Algèbre générale et algèbre linéaire

Préparation à l'Agrégation Interne
Université Paris Diderot (Paris 7) - IREM

Année 2013-2014

Table des matières

Première partie

Algèbre générale

1 Arithmétique dans $\mathbb{Z}$

1.1 Division dans $\mathbb{Z}$

1.1 Définition. Soient $a, b \in \mathbb{Z}$. On dit que a divise b et on écrit $a|b$ s'il existe $c \in \mathbb{Z}$ tel que $b = ac$.

On dit aussi que b est un multiple de a, que b est divisible par a, que a est un diviseur de b...

1.2 Propriétés élémentaires. a) Pour tout $a \in \mathbb{Z}$, on a : $1|a$, $a|a$ et $a|0$.

b) Pour tout $a, b \in \mathbb{Z}$, on a : $(a|b$ et $b|a) \iff |a| = |b|$.

c) Pour tout $a, b, c \in \mathbb{Z}$, on a : $(a|b$ et $b|c) \Rightarrow a|c$.

d) Pour tout $a, b, c \in \mathbb{Z}$, on a : $(a|b$ et $a|c) \Rightarrow a|b + c$.

1.3 Exercice. a) Soient $a, b, c, d \in \mathbb{Z}$. Démontrer que si $a|b$ et $c|d$ alors $ac|bd$.

b) Soient $a, b \in \mathbb{Z}$ et $n \in \mathbb{N}$. Démontrer que si $a|b$, alors $a^n|b^n$.

1.4 Théorème : Division euclidienne. *Soient $a \in \mathbb{Z}$ et $b \in \mathbb{N}$ non nul. Alors il existe un unique couple $(q, r) \in \mathbb{Z}^2$ tels que $a = bq + r$ et $0 \leqslant r < b$.*

1.5 Définition. Un nombre $p \in \mathbb{Z}$ est dit *premier* s'il a exactement 4 diviseurs : $1, p, -1$ et $-p$.

En particulier, 1 (et -1) n'est pas (ne sont) pas premier(s).

1.6 Proposition. *Soit $n \in \mathbb{Z}$ un nombre distinct de 1 et de -1. Alors n admet un diviseur premier.*

Le plus petit diviseur strictement supérieur à 1 de n est un nombre premier.

1.7 Théorème. *Il y a une infinité de nombres premiers.*

Il suffit en effet de remarquer que tout diviseur premier de $n! + 1$ est $\geqslant n + 1$.

1.2 Sous-groupes additifs de $\mathbb{Z}$

1.8 Notation. Soit $a \in \mathbb{Z}$. L'ensemble des multiples de a, c'est à dire l'ensemble $\{ab;\ b \in \mathbb{Z}\}$ est noté $a\mathbb{Z}$.

1.9 Remarque. Pour $a, b \in \mathbb{Z}$ on a l'équivalence entre :

$$\text{(i) } a|b; \qquad \text{(ii) } b \in a\mathbb{Z} \qquad \text{et} \quad \text{(iii) } b\mathbb{Z} \subset a\mathbb{Z}.$$

D'après 1.2.a), on a $a\mathbb{Z} = b\mathbb{Z}$ si et seulement si $|a| = |b|$ (*i.e.* $b = \pm a$).

1.10 Proposition. *Pour tout $a \in \mathbb{Z}$, l'ensemble $a\mathbb{Z}$ est un sous-groupe additif de $\mathbb{Z}$. C'est le plus petit sous-groupe de $\mathbb{Z}$ contenant a.*

1.11 Théorème. *Tout sous-groupe de $\mathbb{Z}$ est de cette forme : si $G \subset \mathbb{Z}$ est un sous-groupe additif, il existe (un unique) $a \in \mathbb{N}$ tel que $G = a\mathbb{Z}$.*

1.3 PGCD, PPCM algorithme d'Euclide

1.12 Corollaire. *Soient $a, b \in \mathbb{Z}$.*

a) *Il existe un unique $m \in \mathbb{N}$ tel que $a\mathbb{Z} \cap b\mathbb{Z} = m\mathbb{Z}$. Le nombre m est un multiple commun de a et de b. Les multiples communs de a et b sont les multiples de m.*

b) *Il existe un unique $d \in \mathbb{N}$ tel que $a\mathbb{Z} + b\mathbb{Z} = d\mathbb{Z}$. Le nombre d est un diviseur commun de a et de b. Les diviseurs communs de a et b sont les diviseurs de d.*

1.13 Définition. Le nombre d de ce corollaire s'appelle le plus grand commun diviseur (PGCD) de a et b ; on le note $pgcd(a, b)$. Le nombre m de ce corollaire s'appelle le plus petit commun multiple (PPCM) de a et b ; on le note $ppcm(a, b)$.

1.14 Remarque. Soient $n \in \mathbb{N}^*$ et $x_1, \ldots, x_n \in \mathbb{Z}$. On définit de même le plus grand commun diviseur d et le plus petit commun multiple m de $x_1, \ldots, x_n$:

- Le nombre $m \in \mathbb{N}$ est un multiple commun des x_i ; les multiples communs des x_i sont les multiples de m. Autrement dit
$$m\mathbb{Z} = x_1\mathbb{Z} \cap x_2\mathbb{Z} \cap \ldots \cap x_n\mathbb{Z} = \bigcap_{i=1}^{n} x_i\mathbb{Z}.$$
- Le nombre $d \in \mathbb{N}$ est un diviseur commun des x_i ; les diviseurs communs des x_i sont les diviseurs de d. On a
$$d\mathbb{Z} = x_1\mathbb{Z} + x_2\mathbb{Z} + \ldots + x_n\mathbb{Z} = \sum_{i=1}^{n} x_i\mathbb{Z}.$$

1.15 Lemme. *Soient $a, b \in \mathbb{Z}$. On suppose que $b \neq 0$. On note r le reste de la division euclidienne de a par $|b|$. On a $pgcd(a, b) = pgcd(b, r)$.*

1.16 Algorithme d'Euclide. Soient $a, b \in \mathbb{N}$. On suppose que $b \neq 0$.

- On pose $r_0 = a$, $r_1 = b$ et on note r_2 le reste de la division euclidienne de a par b.
- Soit $n \in \mathbb{N}$ non nul et supposons r_j construits pour $1 \leqslant j \leqslant n$. Si r_n n'est pas nul, alors on définit r_{n+1} comme le reste de la division euclidienne de r_{n-1} par r_n : $r_{n-1} = q_n r_n + r_{n+1}$. Si r_n est nul, on arrête la construction.

a) La construction s'arrête en un nombre fini d'étapes.

b) Le PGCD de a et b est le dernier reste non nul.

1.17 Remarques. a) On peut majorer le nombre N d'étapes qu'il faut pour trouver le PGCD. Sachant que la suite r_k est strictement décroissante, on trouve évidemment $N \leqslant b$. Mais on peut faire bien mieux !

Remarquons que $r_{N-1} = q_N r_N \geqslant 2r_N$ (puisque et $0 \leqslant r_N < r_{N-1}$), et pour $1 \leqslant k \leqslant N-1$, on a $r_{k-1} = q_k r_k + r_{k+1} \geqslant r_k + r_{k+1}$, de sorte que, par récurrence, $r_{N-k} \geqslant r_N F_{k+2}$, où F_k est le k-ième nombre de Fibonacci (donné par récurrence par les formules $F_0 = 0$, $F_1 = 1$ et $F_{k+1} = F_k + F_{k-1}$ pour $k \geqslant 1$ - on initialise la récurrence avec $k = 0$ et 1 sachant que $F_2 = 1$ et $F_3 = 2$). Rappelons que F_k croît géométriquement : $F_k = \dfrac{\phi^k - (-1)^k \phi^{-k}}{\sqrt{5}}$ où $\phi = \dfrac{1+\sqrt{5}}{2}$ est le nombre d'or. On a donc $b = r_1 \geqslant F_{N+1} > \dfrac{\phi^{N+1} - 1}{\sqrt{5}}$ une estimation pour N logarithmique en b : $N < \dfrac{\ln(1 + b\sqrt{5})}{\ln \phi} - 1$.

b) Pour écrire une relation de Bézout $d = r_N = au + bv$, on peut remonter les opérations : $r_N = r_{N-2} - r_{N-1}q_{N-1} = r_{N-2} - q_{N-1}(r_{N-3} - r_{N-2}q_{N-2}) = (1 + q_{N-1}q_{N-2})r_{N-2} - q_{N-1}r_{N-3}$, puis en écrivant $r_{N-2} = r_{N-4} - r_{N-3}q_{N-3}$ on exprime r_N en fonction de r_{N-3} et r_{N-4}, et on continue... Cela demande de garder en mémoire la suite des quotients q_k.

On peut faire un peu mieux, en écrivant à chaque étape de l'algorithme $r_k = u_k a + v_k b$. On aura $r_{k+1} = r_{k-1} - q_k r_k = (u_{k-1} - q_k u_k)a + (v_{k-1} - q_k v_k)b$. En même temps qu'on trouvera le PGCD, on aura une relation de Bézout !

1.18 Quelques explications sur la suite de Fibonacci. Soient $a, b \in \mathbb{C}$. On considère les suites u_n qui satisfont une propriété de récurrence $u_{n+2} = au_{n+1} + bu_n$. Elles forment un sous-espace vectoriel E de l'espace $\mathbb{C}^{\mathbb{N}}$ des suites complexes. Comme une telle suite est entièrement déterminée par u_0 et u_1, cet espace vectoriel est de dimension 2 (l'application linéaire $(u) \mapsto (u_0, u_1)$ est un isomorphisme de E sur $\mathbb{C}^2$). On cherche une base de E de la forme $u_n = x^n$ (avec $x \in \mathbb{C}$). La suite (x^n) est dans E si et seulement si $x^2 = ax + b$. Si les racines r_1 et r_2 du polynôme $X^2 - aX - b$ sont distinctes, on obtient deux suites indépendantes r_1^n et r_2^n, donc toutes les solutions s'écrivent $u_n = \alpha r_1^n + \beta r_2^n$ (avec $\alpha, \beta \in \mathbb{C}$). Si $r_1 = r_2 = r$, on vérifie que $u_n = nr^n$ est aussi solution ; les solutions s'écrivent donc (si $r \neq 0$) $u_n = (\alpha + n\beta)r^n$ (avec $\alpha, \beta \in \mathbb{C}$).

Dans le cas de la suite de Fibonacci, $a = b = 1$ et les racines du polynôme $X^2 - X - 1$ sont ϕ et $-\phi^{-1}$ où ϕ est le nombre d'or. Donc $F_k = \alpha\phi^k + \beta(-1)^k\phi^{-k}$. On détermine α et β à l'aide des premiers termes.

1.4 Nombres premiers entre eux

1.19 Définition. On dit que a et b sont premiers entre eux si leur plus grand commun diviseur est 1.

Si $a, b \in \mathbb{Z}$, on peut écrire $a = a'd$ et $b = b'd$ où a' et b' sont premiers entre eux et d est le plus grand commun diviseur de a et b.

Soient $n \in \mathbb{N}^*$ et $x_1, \ldots, x_n$ des nombres entiers. On dit que les x_i sont *premiers entre eux dans leur ensemble* si le plus grand commun diviseur de $x_1, \ldots, x_n$ est 1 ; on dit que les x_i sont *premiers entre eux deux à deux*, si pour tout couple d'entiers i, j avec $1 \leqslant i < j \leqslant n$, les nombres x_i et x_j sont premiers entre eux.

1.20 Théorème de Bézout. *Soient $a, b \in \mathbb{Z}$. Alors a et b sont premiers entre eux si et seulement s'il existe $u, v \in \mathbb{Z}$ tels que $au + bv = 1$.*

1.21 Théorème de Gauss. *Soient $a, b, c \in \mathbb{Z}$. Si a divise bc et est premier à b, alors a divise c.*

1.22 Corollaire. *Soient $a, b \in \mathbb{Z}$ et p un nombre premier. Si p divise ab alors p divise a ou b.*

1.23 Lemme. *Soient $p_1, \ldots, p_k \in \mathbb{N}$ des nombres premiers distincts deux à deux et $\beta_1, \ldots \beta_k \in \mathbb{N}^*$. Posons $n = \prod_{j=1}^{k} p_j^{\beta_j}$. L'ensemble des diviseurs premiers de n est $\{p_1, \ldots, p_k\}$ et pour tout j, $p_j^{\beta_j}$ divise n et $p_j^{\beta_j+1}$ ne divise pas n.*

1.24 Théorème. *Tout nombre entier admet une décomposition en produit de nombres premiers unique à permutation des termes près.*

On démontre l'existence à l'aide d'une « récurrence forte » sur n. L'unicité résulte du lemme.

1.5 Congruences, l'anneau $\mathbb{Z}/n\mathbb{Z}$

1.25 Définition. Soient $a, b, n \in \mathbb{Z}$. On dit que a est congru à b modulo n et on écrit $a \equiv b\ [n]$ si n divise $b - a$.

1.26 Proposition. *Soit $n \in \mathbb{Z}$. La relation de congruence modulo n est une relation d'équivalence.*

1.27 Lemme. *Soit p un nombre premier. Pour tout entier k tel que $1 \leqslant k \leqslant p - 1$ le coefficient binomial $\binom{p}{k}$ est divisible par p.*

On a $(p - k)\binom{p}{k} = p\binom{p-1}{k}$.

1.28 Petit théorème de Fermat. *Soit p un nombre premier. Pour tout entier k on a $k^p \equiv k\ [p]$. Si k n'est pas divisible par p, alors $k^{p-1} \equiv 1\ [p]$.*

1.29 Théorème de Wilson. *Pour tout nombre premier p, on a $(p-1)! \equiv -1\ [p]$.*

1.30 Définition. Soit $n \in \mathbb{Z}$. On note $\mathbb{Z}/n\mathbb{Z}$ le *quotient d'équivalence* pour la relation de congruence modulo n.

Pour $n \in \mathbb{N}^*$, on a $a \equiv b\ [n]$ si et seulement si a et b ont même reste dans la division euclidienne par n ; on en déduit que $\mathbb{Z}/n\mathbb{Z}$ a n éléments (autant que des restes possibles).

1.31 Proposition. *Soit $n \in \mathbb{Z}$. L'addition et la multiplication de $\mathbb{Z}$ passent au quotient et définissent une structure d'anneau sur $\mathbb{Z}/n\mathbb{Z}$.*

En d'autres termes, si $a \equiv b\ [n]$ et $a' \equiv b'\ [n]$, alors $a + a' \equiv b + b'\ [n]$ et $aa' \equiv bb'\ [n]$.

1.32 Proposition. *Soit $n \in \mathbb{N}^*$. Les propriétés suivantes sont équivalentes.*

(i) *L'anneau $\mathbb{Z}/n\mathbb{Z}$ est un corps.*

(ii) *Le nombre n est premier.*

1.33 Proposition. *Soient $n \in \mathbb{Z}$. La classe d'un élément $a \in \mathbb{Z}$ est un élément inversible de l'anneau $\mathbb{Z}/n\mathbb{Z}$ si et seulement si a et n sont premiers entre eux.*

Pour $n \in \mathbb{N}^*$, le nombre d'éléments inversibles de $\mathbb{Z}/n\mathbb{Z}$ est donc égal au nombre d'entiers $a \in [0, n-1]$ premiers à n. Ce nombre se note $\varphi(n)$. L'application φ ainsi construite s'appelle *l'indicatrice d'Euler.*

Soit p un nombre premier. Tout nombre non divisible par p est premier à p ; on a donc $\varphi(p) = p - 1$. Soient $n \in \mathbb{N}^*$ et $a \in \mathbb{Z}$; alors a est premier avec p^n si et seulement si a est premier avec p, *i.e.* s'il n'est pas divisible par p. Les nombres $a \in [0, p^n - 1]$ divisibles par p sont les kp avec $0 \leqslant k < p^{n-1}$. Ils sont au nombre de p^{n-1}. Donc $\varphi(p^n) = p^n - p^{n-1} = (p-1)p^{n-1}$.

1.34 Remarque. Soient $m, n \in \mathbb{Z}$ deux nombres entiers. On suppose que $m|n$. Pour $a, b \in \mathbb{Z}$, si $a \equiv b\ [n]$, alors *a fortiori* $a \equiv b\ [m]$. On définit une application naturelle $\pi_{m,n} : \mathbb{Z}/n\mathbb{Z} \to \mathbb{Z}/m\mathbb{Z}$ qui à la classe de a modulo n associe sa classe modulo m. C'est clairement un homomorphisme d'anneaux.

1.35 Théorème « Chinois ». *Soient m, n deux nombres premiers entre eux. L'application*

$$\begin{array}{rcl} \mathbb{Z}/mn\mathbb{Z} & \to & \mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z} \\ a & \mapsto & (\pi_{m,mn}(a), \pi_{n,mn}(a)) \end{array}$$

est bijective ; c'est un isomorphisme d'anneaux.

En particulier, si m, n sont premiers entre eux on a $\varphi(mn) = \varphi(m)\varphi(n)$.

1.36 Proposition. *Soit $n \in \mathbb{N}$, $n \geqslant 2$. Notons $p_1, \ldots, p_k$ les nombres premiers (positifs et) distincts qui divisent n. On a $\varphi(n) = n \prod_{j=1}^{k} \left(1 - \frac{1}{p_j}\right)$.*

1.37 Résolution générale de deux équations type. On va donner une méthode générale pour deux équations : un système de congruences et une équation diophantienne. Chacune de ces équations demande d'abord un calcul de plus grand commun diviseur et une « relation de Bézout ».

a) Résoudre l'équation de congruences : $x \equiv a\ [m]$ et $x \equiv b\ [n]$.

- **On suppose que m et n sont premiers entre eux.** Écrivons une relation de Bézout $mu + nv = 1$. Posons $x_0 = mub + nva$. Alors $x_0 - a = mub + (nv - 1)a = mub - mua$ est un multiple de m et $x_0 - b = (mu - 1)b + nva = nv(a - b)$ est un multiple de n. Notre équation devient
$$x \equiv x_0 \ [m] \ \text{ et } \ x \equiv x_0 \ [n],$$
qui est équivalente à $x \equiv x_0 \ [mn]$. L'ensemble de ses solutions est $\{x_0 + mnk;\ k \in \mathbb{Z}\}$.
- **Cas général.** Notons d le plus grand commun diviseur de m et n. Si x est solution de notre équation, comme d divise $x - a$ et $x - b$, alors $d|b - a$. Si a n'est pas congru à b modulo d, alors notre équation n'a pas de solution. Sinon, écrivons $b - a = \ell d$ et écrivons une relation de Bézout $mu + nv = d$. Posons $x_0 = a + \ell mu = a + \ell(d - nv) = b - n\ell v$. C'est une solution de notre équation. Notre équation devient
$$x \equiv x_0 \ [m] \ \text{ et } \ x \equiv x_0 \ [n],$$
qui est équivalente à $x \equiv x_0 \ [M]$ où $M = \dfrac{|mn|}{d}$ est le plus petit commun multiple de m et n. L'ensemble de ses solutions est $\{x_0 + Mk;\ k \in \mathbb{Z}\}$.

b) Résoudre l'équation diophantienne : $ax + by = c$.

On va supposer que a n'est pas nul. Notons d le plus grand commun diviseur de a et b. Écrivons $a = da'$ et $b = db'$ où a' et b' sont deux nombres premiers entre eux, et donnons une relation de Bézout $a'u + b'v = 1$. L'équation devient $d(a'x + b'y) = c$. Si c n'est pas multiple de d, il n'y a pas de solution. Sinon, écrivons $c = dc'$. L'équation devient $a'x + b'y = c' = c'(a'u + b'v)$, soit $a'(x - c'u) = b'(c'v - y)$. Si (x, y) est solution, alors a' divise $b'(c'v - y)$ et est premier avec b', donc il divise $c'v - y$. Écrivons $c'v - y = ka'$. On doit alors avoir : $a'(x - c'u) = a'b'k$, donc $x - c'u = b'k$. L'ensemble des solutions est contenu dans $\{(c'u + b'k, c'v - ka');\ k \in \mathbb{Z}\}$. On vérifie immédiatement que, inversement, pour tout $k \in \mathbb{Z}$, on a $a(c'u + b'k) + b(c'v - ka') = c$.

Remarquons que dans ces deux équations on a trouvé une *solution particulière* et résolu l'*équation homogène associée.* ***Pourquoi ?***

1.6 Exercices

1.6.1 Divisibilité et congruences

1.1 Exercice. 1. Soient $a, b, \delta \in \mathbb{Z}$. On suppose que δ est un diviseur commun de a et b et qu'il existe $u, v \in \mathbb{Z}$ tels que $\delta = au + bv$. Démontrer que le plus grand commun diviseur de a et b est $|\delta|$.

2. Soient $a, b, c \in \mathbb{N}$. Notons d et m le plus grand commun diviseur et le plus petit commun multiple de a et b. Démontrer que le plus grand commun diviseur de ac et bc est dc et que le plus petit commun multiple de ac et bc est mc.

3. Soient $a, b \in \mathbb{N}$. Démontrer que $dm = ab$ où l'on a noté d et m le plus grand commun diviseur et le plus petit commun multiple de a et b respectivement.

1.2 Exercice. 1. Soient $a, b, c \in \mathbb{Z}$. On suppose que a et b sont premiers entre eux, que $a|c$ et $b|c$. Démontrer que $ab|c$.

2. Soient $a, b, c \in \mathbb{Z}$. On suppose que a est premier à b et à c. Démontrer que a est premier à bc.

3. Soient $a, b \in \mathbb{Z}$ et $n \in \mathbb{N}^*$.

   a) On suppose que a et b sont premiers entre eux. Démontrer que a et b^n sont premiers entre eux. En déduire que a^n et b^n sont premiers entre eux.

b) Démontrer que le plus grand commun diviseur de a^n et b^n est d^n où d est le plus grand commun diviseur de a et b.

4. Soient $a, b, c \in \mathbb{Z}$ tels que $a|bc$. Démontrer qu'il existe $d, e \in \mathbb{Z}$ tels que $a = de$ et $d|b$ et $e|c$.

1.3 Exercice. *Propriétés arithmétiques à la base de RSA.*
Soient p, q deux nombres premiers distincts, on note N un multiple commun de $p-1$ et $q-1$. Soit $e \in \{1, \dots, N\}$ un entier premier avec N.

1. Démontrer qu'il existe un entier $d \in \{1, \dots, N\}$ tel que $ed \equiv 1[N]$.
2. En utilisant le théorème de Fermat, démontrer que pour tout entier n, $n^{ed} \equiv n[p]$ et $n^{ed} \equiv n[q]$.
3. En déduire que l'application $C : \{0, \dots, pq-1\} \to \{0, \dots, pq-1\}$ qui à a associe le reste dans la division de a^e par pq est une bijection de $\{0, \dots, pq-1\}$ sur lui-même.

Sur le système de cryptage à clé appelé RSA (Ron Rivest, Adi Shamir, and Leonard Adleman, 1977) :
Je veux pouvoir recevoir des messages chiffrés de telle sorte que je serai seul à pouvoir les déchiffrer. Pour cela

- Je choisis deux nombres premiers p et q grands (environ 100 chiffres chacun), je calcule leur produit n que je rends public, ainsi que la clé de chiffrement e - un nombre premier à $(p-1)(q-1)$.
- Je calcule aussi un nombre d qui est inverse de e modulo $p-1$ et modulo $q-1$; ce nombre je suis le seul à le connaître, ainsi que les nombres p et q qui m'ont permis de le trouver.

Supposons maintenant que vous vouliez m'envoyer de façon secrète un message qui est un nombre a ayant à peu près 200 chiffres, c'est à dire grand mais inférieur à $n = pq$ (ou une suite a_i de tels nombres si votre message est long). Vous m'envoyez juste le nombre b qui est le reste de a^e dans la division par n (ou la suite des $b_i \equiv a_i^e$ modulo n). Pour retrouver le message d'origine, je n'aurai qu'à calculer le reste de b^d (ou b_i^d) modulo n. Ce système repose sur les faits suivants :

1. Il est « relativement rapide » de vérifier qu'un nombre est premier, et il y a beaucoup de nombres premiers : si je donne un nombre m de 100 chiffres au hasard, d'après le théorème des nombres premiers, le plus petit nombre premier $p > m$ a beaucoup de chances d'être tel que $p - m$ soit du même ordre que $\ln m \sim 100 \ln 10$. Donc je peux trouver des nombres premiers p et q « rapidement ».
2. Le nombre e est en général choisi petit ($e = 3, 5$ ou 7 sont des choix courants). Le nombre d est par contre grand (200 chiffres...). Élever à la puissance d modulo n un nombre x est cependant une opération « rapide » : cela implique d'élever des éléments de $\mathbb{Z}/n\mathbb{Z}$ au carré $\log_2 d$ ($\simeq 700$) fois et de multiplier des nombres par x au plus $\log_2 d$ fois.
3. Par contre, on ne sait pas trouver le nombre d connaissant n et e sans trouver p et q, et on ne sait pas trouver la décomposition $n = pq$ rapidement.

1.4 Exercice. *Équations Diophantiennes*
Soient $a, b \in \mathbb{N}^*$ des nombres entiers.

1. Soit $c \in \mathbb{Z}$. Quelles sont toutes les solutions de l'équation $ax + by = c$ avec $(x, y) \in \mathbb{Z}$?
   On suppose dorénavant que a et b sont premiers entre eux.
2. Quel est le plus petit entier qui s'écrit de deux façons sous la forme $ax + by$ avec $x, y \in \mathbb{N}$?
3. On suppose que a et b sont tous deux distincts de 1. Notons A l'ensemble des entiers naturels qui ne peuvent s'écrire sous la forme $ax + by$ avec $x, y \in \mathbb{N}$.
   a) Quel est le plus grand élément de A ?
   b) Démontrer que $A = \{|ua - vb|; (u, v) \in \mathbb{N}^2;\ 1 \leqslant u \leqslant b-1;\ 1 \leqslant v \leqslant a-1\}$.
   c) Combien d'éléments a A ?
4. Rappelons qu'au rugby un essai transformé vaut 7 points, un essai non transformé en vaut 5, un drop ou une pénalité 3.
   a) Quel est le plus grand score pour lequel on est sûr qu'il n'a pas été obtenu que par des essais - transformés ou non ?
   b) Quels sont les scores impossibles ?

1.5 Exercice. *Théorème Chinois.* Soient $a, b \in \mathbb{N}^*$. Posons $d = \text{pgcd}(a, b)$ et $m = \text{ppcm}(a, b)$.

1. Ecrire la décomposition de d et m en facteurs premiers en fonction de celle de a et de b. Comparer cette méthode de calcul de pgcd avec l'algorithme d'Euclide.

2. Démontrer qu'il existe a_1, a_2, b_1, b_2 tels que
   - $a = a_1 a_2$, $b = b_1 b_2$;
   - $a_1 | b_1$, $b_2 | a_2$;
   - a_2 et b_1 sont premiers entre eux.
3. Démontrer que $a_1 b_2 = d$ et $a_2 b_1 = m$.
4. En déduire que $\mathbb{Z}/a\mathbb{Z} \times \mathbb{Z}/b\mathbb{Z}$ est isomorphe à $\mathbb{Z}/d\mathbb{Z} \times \mathbb{Z}/m\mathbb{Z}$.

1.6 Exercice. *Jouons avec la suite de Fibonacci.*

1. Écrire les premiers nombres de Fibonacci. Lesquels sont pairs ? multiples de 3 ? multiples de 5 ?
2. a) Démontrer que, si m divise n alors F_m divise F_n.
   b) Démontrer que pour tout n, l'ensemble des $k \in \mathbb{N}$ tel que n divise F_k est de la forme $a\mathbb{N}$ où $a \in \mathbb{N}^*$. *(Utiliser l'exercice 1.7.3).*
3. Soit $p \geqslant 7$ un nombre premier. Notons J la matrice $\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$ à coefficients dans $\mathbb{F}_p$.
   a) On suppose que 5 est un carré modulo p. Démontrer que la matrice J est diagonalisable (dans $\mathbb{F}_p$). En déduire que F_{p-1} est multiple de p.
   b) (**) On suppose que 5 n'est pas un carré modulo p. Notons $K = \{aI_2 + bJ;\ a, b \in \mathbb{F}_p\}$. Démontrer que
      (i) K est un sous-anneau commutatif de $M_2(\mathbb{F}_p)$;
      (ii) l'anneau K est un corps ;
      (iii) l'application $x \mapsto x^p$ est un automorphisme de K ;
      (iv) pour $x \in K$ on a $x^p = x \iff x \in \{aI_2;\ a \in \mathbb{F}_p\}$;
      (v) posant $J' = J^p$, on a $J' \neq J$ et $J'^2 = J' + 1$;
      (vi) on a $J^p = -J^{-1}$;
      (vii) p divise F_{p+1} ; de plus $F_p \equiv F_{p+2} \equiv -1 \ \ [p]$.

1.7 Exercice. *Algorithme d'Euclide et matrices* 2×2.
Soient $a, b \in \mathbb{N}$, avec $0 < a < b$. On effectue l'algorithme d'Euclide : on pose $r_0 = b$, $r_1 = a$, et, supposant r_{j-1} et r_j construits, si $r_j \neq 0$ on note $r_{j-1} = r_j q_j + r_{j+1}$ la division euclidienne de r_{j-1} par r_j. On note n l'entier pour lequel l'algorithme s'arrête de sorte que $r_{n+1} = 0$ et r_n est le $PGCD$ de a, b.

1. Démontrer que $q_n \geqslant 2$.
2. a) Démontrer que, pour tout $k \in \{1, \ldots, n\}$, on a
$$\begin{pmatrix} 0 & 1 \\ 1 & q_1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & q_2 \end{pmatrix} \cdots \begin{pmatrix} 0 & 1 \\ 1 & q_k \end{pmatrix} \begin{pmatrix} r_{k+1} \\ r_k \end{pmatrix} = \begin{pmatrix} r_1 \\ r_0 \end{pmatrix}.$$
   b) Démontrer qu'il existe des suites $(a_k)_{1 \leqslant k \leqslant n+1}$ et $(b_k)_{1 \leqslant k \leqslant n+1}$ de nombres entiers telles que pour $k \in \{1, \ldots, n\}$ on ait
$$\begin{pmatrix} 0 & 1 \\ 1 & q_1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & q_2 \end{pmatrix} \cdots \begin{pmatrix} 0 & 1 \\ 1 & q_k \end{pmatrix} = \begin{pmatrix} a_k & a_{k+1} \\ b_k & b_{k+1} \end{pmatrix}.$$
   c) Démontrer que $a_1 = 0,\ a_2 = 1,\ b_1 = 1,\ b_2 = q_1$ et, pour $2 \leqslant j \leqslant n$, on a $a_{j+1} = a_j q_j + a_{j-1}$ et $b_{j+1} = b_j q_j + b_{j-1}$. En déduire que les suites a_k et b_k sont croissantes et que l'on a $a_{n+1} \geqslant 2a_n$ et $b_{n+1} \geqslant 2b_n$. Dans quel cas a-t-on égalité dans l'une de ces inégalités ?
   d) Démontrer que l'on a $a_k b_{k+1} - a_{k+1} b_k = (-1)^k$ pour $1 \leqslant k \leqslant n$.
   e) Démontrer que l'on a une relation de Bézout $r_n = (-1)^n a_n b + (-1)^{n+1} b_n a$.

3. a) Démontrer que $\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}^k = \begin{pmatrix} F_{k-1} & F_k \\ F_k & F_{k+1} \end{pmatrix}$ où F_k est le k-ème nombre de Fibonacci.

   b) Démontrer que $b_k \geqslant F_k$ et $a_k \geqslant F_{k-1}$.

4. Expliquer en quoi cette méthode permet de trouver « rapidement » le $PGCD$ de a et b et une identité de Bézout $d = au + bv$.

5. On suppose que a et b sont premiers entre eux. Démontrer qu'il existe $n \in \mathbb{N}^*$ une suite $q_1, \ldots, q_n$ de nombres entiers strictement positifs et $u, v \in \mathbb{N}$ tels que $q_n \geqslant 2$ et

$$\begin{pmatrix} 0 & 1 \\ 1 & q_1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & q_2 \end{pmatrix} \cdots \begin{pmatrix} 0 & 1 \\ 1 & q_n \end{pmatrix} = \begin{pmatrix} u & a \\ v & b \end{pmatrix}.$$

6. On suppose qu'il existe une suite $q_1, \ldots, q_n$ de nombres entiers strictement positifs et $u, v \in \mathbb{N}$ tels que $q_n \geqslant 2$ et

$$\begin{pmatrix} 0 & 1 \\ 1 & q_1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & q_2 \end{pmatrix} \cdots \begin{pmatrix} 0 & 1 \\ 1 & q_n \end{pmatrix} = \begin{pmatrix} u & a \\ v & b \end{pmatrix}.$$

   Démontrer que a et b sont premiers entre eux et que la suite des quotients successifs de la division euclidienne de b par a est $q_1, q_2, \ldots, q_n$.

1.8 Exercice. *Algorithme de Cornacchia* (**)

1. Soient $a, b \in \mathbb{N}$ tels que $a < b$ et $a^2 + b^2$ soit un nombre premier p.

   a) Démontrer que a et b sont premiers entre eux.

   b) Démontrer qu'il existe $n \in \mathbb{N}^*$, des nombres entiers strictement positifs $q_1, \ldots, q_n$ avec $q_n \geqslant 2$ et des nombres $u, v \in \mathbb{N}$ tels que

$$\begin{pmatrix} 0 & 1 \\ 1 & q_1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & q_2 \end{pmatrix} \cdots \begin{pmatrix} 0 & 1 \\ 1 & q_n \end{pmatrix} = \begin{pmatrix} u & a \\ v & b \end{pmatrix}.$$

   c) Démontrer que $2u \leqslant a$ et $2v \leqslant b$.

   d) Démontrer que $\begin{pmatrix} u & v \\ a & b \end{pmatrix} \begin{pmatrix} u & a \\ v & b \end{pmatrix}$ s'écrit $\begin{pmatrix} x & \ell \\ \ell & p \end{pmatrix}$ où ℓ est l'unique entier tel que $\ell^2 \equiv -1\ [p]$ et $0 \leqslant \ell < p/2$.

2. Soit $p > 2$ un nombre premier tel que -1 est un carré modulo p (*i.e.* congru à 1 modulo 4 - voir exercice 1.11). Supposons qu'on ait trouvé ℓ tel que $0 \leqslant \ell < p/2$ et $\ell^2 = xp - 1$ avec $x \in \mathbb{N}$. Expliquer comment, grâce à l'algorithme d'Euclide, on trouve alors a et b tels que $a^2 + b^2 = p$.

1.6.2 Nombres premiers

1.9 Exercice. *Nombres de Fermat, nombres de Mersenne.*
Pour tout entier $n \geqslant 1$, on note $f_n = 2^n + 1$ et $M_n = 2^n - 1$.

1. Soit $n \geqslant 1$ un entier.
   Démontrer que si M_n est premier, alors n aussi, et que si f_n est premier alors n est une puissance de 2.

   Indication : Remarquer que, pour $a \in \mathbb{Z}$ et $k, m \in \mathbb{N}$, $a^k - 1$ divise $a^{km} - 1$ et, si m est impair $a^k + 1$ divise $a^{mk} + 1$.

   On pose $F_k = f_{2^k}$.

2. Soient k, ℓ deux nombres entiers avec $k < \ell$. Démontrer que $2^{2^\ell} \equiv 1[F_k]$. En déduire que F_k et F_ℓ sont premiers entre eux.

3. Soit $p > 2$ un nombre premier et soit q un diviseur premier de M_p. Quel est l'ordre de 2 dans le groupe $(\mathbb{F}_q^*, \cdot)$? En déduire que q est de la forme $2kp + 1$.

4. Démontrer que M_{13} est premier.
5. De même soit $\ell \in \mathbb{N}$ et q un diviseur premier de F_ℓ.
   a) Quel est l'ordre de 2 dans le groupe $(\mathbb{F}_q^*, \cdot)$?
   b) En déduire que q est de la forme $2^{\ell+1}k + 1$.
   c) On suppose que $\ell \geqslant 2$. Notons ω le classe de $2^{2^{\ell-2}}$ dans $\mathbb{F}_q$. Remarquant que $\omega^4 = -1$, démontrer que $2 = (\omega + \omega^{-1})^2$ est un carré modulo q. En déduire que $2^{\ell+2}$ divise $q - 1$.
   d) Démontrer que le plus petit diviseur de F_5 distinct de 1 est $\geqslant 641$.
   e) En remarquant que $641 = 5^4 + 2^4$, démontrer que $F_5 \equiv 1 - 5^4 2^{28}$ [641].
   f) Démontrer que $641 | F_5$.

1.10 Exercice. *Cas du théorème de Dirichlet.* (*cf.* Combes. *Algèbre et géométrie* 12.6).

Théorème de Dirichlet. *Soient $a, b \in \mathbb{N}^*$ premiers entre eux. Il y a une infinité de nombres premiers congrus à a modulo b.*

Nous étudions ici le cas où $a = 1$.

Le cas $b = 4$: Soit $a \in \mathbb{N}$ et p un diviseur premier de $a^2 + 1$ distinct de 2.

1. Démontrer que a et p sont premiers entre eux.
2. On note x la classe de a dans $\mathbb{F}_p$. Démontrer que $x^4 = 1$.
3. Démontrer que $x^2 \neq 1$.
4. En déduire que p est congru à 1 modulo 4.
5. En prenant a sous-la forme $n!$, démontrer qu'il existe une infinité de nombres premiers congrus à 1 modulo 4.
6. Démontrer que, pour $n \geqslant 4$, $n! - 1$ a au moins un diviseur premier congru à 3 modulo 4. En déduire qu'il existe une infinité de nombres premiers congrus à 3 modulo 4.

Le cas $b = 6$ Soit $a \in \mathbb{N}$ et p un diviseur premier de $a^2 + a + 1$ distinct de 3.

1. Démontrer que a et p sont premiers entre eux.
2. On note x la classe de a dans $\mathbb{F}_p$. Démontrer que $x^3 = 1$.
3. Démontrer que $x \neq 1$.
4. En déduire que p est congru à 1 modulo 3.
5. En prenant a sous-la forme $n!$, démontrer qu'il existe une infinité de nombres premiers congrus à 1 modulo 6.
6. Démontrer que, pour $n \geqslant 3$, $n! - 1$ a au moins un diviseur premier congru à 5 modulo 6. En déduire qu'il existe une infinité de nombres premiers congrus à 5 modulo 6.

Le cas $b = 12$: Soit $a \in \mathbb{N}$ et p un diviseur premier de $a^4 - a^2 + 1$.

1. Démontrer que $p \neq 2$ et $p \neq 3$. Démontrer que a et p sont premiers entre eux.
2. On note x la classe de a dans $\mathbb{F}_p$. Démontrer que $x^{12} = 1$.
3. Démontrer que $x^4 \neq 1$ et $x^6 \neq 1$.
4. En déduire que p est congru à 1 modulo 12.
5. En prenant a sous-la forme $n!$, démontrer qu'il existe une infinité de nombres premiers congrus à 1 modulo 12.

Le cas général [**] Pour $n \in \mathbb{N}^*$, on note Φ_n le n-ième polynôme cyclotomique : $\Phi_n = \prod_{0 \leqslant k < n;\ k \wedge n = 1} X - e^{\frac{2ik\pi}{n}}$. Rappelons que $\Phi_n \in \mathbb{Z}[X]$ et que l'on a l'égalité $X^n - 1 = \prod_{d|n} \Phi_d$.

Soient $n \in \mathbb{N}$, $n \geqslant 2$, $a \in \mathbb{N}$ un multiple de n et p un diviseur premier de $\Phi_n(a)$.

1. Démontrer que $\Phi_n(0) = 1$. En déduire que a et p sont premiers entre eux.
2. On note x la classe de a dans $\mathbb{F}_p$. Démontrer que $x^n = 1$.
3. Démontrer que le polynôme $X^n - 1$ n'a pas de facteur carré dans $\mathbb{F}_p[X]$.

   Indication : Utiliser la dérivée
4. Soit $d \in \mathbb{N}$ un diviseur de n distinct de n. Démontrer que les polynômes $X^d - 1$ et Φ_n sont premiers entre eux dans $\mathbb{F}_p[X]$. En déduire que $x^d \neq 1$.
5. En déduire que p est congru à 1 modulo n.
6. En prenant a sous-la forme $N!$, démontrer qu'il existe une infinité de nombres premiers congrus à 1 modulo n.

1.11 Exercice. *Carrés dans $\mathbb{F}_p$.* (*cf.* COMBES p. 267)
Soit p un nombre premier distinct de 2. Notons $C \subset \mathbb{F}_p^*$ l'ensemble des carrés, *i.e.* l'ensemble des $x \in \mathbb{F}_p^*$ tels qu'il existe $y \in \mathbb{F}_p^*$ avec $x = y^2$.

1. Le cas de -1.
   a) Démontrer que pour tout $x \in C$ il existe un et un seul $c \in \left\{1, \ldots, \dfrac{p-1}{2}\right\}$ tel que x soit la classe de c^2. Combien y a-t-il de carrés dans F_p^* ?
   b) Démontrer que tout $x \in C$, on a $x^{\frac{p-1}{2}} = 1$.
   c) En déduire que, pour $x \in \mathbb{F}_p^*$, on a $x \in C \iff x^{\frac{p-1}{2}} = 1$.
   d) Démontrer que -1 est un carré modulo p si et seulement si p est congru à 1 modulo 4.
2. Le cas de 3.
   a) Soit $P = X^2 + aX + b$ un polynôme à coefficients dans $\mathbb{F}_p$. Démontrer que P a une racine dans $\mathbb{F}_p$ si et seulement si $a^2 - 4b$ est un carré (*i.e.* $a^2 - 4b \in \{0\} \cup C$).
   b) On suppose que $p \notin \{2, 3\}$. Démontrer l'équivalence entre
      (i) $-3 \in C$.
      (ii) Il existe $x \in \mathbb{F}_p^*$, $x^2 + x + 1 = 0$.
      (iii) Il existe x d'ordre 3 dans le groupe $\mathbb{F}_p^*$.
      (iv) $p \equiv 1$ [3] (ce qui signifie encore $p \equiv 1$ [6]).
3. Le polynôme $X^4 + 1$.
   a) Démontrer que si $a, b \in \mathbb{F}_p^* \setminus C$, alors $ab \in C$.
   b) En déduire qu'un au moins des éléments $-1, 2, -2$ est un carré dans $\mathbb{F}_p$.
   c) En écrivant $X^4 + 1 = (X^2 + 1)^2 - 2X^2 = (X^2 - 1)^2 + 2X^2$ en déduire que pour tout p le polynôme $X^4 + 1$ n'est pas irréductible dans $\mathbb{F}_p[X]$.
   d) Quelle est la décomposition dans $\mathbb{R}[X]$ du polynôme $X^4 + 1$ en polynômes irréductibles ?
   e) En déduire que $X^4 + 1$ est irréductible sur $\mathbb{Q}$ (et sur $\mathbb{Z}$).

1.12 Exercice. *Réciprocité quadratique pour 2, pour 5.*
Soit p un nombre premier.

1. Soit L un corps commutatif de caractéristique p, autrement dit une extension de $\mathbb{F}_p$.
   a) Démontrer que $x \mapsto x^p$ est un endomorphisme de corps de L.
   b) Quelles sont les racines du polynôme $X^p - X$ dans L ?
2. On suppose que p est distinct de 2. Soit L une extension de $\mathbb{F}_p$ et $\omega \in L$ tel que $\omega^4 = -1$. Une telle extension existe d'après le corollaire 3.18. Posons $x = \omega + \omega^{-1}$.
   a) Démontrer que $\omega^2 + \omega^{-2} = 0$ et $x^2 = 2$.
   b) Démontrer que les assertions suivantes sont équivalentes :

(i) Il existe $y \in \mathbb{F}_p$ tel que $y^2 = 2$.
(ii) $x \in \mathbb{F}_p$;
(iii) $x^p = x$;
(iv) $\omega^p = \omega$ ou $\omega^p = \omega^{-1}$;
(v) $p \equiv \pm 1$ [8] ;

3. On suppose que p est distinct de 2 et de 5. Soit L une extension de $\mathbb{F}_p$ et $\omega \in L$ tel que $\omega^5 = 1$ et $\omega \neq 1$ (*i.e.* une racine du polynôme $1 + X + X^2 + X^3 + X^4$ - une telle extension L existe d'après le corollaire 3.18). Posons $x = \omega + \omega^{-1}$.
   a) Démontrer que $\omega^2 + \omega^{-2} = -1 - x$ et $x^2 + x - 1 = 0$.
   b) Démontrer que les assertions suivantes sont équivalentes :
      (i) Il existe $y \in \mathbb{F}_p$ tel que $y^2 = 5$.
      (ii) $x \in \mathbb{F}_p$;
      (iii) $x^p = x$;
      (iv) $\omega^p = \omega$ ou $\omega^p = \omega^{-1}$;
      (v) $p \equiv \pm 1$ [5] ;
      (vi) La classe de p est un carré modulo 5.

1.13 Exercice. *Racine carrée de -1 dans $\mathbb{F}_p$.*

Soit p un (grand !) nombre premier. Soit $x \in \mathbb{F}_p^*$.

1. Démontrer que x est un carré dans $\mathbb{F}_p$ si et seulement si $x^{(p-1)/2} = 1$. *(Voir exercice 1.11).*
   On suppose que x est un carré et on veut trouver une racine carrée de x.
2. On suppose que $p \equiv 3$ [4]. Démontrer que, si x est un carré, alors $x^{\frac{p+1}{4}}$ est une racine carrée de x.
3. On suppose que $p \equiv 1$ [4] et on cherche une racine carrée de -1. On écrit $p - 1 = 2^\ell u$ avec u entier impair.
   a) Soit $a \in \mathbb{F}_p^*$; posons $b = a^u$. Démontrer que b est d'ordre 2^k avec $0 \leqslant k \leqslant \ell$.
   b) En choisissant a au hasard, quelle est la probabilité que $b = \pm 1$?
   c) Expliquer comment trouver une racine carrée de -1 si $b \neq \pm 1$.

1.14 Exercice. 1. *Les nombres premiers sont espacés.* Démontrer que pour tout $n \in \mathbb{N}$, il existe une suite de n nombres consécutifs non premiers (*i.e.* il existe $a \in \mathbb{N}$ tel que les nombres entiers k avec $a \leqslant k \leqslant a + n - 1$ ne soient pas premiers).

