soient $\lambda_1, \dots, \lambda_k$ les valeurs propres (distinctes) de u . Montrer que la somme $S = E_{\lambda_1}(u) + \dots + E_{\lambda_k}(u)$ est directe, c'est à dire que : $E_{\lambda_i}(u) \cap E_{\lambda_j}(u) = \{\vec{0}\}$ si $i \neq j$.

Interpréter cette somme S comme le noyau de $P(u)$, où P est un polynôme que l'on explicitera.

- b) Montrer que u est diagonalisable si et seulement si : $E = \bigoplus_{\lambda \in \text{Sp}(u)} E_\lambda(u)$

13. Dans ce problème, on notera id l'identité de E .

14. C'est à dire tel que : $P(u) = 0_{\mathcal{L}(E)}$.

CORRIGÉ

Partie 1 — Polynômes annulateurs d'un endomorphisme

1) a) La matrice de u_1^2 est $A_1^2 = 4I_2$. Ainsi : $A_1^2 - 4I_2 = 0$ d'où $u_1^2 - 4\text{id} = 0$.

Donc $X^2 - 4$ est un polynôme annulateur de u_1 .

b) Quelques exemples de polynômes annulateurs :

☞ $u_2 = \text{id}$ a pour polynôme annulateur $X - 1$

☞ $u_3 = 0_{\mathcal{L}(E)}$ a pour polynôme annulateur X

☞ u_4 a pour polynôme annulateur $X^2 - 1$, puisque u_4 est une involution ($u_4^2 = \text{id}$).

☞ u_5 a pour polynôme annulateur X^n , puisque tout endomorphisme nilpotent u d'un espace vectoriel E de dimension finie n vérifie $u^n = 0$.

☞ u_6 a pour polynôme annulateur $X^2 - X$, puisque c'est un projecteur de E , et à ce titre : $u_6^2 = u_6$.

☞ l'endomorphisme u_7 a pour polynôme annulateur $X^2 - 4X$, puisque la matrice de u_7^2 est $A_7^2 = 4A_7$ (donc $u_7^2 = 4u_7$).

2) Puisque E est de dimension finie n , l'espace vectoriel $\mathcal{L}(E)$ est de dimension finie égale à n^2 . La famille $(\text{id}, u, u^2, \dots, u^{n^2})$ étant de cardinal $n^2 + 1$ dans un ev de dimension n^2 , elle est liée. Il existe donc $n^2 + 1$ scalaires $\alpha_0, \dots, \alpha_{n^2}$ non tous nuls tels que :

$$\sum_{k=0}^{n^2} \alpha_k u^k = 0 \quad \text{d'où : } P(u) = 0 \text{ avec } P = \sum_{k=0}^{n^2} \alpha_k X^k.$$

Le polynôme P est donc un polynôme annulateur de u (et $P \neq 0$).

Conclusion. Tout endomorphisme d'un ev de dimension finie admet un polynôme annulateur (non nul).

3) a) Soient P et Q dans I_u . Alors : $(P + Q)(u) = P(u) + Q(u) = 0$. D'où $(P + Q)$ appartient à I_u .

Conclusion. $\forall (P, Q) \in I_u^2, (P + Q) \in I_u$.

b) Soient P dans I_u et $R \in \mathbb{K}[X]$. Alors : $(RP)(u) = R(u)P(u) = R(u)0 = 0$. D'où (RP) appartient à I_u .

Conclusion. $\forall P \in I_u, \forall R \in \mathbb{K}[X], (RP) \in I_u$.

Partie 2 — Valeurs propres et sous-espaces propres

4) Si λ est une valeur propre de u , alors il existe un vecteur $\vec{v}$ non nul tel que : $u(\vec{v}) = \lambda \vec{v}$. Ainsi : $(u - \lambda \text{id})(\vec{v}) = 0$. Donc l'endomorphisme $u - \lambda \text{id}$ n'est pas injectif. Puisque E est de dimension finie, il est équivalent de dire que $u - \lambda \text{id}$ n'est pas un automorphisme.

Réciproquement, si $u - \lambda \text{id}$ n'est pas un automorphisme, alors (E étant de dimension finie) il n'est pas injectif. Donc il existe un vecteur $\vec{v}$ non nul tel que $(u - \lambda \text{id})(\vec{v}) = 0$, donc il existe un vecteur $\vec{v}$ non nul tel que : $u(\vec{v}) = \lambda \vec{v}$. Ainsi λ est une valeur propre de u .

Conclusion. $[\lambda \in \text{Sp}(u)] \iff [(u - \lambda \text{id}) \notin \text{GL}(E)]$

5) Soit $\vec{v}$ un vecteur propre associé à la valeur propre λ . Alors $u(\vec{v}) = \lambda \vec{v}$ donc $\vec{v} \in \text{Ker}(u - \lambda \text{id})$. Réciproquement, tout vecteur non nul de $\text{Ker}(u - \lambda \text{id})$ est clairement un vecteur propre pour la valeur propre λ .

Conclusion. Les vecteurs propres associés à λ sont les vecteurs non nuls de $\text{Ker}(u - \lambda \text{id})$.

6) Soit $\vec{v} \in E_\lambda(u)$. Alors : $u(\vec{v}) = \lambda \vec{v}$. D'où : $u(u(\vec{v})) = u(\lambda \vec{v}) = \lambda u(\vec{v})$. En résumé : si $\vec{v} \in E_\lambda(u)$, alors : $u(u(\vec{v})) = \lambda u(\vec{v})$ c'est à dire : $u(\vec{v}) \in E_\lambda(u)$.

Conclusion. Les sous-espaces propres de u sont stables par u : $\forall \lambda \in \text{Sp}(u), u(E_\lambda(u)) \subset E_\lambda(u)$

7) En évitant une récurrence qu'il faudrait écrire si l'on était vraiment des gens bien : soient $\vec{v}_1, \dots, \vec{v}_k$ des vecteurs propres pour les valeurs propres $\lambda_1, \dots, \lambda_k$ respectivement. On a :

$$u\left(u\left(\sum_{i=1}^k \vec{v}_i\right)\right) = u\left(\sum_{i=1}^k u(\vec{v}_i)\right) = u\left(\sum_{i=1}^k \lambda_i \vec{v}_i\right) = \sum_{i=1}^k \lambda_i u(\vec{v}_i)$$

Or pour tout entier i compris entre 1 et k , on a $u(\vec{v}_i) \in F_i$ (question précédente), donc $\lambda_i u(\vec{v}_i) \in F_i$ (puisque F_i est un sev de E).

De cette observation et de l'égalité précédente, on déduit que : $u\left(u\left(\sum_{i=1}^k \vec{v}_i\right)\right) \in F_1 + \dots + F_k$.

En résumé : $\forall \vec{V} \in F_1 + \dots + F_k, u(\vec{V}) \in F_1 + \dots + F_k$.

Conclusion. $\forall k \in \mathbb{N}, \forall (\lambda_1, \dots, \lambda_k) \in [\text{Sp}(u)]^k, u(E_{\lambda_1}(u) + \dots + E_{\lambda_k}(u)) \subset E_{\lambda_1}(u) + \dots + E_{\lambda_k}(u)$.

Partie 3 — Valeurs propres et polynômes annulateurs

8) Soit P un polynôme. Il existe un entier naturel n et $n+1$ scalaires $\alpha_0, \dots, \alpha_n$ tels que : $P = \sum_{k=0}^n \alpha_k X^k$.

Soit $\vec{v}$ un vecteur propre de u , associé à une valeur propre λ . Alors :

$$[P(u)](\vec{v}) = \left[\sum_{k=0}^n \alpha_k u^k \right](\vec{v}) = \sum_{k=0}^n \alpha_k u^k(\vec{v}) = \sum_{k=0}^n \alpha_k \lambda^k \vec{v} = \left(\sum_{k=0}^n \alpha_k \lambda^k \right) \vec{v} = P(\lambda) \vec{v}$$

Le point-clef étant que dans la troisième égalité, on utilise la propriété : $\forall k \in \mathbb{N}, u^k(\vec{v}) = \lambda^k \vec{v}$ qui provient du fait que $\vec{v}$ est un vecteur propre pour la valeur propre λ et d'une récurrence immédiate sur k .

Conclusion. $\forall \lambda \in \text{Sp}(u), \forall \vec{v} \in E_\lambda(u), \forall P \in \mathbb{K}[X], [P(u)](\vec{v}) = P(\lambda) \vec{v}$.

9) Soit $\lambda \in \text{Sp}(u)$, et $\vec{v}$ un vecteur propre non nul associé à la valeur propre λ . Supposons en outre que P soit un polynôme annulateur de u . Alors : $P(u) = 0$ donc d'après la question précédente, on a : $P(\lambda) \vec{v} = 0$. Puisque $\vec{v}$ est non nul, on en déduit que $P(\lambda) = 0$, donc λ est racine de P .

Conclusion. $\text{Sp}(u) \subset \text{rac}(P)$ pour tout polynôme P annulateur de u .

10) Soit π_u le polynôme minimal de u , et soit λ une racine de π_u . On peut factoriser le polynôme minimal et écrire $\pi_u = (X - \lambda)Q$, avec $Q \in \mathbb{K}[X]$.

Supposons que λ n'est pas valeur propre de u . Alors pour tout vecteur $\vec{v}$ dans E , on a : $(u - \lambda \text{id})(Q(u)(\vec{v})) = 0$. Puisque λ n'est pas valeur propre de u , alors $(u - \lambda \text{id}) \in \text{GL}(E)$, et on déduit donc de ce fait et de l'égalité précédente que : $\forall \vec{v} \in E, Q(u)(\vec{v}) = 0$. Ainsi : $Q(u) = 0$, et Q est donc un polynôme annulateur de u , de degré strictement inférieur à celui du polynôme minimal de u : contradiction.

Conclusion. Si λ est racine de π_u , alors λ est une valeur propre de u .

11) D'après la question précédente : $\text{rac}(\pi_u) \subset \text{Sp}(u)$. Et d'après la question 9, comme π_u est en particulier annulateur de u , on a : $\text{Sp}(u) \subset \text{rac}(\pi_u)$. **Conclusion.** $\forall u \in \mathcal{L}(E), \text{Sp}(u) = \text{rac}(\pi_u)$.