2. *Il y a beaucoup de nombres premiers.* On désigne par $(p_n)_{n \geqslant 1}$ la suite ordonnée des nombres premiers. On veut démontrer que la série $\displaystyle\sum_{n=1}^{+\infty} 1/p_n$ diverge.

   On suit Combes (p. 269).
   Soit $k \in \mathbb{N}$. Notons $p_1, \ldots, p_k$ les k plus petits nombres premiers et $A_k \subset \mathbb{N}^*$ l'ensemble des nombres entiers dont tous les diviseurs premiers sont $\leqslant p_k$.
   a) Démontrer que tout $a \in A_k$ s'écrit sous la forme $a = b^2 p_1^{\varepsilon_1} \ldots p_k^{\varepsilon_k}$ avec $b \in \mathbb{N}$ et $\varepsilon_j \in \{0, 1\}$. En déduire que, pour tout $x \in \mathbb{N}^*$, le nombre d'éléments de A_k inférieurs à x est $\leqslant \sqrt{x} 2^k$.
   b) Démontrer que, pour $x \in \mathbb{N}^*$, la proportion d'éléments $\mathbb{N} \setminus A_k$ dans $[1, x]$ est plus petite que $\displaystyle\sum_{p \in \mathcal{P},\ p_k < p \leqslant x} 1/p$.
   c) Démontrer que pour $x = 4^{k+1}$ on a $\displaystyle\sum_{p \in \mathcal{P},\ p_k < p \leqslant x} 1/p \geqslant 1/2$. En déduire que la série $\displaystyle\sum_{n=1}^{+\infty} 1/p_n$ diverge.

3. Démontrer que pour tout entier $k \geqslant 1$, $\displaystyle\prod_{i=1}^{k} \frac{p_i}{p_i - 1} \geqslant \sum_{i=1}^{k} \frac{1}{i}$.

4. Démontrer qu'il existe une infinité de nombres premiers comportant au moins un 9 dans leur développement décimal.

D'après le théorème des nombres premiers, $\pi(x)$ est équivalent à $\dfrac{x}{\ln x}$. Les inégalités de Tchebychef, ci-dessous s'approchent de cet équivalent.

1.15 Exercice. *Inégalités de Tchebychef*

1. Pour $N \in \mathbb{Z}^*$ et un nombre premier p, on appelle *valuation* p-adique de N et on note $v_p(N)$ le plus grand entier k tel que $p^k | N$ - de sorte que l'on a $|N| = \displaystyle\prod_p p^{v_p(N)}$.

   Soient $n \in \mathbb{N}$, $n \geqslant 3$ et p un nombre premier.

   a) Démontrer que l'on a $v_p(n!) = \displaystyle\sum_{k=1}^{+\infty} E\big(np^{-k}\big)$ (où E désigne la partie entière).

   b) En déduire que $v_p\binom{2n}{n}$ est le nombre de $k \in \mathbb{N}$ tel que $E\big(2np^{-k}\big)$ soit impair.

   c) Démontrer que
   - $v_p\binom{2n}{n} \leqslant \dfrac{\ln 2n}{\ln p}$.
   - Si $n < p \leqslant 2n$ alors $v_p\binom{2n}{n} = 1$.
   - Si $p \leqslant n < \dfrac{3p}{2}$ alors $v_p\binom{2n}{n} = 0$.

   d) Démontrer que l'on a :

   (i) $\ln\binom{2n}{n} \geqslant \displaystyle\sum_{n \leqslant p \leqslant 2n;\ p \text{ premier}} \ln p$.

   (ii) $\ln\binom{2n}{n} \leqslant (\ln 2n)\Big(\pi(2n/3) + \pi(2n) - \pi(n)\Big) \leqslant (\ln 2n)\pi(2n)$.

2. Soit $n \in \mathbb{N}^*$. Démontrer que $\displaystyle\sum_{k=0}^{n-1} \binom{2n-1}{k} = 2^{2n-2}$. En déduire que $\dfrac{2^{2n-2}}{n} \leqslant \binom{2n-1}{n-1} \leqslant 2^{2n-2}$, puis que $\dfrac{2^{2n-1}}{n} \leqslant \binom{2n}{n} \leqslant 2^{2n-1}$.

3. Démontrer que, pour tout $n \in \mathbb{N}$, $n > 2$, on a

   a) $\displaystyle\sum_{n \leqslant p \leqslant 2n;\ p \text{ premier}} \ln p \leqslant (2n-1)\ln 2$ et en déduire que $\displaystyle\sum_{p \leqslant n;\ p \text{ premier}} \ln p \leqslant n \ln 4$

   b) $\pi(n) \geqslant \dfrac{n(\ln 2)}{\ln n} - 1$.

2 Anneaux

2.1 Généralités

2.1 Définition. Un anneau est un ensemble A muni de deux lois : la première s'appelle en général l'addition et est notée $+$; la deuxième s'appelle en général la multiplication et est notée $(x, y) \mapsto xy$. On suppose que :

- Muni de l'addition A est un groupe abélien ; son élément neutre est noté en général 0 ou 0_A en cas d'ambiguïté ; le symétrique d'un élément $x \in A$ pour $+$ s'appelle l'opposé de x et se note $-x$.
- La multiplication est associative et possède un élément neutre, en général noté 1 ou 1_A en cas d'ambiguïté.
- La multiplication est distributive par rapport à l'addition : pour tout $a, b, c \in A$, on a $a(b+c) = ab+ac$ et $(a + b)c = ac + bc$.
  Lorsque la multiplication est aussi commutative, on dit que l'anneau A est abélien ou commutatif.

2.2 Exemples. a) Munis des opérations (addition et multiplication) usuelles, les ensembles $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ sont des anneaux commutatifs, ainsi que l'anneau $K[X]$ des polynômes sur un corps (ou un anneau) commutatif K.

b) L'ensemble des matrices carrées de taille n à coefficients dans $\mathbb{R}$, muni de l'addition et de la multiplication des matrices est un anneau non commutatif pour $n \geqslant 2$.

Si A et B sont deux anneaux, une application $f : A \to B$ est appelée un *homomorphisme (ou morphisme) d'anneaux* si pour tout $x, y \in A$ on a $f(x + y) = f(x) + f(y)$ et $f(xy) = f(x)f(y)$ et si $f(1_A) = 1_B$. (Remarquons qu'on a automatiquement $f(0_A) = 0_B$).

Soient A un anneau et $x \in A$. On définit nx pour $n \in \mathbb{Z}$ en posant $0x = 0$, $1x = x$, puis, pour tout $n \in \mathbb{N}$, $(n + 1)x = (nx) + x$; enfin pour n négatif $nx = -\big((-n)x\big)$. L'application $n \mapsto nx$ est un homomorphisme d'anneaux de $\mathbb{Z}$ dans A.

L'élément $n1_A$ se note parfois n même lorsque cet homomorphisme n'est pas injectif.

On définit de même x^n pour $x \in A$ et $n \in \mathbb{N}$: on pose $x^0 = 1_A$, $x^1 = x$ puis $x^{n+1} = x^n x (= xx^n)$.

2.3 Formule du binôme. Soient A un anneau et $a, b \in A$ deux éléments *permutables* - *i.e.* tels que $ab = ba$. Alors, pour tout $n \in \mathbb{N}$, on a

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

C'est faux si $ab \neq ba$. Par exemple $(a + b)^2 = a^2 + ab + ba + b^2 \neq a^2 + 2ab + b^2$.

2.4 Définition. Soit A un anneau. Un élément $a \in A$ est dit *inversible* (on dit parfois une unité de A) s'il existe a' dans A (nécessairement unique) tel que $a'a = aa' = 1_A$. Si a est inversible, l'élément a' tel que $a'a = aa' = 1_A$ s'appelle l'inverse de a et se note a^{-1}.

2.5 Proposition. *L'ensemble (noté parfois A^{-1}) des éléments inversibles de A est un groupe pour la multiplication.*

2.6 Définition. Un corps est un anneau K tel que $K^{-1} = K - \{0_K\}$.

2.7 Exercice. Soient A un anneau et $a \in A$. Démontrer que a est inversible si et seulement si l'application $b \mapsto ab$ est bijective de A dans A.

2.2 Anneaux intègres ; anneaux principaux

Dans la suite, tous les anneaux seront supposés commutatifs.

2.8 Définition. On dit qu'un anneau commutatif A est *intègre* si le produit de deux éléments non nuls de A est non nul.

2.9 Division ; éléments associés. Dans un anneau commutatif intègre, on peut définir la divisibilité comme dans $\mathbb{Z}$. On dit que a divise b et on écrit $a|b$ s'il existe c (*nécessairement unique* si a n'est pas nul) tel que $b = ac$. Autrement dit $a|b$ si $b \in aA$.

On dira que deux éléments a et b de A sont *associés* si $a|b$ et $b|a$, c'est à dire s'il existe $u \in A$ inversible tel que $a = ub$.

Le sous-ensemble aA de A est un sous-groupe de A. Mais contrairement au cas de $\mathbb{Z}$, les sous-groupes de A sont loin d'être en général tous de cette forme.

2.10 Définition. Soit A un anneau commutatif. On appelle *idéal* de A une partie I de A qui est un sous-groupe de $(A, +)$ et telle que, pour tout $a \in A$ et tout $x \in I$ on ait $ax \in I$.

A un idéal on peut encore associer une relation d'équivalence et définir un anneau quotient :

2.11 Proposition. *Soient A un anneau commutatif et I un idéal dans A. La relation R définie sur A par $a\,R\,b$ si $b - a \in I$ est une relation d'équivalence.*

2.12 Définition. Soient A un anneau commutatif et I un idéal dans A. On note A/I le *quotient d'équivalence* pour la relation R.

2.13 Proposition. *Soient A un anneau commutatif et I un idéal dans A. L'addition et la multiplication de A passent au quotient et définissent une structure d'anneau sur A/I.*

En effet, si $a, b \in A$ et $x, y \in I$, alors $(a+x)+(b+y)\,R\,a+b$ et $(a+x)(b+y) = ab+(ay+x(b+y))\,R\,ab$.

2.14 A retenir. a) Si I est un idéal d'un anneau (commutatif) A, on peut construire un anneau A/I et un homomorphisme surjectif d'anneaux $\pi : A \to A/I$ de noyau I.

b) Inversement, le noyau d'un homomorphisme d'anneaux $\pi : A \to B$ est un idéal.

Pour $a \in A$, l'ensemble aA est un idéal de A. On l'appelle l'idéal principal associé à a.

2.15 Définition. On dit qu'un anneau commutatif est *principal* s'il est intègre et tous ses idéaux sont principaux.

Un idéal étant en particulier un sous-groupe, l'anneau $\mathbb{Z}$ est principal. Nous verrons que si K est un corps commutatif, l'anneau $K[X]$ des polynômes à coefficients dans K est aussi un anneau principal. D'autres exemples d'anneaux principaux et d'anneaux intègres non principaux seront donnés dans les exercices 2.5, 2.7, 2.8.

Dans un anneau principal, la division se comporte essentiellement comme dans $\mathbb{Z}$.

2.16 Théorème. *Soient A un anneau principal et $a, b \in A$.*

a) *Il existe un élément $m \in A$ tel que $aA \cap bA = mA$. L'élément m est un multiple commun de a et de b. Les multiples communs de a et b sont les multiples de m.*

b) *Il existe un élément $d \in A$ tel que $aA + bA = dA$. L'élément d est un diviseur commun de a et de b. Les diviseurs communs de a et b sont les diviseurs de d.*

2.17 Définition. L'élément d de ce théorème s'appelle un plus grand commun diviseur (PGCD) de a et b. L'élément m s'appelle un plus petit commun multiple (PPCM) de a et b.

Un *PGCD* de a et de b n'est en général pas unique : il est unique à multiplication par un élément inversible de A près. On a fait un choix dans $\mathbb{Z}$ en les prenant dans $\mathbb{N}$ ce qui les a rendus uniques. On fait un tel choix aussi dans $K[X]$, mais il n'y a pas en général un choix « meilleur que les autres ».

Comme dans le cas de $\mathbb{Z}$, on peut définir la notion d'éléments premiers entre eux : leurs seuls diviseurs communs sont les éléments inversibles. Alors 1_A est un PGCD, *i.e.* $aA+bA = A$. On a donc le théorème de Bézout dans ce cadre, dont découle le théorème de Gauss :

2.18 Théorème de Bézout. *Soit A un anneau principal. Soient $a, b \in A$. Alors a et b sont premiers entre eux si et seulement s'il existe $u, v \in A$ tels que $au + bv = 1$.*

2.19 Théorème de Gauss. *Soit A un anneau principal. Soient $a, b, c \in A$. Si a divise bc et est premier à b, alors a divise c.*

Le rôle des nombres premiers est ici joué par les éléments irréductibles.

2.20 Définition. Soit A un anneau intègre. Un élément $a \in A$ est dit *irréductible* s'il n'est pas inversible et dans toute décomposition $a = bc$ un des deux facteurs b ou c est inversible.

2.21 Proposition. *Soient A un anneau principal et $a \in A$ non nul. Alors a est irréductible si et seulement si l'anneau quotient A/aA est un corps.*

Pour établir la décomposition en produit d'éléments irréductibles dans un anneau principal, la difficulté est de démontrer que tout élément non nul et non inversible possède un diviseur irréductible, et qu'il n'en possède qu'un nombre fini. Nous esquissons une preuve ci-dessous :

2.22 Lemme. a) *Soit A un anneau principal. Toute suite croissante d'idéaux de A stationne.*

b) *Toute suite décroissante d'idéaux d'intersection non nulle stationne.*

Démonstration. Soit I_n une suite d'idéaux de A.

a) On suppose que la suite I_n est croissante, c'est à dire que, pour $k, \ell \in \mathbb{N}$ avec $k \leqslant \ell$, on a $I_k \subset I_\ell$. On veut démontrer qu'il existe n tel que, pour $k \geqslant \ell$ on a $I_k = I_n$.
Comme la suite I_n est croissante, on vérifie que la réunion des I_n est un idéal J de A.
Puisque A est principal, il existe $a \in A$ tel que $J = aA$. Alors $a \in J$ et il existe $n \in \mathbb{N}$ tel que $a \in I_n$ (par définition d'une réunion). On a alors $J = aA \subset I_n$ donc $J = I_n$ (puisque J est la réunion de I_k). Pour $k \geqslant n$, on a $I_n \subset I_k \subset J = I_n$.

b) On suppose que la suite I_n est décroissante, c'est à dire que, pour $k, \ell \in \mathbb{N}$ avec $k \leqslant \ell$, on a $I_k \supset I_\ell$. Soit a un élément non nul de l'intersection $\bigcap_{k\in\mathbb{N}} I_k$. Pour $k \in \mathbb{N}$, il existe b_k tel que $I_k = b_k A$ (A étant principal). Comme $a \in I_k$, il existe $c_k \in A$ tel que $b_k c_k = a$. Pour $k \leqslant \ell$, on a $I_k \supset I_\ell$, de sorte que $b_k \in I_\ell$: il existe $x \in A$ tel que $b_k = x b_\ell$. Comme $b_k c_k = a = b_\ell c_\ell$, il vient $x c_k = c_\ell$, soit $c_k | c_\ell$. Posons $J_k = c_k A$. La suite J_k est croissante, donc stationne d'après a). Il existe donc n tel que, pour $k \geqslant n$ on ait $J_k = J_n$. Pour $k \geqslant n$, on a $c_k \in J_n$, donc il existe $y \in A$ tel que $c_k = y c_n$; comme $b_k c_k = a = b_n c_n$ il vient $b_n = y b_k$, donc $I_n \subset I_k$, et l'on a l'égalité.

□

2.23 Théorème. *Soient A un anneau principal et $a \in A$ un élément non nul et non inversible.*

a) *Il existe un élément irréductible $p \in A$ tel que $p|a$.*

b) *Il existe un ensemble fini F d'éléments irréductibles de A tels que tout élément irréductible de A qui divise a est associé à un élément de F ; pour tout irréductible p, il existe $n \in \mathbb{N}$ tel que $p^n \not| a$.*

Une fois ce théorème établi, on en déduit immédiatement l'existence de la décomposition en facteurs irréductibles. L'unicité est plus difficile à énoncer mais se démontre comme dans le cas de $\mathbb{Z}$:

2.24 Théorème. *Soient A un anneau principal et $a \in A$ un élément non nul et non inversible. Il existe un entier $n \geqslant 1$ et des éléments irréductibles $p_1, \ldots, p_n \in A$ tels que $a = \prod_{j=1}^{n} p_j$. Cette décomposition est unique à l'ordre des facteurs près : si $a = \prod_{j=1}^{n} p_j = \prod_{j=1}^{m} q_j$, alors $n = m$ et il existe $\sigma \in \mathfrak{S}_n$, i.e. une bijection $\sigma : \{1, \ldots, n\} \to \{1, \ldots, n\}$ telle que p_j soit associé à $q_{\sigma(j)}$ (pour tout j).*

2.3 Anneaux euclidiens

Les anneaux euclidiens sont ceux pour lesquels on dispose d'une division euclidienne. La même preuve que pour $\mathbb{Z}$ démontre qu'ils sont principaux. De plus, dans un anneau euclidien, comme dans $\mathbb{Z}$, on peut calculer le PGCD, écrire une relation de Bézout, résoudre des équations diophantiennes ou de congruence, *etc.* de façon algorithmique.

2.25 Définition. Soit A un anneau commutatif et intègre. On dit que A est *euclidien* s'il existe une application $v : A - \{0\} \to \mathbb{N}$, - appelée *stathme euclidien* telle que pour tous $a, b \in A - \{0\}$ il existe $q, r \in A$ tels que $a = bq + r$ et $r = 0$ ou $v(r) < v(b)$.

2.26 Remarque. En général, on demande de plus que, pour tous $a, b \in A - \{0\}$ tels que $a|b$ on ait $v(a) \leqslant v(b)$. Cette condition est en pratique toujours vérifiée, mais n'est pas utile dans ce qui suit. On peut démontrer que si A possède un stathme qui ne vérifie pas cette propriété, il en possède un qui la vérifie.

L'anneau $\mathbb{Z}$ est euclidien de stathme $a \mapsto |a|$. Nous verrons plus bas que l'anneau $K[X]$ est aussi euclidien : l'application qui à un polynôme associe son degré est un stathme euclidien sur $K[X]$.

2.27 Théorème. *Tout anneau euclidien est principal.*

Démonstration. Soit A un anneau euclidien ; notons v son stathme. Soit I un idéal non nul de A et $a \in I - \{0\}$ tel que $v(a) = \inf\{v(x);\ x \in I - \{0\}\}$. Puisque $a \in I$, on a $aA \subset I$. Soit x un élément de I ; écrivons $x = aq + r$, avec $q, r \in A$ et $r = 0$ ou $r \neq 0$ et $v(r) < v(a)$. Or $r = x - aq \in I$, et on ne peut avoir $r \neq 0$ et $v(r) < v(a)$ par définition de a. Il vient $r = 0$, donc $x \in aA$. Cela prouve que $I = aA$. □

2.28 Remarque. Dans un anneau euclidien, comme pour le cas de $\mathbb{Z}$, on dispose de l'*algorithme d'Euclide* qui permet de calculer en pratique le plus grand commun diviseur de deux éléments.

2.4 Un exemple

Nous allons développer ici assez en détail une famille d'anneaux.

Soit $\tau \in \mathbb{C} - \mathbb{R}$ un entier quadratique, *i.e.* tel qu'il existe $a, b \in \mathbb{Z}$ avec $\tau^2 + a\tau + b = 0$. Il est alors immédiat que l'ensemble $\mathbb{Z} + \tau\mathbb{Z} = \{m + n\tau;\ (m, n) \in \mathbb{Z}^2\}$ est un sous anneau - noté $\mathbb{Z}[\tau]$ de $\mathbb{C}$. Inversement, si $\mathbb{Z} + \tau\mathbb{Z}$ est un anneau, alors $\tau^2 \in \mathbb{Z} + \tau\mathbb{Z}$, donc τ est racine d'un polynôme $X^2 + aX + b$ avec $a, b \in \mathbb{Z}$.

Les racines du polynôme $X^2 + aX + b$ sont τ et $\overline{\tau}$, de sorte que $\tau + \overline{\tau} = -a$ et $\tau\overline{\tau} = b$. En particulier $\overline{\tau} = -a - \tau \in \mathbb{Z}[\tau]$.

Pour $x \in \mathbb{Z}[\tau]$, on a $\overline{x} \in \mathbb{Z}[\tau]$, donc $|x|^2 = \overline{x}x \in \mathbb{Z}[\tau] \cap \mathbb{R}_+ = \mathbb{N}$ (et, de même, $\overline{x} + x \in \mathbb{Z}[\tau] \cap \mathbb{R} = \mathbb{Z}$). Posons $v(x) = |x|^2$. Nous allons voir que *pour des valeurs très particulières* de τ, l'anneau $\mathbb{Z}[\tau]$ est euclidien de stathme v, et que dans d'autres cas, il n'est pas principal.

Inversibles. *Un élément x de $\mathbb{Z}[\tau]$ est inversible si et seulement si $v(x) = 1$.*

Démonstration. Si $xy = 1$, on a $v(x)v(y) = |x|^2|y|^2 = 1$ donc $v(x)$ est inversible dans $\mathbb{N}$: $v(x) = 1$. Si $v(x) = 1$ alors $x\overline{x} = 1$, donc x est inversible dans $\mathbb{Z}[\tau]$. □

Lemme. *On suppose que $|\mathrm{Im}(\tau)| < \sqrt{3}$. Alors pour tout $z \in \mathbb{C}$, il existe $q \in \mathbb{Z}[\tau]$ tel que $|z - q| < 1$.*

Démonstration. Soit $n \in \mathbb{Z}$ l'entier le plus proche de $\dfrac{\mathrm{Im}(z)}{\mathrm{Im}(\tau)}$, de sorte que $|\mathrm{Im}(z - n\tau)| \leqslant \dfrac{|\mathrm{Im}(\tau)|}{2}$. Soit aussi m le nombre entier le plus proche de la partie réelle de $z - n\tau$, de sorte que $|\mathrm{Re}(z - n\tau - m)| \leqslant \dfrac{1}{2}$. Posons $q = m + n\tau$. On a $|\mathrm{Re}(z - q)| \leqslant \dfrac{1}{2}$ et $|\mathrm{Im}(z - q)| \leqslant \dfrac{|\mathrm{Im}(\tau)|}{2}$, donc $|z - q|^2 \leqslant \dfrac{1 + \mathrm{Im}(\tau)^2}{4} < 1$. □

Anneau Euclidien. *Si $|\mathrm{Im}(\tau)| < \sqrt{3}$, l'anneau $\mathbb{Z}[\tau]$ est euclidien de stathme $v : x \mapsto |x|^2$.*

Démonstration. Soient $a, b \in \mathbb{Z}[\tau] - \{0\}$; posons $z = \dfrac{a}{b}$ et soit $q \in \mathbb{Z}[\tau]$ tel que $|z - q| < 1$. Posons $r = a - bq$. On a $a = bq + r$ et $|r| = |b||z - q| < |b|$ donc $v(r) < v(b)$. □

Sans changer l'anneau $\mathbb{Z}[\tau]$, on peut remplacer τ par $\overline{\tau}$, de sorte que l'on peut supposer que $\mathrm{Im}(\tau) > 0$; on peut aussi remplacer τ par $\tau + n$ (avec n dans $\mathbb{Z}$). On peut donc supposer que la partie réelle de τ est dans $[0, 1[$; comme $\tau + \overline{\tau} \in \mathbb{Z}$, on a $\tau + \overline{\tau} = 0$ ou 1. Cela nous ramène à étudier seulement le cas où τ est racine d'un polynôme $X^2 + b$, ou $X^2 - X + b$ (avec $b \in \mathbb{N}^*$). Dans le premier cas, $\tau = i\sqrt{b}$; dans le deuxième $\tau = \dfrac{1 + i\sqrt{4b - 1}}{2}$.
Le théorème s'applique donc uniquement dans les cinq cas suivants :

$$\tau \in \left\{ i, i\sqrt{2}, \frac{1 + i\sqrt{3}}{2}, \frac{1 + i\sqrt{7}}{2}, \frac{1 + i\sqrt{11}}{2} \right\}.$$

Commentaire. On peut démontrer relativement facilement (*cf.* exerc. 2.5) que dans tous les autres cas, l'anneau $\mathbb{Z}[\tau]$ n'est pas euclidien. Cependant, il y a quelques cas où $\mathbb{Z}[\tau]$ est quand même principal. Cela se produit pour $\tau = \dfrac{1 + i\sqrt{19}}{2}$. (*cf.* exerc. 2.7). Cependant, pour $b \geqslant 3$, l'anneau $\mathbb{Z}[i\sqrt{b}]$ n'est pas factoriel, donc il n'est pas principal (*cf.* exerc. 2.8).

L'équation diophantienne $x^2 + y^2 = z^2$. On cherche à trouver tous les triples $(x, y, z) \in \mathbb{Z}^3$ tels que $x^2 + y^2 = z^2$. Si (x, y, z) est une solution et $k \in \mathbb{Z}$, alors (kx, ky, kz) est aussi une solution. On peut donc supposer que (x, y, z) sont premiers entre eux. Si d divise x et y, alors d^2 divise z^2, donc d divise z. On peut donc supposer que x et y sont premiers entre eux. Remarquons que x et y ne peuvent être tous deux impairs car alors $x^2 \equiv y^2 \equiv 1$ [4], donc $z^2 \equiv 2$ [4] ce qui est impossible. Donc l'un des deux est pair et l'autre impair.

Dans ce cas, l'idéal $(x + iy)\mathbb{Z}[i] + (x - iy)\mathbb{Z}[i]$ de $\mathbb{Z}[i]$ contient $(x + iy) + (x - iy) = 2x$, ainsi que $i\big((x - iy) - (x + iy)\big) = 2y$ donc il contient 2. Or il existe $q \in \mathbb{Z}[i]$ tel que $(x + iy) - 2q$ soit égal à 1 ou à i. Cela prouve que $(x + iy)$ et $(x - iy)$ sont premiers entre eux. Décomposons $z^2 = (x + iy)(x - iy)$ en éléments irréductibles dans $\mathbb{Z}[i]$; puisque c'est un carré, chacun figure un nombre pair de fois. Cela prouve que $x + iy$ est associé à un carré : il existe $(a, b) \in \mathbb{Z}$, tels que $x + iy$ soit associé à $(a + ib)^2$ c'est à dire $x + iy = \pm(a + ib)^2$ (si y est pair) ou $x + iy = \pm i(a + ib)^2$ (si x est pair). On en déduit que les solutions sont nécessairement de la forme $(k(a^2 - b^2), 2kab, k(a^2 + b^2))$ ou $(2kab, k(a^2 - b^2), k(a^2 + b^2))$ (avec $a, b, k \in \mathbb{Z}$).

Irréductibles. *On suppose que $\mathbb{Z}[\tau]$ est principal. Soit q un élément irréductible de $\mathbb{Z}[\tau]$. Alors deux cas sont possibles :*

- *il existe un nombre premier $p \in \mathbb{N}$ tel que $v(q) = p$;*
- *il existe un nombre premier $p \in \mathbb{N}$ tel que q soit associé à p (et l'on a $v(q) = p^2$).*

Démonstration. Décomposons $v(q) = q\overline{q}$ en facteurs premiers dans $\mathbb{Z}$. C'est une décomposition dans $\mathbb{Z}[\tau]$ qui ne peut donc avoir que un ou deux éléments : dans le premier cas $v(q)$ est premier ; dans le deuxième un des facteurs p est associé à q, donc $v(q) = v(p) = p^2$. □

Nous verrons en exercice (2.1) quels nombres premiers de $\mathbb{N}$ ne sont plus irréductibles dans $\mathbb{Z}[i]$.

2.5 Sous-corps

2.5.1 Caractéristique d'un corps ; sous-corps premier

Soit K un corps. Tout morphisme f d'anneaux de K dans un anneau non nul est injectif : si $x \in K^*$, alors $f(x^{-1})f(x) = 1$, donc $f(x) \neq 0$.

Soit K un corps et $f : \mathbb{Z} \to K$ l'unique homomorphisme d'anneaux (défini par $f(n) = n1_K$). Le noyau de f est un idéal $n\mathbb{Z}$ de $\mathbb{Z}$. L'image $f(\mathbb{Z})$ est un sous-anneau commutatif de $\mathbb{K}$ isomorphe à $\mathbb{Z}/n\mathbb{Z}$. Puisque K est un corps, $f(\mathbb{Z})$ est un anneau intègre, donc ou bien n est premier, ou bien f est injective. Ce nombre n s'appelle la *caractéristique* de K.

- Lorsque la caractéristique p n'est pas nulle, l'image $f(\mathbb{Z})$ est un sous-corps de K isomorphe à $\mathbb{Z}/p\mathbb{Z}$.
- Lorsque f est injective, on peut étendre f en un homomorphisme $\tilde{f} : \mathbb{Q} \to K$ en posant $\tilde{f}\left(\dfrac{p}{q}\right) = f(p)f(q)^{-1}$ pour $p, q \in \mathbb{Z}$ avec $q \neq 0$. L'image $\tilde{f}(\mathbb{Q})$ est un sous-corps de K isomorphe à $\mathbb{Q}$.
- Le corps ainsi obtenu, isomorphe selon les cas à $\mathbb{Z}/p\mathbb{Z}$ ou à $\mathbb{Q}$ est le plus petit sous-corps de K. On l'appelle le *sous-corps premier* de K.

2.5.2 Corps des fractions d'un anneau intègre

Soit A un anneau commutatif intègre non nul. On définit un corps K contenant A. Sa construction est la généralisation de la construction de $\mathbb{Q}$ à partir de $\mathbb{Z}$. Les éléments de K sont des fractions $\dfrac{a}{b}$ où $a \in A$ et $b \in A - \{0\}$. On peut alors dire quand deux fractions sont égales, définir l'addition et la multiplication des fractions, et vérifier que l'on obtient ainsi un corps qui contient l'anneau A.

Pour formaliser cela, considérons la relation R sur $A \times (A - \{0\})$ définie par $(a, b)\, R\, (c, d)$ si $ad = bc$. On vérifie sans peine que R est une relation d'équivalence. Notons K l'ensemble quotient. La classe dans K d'un élément $(a, b) \in A \times (A - \{0\})$ se note $\dfrac{a}{b}$.

On définit la somme et le produit d'éléments de K en posant pour des éléments a, b, c, d de A avec b, d non nuls

$$\frac{a}{b} + \frac{c}{d} = \frac{ad + bc}{bd} \quad \text{et} \quad \frac{a}{b}\frac{c}{d} = \frac{ac}{bd}.$$

Ces opérations sont bien définies : si $(a, b)\, R\, (a', b')$ et $(c, d)\, R\, (c', d')$, alors $(ad + bc, bd)\, R\, (a'd' + b'c', b'd')$ et $(ac, bd)\, R\, (a'c', b'd')$. De plus, on a

$$\frac{a}{d} + \frac{c}{d} = \frac{a + c}{d}$$

ce qui permet de démontrer facilement les règles des opérations : K est bien un anneau commutatif. De plus K est un corps : l'inverse de $\dfrac{a}{b}$ est $\dfrac{b}{a}$ (pour $a, b \in A - \{0\}$).

Le corps K s'appelle le *corps de fractions* de A.

Enfin, on plonge A dans K au moyen de l'application $a \mapsto \dfrac{a}{1}$: cette application est un morphisme injectif qui plonge l'anneau A dans K.

2.29 Proposition. *Soit A un anneau commutatif intègre. Notons $K(A)$ son corps des fractions. Pour tout corps L et tout homomorphisme injectif $f : A \to L$, il existe un unique homomorphisme $\tilde{f} : K(A) \to L$ dont la restriction à $A \subset K(A)$ soit f*

2.5.3 Éléments algébriques, éléments transcendants

Soient L un corps commutatif et $K \subset L$ un sous-corps.

Soit $x \in L$. Considérons L comme espace vectoriel sur K et introduisons le sous-espace $K[x] \subset L$ engendré par les éléments x^n pour $n \in \mathbb{N}$. Cet espace est l'image de l'application $f : P \mapsto P(x)$ de $K[X]$ dans L. Cette application étant un homomorphisme d'anneaux, son noyau est un idéal de l'anneau principal $K[X]$. Il existe donc un polynôme $\varpi \in K[X]$ tel que $\ker f = \varpi K[X]$. Deux cas sont possibles :

a) Si $\varpi = 0$, l'application $f : P \mapsto P(x)$ est injective de $K[X]$ dans L. On dit alors que x est *transcendant* sur K.

b) Si $\varpi \neq 0$. Remarquons que f n'est pas l'application nulle, donc ϖ n'est pas inversible ; si $\varpi = PQ$, on trouve $P(x)Q(x) = 0$, ce qui implique que $P(x) = 0$ ou $Q(x) = 0$, *i.e.* l'un des deux est dans $\ker f$ donc multiple de ϖ. Il s'ensuit que ϖ est irréductible. On dit alors que x est *algébrique* sur K et le polynôme ϖ s'appelle le *polynôme minimal* de x.

Citons sans démonstration le résultat suivant (*cf.* exercice 2.9) :

2.30 Proposition. *Soient L un corps commutatif et $K \subset L$ un sous corps. Les éléments de L algébriques sur K forment un sous-corps de L.*

Cela signifie que la somme, le produit, l'inverse d'éléments algébriques est algébrique.

2.31 Proposition. *Le corps des nombres complexes algébriques sur $\mathbb{Q}$ est dénombrable.*

En effet, les éléments algébriques sont les racines de polynômes à coefficients rationnels (non nuls). Or $\mathbb{Q}$ étant dénombrable, l'ensemble des polynômes à coefficients rationnels est dénombrable, et chacun a un nombre fini de racines

On déduit de ce résultat qu'il y a « bien plus » de nombres transcendants que de nombres algébriques. On peut démontrer que les nombres e et π sont transcendants, mais ce n'est pas si facile.

2.6 Exercices

2.1 Exercice. Soit $p \in \mathbb{N}$ un nombre premier. Démontrer que les assertions suivantes sont équivalentes :

(i) il existe $a, b \in \mathbb{Z}$ tels que $a^2 + b^2 = p$;

(ii) l'élément $p \in \mathbb{Z}[i]$ n'est pas irréductible dans $\mathbb{Z}[i]$;

(iii) -1 est un carré modulo p ;

(iv) $p \not\equiv 3$ [4].

2.2 Exercice. *Le groupe $\mathbb{F}_p^*$ est cyclique. (1)*

1. Soit G un groupe commutatif fini.

   a) Soient $a, b \in G$. On note k_a et k_b leurs ordres respectifs. On suppose que k_a et k_b sont premiers entre eux. Démontrer que l'ordre de ab est $k_a k_b$.

   b) Démontrer qu'il existe $n \in \mathbb{N}^*$ tel que $\{k \in \mathbb{Z};\ \forall x \in G;\ x^k = 1\} = n\mathbb{Z}$. Démontrer que n divise le cardinal de G.
   Le nombre n s'appelle l'*exposant* de G.

c) Ecrivons $n = \prod p_j^{m_j}$ la décomposition de n en nombres premiers distincts. Démontrer que pour tout j, il existe $x_j \in G$ d'ordre $p_j^{m_j}$.

d) En déduire qu'il existe $x \in G$ d'ordre n.

2. Soit K un corps commutatif et G un sous-groupe fini à N éléments de K^*. Soit n son exposant.

a) Démontrer que l'équation $x^n = 1$ a au plus n solutions dans K. En déduire que $N \leqslant n$.

b) Démontrer que G est cyclique.

2.3 Exercice. *Le groupe $\mathbb{F}_p^*$ est cyclique. (2)*

1. Soit $n \in \mathbb{N}^*$. On considère l'ensemble $A_n = \left\{ \frac{k}{n};\ k \in \mathbb{N},\ \ 0 \leqslant k < n \right\}$.

a) Soit d un diviseur de n. Combien d'éléments de A_n ont leur écriture irréductible de la forme $\frac{a}{d}$?

b) En déduire l'égalité $\sum_{d|n} \varphi(d) = n$.

2. Soit K un corps commutatif et G un sous-groupe fini à n éléments de K^*. Pour $d \in \mathbb{N}^*$, on note s_d le nombre d'éléments d'ordre d de G.

a) Démontrer que $\sum_{d|n} s_d = n$.

b) Soit $x \in G$; notons d son ordre et H le sous-groupe (cyclique) de G engendré par x. Démontrer que
- H a d éléments et $\varphi(d)$ éléments d'ordre d.
- Tout élément $y \in H$ vérifie $y^d = 1$.
- L'équation $y^d = 1$ a au plus d solutions dans K.
- Tout élément d'ordre d de G est dans H.

c) En déduire que si $s_d \neq 0$, alors $s_d = \varphi(d)$.

d) En déduire que pour tout diviseur d de n on a $s_d = \varphi(d)$, puis que G est cyclique.

2.4 Exercice. *Le groupe $(\mathbb{Z}/n\mathbb{Z})^*$ est-il cyclique ? (***)* Perrin, *Cours d'algèbre* p.24.

Cet exercice complète les précédents.

1. a) Soient G et H deux groupes commutatifs finis. Démontrer que $G \times H$ est cyclique si et seulement si G et H sont cycliques et que leurs ordres sont premiers entre eux.

b) Quels sont les nombres n tels que $\varphi(n)$ soit impair ?

c) Soient m et n deux nombres entiers premiers entre eux distincts de 1 et de 2. Démontrer que $(\mathbb{Z}/nm\mathbb{Z})^*$ n'est pas cyclique.

d) $\mathbb{Z}/8\mathbb{Z}^*$ est-il cyclique ?

2. Soient p un nombre premier distinct de 2 et $n \in \mathbb{N}$, $n \geqslant 2$.

a) Démontrer (par récurrence) que, pour tout $k \in \mathbb{N}$, $(1+p)^{p^k} \equiv 1 + p^{k+1} \ \ [p^{k+2}]$.

b) Quel est l'ordre de $1 + p$ dans le groupe $\mathbb{Z}/p^n\mathbb{Z}^*$?

c) Soit $a \in \mathbb{Z}$ dont la classe dans $\mathbb{Z}/p\mathbb{Z}$ engendre $\mathbb{Z}/p\mathbb{Z}^*$, et soit $x \in \mathbb{Z}/p^n\mathbb{Z}^*$ la classe de a. Démontrer que l'ordre de x dans $\mathbb{Z}/p^n\mathbb{Z}^*$ est un multiple de $p - 1$. En déduire qu'il existe dans $\mathbb{Z}/p\mathbb{Z}$ un élément d'ordre $p - 1$.

d) Démontrer que $\mathbb{Z}/p^n\mathbb{Z}^*$ est cyclique. Démontrer que $\mathbb{Z}/2p^n\mathbb{Z}^*$ est aussi cyclique.

3. Quels sont les entiers n tels que $\mathbb{Z}/n\mathbb{Z}^*$ soit cyclique ?

2.5 Exercice. Soit $a \in \mathbb{N}^*$ et $\tau \in \mathbb{C}$ une racine du polynôme $X^2 + X + a$. On note $\mathbb{Z}[\tau]$ l'anneau $\mathbb{Z} + \tau\mathbb{Z}$.

1. a) Soit $x \in \mathbb{Z}[\tau]$ non nul. Démontrer que $\mathbb{Z}[\tau]/x\mathbb{Z}[\tau]$ est fini. Notons $v(x)$ le nombre de ses éléments.

   b) Soient $x, y \in \mathbb{Z}[\tau]$ non nuls. Donnons nous des représentants $r_1, \ldots, r_n$ des classes d'éléments de $\mathbb{Z}[\tau]$ modulo x et des représentants $s_1, \ldots, s_m$ des classes d'éléments de $\mathbb{Z}[\tau]$ modulo y. Démontrer que tout élément de $\mathbb{Z}[\tau]$ est congru modulo xy à un un et un seul élément de la forme $r_i + xs_j$. En déduire que $v(xy) = v(x)v(y)$.

   c) En calculant $v(k)$ pour $k \in \mathbb{Z}$, démontrer que, pour tout $x \in \mathbb{Z}[\tau]$ non nul, on a $v(x) = |x|^2$.

2. On suppose que $\mathbb{Z}[\tau]$ possède un stathme euclidien V.

   a) On suppose aussi que les seuls éléments inversibles de $\mathbb{Z}[\tau]$ sont ± 1. Soit $x \in \mathbb{Z}[\tau]$ non nul et non inversible de stathme minimal. Démontrer que tout élément de $\mathbb{Z}[\tau]$ est congru modulo x à $0, 1$ ou -1 ; en déduire que $v(x) \leqslant 3$.

   b) Démontrer que $\operatorname{Im} \tau \leqslant \sqrt{3}$.

2.6 Exercice. *Sous-groupes de* $\mathbb{Z}[\tau]$. Soit $\tau \in \mathbb{C}$ racine d'un polynôme $X^2 + X + a$ ou $X^2 + a$ avec $a \in \mathbb{N}^*$. Soit G un sous-groupe non nul de $\mathbb{Z}[\tau]$. Soit $\alpha \in G$ un élément non nul tel que $|\alpha|^2$ soit minimal dans $\{|x|^2;\ x \in G \setminus \{0\}\}$. (Un tel élément existe d'après le « principe de récurrence »- puisque $|x|^2 \in \mathbb{N}$ pour tout $x \in \mathbb{Z}[\tau]$).

1. Démontrer que $G \cap \mathbb{R}\alpha = \mathbb{Z}\alpha$.
   On suppose désormais que $G \not\subset \mathbb{R}\alpha$. Soit $\beta \in G \setminus \mathbb{Z}\alpha$ tel que $|\beta|^2$ soit minimal dans $\{|x|^2;\ x \in G \setminus \mathbb{Z}\alpha\}$. Quitte à remplacer β par $-\beta$, on peut supposer que $\operatorname{Im} \dfrac{\beta}{\alpha} > 0$.
2. Démontrer que $\left|\dfrac{\beta}{\alpha}\right| \geqslant 1$ et $\left|\operatorname{Re} \dfrac{\beta}{\alpha}\right| \leqslant \dfrac{1}{2}$.
3. Soit $x \in G$. Démontrer qu'il existe $m, n \in \mathbb{Z}$ tels que

$$\left|\operatorname{Im} \frac{x - n\beta}{\alpha}\right| \leqslant \frac{1}{2} \operatorname{Im} \frac{\beta}{\alpha} \quad \text{et} \quad \left|\operatorname{Re} \frac{x - (m\alpha + n\beta)}{\alpha}\right| \leqslant \frac{1}{2}.$$

   En déduire que $|x - (m\alpha + n\beta)| < |\beta|$, puis que $x = m\alpha + n\beta$.
   Il s'ensuit que $G = \alpha\mathbb{Z} + \beta\mathbb{Z}$.

2.7 Exercice. *Un anneau principal non euclidien*

Le but de cet exercice est de démontrer que pour $\tau = \dfrac{1 + i\sqrt{19}}{2}$, l'anneau $\mathbb{Z}[\tau]$ est principal mais n'est pas euclidien. Soit J un idéal non nul de $\mathbb{Z}[\tau]$. Puisque J est un sous-groupe de $\mathbb{Z}[\tau]$, non contenu dans un $a\mathbb{R}$ (il contient un élément non nul a et $a\tau$) il existe d'après l'exercice 2.6 $\alpha, \beta \in \mathbb{Z}[\tau]$ tels que

$$\operatorname{Im} \frac{\beta}{\alpha} > 0, \quad \left|\frac{\beta}{\alpha}\right| \geqslant 1, \quad \left|\operatorname{Re} \frac{\beta}{\alpha}\right| \leqslant \frac{1}{2} \quad \text{et} \quad G = \alpha\mathbb{Z} + \beta\mathbb{Z}$$

(remarquons que $\tau\alpha \in J$, donc $J \not\subset \mathbb{R}\alpha$).

1. Démontrer qu'il existe $a, b, c, d \in \mathbb{Z}$ tels que $\tau\alpha = a\alpha + b\beta$ et $\tau\beta = c\alpha + d\beta$.
2. En regardant les signes des parties imaginaires de τ et $\dfrac{\beta}{\alpha}$, démontrer que $b > 0$.
3. Quelles sont les valeurs propres et espaces propres de la matrice $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$? Démontrer que $a + d = 1$ et $ad - bc = 5$. En déduire que $ad \leqslant 0$, que ad est pair, que $bc < 0$, que b et c sont impairs et que $4bc + (a - d)^2 = -19$.
4. Posons $x = \dfrac{\beta}{\alpha}$. Démontrer que $\begin{vmatrix} a + bx & 1 \\ c + dx & x \end{vmatrix} = 0$. En déduire que

   a) x et $\overline{x}$ sont racines du polynôme $bX^2 + (a - d)X - c$,

b) $|x|^2 = -\dfrac{c}{b}$ et $\operatorname{Re} x = \dfrac{d-a}{2b}$.

5. Démontrer que $|a-d| \leqslant b$ et $b \leqslant -c$. En déduire que $3b^2 \leqslant 19$, puis que $b = 1$.
6. En déduire que $(\alpha, \tau\alpha)$ est une $\mathbb{Z}$-base de J et conclure.

On démontre de même que pour $D \in \{19, 43, 67, 163\}$, l'anneau $\mathbb{Z}\left[\dfrac{1+i\sqrt{D}}{2}\right]$ est principal.

2.8 Exercice. 1. Dans $\mathbb{Z}[X]$, démontrer que l'idéal engendré par 2 et X n'est pas principal.

2. Démontrons que pour $\tau = i\sqrt{b}$ avec $b \geqslant 3$ et pour $\tau = \dfrac{1+i\sqrt{15}}{2}$, l'anneau $\mathbb{Z}[\tau]$ n'est pas factoriel (donc n'est pas principal). En utilisant les égalités :
   - pour $\tau = i\sqrt{3}$, on a $(1+\tau)(1+\overline{\tau}) = 4 = 2 \times 2$;
   - pour $\tau = 2i$ ou $\tau = \dfrac{1+i\sqrt{15}}{2}$, on a $\tau\overline{\tau} = 4 = 2 \times 2$;
   - pour $\tau = i\sqrt{5}$, on a $(1+\tau)(1+\overline{\tau}) = 6 = 2 \times 3$;

   démontrer que l'on n'a pas l'unicité dans la décomposition en éléments irréductibles. Pour $b \geqslant 5$, et $\tau = i\sqrt{b}$, écrire une égalité de ce style en discutant la parité de p. En déduire que $\mathbb{Z}[\tau]$ n'est pas principal.