12) D'après la question précédente, le spectre de u a pour cardinal le nombre de racines de π_u . C'est donc un ensemble fini. **Conclusion.** u possède un nombre fini de valeurs propres.

Partie 4 — Diagonalisabilité

13) Soient P et Q deux polynômes premiers entre eux. Alors il existe deux polynômes A et B tels que : $AP + BQ = 1$.

Soit $\vec{v}$ un vecteur de E : $\vec{v} = \text{id}(\vec{v}) = [(AP + BQ)(u)](\vec{v}) = [(AP)(u)](\vec{v}) + [(BQ)(u)](\vec{v})$.

Supposons à présent que : $\vec{v} \in \text{Ker}((PQ)(u))$. Alors :

$$Q([(AP)(u)](\vec{v})) = [(APQ)(u)](\vec{v}) = A(u)[(PQ)(u)](\vec{v})$$

D'où : $Q([(AP)(u)](\vec{v})) = 0$ puisque $\vec{v} \in \text{Ker}((PQ)(u))$. Donc : $[(AP)(u)](\vec{v}) \in \text{Ker}(Q(u))$.

De façon analogue : $[(BQ)(u)](\vec{v}) \in \text{Ker}(P(u))$. On a donc :

$$\boxed{\text{Ker}(PQ(u)) \subset \text{Ker}(P(u)) + \text{Ker}(Q(u)) \quad (\spadesuit)}$$

Réciproquement, si $\vec{v} \in \text{Ker}(P(u)) + \text{Ker}(Q(u))$ alors évidemment ¹⁵ $\vec{v} \in \text{Ker}(PQ(u))$. Conjuguée à l'inclusion précédente, cette inclusion fournit l'égalité : $\boxed{\text{Ker}(PQ(u)) = \text{Ker}(P(u)) + \text{Ker}(Q(u)) \quad (\heartsuit)}$

Il reste à voir que la somme est directe. Soit $\vec{v} \in \text{Ker}(P(u)) \cap \text{Ker}(Q(u))$. Alors d'après la relation de Bezout : $\vec{v} = \underbrace{[(AP)(u)](\vec{v})}_{=0 \text{ car } \vec{v} \in \text{Ker}(P(u))} + \underbrace{[(BQ)(u)](\vec{v})}_{=0 \text{ car } \vec{v} \in \text{Ker}(Q(u))}$. D'où $\vec{v} = 0$. Par suite : $\text{Ker}(P(u)) \cap \text{Ker}(Q(u)) = \{0\}$.

Conclusion. $\text{Ker}((PQ)(u)) = \text{Ker}(P(u)) \oplus \text{Ker}(Q(u))$.

14) a) Soient λ_i et λ_j deux valeurs propres distinctes de u , et soit $\vec{v} \in E_{\lambda_i}(u) \cap E_{\lambda_j}(u)$. Alors : $u(\vec{v}) = \lambda_i \vec{v} = \lambda_j \vec{v}$. D'où : $(\lambda_i - \lambda_j) \vec{v} = 0$. Puisque $\lambda_i \neq \lambda_j$ par hypothèse, on en déduit que : $\vec{v} = 0$.

Conclusion. $E_{\lambda_i}(u) \cap E_{\lambda_j}(u) = \{\vec{0}\}$ (si $i \neq j$).

On a : $S = E_{\lambda_1}(u) + \dots + E_{\lambda_k}(u) = \text{Ker}[P(u)]$, avec : $P = \prod_{i=1}^k (X - \lambda_i)$ (d'après la question 13).

b) Supposons que u soit diagonalisable. Alors il existe une base $\mathcal{B}$ de E telle que $\text{Mat}_{\mathcal{B}}(u)$ soit diagonale (égale à D). Notons $\lambda_1, \dots, \lambda_k$ les coefficients diagonaux de cette matrice, et notons m_i le nombre d'occurrences de la valeur propre λ_i sur la diagonale (les coefficients diagonaux sont nécessairement les valeurs propres de u , puisque $D - \lambda_i I$ est non inversible). Notons encore $C_{1,1}, \dots, C_{1,m_1}$ les m_1 -èmes premières colonnes de D . Ces colonnes correspondent aux images de m_1 vecteurs $\vec{v}_{1,1}, \dots, \vec{v}_{1,m_1}$ formant une base de $E_{\lambda_1}(u)$. On a alors : $E_{\lambda_1}(u) = \text{Vect}(\vec{v}_{1,1}, \dots, \vec{v}_{1,m_1})$, et plus généralement : $E_{\lambda_i}(u) = \text{Vect}(\vec{v}_{i,1}, \dots, \vec{v}_{i,m_i})$ pour tout i entre 1 et k . En regroupant tous ces vecteurs, on obtient alors la base $\mathcal{B}$, et on a donc : $E = \bigoplus_{\lambda \in \text{Sp}(u)} E_{\lambda}(u)$.

Réciproquement, si $E = \bigoplus_{\lambda \in \text{Sp}(u)} E_{\lambda}(u)$, alors pour tout $\lambda \in \text{Sp}(u)$, le sous-espace propre $E_{\lambda}(u)$ admet une base $\mathcal{B}_{\lambda}$. La réunion de ces bases lorsque λ décrit $\text{Sp}(u)$ constitue alors une base de E dans laquelle $\text{Mat}_{\mathcal{B}}(u)$ est diagonale.

Conclusion. $[u \text{ est diagonalisable}] \iff \left[E = \bigoplus_{\lambda \in \text{Sp}(u)} E_{\lambda}(u) \right]$.

15. Pour de vrai !

Chapitre 25

Séries numériques

Tout au long de ce chapitre, $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$.

25.1 Généralités

DÉFINITION 25.1 - Soit $(u_n)_{n \in \mathbb{N}}$ une suite à valeurs dans $\mathbb{K}$. Pour tout entier naturel N , on appelle **somme partielle de rang N** (de (u_n)) le scalaire : $S_N = \sum_{n=0}^N u_n$.

Exemples : dans le cas où la suite (u_n) est télescopique, ie s'il existe une suite (v_n) telle que pour tout entier n on a $u_n = v_{n+1} - v_n$, alors pour tout entier naturel N on a : $S_N = \sum_{n=0}^N u_n = \sum_{n=0}^N (v_{n+1} - v_n) = v_{N+1} - v_0$.

Et dans celui où la suite (u_n) est géométrique, ie s'il existe un scalaire $q \neq 1$ tel que pour tout entier n on a $u_n = u_0 q^n$, alors pour tout entier naturel N on a : $S_N = \sum_{n=0}^N u_0 q^n = u_0 \times \frac{1 - q^{N+1}}{1 - q}$.

Observation banale, mais utile. Pour tout entier $n \in \mathbb{N}^*$ on a : $u_n = S_n - S_{n-1}$.

DÉFINITION 25.2 - Soit $(u_n)_{n \in \mathbb{N}}$ une suite à valeurs dans $\mathbb{K}$. On appelle **série de terme général u_n** la suite des sommes partielles $(S_N)_N$ de la suite (u_n) .

Une série numérique n'est donc qu'un cas particulier de suite.¹

Notation. Soit $(u_n)_{n \in \mathbb{N}}$ une suite à valeurs dans $\mathbb{K}$. On pourra noter $\sum u_n$ la série de terme général u_n .

DÉFINITION 25.3 - Soit $(u_n)_{n \in \mathbb{N}}$ une suite à valeurs dans $\mathbb{K}$. La série $\sum u_n$ est **convergente** (resp. **divergente**) lorsque la suite de ses sommes partielles $(S_N)_N$ est convergente (resp. divergente).

1. Cette remarque n'étant pas faite péjorativement, mais plutôt dans l'optique de vous rassurer.

On obtient comme conséquence immédiate de la propriété de convergence des suites extraites l'énoncé ci-dessous.

PROPRIÉTÉ 25.1 - Si la série $\sum u_n$ converge, alors $\lim_{n \rightarrow +\infty} u_n = 0$.

PREUVE. Avec les notations précédemment introduites, si la série de terme général u_n converge, alors il existe un réel ℓ telle que la suite des sommes partielles (S_n) converge vers ℓ . Il s'ensuit que $S_n - S_{n-1}$ tend vers 0 lorsque n tend vers $+\infty$, d'où la conclusion puisque : $S_n - S_{n-1} = u_n$. ☕

Remarque. La réciproque est évidemment FAUSSE. Il se peut très bien que $\lim_{n \rightarrow +\infty} u_n = 0$, et que la série u_n soit divergente. L'exemple d'illustration le plus fameux de cette situation est fourni par la série harmonique $\sum \frac{1}{n}$, puisque la somme partielle de rang N associée est minorée par une expression tendant vers $+\infty$, explicitement : $S_N = \sum_{n=1}^N \frac{1}{n} \geq \ln(N+1)$.

La contraposée de la propriété précédente donne naissance à la notion de divergence grossière.

PROPRIÉTÉ 25.2 - Si u_n ne tend pas vers 0, alors la série $\sum u_n$ diverge. On parle dans ce cas de **divergence grossière**.

PREUVE. C'est la contraposée de l'assertion énoncée dans la propriété 25.1. ☕

Remarque. Ultime observation générale sur les séries : la notion de convergence d'une série $(\sum u_n)$ est une notion *asymptotique*, et ne dépend pas des premiers termes de (u_n) . En clair, on ne modifie pas la nature (convergente ou divergente) d'une série en modifiant un nombre fini de termes de la suite (u_n) .

PROPRIÉTÉ 25.3 - (**Linéarité de la convergence**). Soient $\sum u_n$ et $\sum v_n$ deux séries.

Si $\sum u_n$ et $\sum v_n$ sont convergentes, alors pour tout couple (λ, μ) de scalaires, la série $\sum (\lambda u_n + \mu v_n)$ est convergente, et :

$$\sum_{n=0}^{+\infty} (\lambda u_n + \mu v_n) = \lambda \sum_{n=0}^{+\infty} u_n + \mu \sum_{n=0}^{+\infty} v_n$$

PREUVE. Immédiate. ☕

Remarque. On déduit de la propriété précédente que la somme d'une série convergente et d'une divergente est divergente. On ne peut en revanche rien dire en général de la somme de deux séries divergentes.²

2. L'expression "somme de deux séries" est un abus de langage, synonyme dans ces lignes de séries dont le terme général est la somme. . .