2.9 Exercice. Soient L un corps commutatif et $K \subset L$ un sous-corps. Remarquons que L est un espace vectoriel sur K et que tout sous-anneau de L contenant K est un sous-K-espace vectoriel de L.

1. Soit K_1 un sous-corps de L contenant K. Démontrer que tout élément algébrique sur K est algébrique sur K_1.
2. Démontrer que pour $x \in L$ les conditions suivantes sont équivalentes :
   (i) x est algébrique sur K ;
   (ii) il existe un sous-anneau K_1 de L contenant K et x et qui soit un espace vectoriel de dimension finie sur K ;
   (iii) il existe un sous-corps K_1 de L contenant K et x et qui soit un espace vectoriel de dimension finie sur K.
3. Soient K_1, K_2 des sous-corps de L tels que $K \subset K_1 \subset K_2$. Démontrer que K_2 est un K-espace vectoriel de dimension finie si et seulement si K_2 est un K_1-espace vectoriel de dimension finie et K_1 est un K-espace vectoriel de dimension finie, et que dans ce cas, on a $\dim_K(K_2) = \dim_{K_1}(K_2)\dim_K(K_1)$.
4. Soient $\alpha, \beta \in L$ des éléments algébriques sur K. Soit K_1 un sous-corps de L contenant K et α et de dimension finie sur K.
   a) On suppose que $\alpha \neq 0$. Démontrer que α^{-1} est algébrique sur K.
   b) Démontrer que $\alpha + \beta$ et $\alpha\beta$ sont algébriques sur K.
5. Démontrer que les éléments de L algébriques sur K forment un sous-corps K' de L. Démontrer que si $x \in L$ est algébrique sur K' alors $x \in K'$.

3 Polynômes et fractions rationnelles

3.1 Polynômes à une indéterminée sur un corps commutatif K

Soit K un corps commutatif. On sait très bien ce qu'est un polynôme à coefficients dans K : c'est une expression abstraite $P = \sum_{k=0}^{n} a_k X^k$ où les a_i sont des éléments de K appelés les coefficients de P. On sait ajouter et multiplier les polynômes, les multiplier par un scalaire : les polynômes forment une K-algèbre.

3.1 Quelques mots sur la définition de l'algèbre $K[X]$. Se donner un polynôme revient à se donner ses coefficients c'est à dire une suite $(a_k)_{k\in\mathbb{N}}$ d'éléments de K qui sont nuls pour k assez grand : il existe $n \in \mathbb{N}$ satisfaisant $a_k = 0$ pour $k > n$. On peut formaliser cela en définissant un polynôme comme la suite abstraite de ses coefficients : l'ensemble des polynômes est alors l'ensemble $K^{(\mathbb{N})}$ des suites $(a_k)_{k\in\mathbb{N}}$ telles qu'il existe $n \in \mathbb{N}$ satisfaisant $a_k = 0$ pour $k > n$. Dans cette vision, le $k^{\text{ème}}$ coefficient du polynôme X^k est égal à 1, et tous les autres sont nuls. L'ensemble $K^{(\mathbb{N})}$ est naturellement un K-espace vectoriel de dimension infinie et $(X^k)_{k\in\mathbb{N}}$ en est une base.

L'algèbre $K[X]$ est donc l'espace vectoriel $K^{(\mathbb{N})}$ muni de l'unique produit tel que $X^k X^\ell = X^{k+\ell}$ (pour tous $k, \ell \in \mathbb{N}$). Enfin, on identifie K avec l'ensemble des polynômes constants (au moyen de $a \mapsto aX^{(0)}$).

Soit $P \in K[X]$ un polynôme non nul. On appelle *degré* de P l'entier $\partial P = n \in \mathbb{N}$ tel que $a_n \neq 0$ et, $a_k = 0$ pour $k > n$ (où les a_k sont les coefficients de P). Le coefficient non nul de plus haut degré (a_n si $\partial P = n$) s'appelle le *coefficient directeur* de P. On dit que P est *unitaire* (ou *monique*) si son coefficient directeur est 1.

3.2 Proposition. *Pour $P, Q \in K[X]$ deux polynômes non nuls, on a $PQ \neq 0$, $\partial(PQ) = \partial P + \partial Q$, et le coefficient directeur de PQ est le produit des coefficients directeurs de P et de Q. En particulier, l'anneau $K[X]$ est intègre.*

L'anneau $K[X]$ est euclidien de stathme ∂. Plus précisément on a (où l'on a convenu $\partial 0 < 0$) :

3.3 Proposition : Division euclidienne dans $K[X]$. *Soient $A, B \in K[X]$ avec $B \neq 0$. Il existe un unique couple $Q, R \in K[X]$ tels que $A = BQ + R$ et $\partial R < \partial B$.*

On en déduit que $K[X]$ est principal, c'est-à-dire que tous les idéaux de $K[X]$ sont de la forme $AK[X]$. On peut alors définir le plus grand commun diviseur (PGCD) et plus petit commun multiple (PPCM) de deux polynômes, établir un théorème de Bézout, un algorithme d'Euclide qui permet de trouver le PGCD et une relation de Bézout ainsi que la décomposition unique d'un polynôme en facteurs irréductibles.

3.4 Exercices. a) Soient L un corps commutatif et K un sous-corps de L. Soient $A, B \in K[X]$; Démontrer que leur PGCD est le même qu'on les considère comme éléments de $K[X]$ ou de $L[X]$.

b) Calculer $PGCD(X^m - 1, X^n - 1)$.

De l'égalité $\partial(PQ) = \partial P + \partial Q$ on déduit :

3.5 Proposition. *a) Les éléments inversibles de $K[X]$ sont les polynômes non nuls de degré nul, i.e. les éléments de K.*

b) Tout polynôme de degré 1 est irréductible.

3.2 Fonctions polynômes

3.2.1 Racines

Soit K un corps. Si $x \in K$ et $P = \sum_{k=0}^{n} a_k X^k \in K[X]$, on pose $P(x) = \sum_{k=0}^{n} a_k x^k$. L'application $x \mapsto P(x)$ s'appelle la fonction polynôme associée à P. L'application $P \mapsto P(x)$ est un homomorphisme d'anneaux de $K[X]$ dans K. On dit que x est une *racine* de P si $P(x) = 0$.

3.6 Proposition. *Le reste de la division euclidienne de P par $X - a$ est $P(a)$. En particulier, $X - a$ divise P si et seulement si $P(a) = 0$.*

En effet, écrivons $P = (X - a)Q + R$ avec $\partial R < 0$, donc $R \in K$. Comme $(X - a)(a) = 0$, on trouve $P(a) = R$.

Cette proposition nous conduit à dire que a est une racine d'ordre k (au moins) si $(X - a)^k$ divise P et d'ordre exactement k si de plus $(X - a)^{k+1}$ ne divise pas P. Si $k = 2, 3$, on dira que a est racine double, triple... de P. Si $k \geqslant 2$ on dira que a est *racine multiple* de P

3.7 Proposition. *Soient $a_1, \ldots, a_k \in K$ des éléments deux à deux distincts et $m_1, \ldots, m_k \in \mathbb{N}$. Si un polynôme non nul P admet les racines a_j avec multiplicité m_j, il est divisible par $\prod_{j=1}^{k}(X - a_j)^{m_j}$. En particulier $\partial P \geqslant \sum m_j$ et si $\partial P = \sum m_j$, alors $P = a\prod_{j=1}^{k}(X - a_j)^{m_j}$ où $a \in K$ est le coefficient directeur de P.*

Les polynômes $X - a_j$ sont premiers entre eux deux à deux, donc il en va de même pour $(X - a_j)^{m_j}$. Si P admet les racines a_j avec multiplicité m_j, il est divisible par $(X - a_j)^{m_j}$, donc par leur produit.

3.8 Exemple. Soit p un nombre premier. D'après le (petit) théorème de Fermat, pour tout $x \in \mathbb{Z}/p\mathbb{Z}$, on a $x^p = x$. En d'autres termes, tout élément de $\mathbb{Z}/p\mathbb{Z}$ est racine du polynôme $X^p - X \in \mathbb{Z}/p\mathbb{Z}[X]$. On en déduit que $\prod_{x \in \mathbb{Z}/p\mathbb{Z}} (X - x) = X^p - X$.

3.9 Corollaire. *Si K est infini, l'homomorphisme qui à un polynôme P associe la fonction polynôme $x \mapsto P(x)$ de K dans K est injectif.*

En effet, un polynôme non nul ne peut avoir qu'un nombre fini de racines. Il ne peut s'annuler sur tout K.

A cause de ce corollaire, on confond souvent les polynômes avec les fonctions polynômes.

3.10 Remarque. Pour $K = \mathbb{Z}/p\mathbb{Z}$, le noyau de l'homomorphisme qui à un polynôme P associe la fonction polynôme $x \mapsto P(x)$ de K dans K est l'idéal engendré par $X^p - X$.

3.11 Exemple. Polynôme d'interpolation de Lagrange. Soient $x_1, x_2 \ldots, x_n$ des éléments distincts de K et $\lambda_1, \lambda_2, \ldots, \lambda_n$ des éléments de K. Il existe un unique polynôme P de degré au plus $n - 1$ tel que $P(x_i) = \lambda_i$ pour tout i.

Existence. Pour $i = 1, \ldots, n$, posons $Q_i = \prod_{1 \leqslant j \leqslant n;\ j \neq i} (X - x_j)$. On prend $P = \sum_{i=1}^{n} \frac{\lambda_i}{Q_i(x_i)} Q_i$.

Unicité. Si P et Q satisfont ces conditions alors $P - Q$ s'annule en les x_i ; comme $\partial(P - Q) < n$, il vient $P - Q = 0$.

3.2.2 Polynômes scindés ; relations entre coefficients et racines

On dit qu'un polynôme P est *scindé* s'il est produit de polynômes du premier degré. Alors P s'écrit $P = a\prod_{k=1}^{n}(X - x_k)$.

Soit $P = \prod_{k=1}^{n}(X - x_k)$ un polynôme unitaire scindé. Écrivons $P = X^n + \sum_{k=0}^{n-1} a_k X^k$. Alors on a

- somme des racines $\sum_{k=1}^{n} x_k = -a_{n-1}$;
- produit des racines $(-1)^n a_0 = \prod_{k=1}^{n} x_k$.
- Plus généralement, $(-1)^\ell a_{n-\ell} = \sum_{1\leqslant k_1<\ldots<k_\ell\leqslant n}\left(\prod_{j=1}^{\ell} x_{k_j}\right)$ est la somme de tous les produits de ℓ racines.
- Pour $n = 2$ on trouve $P = X^2 - SX + P$ où S est la somme et P le produit des racines.
- Pour $n = 3$ on trouve $P = X^3 - SX^2 + \Sigma_2 X - P$, où S est la somme, P le produit des racines et $\Sigma_2 = x_1x_2 + x_1x_3 + x_2x_3$.

3.2.3 Dérivation des polynômes

Soit $P = \sum_{k=0}^{n} a_k X^k$. On définit sa dérivée : c'est le polynôme $P' = \sum_{k=1}^{n} k a_k X^{k-1}$.

3.12 Proposition. a) *Pour $P, Q \in K[X]$, on a $(PQ)' = P'Q + PQ'$.*

b) *Soient $P \in K[X]$ et $a \in K$ une racine de P. Alors a est une racine double de P si et seulement si $P'(a) = 0$.*

a) se vérifie pour $P = X^k$ et $Q = X^\ell$ et s'étend par linéarité.
Pour b), écrivons $P = (X-a)Q$ de sorte que (d'après a) $P' = Q+(X-a)Q'$, donc $Q(a) = P'(a)$. Alors a est racine double de P, si et seulement si c'est une racine de Q, *i.e.* si et seulement si $P'(a) = Q(a) = 0$.

3.13 Dérivées successives ; identité de Taylor. On définit aussi les dérivées successives en posant $P'' = (P')'$ *etc.* La dérivée k-ième se note $P^{(k)}$. On a $P^{(k)}(0) = k!a_k$, de sorte que, si K est de caractéristique nulle (et $\partial P \leqslant n$),

$$P = \sum_{k=0}^{n} \frac{P^{(k)}(0)}{k!} X^k.$$

Plus généralement, soit $a \in K$. Les polynômes $(X - a)^k$ forment une base de $K[X]$ (car ils sont échelonnés). Posons $Q_k = \dfrac{(X-a)^k}{k!}$. On a $Q_k^{(j)} = Q_{k-j}$ si $k \geqslant j$ et $Q_k^{(j)} = 0$ si $k < j$. En particulier, $Q_k^{(j)}(a) = \delta_k^j$, et si P s'écrit $\sum_k b_k Q_k$, il vient $b_j = P^{(j)}(a)$, donc

$$P = \sum_{k=0}^{n} \frac{P^{(k)}(a)}{k!}(X - a)^k.$$

3.2.4 Polynômes irréductibles sur $\mathbb{R}$ et $\mathbb{C}$

Donnons sans démonstration le théorème fondamental suivant :

3.14 Théorème de d'Alembert-Gauss. *Tout polynôme non constant à coefficients complexes admet au moins une racine dans* $\mathbb{C}$.

Tout polynôme non constant est donc divisible par un $X - a$. Il en résulte immédiatement que les polynômes irréductibles dans $\mathbb{C}[X]$ sont les polynômes du premier degré : tout polynôme à coefficients complexes est donc scindé.

Soit maintenant $P \in \mathbb{R}[X]$ un polynôme irréductible. Considérons le comme polynôme à coefficients complexes. Il a une racine $z \in \mathbb{C}$. Si $z \in \mathbb{R}$, P est du premier degré. Si $z = a + ib \notin \mathbb{R}$, alors écrivons $P = BQ + R$ la division euclidienne de P par $B = (X - z)(X - \overline{z}) = X^2 - 2aX + (a^2 + b^2)$ (dans $\mathbb{R}[X]$). Alors $R \in \mathbb{R}[X]$ est de degré au plus 1 et s'annule en z : c'est le polynôme nul.

On trouve :

3.15 Corollaire. *Les polynômes irréductibles de* $\mathbb{R}[X]$ *sont les polynômes du premier degré et ceux du deuxième degré de discriminant strictement négatif.*

3.2.5 Racines et extensions de corps

Soient K un corps commutatif et $P \in K[X]$. Si L est une extension de K, on peut considérer P comme polynôme à coefficients dans L : en d'autres termes on identifie $K[X]$ à un sous-anneau de $L[X]$. En particulier, on peut définir

On note (P) l'idéal $PK[X]$ de $K[X]$. Puisque $K[X]$ est principal, tout idéal de $K[X]$ est de cette forme. Nous utiliserons le résultat suivant.

3.16 Proposition. *Soit* $P \in K[X]$ *un polynôme non nul. On a l'équivalence entre :*

(i) *Le polynôme* P *est irréductible ;*

(ii) *L'anneau quotient* $K[X]/(P)$ *est intègre ;*

(iii) *L'anneau quotient* $K[X]/(P)$ *est un corps.* □

Soit $P \in K[X]$ un polynôme irréductible. Notons $L = K[X]/(P)$ l'anneau quotient.

- On considère l'application $i : K \to L$ qui à un scalaire $a \in K$ associe la classe du polynôme constant $a \in K[X]$ dans le quotient. L'application i est un morphisme de corps (injectif) au moyen duquel on identifie K à un sous-corps de L et donc L à une extension de K.
- Notons aussi x la classe dans L du polynôme X dans le quotient $L = K[X]/(P)$. En d'autres termes, on a $x = \pi(X)$ où $\pi : K[X] \to L = K[X]/(P)$ est l'application quotient.
- Comme π est un homomorphisme d'anneaux, on $\pi(X^2) = x^2$ et plus généralement $\pi(X^n) = x^n$.
- Pour $a \in K$, on a $\pi(aX^0) = i(a)$, en d'autres termes, avec les identifications de $K \subset K[X]$ et $K \subset L$, la restriction de π à K est l'identité.
- On a donc $\pi(\sum_{k=0}^{n} a_k X^k) = \sum_{k=0}^{n} a_k x^k$; autrement dit, pour tout polynôme $Q \in K[X]$, on a $\pi(Q) = Q(x)$.
- En particulier, puisque $P \in \ker \pi = (P)$, on a $P(x) = 0$.

On a démontré :

3.17 Théorème. *Soient* K *un corps et* $P \in K[X]$ *un polynôme irréductible sur* K. *Il existe une extension* L *de* K *dans laquelle* P *a une racine.* □

3.18 Corollaire. *Soient* K *un corps et* $P \in K[X]$ *un polynôme non constant.*

a) Il existe une extension L *de* K *dans laquelle* P *a une racine.*

b) Il existe une extension L *de* K *dans laquelle* P *est scindé.*

Démonstration. a) Soit $P_0 \in K[X]$ un polynôme irréductible dans K divisant P. Par le théorème ci-dessus, il existe une extension L de K dans laquelle P_0 admet une racine ; celle-ci sera une racine de P.

b) On procède par récurrence sur le degré de P. On démontre par récurrence sur n l'énoncé suivant : $S(n)$: pour tout corps commutatif K et tout polynôme $P \in K[X]$ de degré n, il existe une extension L de K telle que P est scindé sur L.
- Pour $n = 1$: tout polynôme de degré 1 est scindé donc $S(1)$ est vraie.
- Supposons $S(n)$ démontrée et soit P un polynôme de degré $n+1$ sur un corps commutatif K. Par (a), il existe une extension L_1 de K dans laquelle P admet une racine α. Alors P vu comme polynôme de $L_1[X]$ s'écrit $P = (X-\alpha)Q$ où $Q \in L_1[X]$ est de degré n. Puisque $S(n)$ est vraie (hypothèse de récurrence), il existe une extension L de L_1 dans laquelle le polynôme Q est scindé. Alors L est une extension de K et le polynôme $P = (X-\alpha)Q$ est scindé dans L. □

3.3 Fractions rationnelles sur un corps commutatif K

3.19 Définition. Le corps des fractions de $K[X]$ se note $K(X)$. Ses éléments s'appellent des *fractions rationnelles.*

Si A, B, D sont des polynômes avec $BD \neq 0$, on a $\dfrac{AD}{BD} = \dfrac{A}{B}$. Donc pour chaque fraction rationnelle F il existe des polynômes A, B premiers entre eux $B \neq 0$ tels que $F = \dfrac{A}{B}\cdot$ Une écriture de $F = \dfrac{A}{B}$ avec A, B premiers entre eux s'appelle une *forme irréductible* de F.

Soit F une fraction rationnelle et $F = \dfrac{A}{B}$ une forme irréductible. Les racines de A s'appellent les *zéros* ou *racines* de F ; les racines de B s'appellent les *pôles* de F. L'*ordre de multiplicité* d'un zéro (*resp.* pôle) a est l'ordre de multiplicité de la racine a de A (*resp.* B).

Soit $F \in K(X)$. Notons $\mathcal{P} \subset K$ l'ensemble de ses pôles. Soit $x \in K - \mathcal{P}$. Il existe une écriture $F = \dfrac{A}{B}$ telle que $B(x) \neq 0$. On pose alors $F(x) = A(x)B(x)^{-1}$. Cet élément de K ne dépend pas de l'écriture $F = \dfrac{A}{B}$ (avec $B(x) \neq 0$).
L'application $x \mapsto F(x)$ de $K - \mathcal{P}$ dans K s'appelle la *fonction rationnelle* associée à F.

Décomposition en éléments simples

On va peu à peu essayer de décomposer une fraction rationnelle en une somme de termes plus simples.

a) ***Partie entière*** Le degré d'une fraction rationnelle $F = \dfrac{A}{B}$ est le nombre $\partial F = \partial A - \partial B (\in \mathbb{Z})$. Ce nombre est indépendant de l'écriture. On a $\partial(FG) = \partial F + \partial G$ et $\partial(F+G) \leqslant \max\{\partial F, \partial G\}$ (avec la convention $\partial 0 = -\infty$).
Soit $F = \dfrac{A}{B}$ une fraction rationnelle. Écrivons $A = BQ + R$ la division euclidienne de A par B. On trouve $F = Q + \dfrac{R}{B}$, où $Q \in K[X] \subset K(X)$ et $\dfrac{R}{B}$ est une fraction rationnelle de degré < 0 (ou nulle). Donc :
Toute fraction rationnelle $F \in K(X)$ se décompose de façon unique en une somme d'un polynôme Q et d'une fraction rationnelle F_1 de degré strictement négatif. Le polynôme Q de cette décomposition s'appelle la partie entière *de F.*

b) ***Parties primaires.*** Soit à présent $F = \dfrac{A}{B}$ une fraction rationnelle de degré < 0. Supposons que B s'écrive $B = B_1B_2$ où B_1 et B_2 sont des polynômes premiers entre eux. D'après le théorème de Bézout, il existe des polynômes C_1 et C_2 tels que $A = B_1C_2 + B_2C_1$. Écrivons $C_1 = QB_1 + A_1$ la division euclidienne de C_1 par B_1 et posons $A_2 = C_2 + QB_2$, de sorte que $A = A_2B_1 + A_1B_2$ avec $\partial A_1 < \partial B_1$. Notons qu'alors $A_2B_1 = A - A_1B_2$, de sorte que $\partial(A_2B_1) < \partial B$; il vient $\dfrac{A}{B} = \dfrac{A_1}{B_1} + \dfrac{A_2}{B_2}$ avec $\partial A_1 < \partial B_1$ et $\partial A_2 < \partial B_2$. On vérifie que cette décomposition est unique.

Décomposant B en produit $\displaystyle\prod_{i=1}^{k} P_i^{m_i}$ où les P_i sont des polynômes irréductibles distincts on obtient (par récurrence sur k) une décomposition unique

$$\frac{A}{B} = \sum_{i=1}^{k} \frac{A_i}{P_i^{m_i}}$$

avec $\partial A_i < m_i \partial P_i$.

c) ***Éléments simples.*** Considérons enfin le cas où $F = \dfrac{A}{P^m}$ où P est irréductible, $\partial A < m\partial P$. Supposons que $m \geqslant 2$, et notons $A = PQ + R$ la division euclidienne de A par P. On trouve $F = \dfrac{R}{P^m} + \dfrac{Q}{P^{m-1}}$. Par récurrence sur m, on trouve donc que F s'écrit

$$\sum_{k=1}^{m} \frac{R_k}{P^k}$$

avec $\partial R_k < \partial P$. Cette décomposition est encore unique.

d) Mettant tout cela ensemble, on trouve que toute fraction rationnelle $F = \dfrac{A}{B}$ s'écrit de façon unique sous la forme :

$$F = E + \sum_{i=1}^{k} \sum_{j=1}^{m_i} \frac{R_{i,j}}{P_i^j}$$

où $B = b\displaystyle\prod P_i^{m_i}$ est la décomposition de B en facteurs irréductibles (et $b \in K^*$), E et $R_{i,j}$ sont des polynômes avec $\partial R_{i,j} < \partial P_i$.

Cette décomposition s'appelle la *décomposition* de F *en éléments simples.*

Lorsque P_i est un polynôme du premier degré - c'est toujours le cas si $K = \mathbb{C}$ - les $R_{i,j}$ sont de degré nul, donc des éléments de K.

Dans le cas où $K = \mathbb{R}$, on peut avoir des P_i du deuxième degré ; on aura alors des termes du premier degré au numérateur.

3.4 Exercices

3.1 Exercice. *Racines rationnelles*

Soit $P = \displaystyle\sum_{k=0}^{n} a_k X^k$ un polynôme à coefficients entiers. Soit $x = \dfrac{p}{q}$ une racine rationnelle de P écrite sous forme irréductible. Démontrer que $p|a_0$ et $q|a_n$.

3.2 Exercice. Factoriser le polynôme $P = X^5 - 4X^4 + 9X^3 - 21X^2 + 20X - 5$ sachant qu'il s'écrit comme un produit de trois polynômes à coefficients entiers.

3.3 Exercice. Décomposer en éléments simples dans $\mathbb{R}[X]$ la fraction rationnelle $F = \dfrac{X^2+2}{X^3(X-1)^2}$.

En déduire une primitive de l'application $t \mapsto \dfrac{t^2+2}{t^3(t-1)^2}$.

3.4 Exercice. Trouver un polynôme $P \in K[X]$ de degré 3 tel que $P(0) = 1$, $P'(0) = 0$, $P(1) = 0$ et $P'(1) = 1$. Quels sont les polynômes $Q \in K[X]$ qui vérifient $Q(0) = 1$, $Q'(0) = 0$, $Q(1) = 0$ et $Q'(1) = 1$?

3.5 Exercice. Calculer des primitives des fonctions

a) $x \mapsto \dfrac{1}{x^4 - x^2 - 2}$; b) $x \mapsto \dfrac{x+1}{(x^2+1)^2}$; c) $x \mapsto \dfrac{x+1}{x(x-1)^6}$; d) $x \mapsto \dfrac{1}{\cos^3 x}$.

3.6 Exercice. Résoudre le système d'équations $\begin{cases} x+y+z=3 \\ xy+yz+zx=1 \\ x^3+y^3+z^3=15 \end{cases}$ d'inconnues $x, y, z \in \mathbb{C}$.

3.7 Exercice. Soit K un corps et $a, b \in \mathbb{N}$. On considère les polynômes $A = X^a - 1$ et $B = X^b - 1$ de $K[X]$.

1. On suppose $b \neq 0$. Quel est le reste de la division euclidienne de A par B ?
2. Quel est le PGCD D de A et B ?
3. Écrire une relation de Bézout $D = AU + BV$.
4. Autre méthode : décomposer A et B en facteurs irréductibles dans $\mathbb{C}$. Pourquoi cela donne-t-il le PGCD de A et B vus comme éléments de $\mathbb{Q}[X]$?

3.8 Exercice. Soit $P \in \mathbb{R}[X]$ un polynôme unitaire sans racines réelles.

1. Démontrer qu'il existe un polynôme $A \in \mathbb{C}[X]$ tel que $P = \overline{A}\,A$ et A et $\overline{A}$ soient premiers entre eux.

   Notons k le degré de A.
2. Démontrer qu'il existe un unique polynôme $J \in \mathbb{C}[X]$ de degré $< 2k$ tel que $J \equiv i\ [A]$ et $J \equiv -i\ [\overline{A}]$.
3. Démontrer que $J \in \mathbb{R}[X]$ et $J^2 \equiv -1\ [P]$.
4. Un espace vectoriel complexe peut être considéré comme $\mathbb{R}$-espace vectoriel. Inversement, soient E un $\mathbb{R}$-espace vectoriel et j un endomorphisme de E tel que $j^2 = -\mathrm{id}_E$.
   a) Démontrer qu'il existe une unique structure d'espace vectoriel sur E telle que, pour $s, t \in \mathbb{R}$ et $x \in E$ on ait $(s+it)x = sx + tj(x)$.
   b) Munissons E de cette structure. Démontrer que les endomorphismes du $\mathbb{C}$-espace vectoriel E sont les endomorphismes f du $\mathbb{R}$ espace vectoriel E tels que $j \circ f = f \circ j$.
5. Soit E un $\mathbb{R}$ espace vectoriel de dimension finie et f un endomorphisme de E sans valeurs propres réelles. Démontrer qu'il existe sur E une structure d'espace vectoriel complexe telle que f soit $\mathbb{C}$-linéaire.

3.9 Exercice. Soit $P \in K[X]$.

1. Décomposer $\dfrac{P'}{P}$ en éléments simples.
2. *Théorème de Lucas.* On suppose $K = \mathbb{C}$. Démontrer que l'ensemble des zeros de P' est inclus dans l'enveloppe convexe de l'ensemble des zéros de P.
3. *Ellipse de Steiner.* Soient $\alpha, \beta, \gamma \in \mathbb{C}$ les affixes de trois points non alignés et notons P le polynôme $P = (X-\alpha)(X-\beta)(X-\gamma)$. On veut démontrer qu'il existe une unique ellipse (appelée ellipse de Steiner) inscrite dans le triangle de sommets (α, β, γ) et tangente aux côtés du triangle en leur milieu dont les foyers sont les racines de P'.

a) Démontrer qu'il existe une unique application affine (sur $\mathbb{R}$) $\ell : \mathbb{C} \to \mathbb{C}$ telle que $\ell(1) = \alpha$, $\ell(j) = \beta$, $\ell(j^2) = \gamma$ (où $j = e^{\frac{2i\pi}{3}}$). Démontrer qu'il existe $a, b, c \in \mathbb{C}^2$ tels $\ell(z) = az + b\overline{z} + c$ pour tout $z \in \mathbb{C}$ et que l'on a $P = (X - c)^3 - 3ab(X - c) - a^3 - b^3$.

b) Démontrer qu'il existe une unique ellipse qui soit tangente au milieu des trois côtés du triangle (α, β, γ). Démontrer que les affixes des foyers de cette ellipse sont les racines de P'.

3.10 Exercice. *Hyperbole et triangle équilatère.* Dans le plan affine euclidien on considère une hyperbole équilatère H. Notons O son centre de symétrie. Soient P un point de H, P' son symétrique par rapport à O et $\mathcal{C}$ le cercle de centre P et de rayon PP'.

1. Démontrer que $\mathcal{C}$ et H se coupent en quatre points (avec la possibilité que P' soit un point double).
2. On note A, B, C les trois autres points d'intersection de $\mathcal{C}$ avec H. Démontrer que le centre de gravité du triangle ABC est P ; en déduire que c'est un triangle équilatéral.

3.11 Exercice. *Polynômes à racines de module* 1
Soit P un polynôme unitaire à coefficients entiers. Notons $x_1, \ldots, x_n$ les racines de P (comptées avec leur multiplicité).

1. On suppose que pour tout k, on a $|x_k| = 1$.
   a) Soit $\ell \in \mathbb{N}$. Démontrer qu'il existe un polynôme unitaire P_ℓ à coefficients entiers dont les racines sont les x_k^ℓ.
   b) Soit $Q = X^n + \sum_{j=0}^{n-1} a_j X^j \in \mathbb{C}[X]$ un polynôme dont toutes les racines sont de module 1. Démontrer que $|a_j| \leqslant \binom{n}{j}$.
   c) En déduire qu'il existe ℓ et m tels que tels que $\ell \neq m$ et $P_\ell = P_m$.
   d) Démontrer qu'il existe une permutation $\sigma \in \mathfrak{S}_n$ telle que, pour tout k on ait $x_k^\ell = x_{\sigma(k)}^m$.
   e) En déduire que pour tout $r \in \mathbb{N}$, on a $x_k^{\ell^r} = x_{\sigma^r(k)}^{m^r}$.
   f) Démontrer que toutes les racines de P sont des racines de 1.
2. On suppose que toutes les racines de P sont réelles comprises entre -2 et 2. Démontrer qu'elles sont de la forme $2\cos q\pi$ avec $q \in \mathbb{Q}$. (Considérer un polynôme Q tel que $Q(x) = x^n P(x + 1/x)$).
3. Soit A une matrice symétrique à coefficients entiers de norme < 2. Démontrer que les valeurs propres de A sont de la forme $2\cos q\pi$ avec $q \in \mathbb{Q}$.

3.12 Exercice. *Résultant de deux polynômes* Soit K un corps. Pour $n \in \mathbb{N}$, notons E_n l'espace vectoriel des polynômes de degré $< n$.

Soient $A, B \in K[X]$ des polynômes non nuls. Posons $m = \partial A$ et $n = \partial B$ et écrivons $A = \sum_{k=0}^{m} a_k X^k$, $B = \sum_{k=0}^{n} b_k X^k$. On considère l'application linéaire $f_{A,B} : E_n \times E_m \to E_{m+n}$ définie par $f_{A,B}(P, Q) =$

$AP + BQ$. Pour $k = 0, \ldots, n-1$, notons C_k la matrice colonne à $n+m$ lignes :

$$C_0 = \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ \vdots \\ a_m \\ 0 \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad C_1 = \begin{pmatrix} 0 \\ a_0 \\ a_1 \\ a_2 \\ \vdots \\ a_m \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \ldots \quad C_k = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ a_0 \\ a_1 \\ a_2 \\ \vdots \\ a_m \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \ldots \quad C_{n-1} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ a_0 \\ a_1 \\ a_2 \\ \vdots \\ a_m \end{pmatrix}.$$

De même, pour $k = 0, \ldots, m-1$, notons D_k la matrice colonne à $n+m$ lignes :

$$D_0 = \begin{pmatrix} b_0 \\ b_1 \\ b_2 \\ \vdots \\ b_n \\ 0 \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad D_1 = \begin{pmatrix} 0 \\ b_0 \\ b_1 \\ b_2 \\ \vdots \\ b_n \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \ldots \quad D_k = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ b_0 \\ b_1 \\ b_2 \\ \vdots \\ b_n \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \ldots \quad D_{m-1} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ b_0 \\ b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}.$$

(La matrice C_k commence par k lignes nulles et se termine par $n-1-k$ lignes nulles ; la matrice D_k commence par k lignes nulles et se termine par $m-1-k$ lignes nulles.)

1. Démontrer l'équivalence :
   (i) les polynômes A et B sont premiers entre eux (pour $K = \mathbb{C}$ les polynômes A et B n'ont pas de racine commune) ;
   (ii) l'application f est bijective ;
   (iii) le déterminant $\mathrm{Res}_{A,B}$ de la matrice carrée de colonnes $C_0, \ldots, C_{n-1}, D_0, \ldots, D_{m-1}$ n'est pas nul. Ce déterminant s'appelle le *résultant* de A et B.
2. Pour $K = \mathbb{C}$ écrire une relation nécessaire et suffisante pour qu'un polynôme A possède des racines multiples.
3. Applications : calculer le résultant de A et B dans les cas suivant
   a) $A = aX^2 + bX + c$ et $B = A'$;
   b) $A = X^3 + pX + q$ et $B = A'$;
   c) A quelconque $B = X - b$.

   Le résultant de A et A' s'appelle le *discriminant* de A.
4. Quelques formules sur le résultant. Démontrer que l'on a :
   a) $\mathrm{Res}_{B,A} = (-1)^{mn}\mathrm{Res}_{A,B}$ pour $A \in K[X]$ de degré m et $B \in K[X]$ de degré n ;
   b) $\mathrm{Res}_{A,bB} = b^m\mathrm{Res}_{A,B}$ pour $b \in K$;
   c) $\mathrm{Res}_{A,B_1B_2} = \mathrm{Res}_{A,B_1}\mathrm{Res}_{A,B_2}$;

d) si $B = \prod_{i=1}^{n} X - y_i$, $\mathrm{Res}(A, B) = \prod_{i=1}^{n} A(y_i)$;

e) si $A = \prod_{i=1}^{m} X - x_i$, $\mathrm{Res}(A, B) = (-1)^{mn} \prod_{i=1}^{m} B(x_i)$;

f) si $A = \prod_{i=1}^{m} X - x_i$ et $B = \prod_{i=1}^{n} X - y_i$, $\mathrm{Res}(A, B) = \prod_{i,j} x_j - y_i$.

3.13 Exercice. On se propose de démontrer le :

Théorème de Sturm. *Soit $P \in \mathbb{R}[X]$ un polynôme sans racines multiples dans $\mathbb{C}$. Posons $P_0 = P$, $P_1 = P'$ et pour $k \geqslant 2$, supposant P_{k-2} et P_{k-1} construits, on écrit la division euclidienne de P_{k-2} par P_{k-1} sous la forme $P_{k-2} = Q_k P_{k-1} - P_k$. On note P_m le dernier P_k non nul. Notons A l'ensemble des nombres réels qui sont racine d'un P_k (au moins) pour $0 \leqslant k \leqslant m$. Pour $x \in \mathbb{R} \setminus A$, notons $n(x)$ le nombre de changements de signes de la suite $P_0(x), P_1(x), \ldots, P_m(x)$, c'est-à-dire le nombre de $i \in \{1, \ldots, m\}$ tels que $P_i(x)$ et $P_{i-1}(x)$ soient de signes contraires. Pour tous $a, b \in \mathbb{R} \setminus A$ tels que $a < b$, le nombre de racines de P dans l'intervalle $[a, b]$ est $n(a) - n(b)$.*

1. Démontrer que, pour tout $k < m$, P_k et P_{k+1} sont premiers entre eux. Démontrer que P_m est constant et non nul.
2. Démontrer que l'application $x \mapsto n(x)$ est constante sur tout intervalle contenu dans le complémentaire de A.
   Pour $x \in A$, on note $n_g(x)$ la limite à gauche de n en x et $n_d(x)$ sa limite à droite.
3. Soit x une racine de P_k avec $k > 0$.
   a) Démontrer que P_{k-1} et P_{k+1} ne s'annulent pas en x et sont de signes contraires. Démontrer qu'il existe un intervalle ouvert J contenant x tel que, pour $y \in J \setminus \{x\}$, le nombre de changements de signe dans la suite $P_{k-1}(y), P_k(y), P_{k+1}(y)$ soit égal à 1.
   b) Démontrer que si $x \in A$ n'est pas racine de $P_0 = P$, alors $n_g(x) = n_d(x)$.
4. Soit x une racine de P. Démontrer que $n_g(x) = n_d(x) + 1$.
   Indication : Les polynômes P et P' ont le même signe à droite de x et des signes contraires à gauche de x.
5. Établir le théorème de Sturm.

3.14 Exercice. *Résolution des équations du quatrième degré*

1. Soit $P \in K[X]$ un polynôme scindé unitaire de degré 4. Notons z_1, z_2, z_3, z_4 ses racines. Trouver un polynôme de degré 3 dont les racines sont $u_1 = z_1z_2 + z_3z_4$, $u_2 = z_1z_3 + z_2z_4$, $u_3 = z_1z_4 + z_2z_3$.
2. Si on sait résoudre les équations du troisième degré, on peut trouver u_1, u_2, u_3. Comment trouver alors les z_i ?

3.15 Exercice. Soit p un nombre premier.

1. Démontrer que dans $\mathbb{F}_p[X]$ on a l'égalité $X^p - X = \prod_{x \in \mathbb{F}_p} (X - x)$.
2. Démontrer le *théorème de Wilson* : $(p-1)! + 1 \equiv 0 \ \ [p]$.

3.16 Exercice. [Contenu d'un polynôme]

1. Soient $A, B \in \mathbb{Z}[X]$.
   Écrivons $A = \sum_{k=0}^{m} a_k X^k$, $B = \sum_{k=0}^{n} b_k X^k$ et $AB = \sum_{k=0}^{m+n} c_k X^k$. Soit p un nombre premier. On suppose que p divise tous les c_k. Démontrer que p divise tous les a_k ou tous les b_k.
   On appelle *contenu* d'un polynôme $P = \sum_{k=0}^{n} p_k X^k$ à coefficients dans $\mathbb{Z}$ et on note $c(P)$ le PGCD de ses coefficients $p_0, \ldots, p_n$.

2. Soient $A, B \in \mathbb{Z}[X]$. Démontrer que si $c(A) = c(B) = 1$, alors $c(AB) = 1$. En déduire que l'on a toujours $c(AB) = c(A)c(B)$.
3. Soit $P \in \mathbb{Z}[X]$ un polynôme non constant. Démontrer que si P est irréductible dans $\mathbb{Z}[X]$ il est irréductible dans $\mathbb{Q}[X]$.

3.17 Exercice. [Critère d'Eisenstein] Soient P un polynôme unitaire à coefficients entiers et p un nombre premier. On suppose que p divise tous les coefficients de P - sauf le coefficient dominant - et que $P(0)$ n'est pas divisible par p^2. Démontrer que P est irréductible sur $\mathbb{Z}$ - donc sur $\mathbb{Q}$.

Application. Démontrer que pour tout nombre premier p le polynôme $\Phi_p = \sum_{k=0}^{p-1} X^k$ est irréductible sur $\mathbb{Q}$.

3.18 Exercice. (**) *Comment trouver les racines d'un polynôme dans $\mathbb{F}_p$?*
On se donne $P \in K[X]$ dont on veut trouver les racines dans K.

1. Trouver une méthode pour isoler les racines multiples.

   Indication : On pourra utiliser la dérivée de P.
2. On suppose que $K = \mathbb{F}_p$ où p est un (grand !) nombre premier. Donner une méthode pour trouver un polynôme scindé à racines simples ayant les mêmes racines que P.

   Indication : Penser au polynôme $X^p - X$.
3. On suppose que P est scindé à racines simples.
   a) Ecrire $P = AB$ où les racines de A sont des carrés dans $\mathbb{F}_p$ et celles de B ne le sont pas.
   b) Soient a, b deux racines (qu'on ne connait pas). On veut les séparer, c'est à dire écrire $P = AB$ avec a racine de A et b de B. Pour cela, on cherche un polynôme Q dont a ou b est racine mais pas l'autre, puis on prend le PGCD de P et Q. (On dit que Q sépare a et b.) Soit $c \in \mathbb{F}_p$ - distinct de a et de b. On pose $Q = (X - c)^{\frac{p-1}{2}} - 1$. Démontrer que Q sépare a et b si et seulement si $\dfrac{c-a}{c-b}$ n'est pas un carré.
   En choisissant c au hasard, on a donc une chance sur 2 de séparer a et b.
   c) Esquisser une méthode qui va nous permettre de trouver toutes les racines de P (le degré de P est ici supposé petit par rapport à p).

3.19 Exercice. (****) *Polynômes irréductibles dans $\mathbb{F}_p[X]$.*

1. *Fonction de Moebius.* On définit la fonction de Moebius $\mu : \mathbb{N}^* \to \mathbb{Z}$ en posant $\mu(n) = 0$ si n a des facteurs carrés, $\mu(1) = 1$ et $\mu(p_1 p_2 \dots p_n) = (-1)^n$ si les p_i sont des nombres premiers distincts.
   a) Démontrer que si m, n sont premiers entre eux, on a $\mu(mn) = \mu(m)\mu(n)$.
   b) Soient $(a_n)_{n\in\mathbb{N}^*}$ et $(b_n)_{n\in\mathbb{N}^*}$ des suites de nombres réels. Démontrer que l'on a $a_n = \sum_{d|n} b_d$ pour tout n si et seulement si on a $b_n = \sum_{d|n} \mu(\frac{n}{d}) a_d$ pour tout n.
2. On note Q l'ensemble des polynômes unitaires à coefficients dans $\mathbb{F}_p$ et P l'ensemble des polynômes unitaires irréductibles. On note N_n le nombre de polynômes unitaires irréductibles de degré n.
   a) Démontrer que pour $t \in]-1/p, 1/p[$ on a
   $$\frac{1}{1-pt} = \sum_{A\in Q} t^{\partial A} = \prod_{R\in P} \frac{1}{1-t^{\partial R}} = \prod_{n=1}^{+\infty} (1-t^n)^{-N_n}.$$
   b) Démontrer que $p^n = \sum_{d|n} d N_d$.

Indication : Prendre le logarithme - ou la dérivée logarithmique.

c) En déduire que $nN_n = \sum_{d|n} \mu(\frac{n}{d}) p^d$.

d) Remarquant que n a au plus $\frac{n}{2}$ diviseurs distincts de n tous $\leqslant \frac{n}{2}$, en déduire que $nN_n \geqslant p^{n/2}(p^{n/2} - n/2)$, puis que $N_n > 0$ pour tout $n > 0$.

e) En déduire l'existence d'un corps à p^n éléments.

Deuxième partie

Algèbre linéaire sur un sous-corps de $\mathbb{C}$

4 Définitions et généralités

4.1 Espaces vectoriels

4.1 Définition. Soit K un corps commutatif. Un *espace vectoriel* sur K (ou K-*espace vectoriel*) est un ensemble E muni
- d'une loi de composition interne notée $+$,
- d'une loi externe (action de K) $K \times E \to E$ notée $(\lambda, x) \mapsto \lambda x$,
  telles que
  * $(E, +)$ soit un groupe commutatif ;
  * pour tous $\lambda, \mu \in K$ et tous $x, y \in E$ on a :

$$\lambda(x+y) = \lambda x + \lambda y, \quad (\lambda+\mu)x = \lambda x + \mu x, \quad (\lambda\mu)x = \lambda(\mu x) \quad \text{et} \quad 1x = x.$$

Rappelons la notion suivante qui prend un sens grâce à l'associativité et la commutativité de la somme :

4.2 Définition. Soient E un K-espace vectoriel et A une partie de E. On appelle *combinaison linéaire* d'éléments de A un élément x de E qui s'écrit sous la forme $x = \sum_{k=1}^{n} \lambda_k x_k$ où $n \in \mathbb{N}$, $\lambda_1, \ldots, \lambda_n$ sont des éléments de K et $x_1, \ldots, x_n$ des éléments de A.

4.3 Exemples.
a) Muni de l'addition et de la multiplication de K, le corps K est un K-espace vectoriel.

b) Muni de l'addition des polynômes et du produit d'un polynôme par un scalaire, $K[X]$ est un K-espace vectoriel.

c) Sur K^n on considère l'addition et l'action de K données par les formules :
$(x_1, \ldots, x_n) + (y_1, \ldots, y_n) = (x_1 + y_1, \ldots, x_n + y_n)$ et $\lambda(x_1, \ldots, x_n) = (\lambda x_1, \ldots, \lambda x_n)$.
Muni de ces opérations, K^n est un K-espace vectoriel.

d) Plus généralement, donnons nous un entier n et une famille $(E_1, \ldots, E_n)$ de K-espaces vectoriels. Notons $\mathcal{E}$ le produit $\prod_{k=1}^{n} E_k = E_1 \times \ldots \times E_n$ des E_k. Les éléments de $\mathcal{E}$ sont des suites $(x_1, \ldots, x_n)$ où $x_k \in E_k$. Sur $\mathcal{E}$ on considère l'addition et l'action de K données par les formules : $(x_1, \ldots, x_n) + (y_1, \ldots, y_n) = (x_1 + y_1, \ldots, x_n + y_n)$ et $\lambda(x_1, \ldots, x_n) = (\lambda x_1, \ldots, \lambda x_n)$. Muni de ces opérations, $\mathcal{E}$ est un K-espace vectoriel. Cet espace s'appelle l'*espace vectoriel produit* des E_i.

4.2 Sous-espaces vectoriels

4.4 Définition. Soit E un K-espace vectoriel. On appelle *sous-espace vectoriel* de E une partie F de E, qui est un sous-groupe de E pour $+$ et telle que pour tout $x \in F$ et tout $\lambda \in K$ on ait $\lambda x \in F$.

Pour vérifier que $F \subset E$ est un sous-espace vectoriel on doit vérifier :
- $0 \in F$ et $\forall \lambda \in K$, $x, y \in F$, on a $x + y \in F$ et $\lambda x \in F$;
  ou
- $F \neq \emptyset$ et $\forall \lambda, \mu \in K$, $x, y \in F$, on a $\lambda x + \mu y \in F$.

Plus généralement, un sous-espace vectoriel est stable par combinaisons linéaires :

4.5 Proposition. *Soient E un K-espace vectoriel et F un sous-espace vectoriel de E. Toute combinaison linéaire d'éléments de F est dans F.*

4.6 Proposition. *Soient E un K-espace vectoriel. L'intersection d'une famille de sous-espaces vectoriels de E est un sous-espace vectoriel de E.*

Autrement dit, si $(E_i)_{i\in I}$ est une famille de sous-espaces vectoriels de E, son intersection *i.e.* l'ensemble $\bigcap_{i\in I} E_i = \{x \in E;\ \forall i \in I,\ x \in E_i\}$ est un sous-espace vectoriel de E.

4.7 Exemples. Soit E un K-espace vectoriel.

a) Soit A une partie de E. Il existe un plus petit sous-espace vectoriel de E qui contient A : l'intersection de tous les sous-espaces de E contenant A.

b) Soit $(E_i)_{i\in I}$ une famille de sous-espaces de E. Il existe un plus petit sous-espace vectoriel de E qui contient tous les E_i : c'est le plus petit sous-espace de E contenant leur réunion $A = \bigcup_{i\in I} E_i$.

4.8 Définition. Soit E un K-espace vectoriel.

a) Soit A une partie de E. Le plus petit sous-espace vectoriel de E qui contient A s'appelle le *sous-espace vectoriel engendré* par A. Nous le noterons $\mathrm{Vect}(A)$

b) Soit $(E_i)_{i\in I}$ une famille de sous-espaces vectoriels de E. Le plus petit sous-espace vectoriel de E qui contient tous les E_i s'appelle la *somme* des E_i et se note $\sum_{i\in I} E_i$. Si $I = \{1, \ldots, n\}$ cette somme se note aussi $E_1 + \cdots + E_n$.

4.9 Proposition. *Soit E un K-espace vectoriel.*

a) Soit A une partie de E. Le sous-espace engendré par une partie A de E est l'ensemble des combinaisons linéaires d'éléments de A.

b) Soit $(E_1, \ldots, E_n)$ des sous-espaces vectoriels de E. La somme $\sum_{i=1}^{n} E_i$ des E_i est l'ensemble

$$\Big\{ \sum_{k=1}^{n} x_k;\ (x_1, \ldots, x_n) \in E_1 \times \ldots \times E_n \Big\}.$$

Pour établir cette proposition, on démontre que les combinaisons linéaires d'éléments de A (*resp.* des sommes d'éléments des E_k) forment un sous-espace vectoriel et on utilise la proposition 4.5 pour démontrer que c'est le plus petit.

4.10 Définition. Soit E un K-espace vectoriel.

a) Deux sous-espaces F, G de E sont dits *supplémentaires* si $F + G = E$ et $F \cap G = \{0\}$.

b) Soient $n \in \mathbb{N}$ ($n \geqslant 2$) et $(E_1, \ldots, E_n)$ des sous-espaces de E. On dit que les sous-espaces E_j sont *en somme directe* si l'application $(x_1, \ldots, x_n) \mapsto x_1 + \ldots + x_n$ est injective.

4.11 Exercices. Soit E un K-espace vectoriel.

a) Soient F, G deux sous-espaces vectoriels de E. Démontrer qu'ils sont supplémentaires si et seulement si l'application $(x, y) \mapsto x + y$ de $F \times G$ dans E est bijective.

b) Soient $n \in \mathbb{N}$ et $(E_1, \ldots, E_n)$ des sous-espaces de E. Démontrer que les espaces E_j sont en somme directe si et seulement si pour tout $k \in \{1, \ldots, n-1\}$ on a $(E_1 + \ldots + E_k) \cap E_{k+1} = \{0\}$.

4.3 Applications linéaires

4.12 Définition. Soient E, F deux K-espaces vectoriels.
- Une application $f : E \to F$ est dite *linéaire* (ou *K-linéaire*) si $\forall x, y \in E$ et $\forall \lambda \in K$ on a $f(x+y) = f(x) + f(y)$ et $f(\lambda x) = \lambda f(x)$.
- Une application linéaire de E dans E s'appelle aussi un *endomorphisme* de E.
- On appelle *isomorphisme* (d'espaces vectoriels) une application linéaire bijective.
- On appelle *automorphisme* un endomorphisme bijectif.

4.13 Proposition. *a) La composée d'applications linéaires est linéaire.*

b) La réciproque d'un isomorphisme est linéaire : c'est un isomorphisme.

4.14 Exemple. Soient E un espace vectoriel et F, G deux sous-espaces supplémentaires de E. Pour tout $x \in E$, il existe un unique élément $y \in F$ tel que $x - y \in G$. Notons $P(x)$ cet unique élément. L'application $P : E \to E$ ainsi définie (d'image égale à F) est linéaire. On l'appelle le *projecteur sur F parallèlement à G*. Pour $y \in F$ et $z \in G$, on a $P(y+z) = y$. Remarquons que $P \circ P = P$: en d'autres termes P est idempotent.

Inversement, tout endomorphisme idempotent est de cette forme.

4.15 Proposition. *Soient E, F des espaces vectoriels et $f : E \to F$ une application linéaire. L'image $f(E_1)$ d'un sous-espace vectoriel E_1 de E est un sous-espace vectoriel de F ; l'image réciproque $f^{-1}(F_1)$ d'un sous-espace vectoriel F_1 de F est un sous-espace vectoriel de E.*

4.16 Définition. Soient E, F des espaces vectoriels et $f : E \to F$ une application linéaire. On appelle *image* de f et on note $\operatorname{im} f$ le sous-espace $f(E)$ de F ; on appelle *noyau* de f et on note $\ker f$ le sous-espace $f^{-1}(\{0\})$ de E.

4.17 Proposition. *Soient E, F des espaces vectoriels et $f : E \to F$ une application linéaire. L'application f est injective si et seulement si $\ker f = \{0\}$; l'application f est surjective si et seulement si $\operatorname{im} f = F$.*

4.18 Proposition. *Soient E, F des espaces vectoriels et $f : E \to F$ une application linéaire. Soit E_1 un supplémentaire de $\ker f$ dans E et notons $g : E_1 \to \operatorname{im} f$ l'application $x \mapsto f(x)$. Alors g est un isomorphisme.*

4.4 Ensembles d'applications linéaires

4.19 Proposition. *Soient E, F des K-espaces vectoriels.*

a) Soient f, g des applications linéaires de E dans F et $\lambda \in K$. Les applications $f + g : x \mapsto f(x) + g(x)$ et $\lambda f : x \mapsto \lambda f(x)$ (de E dans F) sont linéaires.

b) Muni de ces opérations, l'ensemble des application linéaires de E dans F est un K-espace vectoriel noté $L(E, F)$.

Lorsque $E = F$, l'espace $L(E, F)$ se note $L(E)$. Muni de l'addition et de la composition des applications linéaires, $L(E)$ est un anneau. De plus, les structures d'espace vectoriel et d'anneau sont compatibles au sens suivant : pour $g \in L(E)$, les applications $f \mapsto g \circ f$ et $f \circ g$ sont linéaires. En d'autres termes, $L(E)$ est une K-algèbre.

Rappelons pour mémoire la définition d'une K-algèbre [1] :

4.20 Définition. Une K-algèbre est un ensemble A muni de trois lois :
- une addition $+$ (qui est une loi interne $A \times A \to A$) ;

1. Nos algèbres sont supposées associatives. Cette convention n'est pas prise par tous les ouvrages.

- une multiplication $(a, b) \mapsto ab$ (interne) ;
- une loi externe $K \times A \to A$, notée $(\lambda, a) \mapsto \lambda a$.

Ces lois doivent satisfaire :

a) muni des lois internes (addition et multiplication) A est un anneau ;

b) muni de l'addition et de la loi externe A est un espace vectoriel ;

c) les multiplications interne et externe vérifient : $\forall a, b \in A$ et $\forall \lambda \in K$ on a $(\lambda a)b = \lambda(ab) = a(\lambda b)$.

4.21 Proposition. *L'ensemble des automorphismes d'un K-espace vectoriel E est un groupe (pour la composition). On l'appelle* groupe linéaire *de E et on le note $GL(E)$.*

Pour démontrer que $GL(E)$ est un groupe, on démontre en fait que c'est un sous-groupe du groupe des bijections de E dans E : l'application identité id_E de E est linéaire, et on applique la proposition 4.13.

4.5 Familles libres, génératrices, bases

4.5.1 Familles, familles de vecteurs

Soient X et I des ensembles. On appelle *famille* d'éléments de X indicée par I une application de I dans X. Cependant, la notation est un peu modifiée. Si f est une famille d'éléments de X indicée par I, l'élément $f(i)$ (pour un point $i \in I$) se note avec i en indice, par exemple x_i. La famille f elle-même se note $(x_i)_{i\in I}$.

Soient $(x_i)_{i\in I}$ et $(y_j)_{j\in J}$ des familles d'éléments de X. On dira que $(x_i)_{i\in I}$ est une *sous-famille* de $(y_j)_{j\in J}$ (on dit parfois que $(y_j)_{j\in J}$ est une sur-famille de $(x_i)_{i\in I}$) si $I \subset J$ et pour tout $i \in I$, on a $x_i = y_i$.

Soit E un K-espace vectoriel. Une *famille de vecteurs* (de E) est donc une application d'un ensemble I dans E notée $(x_i)_{i\in I}$. On parle aussi parfois de *système de vecteurs*. On s'intéressera ici surtout au cas des familles finies $(x_i)_{i\in I}$, c'est à dire au cas où I est un ensemble fini. Le plus souvent, on prendra $I = \{1, \ldots, n\}$.

4.22 Remarque. Une partie A de E détermine une famille de vecteurs : $(x)_{x\in A}$ (qui est l'application de A dans E qui à $x \in A$ associe $x \in E$).

4.5.2 Applications linéaires de K^I dans E

Fixons un ensemble fini I. Rappelons que K^I est le K-espace vectoriel des familles d'éléments de K indicées par I, *i.e.* des applications de I dans K Pour $i \in I$, notons $e_i \in K^I$ l'élément $(s_j)_{j\in I}$ déterminé par $s_i = 1$ et $s_j = 0$ pour $j \neq i$.

Remarquons qu'un élément $(\lambda_i)_{\in I}$ s'écrit alors $\sum_{i\in I} \lambda_i e_i$.

4.23 Proposition. *Soit E un K-espace vectoriel.*

a) Soit $(x_i)_{i\in I}$ une famille finie de vecteurs de E. L'application $f : (\lambda_i)_{i\in I} \mapsto \sum_{i\in I} \lambda_i x_i$ de K^I dans E linéaire.

b) Toute application linéaire f de K^I dans E est de cette forme : il existe une unique famille $(x_i)_{i\in I}$ d'éléments de E telle que pour tout $(\lambda_i)_{i\in I} \in K^I$ on ait $f((\lambda_i)_{i\in I}) = \sum_{i\in I} \lambda_i x_i$.

On a $x_i = f(e_i)$ pour tout $i \in I$.

4.5.3 Familles libres, génératrices, bases

4.24 Définition. Soient E un K-espace vectoriel $(x_i)_{i\in I}$ une famille finie de vecteurs de E. Notons $f : K^I \to E$ l'application $(\lambda_i)_{i\in I} \mapsto \sum_{i\in I} \lambda_i x_i$.

- On dit que la famille $(x_i)_{i\in I}$ est *libre* si f est injective (sinon, on dit qu'elle est *liée*) ;
- on dit que la famille $(x_i)_{i\in I}$ est *génératrice* si f est surjective ;
- on dit que la famille $(x_i)_{i\in I}$ est une *base* de E si f est bijective.

4.25 Remarques. a) La famille $(x_i)_{i\in I}$ est libre si le noyau de f est réduit à $\{0\}$, c'est-à-dire si la condition $\sum_{i\in I} \lambda_i x_i = 0$ implique que tous les λ_i sont nuls.

b) L'image de l'application f est le sous-espace vectoriel de E engendré par $\{x_i;\ i \in I\}$; la famille $(x_i)_{i\in I}$ est génératrice si ce sous-espace est E.

c) La famille $(x_i)_{i\in I}$ est une base si elle est à la fois libre et génératrice.

4.26 Proposition. *a) Une sous-famille d'une famille libre est libre.*

b) Une sur-famille d'une famille génératrice est génératrice.

4.27 Généralisation. Soient E un K-espace vectoriel et $(x_i)_{i\in I}$ une famille quelconque de vecteurs de E.

a) La famille $(x_i)_{i\in I}$ est dite génératrice si le sous-espace engendré par $\{x_i;\ i \in I\}$ est E ;

b) la famille $(x_i)_{i\in I}$ est dite libre si toute sous-famille finie est libre ;

c) la famille $(x_i)_{i\in I}$ est une base si elle est à la fois libre et génératrice.

4.28 Proposition. *Soit $(x_i)_{i\in I}$ une famille de vecteurs de E. Les propriétés suivantes sont équivalentes.*

(i) *$(x_i)_{i\in I}$ est libre maximale : toute sur-famille libre de $(x_i)_{i\in I}$ est égale à $(x_i)_{i\in I}$;*

(ii) *$(x_i)_{i\in I}$ est génératrice minimale : toute sous-famille génératrice de $(x_i)_{i\in I}$ est égale à $(x_i)_{i\in I}$;*

(iii) *$(x_i)_{i\in I}$ est une base.*

4.6 Matrices

Soient I, J deux ensembles finis et K un corps. Une *matrice* de type $I \times J$ à coefficients dans K est une famille d'éléments de K indicée par $I \times J$. On la représente par un tableau dont les lignes sont indicées par I et les colonnes par J. On note $\mathcal{M}_{I,J}(K)$ l'espace des matrices de type $I \times J$ à coefficients dans K. Si $I = \{1, \dots, m\}$ et $J = \{1, \dots, n\}$, les matrices de type $I \times J$ s'appellent des matrices d'ordre m, n et on écrit $\mathcal{M}_{m,n}(K)$ plutôt que $\mathcal{M}_{I,J}(K)$. Enfin lorsque $I = J$ (ou $m = n$) on parle de matrices carrées de type I (ou d'ordre m). On note $\mathcal{M}_n(K)$ l'ensemble des matrices carrées d'ordre n.

4.6.1 Applications linéaires de K^J dans K^I

Soient I et J deux ensembles finis. Il résulte en particulier de la proposition 4.23 qu'une application K-linéaire f de K^J dans K^I est déterminée par une famille $(x_j)_{j\in J}$ d'éléments de K^I. Une telle famille est donnée par une matrice $(a_{i,j})_{(i,j)\in I\times J}$ à coefficients dans K, appelée *matrice de f*, et l'on a donc

- $f(e_j) = (a_{i,j})_{i\in I}$ pour tout $j \in J$;
- $f((\lambda_j)_{j\in J}) = (\mu_i)_{i\in I}$ avec $\mu_i = \sum_{j\in J} a_{i,j}\lambda_j$, pour tout $(\lambda_j)_{j\in J} \in K^J$.

4.6.2 Produit matriciel

Soient I, J, L trois ensembles finis, $f : K^L \to K^J$ et $g : K^J \to K^I$ des applications linéaires. Notons $A = (a_{j,\ell})_{(j,\ell)\in J\times L}$ et $B = (b_{i,j})_{(i,j)\in I\times J}$ leurs matrices respectives. La matrice de $g \circ f$ est $(c_{i,\ell})_{(i,\ell)\in I\times L}$ où $c_{i,\ell} = \sum_{j\in J} b_{i,j}a_{j,\ell}$. Cette matrice s'appelle le produit des matrices B et A et se note BA.

4.6.3 Matrices inversibles. Groupe $GL(n, K)$

Soit J un ensemble fini. La matrice de l'application identité de K^J dans lui même est la matrice carrée appelée matrice identité $(\delta_{i,j})_{(i,j)\in J\times J}$ telle que, pour tout $i, j \in I$ on ait $\delta_{i,j} = 1$ si $i = j$ et $\delta_{i,j} = 0$ si $i \neq j$. Convenons de noter 1_J cette matrice. Pour $J = \{1, \ldots, n\}$ on écrira plutôt 1_n ou I_n.

Une matrice $A \in \mathcal{M}_{I,J}(K)$ est dite inversible si elle représente un isomorphisme, *i.e.* s'il existe $B \in \mathcal{M}_{J,I}(K)$ telle que $AB = 1_I$ et $BA = 1_J$.

On appelle *groupe linéaire (d'ordre n)* et l'on note $GL(n, K)$ le groupe des matrices carrées d'ordre n inversibles.

4.7 Exercices

4.1 Exercice. On note E l'ensemble des fonctions continues de $\mathbb{R}$ dans $\mathbb{R}$.

1. Démontrer que E est naturellement muni d'une structure de $\mathbb{R}$-espace vectoriel.
2. On note $P \subset E$ et $I \subset E$ les sous-ensembles formés des fonctions continues paires et impaires respectivement. Démontrer que P et I sont des sous-espaces vectoriels supplémentaires de E.
3. Donner quelques exemples de décomposition $f = p + i$ avec $f \in E$, $p \in P$ et $i \in I$.

4.2 Exercice. Soient F, G, H des sous-espaces vectoriels d'un espace vectoriel E. On suppose que $G \subset H$, $F \cap H \subset F \cap G$ et $F + H \subset F + G$. Démontrer que $G = H$.

4.3 Exercice.

1. Soient F, G des sous-espaces vectoriels d'un espace vectoriel E. On suppose que $F \cup G$ est un sous-espace vectoriel de E. Démontrer que l'on a $F \subset G$ ou $G \subset F$?
2. Soient $F_1, \ldots, F_n$ des sous-espaces vectoriels stricts ($F_j \neq E$) d'un K-espace vectoriel E.
   a) Soit $k \in \{1, \ldots, n-1\}$ et soient $x, y \in E$ tels que $x \notin \bigcup_{i=1}^{k} F_i$ et $y \notin F_{k+1}$. Démontrer que pour $j \in \{1, \ldots, k\}$ il existe au plus un $\lambda_j \in K$ tel que $y + \lambda_j x \in F_j$.
   b) On suppose que K est infini. Démontrer que $\bigcup_{i=1}^{n} F_i \neq E$.

4.4 Exercice. Soient E, F des espaces vectoriels et $f : E \to F$ une application. Démontrer que f est linéaire si et seulement si son graphe $G_f = \{(x, f(x));\ x \in E\}$ est un sous-espace vectoriel de l'espace vectoriel produit $E \times F$.

4.5 Exercice. Soient E, F, G des espaces vectoriels, $f : E \to F$ et $g : F \to G$ des applications linéaires. Démontrer que $\ker g \cap \operatorname{im} f = f(\ker g \circ f)$.

4.6 Exercice.

Soient $a, b \in K$. Notons $E \subset K^{\mathbb{N}}$ l'ensemble des suites $(x_n)_{n\in K}$ telles que, pour tout $n \geqslant 2$ on ait $x_n = ax_{n-1} + bx_{n-2}$.

1. Démontrer que E est un sous-espace vectoriel de $K^{\mathbb{N}}$.

2. Soit $(x_n)_{n\in\mathbb{N}}$ un élément de E tel que $x_0 = x_1 = 0$. Démontrer que l'on a $x_n = 0$ pour tout $n \in \mathbb{N}$.
3. Posons $A = \begin{pmatrix} a & b \\ 1 & 0 \end{pmatrix}$. Écrivons $A^n \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} u_n \\ v_n \end{pmatrix}$. Démontrer que
   a) pour tout $n \in \mathbb{N}$ on a $v_{n+1} = u_n$;
   b) $(u_n)_{n\in\mathbb{N}}$ et $(v_n)_{n\in\mathbb{N}}$ sont des éléments de E ;
   c) $(u_n)_{n\in\mathbb{N}}$ et $(v_n)_{n\in\mathbb{N}}$ forment une base de E.
4. Posons $F = \{(x_n)_{n\in\mathbb{N}};\ x_0 = x_1 = 0\}$. Démontrer que E et F sont des sous-espaces supplémentaires de $K^{\mathbb{N}}$.

5 Théorie de la dimension

5.1 Espaces vectoriels de dimension finie

5.1 Définition. On dit qu'un K-espace vectoriel est de *dimension finie* s'il possède une famille génératrice finie.

Il sera ici plus commode de raisonner en termes de *parties* libres ou génératrices d'un espace vectoriel :

5.2 Remarque. Soit E un K-espace vectoriel. Une partie X de E détermine une famille $(x)_{x\in X}$ et l'on peut donc parler de parties libres ou génératrices. Soit$(x_i)_{i\in I}$ une famille de vecteurs de E. Posons $X = \{x_i;\ i \in I\}$. La famille $(x_i)_{i\in I}$ est une famille génératrice de E si et seulement si X est une partie génératrice de E ; la famille $(x_i)_{i\in I}$ est une famille libre de E si et seulement si X est une partie libre de E et l'application $i \mapsto x_i$ est injective.

5.3 Théorème de la base incomplète. *Soient E un K-espace vectoriel, $G \subset E$ une partie génératrice finie de E et $L \subset G$ une partie libre de E. Alors il existe une base B de E telle que $L \subset B \subset G$.*

L'ensemble $\mathcal{G} = \{X$ partie génératrice de $E;\ L \subset X \subset G\}$ n'est pas vide puisque $G \in \mathcal{G}$. Un élément B de $\mathcal{G}$ qui possède un nombre minimum d'éléments est une base de E.

5.4 Corollaire. *Soit E un K-espace vectoriel de dimension finie.*

a) Toute partie génératrice finie de E contient une base ;

b) toute partie libre finie est contenue dans une base.

Pour (a) on applique le théorème de la base incomplète à $L = \emptyset$; pour (b) si L est une partie libre finie et G_1 est une partie génératrice finie, on applique le théorème de la base incomplète à L et $G = L \cup G_1$.

5.2 Dimension d'un espace vectoriel

5.5 Lemme d'échange (Steinitz). *Soit $G \subset E$ une partie génératrice finie, $x \in G$ et $y \in E$. Écrivons $y = \sum_{z\in G} \lambda_z z$, et supposons que $\lambda_x \neq 0$. Alors $(G \setminus \{x\}) \cup \{y\}$ est génératrice.*

Écrivons $G_1 = G \setminus \{x\}$ et $G_2 = G_1 \cup \{y\}$. On a $x = \lambda_x^{-1}\left(y - \sum_{z\in G_1} \lambda_z z\right)$, donc $x \in \mathrm{Vect}(G_2)$. Il vient $G \subset \mathrm{Vect}(G_2)$, donc $E = \mathrm{Vect}(G) \subset \mathrm{Vect}(G_2)$.

5.6 Théorème. *Soient E un espace vectoriel, G une partie génératrice finie de E et L une partie libre de E. Alors L est finie et* $\mathrm{Card}(L) \leqslant \mathrm{Card}(G)$.

Supposons d'abord que L soit finie et démontrons que $\mathrm{Card}(L) \leqslant \mathrm{Card}(G)$.

On raisonne par récurrence sur le nombre N d'éléments de L qui ne sont pas dans G.

- Si $N = 0$, alors $L \subset G$, donc $\mathrm{Card}(L) \leqslant \mathrm{Card}(G)$.
- Soit $N \in \mathbb{N}$ et supposons que l'on ait démontré que si L et G sont deux parties finies de E avec L libre et G génératrice et $\mathrm{Card}(L \setminus G) = N$, alors $\mathrm{Card}(G) \geqslant \mathrm{Card}(L)$.
  Soient maintenant L une partie libre et G une partie génératrice de E telles que $\mathrm{Card}(L\setminus G) = N+1$. Posons $L_1 = (G \cap L)$. Soit $y \in L \setminus G$. Comme G est génératrice, on peut écrire $y = \sum_{z\in G} \lambda_z z$. Comme L est libre, il vient $y \notin \mathrm{Vect}(L_1)$, donc il existe $x \in G \setminus L_1$ tel que $\lambda_x \neq 0$. Par le lemme d'échange $G_1 = (G \setminus \{x\}) \cup \{y\}$ est génératrice et $L \setminus G_1$ a N éléments. Par l'hypothèse de récurrence, on a $\mathrm{Card}(L) \leqslant \mathrm{Card}(G_1) = \mathrm{Card}(G)$.

Enfin, une partie infinie T de E contient une partie finie de cardinal $\mathrm{Card}(G)+1$, qui ne peut être libre ; donc T n'est pas libre.

5.7 Corollaire. *Dans un espace vectoriel de dimension finie, toutes les bases sont finies et ont même nombre d'éléments.*

5.8 Définition. Soit E un espace vectoriel de dimension finie. On appelle *dimension* de E le nombre $\dim E$ d'éléments d'une base quelconque de E.

5.9 Théorème. *Soit E un espace vectoriel de dimension finie. Posons $n = \dim E$.*

a) Toutes les bases de E ont n éléments.

b) Toute famille libre de E a au plus n éléments ; une famille libre à n éléments est une base.

c) Toute famille génératrice de E a au moins n éléments ; une famille génératrice à n éléments est une base.

5.10 Théorème. *Soit E un espace vectoriel de dimension finie. Tout sous-espace vectoriel F de E est de dimension finie. On a $\dim F \leqslant \dim E$; si $\dim F = \dim E$, alors $F = E$.*

Toute partie libre de F est libre dans E donc a au plus n éléments. Notons $k(\leqslant n)$ le plus grand nombre d'éléments d'une partie libre de F et soit L une partie libre de F à k éléments. C'est une partie libre maximale, donc une base de F. Donc $\dim F = k \leqslant n$. Si $k = n$, alors L est une base de E, donc $F = E$.

5.11 Corollaire. *Tout sous-espace vectoriel d'un espace vectoriel de dimension finie possède un supplémentaire.*

En effet, soient E un espace vectoriel de dimension finie, F un sous-espace vectoriel de E et L une base de F. On peut la compléter en une base B de E. Le sous-espace vectoriel engendré par $B \setminus L$ est un supplémentaire de F.

5.12 Proposition. *Soient F, G des sous-espaces vectoriels de dimension finie d'un espace vectoriel E. Alors $F+G$ et $F \cap G$ sont de dimension finie et l'on a $\dim F + \dim G = \dim(F+G) + \dim(F \cap G)$.*

Supposons d'abord que $F \cap G = \{0\}$. On obtient une base de $F+G$ en réunissant une base de F avec une base de G : on a donc $\dim(F+G) = \dim F + \dim G$.

Dans le cas général, choisissons un supplémentaire G_1 de $F \cap G$ dans G. Alors G_1 est un supplémentaire de F dans $F+G$ et par le premier cas, on a $\dim G = \dim(F \cap G) + \dim G_1$, et $\dim(F+G) = \dim F + \dim G_1$.

5.3 Rang

5.13 Définition. a) Une famille $(x_i)_{i \in I}$ de vecteurs d'un espace vectoriel est dite de *rang fini* si l'espace vectoriel qu'elle engendre est de dimension finie. La dimension de cet espace s'appelle *rang* de la famille $(x_i)_{i \in I}$ et se note $\mathrm{rg}(x_i)_{i \in I}$.

b) Une application linéaire f est dite de *rang fini* si son image est de dimension finie. La dimension de $\mathrm{im}\, f$ s'appelle *rang* de f et se note $\mathrm{rg}(f)$.

c) Le *rang* d'une matrice d'ordre m, n est le rang de l'application linéaire de K^n dans K^m qu'elle représente.

5.14 Théorème du rang. *Soient E, F des espaces vectoriels et $f : E \to F$ une application linéaire. Si E est de dimension finie on a $\dim E = \dim(\ker f) + \mathrm{rg}(f)$.*

Cela résulte immédiatement de la proposition 4.18.

5.15 Théorème. *Soient E, F des espaces vectoriels de dimension finie et $f : E \to F$ une application linéaire.*

a) Si $\dim E < \dim F$, l'application f n'est pas surjective.

b) Si $\dim E > \dim F$, l'application f n'est pas injective.

c) Si $\dim E = \dim F$ et en particulier si $E = F$, i.e. si f est un endomorphisme, on a l'équivalence entre : (i) f est surjective ; (ii) f est injective ; (iii) f est bijective.

Pour vérifier qu'un endomorphisme d'un espace vectoriel de dimension finie est un automorphisme, il suffit de vérifier son injectivité ou sa surjectivité.

5.16 Théorème. *Deux K-espaces vectoriels de dimension finie sont isomorphes si et seulement s'ils ont même dimension.*

5.4 Exercices

5.1 Exercice. Soit E un espace vectoriel et $(x_1, \dots, x_n)$ une famille génératrice de E. Posons

$$I = \{i \in \{1, \dots n\};\ x_i \notin \mathrm{Vect}\{x_j;\ j < i\}\}.$$

1. Soit $(\lambda_j) \in K^n$ une famille non nulle telle que $\sum_{j=1}^{n} \lambda_j x_j = 0$. Soit i_0 le plus grand élément de $\{j \in \{1, \dots, n\};\ \lambda_j \neq 0\}$. Démontrer que $i_0 \notin I$.
2. Démontrer que, pour tout $j \in \{1, \dots, n\}$ on a $x_j \in \mathrm{Vect}\{x_i;\ i \in \{1, \dots j\} \cap I\}$.
3. Démontrer que $(x_i)_{i \in I}$ est une base de E.

5.2 Exercice. Soient E un espace vectoriel de dimension finie et F, G des sous-espaces vectoriels de E. Démontrer qu'il existe un automorphisme f de E tel que $f(F) = G$ si et seulement si F et G ont même dimension.

5.3 Exercice. Soit F un sous-espaces vectoriel d'un espace vectoriel E.

1. Soient G_1 et G_2 deux sous-espaces de E. On suppose qu'ils sont tous les deux supplémentaires de F. Démontrer que G_1 et G_2 sont isomorphes.
   On dit que F est de codimension finie dans E s'il admet un supplémentaire de dimension finie. On appelle alors codimension de F et l'on note $\mathrm{codim}\, F$ la dimension d'un tel supplémentaire.
2. On suppose que E est de dimension finie. Démontrer que la dimension et la codimension de F sont finies et que l'on a $\mathrm{codim}\, F = \dim E - \dim F$.
3. Démontrer que si F est un sous-espace de E de codimension finie, tout sous-espace de E contenant F est de codimension finie inférieure ou égale à celle de F.
4. Soit f une application linéaire de E dans un espace vectoriel G. Démontrer que f est de rang fini si et seulement si $\ker f$ est de codimension finie et que, dans ce cas on a $\mathrm{codim}(\ker f) = \mathrm{rg} f$.

5.4 Exercice. Soient E, F, G des espaces vectoriels $f : E \to F$ et $g : F \to G$ des applications linéaires.

1. On suppose que f est de rang fini. Démontrer que $g \circ f$ est de rang fini et que l'on a $\mathrm{rg}(g \circ f) \leqslant \mathrm{rg} f$. Plus précisément, démontrer que l'on a $\mathrm{rg}(g \circ f) = \mathrm{rg} f - \dim(\ker g \cap \mathrm{im}\, f)$.
2. On suppose que g est de rang fini. Démontrer que $g \circ f$ est de rang fini et que l'on a $\mathrm{rg}(g \circ f) \leqslant \mathrm{rg} g$. Plus précisément, démontrer que l'on a $\mathrm{rg}(g \circ f) = \mathrm{rg} g - \mathrm{codim}(\ker g + \mathrm{im}\, f)$.

6 Matrices et bases

Dorénavant, (presque toutes) nos bases seront indicées par $\{1, \ldots, n\}$.

6.1 Matrice d'une application linéaire

6.1.1 Matrice d'une application linéaire entre espaces vectoriels munis de bases

Soient E, F des espaces vectoriels de dimension finie. Choisissons des bases $B = (e_1, \ldots, e_n)$ et $B' = (e'_1, \ldots, e'_m)$ de E et F respectivement. Par définition (*cf.* déf. 4.24), elles donnent lieu à des isomorphismes $p_B : K^n \to E$ et $p_{B'} : K^m \to F$. Soit $f : E \to F$ une application linéaire. On lui associe l'application linéaire $p_{B'}^{-1} \circ f \circ p_B : K^n \to K^m$. Notons $A = (a_{i,j})_{1 \leqslant i \leqslant m;\ 1 \leqslant j \leqslant n}$ la matrice de cette application. On a donc $f(e_j) = \sum_{i=1}^{m} a_{i,j} e'_i$.

6.1 Définition. La matrice ainsi définie s'appelle *matrice de f dans les bases B et B'*. On la note $M_{B',B}(f)$.

La matrice d'une composée est le produit des matrices :

6.2 Proposition. *Soient E, F, G des K espaces vectoriels de dimension finie munis de bases B, B', B'' respectivement. Soient $f : E \to F$ et $g : F \to G$ des applications linéaires. On a $M_{B'',B}(g \circ f) = M_{B'',B'}(g) M_{B',B}(f)$.*

6.1.2 Changements de base, matrices de passage

Donnons-nous à présent deux bases B et B_1 de E. Il est commode (usuel) d'appeler B l'« ancienne base » et B_1 la « nouvelle base ».

6.3 Définition. On appelle *matrice de passage* de la base B à la base B_1 la matrice dont les vecteurs-colonnes sont les coefficients des vecteurs de B_1 (la nouvelle base) dans la base B (l'ancienne base).

La matrice de passage de B à B_1 est $M_{B,B_1}(\mathrm{id}_E)$.

6.4 Remarques. a) Une matrice de passage est inversible : la matrice de passage de B_1 à B est l'inverse de la matrice de passage de B à B_1.

b) Toute matrice inversible est la matrice de passage de B à une autre base : une matrice inversible P représente un automorphisme f de E dans la base B (*i.e.* $P = M_{B,B}(f)$). C'est aussi la matrice de passage de B à la base $f(B)$.

6.5 Formule de changement de base. *Soient E, F des K-espaces vectoriels de dimension finie. Donnons nous des bases B et B_1 de E et des bases B' et B'_1 de F. Notons P la matrice de passage de B à B_1 et Q la matrice de passage de B' à B'_1. Soit $f : E \to F$ une application linéaire. Les matrices M et M_1 de f dans les bases B, B' et B_1, B'_1 sont reliées par la formule $M_1 = Q^{-1} M P$.*

En effet, on écrit $M = M_{B',B}(f)$, $M_1 = M_{B'_1,B_1}(f)$, $P = M_{B,B_1}(\mathrm{id}_E)$ et $Q = M_{B',B'_1}(\mathrm{id}_F)$. Par la prop. 6.2, on a

$$QM_1 = M_{B',B'_1}(\mathrm{id}_F) M_{B'_1,B_1}(f) = M_{B',B_1}(f) = M_{B',B}(f) M_{B,B_1}(\mathrm{id}_E) = MP.$$

6.2 Matrices équivalentes, matrices semblables

6.2.1 Matrices équivalentes

6.6 Définition. Soient $m, n \in \mathbb{N}$. Deux matrices $A, B \in \mathcal{M}_{m,n}(K)$ sont dites *équivalentes* s'il existe des matrices inversibles $P \in \mathcal{M}_m(K)$ et $Q \in \mathcal{M}_n(K)$ telles que $B = P^{-1}AQ$.

Deux matrices sont donc équivalentes si elles représentent la même application linéaire (d'un espace vectoriel dans un autre) dans des bases différentes.

6.7 Proposition. *Deux matrices équivalentes ont le même rang.*

6.8 Proposition. *Soient E et F des espaces vectoriels et $f : E \to F$ une application linéaire. Notons r son rang. Il existe des bases B de E et B' de F telles que la matrice de f soit $(a_{i,j})$ avec $a_{i,j} = 1$ si $i = j \leqslant r$ et $a_{i,j} = 0$ sinon.*

En d'autres termes, $M_{B',B}(f)$ admet la décomposition par blocs $\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$.

Il suffit de prendre
- pour B une base $(e_1, \ldots, e_n)$ où $(e_1, \ldots, e_r)$ est la base d'un supplémentaire de $\ker f$ et $(e_{r+1}, \ldots, e_n)$ une base de $\ker f$;
- pour $j \leqslant r$ on pose $e'_j = f(e_j)$; d'après la proposition 4.18, la famille $(e'_1, \ldots, e'_r)$ est une base de $\operatorname{im} f$; on la complète en une base $B' = (e'_1, \ldots, e'_m)$ de F.

6.9 Théorème. *Deux matrices d'ordre m, n sont équivalentes si et seulement si elles ont le même rang.*

6.2.2 Transposée d'une matrice

6.10 Définition. Si $A = (a_{i,j})_{(i,j)\in I\times J}$ est une matrice de type $I \times J$, on appelle *transposée* de A la matrice ${}^tA = (b_{j,i})_{(j,i)\in J\times I}$ de type $J \times I$ donnée par $b_{j,i} = a_{i,j}$ pour tout $(j, i) \in J \times I$.

Regroupons dans la proposition suivante les principales propriétés de la transposée des matrices :

6.11 Proposition. *a) Soient I, J, L des ensembles finis. On a ${}^t1_I = 1_I$. Pour tout $A \in \mathcal{M}_{I,J}(K)$ et tout $B \in \mathcal{M}_{J,L}(K)$, on a ${}^t(AB) = {}^tB\,{}^tA$.*

b) La transposée d'une matrice carrée inversible est inversible.

c) Le rang d'une matrice est égal au rang de sa transposée.

6.2.3 Matrices extraites

6.12 Définition. Soient I, J des ensembles finis et $A = (a_{i,j})_{(i,j)\in I\times J} \in \mathcal{M}_{I,J}(K)$. Soient $I' \subset I$ et $J' \subset J$. La matrice $(a_{i,j})_{(i,j)\in I'\times J'} \in \mathcal{M}_{I',J'}(K)$ s'appelle une *matrice extraite* de A.

6.13 Exercice. Décrire l'application linéaire de $K^{J'}$ dans $K^{I'}$ associée à une matrice extraite. On sera amené à construire des applications linéaires $K^{J'} \to K^J$ et $K^I \to K^{I'}$.

6.14 Proposition. *a) Le rang d'une matrice extraite de A est inférieur ou égal à celui de A.*

b) Si le rang de A est $\geqslant r$, il existe une matrice carrée d'ordre r inversible extraite de A.

En d'autres termes, le rang de A est l'ordre de la plus grande matrice carrée inversible extraite de A.

Le (a) résulte... de l'exercice.

Pour (b), on sait que le rang de A est le rang de ses vecteurs-colonnes. Dans le système $(x_j)_{j\in J}$ qui engendre l'image de A qui est de dimension $\geqslant r$, on choisit une base $(x_j)_{j\in J_1}$ de cette image ; puisque J_1 a $\mathrm{rg}A$ éléments, on peut choisir une sous-famille $(x_j)_{j\in J'}$ a r éléments, qui sera donc libre. La matrice extraite B, de type $I \times J'$, est de rang r, donc sa transposée aussi. Raisonnant de même avec ${}^tB$, on trouve une partie $I' \subset I$ dans les colonnes de B à r éléments telle que la matrice extraite (carrée d'ordre r) soit de rang r.

6.2.4 Matrices d'endomorphismes

Dans le cas d'un endomorphisme f d'un espace vectoriel E, on va en général choisir la même base au départ et à l'arrivée, *i.e.* considérer la matrice $M_{B,B}(f)$ où B est une base de E.

6.15 Formule de changement de base. *Soient E un K-espace vectoriel de dimension finie et f un endomorphisme de E. Donnons nous des bases B et B_1 de E. Notons P la matrice de passage de B à B_1. Les matrices M et M_1 de f dans les bases B et B_1 sont reliées par la formule $M_1 = P^{-1}MP$.*

6.16 Définition. Soit $n \in \mathbb{N}$. Deux matrices $A, B \in \mathcal{M}_n(K)$ sont dites *semblables* s'il existe une matrice inversible $P \in \mathcal{M}_n(K)$ telle que $B = P^{-1}AP$.

Deux matrices sont donc semblables si elles représentent le même *endomorphisme* dans des bases différentes.

6.17 Définition. Soit $A = (a_{i,j}) \in \mathcal{M}_n(K)$ une matrice carrée. On appelle *trace* de A le nombre $\mathrm{Tr}(A) = \sum_{i=1}^{n} a_{i,i}$.

6.18 Proposition. *Soient $A \in \mathcal{M}_{m,n}(K)$ et $B \in \mathcal{M}_{n,m}(K)$ deux matrices. On a $\mathrm{Tr}\, AB = \mathrm{Tr}\, BA$.*

6.19 Corollaire. *Deux matrices semblables ont même trace.*

Cela nous permet de définir la trace d'un endomorphisme : c'est la trace de sa matrice dans n'importe quelle base.

Nous reviendrons plus longuement sur les endomorphismes au paragraphe 8.

6.3 Dualité, base duale

6.3.1 Formes linéaires ; dual d'un espace vectoriel

Soit E un K-espace vectoriel.

6.20 Définition. On appelle *hyperplan* de E un sous-espace vectoriel de codimension 1 de E. On appelle *forme linéaire* sur E une application linéaire de E à valeurs dans K. On appelle *dual* de E et l'on note E^* l'espace vectoriel $L(E, K)$ formé des formes linéaires sur E.

6.21 Proposition. *Le noyau d'une forme linéaire non nulle est un hyperplan. Inversement, tout hyperplan est noyau d'une forme linéaire.*

Soit H un hyperplan de E. Il existe donc une forme linéaire f, unique à un multiple scalaire près dont c'est le noyau. Soit $x \in E$. On a donc $x \in H \iff f(x) = 0$. On dit donc que l'équation $f(x) = 0$ est une *équation* de H.

Base duale

6.22 Proposition. *Soit E un espace vectoriel de dimension finie et soit $(e_1, \dots, e_n)$ une base de E.*

a) Pour tout $k \in \{1, \dots, n\}$, il existe une unique forme linéaire $e_k^* \in E^*$ telle que $e_k^*(e_j) = 0$ si $j \neq k$ et $e_k^*(e_k) = 1$.

b) La famille $(e_1^*, \dots, e_n^*)$ est une base de E^*.

En particulier, $\dim E^* = \dim E$.

6.23 Définition. La base $(e_1^*, \dots, e_n^*)$ de E^* s'appelle la *base duale* de $(e_1, \dots, e_n)$.

6.24 Proposition. *Soient E un espace vectoriel de dimension finie, F un sous-espace vectoriel et $x \in E \setminus F$. Il existe $f \in E^*$ telle que $F \subset \ker f$ et $x \notin \ker f$.*

Soit $(e_1, \dots e_k)$ une base de F. Posons $e_{k+1} = x$; puisque $x \notin F$, la famille $(e_1, \dots, e_{k+1})$ est libre. Complétons-la en une base $(e_1, \dots e_n)$. Il suffit alors de poser $f = e_{k+1}^*$.

6.25 Exemple. Soit $n \in \mathbb{N}$. Notons E_n le sous-espace vectoriel de $K[X]$ formé des polynômes de degré $< n$. C'est un espace vectoriel de dimension n. Soient $(x_1, \dots, x_n)$ des points distincts de K. Pour $k \in \{1, \dots, n\}$, notons f_k la forme linéaire $P \mapsto P(x_k)$. Considérons l'application linéaire $\varphi : E_n \to K^n$ définie par $\varphi(P) = (f_1(P), \dots, f_n(P))$

Si $P \in \ker \varphi$, alors P admet les racines $x_1 \dots, x_n$, donc P est le polynôme nul. L'application φ est donc injective ; puisque $\dim E_n = \dim K^n = n$, elle est bijective.

Pour $k \in \{1, \dots, n\}$, notons Q_k le polynôme $Q_k = \prod_{1 \leqslant j \leqslant n;\ j \neq k} (X - x_j)$ et posons $P_k = \dfrac{1}{Q_k(x_k)} Q_k$. La famille $(P_1, \dots, P_n)$ est une base de E_n. Sa base duale est $(f_1, \dots, f_n)$.

En particulier, si $\lambda_1, \dots, \lambda_n$ sont des éléments de K, il existe un et un seul polynôme de degré $< n$ tel que $P(x_k) = \lambda_k$ pour tout k : ce polynôme est $\sum_{k=1}^{n} \lambda_k P_k$ (*formule d'interpolation de Lagrange*).

Transposée d'une application linéaire

6.26 Définition. Soient E, F des espaces vectoriels et $\varphi : E \to F$ une application linéaire. On appelle *transposée* de φ l'application linéaire ${}^t\varphi : f \mapsto f \circ \varphi$ de F^* dans E^*.

Supposons que E et F soient des espaces vectoriels de dimension finie. Soient B et B' des bases de E et F respectivement. La matrice $M_{B^*,(B')^*}({}^t\varphi)$ de ${}^t\varphi$ de la base duale de B' dans la base duale de B est la transposée de la matrice $M_{B',B}(\varphi)$ de φ dans de la base B dans la base B'. En particulier, on a $\mathrm{rg}\, {}^t\varphi = \mathrm{rg}\, \varphi$.

6.3.2 Espaces vectoriels en dualité

Le point de vue que nous adoptons pour ce qui concerne l'orthogonalité est celui de deux espaces « en dualité ». Cela nous permet

- de traiter de façon symétrique un espace et son dual ;
- de traiter en même temps le cas d'un espace muni d'une forme bilinéaire symétrique non dégénérée comme un espace euclidien.

Soient E, F des espaces vectoriels. Il revient au même de se donner

∗ une forme bilinéaire $b : E \times F \to K$;

∗ une application linéaire $\varphi : E \to F^*$;

$*$ une application linéaire $\psi : F \to E^*$.
Ces applications sont reliées par la formule $b(x,y) = \varphi(x)(y) = \psi(y)(x)$.

Lorsqu'on s'est donné de telles applications, on dit que E et F sont des espaces vectoriels *en dualité*.