25.2 Premières séries de référence

PROPRIÉTÉ 25.4 - (Séries géométriques). Soit $q \in \mathbb{C}$.

La série $\sum q^n$ converge si et seulement si $|q| < 1$.

Dans ce cas : $\sum_{n \in \mathbb{N}} q^n = \frac{1}{1-q}$.

PREUVE. Lorsque q est de module supérieur ou égal à 1, la série de terme général q^n diverge grossièrement.

Lorsque $|q| < 1$, on a : $\sum_{n=0}^N q^n = \frac{1-q^{N+1}}{1-q}$. Puisque : $\lim_{N \rightarrow +\infty} q^{N+1} = 0$, on en déduit que la série de terme général q^n converge, et que : $\sum_{n=0}^{+\infty} q^n = \frac{1}{1-q}$. ☕

Exemple. La série $\sum_{n \in \mathbb{N}} \frac{1}{2^n}$ est convergente, et : $\sum_{n \in \mathbb{N}} \frac{1}{2^n} = 2$.

PROPRIÉTÉ 25.5 - (Séries télescopiques). Soit $(u_n) \in \mathbb{C}^{\mathbb{N}}$.

La série $\sum (u_{n+1} - u_n)$ converge si et seulement si la suite $(u_n)_n$ converge.

Dans ce cas : $\sum_{n \in \mathbb{N}} (u_{n+1} - u_n) = \ell - u_0$, où l'on a noté $\ell = \lim_{n \rightarrow +\infty} u_n$.

On a aussi, pour tout $n_0 \in \mathbb{N}$: $\sum_{n=n_0}^{+\infty} (u_{n+1} - u_n) = \ell - u_{n_0}$.

PREUVE. La suite (S_N) des sommes partielles de la série $\sum (u_{n+1} - u_n)$ a pour terme général $S_N = u_{N+1} - u_0$. Celle-ci converge si et seulement si u_{N+1} admet une limite finie lorsque N tend vers $+\infty$, c'est-à-dire si et seulement si la suite $(u_n)_n$ converge. ☕

Exemple. Considérons la série de terme général $\frac{1}{n(n-1)}$.

On a pour tout entier naturel $n \geq 2$: $\frac{1}{n(n-1)} = \frac{1}{n-1} - \frac{1}{n}$.

Il s'ensuit que : $\forall N \geq 2, \quad S_N = \sum_{n=2}^N \frac{1}{n(n-1)} = \sum_{n=2}^N \left(\frac{1}{n-1} - \frac{1}{n} \right) = 1 - \frac{1}{N}$.

On en déduit que la série converge, et que : $\sum_{n=2}^{+\infty} \frac{1}{n(n-1)} = 1$.

Place maintenant à un critère d'importance CRUCIALE en pratique.

PROPRIÉTÉ 25.6 - (Séries de Riemann). Soit $\alpha \in \mathbb{R}$.

La série $\sum \frac{1}{n^\alpha}$ converge si et seulement si $\alpha > 1$.

PREUVE. Cette propriété a déjà été établie comme conséquence du théorème des accroissements finis (corollaire 17.8, page 444). Nous en verrons une autre démonstration, plus courte, au cours du présent chapitre. ☕

Exemple. La série de terme général $u_n = \frac{1}{n\sqrt{n}}$ est convergente, puisque $u_n = \frac{1}{n^{3/2}}$ et que $3/2 > 1$.

25.3 Séries à termes positifs

Dans ce paragraphe, les séries considérées ont des termes généraux dans $\mathbb{R}_+$. Dans ce contexte, les suites des sommes partielles seront positives et croissantes.

PROPRIÉTÉ 25.7 - (Comparaison). Soient (u_n) et (v_n) deux suites réelles positives telles que $u_n \leq v_n$ à partir d'un certain rang.

Si la série $\sum v_n$ converge, alors la série $\sum u_n$ converge.

Si la série $\sum u_n$ diverge, alors la série $\sum v_n$ diverge.

PREUVE. Notons (S_N) et (T_N) les suites de sommes partielles associées aux suites (u_n) et (v_n) respectivement. Par hypothèse, (S_N) et (T_N) sont deux suites réelles croissantes, et quitte à tronquer un nombre fini de termes on peut supposer que $S_N \leq T_N$.

Si la série $\sum u_n$ diverge, alors la suite (S_N) tend vers $+\infty$, et on en déduit par comparaison (sur les suites) que (T_N) tend vers $+\infty$. Par conséquent la série de terme général v_n diverge.

Si la série $\sum v_n$ converge, alors la suite (T_N) converge. En particulier elle est bornée, et encore plus particulièrement elle est majorée. Puisque $S_N \leq T_N$, on en déduit que (S_N) est majorée ; puisqu'en outre elle est croissante, le théorème de la limite monotone permet d'affirmer qu'elle converge. Ce qui signifie que la série $\sum u_n$ est convergente. ☕

Exemple 1. Pour tout entier naturel $n \geq 2$, on a : $\frac{1}{n^2} \leq \frac{1}{n(n-1)}$. Puisque les séries de termes généraux $\frac{1}{n^2}$ et $\frac{1}{n(n-1)}$ sont à termes positifs, et que la série de terme général $\frac{1}{n(n-1)}$ converge (voir plus haut), on en déduit par comparaison que la série de terme général $\frac{1}{n^2}$ converge (sans avoir recours au critère de Riemann... mais c'est plus long!).

Exemple 2. Pour tout entier $n \geq 2$, on a : $\frac{1}{\ln(n)} \geq \frac{1}{n}$. Puisque les séries de termes généraux $\frac{1}{\ln(n)}$ et $\frac{1}{n}$ sont à termes positifs, et que la série de terme général $1/n$ diverge (série harmonique), on en déduit que la série $\sum \frac{1}{\ln(n)}$ diverge.

PROPRIÉTÉ 25.8 - (Equivalents). Soient (u_n) et (v_n) deux suites réelles positives.

Si $u_n \sim_{+\infty} v_n$, alors les séries $\sum u_n$ et $\sum v_n$ sont de même nature.

PREUVE. Puisque $u_n \sim_{+\infty} v_n$, il existe une suite (φ_n) qui tend vers 1 telle que : $v_n = u_n \varphi_n$. Or, (φ_n) étant de limite 1, on a : $\frac{1}{2} \leq \varphi_n \leq \frac{3}{2}$ à partir d'un certain rang.

On en déduit que $\frac{1}{2} u_n \leq v_n \leq \frac{3}{2} u_n$ à partir d'un certain rang.

Si la série $\sum u_n$ converge, alors la série de terme général $\frac{3}{2} u_n$ converge (d'après la propriété 25.3). Puisque les séries considérées sont à termes positifs, on en déduit par comparaison que la série de terme général v_n converge.

Inversement, si la série $\sum v_n$ diverge, alors la série de terme général $\frac{3}{2} u_n$ diverge (par comparaison). On en déduit que la série de terme général u_n diverge. ☕

Exemple 1. La série $\sum \sin\left(\frac{1}{n}\right)$ est divergente.

Exemple 2. La série $\sum \ln\left(1 + \frac{1}{n^2}\right)$ est convergente.

Exemple 3. La série $\sum n^\alpha \left(\sqrt{1 + \frac{1}{n}} - 1\right)$ est convergente SSI $\alpha < 0$.

Exemple 4. La série $\sum \frac{\text{ch}(n)}{\text{ch}(2n)}$ est convergente, puisque $\frac{\text{ch}(n)}{\text{ch}(2n)} \sim_{+\infty} e^{-n}$ (et que e^{-n} est le terme général d'une série géométrique convergente).

Exemple 5. La série de terme général $\frac{\ln(n+1) - \ln(n)}{\sqrt{n}}$ est convergente.

Exemple 6. La série de terme général $n \left(1 - \cos\left(\frac{1}{n}\right)\right)$ est divergente.

PROPRIÉTÉ 25.9 - (Négligeabilité). Soient (u_n) et (v_n) deux suites réelles positives.

On suppose que $u_n = o(v_n)$. Si $\sum v_n$ converge, alors $\sum u_n$ converge.

PREUVE. Par hypothèse, on a : $0 \leq u_n \leq \frac{1}{2} v_n$ à partir d'un certain rang. On déduit la propriété de cet encadrement et du critère de comparaison. ☕

Remarque. Evidemment, l'énoncé précédent peut encore être lu "dans l'autre sens" : si $u_n = o(v_n)$ et $\sum u_n$ diverge alors $\sum v_n$ est divergente.

Ce critère de négligeabilité induit une méthode bien pratique pour montrer la convergence d'une série.

COROLLAIRE 25.1 - Soit (u_n) une suite réelle positive.

Si $\lim_{n \rightarrow +\infty} n^2 u_n = 0$, alors la série $\sum u_n$ converge.

Plus généralement d'ailleurs, si $\lim_{n \rightarrow +\infty} n^\alpha u_n = 0$ pour un réel $\alpha > 1$, alors la série $\sum u_n$ converge.

PREUVE. C'est une conséquence immédiate de la propriété précédente et du critère de convergence pour les séries de Riemann, puisque les hypothèses impliquent que u_n est négligeable devant $1/n^2$ (ou devant $1/n^\alpha$ avec $\alpha > 1$). ☕

Exemple. La série $\sum ne^{-n}$ est convergente, puisque $n^2 \times (ne^{-n}) = n^3 e^{-n}$ a pour limite 0 en $+\infty$ (croissances comparées).

PROPRIÉTÉ 25.10 - (Domination). Soient (u_n) et (v_n) deux suites réelles positives.

On suppose que $u_n = O(v_n)$. Si $\sum v_n$ converge, alors $\sum u_n$ converge.

PREUVE. Par hypothèse, il existe un réel positif M tel que : $0 \leq u_n \leq M \times v_n$ à partir d'un certain rang. On déduit la propriété de cet encadrement et du critère de comparaison. ☕

PROPRIÉTÉ 25.11 - (Comparaison série/intégrale). Soit f une fonction à valeurs réelles, positive et décroissante sur $\mathbb{R}_+$.

La série $\sum f(n)$ et l'intégrale $\int_a^{+\infty} f$ sont de même nature (avec a réel positif quelconque).

PREUVE. Sous les hypothèses de l'énoncé, on a :

$$\forall n \in \mathbb{N}, f(n+1) \leq f(x) \leq f(n) \quad \text{d'où} \quad \forall n \in \mathbb{N}, f(n+1) \leq \int_n^{n+1} f \leq f(n)$$

Par sommation des encadrements précédents entre 0 et N on obtient : $\forall N \in \mathbb{N}, \sum_{n=0}^N f(n+1) \leq \int_0^N f \leq$

$$\sum_{n=0}^N f(n)$$

Supposons que la série $\sum f(n)$ converge. Alors la fonction $x \mapsto \int_0^x f$ est croissante (f étant positive par hypothèse) et majorée, donc convergente ; en d'autres termes, l'intégrale $\int_0^{+\infty} f$ est convergente.

Inversement si la série $\sum f(n)$ diverge, alors la série $\sum f(n+1)$ diverge également, donc la fonction $x \mapsto \int_0^x f$ est non majorée ; en d'autres termes, l'intégrale $\int_0^{+\infty} f$ est divergente.

Conclusion. Soit f une fonction à valeurs réelles, positive et décroissante sur $\mathbb{R}_+$. La série $\sum f(n)$ et l'intégrale $\int_0^{+\infty} f$ sont de même nature. ☕

Exemple. Etudions la série $\sum \frac{1}{n \ln(n)}$. Celle-ci est à termes positifs et décroissante.

En outre pour tout réel $A \geq 2$ on a : $\int_2^A \frac{dx}{x \ln(x)} = \ln(\ln(A)) - \ln(\ln(2)) \xrightarrow{A \rightarrow +\infty} +\infty$.

On en déduit que l'intégrale $\int_2^{+\infty} \frac{dx}{x \ln(x)}$ est divergente. D'après le critère de comparaison série/intégrale, on peut conclure.

Conclusion. La série $\sum \frac{1}{n \ln(n)}$ est divergente.

25.4 Convergence absolue

Dans ce paragraphe, on revient au cas général, en considérant des suites (u_n) à valeurs dans $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$.

DÉFINITION 25.4 - Soit $(u_n)_{n \in \mathbb{N}}$ une suite à valeurs dans $\mathbb{K}$.

La série $\sum u_n$ est **absolument convergente** si la série $\sum |u_n|$ converge.

Exemple. Pour tout réel θ , et pour tout réel $\alpha > 1$, la série $\sum \frac{e^{in\theta}}{n^\alpha}$ est absolument convergente.

PROPRIÉTÉ 25.12 - (“ACv $\implies$ Cv”). Soit $(u_n)_{n \in \mathbb{N}}$ une suite à valeurs dans $\mathbb{K}$.

Si la série $\sum u_n$ est absolument convergente, alors elle est convergente.

PREUVE. Soit (u_n) une suite à valeurs dans $\mathbb{C}$, telle que la série $\sum |u_n|$ converge.

➤ **1er cas — Si (u_n) est réelle positive.** Alors $\sum u_n$ converge, puisque $|u_n| = u_n \dots^3$.

➤ **2ème cas — Si (u_n) est réelle.** Posons pour tout entier n :

$$u_n^+ = \max(0, u_n) \quad \text{et} \quad u_n^- = \min(0, u_n)$$

Pour tout entier n , on a :

$$u_n = u_n^+ + u_n^- \quad \text{et} \quad |u_n| = u_n^+ - u_n^-$$

La série de terme général u_n^+ converge (hypothèse + positivité + comparaison), donc la série de terme général u_n^- converge (ce qui précède + linéarité). On en déduit que la série de terme général u_n converge.

3. Non, j'insiste, pas d'applaudissements

➤ **3ème cas — Si (u_n) est complexe.** Pour tout entier n , on a :

$$|\operatorname{Re}(u_n)| \leq |u_n| \quad \text{et} \quad |\operatorname{Im}(u_n)| \leq |u_n|$$

Puisque la série de terme général $|u_n|$ converge, on en déduit par comparaison que les séries de termes généraux respectifs $|\operatorname{Re}(u_n)|$ et $|\operatorname{Im}(u_n)|$ convergent, puis (d'après le second cas) que les séries de termes généraux respectifs $\operatorname{Re}(u_n)$ et $\operatorname{Im}(u_n)$ convergent. ☕

Remarque 1. Cette propriété est d'une grande importance pratique, puisqu'elle permet de ramener l'étude des séries en général à celle des séries à termes positifs (pour lesquelles on dispose d'un “arsenal” impressionnant, cf paragraphe précédent).

Remarque 2. La remarque précédente a cependant ses limites, dans le sens où la réciproque de la propriété est fausse. Il existe en effet des séries qui sont convergentes, mais non absolument convergentes ; de telles séries sont dites **semi-convergentes**.

L'exemple le plus célèbre de série semi-convergente est la série $\sum_{n \in \mathbb{N}^*} \frac{(-1)^{n+1}}{n}$: en effet, celle-ci converge et a pour somme $\ln 2$ (conséquence de Taylor avec reste intégrale), mais elle n'est évidemment pas absolument convergente (la série des modules est la série harmonique).

25.5 Epilogue

25.5.1 Stratégie d'étude : le point sur ce qu'il faut retenir sur les séries

PRÉSENTATION — *S'il y a vraiment un chapitre à maîtriser avant d'arriver en Spé⁴, c'est bien celui portant sur les séries numériques. Les exercices ci-dessous à la frontière entre Sup et Spé, ne demandent bien entendu que des connaissances de Sup (celles que nous avons vues en cours donc), mais vous reverrez probablement l'an prochain un certain nombre des situations présentées dans cette feuille.*

LE POINT SUR LES SÉRIES EN SUP — STRATÉGIE D'ÉTUDE

Pour commencer, gardez présent à l'esprit qu'une série numérique n'est finalement rien d'autre qu'une suite (une suite de sommes partielles). Or dans le contexte général des suites, on peut se poser essentiellement deux questions : celle de sa nature (convergente ou divergente), et lorsqu'elle converge celle de sa limite.⁵

4. En dehors de celui sur la trigonométrie, celui sur les nombres complexes, celui sur les fonctions de référence, celui sur les suites réelles ou complexes, celui sur les polynômes, celui sur le calcul matriciel, celui sur la continuité, celui sur la dérivabilité et celui sur ses applications, celui sur les équations différentielles, celui sur l'analyse asymptotique et en particulier les développements limités, ceux sur le calcul intégral, ceux d'algèbre linéaire et si vous vous destinez à une MP, celui sur les groupes et le chapitre d'arithmétique.

5. On pourrait en rajouter une troisième qui consiste à savoir, dans le cas où la suite diverge vers $+\infty$, “comment” diverge t-elle vers $+\infty$? A quelle vitesse ? Comme $\ln(n)$ ou comme n^3 ou comme e^n ? ...

Les exercices de cette feuille tourneront donc autour des deux problèmes suivants : étant donnée une série numérique $\sum u_n$,

Question 1 — quelle est la nature (convergente ou divergente) de la série $\sum u_n$?

Question 2 — si la série $\sum u_n$ converge, que vaut sa somme ?

► Un algorithme pour répondre à la question 1

Etape 1 : a-t-on $\lim_{n \rightarrow +\infty} u_n = 0$? Si oui, passer à l'étape 2. Sinon, la série **diverge grossièrement** (et c'est fini).

Etape 2 : la suite (u_n) est-elle à termes positifs ? Si oui, passer à l'étape 3 ; sinon passer à l'étape 4.

Etape 3 : essayer de déterminer la nature de la série en utilisant l'un des critères suivants, valides pour les séries à termes REELS POSITIFS :

→ Critère de **comparaison** : si $u_n \leq v_n$, $[\sum v_n \text{ converge}] \implies [\sum u_n \text{ converge}]$

→ Critère des **équivalents** : si $u_n \sim_{+\infty} v_n$, $\sum u_n$ et $\sum v_n$ sont de même nature

→ Critère des “o” ou des “O” : si $u_n = o(v_n)$ (ou $u_n = O(v_n)$), $[\sum v_n \text{ converge}] \implies [\sum u_n \text{ converge}]$

→ “**Règle de Riemann**” : si $\exists \alpha > 1$ tq $\lim_{n \rightarrow +\infty} n^\alpha u_n = 0$, alors $[\sum u_n \text{ converge}]$

→ Critère de **comparaison série-intégrale** :

si $f \geq 0$, $\mathcal{C}^0$ et $\searrow$, $\sum f(n)$ et $\int_a^{+\infty} f$ sont de même nature

où u_n est le terme général de la série étudiée, et v_n est le terme général d'une série de référence parmi les suivantes :

→ Série de **Riemann** : $[\sum \frac{1}{n^\alpha} \text{ converge}] \iff [\alpha > 1]$

→ Série **télescopique** : $[\sum (w_{n+1} - w_n) \text{ converge}] \iff [(w_n) \text{ converge}]$

→ Série **géométrique** : soit $q \in \mathbb{C}$. $[\sum q^n \text{ converge}] \iff [|q| < 1]$

Etape 4 : on considère la série de terme général $|u_n|$ (qui est un réel positif, si, si!), et on retourne à l'étape 3. Deux cas peuvent se présenter :

→ la série $\sum |u_n|$ converge, et alors la série $\sum u_n$ converge puisque :

$[\text{Absolument convergente}] \implies [\text{Convergente}]$

→ la série $\sum |u_n|$ diverge, et on ne peut alors pas conclure sur la nature de $\sum u_n$ (en Sup).

► Trois méthodes pour répondre à la question 2

Le calcul de la somme d'une série convergente n'est pas chose aisée, sauf cas particuliers ; pour vous en convaincre, revoir le DS consacré à la détermination de $\zeta(2)$, où un problème entier avait pour but d'établir

$$\text{que } \sum_{n=1}^{+\infty} \frac{1}{n^2} = \frac{\pi^2}{6}.$$

Néanmoins, dans quelques cas particuliers, le calcul de la somme est possible, explicitement :

→ **Somme d'une série géométrique** : si $q \in \mathbb{C}$ et si $|q| < 1$, alors $\sum_{n=0}^{+\infty} q^n = \frac{1}{1-q}$

→ **Somme d'une série télescopique** : si (w_n) converge : $\sum_{n=0}^{+\infty} (w_{n+1} - w_n) = \lim_{n \rightarrow +\infty} w_n - w_0$

→ **Formule de Taylor avec reste intégrale** : voir exemples d'utilisation dans le programme de la colle 28. Elle permet notamment d'établir que : $\forall x \in \mathbb{R}, e^x = \sum_{n=0}^{+\infty} \frac{x^n}{n!}$. En particulier : $e = \sum_{n=0}^{+\infty} \frac{1}{n!}$.