Cette généralité permet de recouvrir deux cas particuliers fondamentaux :
- $F = E^*$ et $\psi = \mathrm{id}_F$;
- $F = E$ et b est une forme bilinéaire symétrique (*i.e.* $\varphi = \psi$) voire antisymétrique (*i.e.* $\varphi = -\psi$).

Choisissons des bases $B = (e_1, \ldots, e_m)$ et $B' = (e'_1, \ldots, e'_n)$ de E et F respectivement. On appelle *matrice de la forme bilinéaire* b la matrice $A = (a_{i,j}) \in \mathcal{M}_{m,n}(K)$ définie par $a_{i,j} = b(e_i, e'_j)$. C'est la matrice $M_{B^*,B'}(\psi)$ de ψ de la base B' dans la base duale de B et la transposée de la matrice $M_{(B')^*,B}(\varphi)$ de φ dans de la base B dans la base duale de B'. On en déduit :

6.27 Proposition. *Soient E, F des espaces vectoriels de* dimension finie *en dualité. On a* $\mathrm{rg}\varphi = \mathrm{rg}\psi$. *En particulier, φ est un isomorphisme si et seulement si ψ est un isomorphisme.*

6.28 Exemple. Supposons que $F = E^*$ et $\psi = \mathrm{id}_F$. Alors $\varphi : E \to (E^*)^*$ est une application linéaire appelée *homomorphisme canonique* de E dans son *bidual* que l'on note E^{**} caractérisé par l'égalité $\varphi(x)(f) = f(x)$ pour $x \in E$ et $f \in E^*$. Lorsque E est de dimension finie, φ est un isomorphisme appelé *isomorphisme canonique*.

6.3.3 Orthogonalité

6.29 Définition. Soient E et F des espaces vectoriels en dualité. Des vecteurs $x \in E$ et $y \in F$ sont dits *orthogonaux* si $b(x,y) = 0$. L'*orthogonal* d'une partie A de E (*resp.* d'une partie B de F) est l'ensemble $A^\perp = \{y \in F;\ \forall x \in A;\ b(x,y) = 0\}$ (*resp.* l'ensemble $B^o = \{x \in E;\ \forall y \in B;\ b(x,y) = 0\}$).

Regroupons dans le prochain énoncé les principales propriétés de l'orthogonalité.

6.30 Proposition. *Soient E et F des espaces vectoriels en dualité.*

a) Pour toute partie $A \subset E$ (resp. $B \subset F$), l'ensemble $A^\perp$ (resp. B^o) est un sous-espace vectoriel de E (resp. de F). Si $A_1 \subset A_2$, on a $A_2^\perp \subset A_1^\perp$ (resp. si $B_1 \subset B_2$, on a $B_2^o \subset B_1^o$).

On suppose de plus que E et F sont de dimension finie, et que φ et ψ sont des isomorphismes.

b) Pour toute partie $A \subset E$ (resp. $B \subset F$) l'espace $(A^\perp)^o$ (resp. $(A^o)^\perp$) est le sous-espace de E (resp. F) engendré par A (resp. B).

c) L'application $E_1 \mapsto E_1^\perp$ est une bijection décroissante de l'ensemble des sous-espaces vectoriels de E sur l'ensemble des sous-espaces vectoriels de F ; la bijection réciproque est $F_1 \mapsto F_1^o$.

d) Pour tout sous-espace vectoriel E_1 de E (resp. F_1 de F), on a $\dim E_1^\perp = \dim E - \dim E_1$ *(resp.* $\dim F_1^o = \dim F - \dim F_1$*).*

e) Soient E_1, E_2 (resp. F_1, F_2) des sous-espaces vectoriels de E (resp. F). On a $(E_1+E_2)^\perp = E_1^\perp \cap E_2^\perp$ et $(E_1 \cap E_2)^\perp = E_1^\perp + E_2^\perp$ (resp. $(F_1 + F_2)^o = F_1^o \cap F_2^o$ et $(F_1 \cap F_2)^o = F_1^o + F_2^o$).

Notons que les assertions « *resp.* » se déduisent des autres en échangeant les rôles de E et F et en remplaçant b par l'application $b' : (y,x) \mapsto b(x,y)$ de $F \times E$ dans K.

6.4 Exercices

6.1 Exercice. Soit A une matrice décomposée par blocs $A = \begin{pmatrix} A_1 & A_2 \\ A_3 & A_4 \end{pmatrix}$. On suppose que A_1 est une matrice carrée d'ordre r inversible. Démontrer que $\mathrm{rg}A = r \iff A_4 = A_3 A_1^{-1} A_2$.

6.2 Exercice. Soit E un espace vectoriel de dimension finie B, B' deux bases de E. Notons P la matrice de passage de B à B'. Quelle est la matrice de passage de la base duale B^* de B à la base duale $(B')^*$ de B'.

6.3 Exercice. Soit E un K espace vectoriel.

1. Démontrer que deux formes linéaires sur E dont les noyaux sont égaux sont proportionnelles.
2. Soient $f_1, \dots, f_k$ des formes linéaires sur E et $f \in E^*$. Démontrer que $f \in \mathrm{Vect}\{f_1, \dots, f_k\}$ si et seulement si $\bigcap_{j=1}^{k} \ker f_j \subset \ker f$.

6.4 Exercice. Soit E un espace vectoriel de dimension finie. Soit $(f_1, \dots, f_n)$ une famille d'éléments de E^*. Notons $\varphi : E \to K^n$ l'application $x \mapsto (f_1(x), \dots, f_n(x))$.

1. On suppose que la famille $(f_1, \dots, f_n)$ est une base de E^*.
   a) Démontrer que φ est injective. En déduire qu'elle est bijective.
   b) Démontrer qu'il existe une base B de E telle que $\varphi(B)$ soit la base canonique de K^n.
2. En déduire que toute base de E^* est duale d'une base de E.
3. Démontrer que l'on a les équivalences suivantes :
   φ est injective si et seulement si la famille $(f_1, \dots, f_n)$ est génératrice ;
   φ est surjective si et seulement si la famille $(f_1, \dots, f_n)$ est libre.

6.5 Exercice. Soient E et F des espaces vectoriels de dimension finie et f une application linéaire de E dans F.

1. Démontrer que ${}^t f$ est injective si et seulement si f est surjective et ${}^t f$ est surjective si et seulement si f est injective.
2. Démontrer que $\ker {}^t f = (\mathrm{im}\, f)^\perp$ et $\mathrm{im}\, {}^t f = (\ker f)^\perp$.

6.6 Exercice. Soient E et F deux espaces vectoriels de dimension finie en dualité. Démontrer que $\ker \varphi = F^o$ et $\ker \psi = E^\perp$.

Notons $b : \mathcal{M}_n(K) \times \mathcal{M}_n(K) \to K$ l'application $(A, B) \mapsto \mathrm{Tr}(AB)$.

1. Démontrer que b est une forme bilinéaire symétrique non dégénérée (ces termes sont définis page 76).
2. On suppose que $n \geqslant 2$. Démontrer que tout hyperplan de $\mathcal{M}_n(K)$ contient une matrice inversible.
3. On suppose que $K = \mathbb{R}$. Quelle est la signature de b ?

6.7 Exercice. Soient a et b deux points distincts de K. Sur le K espace vectoriel E des polynômes de degré $\leqslant 3$, on considère les formes linéaires $f_1 : P \mapsto P(a)$, $f_2 : P \mapsto P'(a)$, $f_3 : P \mapsto P(b)$, $f_4 : P \mapsto P'(b)$.

1. Calculer $\{f_1, f_2, f_3, f_4\}^o$.
2. Démontrer que (f_1, f_2, f_3, f_4) est une base de E^*.
3. Quelle est la base de E dont (f_1, f_2, f_3, f_4) est la base duale ?

6.8 Exercice. On se propose de donner deux démonstration du
Lemme de Schur. *Un endomorphisme u d'un espace vectoriel E de dimension finie qui laisse stable tout hyperplan est une homothétie.*

1. Rappel : Démontrer qu'un endomorphisme qui laisse invariante toute droite vectorielle est une homothétie.
2. *Première méthode.*

a) Démontrer que la transposée de u laisse fixe toute droite - donc c'est est une homothétie.

b) En déduire que u est une homothétie.

3. *Deuxième méthode.* Démontrer que u laisse stable toute droite - donc c'est une homothétie.

6.9 Exercice. [Dual d'un espace vectoriel complexe] Remarquons que tout espace vectoriel complexe est naturellement un espace vectoriel réel. Soit E un espace vectoriel complexe. Notons $E^*_{\mathbb{C}}$ son dual et $E^*_{\mathbb{R}}$ le dual de E considéré comme espace vectoriel réel. Pour $\ell \in E^*_{\mathbb{C}}$, notons $\mathrm{Re}(\ell)$ l'application $x \mapsto \mathrm{Re}(\ell(x))$.

1. Démontrer que $\ell \mapsto \mathrm{Re}(\ell)$ est une bijection de $E^*_{\mathbb{C}}$ sur $E^*_{\mathbb{R}}$.
2. En particulier, $E^*_{\mathbb{R}}$ s'identifie à l'espace vectoriel complexe $E^*_{\mathbb{C}}$. Décrire directement la structure d'espace vectoriel complexe sur $E^*_{\mathbb{R}}$, *i.e.* la multiplication d'un élément de $E_{\mathbb{R}}$ par un nombre complexe.

6.10 Exercice. Soient E un espace vectoriel de dimension finie sur un corps K et f un endomorphisme de E.

1. Démontrer qu'un sous-espace F de E est stable par f si et seulement son orthogonal $F^\perp$ est stable par ${}^t f$.
2. Démontrer que f possède une valeur propre (dans K) si et seulement s'il existe un hyperplan de E stable par f.

On en déduit par récurrence sur $\dim E$ que si le polynôme caractéristique de f est scindé, alors f est trigonalisable - *cf.* fin de la démonstration du théorème 8.7.

6.11 Exercice. Soient L un corps commutatif, $K \subset L$ in sous-corps de L.

1. Démontrer que toute matrice $A \in M_{m,n}(K)$ a même rang considérée comme matrice à coefficients dans K ou dans L. dans K et dans L.
2. En déduire qu'un système de vecteurs dans K^n libre sur K implique est libre sur L.
3. Soit $M \in \mathcal{M}_n(K)$. Démontrer que le polynôme minimal de M est le même sur K et sur L

6.12 Exercice. 1. Soit $A \in M_k(\mathbb{Q})$. Pour $K = \mathbb{Q}, \mathbb{R}$ ou $\mathbb{C}$, notons $f_K : K^k \to K^k$ l'application linéaire de matrice A. Démontrer que $\ker f_{\mathbb{R}} = \overline{\ker f_{\mathbb{Q}}}$ et $\ker f_{\mathbb{C}} = \{x+iy;\ x, y \in \ker f_{\mathbb{R}}\}$. Démontrer que $\mathrm{im}\, f_{\mathbb{R}} = \overline{\mathrm{im}\, f_{\mathbb{Q}}}$ et $\mathrm{im}\, f_{\mathbb{C}} = \{x + iy;\ x, y \in \mathrm{im}\, f_{\mathbb{R}}\}$.

2. Soient $A, B \in M_n(\mathbb{Q})$. Posons $E_{\mathbb{Q}} = \{M \in M_n(\mathbb{Q});\ AM = MB\}$, $E_{\mathbb{R}} = \{M \in M_n(\mathbb{R});\ AM = MB\}$ et $E_{\mathbb{C}} = \{M \in M_n(\mathbb{C});\ AM = MB\}$.

a) Démonter que $E_{\mathbb{Q}}$ est dense dans $E_{\mathbb{R}}$ et que $E_{\mathbb{C}} = \{M + iN;\ M, N \in E_{\mathbb{R}}\}$.

b) En déduire que A et B sont semblables sur $\mathbb{Q}$ si et seulement si elles le sont sur $\mathbb{C}$.

7 Systèmes d'équations linéaires, nants

7.1 Systèmes d'équations linéaires

Dans un sens, l'algèbre linéaire consiste à expliquer la structure des systèmes linéaires. Inversement, les questions d'algèbre linéaire se résolvent à l'aide de systèmes.

Un système linéaire de m équations aux inconnues $(x_1, \dots, x_n)$ est de la forme

$$\begin{cases} a_{1,1}x_1 & + & a_{1,2}x_2 & + & \dots & + & a_{1,n}x_n & = & b_1 \\ a_{2,1}x_1 & + & a_{2,2}x_2 & + & \dots & + & a_{2,n}x_n & = & b_2 \\ \vdots & & \vdots & & \ddots & & \vdots & & \vdots \\ a_{m,1}x_1 & + & a_{m,2}x_2 & + & \dots & + & a_{m,n}x_n & = & b_m \end{cases}$$

On dit qu'on a résolu ce système si on a décrit l'ensemble des n-uplets $(x_1, \dots, x_n)$ qui satisfont les m égalités ci-dessus.

Ce système est équivalent à l'équation matricielle $AX = B$ où A est la matrice $A = (a_{i,j}) \in \mathcal{M}_{m,n}(K)$, X est la matrice colonne inconnue $(x_j) \in K^n$ et B est la matrice colonne $(b_i) \in K^m$. Le rang de la matrice A s'appelle aussi le *rang du système.*

7.1 Définition. Un système d'équations linéaires est dit *de Cramer* s'il s'écrit de façon matricielle $AX = B$ où A est une matrice inversible. Un système de Cramer a une et une seule solution : $X = A^{-1}B$.

Nous verrons plus loin (au paragraphe 7.3) comment résoudre *en pratique* un système linéaire.

Résolution théorique. Notons r le rang de A. Par la proposition 6.14, on peut choisir $I \subset \{1, \dots, m\}$ et $J \subset \{1, \dots, n\}$ tels que la matrice extraite $(a_{i,j})_{(i,j)\in I\times J}$ soit carrée d'ordre r et inversible. Les x_j pour $j \in J$ s'appellent alors les *inconnues principales*, et les équations d'indice $i \in I$ s'appellent les *équations principales.*

Quitte à échanger l'ordre des équations et des inconnues, nous allons supposer que $I = J = \{1, \dots, r\}$. Décomposons A par blocs $A = \begin{pmatrix} A_1 & A_2 \\ A_3 & A_4 \end{pmatrix}$ (où A_1 est inversible d'ordre r).

7.2 Proposition. *Soit A une matrice d'ordre (m, n) de rang r ; supposons qu'elle admet une décomposition par blocs $A = \begin{pmatrix} A_1 & A_2 \\ A_3 & A_4 \end{pmatrix}$ où A_1 est inversible d'ordre r. Soit $B = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$. L'équation $AX = B$ en les inconnues $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ admet des solutions si et seulement si $\begin{pmatrix} b_{r+1} \\ \vdots \\ b_m \end{pmatrix} = A_3A_1^{-1} \begin{pmatrix} b_1 \\ \vdots \\ b_r \end{pmatrix}$; dans ce cas l'ensemble des solutions est $\{X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} ; \begin{pmatrix} x_1 \\ \vdots \\ x_r \end{pmatrix} = A_1^{-1} \left(\begin{pmatrix} b_1 \\ \vdots \\ b_r \end{pmatrix} - A_2 \begin{pmatrix} x_{r+1} \\ \vdots \\ x_n \end{pmatrix} \right) ; x_{r+1}, \dots, x_n \in K\}$.*

La matrice carrée $U = \begin{pmatrix} A_1 & 0 \\ A_3 & I_{m-r} \end{pmatrix}$ d'ordre m est inversible d'inverse $U^{-1} = \begin{pmatrix} A_1^{-1} & 0 \\ -A_3A_1^{-1} & I_{m-r} \end{pmatrix}$. Le système est donc équivalent à $U^{-1}AX = U^{-1}B$. Or $U^{-1}A$ est de la forme $\begin{pmatrix} I_r & A_1^{-1}A_2 \\ 0 & C_4 \end{pmatrix}$, et puisque $\mathrm{rg}U^{-1}A = \mathrm{rg}A = r$, il vient $C_4 = 0$. Le système devient $\begin{pmatrix} I_r & A_1^{-1}A_2 \\ 0 & 0 \end{pmatrix} X = U^{-1}B$, soit

$$0 = \begin{pmatrix} -A_3A_1^{-1} & I_{m-r} \end{pmatrix} B \quad \text{et} \quad \begin{pmatrix} I_r & A_1^{-1}A_2 \end{pmatrix} X = \begin{pmatrix} A_1^{-1} & 0 \end{pmatrix} B.$$

Écrivons enfin $X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$ et $B = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$ où B_1 et X_1 sont des matrices-colonne à r lignes. Le système devient :

$$B_2 = A_3 A_1^{-1} B_1 \quad \text{et} \quad X_1 + A_1^{-1} A_2 X_2 = A_1^{-1} B_1.$$

Quelques applications

1. Soient E, F des espaces vectoriels de dimension finie, f une application linéaire et A sa matrice dans des bases données. La résolution du système $AX = B$ donne :
   - Des équations de l'image, *i.e.* l'ensemble des B pour lesquels ce système admet des solutions (on dit qu'il est *compatible*) : pour $i > r$, (ou plus généralement pour $i \in \{1, \dots, n\} \setminus I$ où I est l'ensemble des équations principales) on obtient une équation du type $b_i = \sum \alpha_{i,k} b_k$ la somme étant prise de 1 à r (sur J dans le cas général). Les $\alpha_{i,k}$ sont les coefficients de la matrice $A_3 A_1^{-1}$.
   - Une base de l'image : les vecteurs-colonne d'indice $j \in \{1, \dots, r\}$ ($j \in J$ dans le cas général).
   - Pour chaque élément b de l'image, une paramétrisation (par x_j, $j > r$ - ou $j \notin J$ dans le cas général) de l'ensemble des $x \in E$ tels que $f(x) = b$.
   - En particulier, on obtient une base du noyau indexée par $j \in \{r+1, \dots, n\}$ (en considérant le système $AX = 0$, *i.e.* en prenant les b_i tous nuls) : elle est formée par les vecteurs-colonne de la matrice $\begin{pmatrix} -A_1^{-1} A_2 \\ I_{n-r} \end{pmatrix}$.
2. ***Applications « géométriques »***
   Soient E un espace vectoriel de dimension finie et F un sous-espace vectoriel (*resp.* affine) de E. Une *représentation paramétrique* de F est donnée par une application linéaire (*resp.* affine) f de $\mathbb{R}^k$ dans E d'image F. Une telle représentation paramétrique sera minimale si f est injective.

   Un *système d'équations cartésiennes* (ou *représentation cartésienne*) de F est donné par une application linéaire $g : E \to \mathbb{R}^\ell$ telle que $F = \ker g$ (*resp.* $F = g^{-1}(B)$ où B est un point de $\mathbb{R}^\ell$). Une telle représentation cartésienne sera minimale si g est surjective.

   La proposition 7.2 nous permet de passer d'une représentation à l'autre : elle donne une représentation paramétrique minimale de l'ensemble des solutions de l'équation $g(X) = B$; elle donne un système minimal d'équations cartésiennes de l'ensemble des X tels que le système $f(Y) = X$ admet des solutions.

   En particulier, si F et G sont des sous-espaces vectoriels de E donnés par des représentations cartésiennes, on a immédiatement un système d'équations cartésiennes de $F \cap G$; on peut de même facilement donner une représentation paramétrique de $F + G$ si F et G sont donnés par une représentation paramétrique. La résolution de systèmes nous permet donc de donner des équations cartésiennes et paramétriques d'une intersection et d'une somme de sous-espaces.

7.2 Déterminants

7.2.1 Formes multilinéaires alternées ; déterminant relatif à une base

7.3 Définition. Soient E, F des K-espaces vectoriels et $n \in \mathbb{N}$. Une application $D : E^n \to F$ est appelée *multilinéaire* ou *n-linéaire* si elle est linéaire par rapport à chacune des variables. Une application multilinéaire $D : E^n \to F$ est dite *alternée* si $D(x_1, \dots, x_n) = 0$ dès que deux x_i sont égaux, *i.e.* dès qu'il existe $i, j \in \{1, \dots, n\}$ avec $i \neq j$ et $x_i = x_j$. Lorsque $F = K$ on parle de *forme multilinéaire* et de *forme multilinéaire alternée.*

L'ensemble des formes n-linéaires alternées est naturellement muni d'une structure d'espace vectoriel (sous-espace vectoriel de l'espace vectoriel F^{E^n} de toutes les applications de E^n dans F).

Soit D une application n-linéaire alternée. Soient $x_1, \ldots, x_n \in E$. On a

$$0 = D(x_1 + x_2, x_1 + x_2, x_3, \ldots, x_n) = D(x_1, x_2, x_3, \ldots, x_n) + D(x_2, x_1, x_3, \ldots, x_n),$$

donc $D(x_2, x_1, x_3, \ldots, x_n) = -D(x_1, x_2, x_3, \ldots, x_n)$. Plus généralement, si i, j sont deux éléments distincts de $\{1, \ldots, n\}$, on a $D(\ldots, x_i, \ldots, x_j, \ldots) = -D(\ldots, x_j, \ldots, x_i, \ldots)$.

On en déduit que pour toute permutation $\sigma \in \mathfrak{S}_n$ on a $D(x_{\sigma(1)}, x_{\sigma(2)}, \ldots, x_{\sigma(n)}) = \varepsilon(\sigma) D(x_1, x_2, \ldots, x_n)$ où ε est la signature.

7.4 Théorème. *Soit E un K-espace vectoriel de dimension n. Soit $B = (e_1, \ldots, e_n)$ une base de E. Il existe une unique forme n-linéaire alternée $\det_B : E^n \to K$ telle que $\det_B(e_1, \ldots, e_n) = 1$.*

Notons $(e_1^*, \ldots, e_n^*)$ la base duale de $(e_1, \ldots, e_n)$.

Existence Une formule pour une telle forme D est $D(x_1, \ldots, x_n) = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) \prod_{j=1}^{n} e^*_{\sigma(j)}(x_j)$.

Unicité Par multilinéarité, pour connaître une forme n-linéaire $D : E^n \to K$, il suffit de la connaître sur les vecteurs de la base, *i.e.* de connaître les nombres $D(e_{s(1)}, \ldots, e_{s(n)})$ pour toute application $s : \{1, \ldots, n\} \to \{1, \ldots, n\}$. On aura $D(x_1, \ldots, x_n) = \sum_s D(e_{s(1)}, \ldots, e_{s(n)}) \prod_{j=1}^{n} e^*_{s(j)}(x_j)$ où la somme est prise sur l'ensemble de toutes les applications $s : \{1, \ldots, n\} \to \{1, \ldots, n\}$, puisqu'on a $x_j = \sum_{i=1}^{n} e_i^*(x_j) e_i$.

Si D est alternée, alors $D(e_{s(1)}, \ldots, e_{s(n)}) = 0$ si s n'est pas injective. De plus, pour tout $\sigma \in \mathfrak{S}_n$, on a $D(e_{\sigma(1)}, \ldots, e_{\sigma(n)}) = \varepsilon(\sigma) D(e_1, \ldots, e_n)$. En d'autres termes, l'application $D \mapsto D(e_1, \ldots, e_n)$ est injective.

L'application $\det_B$ s'appelle le *déterminant associé à la base B*.

7.5 Proposition. *Soient B, B' deux bases d'un espace vectoriel E de dimension finie. Posons $n = \dim E$. Pour tout $(x_1, \ldots, x_n) \in E^n$, on a $\det_B(x_1, \ldots, x_n) = \det_B(B') \det_{B'}(x_1, \ldots, x_n)$.*

En effet, il résulte (de la démonstration) du théorème 7.4 que l'application $D \mapsto D(B')$ est une bijection de l'espace vectoriel des applications n-linéaires alternées sur K. Les applications n-linéaires alternées $\det_B$ et $\det_B(B') \det_{B'}$ qui coïncident en B' sont égales.

7.6 Proposition. *Soit B une base d'un espace vectoriel E de dimension finie. Posons $n = \dim E$. Pour tout $(x_1, \ldots, x_n) \in E^n$, on a $\det_B(x_1, \ldots, x_n) \neq 0$ si et seulement si $(x_1, \ldots, x_n)$ est une base de E.*

Posons $B' = (x_1, \ldots, x_n)$. Si B' est une base de E, on a $1 = \det_B(B) = \det_B(B') \det_{B'}(B)$ d'après la prop. 7.5, donc $\det_B(B') \neq 0$.

Puisque $\det_B$ est n-linéaire et alternée, le nombre $\det_B(x_1, \ldots, x_n)$ ne change pas lorsque on ajoute à un x_k une combinaison linéaire des autres - sans changer les autres x_j. Si la famille $(x_1, \ldots, x_n)$ n'est pas libre, par une telle opération, on peut changer un x_k en 0, donc $\det_B(x_1, \ldots, x_n) = 0$.

7.7 Formules de Cramer. Soit $B = (e_1, \ldots, e_n)$ une base de E. Soit $v \in E$. Écrivons $v = \sum_{k=1}^{n} \lambda_k e_k$.

On a $\det_B(e_1, \ldots, e_{k-1}, v, e_{k+1}, \ldots, e_n) = \lambda_k$.

Soit maintenant $(u_1, \ldots, u_n) = B'$ une autre base de E. L'équation $x_1 u_1 + \ldots + x_n u_n = v$ en les inconnues $x_1, \ldots, x_n \in K$ admet une et une seule solution donnée par

$$x_k = \det_{B'}(u_1, \ldots, u_{k-1}, v, u_{k+1}, \ldots, u_n) = \frac{\det_B(u_1, \ldots, u_{k-1}, v, u_{k+1}, \ldots, u_n)}{\det_B(u_1, \ldots, u_n)}.$$

7.2.2 Déterminant d'un endomorphisme

7.8 Proposition. *Soient E un espace vectoriel de dimension finie et soit $f \in L(E)$ un endomorphisme de E. Il existe un unique élément de K appelé* déterminant de l'endomorphisme f *et noté* $\det f$ *tel que pour toute forme n-linéaire alternée D sur E et tout $(x_1, \ldots, x_n) \in E$ on ait $D(f(x_1), \ldots, f(x_n)) = (\det f)\, D(x_1, \ldots, x_n)$.*

Soit $B = (e_1, \ldots, e_n)$ une base de E. L'application $D_0 : (x_1, \ldots, x_n) \mapsto \det_B(f(x_1), \ldots, f(x_n))$ est une forme n-linéaire alternée sur E. On a donc $D_0 = D_0(B)\det_B$. Soit D une forme n-linéaire alternée sur E ; il existe $\lambda \in K$ tel que $D = \lambda\det_B$; notons D_f l'application $D_f : (x_1, \ldots, x_n) \mapsto D(f(x_1), \ldots, f(x_n)) = \lambda\det_B(f(x_1), \ldots, f(x_n))$. Il vient $D_f = \lambda D_0 = \lambda D_0(B)\det_B = D_0(B)D$.
Il suffit donc de poser $\det f = D_0(B)$.

Si $B = (e_1, \ldots, e_n)$ est une base de E, on a $\det f = \det_B(f(e_1), \ldots, f(e_n))$.

Donnons les principales propriétés du déterminant des endomorphismes.

7.9 Théorème. *Soient E un espace vectoriel de dimension finie et $f, g \in L(E)$.*

a) On a $\det f \neq 0$ si et seulement si f est inversible.

b) On a $\det(g \circ f) = \det f \det g$.

a) Soit $B = (e_1, \ldots, e_n)$ une base de E. On a $\det f = \det_B(f(e_1), \ldots, f(e_n))$; donc $\det f \neq 0$ si et seulement si $(f(e_1), \ldots, f(e_n))$ est une base de E, *i.e.* si et seulement si f est inversible.

b) Si D est une forme n-linéaire alternée sur E et $h \in L(E)$, notons D_h la forme n-linéaire alternée $D_h : (x_1, \ldots, x_n) \mapsto D(h(x_1), \ldots, h(x_n))$. Par la proposition ci-dessus, on a $D_h = (\det h)D$. On trouve $\det(g \circ f)\, D = D_{g\circ f} = (D_g)_f = (\det f)\, D_g = (\det f)(\det g)\, D$. Il suffit de choisir D non nulle pour conclure.

Il résulte de ce théorème que $f \mapsto \det f$ est un homomorphisme de groupes de $GL(E)$ dans K^*. Son noyau $\{f \in L(E);\ \det f = 1\}$ s'appelle le *groupe spécial linéaire* de E et se note $SL(E)$.

7.2.3 Déterminant d'une matrice carrée

Soit $n \in \mathbb{N}$. L'espace vectoriel des vecteurs-colonnes $\mathcal{M}_{n,1}(K)$ est naturellement muni d'une base B dite « canonique » formée des vecteurs-colonnes e_j ayant un 1 dans la $j^{\text{ème}}$ ligne et toutes les autres lignes nulles.

7.10 Définition. Le *déterminant d'une matrice* carrée $A \in \mathcal{M}_n(K)$ est le déterminant de ses vecteurs-colonnes relativement à la base canonique. C'est le déterminant de l'endomorphisme $X \mapsto AX$ de $\mathcal{M}_{n,1}(K)$ défini par A.

De la formule donnée dans le théorème 7.4, il résulte que si $A = (a_{i,j})$, on a $\det(A) = \displaystyle\sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) \prod_{j=1}^{n} a_{\sigma(j),j}$.

7.11 Remarque. Si K est juste un anneau *commutatif*, cette formule garde un sens et permet de définir le déterminant d'une matrice à coefficients dans K.

Si $\varphi : K_1 \to K_2$ est un homomorphisme d'anneaux et $A = (a_{i,j}) \in \mathcal{M}_n(K_1)$ est une matrice à coefficients dans K_1, notons $\varphi(A) \in \mathcal{M}_n(K_2)$ la matrice $(\varphi(a_{i,j}))$. On a $\det(\varphi(A)) = \varphi(\det(A))$.

Si P est la matrice de passage d'une base B à une base B', on a $\det(P) = \det_B(B')$.

Donnons les principales propriétés du déterminant des matrices carrées.

7.12 Théorème. *Soient $P, Q \in \mathcal{M}_n(K)$.*

a) On a $\det(P) \neq 0$ si et seulement si P est inversible.

b) On a $\det(PQ) = \det P \det Q$.

c) On a $\det({}^tP) = \det(P)$.

a) et b) résultent du théorème 7.9.

Ecrivons $P = (a_{i,j})$. On a $\det({}^tP) = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) \prod_{j=1}^{n} a_{j,\sigma(j)}$.

Puisque le produit dans K est associatif et commutatif, pour $\sigma \in \mathfrak{S}_n$ et $(\lambda_1, \ldots, \lambda_n) \in K^n$, on a $\prod_{j=1}^{n} \lambda_{\sigma(j)} = \prod_{j=1}^{n} \lambda_j$. En particulier, posant $\lambda_j = a_{j,\sigma^{-1}(j)}$, on trouve $\prod_{j=1}^{n} a_{j,\sigma^{-1}(j)} = \prod_{j=1}^{n} a_{\sigma(j),j}$. Donc $\det(P) = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) \prod_{j=1}^{n} a_{j,\sigma^{-1}(j)}$. Notons que $\sigma \mapsto \sigma^{-1}$ est une bijection de $\mathfrak{S}_n$ dans lui-même. Il vient $\det(P) = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma^{-1}) \prod_{j=1}^{n} a_{j,\sigma(j)} = \det({}^tP)$, puisque $\varepsilon(\sigma^{-1}) = \varepsilon(\sigma)$.

Il résulte de ce théorème que $P \mapsto \det(P)$ est un homomorphisme de groupes de $GL_n(K)$ dans K^*. Son noyau $\{P \in \mathcal{M}_n(K);\ \det P = 1\}$ s'appelle le *groupe spécial linéaire* et se note $SL_n(K)$.

7.13 Formules de Cramer. Soit $A \in \mathcal{M}_n(K)$ une matrice inversible et $B \in \mathcal{M}_{n,1}(K)$ une matrice-colonne. Écrivons $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$. D'après 7.7, la solution du système $AX = B$ est donnée par $x_j = \dfrac{\Delta_j}{\Delta}$ où $\Delta = \det(A)$ et Δ_j est le déterminant de la matrice obtenue en remplaçant la $j^{\text{ème}}$ colonne de A par la colonne B.

Mineurs. Soit $A \in \mathcal{M}_{m,n}(K)$. On appelle *mineurs* de A les déterminants des matrices carrées extraites de A.

La proposition 6.14 s'énonce :

7.14 Proposition. *Le rang d'une matrice est l'ordre de son plus grand mineur non nul.*

Coffacteurs. Soit $A \in \mathcal{M}_n(K)$ une matrice carrée et $i, j \in \{1, \ldots n\}$. On note $A_{i,j} \in \mathcal{M}_{n-1}(K)$ la matrice obtenue en enlevant de A sa $i^{\text{ème}}$ ligne et sa $j^{\text{ème}}$ colonne.

7.15 Lemme. *Soit $A = (a_{k,\ell}) \in \mathcal{M}_n(K)$ une matrice carrée et $i, j \in \{1, \ldots n\}$. On suppose que $a_{i,j} = 1$ et que pour $k \neq i$, on a $a_{k,j} = 0$. Alors $\det A = (-1)^{i+j} \det A_{i,j}$.*

Supposons d'abord que $i = j = n$. Dans la somme $\det A = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) \prod_{k=1}^{n} a_{\sigma(k),k}$ les seuls termes non nuls sont obtenus par les $\sigma \in \mathfrak{S}_n$ tels que $\sigma(n) = n$. Identifions avec $\mathfrak{S}_{n-1}$ l'ensemble de ces permutations. On a $\det A = \sum_{\sigma \in \mathfrak{S}_{n-1}} \varepsilon(\sigma) a_{n,n} \prod_{k=1}^{n-1} a_{\sigma(k),k} = \det(A_{n,n})$.

Pour le cas général, considérons les permutations $c_i = (i, i+1, \ldots, n)$ et $c_j = (j, j+1, \ldots, n)$ et P_{c_i} les matrices de permutation associées. La matrice $B = P_{c_i}^{-1} A P_{c_j}$ s'écrit par blocs $B = \begin{pmatrix} A_{i,j} & 0 \\ * & 1 \end{pmatrix}$, donc par le cas $i = j = n$, on a $\det B = \det A_{i,j}$. Or c_i étant un cycle de longueur $n - i + 1$, on a $\det(P_{c_i}) = \varepsilon(c_i) = (-1)^{n-i}$ et $\det(P_{c_j}) = (-1)^{n-j}$, d'où le résultat.

7.16 Proposition: développement relativement à une colonne ou une ligne.
Soit $A = (a_{i,j}) \in \mathcal{M}_n(K)$ une matrice carrée.

a) Pour tout $j \in \{1, \dots, n\}$, on a $\det A = \sum_{i=1}^{n} (-1)^{i+j} a_{i,j} \det A_{i,j}$.

b) Pour tout $i \in \{1, \dots, n\}$, on a $\det A = \sum_{j=1}^{n} (-1)^{i+j} a_{i,j} \det A_{i,j}$.

Pour $i, j \in \{1, \dots, n\}$ notons $B_{i,j} \in \mathcal{M}_n(K)$ la matrice qui a les mêmes colonnes que A sauf la $j^{\text{ème}}$ qui est égale à e_i. Par linéarité en la $j^{\text{ème}}$ colonne, on a $\det A = \sum_{i=1}^{n} a_{i,j} \det B_{i,j}$. L'assertion (a) résulte donc du lemme 7.15. En remplaçant A par sa matrice transposée, on en déduit (b).

Comatrice

7.17 Définition. Soit $A = (a_{i,j}) \in \mathcal{M}_n(K)$ une matrice carrée. On appelle *coffacteur* associé à (i, j) pour $i, j \in \{1, \dots, n\}$ le terme $(-1)^{i+j} \det A_{i,j}$. On appelle *comatrice* de A la matrice $\mathrm{com}(A)$ de terme général $(-1)^{i+j} \det A_{i,j}$.

7.18 Proposition. *Soit $A \in \mathcal{M}_n(K)$ une matrice carrée. On a $A\,{}^t\mathrm{com}(A) = {}^t\mathrm{com}(A)A = \det(A) I_n$. En particulier, si A est inversible, on a $A^{-1} = (\det A)^{-1}\,{}^t\mathrm{com}(A)$.*

Remarquons que cette formule pour l'inverse de A est très peu praticable - sauf en dimension 2... Si $ad - bc \neq 0$, l'inverse de $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ est $\dfrac{1}{ad-bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$.

7.2.4 Interprétation du déterminant lorsque le corps de base est $\mathbb{R}$

Supposons que le corps de base soit $\mathbb{R}$.

Signe du déterminant et *orientation*. On dit que deux bases B et B' ont *même orientation* si $\det_B(B') \in \mathbb{R}_+^*$; on dit sinon qu'elles ont une *orientation opposée*. La relation « avoir même orientation » est une relation d'équivalence. Ainsi, les bases de E se séparent en deux classes d'équivalence. Choisir une orientation de l'espace c'est choisir une de ces deux classes.

Valeur absolue du déterminant et *volume*. Pour fixer une mesure des volumes sur E - c'est à dire une mesure de Lebesgue sur E, choisissons une base $B = (e_1, \dots, e_n)$ et normalisons le volume en décidant que le volume du « cube » $\{\sum_{i=1}^{n} t_i e_i;\ (t_1, \dots, t_n) \in [0,1]^n\}$ est 1. Si $(x_1, \dots, x_n)$ est une famille de vecteurs, le volume du parallélépipède $\{\sum_{i=1}^{n} t_i x_i;\ (t_1, \dots, t_n) \in [0,1]^n\}$ est $|\det_B(x_1, \dots, x_n)|$.

Un endomorphisme f de E multiplie le volume par $|\det f|$: si $A \subset E$ est une partie « mesurable » on a $vol(f(A)) = |\det f|\, vol(A)$.

Plus généralement, la formule de changement de variable d'une intégrale multiple pour un difféomorphisme fait aussi intervenir la valeur absolue d'un déterminant : si U et V sont des ouverts de $\mathbb{R}^n$ et $f : U \to V$ est un difféomorphisme de classe C^1, pour toute fonction intégrable $g : V \to \mathbb{R}$ on a

$$\int_V g(x_1, \dots, x_n)\, dx_1 \dots dx_n = \int_U g \circ f(x_1, \dots, x_n) \left| J_f(x_1, \dots, x_n) \right| dx_1 \dots dx_n$$

où J_f est le déterminant de la matrice jacobienne.

7.3 Opérations élémentaires sur les matrices

Nous expliquons à présent comment de façon algorithmique
- calculer le rang et le déterminant d'une matrice ;
- trouver l'inverse d'une matrice inversible ;
- résoudre un système d'équations linéaires...

7.3.1 Matrices élémentaires

On fixe un corps commutatif K et $n \in \mathbb{N}$ ($n \geqslant 2$). Pour $i, j \in \{1, \dots n\}$, notons $E_{i,j} \in \mathcal{M}_n(K)$ la matrice dont tous les coefficients sont nuls sauf celui d'indice (i, j) (i-ième ligne, j-ième colonne) qui vaut 1. Les $E_{i,j}$ forment une base de $\mathcal{M}_n(K)$.

On appelle matrices élémentaires trois types de matrices carrées :

Transvections. Soient $i, j \in \{1, \dots, n\}$ avec $i \neq j$ et soit $\lambda \in K^*$. Posons $T_{i,j}(\lambda) = I_n + \lambda E_{i,j}$. Cette matrice a donc tous ses coefficients diagonaux égaux à 1, ses coefficients hors-diagonaux nuls sauf celui d'indice (i, j) qui vaut λ.

Les matrices $T_{i,j}(\lambda)$ s'appellent des *matrices de transvection*.

Dilatations. Soit $i \in \{1, \dots, n\}$ et soit $\lambda \in K^*$, $\lambda \neq 1$. Posons $D_i(\lambda) = I_n + (\lambda - 1)E_{i,i} \in \mathcal{M}_n(K)$. Cette matrice a donc tous ses coefficients hors-diagonaux nuls, ses coefficients diagonaux égaux à 1 sauf celui d'indice (i, i) qui vaut λ.

Les matrices $D_i(\lambda)$ s'appellent des *matrices de dilatation*.

Transpositions. Soient $i, j \in \{1, \dots, n\}$ avec $i \neq j$. Posons $P_{i,j} = I_n - E_{i,i} - E_{j,j} + E_{i,j} + E_{j,i}$.

Cette matrice a donc tous ses coefficients diagonaux égaux à 1 sauf ceux d'indice (i, i) et (j, j) qui sont nuls, et ses coefficients hors-diagonaux nuls sauf ceux d'indice (i, j) et (j, i) qui valent 1.

Les matrices $P_{i,j}$ s'appellent des *matrices de transposition*.

Les matrices de transposition sont des cas particuliers des matrices de permutation : soit $\sigma \in \mathfrak{S}_n$ une permutation ; on appelle matrice de permutation associée la matrice $P_\sigma = (a_{k,\ell}) \in \mathcal{M}_n(K)$ telle que
- $a_{k,\ell} = 1$ si $k = \sigma(\ell)$;
- $a_{k,\ell} = 0$ si $k \neq \sigma(\ell)$.

Remarquons que $\sigma \mapsto P_\sigma$ est un homomorphisme de groupes de $\mathfrak{S}_n$ dans $GL_n(K)$.

Les matrices élémentaires sont inversibles : on a $T_{i,j}(\lambda)^{-1} = T_{i,j}(-\lambda)$, $D_i(\lambda)^{-1} = D_i(1/\lambda)$ et $P_{i,j}^{-1} = P_{i,j}$.

7.3.2 Opérations sur les lignes et les colonnes

Soit $A \in M_{m,n}(K)$ une matrice. On appelle *opération élémentaire sur les lignes* de A une opération d'un des trois types suivants :

(L1) Ajouter à une ligne un multiple d'une autre ligne.

(L2) Multiplier une ligne par un scalaire non nul.

(L3) Intervertir deux lignes.

Ces opérations reviennent à multiplier A à gauche par une matrice élémentaire. Ainsi

1. la matrice $T_{i,j}(\lambda)A$ s'obtient à partir de A en ajoutant à la $i^{\text{ème}}$ ligne λ fois la $j^{\text{ème}}$ ligne.
2. la matrice $D_i(\lambda)A$ s'obtient à partir de A en multipliant la $i^{\text{ème}}$ ligne par λ.
3. la matrice $P_{i,j}A$ s'obtient à partir de A en intervertissant la $i^{\text{ème}}$ et la $j^{\text{ème}}$ lignes.

De même, on appelle *opération élémentaire sur les colonnes* de A une opération d'un des trois types suivants :

(C1) Ajouter à une colonne un multiple d'une autre colonne.

(C2) Multiplier une colonne par un scalaire non nul.

(C3) Intervertir deux colonnes.

Ces opérations reviennent à multiplier A à droite par une matrice élémentaire.

7.3.3 Opérations sur les lignes : Algorithme de Gauss

Soit $A \in \mathcal{M}_{m,n}(K)$.

L'algorithme de Gauss consiste à effectuer des opérations élémentaires sur les lignes d'une matrice jusqu'à ce qu'elle obtienne une forme « simple ». Commençons par définir ces matrices simples. Notons $(E_1, \dots E_m)$ la base canonique de l'espace vectoriel $\mathcal{M}_{m,1}(K)$ des matrices-colonnes à m lignes.

7.19 Définition. Convenons d'appeler *matrice échelonnée réduite* ou *matrice à pivots* une matrice $A \in M_{m,n}(K)$, telle qu'il existe $r \in \{0, \dots m\}$ et une application strictement croissante $k \mapsto j(k)$ de $\{1, \dots, r\}$ dans $\{1, \dots, n\}$ satisfaisant :
- pour $k \in \{1, \dots, r\}$, la colonne d'ordre $j(k)$ de A est égale à E_k ;
- pour $i \in \{1, \dots, r\}$ et $j < j(i)$ on a $a_{i,j} = 0$;
- pour $i \in \{r+1, \dots, m\}$ et tout $j \in \{1, \dots n\}$ on a $a_{i,j} = 0$.

Au bout du k-ième pas de l'algorithme de Gauss on aura choisi $j(k)$ et les $j(k)$ premières colonnes de $A^{(k)}$ formeront une matrice échelonnée réduite (qui ne changera plus dans la suite).

L'algorithme de Gauss est le suivant :

- On pose $A^{(0)} = A$ et on construit successivement des matrices $A^{(k)} = (a_{i,j}^{(k)})$ pour $1 \leqslant k \leqslant m$.
- Soit $k \in \{1, \dots, m\}$ et supposons $A^{(k-1)}$ construit.
  Si les $m - k + 1$ dernières lignes sont nulles, la matrice $A^{(k-1)}$ est échelonnée réduite.
  S'il existe $i \geqslant k$ et j tels que $a_{i,j}^{(k-1)} \neq 0$, on note $j(k)$ le plus petit j tel qu'il existe $i \geqslant k$ avec $a_{i,j}^{(k-1)} \neq 0$ et on choisit $i(k) \geqslant k$ tel que $a_{i(k),j(k)}^{(k-1)} \neq 0$. Le couple $(i(k), j(k))$ ainsi choisi s'appelle un *pivot*.
  Maintenant on fait subir à $A^{(k-1)}$ successivement les opérations élémentaires suivantes :
  (E1) on divise la ligne d'ordre $i(k)$ par $a_{i(k),j(k)}^{(k-1)}$; (opération de type (L2))
  (E2) on intervertit la ligne d'ordre k et la ligne d'ordre $i(k)$; (opération de type (L3))
  (E3) maintenant $a_{k,j(k)} = 1$. On annule tous les autres termes de la colonne $j(k)$ ([2]) : pour $i \neq k$ on retranche $a_{i,j(k)}$ fois la ligne d'ordre k à la ligne d'ordre i. (opérations de type (L1))
- Lorsque $k = m$ ou lorsque toutes les lignes d'ordre $i \geqslant k+1$ de $A^{(k)}$ sont nulles, on a obtenu une matrice échelonnée réduite : on arrête l'algorithme.

7.20 Remarque. Remarquons que dans la k-ième étape, lorsque $k < m$, on peut utiliser uniquement des opérations de type (L1). En effet,

a) si $a_{k,j(k)}^{(k-1)} = 1$ on effectue directement l'étape (E3) qui n'utilise que les opérations (L1) ;

b) si un coefficient $i > k$ est non nul, en ajoutant un multiple convenable de la i-ième ligne à la k-ième on arrive à $a_{k,j(k)} = 1$;

c) si $a_{k,j(k)} \neq 1$ et $a_{i,j(k)} = 0$ pour tout $i > k$, on commence par ajouter la k-ième ligne à la suivante (on peut puisque $k < m$), et on est ramené au cas (b).