Autre application : $\ln(2) = \sum_{n=1}^{+\infty} \frac{(-1)^{n+1}}{n}$.

► Exemples d'application directe

Point par point, chacune des questions ci-dessous illustre les critères mentionnés ci-dessus.

EXERCICE 25.1 - (Divergence grossière) Montrer que la série $\sum \frac{\sqrt{n}}{\ln(n)}$ est grossièrement divergente.

Corrigé. Immédiat, puisque : $\lim_{n \rightarrow +\infty} \frac{\sqrt{n}}{\ln(n)} = +\infty$ (croissances comparées).

EXERCICE 25.2 - (Critère de comparaison) Montrer que la série $\sum \frac{1}{\ln(n)}$ est divergente, mais pas grossièrement divergente.

Corrigé. Le terme général $1/\ln(n)$ a pour limite 0 : la série ne diverge donc pas grossièrement. En revanche on a $1/\ln(n) \geq 1/n$ pour $n \geq 2$, d'où la conclusion (comparaison + critère de Riemann).

EXERCICE 25.3 - (Critère des équivalents) Montrer que la série $\sum n \arctan\left(\frac{1}{n^4}\right)$ est convergente.

Corrigé. Pour tout $n \in \mathbb{N}^*$, on a : $n \arctan\left(\frac{1}{n^4}\right) \geq 0$. En outre : $n \arctan\left(\frac{1}{n^4}\right) \sim_{n \rightarrow +\infty} \frac{1}{n^3}$.

Puisque $\frac{1}{n^3}$ est le terme général d'une série de Riemann convergente, on en déduit via le critère des équivalents que la série $\sum n \arctan\left(\frac{1}{n^4}\right)$ est convergente.

EXERCICE 25.4 - (Critère de négligeabilité) Montrer que la série $\sum \frac{\ln(n)}{n^3}$ est convergente.

Corrigé. Pour tout $n \in \mathbb{N}^*$, on a : $n^2 \times \frac{\ln(n)}{n^3} = \frac{\ln(n)}{n}$. Il s'ensuit que $\frac{\ln(n)}{n^3} = o\left(\frac{1}{n^2}\right)$. Puisque la série étudiée est à termes positifs, on peut alors conclure qu'elle converge via le critère de négligeabilité.

EXERCICE 25.5 - (Règle de Riemann) Montrer que pour tout réel $\alpha > 1$, la série $\sum \frac{\ln(n)}{n^\alpha}$ est convergente.

Corrigé. La série étudiée est à termes positifs. Considérons un réel β tel que : $1 < \beta < \alpha$.

Alors : $n^\beta \times \frac{\ln(n)}{n^\alpha} = \frac{\ln(n)}{n^{\alpha-\beta}}$. Puisque $\alpha - \beta > 0$, on a : $\lim_{n \rightarrow +\infty} n^\beta \frac{\ln(n)}{n^\alpha} = 0$.

On en déduit que : $\frac{\ln(n)}{n^\alpha} = o\left(\frac{1}{n^\beta}\right)$, d'où la conclusion puisque $\beta > 1$ (critère de Riemann + négligeabilité).

EXERCICE 25.6 - (Comparaison série-intégrale) Soit f la fonction définie sur $[2, +\infty[$ en posant :

$$\forall x \geq 2, f(x) = \frac{1}{x \ln(x)}.$$

1/ Justifier brièvement que f est positive et de classe $\mathcal{C}^\infty$ sur $[2, +\infty[$.

2/ Etudier le sens de variation de f sur $[2, +\infty[$.

3/ Pour tout réel $A \geq 2$, calculer : $I(A) = \int_2^A \frac{1}{x \ln(x)} dx$.

4/ Dédurre de ce qui précède la nature de $\sum \frac{1}{n \ln(n)}$.

Corrigé. Voir un peu plus haut dans ces lignes.

EXERCICE 25.7 - (Somme d'une série géométrique) Montrer que : $1 + \frac{1}{3} + \frac{1}{9} + \frac{1}{27} + \dots = \frac{3}{2}$.

Corrigé. La série de terme général $1/3^n$ est géométrique de raison $1/3$. Puisque $|1/3| < 1$, la série converge et $\sum_{n=0}^{+\infty} 1/3^n = 3/2$.

EXERCICE 25.8 - (Somme d'une série télescopique) Calculer $\sum \frac{1}{n^3 - n}$ (factoriser $n^3 - n$, et utiliser une décomposition en éléments simples).

Corrigé. Pour tout entier naturel $N \geq 2$, on a :

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

$$\text{D'où : } \sum_{n=2}^N \frac{1}{n^3 - n} = \frac{1}{2} \left[1 - \frac{1}{N} + \frac{1}{N+1} - \frac{1}{2} \right]. \text{ On en déduit que : } \sum_{n=2}^{+\infty} \frac{1}{n^3 - n} = \frac{1}{4}.$$

► **ET MAINTENANT, À VOUS DE JOUER !**

EXERCICE 25.9 - Pour tout réel $x \in]-1, 1[$, calculer $\sum_{n=0}^{+\infty} (-1)^n x^n$.

EXERCICE 25.10 - Soit x un réel tel que : $0 \leq x < \frac{1}{e}$. Montrer que la série $\sum \frac{x^n \text{sh}(n)}{n(n+1)}$ est convergente.

EXERCICE 25.11 - (**Exemples de séries de Bertrand**) Déterminer la nature de chacune des séries suivantes :

$$1) \sum \frac{\ln(n)}{n} \qquad 2) \sum \frac{\ln(n)}{n^\alpha} \ (\alpha > 1) \qquad 3) \sum \frac{1}{\sqrt{n} \ln(n)} \qquad 4) \sum \frac{1}{n \ln^2(n)}$$

EXERCICE 25.12 - Soient a et b deux réels. Le but est d'étudier la série $\sum (\ln(n) + a \ln(n+1) + b \ln(n+2))$.

a) Montrer qu'il existe deux réels α et β , que l'on exprimera en fonction de a et b tels que :

$$\forall n \in \mathbb{N}^*, \ln(n) + a \ln(n+1) + b \ln(n+2) = \alpha \ln(n) + \frac{\beta}{n} + O\left(\frac{1}{n^2}\right)$$

b) Dédurre de la question précédente une condition nécessaire et suffisante sur les réels a et b pour que la série de terme général $\ln(n) + a \ln(n+1) + b \ln(n+2)$ converge.

c) Dans le cas où elle converge, calculer la somme de la série $\sum (\ln(n) + a \ln(n+1) + b \ln(n+2))$.

EXERCICE 25.13 - On rappelle que $\sum_{n=1}^{+\infty} \frac{1}{n^2} = \frac{\pi^2}{6}$.

Etablir que la série $\sum \frac{1}{(2n+1)^2}$ converge, puis calculer sa somme : $\sum_{n=0}^{+\infty} \frac{1}{(2n+1)^2}$.

EXERCICE 25.14 - On rappelle que $\sum_{n=1}^{+\infty} \frac{1}{n^2} = \frac{\pi^2}{6}$.

Etablir que la série $\sum \frac{1}{n^2(n+1)^2}$ converge, puis calculer sa somme : $\sum_{n=1}^{+\infty} \frac{1}{n^2(n+1)^2}$.

On pourra utiliser une décomposition en éléments simples : $\frac{1}{n^2(n+1)^2} = \frac{a}{n} + \frac{b}{n^2} + \frac{c}{n+1} + \frac{d}{(n+1)^2}$.

EXERCICE 25.15 - On rappelle que $\sum_{n=0}^{+\infty} \frac{1}{n!} = e$. Calculer les sommes $\sum_{n=0}^{+\infty} \frac{n+1}{n!}$ et $\sum_{n=0}^{+\infty} \frac{n^2-1}{n!}$.

EXERCICE 25.16 - On pose : $\forall n \in \mathbb{N}, u_n = \frac{(2n)!}{(2^n n!)^2}$.

1) Déterminer un équivalent de : $w_n = \ln(u_{n+1}) - \ln(u_n)$.

2) Dédurre de la question précédente que : $\lim_{n \rightarrow +\infty} u_n = 0$.

3) Etablir qu'il existe un réel $C > 0$ tel que : $\lim_{n \rightarrow +\infty} \sqrt{n} u_n = C$ (sans chercher à expliciter la valeur de C).

4) En déduire la nature de la série $\sum \sqrt{n} u_n$.

25.5.2 Séries de Riemann

On a déjà vu le critère des convergences des séries de Riemann comme conséquence du théorème des accroissements finis au cours du chapitre portant sur la dérivabilité. On propose ci-dessous une preuve alternative (et sensiblement plus courte!) de ce résultat reposant sur le critère de comparaison série/intégrale.

THÉORÈME 25.1 - (Critère de convergence des séries de Riemann).

Soit α un réel.

$$[\text{La série } \sum \frac{1}{n^\alpha} \text{ converge}] \iff [\alpha > 1]$$

PREUVE. On distingue essentiellement deux cas, suivant que α est strictement positif ou non.

► **1er cas** — Si $\alpha \leq 0$. Alors $n^{-\alpha}$ ne tend pas vers 0, et la série de terme général $n^{-\alpha}$ diverge donc grossièrement.

► **2ème cas** — Si $\alpha > 0$. Introduisons alors la fonction $f : x \in \mathbb{R}_+^* \mapsto x^{-\alpha}$. La fonction f est continue, décroissante et positive sur $\mathbb{R}_+^*$.

Soit A un réel assez grand, et supposons $\alpha \neq 1$. On a : $I(A) = \int_1^A x^{-\alpha} dx = \frac{1}{1-\alpha} (A^{1-\alpha} - 1)$.

$$\text{On note alors que : } \lim_{A \rightarrow +\infty} A^{1-\alpha} = \begin{cases} 0 & \text{si } \alpha > 1 \\ +\infty & \text{si } \alpha < 1 \end{cases}.$$

On en déduit que $I(A)$ est convergente si $\alpha > 1$, et divergente si $\alpha < 1$. D'après le critère de comparaison série/intégrale, on en déduit que la série $\sum n^{-\alpha}$ est convergente si $\alpha > 1$, et divergente si $\alpha < 1$.