2. Souvent, on ne fait pas la dernière étape de cette transformation. On obtient ainsi une matrice échelonnée « non réduite ». On a toujours des « pivots » qui sont des 1 en position $(k, j(k))$, avec des 0 à gauche et en dessous, mais on n'impose pas qu'il ait aussi des 0 au dessus.

7.3.4 Applications

Résolution pratique de systèmes linéaires

On veut résoudre un système d'équations $AX = B$, où $A \in \mathcal{M}_{m,n}(K)$. On crée une matrice $C = (A\ B)$ à m lignes et $n+1$ colonnes. En lui appliquant l'algorithme de Gauss *sur les lignes*, on obtient une matrice échelonnée réduite $C' = (A'\ B')$ avec $A' = UA$ et $B' = UB$ où U est une matrice inversible, donc un système $A'X = B'$ équivalent à celui du départ.

Notons r le rang de C. Deux cas sont possibles :

a) On a $j(r) = n+1$, c'est-à-dire $\operatorname{rg} A = r - 1 < r = \operatorname{rg}(A\ B)$ et l'équation d'ordre r du système $A'X = B'$ est $0 = 1$: il n'y a pas des solutions.

b) On a $j(r) \leqslant n$, *i.e.* $\operatorname{rg} A = \operatorname{rg} C$ et le système admet des solutions qui sont très facilement paramétrées par les x_j pour j qui n'est pas de la forme $j(k)$.

Inversion de matrices carrées

Soit A une matrice carrée d'ordre n inversible. Lorsqu'on fait subir à A l'algorithme de Gauss *sur les lignes*, on obtient nécessairement une échelonnée réduite inversible. La seule telle matrice est I_n.

Inverser la matrice A, revient à résoudre un système $AX = B$ pour toute matrice B. Si on pose $B = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$ et que l'on résout le système $AX = B$ *i.e.* on fait subir à $(A\ B)$ l'algorithme de Gauss, la solution sera donnée sous la forme $X = A^{-1}B$.

Une autre façon d'effectuer le même calcul est la suivante : il suffit d'opérer l'algorithme de Gauss aux ***lignes*** de la matrice $(A\, I_n)$ $(\in \mathcal{M}_{n,2n}(K))$. À la fin de l'algorithme on aura la matrice $(I_n\, A^{-1})$.

Génération de $GL(n,K)$ et $SL(n,K)$

Si $A \in GL(n,K)$, on a démontré que l'on peut la multiplier par des matrices élémentaires et obtenir I_n. En particulier, A^{-1} est produit de matrices élémentaires. On en déduit immédiatement :

7.21 Proposition. *$GL(n,K)$ est engendré par les matrices élémentaires.*

Grâce à la remarque 7.20, on peut faire mieux :

7.22 Théorème. *a) $SL(n,K)$ est engendré par les transvections.*

b) $GL(n,K)$ est engendré par les transvections et les dilatations.

Calculs de rang, de déterminants

Pour calculer le rang d'une matrice A, on peut simplement lui faire subir l'algorithme de Gauss. Par contre, comme on ne change pas le rang en multipliant à gauche ou à droite par une matrice inversible, on peut utiliser à loisir les opérations à la fois sur les lignes et sur les colonnes.

On peut calculer le déterminant d'une matrice carrée A en utilisant les opérations sur les lignes et les colonnes. La seule chose à retenir est :

- Lorsqu'on ajoute à une ligne (*resp.* colonne) un multiple d'une ligne (*resp.* colonne) on ne change pas le déterminant.
- Lorsqu'on multiplie une ligne (ou une colonne) par un scalaire on multiplie le déterminant par ce scalaire.

- Lorsqu'on intervertit deux lignes ou deux colonnes on multiplie le déterminant par -1.

7.23 Remarques. a) Lorsqu'on effectue des calculs de manière approchée, on ne pourra jamais démontrer que le rang est strictement inférieur à $\min(m,n)$. En effet, l'ensemble des matrices de rang $\min(m,n)$ est dense. Par contre, on arrivera à *minorer le rang* puisque les matrices de rang $\geqslant k$ forment un ouvert, par exemple, si on arrive à démontrer qu'un déterminant extrait d'ordre k est égal à D à ε près et que $0 \notin]D-\varepsilon, D+\varepsilon[$.

b) On peut effectuer de façon très efficace des opérations élémentaires sur les matrices à coefficients entiers - et plus généralement sur un anneau euclidien A. Dans ce cas, les seules dilatations « autorisées » sont les $D_i(\lambda)$ avec λ inversible dans A. La division euclidienne est après tout l'opération élémentaire $\begin{pmatrix} a \\ b \end{pmatrix} \to \begin{pmatrix} a-bq \\ b \end{pmatrix}$. L'algorithme d'Euclide nous dit que par opérations élémentaires on peut passer de $\begin{pmatrix} a \\ b \end{pmatrix}$ à $\begin{pmatrix} d \\ 0 \end{pmatrix}$ où d est le PGCD de a et b.

Nous ne pousserons pas plus loin ces considérations - qui sont hors programme - mais signalons juste que l'on démontre ainsi que pour un anneau euclidien A :

- $SL_n(A)$ est engendré par les matrices de transvection ;
- toute matrice de $M_{m,n}(A)$ est équivalente à une matrice $\begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix}$ où D est une matrice carrée d'ordre r diagonale $D = \operatorname{diag}(d_i)$ avec $d_1 | d_2 | \dots | d_r$.

7.4 Exercices

7.4.1 Calculs de déterminants

7.1 Exercice. Soient $A \in \mathcal{M}_p(K)$, $B \in \mathcal{M}_{p,q}(K)$ et $C \in \mathcal{M}_q(K)$. Notons $M \in \mathcal{M}_{p+q}(K)$ la matrice qui admet la décomposition par blocs $M = \begin{pmatrix} A & B \\ 0_{q,p} & C \end{pmatrix}$. Démontrer que l'on a $\det M = \det A \det C$.

Indication : Distinguer le cas où A est inversible et celui où elle ne l'est pas.

Par récurrence, on en déduit que si M est une matrice triangulaire par blocs et si l'on note $A_1, \dots, A_k$ ses blocs diagonaux, alors $\det M = \det A_1 \dots \det A_k$.

7.2 Exercice. Soient $a_1, \dots, a_n \in K$. Considérons le déterminant

$$\Delta_n(a_1, \dots, a_n) = \begin{vmatrix} 1 & a_1 & a_1^2 & \dots & a_1^{n-1} \\ 1 & a_2 & a_2^2 & \dots & a_2^{n-1} \\ 1 & a_3 & a_3^2 & \dots & a_3^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & a_n & a_n^2 & \dots & a_n^{n-1} \end{vmatrix}$$

appelé *déterminant de Vandermonde.*

1. Calculer $\Delta_2(a_1, a_2)$.
2. En considérant $\Delta_n(a_1, \dots, a_n)$ comme polynôme en a_n, établir l'égalité
$$\Delta_n(a_1, \dots, a_n) = \Delta_{n-1}(a_1, \dots, a_{n-1}) \prod_{k=1}^{n-1} (a_n - a_k).$$
3. En déduire que $\Delta_n(a_1, \dots, a_n) = \displaystyle\prod_{1 \leqslant j < k \leqslant n} (a_k - a_j)$.
4. Pouvait-on prévoir dès le départ que si les a_i sont distincts $\Delta_n(a_1, \dots, a_n) \neq 0$? Quelle application linéaire représente cette matrice ?

7.3 Exercice. Soit $P = \sum_{k=0}^{n} a_k X^k \in K[X]$ un polynôme unitaire ($a_n = 1$). En développant relativement à une ligne ou une colonne, démontrer que l'on a

$$\begin{vmatrix} \lambda & 0 & 0 & \dots & 0 & a_0 \\ -1 & \lambda & 0 & \dots & 0 & a_1 \\ 0 & -1 & \lambda & \ddots & 0 & a_2 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \lambda & a_{n-2} \\ 0 & 0 & 0 & \dots & -1 & \lambda + a_{n-1} \end{vmatrix} = P(\lambda).$$

7.4 Exercice. Soient K un corps commutatif et $n \in \mathbb{N}^*$. Soit $P = X^n + \sum_{k=0}^{n-1} a_k X^k$ un polynôme unitaire de degré n. Notons E_n l'espace vectoriel des polynômes de degré $< n$. Soit $T \in L(E_n)$ l'application linéaire qui à $Q \in E_n$ associe le reste de la division euclidienne de XQ par P.

1. Quelle est la matrice de T dans la base $(1, X, \dots, X^{n-1})$ de E ?
2. Calculer $\det T$.
3. Quelle est la matrice de $T - \lambda \mathrm{id}_{E_n}$ dans la base $(1, X - \lambda, \dots, (X - \lambda)^{n-1})$.
4. Calculer $\det(T - \lambda \mathrm{id}_{E_n})$.

7.5 Exercice. *Déterminant de Cauchy.* Soient $x_1, \dots x_n, y_1 \dots, y_n \in K$. On suppose que pour tout i, j, on a $x_i + y_j \neq 0$. Calculer le déterminant de la matrice $\left(\dfrac{1}{x_i + y_j}\right)$.

7.6 Exercice. Pour $n \in \mathbb{N}^*$, calculer la différence entre le nombre de dérangements pairs et le nombre des dérangements impairs.

7.7 Exercice. [Semi-continuité du rang] Soit $m, n, r \in \mathbb{N}$. Démontrer que les matrices de rang $\leqslant r$ forment un fermé de $\mathcal{M}_{m,n}(\mathbb{K})$ (pour $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$). On suppose que $r \leqslant \min(m, n)$. Quelle est l'adhérence de l'ensemble des matrices de rang r ?

7.4.2 Opérations élémentaires

Il est certainement utile de faire un certain nombre d'exercices pratiques de résolution de systèmes linéaires - que l'on trouve dans de nombreux ouvrages. On ne présente ici que quelques exercices un peu plus théoriques.

7.8 Exercice. Calculer le déterminant de Vandermonde par opérations élémentaires.

Indication : On commence par retrancher la première ligne de toutes les autres - et on développe par la première colonne ; puis on met un terme en facteur dans toutes les lignes obtenues ; enfin on retranche chaque colonne à la suivante.

7.9 Exercice. Soient $A \in \mathcal{M}_{m,n}(K)$ une matrice et U une matrice carrée d'ordre m inversible. Posons $A' = UA$. Notons $C_1, \dots, C_n$ les colonnes de A, r son rang, et $C'_1, \dots, C'_n$ les colonnes de A' et r' son rang.

1. Démontrer que $r = r'$ et que pour toute partie J de $\{1, \dots n\}$, on a $\mathrm{rg}\{C_j;\ j \in J\} = \mathrm{rg}\{C'_j;\ j \in J\}$. On suppose que $A' = UA$ est échelonnée réduite.
2. Pour $k = 1, \dots, r$, notons $j(k)$ la place du $k^{\text{ème}}$ pivot de A'. Démontrer que

$$j(k) = \inf\{j;\ \mathrm{rg}(C_1, \dots, C_j) = k\}.$$

3. Écrivons $A' = (a'_{i,j})$. Démontrer que l'on a $C_j = \sum_{k=1}^{r} a'_{k,j} C_{j(k)}$.

4. En déduire que pour toute matrice A il existe une et une seule matrice échelonnée réduite qui s'écrive $A' = UA$ avec U inversible.

   En ce sens, les matrices échelonnée réduite représentent les classes de l'action de $GL(n)$ par multiplication à gauche sur $M_{n,p}$.

7.10 Exercice. 1. Démontrer que $SL(n, K)$ est le sous-groupe des commutateurs de $GL(n, K)$ (sauf pour $n = 2$ et $K = \mathbb{F}_2$).

2. Démontrer que $SL(n, \mathbb{R})$ et $SL(n, \mathbb{C})$ sont connexes. En déduire que $GL(n, \mathbb{R})$ a deux composantes connexes et $GL(n, \mathbb{C})$ est connexe.

7.11 Exercice. (****) Soient $p \in \mathbb{N}$ et E un K-espace vectoriel de dimension finie. Notons Λ_p l'espace vectoriel des formes p-linéaires alternées $D : E^p \to K$.

1. Soient F un espace vectoriel de dimension p, $f : E \to F$ une application linéaire et B une base de F. Démontrer que l'application $(x_1, \ldots, x_p) \mapsto \det_B(f(x_1), \ldots, f(x_p))$ est un élément de Λ_p.

2. Fixons une base $(e_1, \ldots, e_n)$ de E. Notons J_p l'ensemble des applications strictement croissantes $s : \{1, \ldots, p\} \to \{1, \ldots, n\}$. Démontrer que l'application $\Phi : \Lambda_p \to K^{J_p}$ définie par $\Phi(D)(s) = D(e_{s(1)}, e_{s(2)}, \ldots, e_{s(p)})$ est linéaire et bijective.

3. En déduire que pour $p \leqslant n$, l'espace vectoriel Λ_p est de dimension $\binom{n}{p}$, et que, pour $p > n$, l'espace vectoriel des formes p-linéaires alternées est réduit à 0.

8 Réduction des endomorphismes

Nous énonçons ici les définitions, propriétés, résultats en termes d'endomorphismes. Ils peuvent bien sûr être aussi énoncés en termes de similitude de matrices carrés. On peut en effet identifier une matrice carrée $A \in \mathcal{M}_n(K)$ avec l'endomorphisme $X \mapsto AX$ de $\mathcal{M}_{n,1}(K)$ qu'elle définit.

8.1 Vecteurs propres et valeurs propres

8.1.1 Sous-espaces stables par un endomorphisme

8.1 Définition. Soit u un endomorphisme d'un espace vectoriel E. Un sous-espace vectoriel F est dit *stable* par u si $u(F) \subset F$. L'application $x \mapsto u(x)$ de F dans F est un endomorphisme de f appelé *endomorphisme* de F *induit* par u.

Soit E un espace vectoriel de dimension finie, u un endomorphisme de E et F un sous-espace vectoriel de E. Soit $(e_1, \ldots, e_k)$ une base de F que l'on complète en une base $B = (e_1, \ldots, e_n)$ de E. Alors F est stable par u si et seulement si la matrice de u dans la base B est de la forme $\begin{pmatrix} * & * \\ 0 & * \end{pmatrix}$.

8.2 Proposition. *Si les endomorphismes u et v commutent, alors* $\operatorname{im} u$ *et* $\ker u$ *sont stables par v.*

8.1.2 Vecteurs propres et valeurs propres

On cherche à présent les espaces stables de dimension 1.

8.3 Définition. Soient E un K-espace vectoriel et u un endomorphisme de E.

a) Soient $\lambda \in K$ et x un vecteur non nul de E. On dit que λ et x sont une *valeur propre et un vecteur propre associés* si $u(x) = \lambda x$.

b) Soit $\lambda \in K$. On dit que λ est une *valeur propre* de u s'il existe un vecteur $x \in E$ non nul tel que $u(x) = \lambda x$.

c) Soit $x \in E$ un vecteur non nul. On dit que x est *vecteur propre* de u si $u(x)$ est proportionnel à x.

d) Soit $\lambda \in K$ une valeur propre de u. L'espace propre associé est $\ker(u - \lambda \mathrm{id}_E) = \{x \in E;\ u(x) = \lambda x\}$. On le note $E_\lambda(u)$.

Il est parfois commode de poser $E_\lambda(u) = \ker(u - \lambda \mathrm{id}_E)$ même lorsque λ n'est pas une valeur propre de u. Dans ce cas, on a $E_\lambda(u) = \{0\}$.

8.4 Proposition. *Les espaces propres d'un endomorphisme sont en somme directe.*

On doit démontrer que pour tout N et tout n-uplet $(\lambda_1, \ldots, \lambda_N)$ de valeurs propres distinctes de u les espaces propres $E_{\lambda_1}, \ldots, E_{\lambda_N}$ sont en somme directe.

- Pour $n = 1$, il n'y a rien à démontrer.
- Soit $N \geqslant 2$ et $\lambda_1, \ldots, \lambda_N$ des valeurs propres distinctes de u. Supposons que les espaces propres $E_{\lambda_1}, \ldots, E_{\lambda_{N-1}}$ soient en somme directe. Soient $x_i \in E_{\lambda_i}$ tels que $\sum_{i=1}^{N} x_i = 0$. Alors $0 = u(\sum_{i=1}^{N} x_i) = \sum_{i=1}^{N} \lambda_i x_i$, donc $\sum_{i=1}^{N-1} (\lambda_N - \lambda_i) x_i = \lambda_N \sum_{i=1}^{N} x_i - \sum_{i=1}^{N} \lambda_i x_i = 0$. Puisque $E_{\lambda_1}, \ldots, E_{\lambda_{N-1}}$ sont en somme directe, il vient $(\lambda_N - \lambda_1)x_1 = \ldots = (\lambda_N - \lambda_{N-1})x_{N-1} = 0$ et puisque les λ_i sont distincts, il vient $x_1 = \ldots = x_{N-1} = 0$. Enfin l'égalité $\sum_{i=1}^{N} x_i = 0$ permet de conclure que $x_N = 0$ aussi.

8.1.3 Polynôme caractéristique

Soient E un K-espace vectoriel de dimension finie et u un endomorphisme de E. Notons $A = (a_{i,j}) \in \mathcal{M}_n(K)$ la matrice de u dans une base B de E. Notons $M = (m_{i,j}) \in \mathcal{M}_n(K[X])$ la matrice définie par $m_{i,j} = a_{i,j}$ si $i \neq j$ et $m_{i,i} = a_{i,i} - X$. On pose $\chi_u^B = \det(M) \in K[X]$) (*cf.* remarque 7.11).

Soit B_1 une autre base de E, notons $P \in \mathcal{M}_n(K)$ la matrice de passage de B à B_1 et $A_1 = P^{-1}AP$ la matrice de l'endomorphisme u dans la base B_1. On plonge K dans le corps $K(X)$ des fractions rationnelles sur K. Par définition, $\chi_u^B = \det(A - XI_n)$ et $\chi_u^{B_1} = \det(A_1 - XI_n) = \det(P^{-1}(A - XI_n)P) = \det(P^{-1})\chi_u^B \det P$. En d'autres termes, on a démontré que le polynôme χ_u^B ne dépend pas de la base B.

8.5 Définition. Soient E un K-espace vectoriel de dimension finie et u un endomorphisme de E. On appelle *polynôme caractéristique* de l'endomorphisme u et l'on note χ_u le polynôme χ_u^B pour n'importe quelle base B de E.

Pour tout $\lambda \in K$, on a donc $\det(u - \lambda \mathrm{id}_E) = \chi_u(\lambda)$ (on applique la remarque 7.11 à l'homomorphisme $P \mapsto P(\lambda)$ de $K[X]$ dans K). Remarquons que cette égalité définit le polynôme caractéristique si K est infini.

On a immédiatement :

8.6 Proposition. *Soient E un K-espace vectoriel de dimension finie et u un endomorphisme de E. Les valeurs propres de u sont les racines de χ_u.*

8.1.4 Triangulation d'un endomorphisme

Une matrice carrée $A = (a_{i,j} \in \mathcal{M}_n(K)$ est dite *triangulaire supérieure* (*resp. inférieure*) si pour $i > j$ (*resp.* $i < j$) on a $a_{i,j} = 0$.

On dit qu'un endomorphisme u d'un espace vectoriel de dimension finie E est *triangulable* ou *trigonalisable* s'il existe une base de E dans laquelle la matrice de u est triangulaire. On dit qu'une matrice carrée M est *triangulable* ou *trigonalisable* si c'est la matrice d'un endomorphisme trigonalisable, *i.e.* s'il existe une matrice inversible P telle que $P^{-1}MP$ soit triangulaire.

8.7 Théorème. *Un endomorphisme (une matrice carrée) est trigonalisable si et seulement si son polynôme caractéristique est scindé.*

Le polynôme caractéristique d'une matrice triangulaire $A = (a_{i,j})$ est $\prod_{i=1}^{n}(a_{i,i} - X)$. Il est scindé. Si un endomorphisme (une matrice carrée) est trigonalisable, son polynôme caractéristique est égal au polynôme caractéristique d'une matrice triangulaire : il est scindé.

Démontrons la réciproque par récurrence sur la dimension de l'espace :

Si $n = 1$, il n'y a rien à démontrer : toute matrice est triangulaire ! !

Soit $n > 1$. Supposons que tout endomorphisme d'un espace vectoriel de dimension $n-1$ (toute matrice carrée d'ordre $n-1$) à polynôme caractéristique scindé soit trigonalisable. Soit E un espace vectoriel de dimension n et u un endomorphisme dont le polynôme caractéristique χ_u s'écrit $\chi_u = \prod_{i=1}^{n}(\lambda_i - X)$. Comme λ_1 est une valeur propre de E, il existe un vecteur propre e_1 associé à la valeur propre λ_1. Complétons e_1 en une base B de E. La matrice de u dans la base B est de la forme $M = \begin{pmatrix} \lambda_1 & L \\ 0 & N \end{pmatrix}$.

On a $\chi_M = \chi_u = (\lambda_1 - X)\chi_N$, donc $\chi_N = \prod_{i=2}^{n}(\lambda_i - X)$. D'après l'hypothèse de récurrence, N est

trigonalisable : il existe une matrice carrée inversible Q d'ordre $n-1$ telle que $Q^{-1}NQ$ soit triangulaire supérieure. Posons alors $P = \begin{pmatrix} 1 & 0 \\ 0 & Q \end{pmatrix}$. C'est une matrice inversible (d'inverse $\begin{pmatrix} 1 & 0 \\ 0 & Q^{-1} \end{pmatrix}$). La matrice $P^{-1}MP = \begin{pmatrix} \lambda_1 & LQ \\ 0 & Q^{-1}NQ \end{pmatrix}$ est triangulaire, d'où le résultat.

D'après la preuve ci-dessus, si u est un endomorphisme dont le polynôme caractéristique est $\chi_u = \prod_{i=1}^{n}(\lambda_i - X)$ on peut choisir une base dans laquelle la matrice de u sera triangulaire de diagonale $(\lambda_1, \ldots, \lambda_n)$, c'est-à-dire : on peut choisir l'ordre dans lequel apparaissent les éléments diagonaux.

8.1.5 Diagonalisation d'un endomorphisme

On dit qu'un endomorphisme d'un espace vectoriel de dimension finie E est *diagonalisable* s'il existe une base de E dans laquelle la matrice de E est diagonale. On dit qu'une matrice carrée M est *diagonalisable* si c'est la matrice d'un endomorphisme diagonalisable, *i.e.* s'il existe une matrice inversible P telle que $P^{-1}MP$ soit diagonale.

8.8 Proposition. *Soit u un endomorphisme d'un espace vectoriel de dimension finie E. Soit λ une valeur propre de E. Alors* $\dim E_\lambda$ *est inférieure ou égale à l'ordre de multiplicité de la racine λ dans le polynôme caractéristique χ_u.*

Soit $(e_1, \ldots, e_k)$ une base de E_λ. Complétons-la en une base de E. Dans cette base, la matrice de u est de la forme $M = \begin{pmatrix} \lambda I_k & * \\ 0 & N \end{pmatrix}$, donc $\chi_u = (\lambda - X)^k \chi_N$.

8.9 Théorème. *Soit u un endomorphisme d'un espace vectoriel de dimension finie E. Alors u est diagonalisable si et seulement si son polynôme caractéristique est scindé et la dimension de tout sous-espace propre est égale à l'ordre de multiplicité de la valeur propre associée.*

Si la matrice de u dans une base est la matrice diagonale $\mathrm{diag}(\lambda_1, \ldots, \lambda_n)$, la dimension de l'espace propre associé à une valeur propre λ est égale au nombre de j tels que $\lambda_j = \lambda$ qui est lui-même égal à la multiplicité de la racine λ de $\chi_u = \prod_{j=1}^{n}(\lambda_j - X)$.

Inversement, si χ_u est scindé et la dimension de tout sous-espace propre est égale à l'ordre de multiplicité de la valeur propre associée, en mettant ensemble des bases des espaces propres, on obtient une famille libre d'après la prop. 8.4. Le nombre d'éléments de cette famille libre est $\sum_\lambda \dim E_\lambda = \sum_\lambda \mathrm{ordre}(\lambda) = \partial\chi_u = \dim E$: c'est donc une base. La matrice de u dans cette base est diagonale.

Remarquons que toute racine λ de χ_u est une valeur propre de u : on a donc $1 \leqslant \dim E_\lambda \leqslant \mathrm{ordre}(\lambda)$. En particulier, si λ est racine simple de χ_u on aura $\dim E_\lambda = \mathrm{ordre}(\lambda)$. Pour voir si u est diagonalisable, on ne doit donc se préoccuper que des racines multiples de χ_u.

8.10 Corollaire. *Un endomorphisme dont le polynôme caractéristique est scindé à racines simples est diagonalisable.*

8.2 Polynômes d'endomorphismes

8.2.1 Polynômes annulateurs, polynôme minimal

Soient E un espace vectoriel et u un endomorphisme de E. Pour un polynôme $P = \sum_{k=0}^{N} a_k X^k$, on pose $P(u) = \sum_{k=0}^{N} a_k u^k$. L'application $P \mapsto P(u)$ est un homomorphisme d'algèbres de $K[X]$ dans $L(E)$.

Si E est de dimension finie, cet homomorphisme n'est pas injectif. Son noyau est un idéal de l'anneau principal $K[X]$.

8.11 Définition. Un polynôme P est dit *annulateur* pour u si $P(u) = 0$. On appelle *polynôme minimal* de u l'unique polynôme unitaire ϖ_u qui engendre l'idéal formé par les polynômes annulateurs pour u.

En d'autres termes, les polynômes annulateurs sont les multiples du polynôme minimal. Pour $P \in K[X]$, on a donc $P(u) = 0 \iff \varpi_u | P$.

8.2.2 Le théorème de Cayley-Hamilton

8.12 Théorème de Cayley-Hamilton. *Soient E un espace vectoriel de dimension finie et u un endomorphisme de E. On a $\chi_u(u) = 0$.*

En d'autres termes ϖ_u divise χ_u.

Il y a de nombreuses démonstrations de ce théorème. En voici une relativement simple basée sur les matrices Compagnon.

Soit $x \in E$ un vecteur non nul. Soit k le plus petit entier tel que $u^k(x)$ soit contenu dans le sous-espace engendré par les $u^j(x)$ pour $0 \leqslant j < k$. Écrivons $u^k(x) = \sum_{j=0}^{k-1} a_j u^j(x)$. Pour $j = 1, \ldots, k$, posons $e_j = u^{j-1}(x)$. Par définition de k, la famille $(e_1, \ldots, e_k)$ est libre. Complétons-la en une base $(e_1, \ldots, e_n)$ de E. Dans cette base, la matrice de u est sous la forme $\begin{pmatrix} C & D \\ 0 & N \end{pmatrix}$ où C est la matrice carrée d'ordre k (appelée matrice compagnon) : $C = \begin{pmatrix} 0 & 0 & \ldots & 0 & a_0 \\ 1 & 0 & \ldots & 0 & a_1 \\ 0 & 1 & \ldots & 0 & a_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & 1 & a_{k-1} \end{pmatrix}$. Il en résulte que $\chi_u = \chi_C \chi_N$.

Or $(-1)^k \chi_C = X^k - \sum_{j=0}^{k-1} a_j X^j$ (*cf.* Exerc. .7.3). On a donc $(-1)^k \chi_C(u)(x) = u^k(x) - \sum_{j=0}^{k-1} a_j u^j(x) = 0$.

Il vient $\chi_u(u)(x) = \chi_N(u) \circ \chi_C(u)(x) = 0$. Cela étant vrai pour tout x, il vient $\chi_u(u) = 0$.

8.2.3 Théorème de décomposition des noyaux

8.13 Théorème de décomposition des noyaux. *Soient $P_1, \ldots, P_k$ des polynômes premiers entre eux deux à deux. Posons $P = \prod_{j=1}^{k} P_j$. Soient E un espace vectoriel et u un endomorphisme de E.*

a) On a $\ker P(u) = \bigoplus_{j=1}^{k} \ker(P_j(u))$.

b) En particulier, si P est un polynôme annulateur de u, on a $E = \bigoplus_{j=1}^{k} \ker(P_j(u))$. De plus, pour tout j, le projecteur d'image $\ker P_j(u)$ et de noyau $\bigoplus_{i\neq j} \ker(P_i(u))$ est un polynôme en u (i.e. un élément de l'algèbre $K[u]$).

Pour $j \in \{1, \ldots, k\}$, posons $Q_j = P/P_j = \prod_{i\neq j} P_i$. Comme P_j et Q_j sont premiers entre eux, il existe un polynôme R_j tel que $R_jQ_j \equiv 1 \quad [P_j]$. Posons enfin $S_j = R_jQ_j$. Pour $i \neq j$, P_i divise S_j, donc $\sum_{i=1}^{k} S_i \equiv 1 \quad [P_j]$. Alors $1 - \sum_{i=1}^{k} S_i$ est divisible par les P_j, donc par leur PPCM, c'est-à-dire P. Remarquons aussi que, pour tout j, P divise P_jS_j.

- Puisque $P = Q_jP_j$, on a $P(u) = Q_j(u) \circ P_j(u)$, donc $\ker P_j(u) \subset \ker P(u)$. Il vient $\sum_{j=1}^{k} \ker P_j(u) \subset \ker P(u)$.
- Soit $x \in \ker P(u)$. Comme P divise P_jS_j, on a $P_j(u) \circ S_j(u)(x) = 0$, donc $S_j(u)(x) \in \ker P_j(u)$. Comme P divise $1 - \sum_{i=1}^{k} S_i$, si $x \in \ker P(u)$, alors $x = \sum_{j=1}^{k} S_j(u)(x)$. Donc $\ker P(u) \subset \sum_{i=1}^{k} \ker P_j(u)$.
- Donnons-nous des x_j pour $j = 1, \ldots, n$ avec $x_j \in \ker P_j(u)$ tels que $\sum_{j=1}^{k} x_j = 0$. Comme P_j divise $1 - S_j$, alors $S_j(u)(x_j) = x_j$; pour $i \neq j$, P_i divise S_j, donc $S_j(u)(x_i) = 0$. Il vient $x_j = S_j(u)(\sum_{i=1}^{k} x_i) = 0$. Donc les $\ker P_j(u)$ sont en somme directe.

L'assertion (b) résulte aussi des calculs ci-dessus : le projecteur d'image $\ker P_j(u)$ et de noyau $\sum_{i\neq j} \ker(P_i(u))$ est $S_j(u)$.

8.2.4 Endomorphismes diagonalisables

8.14 Proposition. *Soient E un espace vectoriel de dimension finie et u un endomorphisme de E. Les propriétés suivantes sont équivalentes :*

(i) *u est diagonalisable ;*

(ii) *u admet un polynôme annulateur scindé à racines simples ;*

(iii) *le polynôme minimal de u est scindé à racines simples.*

Tout diviseur d'un polynôme scindé à racines simples est scindé à racines simples ; donc $(ii) \iff (iii)$.

Soient $\lambda_1, \ldots, \lambda_k$ les valeurs propres de u et posons $P = \prod_{i=1}^{k} (X - \lambda_i)$. Si u est diagonalisable, il admet une base B de vecteurs propres. L'endomorphisme $P(u)$ est nul sur la base B, donc il est nul. Le polynôme P est donc un polynôme annulateur.

Inversement, s'il existe un polynôme annulateur scindé à racines simples $P = \prod_{i=1}^{k} (X - \lambda_i)$, alors $E = \bigoplus_{i=1}^{k} \ker(u - \lambda_i \mathrm{id}_E)$ d'après le « lemme des noyaux » (théorème 8.13).

8.15 Corollaire. *Soient E un espace vectoriel de dimension finie et u un endomorphisme diagonalisable de E. La restriction de u à tout sous-espace de E stable par u est diagonalisable.*

8.16 Proposition. *Soient E un espace vectoriel et $(u_i)_{i\in I}$ une famille d'endomorphismes de E. On suppose que tous les u_i sont diagonalisables et que pour tout $i, j \in I$, on a $u_i \circ u_j = u_j \circ u_i$. Alors il existe une base $(e_1, \ldots, e_n)$ de E dans laquelle tous les u_i sont diagonaux.*

Procédons une récurrence « forte » sur $\dim E$. Si $\dim E$ est 1, il n'y a rien à démontrer, de même que si tous les u_i sont des homothéties.

Supposons donc que u_{i_0} n'est pas une homothétie. Alors $E = \bigoplus_{k=1}^{m} E_k$ où les E_k sont les espaces propres de u_{i_0} ; chacun de ces espaces est invariant par tous les u_i, et les restrictions des u_i à ces espaces sont diagonalisables et deux à deux permutables. Par l'hypothèse de récurrence, il existe une base de chacun des E_k qui diagonalise toutes les restrictions des u_i. Mettant ensemble toutes ces bases, on trouve une base de E qui diagonalise les u_i.

8.2.5 Sous-espaces caractéristiques

8.17 Définition. Soient E un espace vectoriel de dimension finie, u un endomorphisme de E et λ une valeur propre de u. Soit r l'ordre de multiplicité de λ dans χ_u. On appelle *sous-espace caractéristique* de u associé à la valeur propre λ le noyau de $(\lambda \mathrm{id}_E - u)^r$.

8.18 Proposition. *Soient E un espace vectoriel de dimension finie et u un endomorphisme de E. On suppose que le polynôme caractéristique de u est scindé. Alors E est somme directe des espaces caractéristiques de u.*

Résulte du lemme de décomposition des noyaux (théorème 8.13) à l'aide du théorème de Cayley-Hamilton.

8.19 Définition. Soit E un espace vectoriel. Un endomorphisme u de E est dit *nilpotent* s'il existe $p \in \mathbb{N}$ tel que $u^p = 0$.

Si n est nilpotent, toute valeur propre de u est nulle. En fait $\chi_u = (-X)^{\dim E}$.

8.20 Théorème: Décomposition de Dunford. *Soient E un espace vectoriel de dimension finie et u un endomorphisme de E. On suppose que le polynôme caractéristique de u est scindé. Il existe un unique couple (d, n) d'endomorphismes tels que $u = d + n$ avec d diagonalisable n nilpotent et satisfaisant $d \circ n = n \circ d$.*

Existence. Soient $(\lambda_1, \ldots, \lambda_k)$ les valeurs propres distinctes de u. Pour tout j, notons N_j l'espace caractéristique associé à λ_j. D'après le théorème de décomposition des noyaux, $E = \bigoplus_{j=1}^{k} N_j$.

Notons p_j le projecteur d'image N_j et de noyau $\bigoplus_{i\neq j} N_i$. Posons $d = \sum_{j=1}^{n} \lambda_j p_j$. Il est diagonalisable : son espace propre pour la valeur propre λ_j est N_j. Posons $n = u - d$. Chaque N_j est stable par n et l'endomorphisme de N_j induit par n coïncide avec $u - \lambda_j \mathrm{id}_E$. Il est nilpotent. On en déduit que n est nilpotent. Enfin, les p_j sont des polynômes en u, donc d et n sont des polynômes en u : ils commutent.

Unicité. Soient (d, n) le couple construit dans la partie existence, et (d', n') un autre couple satisfaisant les conditions ci-dessus. Alors d' commute à u, donc à tout polynôme en u. Il commute avec d. D'après la prop. 8.16, d et d' sont simultanément diagonalisables. De même, n' et n commutent. On en déduit (d'après la formule du binôme) que $n' - n$ est nilpotent. Or $d - d' = n' - n$. Cet endomorphisme est diagonalisable et nilpotent. Il est nul.

8.3 Applications ; considérations topologiques dans le cas où le corps K est $\mathbb{R}$ ou $\mathbb{C}$

8.3.1 Puissances de matrices ; suites récurrentes

Donnons-nous une suite (X_k) de vecteurs-colonne définie par une relation de récurrence $X_{k+1} = AX_k$ où A est une matrice carrée donnée. Il vient immédiatement $X_k = A^k X_0$, d'où la nécessité de calculer les puissances de A.

Une telle formule de récurrence peut être un peu plus cachée : fixons $a_0, \dots, a_{n-1} \in K$ et considérons une suite $(x_k)_{k\in\mathbb{N}}$ vérifiant pour tout $k \in \mathbb{N}$, $x_{k+n} = \sum_{j=0}^{n-1} a_j x_{k+j}$. On pose alors $X_k = \begin{pmatrix} x_k \\ x_{k+1} \\ \vdots \\ x_{k+n-1} \end{pmatrix}$. La suite X_k vérifie la relation de récurrence $X_{k+1} = AX_k$ où $A = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ a_0 & a_1 & a_2 & \dots & a_{n-1} \end{pmatrix}$ est (la transposée d')une matrice Compagnon

Il est bien plus facile de calculer les puissances d'une matrice si elle est sous forme diagonale ! Si A est diagonalisable, elle s'écrit $A = PDP^{-1}$ avec $D = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$ diagonale. Alors, on a $A^k = PD^kP^{-1}$ et $D^k = \mathrm{diag}(\lambda_1^k, \dots, \lambda_n^k)$.

Faisons quelques remarques dans le cas où $K = \mathbb{R}$ ou $K = \mathbb{C}$. Rappelons que, sur un espace vectoriel réel ou complexe de dimension finie, toutes les normes sont équivalentes et définissent donc la même topologie, en particulier la même notion de convergence des suites, de continuité, différentiabilité, *etc.*

1. Si $K = \mathbb{R}$, il peut être utile de considérer une matrice $A \in \mathcal{M}_n(\mathbb{R})$ comme matrice à coefficients complexes et de la diagonaliser sur $\mathbb{C}$.
2. Si $A \in \mathcal{M}_n(\mathbb{C})$ n'est pas diagonalisable, on peut utiliser la décomposition de Dunford : $A = D+N$ avec $DN = ND$ où A est diagonalisable et N est nilpotente. Comme $N^n = 0$, la formule du binôme qui calcule $(D+N)^k$ (pour k grand) n'a que n termes.
3. On peut étudier le comportement à l'infini d'une suite $X_k = A^k X_0$. Par les deux remarques précédentes, l'étude d'une telle suite se ramène (presque) à une étude de suites géométriques.

8.3.2 Exponentielles de matrices et applications

8.21 Proposition. *Soit E un espace vectoriel réel ou complexe de dimension finie.*

a) Soit $u \in L(E)$. La série de terme général $\dfrac{u^n}{n!}$ est convergente.

On pose $\exp(u) = \sum_{n=0}^{+\infty} \dfrac{u^n}{n!}$.

b) Soit $u, v \in L(E)$ tels que $u \circ v = v \circ u$. On a $\exp(u+v) = \exp(u)\exp(v)$.

c) Soit $u \in L(E)$. L'équation différentielle $x'(t) = u\,x(t)$ admet comme solution $x(t) = \exp(tu)x_0$.

Considérons le système différentiel $x'(t) = ux(t) + b(t)$, où b est une fonction continue définie sur un intervalle ouvert I à valeurs dans E. Cherchons la solution sous la forme $x(t) = \exp(tu)y(t)$. Le système devient $\exp(tu)y'(t) = b(t)$, soit $y(t) = y(t_0) + \int_{t_0}^{t} \exp(-tu)b(t)\,dt$.

Encore une fois, il est bien plus facile de calculer l'exponentielle d'une matrice si elle est diagonalisée : remarquons que $\exp(PDP^{-1}) = P\exp(D)\,P^{-1}$. Si l'endomorphisme u n'est pas diagonalisable, on utilisera sa décomposition de Dunford pour calculer $\exp(tu)$. Remarquons que si n est un endomorphisme nilpotent, $\exp(tn)$ est polynomiale en t.

8.3.3 Exemples de parties denses de $L(E)$

8.22 Proposition. *On suppose que $K = \mathbb{R}$ ou $\mathbb{C}$. Soit E un K-espace vectoriel de dimension finie.*

a) $GL(E)$ est ouvert et dense dans $L(E)$.

b) Si $K = \mathbb{C}$, l'ensemble des endomorphismes diagonalisables est dense dans $L(E)$.

Démonstration. a) L'application déterminant est polynomiale donc continue, donc $GL(E)$ image inverse par det de l'ouvert K^* de K est ouvert dans $L(E)$. Soit $u \in L(E)$. Comme χ_u a un nombre fini de racines, il y a un nombre fini de $k \in \mathbb{N}$ tels que $u_k = u - (1+k)^{-1}\mathrm{id}_E$ soit non inversible. Donc pour k assez grand, $u_k \in GL(E)$, et la suite u_k converge vers u, donc u est dans l'adhérence de $GL(E)$: donc $GL(E)$ est dense dans $L(E)$.

b) On peut trianguler u dans une base $(e_1, \dots, e_n)$ de E. Notons v l'endomorphisme de E tel que $ve_j = je_j$ pour $j \in \{1, \dots, n\}$. Soit $k \in \mathbb{N}^*$ un nombre tel que $\dfrac{n-1}{k}$ soit strictement inférieur à $\inf|\lambda_i - \lambda_j|$, cet « inf » étant pris sur tous les couples de valeurs propres distinctes. Alors les nombres $\lambda_j + \dfrac{j}{k}$ sont deux à deux distincts, donc $u + k^{-1}v$ a toutes ses valeurs propres distinctes : il est diagonalisable. □

8.4 Exercices

8.1 Exercice. Soit E un espace vectoriel de dimension finie. Soit $(E_1, \dots, E_k)$ une famille de sous-espaces vectoriels de E telle que $E = \bigoplus_{j=1}^{k} E_j$. Choisissons une base de E formée de bases de E_j. Caractériser les matrices des endomorphismes pour lesquels les sous-espaces E_j sont stables.

8.2 Exercice. Soient $a_0, \dots, a_n$ des nombres complexes. On définit les matrices $A, J \in M_n(\mathbb{C})$ en posant

$$A = \begin{pmatrix} a_0 & a_1 & a_2 & \dots & a_{n-1} \\ a_{n-1} & a_0 & a_1 & \dots & a_{n-2} \\ a_{n-2} & a_{n-1} & a_0 & \dots & a_{n-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_1 & a_2 & a_3 & \dots & a_0 \end{pmatrix} \qquad \text{et} \qquad J = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ 1 & 0 & 0 & \dots & 0 \end{pmatrix}$$

1. Démontrer que $A = \sum_{k=0}^{n-1} a_k J^k$.
2. Calculer le polynôme minimal et le polynôme caractéristique de J.
3. Diagonaliser J puis A.

Endomorphismes trigonalisables

8.3 Exercice. Soit E un espace vectoriel de dimension finie. Posons $n = \dim E$. Un *drapeau* de E est une suite $(E_k)_{0 \leqslant k \leqslant n}$ de sous-espaces de E telle que pour tout k on ait $E_k \subset E_{k+1}$ et $\dim E_k = k$. Une base $(e_1, \dots, e_n)$ est dite *adaptée* au drapeau $(E_k)_{0 \leqslant k \leqslant n}$ si pour tout k, on a $e_k \in E_k$.

1. Démontrer que toute base de E est adaptée à un et un seul drapeau de E. Démontrer que tout drapeau possède des bases adaptées.
2. Soient $(E_k)_{0\leqslant k\leqslant n}$ un drapeau, $(e_1,\dots,e_n)$ une base adaptée et u un endomorphisme de E. Démontrer que le drapeau $(E_k)_{0\leqslant k\leqslant n}$ est stable par u (*i.e.* tous les E_k sont stables par u) si et seulement si la matrice de u dans la base $(e_1,\dots,e_n)$ est triangulaire supérieure.
3. Soient u un endomorphisme de E et $(E_k)_{0\leqslant k\leqslant n}$ un drapeau stable par u. Notons $\lambda_1,\dots,\lambda_n$ les éléments diagonaux de la matrice de u dans une base $(e_1,\dots,e_n)$ adaptée à $(E_k)_{0\leqslant k\leqslant n}$. Démontrer que $(u-\lambda_k\mathrm{id}_E)(E_k)\subset E_{k-1}$. En déduire une démonstration du théorème de Cayley-Hamilton pour les endomorphismes triangulables.

8.4 Exercice.
1. Soient E un espace euclidien et u un endomorphisme trigonalisable de E. Démontrer qu'il existe une base orthonormale de E dans laquelle la matrice de u est triangulaire.
2. En déduire que l'ensemble des matrices trigonalisables est fermé dans $M_n(\mathbb{R})$.
3. Quelle est l'adhérence de l'ensemble des matrices diagonalisables réelles ?
4. Démontrer que l'intérieur de l'ensemble des matrices diagonalisables est formé des matrices diagonalisables à valeurs propres distinctes.

8.5 Exercice.
1. Soit $A\in M_n(K)$ une matrice carrée. On suppose que son polynôme caractéristique est scindé : $\chi_A=\prod_{k=1}^{n}(\lambda_k-X)$. Soit $Q\in K[X]$. Démontrer que $\chi_{Q(A)}=\prod_{k=1}^{n}(Q(\lambda_k)-X)$.
2. Soit P un polynôme unitaire à coefficients entiers. Soient $\lambda_1,\dots,\lambda_n$ ses racines dans $\mathbb{C}$ comptées avec leur multiplicité. Démontrer que pour tout entier $q\in\mathbb{N}$, le polynôme $\prod_{k=1}^{n}(X-\lambda_k^q)$ est à coefficients entiers.

   Indication : Utiliser une matrice compagnon.

8.6 Exercice. Démontrer qu'un endomorphisme d'un espace vectoriel de dimension finie est trigonalisable si et seulement s'il admet un polynôme annulateur scindé.

8.7 Exercice. Endomorphismes nilpotents. Soient E un espace vectoriel de dimension finie non nulle et u un endomorphisme nilpotent de E. Notons m le plus petit entier tel que $u^m=0$.
1. Quel est le polynôme minimal de u ?
2. Démontrer que u n'est pas surjective.
3. Démontrer que $\operatorname{im} u$ est stable par u. En déduire par récurrence sur $\dim E$ que u est triangulable. Quel est le polynôme caractéristique de u ?
4. Pour $k=0,\dots,m$, posons $N_k=\ker u^k$ et $I_k=\operatorname{im} u^k$ (on a $N_0=\{0\}=I_m$ et $N_m=E=I_0$). Démontrer que la suite N_k est croissante et la suite I_k est décroissante.
5. Démontrer que la suite $(\dim I_k-\dim I_{k+1})=(\dim N_{k+1}-\dim N_k)$ est décroissante.

Endomorphismes cycliques

8.8 Exercice. Soient E un espace vectoriel de dimension finie et u un endomorphisme de E.
1. Soit $x_0\in E$. Pour $k\in\mathbb{N}^*$, on pose $x_k=u^k(x_0)$. Soit $k\in\mathbb{N}^*$ tel que $x_k\in\operatorname{Vect}(x_0,\dots,x_{k-1})$. Démontrer que $\operatorname{Vect}(x_0,\dots,x_{k-1})$ est stable par u ; en déduire que pour tout $\ell\in\mathbb{N}$, on a $x_\ell\in\operatorname{Vect}(x_0,\dots,x_{k-1})$.
2. Un vecteur $x\in E$ est dit *cyclique pour l'endomorphisme* u si $(x,u(x),u^2(x),\dots,u^k(x),\dots)$ engendre E. On dit que u est *cyclique* s'il existe un vecteur $x\in E$ cyclique pour u. Démontrer que u est cyclique si et seulement s'il existe une base de E dans laquelle la matrice de u est une matrice compagnon. Démontrer que dans ce cas $\chi_u=\varpi_u$.

8.9 Exercice. Soient E un espace vectoriel de dimension finie et u un endomorphisme de E. Pour $x \in E$, notons $J_x = \{P \in K[X];\ P(u)(x) = 0\}$.

1. Démontrer que, pour tout $x \in E$, J_x est un idéal de $K[X]$.
2. Démontrer que le vecteur x est cyclique pour u si et seulement si J_x est l'idéal engendré par le polynôme caractéristique de u.

   Écrivons $\varpi_u = \prod_{j=1}^{k} P_j^{\alpha_j}$ la décomposition de ϖ_u en produit de polynômes irréductibles unitaires.

   Pour $j \in \{1, \ldots, k\}$, soit $Q_j \in K[X]$ le polynôme tel que $P_j Q_j = \varpi_u$.
3. Démontrer que pour tout j, il existe $x \in E$ tel que $(Q_j(u))(x) \neq 0$. En déduire qu'il existe $x_j \in E$ tel que $P_j^{\alpha_j}(u)(x_j) = 0$ mais $P_j^{\alpha_j - 1}(u)(x_j) \neq 0$. Démontrer que J_{x_j} est l'idéal engendré par $P_j^{\alpha_j}$.
4. On pose $y = \sum_{j=1}^{k} x_j$. Démontrer que J_y est l'idéal engendré par ϖ_u.
5. Démontrer que u est cyclique si et seulement si $\varpi_u = \chi_u$.

8.10 Exercice. Soient E un espace vectoriel de dimension finie et u un endomorphisme de E.

1. On suppose que u est cyclique. Soient x un vecteur cyclique pour u et F un sous-espace vectoriel de E stable par u.
   a) Démontrer que l'ensemble $\mathcal{J} = \{P \in K[X];\ P(u)(x) \in F\}$ est un idéal dans $K[X]$.
   Il existe donc un unique polynôme unitaire P_F qui engendre $\mathcal{J}$.
   b) Démontrer que P_F divise le polynôme minimal ϖ_u de u (qui est égal au polynôme caractéristique de u - au signe près).
   On écrit $\varpi_u = P_F Q_F$.
   c) Démontrer que $F = \operatorname{im} P_F(u) = \ker Q_F(u)$.
   d) Démontrer que la restriction de u à F est cyclique et que $\dim F = \partial Q_F$.
2. On suppose que le corps K est infini. Démontrer que u est cyclique si et seulement s'il y a un nombre fin de sous-espaces de E stables par u *(on pourra utiliser l'exercice 4.3)*.
3. On suppose que de E n'est pas nul. Démontrer que E ne possède pas de sous-espaces invariants par u autres que $\{0\}$ et E si et seulement si son polynôme caractéristique est irréductible.

Décomposition de Dunford

8.11 Exercice. Quelle est la décomposition de Dunford de la matrice $\begin{pmatrix} 1 & 1 \\ 0 & t \end{pmatrix}$?

8.12 Exercice. Soit $A \in \mathcal{M}_n(\mathbb{R})$. Considérons-la comme matrice à coefficients complexes et soit $A = D + N$ sa décomposition de Dunford (vue comme matrice à coefficients complexes). Démontrer que D et N sont des matrices réelles.

8.13 Exercice. Soient E un espace vectoriel de dimension finie et u un endomorphisme de E ; notons ϖ_u son polynôme minimal. Soit $P \in K[X]$. Démontrer que $P(u)$ est inversible si et seulement si ϖ_u et P sont premiers entre eux et que, dans ce cas, il existe un polynôme $Q \in K[X]$ tel que $P(u)^{-1} = Q(u)$.

8.14 Exercice. Soit u un endomorphisme d'un K-espace vectoriel E de dimension finie.

1. On suppose que le polynôme minimal de u est de la forme P^k où P est un polynôme irréductible de $K[X]$. Soient $x \in E$ non nul et F le sous-espace vectoriel de E engendré par les $u^j(x)$ $(j \in \mathbb{N})$. Notons v l'endomorphisme de F déduit de u par restriction. Démontrer que $\chi_v = \varpi_v$ est une puissance de P. En déduire (à l'aide d'une récurrence) que le polynôme caractéristique de u est une puissance de P.

2. En utilisant le « lemme des noyaux », démontrer que le polynôme minimal et le polynôme caractéristique d'un endomorphisme ont mêmes diviseurs irréductibles.
3. Démontrer que pour $P \in K[X]$ les assertions suivantes sont équivalentes :
   (i) Les polynômes P et χ_u sont premiers entre eux.
   (ii) Les polynômes P et ϖ_u sont premiers entre eux.
   (iii) L'endomorphisme $P(u)$ est inversible.

8.15 Exercice. Nous proposons une autre méthode pour démontrer que χ_u et ϖ_u ont mêmes diviseurs irréductibles (exercice 8.14). Soit P un diviseur irréductible de χ_u. Soit L une extension de K dans laquelle P a une racine. En considérant la matrice de u dans une base comme matrice à coefficients dans L, démontrer que $\det(P(u)) = 0$. En déduire que P divise ϖ_u.

8.16 Exercice. Soient E un espace vectoriel de dimension finie non nulle sur un corps K et u un endomorphisme de E.

1. Soit ϖ le polynôme minimal de u et soit P un polynôme irréductible divisant ϖ. Notons k le degré de P. Démontrer qu'il existe un sous-espace de dimension k stable par u.
2. On suppose que $K = \mathbb{R}$. Démontrer que u possède un sous-espace stable de dimension 1 ou un sous-espace stable de dimension 2.

Exponentielle de matrices

8.17 Exercice. Démontrer que pour tout endomorphisme u d'un espace vectoriel réel (ou complexe) de dimension finie, il existe un polynôme $P \in \mathbb{K}[X]$ tel que $\exp(u) = P(u)$.

8.18 Exercice. Un endomorphisme u est dit *unipotent* si $u - \mathrm{id}$ est nilpotent. Soit E un espace vectoriel réel ou complexe de dimension finie.

1. Démontrer que l'exponentielle d'un endomorphisme nilpotent de E, est un endomorphisme unipotent.
2. Pour $k \in \mathbb{N}$, notons E_k et L_k les polynômes donnés par $E_k = \displaystyle\sum_{j=1}^{k} \frac{X^j}{j!}$ et $L_k = \displaystyle\sum_{j=1}^{k} \frac{(-1)^{j+1} X^j}{j}$.
   Démontrer que $E_k \circ L_k$ et $L_k \circ E_k$ admettent en 0 les développements limités $E_k \circ L_k(x) = x + o(x^k)$ et $L_k \circ E_k(x) = x + o(x^k)$. En déduire que si u est un endomorphisme de E tel que $u^{k+1} = 0$, on a $\exp(L_k(u)) = \mathrm{id} + u$ et $L_k(\exp(u) - \mathrm{id}) = u$.
3. Démontrer que exp est un homéomorphisme de l'ensemble des matrices carrées d'ordre n nilpotentes sur l'ensemble des matrices carrées d'ordre n unipotentes.
4. Démontrer que l'application $\exp : M_n(\mathbb{C}) \to GL_n(\mathbb{C})$ est surjective.

9 Formes quadratiques

9.1 Formes bilinéaires, formes quadratiques

9.1.1 Définitions et généralités

9.1 Définition. Soient K un corps commutatif et E un K-espace vectoriel.
- Rappelons qu'une *forme bilinéaire* sur E est une application $b : E \times E \to K$ linéaire en chacune des variables, *i.e.* telle que, pour tout $x \in E$, les applications $y \mapsto b(x, y)$ et $y \mapsto b(y, x)$ sont des formes linéaires sur E.
- Une forme bilinéaire $b : E \times E \to K$ est dite *symétrique* si pour tout $(x, y) \in E \times E$ on a $b(y, x) = b(x, y)$.
- Une forme bilinéaire $b : E \times E \to K$ est dite *antisymétrique* si pour tout $(x, y) \in E \times E$ on a $b(y, x) = -b(x, y)$.
- Une forme bilinéaire $b : E \times E \to K$ est dite *alternée* si pour tout $x \in E$ on a $b(x, x) = 0$.

9.2 Proposition. *a) Toute forme bilinéaire alternée est antisymétrique.*

b) Si la caractéristique du corps K est différente de 2, on a la réciproque : toute forme bilinéaire antisymétrique est alternée.

c) Si la caractéristique du corps K est différente de 2, toute forme bilinéaire $b : E \times E \to K$ se décompose de manière unique sous la forme $b = b_s + b_a$ où b_s est une forme bilinéaire symétrique et b_a est une forme bilinéaire alternée.

Démonstration. Soient E un K-espace vectoriel et $b : E \times E \to K$ une forme bilinéaire sur E.

a) Pour tout $x, y \in E$, on a $b(x + y, x + y) = b(x, x) + b(y, y) + b(x, y) + b(y, x)$. Si b est alternée, il vient $0 = b(x, y) + b(y, x)$.

b) Si b est antisymétrique, prenant $x = y$, il vient $b(x, x) = -b(x, x)$. Si la caractéristique de K n'est pas 2, il vient $b(x, x) = 0$.

c) **Unicité.** Si $b = b_s + b_a$, il vient $b(x, y) + b(y, x) = 2b_s(x, y)$ et $b(x, y) - b(y, x) = 2b_a(x, y)$, donc
$$b_s(x, y) = \frac{1}{2}(b(x, y) + b(y, x)) \text{ et } b_a(x, y) = \frac{1}{2}(b(x, y) - b(y, x)).$$

Existence. Il suffit de poser $b_s(x, y) = \frac{1}{2}(b(x, y) + b(y, x))$ et $b_a(x, y) = \frac{1}{2}(b(x, y) - b(y, x))$. On vérifie immédiatement que b_s est une forme bilinéaire symétrique, que b_a est une forme bilinéaire alternée et que l'on a $b = b_s + b_a$. □

9.3 Définition. Soient K un corps commutatif et E un K-espace vectoriel. On appelle *forme quadratique* sur E une application $q : E \to K$ telle qu'il existe une forme bilinéaire $b : E \times E \to K$ satisfaisant $q(x) = b(x, x)$ pour tout $x \in E$.

9.4 Proposition. *Soient K un corps commutatif* de caractéristique différente de 2*, E un K-espace vectoriel et q une forme quadratique sur E. Il existe une unique forme bilinéaire* symétrique $\varphi : E \times E \to K$ *satisfaisant $q(x) = \varphi(x, x)$ pour tout $x \in E$.*

Démonstration. Par définition, il existe une forme bilinéaire $b : E \times E \to K$ telle que pour tout $x \in E$ on ait $b(x, x) = q(x)$. Soit $\varphi : E \times E \to K$ une forme bilinéaire symétrique. Alors $\varphi(x, x) = q(x)$ pour tout $x \in E$, si et seulement si la forme bilinéaire $b - \varphi$ est alternée, c'est à dire si et seulement si $\varphi = b_s$ où $b = b_s + b_a$ est la décomposition de b de la proposition ci-dessus (*cf.* c) de la proposition 9.2). □

9.5 Définition. Soient K un corps commutatif *de caractéristique différente de* 2, E un K-espace vectoriel, q une forme quadratique sur E et $\varphi : E \times E \to K$ une forme bilinéaire symétrique. Si pour tout $x \in E$ on a $\varphi(x, x) = q(x)$ on dit que q est la *forme quadratique associée* à φ et que φ est la *forme polaire* de q.

9.6 Identités de polarisation. Soient K un corps commutatif *de caractéristique différente de* 2, E un K-espace vectoriel, q une forme quadratique sur E et $\varphi : E \times E \to K$ sa forme polaire. En développant $\varphi(x \pm y, x \pm y)$ on trouve les *identités de polarisation* :

$$\varphi(x,y) = \frac{1}{2}(q(x+y) - q(x) - q(y)) = \frac{1}{4}(q(x+y) - q(x-y)) \text{ pour tous } x, y \in E.$$

9.7 Définition. Soient K un corps commutatif, E un K-espace vectoriel de dimension finie et $B = (e_1, \ldots, e_n)$ une base de E. Soit b sur E. La *matrice de la forme bilinéaire* b dans la base B est la matrice $A = (a_{i,j})$ où $a_{i,j} = b(e_i, e_j)$. Si la caractéristique de K est différente de 2, on appelle *matrice d'une forme quadratique* dans la base B la matrice (dans la base B) de sa forme polaire.

Soient $x, y \in E$. Notons $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ et $Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$ les vecteurs-colonnes formés des coordonnées de x et y dans la base B. Soit b une forme bilinéaire sur E et notons $A = (a_{i,j})$ sa matrice dans la base B. Par bilinéarité de b, on a $b(x,y) = \sum_{i,j} a_{i,j} x_i y_j = {}^tXAY$. Notons que la matrice A est symétrique (égale à sa transposée) si et seulement si la forme bilinéaire b est symétrique ; de même la matrice A est antisymétrique si et seulement si la forme bilinéaire b est antisymétrique.

Supposons que la caractéristique de K soit différente de 2. Si $A = (a_{i,j})$ est la matrice dans la base B d'une forme quadratique q, on a $q(x) = \sum_{i,j} a_{i,j} x_i x_j = \sum_i a_{i,i} x_i^2 + \sum_{i<j} 2a_{i,j} x_i x_j$.

9.8 Définition. Soient K un corps commutatif, E un K-espace vectoriel de dimension finie et q une forme quadratique sur E. Un vecteur x de E est dit *isotrope* pour q si $q(x) = 0$. L'ensemble des vecteurs isotropes pour q s'appelle le *cône isotrope* de q.

9.1.2 Orthogonalité

Soient K un corps commutatif *de caractéristique différente de* 2, E un K-espace vectoriel et q une forme quadratique sur E. Notons φ sa forme polaire.

- On dit que deux éléments $x, y \in E$ sont *orthogonaux* pour q si $\varphi(x,y) = 0$. On dit qu'une famille $(x_i)_{i \in I}$ est *orthogonale* pour q si pour tout $i, j \in I$ avec $i \neq j$, on a $\varphi(x_i, x_j) = 0$.
- L'orthogonal pour q d'une partie A de E est son orthogonal pour la forme bilinéaire φ (*cf.* 6.29) : c'est l'ensemble $A^\perp = \{y \in E;\ \forall x \in A;\ \varphi(x,y) = 0\}$. C'est un sous-espace vectoriel de E.
- En particulier, l'ensemble $\{x \in E;\ \forall y \in E;\ \varphi(x,y) = 0\}$ est un sous-espace vectoriel de E (c'est l'orthogonal $E^\perp$ de E tout entier). On l'appelle *noyau* de q ou noyau de φ et on le note $\ker q$ ou $\ker \varphi$.
- On dit que la forme quadratique q est *non dégénérée* si $\ker q = \{0\}$. On dit aussi que la forme bilinéaire φ est non dégénérée.

9.9 Définition. Supposons que la dimension de E soit finie et soit q une forme quadratique sur E. On appelle *rang* de q et l'on note $\operatorname{rg} q$ la codimension de $\ker q$.

On a donc $\operatorname{rg} q = \dim E - \dim \ker q$.

Si φ est la forme polaire de q, le rang de q s'appelle aussi le rang de φ et se note aussi $\operatorname{rg} \varphi$.

Supposons que la dimension de E soit finie et soit B une base de E. Notons A la matrice de q dans la base B. Soit $x \in E$ et X le vecteur-colonne formé des coordonnées de x dans la base B. On a $x \in \ker q \iff AX = 0$. Il vient $\operatorname{rg} q = \operatorname{rg} A$.

9.10 Proposition. *Si E est de dimension finie et q est non dégénérée, pour toute forme linéaire $\ell \in E^*$, il existe un unique $x \in E$ tel que l'on ait $\ell(y) = \varphi(x, y)$ pour tout $y \in E$.*

Démonstration. L'application $L : E \to E^*$ qui à $x \in E$ associe la forme linéaire $y \mapsto \varphi(x, y)$ est injective puisque q est non dégénérée, donc bijective puisque $\dim E^* = \dim E$ (*cf.* 6.23). □

Nous utiliserons plus loin le résultat suivant

9.11 Lemme. *Soit F un sous-espace vectoriel de E de dimension finie tel que la restriction de q à F soit non dégénérée. Alors $E = F \oplus F^\perp$.*

Démonstration. Soit $x \in E$. Par la proposition 9.10 (appliquée à la restriction de q à F), il existe un unique $y \in F$ tel que pour tout $z \in F$ on ait $\varphi(x, z) = \varphi(y, z)$, *i.e.* tel que $x - y \in F^\perp$. □

9.1.3 Décomposition de Gauss

Dans toute cette partie on fixe un corps commutatif K *de caractéristique différente de* 2.

9.12 Théorème. *Soit q une forme quadratique sur un K-espace vectoriel de dimension finie E. Alors il existe une base de E orthogonale pour q.*

Démonstration. On raisonne par récurrence sur la dimension n de E.
- Si $n \leqslant 1$ toute base de E est orthogonale.
- Supposons le théorème démontré pour toute forme quadratique sur un K-espace de dimension $n-1$. Si q est nulle, il n'y a rien à démontrer : toute base de E est orthogonale. Sinon, il existe un vecteur $e \in E$ tel que $q(e) \neq 0$. La restriction de q à Ke est alors non dégénérée. Par le lemme 9.11, on a alors $E = Ke \oplus (Ke)^\perp$. D'après l'hypothèse de récurrence, la restriction de q à $(Ke)^\perp$ admet une base orthogonale $(e_1, \ldots, e_{n-1})$. Posons $e_n = e$. La base $(e_1, \ldots, e_n)$ de E est orthogonale. □

Soient E un K-espace vectoriel de dimension finie, q une forme quadratique sur E et $B = (e_1, \ldots, e_n)$ une base de E. La base B est orthogonale pour q si et seulement si la matrice de q dans la base B est diagonale.

Supposons que la base $(e_1, \ldots, e_n)$ soit orthogonale. Quitte à intervertir les éléments de la base, on peut supposer qu'il existe $r \in \{0, \ldots, n\}$ tel que $q(e_i) \neq 0$ pour $i \leqslant r$ et $q(e_i) = 0$ pour $i > r$. Alors $e_{r+1}, \ldots, e_n$ est une base de $\ker q$. En particulier $r = \operatorname{codim} \ker q$. Pour $x = \sum_{i=1}^{n} x_i e_i$ et $y = \sum_{i=1}^{n} y_i e_i$, on a $\varphi(x, y) = \sum_{i=1}^{r} q(e_i) x_i y_i$ et $q(x) = \sum_{i=1}^{r} q(e_i) x_i^2 = \sum_{i=1}^{r} q(e_i) e_i^*(x)^2$ où (e_i^*) désigne la base duale de $(e_1, \ldots, e_n)$.

On a démontré :

9.13 Corollaire : Décomposition de Gauss. *Soit q une forme quadratique sur un K-espace vectoriel de dimension finie E. Alors il existe des formes linéaires* indépendantes *$\ell_1, \ldots, \ell_r$ et des scalaires non nuls $a_1, \ldots, a_r$ tels que $q(x) = \sum_{i=1}^{r} a_i \ell_i(x)^2$. On a $r = \operatorname{rg} q$.*

□

Méthode de Gauss. Donnons-nous une forme quadratique q sur un K-espace vectoriel de dimension finie E. Notons φ sa forme polaire

Notons r le rang de q. Si on trouve une base orthogonale $(f_1, \dots, f_n)$ vérifiant $q(f_i) \neq 0$ pour $i \leqslant r$ et $q(f_i) = 0$ pour $i > r$, on aura $q = \sum_{i=1}^{r} a_i (f_i^*)^2$, où l'on a posé $a_i = q(f_i)$. Remarquons que, pour tout i, les formes linéaires $a_i f_i^*$ et $x \mapsto \varphi(x, f_i)$ coïncident en $x = f_i$ et sont nulles pour $x = f_j$ pour $j \neq i$. Puisqu'elles coïncident sur la base (f_j), elles sont égales.

Partons d'une base quelconque $(e_1, \dots, e_n)$ de E.

- Si $q(e_1) \neq 0$, on va prendre $f_1 = e_1$, donc poser $\ell_1(x) = \dfrac{1}{q(e_1)}\varphi(x, e_1)$. Notons que le noyau de la forme quadratique $q_1 = q - q(e_1)\ell_1^2$ est $\ker q \oplus Ke_1$, donc $\operatorname{rg} q_1 = r - 1$ et q_1 est une combinaison linéaire des formes $e_i^* e_j^*$ pour $2 \leqslant i \leqslant j \leqslant n$, en d'autres termes, on aura une expression de la forme $q_1(\sum_{i=1}^{n} x_i e_i) = \sum_{2 \leqslant i \leqslant j \leqslant n} b_{i,j} x_i x_j$.
  On aura donc à « réduire » une forme quadratique avec une variable de moins.
- Plus généralement, s'il existe k tel que $q(e_k) \neq 0$, nous pourrons appliquer cette recette.
- Si tous les $q(e_k)$ sont nuls, mais q n'est pas nulle, quitte à intervertir les vecteurs de la base, on peut supposer que $\varphi(e_1, e_2) \neq 0$. La restriction de q à $F = Ke_1 \oplus Ke_2$ est non dégénérée. En trouvant une base orthogonale de F (par exemple $(e_1 + e_2, e_1 - e_2)$), il nous restera à « réduire » la forme q restreinte à $F^\perp$, qui est de rang $r - 2$ et qui ne fait intervenir que les variables $(x_3, \dots, x_n)$.

En pratique, notons $A = (a_{i,j})$ la matrice de q dans cette base. Pour $x = \sum_{i=1}^{n} x_i e_i$, on a donc $q(x) = \sum_{1 \leqslant i,j \leqslant n} a_{ij} x_i x_j$.

La méthode expliquée ci-dessus revient à construire pas à pas, les formes ℓ_i de la façon suivante :

a) Si $a_{11} \neq 0$ (*i.e.* $q(e_1) \neq 0$), on écrit

$$\begin{aligned} q(x) &= a_{11}x_1^2 + 2\sum_{i=2}^{n} a_{1i}x_1x_i + \sum_{2 \leqslant i,j \leqslant n} a_{ij}x_ix_j \\ &= a_{11}\big(x_1 + \sum_{i=2}^{n} \frac{a_{1i}}{a_{11}}x_i\big)^2 - \frac{1}{a_{11}}\big(\sum_{i=2}^{n} a_{1i}x_i\big)^2 + \sum_{2 \leqslant i,j \leqslant n} a_{ij}x_ix_j \\ &= a_{1,1}\ell_1(x)^2 + q'(x) \end{aligned}$$

où $\ell_1(x) = x_1 + \sum_{i=2}^{n} \dfrac{a_{1i}}{a_{11}}x_i$ et q_1 est une forme quadratique qui ne fait plus intervenir x_1, *i.e.* telle que $e_1 \in \ker(q_1)$.

- Si l'un des coefficients diagonaux a_{ii} est non nul, on se ramène au premier cas en permutant les éléments de la base.

b) On a $a_{11} = a_{22} = 0$ mais $a_{12} \neq 0$

$$\begin{aligned} q(x) &= 2a_{12}x_1x_2 + 2x_1(\sum_{i=3}^{n} a_{1i}x_i) + 2x_2(\sum_{i=3}^{n} a_{2i}x_i) + \sum_{3 \leqslant i,j \leqslant n} a_{ij}x_ix_j \\ &= 2(a_{12}x_1 + \sum_{i=3}^{n} a_{2i}x_i)(x_2 + \frac{1}{a_{12}}\sum_{i=3}^{n} a_{1i}x_i) - \frac{2}{a_{12}}(\sum_{i=3}^{n} a_{2i}x_i)(\sum_{i=3}^{n} a_{1i}x_i) + \sum_{3 \leqslant i,j \leqslant n} a_{ij}x_ix_j. \\ &= \ell_1(x)\ell_2(x) + q_1(x) \\ &= \ell_1'(x)^2 - \ell_2'(x)^2 + q_1(x) \end{aligned}$$

où $\ell'_1(x) = \dfrac{\ell_1(x) + \ell_2(x)}{2}$ et $\ell'_2(x) = \dfrac{\ell_1(x) - \ell_2(x)}{2}$, et où q_1 est une forme quadratique qui ne fait plus intervenir x_1 et x_2, *i.e.* telle que $e_1, e_2 \in \ker(q_1)$.

9.1.4 Formes quadratiques positives - $K = \mathbb{R}$

On suppose ici que le corps de base K est le corps des réels.

9.14 Définition. Soit E un espace vectoriel réel. Une forme quadratique q sur E est dite *positive* si pour tout $x \in E$ on a $q(x) \geqslant 0$.

Une forme bilinéaire symétrique φ est dite positive si la forme quadratique associée $x \mapsto \varphi(x, x)$ est positive.

9.15 Inégalité de Cauchy-Schwarz. *Soit q une forme quadratique* positive *et φ sa forme polaire. Pour tout $x, y \in E$, on a $\varphi(x, y)^2 \leqslant q(x)q(y)$.*

Démonstration. Pour $t \in \mathbb{R}$ on a $q(tx + y) \geqslant 0$. Or $q(tx + y) = at^2 + 2bt + c$ avec $a = q(x)$, $b = \varphi(x, y)$ et $c = q(y)$. Le trinôme $at^2 + 2bt + c$ garde un signe constant, donc son discriminant $4(b^2 - ac)$ est négatif ou nul, *i.e.* $b^2 \leqslant ac$.

Notons que si $a = 0$, l'application affine $t \mapsto 2bt + c$ ne peut garder un signe constant que si elle est constante *i.e.* si $b = 0$. On trouve encore $b^2 \leqslant ac$. □

9.16 Corollaire. *Le cône isotrope d'une forme quadratique* positive *est égal à son noyau.*

Démonstration. Soient q une forme quadratique positive sur un espace vectoriel réel E et x un vecteur isotrope pour q. Notons φ la forme polaire de q. Pour tout $y \in E$, on a, d'après l'inégalité de Cauchy-Schwarz, $\varphi(x, y)^2 \leqslant q(x)q(y) = 0$. Donc $\varphi(x, y) = 0$. Cela prouve que $x \in \ker q$. □

En particulier, une forme quadratique positive non dégénérée n'admet pas de vecteur isotrope non nul : on dit qu'elle est *anisotrope* ou *définie*.

Rappelons une conséquence importante de l'inégalité de Cauchy-Schwarz.

9.17 Théorème. *Soit E un espace vectoriel réel de dimension finie muni d'une forme bilinéaire symétrique positive et non dégénérée $(x, y) \mapsto \langle x|y\rangle$. L'application $x \mapsto \|x\| = \sqrt{\langle x|x\rangle}$ est une norme sur E.*

Démonstration. Soient $x, y \in E$ et $\lambda \in \mathbb{R}$.
- D'après le corollaire précédent, on a $\|x\| = 0 \iff x = 0$;
- il est clair que $\|\lambda x\| = |\lambda|\|x\|$;
- On a, en utilisant l'inégalité de Cauchy-Schwarz

$$\|x + y\|^2 = \|x\|^2 + \|y\|^2 + 2\langle x|y\rangle \leqslant \|x\|^2 + \|y\|^2 + 2\|x\|\|y\| = (\|x\| + \|y\|)^2,$$

d'où l'inégalité triangulaire. □

9.1.5 Signature ($K = \mathbb{R}$)

9.18 Théorème d'inertie de Sylvester. *Soient E un espace vectoriel* réel *de dimension finie et q une forme quadratique sur E. Donnons nous deux bases q-orthogonales $(e_1, \dots, e_n)$ et $(f_1, \dots, f_n)$. Le nombre des $i \in \{1, \dots, n\}$ tels que $q(e_i) > 0$ (resp. $q(e_i) < 0$, $q(e_i) = 0$) est égal au nombre des $i \in \{1, \dots, n\}$ tels que $q(f_i) > 0$ (resp. $q(f_i) < 0$, $q(f_i) = 0$).*

Démonstration. Notons E_+ (*resp.* E_-, E_0) le sous-espace vectoriel de E engendré par les e_j tels que $q(e_j) > 0$ (*resp.* $q(e_j) < 0$, $q(e_j) = 0$). De même, notons F_+ (*resp.* F_-, F_0) le sous-espace vectoriel de E engendré par les f_j tels que $q(f_j) > 0$ (*resp.* $q(f_j) < 0$, $q(f_j) = 0$). Remarquons que l'on a $E = E_+ \oplus E_- \oplus E_0 = F_+ \oplus F_- \oplus F_0$. Si $x \in E_+$ n'est pas nul, on a $q(x) > 0$; si $x \in F_- \oplus F_0$, on a $q(x) \leqslant 0$. Il vient $E_+ \cap (F_- \oplus F_0) = \{0\}$, donc $\dim E_+ \leqslant \operatorname{codim}(F_- \oplus F_0) = \dim F_+$.

De même, on a
- $F_+ \cap (E_- \oplus E_0) = \{0\}$, donc $\dim F_+ \leqslant \dim E_+$;
- $E_- \cap (F_+ \oplus F_0) = \{0\}$, donc $\dim E_- \leqslant \dim F_-$;
- $F_- \cap (E_+ \oplus E_0) = \{0\}$, donc $\dim F_- \leqslant \dim E_-$. □

Quelques remarques à propos de cette démonstration. Si $(e_1, \dots, e_n)$ est une base de E orthogonale pour q,

a) on a vu que les e_j tels que $q(e_j) = 0$ engendrent le noyau de q - autrement dit, avec les notations ci-dessus, on a $E_0 = F_0 = \ker q$;

b) cette démonstration démontre que si F est un sous-espace de E tel que la restriction de q à F soit non dégénérée et positive (*resp.* non dégénérée et négative), on a $\dim F \leqslant \dim E_+$ (*resp.* $\dim F \leqslant \dim E_-$).

9.19 Définition. Soient E un espace vectoriel réel de dimension finie et q une forme quadratique sur E. On appelle *signature* de q le couple (k, ℓ) où k et ℓ désignent respectivement le nombre des e_j tels que $q(e_j) > 0$ et $q(e_j) < 0$ pour une base de E orthogonale pour q.

9.20 Conséquence : recherche d'extréma locaux. Soient U un ouvert d'un espace vectoriel réel E de dimension finie, a un point de U et $f : U \to \mathbb{R}$ une application. Rappelons que f est différentiable en a si et seulement si elle admet au voisinage de a un développement limité de la forme $f(x) = f(a) + \ell(x - a) + o(\|x - a\|)$ où ℓ est une forme linéaire sur E (la différentielle df_a de f en a). Si elle est deux fois différentiable en a, elle admet au voisinage de a un développement limité de la forme $f(x) = f(a) + \ell(x - a) + q(x - a) + o(\|x - a\|^2)$ où ℓ est une forme linéaire et q une forme quadratique sur E ($q/2$ est la différentielle seconde de f). On a :

a) Si f admet un extrémum local en a et est différentiable en a, alors $df_a = 0$.

b) Supposons que f admette au voisinage de a un développement limité d'ordre 2 de la forme $f(x) = f(a) + q(x - a) + o(\|x - a\|^2)$.
   - Si f présente en a un minimum (*resp.* maximum) local, alors la forme quadratique q est positive (*resp.* négative).
   - Si la forme quadratique q est définie positive (*resp.* définie négative), alors f présente en a un minimum (*resp.* maximum) local.

9.2 Formes quadratiques sur un espace vectoriel euclidien

9.2.1 Bases orthonormales

Soit E un espace vectoriel euclidien. Autrement dit E est un espace vectoriel réel de dimension finie muni d'une forme bilinéaire symétrique positive et non dégénérée $\langle \ | \ \rangle$. Rappelons que E possède une *base orthonormale* (on dit aussi *base orthonormée*), *i.e.* une base orthogonale (e_i) telle que, pour tout

i on ait $\langle e_i|e_i\rangle = 1$. L'existence d'une base orthonormale résulte immédiatement de l'existence d'une base orthogonale (théorème 9.12). En effet, si $(f_1, \ldots, f_n)$ est une base orthogonale, on a $\langle f_i|f_i\rangle \in \mathbb{R}_+^*$. Posons $e_i = \langle f_i|f_i\rangle^{-1/2} f_i$; la base $(e_1, \ldots, e_n)$ est orthonormale.

Pour construire des bases orthonormales, on utilise le procédé d'orthonormalisation de Gram-Schmidt :

9.21 Orthonormalisation de Gram-Schmidt. *Soit $(x_1, ..., x_n)$ base de E. Il existe une unique base orthogonale $(f_1, ..., f_n)$ et une unique base orthonormale $(e_1, ..., e_n)$ telles que pour tout $k \in \{1, \ldots, n\}$ on ait :*

a) $\mathrm{Vect}(x_1, \ldots, x_k) = \mathrm{Vect}(f_1, \ldots, f_k) = \mathrm{Vect}(e_1, \ldots, e_k)$

b) $x_k - f_k \in \mathrm{Vect}(x_1, \ldots, x_{k-1})$ *(en particulier $f_1 = x_1$) et $\langle x_i|e_i\rangle \geqslant 0$ (en fait $\langle x_i|e_i\rangle > 0$).*

On construit les f_i et les e_i de la manière suivante :

$$f_1 = u_1\,; \quad f_2 = u_2 - \frac{\langle f_1|u_2\rangle}{\langle f_1|f_1\rangle} f_1\,; \quad f_k = u_k - \sum_{j=1}^{k-1} \frac{\langle f_j|u_k\rangle}{\langle f_j|f_j\rangle} f_j\,; \quad e_k = \|f_k\|^{-1} f_k.$$

9.2.2 Endomorphismes et formes bilinéaires

9.22 Proposition. *Pour toute forme bilinéaire φ sur E, il existe un unique endomorphisme f de E tel que, pour tout $(x, y) \in E^2$ on ait $\varphi(x, y) = \langle f(x)|y\rangle$.*

Démonstration. Pour $x \in E$, notons ℓ_x l'application linéaire $y \mapsto \langle x|y\rangle$. Puisque $\langle\ |\ \rangle$ est non dégénérée, l'application linéaire $x \mapsto \ell_x$ est injective : c'est une bijection de E sur E^* puisque $\dim E = \dim E^*$.

Soient φ une forme bilinéaire sur E et $x \in E$. L'application $y \mapsto \varphi(x, y)$ est linéaire ; il existe donc un unique $z_x \in E$ tel que pour tout $y \in E$ on ait $\varphi(x, y) = \langle z_x|y\rangle$; on vérifie aisément que l'application $f : x \mapsto z_x$ est linéaire. □

9.23 Proposition. *Soient E un espace vectoriel euclidien et f un endomorphisme de E. Il existe un unique endomorphisme f^* de E tel que pour $x, y \in E$ on ait $\langle x|f(y)\rangle = \langle f^*(x)|y\rangle$.*

Démonstration. La forme $(x, y) \mapsto \langle x|f(y)\rangle$ est bilinéaire, donc s'écrit sous la forme $\langle f^*(x)|y\rangle$. □

9.24 Définition. Soient E un espace euclidien et f un endomorphisme de E. L'unique f^* de E tel que pour $x, y \in E$ on ait $\langle x|f(y)\rangle = \langle f^*(x)|y\rangle$ s'appelle l'*adjoint* de f. On dit que f est *symétrique* ou *autoadjoint* si $f^* = f$; on dit qu'il est *orthogonal* s'il est bijectif et $f^* = f^{-1}$; on dit qu'il est *normal* si $f^* \circ f = f \circ f^*$.

Remarquons que tout endomorphisme symétrique est normal et tout endomorphisme orthogonal est normal.

9.25 Proposition. *Soient E un espace euclidien et f un endomorphisme de E. La matrice de f^* dans une base orthonormale B de E est la transposée de la matrice de E dans la base B.*

Démonstration. Écrivons $B = (e_1, \ldots, e_n)$. Notons $(a_{i,j})$ la matrice de f dans la base B. On a $f(e_j) = \sum_{i=1}^{n} a_{i,j} e_i$, donc $a_{i,j} = \langle f(e_j)|e_i\rangle$. La matrice de f^* dans la base B est donc $(c_{i,j})$ avec $c_{i,j} = \langle f^*(e_j)|e_i\rangle = \langle e_j|f(e_i)\rangle = a_{j,i}$. C'est la transposée de $(a_{i,j})$. □

9.26 Remarque. Soit q une forme quadratique d'un espace euclidien E et φ sa forme polaire. Il existe un unique endomorphisme f symétrique tel que, pour tout $x, y \in E$ on ait $\varphi(x,y) = \langle f(x)|y\rangle$. En particulier, $q(x) = \langle f(x)|x\rangle$.

L'application qui à un endomorphisme symétrique f associe la forme bilinéaire symétrique $(x,y) \mapsto \langle f(x)|y\rangle$ et l'application qui à une forme bilinéaire symétrique φ associe la forme quadratique $x \mapsto \varphi(x,x)$ sont bijectives, donc l'application qui à un endomorphisme symétrique f associe la forme quadratique $x \mapsto \langle f(x)|x\rangle$ est bijective.

Notons que si un endomorphisme symétrique f, une forme bilinéaire symétrique φ et une forme quadratique q se correspondent à travers ces bijections, on a $\ker f = \ker \varphi = \ker q$, donc $\operatorname{rg} f = \operatorname{rg} \varphi = \operatorname{rg} q$.

9.2.3 Diagonalisation simultanée

9.27 Théorème. *Soient E un espace vectoriel euclidien et q une forme quadratique sur E. Il existe une base orthonormale de E qui soit orthogonale pour q.*

Démonstration. On raisonne par récurrence sur la dimension n de E.
- Si $n = 1$, toute base est orthogonale pour toute forme quadratique !
- Supposons $n \geqslant 2$ et le résultat démontré pour toute forme bilinéaire symétrique d'un espace euclidien de dimension $n-1$.
  Soit S la sphère unité de E. C'est une partie compacte de E. L'application continue q atteint son maximum en un vecteur $e_1 \in S$. Posons $\lambda = q(e_1)$ et $q_1(x) = \lambda\langle x|x\rangle - q(x)$. La forme polaire φ_1 de q_1 est donnée par $\varphi_1(x,y) = \lambda\langle x|y\rangle - \varphi(x,y)$ où φ est la forme polaire de q. Notons enfin H l'orthogonal de e_1 pour le produit scalaire de E.
  Par définition de e_1, la forme q_1 est positive et e_1 est isotrope pour q_1. Il appartient donc au noyau de q_1 : pour tout $y \in E$ on a $\varphi_1(e_1, y) = 0$, soit $\varphi(e_1, y) = \lambda\langle e_1|y\rangle$. En particulier, pour $y \in H$ on trouve $\varphi(e_1, y) = 0$.
  Par l'hypothèse de récurrence, il existe une base orthonormale $(e_2, \ldots, e_n)$ de H qui soit orthogonale pour la restriction de q à H. La base $(e_1, \ldots, e_n)$ convient. □

On peut reformuler le théorème de façon plus abstraite :

9.28 Corollaire : Diagonalisation simultanée « abstraite ». *Soient E un espace vectoriel réel de dimension finie et q_1, q_2 deux formes quadratiques avec q_1 définie positive. Il existe une base orthonormale pour q_1 qui soit orthogonale pour q_2.*

Démonstration. Muni de la forme quadratique q_1, E est un espace euclidien. On peut donc appliquer directement le théorème. □

9.2.4 Diagonalisation des endomorphismes symétriques

Soit f un endomorphisme symétrique de E ; posons $\varphi(x,y) = \langle f(x)|y\rangle$. Par la diagonalisation simultanée, il existe une base orthonormale $(e_1, \ldots, e_n)$ qui soit orthogonale pour φ. On a donc $\langle f(e_i)|e_j\rangle = 0$ pour $i \neq j$, donc la matrice de f est diagonale dans la base $(e_1, \ldots, e_n)$. On a donc :

9.29 Théorème. *Tout endomorphisme autoadjoint d'un espace euclidien se diagonalise dans une base orthonormale.*

□

Nous trouverons en exercice (9.3.9.13) une généralisation de ce théorème pour les endomorphismes normaux.

Donnons une autre démonstration de ce théorème. Nous allons raisonner par récurrence. Notons P_n la propriété : Tout endomorphisme autoadjoint d'un espace euclidien de dimension n se diagonalise dans une base orthonormale.

Soit donc T un endomorphisme autoadjoint d'un espace de dimension n.

Pour $n = 1$, il n'y a rien à démontrer : tout endomorphisme est diagonal dans toute base !

Supposons donc que $n \geqslant 2$ et que P_{n-1} soit vraie.

Nous utilisons le résultat suivant :

9.30 Lemme. *Soit F un sous-espace de E invariant par T. Alors :*

a) L'orthogonal $F^\perp$ de F est invariant par T.

b) La restriction de T à F et $F^\perp$ est auto-adjointe.

Démonstration. a) Pour $x \in F^\perp$ et $y \in F$, on a - puisque T est autoadjoint $\langle T(x)|y\rangle = \langle x|T(y)\rangle = 0$ puisque $y \in F$ et F est invariant. Cela prouve que $T(x)$ est orthogonal à tout $y \in F$: donc $T(x) \in F^\perp$.

b) Pour $x, y \in E$ on a $\langle T(x)|y\rangle = \langle x|T(y)\rangle$. Cela est *a fortiori* vrai pour $x, y \in F$, ou pour $x, y \in F^\perp$. □

Pour finir la démonstration du théorème par récurrence, il suffit de démontrer que tout endomorphisme autoadjoint d'un espace vectoriel euclidien non nul possède un vecteur propre e_1 que l'on peut supposer de norme 1. Nous poserons alors $F = \mathbb{R}e_1$. Par l'hypothèse de récurrence, la restriction de f à $F^\perp$ admet une base orthonormale $(e_2, \ldots, e_n)$ de vecteurs propres. Alors la base $(e_1, e_2, \ldots, e_n)$ est orthonormale et formée de vecteurs propres de T.

Pour $n = 2$, choisissons une base orthonormale (e_1, e_2) et écrivons la matrice de T dans cette base : $A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$. Le polynôme caractéristique de A est $X^2 - (a+c)X + (ac - b^2)$ dont le discriminant est $(a+c)^2 - 4ac + 4b^2 = (a-c)^2 + 4b^2$. Il est positif ou nul, donc T possède un vecteur propre.

Pour finir, on utilise le résultat suivant :

9.31 Lemme. *Soient E un espace vectoriel réel non nul de dimension finie. Tout endomorphisme de E possède un sous-espace stable de dimension 1 ou un sous-espace stable de dimension 2.*

Démonstration. Soit u un endomorphisme de E. Notons $M \in \mathcal{M}_n(\mathbb{R})$ la matrice de u dans une base quelconque. La matrice M possède des valeurs propres dans $\mathbb{C}$. Il existe donc $\lambda = s + it \in \mathbb{C}$ (avec $s, t \in \mathbb{R}$) et un vecteur-colonne $Z \in \mathbb{C}^n$ non nul tel que $MZ = \lambda Z$. Écrivons $Z = X + iY$ où X et Y sont des vecteurs-colonne réels. On a $M(X + iY) = (s + it)(X + iY)$, soit $MX = sX - tY$ et $MY = tX + sY$. Cela prouve que le sous-espace F engendré par les vecteurs x et y de composantes X et Y est stable par u. Comme $Z \neq 0$ ce sous-espace F n'est pas nul. Comme il est engendré par x, y, on a $\dim F \leqslant 2$. □

9.32 Interprétation matricielle. *Si A est une matrice symétrique (réelle), il existe une matrice orthogonale U et une matrice diagonale D telles que $A = U^{-1}DU$.*

9.2.5 Conséquences géométriques : quadriques

Généralités sur les quadriques : E est un K espace vectoriel

a) Définition des quadriques :
Une quadrique est un sous-ensemble D de l'espace E admettant une équation du type $\psi(x) = 0$ où $\psi(x) = q(x) + \ell(x) + c$, avec q forme quadratique, ℓ forme linéaire et $c \in \mathbb{R}$.

b) Quadriques non dégénérées.
Posons $K = \ker q \cap \ker \ell$ et soit $p : E \to E$ un projecteur de noyau K.
Pour $x \in E$ et $y \in K$, on a $\psi(x+y) = \psi(x)$ de sorte que $\psi = \psi \circ p$. Dans ce cas notre quadrique s'écrit $D_1 \times K$ où D_1 est une quadrique d'un supplémentaire de K (l'image de p).
On dira que la quadrique est *non dégénérée* si $K = \{0\}$. Comme $\dim \ker \ell \geqslant n-1$, cela impose $\dim \ker q \leqslant 1$.

c) Centre d'une quadrique non dégénérée.
On cherche les symétries centrales laissant invariante (l'équation d') une quadrique.
Soit $a \in E$. On devra avoir $\psi(a-x) = \psi(x+a)$. Or $q(x+a) - q(a-x) = 4\varphi(x,a)$ où φ est la forme polaire de q. On a donc $\psi(a+x) - \psi(a-x) = 4\varphi(x,a) + 2\ell(x)$.
Si q est non dégénérée, il existe un unique a tel que pour tout x on ait $\varphi(x,a) = -\frac{1}{2}\ell(x)$.
Si q est dégénérée, prenant $x \in \ker q$ non nul (et donc $x \notin \ker \ell$ car notre quadrique est non dégénérée) on ne peut avoir $4\varphi(x,a) + 2\ell(x) = 0$.