Il reste à voir que dans le cas $\alpha = 1$ la série est divergente, ce qui est immédiat. ☕

25.5.3 Séries de Bertrand

On appelle **série de Bertrand** une série dont le terme général de la forme $\sum \frac{1}{n^\alpha (\ln(n))^\beta}$ où α et β sont deux réels.

Aucun résultat concernant ces séries ne peut être utilisé sans démonstration. Néanmoins, il n'est pas rare que ces séries interviennent au détour d'un problème, et il peut donc être utile de connaître les méthodes permettant d'étudier leur nature.

EXERCICE CLASSIQUE 25.1 - Si $\alpha > 1$, alors la série $\sum \frac{1}{n^\alpha (\ln(n))^\beta}$ converge.

PREUVE. Soit γ un réel tel que : $1 < \gamma < \alpha$. Alors : $\lim_{n \rightarrow +\infty} \frac{n^{\overbrace{\gamma - \alpha}^{<0}}}{(\ln(n))^\beta} = 0$ (au pire, via les croissances comparées).

On en déduit que $\frac{1}{n^\alpha (\ln(n))^\beta} = o\left(\frac{1}{n^\gamma}\right)$.

Puisque la série étudiée est à termes positifs, on en déduit (critère de négligeabilité + $\gamma > 1$) qu'elle est convergente. ☕

EXERCICE CLASSIQUE 25.2 - Si $\alpha < 1$, alors la série $\sum \frac{1}{n^\alpha (\ln(n))^\beta}$ diverge.

PREUVE. Soit γ un réel tel que : $\alpha < \gamma < 1$. Alors : $\lim_{n \rightarrow +\infty} \frac{\overbrace{n^{\gamma-\alpha}}^{>0}}{(\ln(n))^\beta} = +\infty$ (au pire, via les croissances comparées).

On en déduit que $\frac{1}{n^\gamma} = o\left(\frac{1}{n^\alpha (\ln(n))^\beta}\right)$.

La série de terme général $n^{-\gamma}$ est divergente (série de Riemann avec $\gamma < 1$). Puisque la série étudiée est à termes positifs, on en déduit (critère de négligeabilité “à l'envers”) qu'elle est divergente. ☕

EXERCICE CLASSIQUE 25.3 - La série $\sum \frac{1}{n (\ln(n))^\beta}$ converge SSI $\beta > 1$.

PREUVE. Dans le cas où $\beta = 0$, la série étudiée est la série harmonique, qui est divergente.

Supposons à présent $\beta \neq 0$, et introduisons la fonction $f : x \in [2, +\infty[\mapsto \frac{1}{x \ln^\beta(x)}$.

La fonction f est dérivable (théorèmes généraux) sur $[2, +\infty[$ et pour tout réel $x \geq 2$ on a :

$$f'(x) = -\frac{\ln^\beta(x) + \beta \frac{\ln^\beta(x)}{x \ln(x)}}{x^2 \ln^{2\beta}(x)}$$

On en déduit que f est décroissante sur $[2, +\infty[$. En outre, f est continue (car dérivable) et positive (immédiat) sur $[2, +\infty[$.

Il s'ensuit que l'intégrale $\int_2^{+\infty} f(x) dx$ et la série $\sum_{n=2}^{+\infty} f(n)$ ont même nature, d'après le critère de comparaison série/intégrale.

Or, pour tout réel $A \geq 2$ on a, en supposant $\beta \neq 1$:

$$I(A) = \int_2^A \frac{1}{x \ln^\beta(x)} dx = \frac{1}{1-\beta} [\ln^{1-\beta}(x)]_2^A = \frac{1}{1-\beta} [\ln^{1-\beta}(A) - \ln^{1-\beta}(2)]$$

Or : $\ln^{1-\beta}(A)$ admet une limite finie lorsque A tend vers $+\infty$ si et seulement si $\beta > 1$.

Il reste à traiter le cas $\beta = 1$. Dans cette situation on a : $I(A) = \int_2^A \frac{1}{x \ln(x)} dx = [\ln(\ln(x))]_2^A$. Il s'ensuit immédiatement que $I(A)$ tend vers $+\infty$ lorsque A tend vers $+\infty$, ce qui achève la preuve. ☕

25.5.4 Séries alternées

PROPRIÉTÉ 25.13 - (Critère spécial des séries alternées).

Soit (u_n) une suite réelle positive et décroissante, telle que $\lim_{n \rightarrow +\infty} u_n = 0$.

La série $\sum (-1)^n u_n$ est convergente.

PREUVE. On introduit la suite $(S_N)_N$ des sommes partielles de cette série, en posant :

$$\forall N \in \mathbb{N}, S_N = \sum_{n=0}^N (-1)^n u_n$$

L'idée de la preuve est d'établir que les suites des termes de rang pair des termes de rang impair extraites de (S_N) sont adjacentes.

► **Etape 1. Décroissance de la suite (S_{2N}) .**

Pour tout entier naturel N on a : $S_{2N+2} - S_{2N} = u_{2N+2} - u_{2N+1}$. Puisque la suite (u_n) est décroissante, on en déduit que : $\forall N \in \mathbb{N}, S_{2N+2} - S_{2N} \leq 0$. Par conséquent : la suite (S_{2N}) est décroissante (♠).

► **Etape 2. Croissance de la suite (S_{2N+1}) .**

Pour tout entier naturel N on a : $S_{2N+3} - S_{2N+1} = u_{2N+2} - u_{2N+3}$. Puisque la suite (u_n) est décroissante, on en déduit que : $\forall N \in \mathbb{N}, S_{2N+3} - S_{2N+1} \geq 0$. Par conséquent : la suite (S_{2N+1}) est croissante (♣).

► **Etape 3. Limite de la différence $(S_{2N+1} - S_{2N})$.**

Pour tout entier naturel N on a : $S_{2N+1} - S_{2N} = -u_{2N+1}$. Puisque par hypothèse la suite (u_n) a pour limite 0, il en est de même pour la suite (u_{2N+1}) (propriété des suites extraites).

On en déduit que : $\lim_{N \rightarrow +\infty} (S_{2N+1} - S_{2N}) = 0$ (♥).

► **Conclusion.**

On déduit de (♠), (♣) et (♥) que les suites (S_{2N}) et (S_{2N+1}) sont adjacentes. A ce titre elles convergent et ont la même limite. On en déduit que la suite (S_N) converge également (vers la même limite).

Par conséquent, la série $\sum (-1)^n u_n$ est convergente. ☕

Exemple d'application. Montrer que la série $\sum \frac{(-1)^n}{\sqrt{n+1}}$ est semi-convergente, c'est-à-dire qu'elle est convergente, mais qu'elle n'est pas absolument convergente.

La suite de terme général $u_n = \frac{1}{\sqrt{n+1}}$ est positive, décroissante et converge vers 0. D'après le critère des séries alternées, la série $\sum \frac{(-1)^n}{\sqrt{n+1}}$ est convergente.

En revanche $|u_n| \sim_{n \rightarrow +\infty} n^{-1/2}$. Il s'ensuit (critère des séries de Riemann) que la série de terme général $|u_n|$ est divergente.

Ainsi la série $\sum \frac{(-1)^n}{\sqrt{n+1}}$ est semi-convergente.

Index

- $i\mathbb{R}$, 61
- $\mathbb{K}$, 35
- $\mathbb{U}_n$, 83
- ℓci , 273
- $\binom{n}{p}$, 44
- $\mathbb{C}$, 61
- $\mathbb{Q}$ est dense dans $\mathbb{R}$, 232
- $\mathbb{U}$, 69
- $\sqrt{2}$ est irrationnel, 23
- e^{ix} , 70
- $\zeta(2)$, 196
- n -uplets, 21
- j , 83
- affixe
 - d'un point, 62
 - d'un vecteur, 62
- angle-moitié, 73
- anneau, 282
 - commutatif, 282
 - des entiers de Gauss, 295
 - des matrices carrées, 308
 - des matrices diagonales, 312
 - des matrices scalaires, 312
 - des matrices triangulaires
 - inférieures, 312
 - supérieures, 312
 - des polynômes, 467
 - intègre, 287
- antécédent, 107
- application, 139
 - bijective, 141
 - injective, 141
 - linéaire, 552
 - canoniquement associée à une matrice, 650
 - surjective, 141
- approximation
 - affine, 117
- argument, 76
- arrangement, 390
- assertion logique, 15
- assertions
 - logiquement équivalentes, 15
- automorphisme, 557
- axe
 - imaginaire pur, 62
 - réel, 62
- axiome, 22
- base
 - d'un espace vectoriel, 573
- Bernoulli
 - Daniel, 223
 - Jakob, 223
 - Jean, 223
 - Jean II, 223
 - Nicolas, 223
- bijection, 141
 - réciproque, 143
- binôme de Newton, 46, 284, 316
- borne
 - inférieure, 228
 - supérieure, 228
- caractérisation
 - de la borne supérieure, 229
 - séquentielle
 - de la densité, 231
- cardinal
 - d'un ensemble fini, 384
- cercle unité, 69
- changement de variable, 170, 171
- codage RSA, 429
- codomaine, 139
- coefficient

- dominant, 470
- colinéaires, 571
- combinaison, 44
 - linéaire, 550
- commutant
 - d'une matrice
 - carrée, 313
 - diagonale, 313
- complémentaire
 - d'une partie, 20
- composée
 - de deux applications, 140
 - de deux fonctions, 108
- condition
 - nécessaire, 16
 - nécessaire et suffisante, 16
 - suffisante, 16
- congru, 51
- congruence, 237, 417
 - modulo un nombre réel, 51
- conjecture, 22
- conjonction, 16
- conjugué, 63
- constante
 - d'Euler, 271
- continûment
 - dérivable, 163
 - n fois, 163
 - indéfiniment, 163
- contraposée, 22
- convergence
 - absolue, 689
- coordonnées
 - d'un vecteur dans une base, 582
- corps, 290
- cosinus
 - d'un nombre réel, 52
 - hyperbolique, 24, 126
- critère
 - de comparaison
 - pour les séries, 686
 - de comparaison série/intégrale, 688
 - de négligeabilité
 - pour les séries, 687
- des équivalents
 - pour les séries, 687
- croissance
 - de l'intégrale, 612
- cycle, 392
- cycles
 - à supports disjoints, 393
- degré
 - d'un polynôme, 467
- demi-vie, 200
- différence
 - de deux ensembles, 20
 - symétrique de deux ensembles, 20
- dilatation, 657
- dimension
 - d'un espace vectoriel, 575
 - d'un sous-espace affine, 660
- direction
 - d'un sous-espace affine, 660
- disjonction, 16
 - de cas, 22
- divergence
 - grossière, 684
- diviseur, 411
 - d'un polynôme, 472
 - de zéro, 286
- division euclidienne
 - dans $\mathbb{K}[X]$, 472
 - dans $\mathbb{Z}$, 413
- domaine, 139
- droite
 - affine, 660
- décomposition
 - en facteurs premiers, 425
 - en éléments simples, 177
- délinéarisation, 75
- démonstration
 - par analyse synthèse, 23
 - par l'absurde, 23
 - par récurrence, 25
- dérivée
 - n -ième
 - d'un polynôme formel, 480
 - n -ième d'une fonction, 121

- déterminant
 - d'une matrice de $M_2(\mathbb{K})$, 325
- développement
 - asymptotique
 - de la série harmonique, 271
 - limité
 - à l'ordre n , 453
 - à l'ordre 1, 117
- EDL1, 197
 - avec second membre, 197
 - charge d'un condensateur, 201
 - désintégration radioactive, 199
 - homogène, 197
 - associée, 197
 - sans second membre, 197
- EDL2, 205
 - avec second membre, 205
 - homogène, 205
 - associée, 205
 - sans second membre, 205
- endomorphisme, 552
- ensemble
 - de définition, 107, 139
 - des parties, 19
 - des points de discontinuité, 376
 - dénombrable, 142
 - en compréhension, 17
 - en extension, 17
 - fini, 382
 - habité, 32
 - ordonné, 228
 - totalement ordonné, 228
- ensembles
 - disjoints, 20
 - distincts, 19
 - équipotents, 142, 381
- entier
 - de Gauss
 - inversible, 295
 - irréductible, 296
- entiers
 - de Gauss
 - premiers entre eux, 296
 - premiers entre eux, 418
- espace
 - probabilisé, 516
 - vectériel, 545
- espace vectoriel
 - de dimension finie, 573
- espérance, 517
 - mathématique, 517
- exponentielle
 - complexe, 79
- expérience aléatoire, 505
- extremum
 - d'une fonction, 113
- factorielle d'un entier, 43
- famille
 - génératrice, 550
 - libre, 572
 - liée, 572
 - échelonnée
 - de polynômes, 578
- fonction
 - arccosinus, 149
 - arcsinus, 150
 - arctangente, 152
 - bornée, 112
 - composée, 108
 - continue
 - en un point, 367
 - par morceaux, 605
 - sur un intervalle, 368
 - cosinus, 52
 - croissante, 113
 - de classe $\mathcal{C}^1$, 163
 - de classe $\mathcal{C}^n$, 163
 - de classe $\mathcal{C}^{+\infty}$, 163
 - dominée, 450
 - décroissante, 113
 - dérivable, 115, 435
 - à valeurs complexes, 443
 - dérivée, 115
 - en escalier, 604
 - exponentielle, 125
 - de base a , 129
 - génératrice d'une VAR, 528
 - identité, 108

- impaire, 111
- indicatrice, 144
- logarithme
 - de base a , 128
- logarithme népérien, 125
- majorée, 112
- minorée, 112
- monotone, 113
- numérique, 107
- paire, 111
- polynomiale
 - associée à un élément de $\mathbb{K}[X]$, 475
- prolongée par continuité, 368
- puissance, 131
- sinus, 52
- tangente, 53
- uniformément continue, 601
- valeur absolue, 124
- fonctions
 - équivalentes, 451
- formule
 - de Bayes, 516
 - de Koenig-Huygens, 519
 - de la moyenne, 374, 614
 - de Moivre, 71
 - de Stirling, 188
 - de Taylor
 - avec reste intégrale, 617
 - dans $\mathbb{K}[X]$, 482
 - Young, 454
 - de transfert, 518
 - de Vandermonde, 467, 529
 - du changement de base, 651
- formules
 - d'addition, 56, 134
 - d'Euler, 71
 - généralisées, 75
 - de duplication, 56, 134
 - de soustraction, 56, 134
- Galois, Evariste, 402
- groupe, 276
 - abélien, 276
 - additif des polynômes, 466
 - alterné, 398
 - cyclique, 280, 294
 - des inversibles d'un anneau, 289
 - des permutations, 281
 - des similitudes directes, 95, 277
 - linéaire, 324, 558
 - symétrique S_n , 390
- heptakaïdécagone régulier, 95
- homothétie, 92
- hyperplan, 581
 - affine, 660
- identité
 - d'un ensemble, 139
- image, 107
 - d'un élément par une application, 139
 - d'une application, 140
 - d'une application linéaire, 554
 - d'une fonction, 107
 - d'une partie, 140
 - réciproque d'une partie, 140
- imaginaire pur, 61
- implication, 16
- inclusion, 19
- injection, 141
- intersection, 19
 - de deux évènements, 505
- intervalle
 - d'entiers, 18
 - de résolution d'une EDL, 198
- intégrale
 - d'une fonction continue par morceaux, 611
 - d'une fonction continue sur un segment, 166
 - d'une fonction en escalier, 608
 - d'une fonction à valeurs complexes, 615
 - de Wallis, 180
- intégration
 - par parties, 168
- invariant de similitude, 336
- inverse
 - d'un entier modulo n , 418
 - d'un élément pour une ℓ ci, 275
- inversion, 397
- involution, 144
- inégalité

- de Марков, 519
- de Bienaymé-Чебышев, 519
- de Cauchy-Schwarz-Буняковски, 612
- de Taylor-Lagrange, 618
- des accroissements finis, 440
- triangulaire, 67, 114
 - renversée, 114
 - triangulaire (cas d'égalité), 68
- isomorphe, 557
- isomorphisme, 557
- issue, 505
- John Machin, 159
- Lao-Tseu, 231
- lemme
 - de Cauchy, 369
 - de Cesàro, 269
 - de Gauss, 419
 - de Riemann-Lebesgue, 189
- limite
 - d'une fonction
 - en $+\infty$, 362
 - en $-\infty$, 362
 - en un réel, 361
 - à droite d'un réel, 363
 - à gauche d'un réel, 363
 - d'une suite, 243
- linéarisation, 72
- linéarité
 - de l'intégrale, 166, 611
 - de la dérivation, 120
 - de la dérivée n -ième, 122
 - de la somme, 37
- lipschitzienne, 441
- loi
 - associative, 274
 - binomiale, 522
 - commutative, 274
 - de Bernoulli, 522
 - de composition interne, 273
 - de probabilité, 521
 - uniforme, 521
- lois
 - de Morgan, 17, 32
- majorant, 112, 228
- matrice, 301
 - antisymétrique, 318
 - augmentée, 330
 - carrée, 306
 - colonne, 302
 - de passage, 582
 - de permutation, 657
 - de transvection, 657
 - diagonale, 309
 - idempotente, 528
 - identité, 302, 307
 - inversible, 323
 - ligne, 302
 - nilpotente, 316
 - nulle, 301
 - opposée, 302
 - produit, 304
 - scalaire, 309
 - symétrique, 318
 - transposée, 306
 - triangulaire
 - inférieure, 309
 - supérieure, 309
- matrices
 - semblables, 334, 654
- maximum
 - d'une fonction, 113
- minimum
 - d'une fonction, 113
- minorant, 112, 228
- module, 65
- modulo, 51
- multiple, 411
 - d'un polynôme, 472
- multiplicité
 - d'une racine, 485
- nombre
 - composé, 422
 - de Carmichael, 427
 - dérivé, 115, 435
 - à droite, 435
 - à gauche, 435
 - premier, 422

- nombre complexe, 61
- noyau
 - d'une application linéaire, 554
- négation, 15
 - d'une implication, 17
- négligeable, 448
- opération élémentaire, 657
- ordre
 - d'un élément dans un groupe, 291
- oscillateur harmonique, 217
- partie
 - bornée, 228
 - dense dans $\mathbb{R}$, 231
 - entière, 233
 - imaginaire
 - d'une fonction, 178
 - majorée, 228
 - minorée, 228
 - réelle
 - d'une fonction, 178
- partie régulière, 453
- partition, 506
- pas d'une subdivision, 603
- pentagone régulier, 95
- permutation, 280, 390, 657
 - impaire, 397
 - paire, 397
- PGCD
 - dans $\mathbb{Z}[i]$, 296
 - de deux entiers, 414
- pivot de Gauss, 330
- plan
 - affine, 660
 - complexe, 62
- plus
 - grand élément, 228
 - petit élément, 228
- point
 - image d'un complexe, 62
 - image d'un réel, 52
- polygone régulier à n côtés, 85
- polynôme, 465
 - dérivé, 480
 - formel, 465
 - irréductible, 487
 - nul, 465
 - produit, 466
 - scindé, 493
 - trigonométrique, 176
 - unitaire, 470
- polynômes
 - associés, 474
 - de Tchebychev, 498
 - interpolateurs de Lagrange, 478
- pont $\mathbb{R} \longleftrightarrow \mathbb{C}$
 - pour les EDL, 214
 - pour les fonctions continues), 375
- positivité
 - de l'intégrale, 166, 611
- postulat, 22
- PPCM
 - de deux entiers, 421
- primitive, 164
- principe
 - du prolongement algébrique, 477
- principe de superposition des solutions, 212
 - version générale, 213
- probabilité, 507
 - conditionnelle, 511
- problème
 - de Cauchy, 199
- problème de Cauchy, 70
- produit, 42
 - cartésien, 21
 - de deux fonctions, 108
 - de deux matrices, 304
 - des racines de l'unité, 88
 - télescopique, 42
- projecteur, 562
- projection
 - sur un sev, 561
- prolongement
 - d'une application, 140
 - par continuité, 368
- propriété
 - de continuité séquentielle, 369
 - de la borne supérieure, 229

- de limite séquentielle, 364
- suites extraites
 - des termes de rangs pairs et impairs, 256, 261
- puissance
 - d'une matrice carrée, 314
- quantificateur
 - d'unicité, 18
 - existentiel, 18
 - universel, 18
- quantificateurs, 18
- quotient
 - de deux fonctions, 108
- racine
 - n -ième
 - d'un complexe, 90
 - de l'unité, 83
 - carrée
 - complexe, 80
 - d'un polynôme, 475
 - multiple, 485
- raison
 - d'une suite arithmétique, 35
 - d'une suite géométrique, 36
- rang
 - d'une application linéaire, 635
 - d'une famille de vecteurs, 635
- relation
 - binaire, 227
 - antisymétrique, 227
 - d'équipotence, 381
 - de finesse sur les subdivisions, 604
 - réflexive, 227
 - symétrique, 227
 - transitive, 227
 - d'ordre, 227
 - totale, 228
 - d'équivalence, 51, 227
 - de Chasles
 - pour les intégrales, 612
 - de congruence, 51, 237, 417
 - de divisibilité, 412
 - de domination, 450
 - de Graßmann, 580
 - de négligeabilité, 448
 - de similitude dans $M_n(\mathbb{K})$, 335
 - fondamentale de la trigonométrie, 55
 - fondamentale de la trigonométrie hyperbolique, 134
- relation de Chasles
 - pour les intégrales, 166
 - pour les produits, 42
 - pour les sommes, 37
- relation de Pascal, 45
- relations
 - coefficients/racines, 83
- restriction
 - d'une application, 140
- Riemann-Lebesgue, 189
- rotation, 92
- règle
 - de double inclusion, 19
 - de non-contradiction, 15
 - du tiers exclu, 15
- règles
 - de négation (Morgan), 17
 - de négation (quantificateurs), 18
- récurrence
 - double, 25
 - décalée, 25
 - faible, 25
 - finie, 25
 - forte, 25
- réflexion, 92
- régime
 - apériodique, 218
 - critique, 218
 - pseudo-périodique, 218
 - périodique, 217
- résonance, 219
- réunion
 - de deux subdivisions, 604
 - disjointe, 507
- scalaire, 35
- sev
 - supplémentaires, 559
- signature, 397
- similitude directe, 92

- singleton, 18
- sinus
 - d'un nombre réel, 52
 - hyperbolique, 24, 126
 - complexe, 104
- somme, 37
 - d'une suite arithmétique, 26, 41
 - d'une suite géométrique, 42
 - de deux fonctions, 108
 - de deux matrices, 302
 - de Riemann, 614
 - de sev, 558
 - des cubes, 27
 - des premiers carrés, 40
 - des premiers cubes, 41
 - des premiers entiers, 40
 - des racines de l'unité, 87
 - directe de deux sev, 559
 - partielle associée à une suite, 683
 - télescopique, 37
- sous-anneau, 286
- sous-corps, 290
- sous-espace
 - affine, 658
- sous-espace vectoriel, 547
- sous-groupe, 281
 - additif de $\mathbb{Z}$, 412
 - des permutations paires, 282
 - engendré, 291
 - trivial, 281
- stabilité
 - des inégalités larges par passage à la limite, 248
- Stirling, 188
- subdivision, 603
 - adaptée
 - à une fonction continue par morceaux, 605
 - à une fonction en escalier, 604
 - plus fine, 604
 - régulière, 603
 - à pas constants, 603
- suitarit, 36
- suite
 - arithmético-géométrique, 262
 - arithmétique, 35, 261
 - bornée, 242
 - complexe, 241
 - bornée, 259
 - convergente, 259
 - constante, 242
 - convergente, 243
 - croissante, 242
 - de Fibonacci, 266
 - des approximations décimales
 - par défaut, 235
 - par excès, 235
 - divergente, 243
 - vers $+\infty$, 247
 - décroissante, 242
 - extraite, 254
 - géométrique, 36, 261
 - majorée, 242
 - minorée, 242
 - monotone, 242
 - périodique, 242
 - récurrente
 - non-linéaire, 267
 - récurrente linéaire d'ordre 2, 263
 - réelle, 241
 - stationnaire, 242
- suites
 - adjacentes, 251
- support, 392
 - d'une subdivision, 604
- surjection, 141
- symbole de Kronecker, 307
- symétrie
 - axiale, 92
 - par rapport à un sev, 561
- symétrique
 - d'un élément pour une ℓ ci, 275
 - par rapport à zéro, 111
- système
 - complet d'évènements, 506
 - de Cramer, 326
- série, 683
 - absolument
 - convergente, 689
 - convergente, 683

- de Bertrand, 695
- de Riemann, 686
 - critère de convergence, 444
- divergente, 683
- grossièrement divergente, 684
- géométrique, 685
- harmonique, 39
- semi-convergente, 690
- éléscopique, 685
- table
 - de vérité, 15
- TAF, 438
- tangente
 - d'un nombre réel, 53
 - hyperbolique, 126
- tautologie, 22
- terme général, 683
 - d'une suite, 35
- TFD, 676
- théorème, 22
 - $\mathbb{R}$ est archimédien, 231
 - approximation uniforme par des fonctions en escalier, 606
 - caractérisation des éléments de $GL_n(\mathbb{K})$, 326
 - d'encadrement
 - pour les suites, 249
 - de Bezout, 415
 - de Bolzano-Weierstrass, 258
 - cas complexe, 261
 - de Cantor-Bernstein, 409
 - de classification
 - des ev de dimension finie, 586
 - des similitudes directes, 94
 - de comparaison
 - pour les suites, 249
 - de d'Alembert-Gauss, 488
 - de décomposition (paire+impaire), 23, 112
 - de Fermat, 423
 - de Heine, 602
 - de la base extraite, 576
 - de la base incomplète, 577
 - de la bijection, 373
 - de la division euclidienne
 - dans $\mathbb{K}[X]$, 472
 - dans $\mathbb{Z}$, 413
 - de la limite monotone, 40, 250
 - de Lagrange, 280, 293
 - de Rolle, 437
 - de Taylor-Young, 454
 - de Wilson, 432
 - des 4 dimensions, 580
 - des accroissements finis, 438
 - des bornes atteintes, 373
 - des gendarmes
 - pour les suites, 249
 - des intervalles emboîtés, 253
 - des suites adjacentes, 251
 - des suites extraites, 255
 - limite infinie, 256
 - des valeurs intermédiaires, 372
 - description de $GL_2(\mathbb{K})$, 324
 - donnant les racines carrées de z , 80
 - du point fixe, 372
 - du prolongement
 - de la dérivée, 440
 - du rang, 636
 - décomposition "Symétrique + Antisymétrique", 321
 - décomposition en facteurs premiers, 425
 - décomposition en irréductibles dans $\mathbb{K}[X]$, 487
 - dérivée d'une fonction composée, 120
 - ensemble des solutions de $e^z = \omega$, 79
 - espérance et variance de $\mathcal{B}(n, p)$, 523
 - existence d'un supplémentaire en dimension finie, 579
 - existence d'une solution pour une EDL1 avec second membre, 203
 - existence et unicité d'une solution pour un problème de Cauchy d'ordre 1, 205
 - existence et unicité des coordonnées, 582
 - fondamental
 - de l'intégration, 616
 - du calcul intégral, 164
 - formule
 - de Stirling, 188
 - formule d'intégration par parties, 168
 - formule de Koenig-Huygens, 519
 - formule de Leibnitz, 122

- formule de transfert, 518
- formule des probabilités totales, 510, 513
- formule du changement de variable dans une intégrale, 170, 171
- image continue d'un intervalle, 372
- irréductibles de $\mathbb{C}[X]$, 488
- lien entre multiplicité et dérivées successives, 485
- loi binomiale, 523
- nombre maximal de racines d'un polynôme, 477
- noyau-image, 636
- pont $\mathbb{R} \longleftrightarrow \mathbb{C}$
 - pour la dérivabilité, 443
 - pour les suites, 259
- racines n -ièmes
 - d'un complexe, 90
 - de l'unité, 84
- solution générale
 - d'une EDL1 homogène, 198
 - d'une EDL2 homogène (cas complexe), 207
 - d'une EDL2 homogène (cas réel), 209
- spécial des séries alternées, 257
- structure de l'ensemble des solutions d'une EDL1, 197
- structure de l'ensemble des solutions d'une EDL2, 206
- suites extraites
 - cas complexe, 260
- terme général d'une SRL2
 - cas complexe, 264
 - cas réel, 265
- équations du second degré à coefficients complexes, 81
- théorèmes généraux
 - sur la continuité, 368
 - sur les espaces vectoriels, 546
 - sur les fonctions de classe $\mathcal{C}^N$, 163
- trace
 - d'un endomorphisme, 656
 - d'une matrice carrée, 322
- transformée de Fourier discrète, 676
- translation, 92
- transposition, 391
- transvection, 657
- triangle de Pascal, 46
- union, 19
 - de deux évènements, 505
- univers, 505
- valeur
 - absolue, 113
 - moyenne, 374, 614
 - d'une VAR, 517
- valeurs remarquables, 53
- valuation p -adique, 426
- variable aléatoire réelle (VAR), 516
- variance, 517
- vecteur, 546
 - image d'un complexe, 62
- vecteurs
 - colinéaires, 571
 - linéairement indépendants, 571
- voisinage, 361
- Wallis, 180
- échelle de négligeabilité, 450
- élément
 - d'ordre fini, 291
 - inversible, 275
 - neutre, 274
 - nilpotent, 287
 - symétrisable, 275
- éléments
 - commutants, 284
- épreuve de Bernoulli, 522
- équation
 - caractéristique
 - d'une EDL2, 207
 - d'une SRL2, 263
 - de Bernoulli, 223
 - différentielle
 - ordinaire, 223
 - homogène
 - associée, 661
 - linéaire, 661
- équation différentielle
 - linéaire
 - d'ordre 1, 197
 - d'ordre 2, 205
- équiprobabilité, 509

équivalence, 16

équivalent, 451

éventualité, 505

évènement, 505

 certain, 505

 contraire, 505

 impossible, 505

 élémentaire, 505

événements

 disjoints, 506

 incompatibles, 506

événements indépendants, 512