Symétries d'une quadrique, axes principaux d'une quadrique à centre *(juste quelques mots...)*

En se plaçant au centre d'une quadrique à centre, l'équation devient $\psi(x) = q(x) + c = 0$. On dira alors qu'elle est *propre* si elle est non dégénérée et ne contient pas son centre, *i.e.* si q est non dégénérée et $c \neq 0$. L'équation devient donc $q(x) = 1$ avec q non dégénérée, soit $\langle x | f(x) \rangle = 1$ où f est un endomorphisme symétrique inversible.

Les droites propres de f s'appellent les *axes principaux* de notre quadrique.

En diagonalisant, l'équation s'écrira $\displaystyle\sum_{i=1}^{k} \frac{x_i^2}{a_i^2} - \sum_{i=k+1}^{n} \frac{x_i^2}{a_i^2} = 1$ (avec $k \geqslant 1$ si notre quadrique n'est pas vide).

Pour $k = n$, notre quadrique est un ellipsoïde de *demi axes principaux* a_i.

On peut chercher les symétries de notre quadrique, *i.e.* les isométries g de l'espace la laissant invariante. On devra avoir $q \circ g = q$, soit $g^* \circ f \circ g = f$, soit encore $g \circ f = f \circ g$. Si toutes les valeurs propres de f sont distinctes, g devra être diagonale, et comme c'est une isométrie, ses valeurs diagonales sont des ± 1.

Dans le cas général, g devra fixer les espaces propres de f. Sa matrice sera donc diagonale par blocs avec des blocs qui représentent des isométries des espaces propres de f.

9.3 Exercices

9.1 Exercice. Soit K un corps commutatif de caractéristique différente de 2. Soient E un K-espace vectoriel, q une forme quadratique sur E et $(e_1, \ldots, e_n)$ une base de E orthogonale pour q.

1. Soit $i \in \{1, \ldots, n\}$ tel que e_i soit isotrope. Démontrer que $e_i \in \ker q$.
2. Posons $J = \{j \in \{1, \ldots, n\};\ e_j \text{ isotrope}\}$. Démontrer que $(e_j)_{j \in J}$ est une base de $\ker q$.

9.2 Exercice. Quelle est la dimension de l'espace vectoriel des formes quadratiques sur K^n ?

9.3 Exercice. Notons Q l'espace vectoriel des formes quadratiques sur $\mathbb{R}^n$ et S la sphère unité de l'espace euclidien $\mathbb{R}^n$.

1. Démontrer que l'application $N : q \mapsto \sup\{|q(x)|;\ x \in S\}$ est une norme sur Q.
2. Démontrer que les formes quadratiques non dégénérées sur $\mathbb{R}^n$ forment un ouvert Q^* dense dans Q.

3. Démontrer que les formes quadratiques signature $(p, n-p)$ forment un ouvert dans Q.
4. Quelles sont les composantes connexes de Q^* ?

9.4 Exercice. Soit E un espace vectoriel réel de dimension finie n et q une forme quadratique non dégénérée sur E.

1. Notons (r, s) la signature de q. Quelles sont les signatures possibles des restrictions de q à des hyperplans de E ?
2. On se donne une suite $(E_k)_{0\leqslant k\leqslant n}$ de sous-espaces de E tels que $\dim E_k = k$ et $E_k \subset E_{k+1}$. On suppose que pour tout k la restriction q_k de q à E_k est non dégénérée.
   a) Démontrer qu'il existe une base orthogonale $(e_1, \ldots, e_n)$ de E telle que pour $k \in \{1, \ldots, n\}$, les vecteurs $(e_1, \ldots, e_k)$ forment une base de E_k.
   b) Notons A_k la matrice de q_k dans une base de E_k. Démontrer que la signature de q est $(n-\ell, \ell)$ où ℓ est le nombre de changements de signes dans la suite $(\det A_k)_{0\leqslant k\leqslant n}$ (le signe de $\det A_k$ ne dépend pas de la base choisie.)

9.5 Exercice. Nature de la quadrique d'équation $xy + yz + zx + 1 = 0$.

9.6 Exercice. Démontrer que l'application $M \mapsto Tr(M^2)$ est une forme quadratique non dégénérée sur $M_n(\mathbb{R})$. Quelle est sa signature ?

9.7 Exercice. Soit q une forme quadratique sur un espace vectoriel réel. Quelle est en fonction de la signature de q la plus grande dimension de sous-espace totalement isotrope de E (*i.e.* sous-espace vectoriel de E formé de vecteurs isotropes) ?

9.8 Exercice. Soient K un corps de caractéristique différente de 2, E un K-espace vectoriel de dimension finie et q une forme quadratique sur E.

1. Soit F un sous espace vectoriel de E démontrer que $\dim F^\perp = \dim E - \dim F + \dim(F \cap \ker q)$.
2. Plus généralement, soient F et G deux sous-espaces de E. Démontrer que

$$\dim G - \dim(F^\perp \cap G) = \dim F - \dim(F \cap G^\perp).$$

9.9 Exercice. Soient K un corps de caractéristique différente de 2, E un K-espace vectoriel de dimension finie et q une forme quadratique sur E.

Un sous-espace F de E est dit *totalement isotrope* si la restriction de q à F est nulle. Un sous-espace F de E est dit *totalement isotrope maximal* s'il est totalement isotrope et s'il n'y a pas de sous-espace de E totalement isotrope contenant F et distinct de F. Le but de cet exercice est de démontrer que tous les sous-espaces totalement isotropes maximaux de E ont même dimension.

Fixons un sous-espace totalement isotrope maximal F.

1. Soit $x \in F^\perp$ un vecteur isotrope. Démonter que $x \in F$.
2. Soit G un autre sous-espace totalement isotrope maximal. Démontrer que $F^\perp \cap G = F \cap G$.
3. Conclure à l'aide de l'exercice 9.8.

9.10 Exercice. Théorème de Witt. Soient E un espace vectoriel de dimension finie sur un corps K de caractéristique $\neq 2$ et q une forme quadratique non dégénérée sur E. Notons $O(q)$ le groupe orthogonal de q, c'est à dire l'ensemble des applications linéaires bijectives $\tau : E \to E$ telles que $q \circ \tau = q$.

Le but de cet exercice est d'établir le théorème de Witt qui affirme que pour tout sous-espace vectoriel F de E toute application linéaire injective $\sigma : F \to E$ satisfaisant $q \circ f(x) = q(x)$ pour tout $x \in F$, il existe $\tau \in O(q)$ qui prolonge σ.

On procède par récurrence sur $\dim E$.

1. Examiner le cas où $\dim E = 1$ ou $\dim F = 0$.

   On suppose à présent que $\dim E = n \geqslant 2$, que $\dim F \neq 0$ et le résultat établi pour un espace vectoriel de dimension $n-1$.
2. On suppose qu'il existe $x \in F$ tel que $q(x) \neq 0$ et $\sigma(x) = x$. Posons $E_1 = x^\perp$, $F_1 = F \cap E_1$.
   a) Démontrer que $\sigma(F_1) \subset E_1$.
   b) Démontrer qu'il existe $\tau \in O(q)$ qui prolonge σ.
3. On suppose que F n'est pas totalement isotrope. Soit $x \in F$ tel que $q(x) \neq 0$ et posons $y = \sigma(x)$.
   a) Calculer $q(x+y)+q(x-y)$ et en déduire que $x+y$ et $x-y$ ne sont pas tous deux isotropes.
   b) Démontrer qu'il existe un sous-espace $G \subset E$ tel que l'on ait $E = G \oplus G^\perp$, $x+y \in G$ et $x-y \in G^\perp$
   c) En déduire qu'il existe $\tau_1 \in O(q)$ telle que $\tau_1(x) = y$.
   d) Conclure dans ce cas.
4. On suppose que F est totalement isotrope. Notons φ la forme polaire de q. Soit ℓ une forme linéaire non nulle sur F.
   a) Démontrer qu'il existe $x \in E$ tel que, pour tout $y \in F$, on ait $\ell(y) = \varphi(x, y)$.
   b) Démontrer que l'élément x de la question précédente peut être choisi isotrope.
   c) Démontrer que l'on peut prolonger σ en une application linéaire injective $\overline{\sigma} : F \oplus Kx \to E$ telle que l'on ait encore $q(\overline{\sigma}(z)) = q(z)$ pour tout $z \in F \oplus Kx$.
   d) Conclure.

9.11 Exercice. On note $M_+(n, \mathbb{R})$ des matrices carrées d'ordre n symétriques positives. Démontrer que l'application $T \mapsto T^2$ est bijective de $M_+(n, \mathbb{R})$ dans lui même.

9.12 Exercice. Démontrer que pour tout $A \in GL_n(\mathbb{K})$ il existe une unique matrice U orthogonale (unitaire) et une unique matrice triangulaire supérieure T dont les coefficients diagonaux sont (réels et) strictement positifs telles que $A = UT$.

9.13 Exercice. ***Réduction des endomorphismes normaux.*** Une matrice $M \in M_n(\mathbb{R})$ est dite *normale* si ${}^tMM = M{}^tM$. En particulier, les matrices symétriques, les matrices antisymétriques et les matrices orthogonales sont normales. Soit M une matrice normale.

1. On suppose que $n = 2$. Démontrer que M est soit symétrique soit de la forme $\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$ $(a, b \in \mathbb{R})$ - (*i.e.* une matrice de similitude directe).
2. On suppose que M se décompose par blocs sous la forme $M = \begin{pmatrix} A & B \\ 0 & C \end{pmatrix}$ où A et C sont des matrices carrées. Démontrer que ${}^tAA = A{}^tA + B{}^tB$ et en déduire (à l'aide d'un calcul de trace) que $B = 0$, puis que A et C sont des matrices normales.
3. À l'aide du lemme 9.31 démontrer qu'il existe une matrice orthogonale U telle que ${}^tUMU$ s'écrive $M = \begin{pmatrix} D & 0 \\ 0 & D_1 \end{pmatrix}$ où $D = \mathrm{diag}(\lambda_i)$ est diagonale et $D_1 = \mathrm{diag}(S_i)$ est diagonale par blocs 2×2, les S_i étant des matrices de similitudes directes.
4. Rappelons qu'un endomorphisme u d'un espace euclidien E est dit *normal* si $uu^* = u^*u$. Énoncer un théorème de réduction des endomorphismes normaux. En déduire des théorèmes de réduction pour les endomorphismes orthogonaux et pour les endomorphismes antisymétriques.

10 Géométrie affine en dimension finie

10.1 Espaces affines, sous-espaces affines

Soit K un corps. Un espace affine est un espace vectoriel dont on aurait perdu le vecteur nul...

Rappelons pour commencer quelques Définitions sur les actions de groupes.

10.1 Définition. Soient G un groupe et X un *G-espace*, *i.e.* un ensemble muni d'une *action* ou *opération* de G. Notons e l'élément neutre de G.
- On dit que l'action est *libre* ou que X est un *G-espace principal* si pour tout $x \in X$ et $g \in G$, on a $gx = x \iff g = e$;
- on dit que l'action est *transitive* si pour tout $(x, y) \in X^2$, il existe $g \in G$ tel que $gx = y$. Si $X \neq \emptyset$ on dit aussi que X est un *G-espace homogène.*

L'application $(g, x) \mapsto (gx, x)$ de $G \times X$ dans $X \times X$ est injective si et seulement si l'action est libre et surjective si et seulement si l'action est transitive.

10.2 Définitions et conventions.

Espace affine. Un *espace affine* est un ensemble non vide E muni d'une action libre et transitive d'un espace vectoriel $\overrightarrow{E}$. L'espace vectoriel $\overrightarrow{E}$ s'appelle l'*espace vectoriel associé* à E, ou la *direction* de E. On dit que E est de dimension finie si $\overrightarrow{E}$ est de dimension finie et on pose $\dim E = \dim \overrightarrow{E}$. Une droite (*resp.* un plan) affine est un espace affine de dimension 1 (*resp.* 2).

Translations. L'action d'un vecteur $\vec{u} \in \overrightarrow{E}$ sur un point A de E se note $A+\vec{u}$ ou $T_{\vec{u}}(A)$. L'application $T_{\vec{u}}$ s'appelle la *translation* de vecteur $\vec{u}$. On a bien sûr $T_{\vec{u}} \circ T_{\vec{v}} = T_{\vec{u}+\vec{v}}$.

Notation $\overrightarrow{AB}$. Si $A, B \in E$, l'unique élément $\vec{u}$ de $\overrightarrow{E}$ tel que $A + \vec{u} = B$ se note $\overrightarrow{AB}$.

Relation de Chasles. Pour $A, B, C \in E$, on a $\overrightarrow{AB} + \overrightarrow{BC} = \overrightarrow{AC}$, $\overrightarrow{AA} = \overrightarrow{0}$.

Parallélogramme. Pour A, B, C, D, on a $\overrightarrow{AB} = \overrightarrow{DC}$ si et seulement si $\overrightarrow{AD} = \overrightarrow{BC}$. On dit alors que A, B, C, D est un *parallélogramme.*

Sous-espace affine. Soit E un espace affine ; notons $\overrightarrow{E}$ sa direction. Soit $\overrightarrow{F}$ un sous-espace vectoriel de $\overrightarrow{E}$. On appelle *sous-espace affine* de direction $\overrightarrow{F}$ une orbite de l'action de $\overrightarrow{F}$ dans E ; c'est alors un espace affine de direction $\overrightarrow{F}$. Un hyperplan affine est un sous-espace affine de co-dimension 1.

Il est d'usage de considérer que l'ensemble vide n'est pas un espace affine, mais que c'est un sous-espace affine dont tout sous-espace vectoriel est une direction. Avec cette convention, l'intersection d'une famille quelconque $(F_i)_{i\in I}$ de sous-espaces affines de E est un sous-espace affine de direction $\bigcap_{i\in I} \overrightarrow{F_i}$.

Sous-espace affine engendré. Soit E un espace affine et P une partie de E. Il existe un plus petit sous-espace affine de E contenant P (l'intersection de tous les sous-espaces affines de E contenant P).

Parallélisme. Deux sous-espaces affines ayant même direction sont dits *parallèles.* On dit encore qu'un sous-espace affine F est parallèle à un sous-espace affine G si la direction de F est contenue dans celle de G.

10.2 Applications affines

10.3 Définition. Soient E et F des espaces affines. Une application $f : E \to F$ est dite *affine* s'il existe une application linéaire $\vec{f} : \overrightarrow{E} \to \overrightarrow{F}$ telle que, pour tout $A, B \in E$ on ait $\overrightarrow{f(A)f(B)} = \vec{f}(\overrightarrow{AB})$. On dit que $\vec{f}$ est l'*application linéaire associée* à f.

10.4 Proposition. *Soient E et F des espaces affines et $f : E \to F$ une application. Pour $A \in E$, notons φ_A l'application $\varphi_A : \vec{u} \mapsto \overrightarrow{f(A)f(A+\vec{u})}$. On équivalence entre :*

(i) *Il existe $A \in E$ tel que l'application φ_A soit linéaire ;*

(ii) *pour tout $A \in E$ l'application φ_A soit linéaire ;*

(iii) *l'application f est affine - et dans ce cas, on a $\varphi_A = \vec{f}$ pour tout $A \in E$.*

L'image (*resp.* l'image réciproque) d'un sous-espace affine de E (*resp.* de F) par une application affine $f : E \to F$ est un sous-espace affine de F (*resp.* de E).

Soit $f : E \to E$ une application affine. Les *points fixes* de f sont les $A \in E$ tels que $f(A) = A$. L'ensemble des points fixes de f est un sous-espace affine de E. S'il n'est pas vide, sa direction est $\ker(\vec{f} - \mathrm{id}_{\vec{E}})$.

Une composée d'applications affines est affine ; la réciproque d'une application affine bijective est affine.

10.5 Exemples. Soit E un espace affine. Soient F un sous-espace affine non vide de E et $\overrightarrow{G}$ un sous-espace supplémentaire de $\overrightarrow{F}$.

Projecteurs. Pour $M \in E$ il existe un unique point $P \in F$ tel que $\overrightarrow{PM} \in \overrightarrow{G}$. L'application $f : M \mapsto P$ ainsi construite est affine. On l'appelle le *projecteur* ou *projection* sur F parallèlement à $\overrightarrow{G}$. On a $f \circ f = f$. Inversement, toute application affine idempotente est de cette forme.

Symétries. Soit $M \in E$; notons P le projeté de M sur F parallèlement à $\overrightarrow{G}$. Posons $M' = P + \overrightarrow{MP}$. L'application $g : M \mapsto M'$ ainsi construite est affine. On l'appelle la *symétrie* par rapport à F parallèlement à $\overrightarrow{G}$. On a $g \circ g = \mathrm{id}_E$. Inversement, (si la caractéristique de K n'est pas 2) toute application affine involutive est de cette forme.
Question. Que se passe-t-il en caractéristique 2 ?

10.3 Barycentres

Soit E un espace affine. Un *point pondéré* est un couple $(A, \lambda) \in E \times K$.

10.6 Proposition. *Soient $(A_i, \lambda_i)_{i \in I}$ une famille finie de points pondérés.*

a) Si $\sum_{i \in I} \lambda_i = 0$, le vecteur $\sum_{i \in I} \lambda_i \overrightarrow{MA_i}$ ne dépend pas de M. On le note $\sum_{i \in I} \lambda_i A_i$

b) On suppose que $\sum_{i \in I} \lambda_i \neq 0$. Pour $G \in E$, les conditions suivantes sont équivalentes :

(i) *On a $\sum_{i \in I} \lambda_i \overrightarrow{GA_i} = \overrightarrow{0}$.*

(ii) *Il existe $M \in E$ tel que $\sum_{i \in I} \lambda_i \overrightarrow{MA_i} = \Big(\sum_{i \in I} \lambda_i\Big) \overrightarrow{MG}$.*

(iii) *Pour tout $M \in E$ on a $\sum_{i \in I} \lambda_i \overrightarrow{MA_i} = \Big(\sum_{i \in I} \lambda_i\Big) \overrightarrow{MG}$.*

Il existe un unique point G de E vérifiant ces conditions.

10.7 Définition. Le point G défini dans la proposition précédente s'appelle le *barycentre* des « points pondérés » $((A_1, \lambda_1), \ldots, (A_n, \lambda_n))$.

Lorsque $\lambda_1 = \ldots = \lambda_n$, on dit que M est l'*isobarycentre* de $(A_1, \ldots, A_n)$.

10.8 Propriétés des barycentres. Soient $((A_1, \lambda_1), \ldots, (A_n, \lambda_n))$ des points pondérés. On suppose que $\sum_{i=1}^{n} \lambda_i \neq 0$. Notons G le barycentre de $((A_1, \lambda_1), \ldots, (A_n, \lambda_n))$.

Homogénéité. Pour $\lambda \in K^*$, le barycentre de $((A_1, \lambda\lambda_1), \ldots, (A_n, \lambda\lambda_n))$ est encore G. On peut ainsi se ramener au cas où $\sum_{i=1}^{n} \lambda_i = 1$. Dans ce cas, le barycentre de $((A_1, \lambda_1), \ldots, (A_n, \lambda_n))$ se note $\sum_{i=1}^{n} \lambda_i A_i$.

Commutativité des barycentres. Pour toute permutation σ de $\{1, \ldots, n\}$, G est le barycentre de $((A_{\sigma(1)}, \lambda_{\sigma(1)}), \ldots, (A_{\sigma(n)}, \lambda_{\sigma(1)}))$.

Associativité des barycentres. Soit k un entier compris entre 1 et $n-1$. Posons $\mu = \sum_{i=k+1}^{n} \lambda_i$, et supposons que $\mu \neq 0$. Notons G_k le barycentre de $((A_{k+1}, \lambda_{k+1}), \ldots, (A_n, \lambda_n))$. Alors le barycentre de $((A_1, \lambda_1), \ldots, (A_k, \lambda_k), (G_k, \mu))$ est G.

10.9 Proposition. *a) Une partie d'un espace affine est un sous-espace affine si et seulement si elle est* stable par barycentres*.*

b) Une application entre espaces affines est affine si et seulement si elle respecte les barycentres*.*

Soit E un espace affine. Une partie F de E est dite *stable par barycentres* si pour tous $A_1, \ldots, A_n \in F$ et $\lambda_1, \ldots, \lambda_n \in K$ tels que $\sum_{i=1}^{n} \lambda_i \neq 0$, le barycentre de $((A_1, \lambda_1), \ldots, (A_n, \lambda_n))$ appartient à F.

Soient E, F des espaces affines et $f : E \to F$ une application. On dit que f *respecte les barycentres* si pour tous $A_1, \ldots, A_n \in E$ et $\lambda_1, \ldots, \lambda_n \in K$ tels que $\sum_{i=1}^{n} \lambda_i \neq 0$, l'image par f du barycentre de $((A_1, \lambda_1), \ldots, (A_n, \lambda_n))$ est le barycentre de $((f(A_1), \lambda_1), \ldots, (f(A_n), \lambda_n))$.

10.10 Corollaire. *Le sous-espace affine engendré par une partie d'un espace affine est l'ensemble des barycentres de points de cette partie.*

Tout ce qui concerne les barycentres devient « évident » si on suppose que E est un espace vectoriel, *i.e.* si on choisit une origine. Reste que ce choix n'est pas « canonique ». On peut par contre toujours considérer E comme sous-espace affine d'un espace vectoriel $\overrightarrow{H}$. Dans ce cas, G est le barycentre de $((A_1, \lambda_1), \ldots, (A_n, \lambda_n))$ si $\left(\sum_{i\in I} \lambda_i\right) G = \sum_{i\in I} \lambda_i A_i$ (dans $\overrightarrow{H}$) - et lorsque $\sum_{i\in I} \lambda_i = 1$, on a bien $\sum_{i\in I} \lambda_i A_i = G$.

10.11 Proposition. *Soit E un espace affine.*

a) Il existe un espace vectoriel $\overrightarrow{H}$, et une application affine injective $\varphi : E \to \overrightarrow{H}$ telle que $\overrightarrow{0} \notin \varphi(E)$. Quitte à remplacer $\overrightarrow{H}$ par le sous-espace engendré par $\varphi(E)$, on peut supposer que $\varphi(E)$ engendre $\overrightarrow{H}$, ce que l'on suppose dans la suite.

b) Il existe une unique forme linéaire $f : \overrightarrow{H} \to K$ telle que $\varphi(E) = \{x \in \overrightarrow{H};\ f(x) = 1\}$.
On a alors une identification canonique de $\overrightarrow{E}$ avec $\ker f$, qui envoie $\overrightarrow{AB}$ sur $\varphi(B) - \varphi(A)$.

c) Si on se donne $(\overrightarrow{H}_1, f_1, \varphi_1)$ et $(\overrightarrow{H}_2, f_2, \varphi_2)$ il existe un unique isomorphisme $u : \overrightarrow{H}_1 \to \overrightarrow{H}_2$ tel que $f_1 = f_2 \circ u$ et $\varphi_2 = u \circ \varphi_1$.

10.4 Repères

10.4.1 Repère cartésien

10.12 Définition. On appelle *repère cartésien* de E un $(n+1)$-uplet $(O, \vec{e}_1, \ldots, \vec{e}_n)$ où O est un point de E et $(\vec{e}_1, \ldots, \vec{e}_n)$ une base de $\overrightarrow{E}$.

Coordonnées cartésiennes. Lorsqu'on a fixé un repère cartésien $(O, \vec{e}_1, \ldots, \vec{e}_n)$ d'un espace affine E, on peut *repérer* un point M par ses *coordonnées cartésiennes*, *i.e.* les composantes du vecteur $\overrightarrow{OM}$ dans la base $(\vec{e}_1, \ldots, \vec{e}_n)$. En d'autres termes un repère cartésien nous donne un isomorphisme d'espaces affines de K^n sur E.

Changement de repère. Un changement de repère est donné par un changement d'origine (*i.e.* les coordonnées de la nouvelle origine dans l'ancien repère) et un changement de base (*i.e.* une matrice de passage).

Un repère cartésien sur E fixe, une base de l'espace vectoriel $\overrightarrow{E}$, donc, d'après 7.2.4 si $K = \mathbb{R}$:

- Une *orientation* :
- Une notion de *volume*, *i.e.* une mesure de Lebesgue.

Une transformation affine f de E

- préserve l'orientation si l'application linéaire tangente $\vec{f}$ a un déterminant positif, sinon elle la renverse ;
- multiplie les volumes par $|\det \vec{f}|$.

10.4.2 Repère affine

Equivalents affines de parties libres et génératrices : On a déjà vu l'équivalent des parties génératrices : une partie est (affinement) génératrice si le plus petit sous-espace affine qui la contient est E. Des points $(A_1, \ldots, A_n)$ sont dits *affinement indépendants* si pour tout $(\lambda_1, \ldots, \lambda_n) \in K^n$ les égalités $\sum \lambda_i = 0$ et $\sum \lambda_i A_i = \overrightarrow{0}$ impliquent $\lambda_1 = \ldots = \lambda_n = 0$.

10.13 Définition. Un *repère affine* ou *repère barycentrique* est une famille affinement indépendante et génératrice de points de E.

Soit $(A_0, A_1, \ldots, A_n)$ une famille de points de E. Cette famille est un repère affine si et seulement si la famille $(A_0, \overrightarrow{A_0A_1}, \ldots, \overrightarrow{A_0A_n})$ est un repère cartésien, *i.e.* si la famille $(\overrightarrow{A_0A_1}, \ldots, \overrightarrow{A_0A_n})$ est une base de $\overrightarrow{E}$.

Soit $(A_0, A_1, \ldots, A_n)$ un repère barycentrique de E. Pour tout point M de E il existe $(\lambda_0, \lambda_1, \ldots, \lambda_n) \in K^{n+1}$ tels que $\sum_{i=0}^{n} \lambda_i \neq 0$ et M soit le barycentre de $((A_0, \lambda_0), (A_1, \lambda_1), \ldots, (A_n, \lambda_n))$. Le $(n+1)$-uplet $(\lambda_0, \lambda_1, \ldots, \lambda_n)$ est unique à multiplication par un scalaire non nul près ; il s'appelle un système de *coordonnées barycentriques de M* dans le repère $(A_0, A_1, \ldots, A_n)$.

Réflexion. Qu'est-ce qu'un « repère » ? On veut que :

a) une application affine soit déterminée par ses valeurs dans le repère ;

b) « toutes les valeurs » soient possibles.

En particulier :

a) Une application affine qui fixe un repère est l'identité ;

b) étant donnés deux repères, il y a une (unique) application affine qui passe de l'un à l'autre.

10.5 Convexité

10.5.1 Généralités

Soit E un espace affine *réel* (ou complexe). Rappelons la Définition suivante.

10.14 Définition. Soit E un espace affine. Une partie C de E est dite *convexe* si pour tous $A_1, \ldots, A_n \in F$ et $t_1, \ldots, t_n \in \mathbb{R}_+$ non tous nuls, le barycentre de $((A_1, t_1), \ldots, (A_n, t_n))$ appartient à C.

10.15 Proposition. *Soit C une partie de E. La partie C est convexe dans E si et seulement si, pour tout $A, B \in C$ et tout $t \in [0,1]$, on a $(1-t)A + tB \in C$.*

Démonstration. Supposons que pour tout $A, B \in C$ et tout $t \in [0,1]$, on ait $(1-t)A + tB \in C$.
Nous devons démontrer que pour tout $n \in \mathbb{N}$, pour toute suite $A_1, \ldots, A_n$ d'éléments de C et toute suite $t_1, \ldots, t_n$ d'éléments de $\mathbb{R}_+$ tels que $\sum_{i=1}^{n} t_i = 1$, on a $\sum_{i=1}^{n} t_i A_i \in C$.
Démontrons cela par récurrence sur n. Cette propriété est vraie pour $n = 1$ (et $n = 2$). Si elle est vraie pour $n \geqslant 1$, soient $A_1, \ldots, A_n, A_{n+1}$ des points de C et $t_1, \ldots, t_n, t_{n+1} \in \mathbb{R}_+$ tels que $\sum_{i=1}^{n+1} t_i = 1$. Démontrons que $\sum_{i=1}^{n+1} t_i A_i \in C$; posons $t = \sum_{i=1}^{n} t_i = 1 - t_{n+1}$ et soient $s_1, \ldots, s_n \in \mathbb{R}_+$ tels que $\sum_{i=1}^{n} s_i = 1$ et $s_i t = t_i$. Par l'hypothèse de récurrence, on a $B = \sum_{i=1}^{n} s_i A_i \in C$; par le cas $n = 2$, on a aussi $\sum_{i=1}^{n+1} t_i A_i = tB + (1-t)A_{n+1} \in C$.
La réciproque est immédiate. □

Regroupons ci-dessous un certain nombre d'énoncés concernant la convexité.

Quelques propriétés de la convexité

a) Les parties convexes de l'espace $\mathbb{R}$ sont les intervalles.

Soit E un espace affine.

b) Tout sous-espace affine de E est convexe (dans E).

c) Une intersection de parties convexes de E est convexe. En particulier, si A est une partie quelconque de E, l'intersection de toutes les parties convexes de E contenant A est *le plus petit convexe de E contenant A* ; cette partie s'appelle l'*enveloppe convexe* de A. C'est aussi l'ensemble $\mathrm{conv}(A)$ des combinaisons convexes d'éléments de A, c'est-à-dire des $x \in E$ tels qu'il existe $p \in \mathbb{N}$ des points $x_0, \ldots, x_p \in A$ et $(t_0, \ldots, t_p) \in \mathbb{R}_+$ tels que $\sum_{i=1}^{p} t_i = 1$ et $x = \sum_{i=1}^{p} t_i x_i$. En effet cet ensemble est convexe (par associativité des barycentres) et tout ensemble convexe contenant A contiendra toutes les combinaisons convexes d'éléments de A.

d) Soit F un espace affine. L'image d'une partie convexe de E par une application affine $E \to F$ est convexe dans F.
De même, l'image réciproque d'une partie convexe de F par une application affine $E \to F$ est convexe dans E.

10.5.2 Théorème de Caratheodory

10.16 Théorème de Caratheodory. *Soient E un espace affine de dimension n et A une partie de E. Pour tout $x \in \mathrm{conv}(A)$, il existe $x_0, \dots, x_n \in A$ tels que x soit barycentre à coefficients positifs ou nuls de $x_0, \dots, x_n$.*

Démonstration. Soit $x \in \mathrm{conv}(A)$. Il suffit de démontrer qu'il existe $p \leqslant n$ tel que l'on puisse trouver $x_0, \dots, x_p \in A$ et $t_0, \dots, t_p \in \mathbb{R}_+$ de somme 1 tels que $x = \sum_{j=0}^{p} t_j x_j$ (quitte à poser $x_k = x_0$ et $t_k = 0$ pour $p < k \leqslant n$).

Notons $\mathcal{F}$ l'ensemble des parties finies F de A telles que x soit dans l'enveloppe convexe de F. Par Définition de $\mathrm{conv}(A)$, x est barycentre d'un nombre fini de points de A, autrement dit $\mathcal{F} \neq \emptyset$. Parmi les éléments $F \in \mathcal{F}$, soit J une partie possédant le plus petit nombre d'éléments. Démontrons que J a au plus $n+1$ éléments.

Soit $F = \{x_0, \dots, x_p\} \in \mathcal{F}$. Écrivons $x = \sum_{j=0}^{p} t_j x_j$ où $(t_j) \in \mathbb{R}_+^{p+1}$.

Supposons que les points $(x_0, \dots, x_p)$ sont affinement liés et démontrons que $F \neq J$.

Il existe $(\lambda_0, \dots, \lambda_p) \in \mathbb{R}^{p+1}$, non tous nuls tels que $\sum_{j=0}^{p} \lambda_j = 0$ et $\sum_{j=0}^{p} \lambda_j x_j = \vec{0}$.

Alors, pour tout $s \in \mathbb{R}$, on a $x = \sum_{j=0}^{p} (t_j + s\lambda_j) x_j$.

Pour $s \in \mathbb{R}$, posons $\varphi(s) = \inf\{t_j + s\lambda_j,\ 0 \leqslant j \leqslant p\}$. L'application φ est continue, positive en 0, et puisque les λ_j ne sont pas tous nuls et leur somme est nulle, il existe j tel que $\lambda_j < 0$, donc $\lim_{s \to +\infty} \varphi(s) = -\infty$. Par le théorème des valeurs intermédiaires, il existe $u \in \mathbb{R}$ tel que $\varphi(u) = 0$. Posons $\mu_j = t_j + u\lambda_j$. Les μ_j sont positifs ou nuls et il existe j tel que $\mu_j = 0$. Alors $x = \sum_{j=0}^{p} \mu_j x_j$, donc $F' = \{x_j \in F;\ \mu_j \neq 0\} \in \mathcal{F}$. En particulier, F n'a pas un nombre minimum d'éléments, donc $F \neq J$. On en déduit que J est affinement libre. En particulier $\mathrm{card} J \leqslant n+1$. □

10.17 Corollaire. *L'enveloppe convexe d'une partie compacte d'un espace affine de dimension finie est compacte.*

Démonstration. Soit A une partie compacta d'un espace affine E de dimension n.

Posons $\Sigma = \{(t_0, \dots, t_n) \in [0,1]^{n+1};\ \sum_{j=0}^{n} t_j = 1\}$. C'est un fermé de $[0,1]^{n+1}$, donc Σ est compact.

Posons $K = A^{n+1} \times \Sigma$; c'est un compact de (l'espace affine de dimension finie) $E^{n+1} \times \mathbb{R}^{n+1}$. Notons que nous pouvons choisir une origine dans E de sorte que E est un espace vectoriel de dimension finie. L'application $\varphi : A^{n+1} \times \Sigma \to E$ qui à $((x_0, \dots, x_n), (t_0, \dots, t_n))$ associe $\sum_{j=0}^{n} t_j x_j$ est continue. Son image est $\mathrm{conv}(A)$ d'après le théorème de Caratheodory. Elle est compacte. □

10.5.3 Fonctions convexes

10.18 Définition. Soient E un espace affine réel et C une partie convexe de E. Une application $f : C \to \mathbb{R}$ est dite *convexe* si son *surgraphe* $\{(x,u) \in E \times \mathbb{R}; f(x) \leqslant u\}$ est une partie convexe de $E \times \mathbb{R}$.

10.19 Proposition. *Soient E un espace affine réel, C une partie convexe de E et $f : C \to \mathbb{R}$ une application. Les conditions suivantes sont équivalentes :*

(i) *l'application f est convexe ;*

(ii) *pour tout $x, y \in C$ et tout $t \in [0,1]$, on a $f\big(tx + (1-t)y\big) \leqslant tf(x) + (1-t)f(y)$;*

(iii) *pour tout $n \in \mathbb{N}$, pour toute suite $x_1, \dots, x_n$ d'éléments de C et toute suite $t_1, \dots, t_n$ d'éléments de $\mathbb{R}_+$ tels que $\sum_{i=1}^{n} t_i = 1$, on a $f\Big(\sum_{i=1}^{n} t_i x_i\Big) \leqslant \sum_{i=1}^{n} t_i f(x_i)$.*

Démonstration. Le cas $n = 2$ dans (iii) est (ii) ; donc (iii)⇒(ii).

Notons S_f le surgraphe de f.

Démontrons que (ii)⇒(i). Soient (x, u) et (y, v) des éléments de S_f et $t \in [0,1]$; si (ii) est satisfaite, on a $f\big(tx + (1-t)y\big) \leqslant tf(x) + (1-t)f(y) \leqslant tu + (1-t)v$ puisque (x, u) et (y, v) sont dans S_f ; cela montre que $t(x, u) + (1-t)(y, v) \in S_f$; donc S_f est convexe d'après la prop. 10.15.

Enfin, soient $n \in \mathbb{N}$, $x_1, \dots, x_n$ une suite d'éléments de C et $t_1, \dots, t_n$ une suite d'éléments de $\mathbb{R}_+$ tels que $\sum_{i=1}^{n} t_i = 1$. Pour tout $i \in \{1, \dots, n\}$, on a $\big(x_i, f(x_i)\big) \in S_f$; si S_f est convexe, on a $\sum_{i=1}^{n} t_i\big(x_i, f(x_i)\big) \in S_f$, d'où l'on déduit l'assertion (i)⇒(iii). □

Soit E un espace affine réel.

- Une application affine $f : E \to \mathbb{R}$ est convexe (on a alors égalité dans la condition (ii) de la prop. 10.19).
- La somme de deux fonctions convexes est convexe.

10.6 Espaces affines euclidiens

10.20 Définition. Soit E un espace affine dont la direction est un espace vectoriel réel $\overrightarrow{E}$. Lorsque $\overrightarrow{E}$ est un espace vectoriel euclidien, on dit que E est un *espace affine euclidien.*

Soit E un espace affine euclidien de direction $\overrightarrow{E}$.

- La distance entre deux points A et B de E est $AB = \|\overrightarrow{AB}\|$.
- On dit qu'un repère cartésien $(O, \vec{e}_1, \dots, \vec{e}_n)$ est orthonormé si la base $(\vec{e}_1, \dots, \vec{e}_n)$ de $\overrightarrow{E}$ est orthonormée.

Projection orthogonale. Soit F un sous espace affine (non vide) de E. Puisque $\overrightarrow{E} = \overrightarrow{F} \oplus \overrightarrow{F}^{\perp}$, on peut définir la projection sur F parallèlement à $\overrightarrow{F}^{\perp}$ (*cf.* exemple 10.5) : on l'appelle *projection orthogonale* sur F. Pour $A \in E$, le projeté orthogonal P de A sur F est l'unique point de F minimisant la distance : pour $M \in F$, puisque $\overrightarrow{PM} \in \overrightarrow{F}$ et $\overrightarrow{AP} \in \overrightarrow{F}^{\perp}$, on a $AM^2 = AP^2 + PM^2 \geqslant AP^2$.

Symétrie orthogonale. On définit de même la *symétrie orthogonale* par rapport à un sous espace à F : c'est la symétrie par rapport à F parallèlement à $\overrightarrow{F}^{\perp}$. La symétrie orthogonale s_F par rapport à F est une isométrie : elle est bijective et pour $A, B \in E$, on a $s(A)s(B) = AB$.

Réflexion. Une *réflexion* est une symétrie orthogonale par rapport à un hyperplan.

Hyperplan médiateur. Soient A, B deux points distincts de E. L'ensemble $\{M \in E;\ AM = BM\}$ est un hyperplan affine : l'hyperplan médiateur de A et B : il passe par le milieu du segment $[A, B]$ et sa direction est $\overrightarrow{AB}^{\perp}$. La réflexion par rapport à cet hyperplan est l'unique réflexion s échangeant A et B.

Rappelons qu'une isométrie de E est une bijection de E sur E qui préserve les distances $f(A)f(B) = AB$ pour tout $A, B \in E$.

10.21 Théorème. *Toute isométrie d'un espace affine euclidien de dimension n est composée d'au plus $n+1$ réflexions. En particulier, elle est affine.*

Démonstration. Soit $f : E \to E$ une application qui préserve les distances. Soit $(A_1, A_2, \ldots, A_n, A_{n+1})$ un repère affine de E. Nous allons démontrer :

a) pour tout $k \in \{0, \ldots, n+1\}$, il existe une isométrie affine σ_k, produit d'au plus k réflexions tels que $\sigma_k \circ f(A_j) = A_j$ pour $1 \leqslant j < k$ (par convention l'identité est un produit de 0 réflexions) ;

b) $\sigma_{n+1} \circ f = \mathrm{id}_E$.

Il en résultera que $f = \sigma_{n+1}^{-1}$ d'où le théorème.

a) On pose $\sigma_0 = \mathrm{id}_E$. Soit $k \in \{1, \ldots, n+1\}$ et supposons σ_{k-1} construit ; si $\sigma_{k-1} \circ f(A_k) = A_k$, on pose $\sigma_k = \sigma_{k-1}$. Sinon, notons τ_k la réflexion par rapport à l'hyperplan médiateur de A_k et $B_k = \sigma_{k-1} \circ f(A_k)$ et posons $\sigma_k = \tau_k \circ \sigma_{k-1}$. On a évidemment $\sigma_k \circ f(A_k) = A_k$. Pour $j < k$, comme $\sigma_{k-1} \circ f$ est une isométrie, la distance A_j et A_k est égale à la distance entre leurs images A_j et B_k : en d'autres termes, A_j est dans l'hyperplan médiateur de A_k et B_k et est donc fixe par τ_k ; il en résulte que $\sigma_k \circ f(A_j) = A_j$.

b) Soit $M \in E$ et posons $N = \sigma_{n+1} \circ f(M)$. Comme tous les A_j sont fixes par l'isométrie $\sigma_{n+1} \circ f$, on a $MA_j = NA_j$. L'ensemble des points P tels que $MP = NP$ contient un repère, donc n'est pas contenu dans un hyperplan. Cela impose que $M = N$. □

Remarquons que, dans cette preuve, nous n'avons pas eu besoin de supposer que f est surjective : il suffit de supposer qu'elle préserve les distances.

10.22 Décomposition canonique. Soient E un espace vectoriel euclidien et $f : E \to E$ une isométrie. Il existe une unique décomposition $f = T \circ g = g \circ T$ où T est une translation et g est une isométrie possédant des points fixes. Cette décomposition de f s'appelle sa *décomposition canonique.*

Pour voir cela, remarquons que si $T_{\vec{v}}$ et la translation de vecteur $\vec{v} \in \overrightarrow{E}$ et g est une application affine, alors $g \circ T_{\vec{v}} = T_{\vec{g}(\vec{v})} \circ g$; donc $T_{\vec{v}}$ et g commutent si et seulement si $\vec{g}(\vec{v}) = \vec{v}$. Remarquons aussi que si $f = T_{\vec{v}} \circ g$ alors $\vec{f} = \vec{g}$. Soit alors $A \in E$ et posons $\vec{w} = \overrightarrow{Af(A)}$.

Si g est telle que $f = T_{\vec{v}} \circ g$ et B est un point fixe de g, on a $g(A) = g(B) + \vec{f}(\overrightarrow{BA})$, donc $f(A) = B + \vec{f}(\overrightarrow{BA}) + \vec{v}$ et enfin $\vec{w} = \overrightarrow{AB} + \vec{f}(\overrightarrow{BA}) + \vec{v} = (\mathrm{id}_{\vec{E}} - \vec{f})(\overrightarrow{AB}) + \vec{v}$. Or $\ker(\mathrm{id}_{\vec{E}} - \vec{f}) = \mathrm{im}(\mathrm{id}_{\vec{E}} - \vec{f})^{\perp}$, donc le seul $\vec{v}$ possible est le projeté orthogonal de $\vec{w}$ sur $\ker(\mathrm{id}_{\vec{E}} - \vec{f})$, d'où l'unicité.
Inversement, il existe un unique $\vec{u} \in \mathrm{im}(\mathrm{id}_{\vec{E}} - \vec{f})$ et $\vec{v} \in \ker(\mathrm{id}_{\vec{E}} - \vec{f})$ tels que $\vec{w} = \vec{u} + \vec{v}$. Il existe alors B tel que $\vec{u} = (\mathrm{id}_{\vec{E}} - \vec{f})(\overrightarrow{AB})$ et l'on vérifie que B est bien point fixe de g.

Classifications des isométries du plan euclidien

$\dim(\ker(\vec{f} - \mathrm{id}))$	avec points fixes	sans points fixes
2	id_E	translation
1	symétrie par rapport à une droite	symétrie glissée
0	rotation	*impossible*

Classifications des isométries de l'espace euclidien (dimension 3)

$\dim(\ker(\vec{f}-\mathrm{id}))$	avec points fixes	sans points fixes
3	id_E	translation
2	symétrie par rapport à un plan (réflexion)	symétrie glissée
1	rotation	vissage
0	antirotation	*impossible*

10.23 Exemple : le groupe du cube. Notons G le groupe des isométries directes du cube, *i.e.* les éléments de $SO(3)$ qui laissent invariant le cube $[-1,1]^3 \subset \mathbb{R}^3$. Le groupe G opère sur les 8 sommets les 12 arêtes et les 6 faces du cube. Quiconque a déjà vu un dé, sait que l'action sur les faces est transitive (toutes les faces du dé peuvent apparaître lorsqu'on lance le dé !). Le stabilisateur d'une face consiste en le groupe des quatre rotations d'axe perpendiculaire à cette face et d'angle $k\pi/2$. On en déduit que G a 24 éléments.

Le groupe du cube opère aussi sur les 4 grandes diagonales (qui passent par deux sommets opposés). On en déduit un homomorphisme de groupes f de G dans $\mathfrak{S}_4$. Nous allons démontrer que f est bijectif. Démontrons donc que f est injectif et surjectif - en sachant que, puisque G et $\mathfrak{S}_4$ ont même nombre d'éléments (24), il suffit de démontrer l'une de ces deux propriétés.

Injectivité. Soit $g \in G$. Si $g \in \ker f$, alors g fixe les 4 grandes diagonales ; en d'autres termes, leurs vecteurs directeurs $e_1 \pm e_2 \pm e_3$ sont propres pour g de valeur propre ± 1 - où (e_1, e_2, e_3) désigne la base canonique de $\mathbb{R}^3$. Les espaces propres étant en somme directe, on ne peut avoir deux espaces propres de dimension 2 ; trois quelconques de ces quatre vecteurs forment une base : on en déduit que $g = \pm \mathrm{id}_{\mathbb{R}^3}$; comme $-\mathrm{id}_{\mathbb{R}^3}$ n'est pas directe, il vient $g = \mathrm{id}_{\mathbb{R}^3}$.

Surjectivité. Notons g le demi tour d'axe $e_1 + e_2$. Les vecteurs $e_1 - e_2 \pm e_3$ sont orthogonaux à $e_1 + e_2$ donc sont des vecteurs propres de g pour la valeur propre -1 et g échange les deux vecteurs $e_1 + e_2 \pm e_3$. En d'autres termes, l'image de g est une transposition. Les 6 transpositions de $\mathfrak{S}_4$ sont ainsi obtenues comme images des demi-tours d'axes $e_i \pm e_j$ avec $1 \leqslant i < j \leqslant 3$, *i.e.* les stabilisateurs des arêtes. Comme ces transpositions engendrent $\mathfrak{S}_4$, on en déduit que f est surjective.

On trouvera en exercice (exerc. 10.11) quelques autres éléments sur le groupe du cube.

10.7 Exercices

10.1 Exercice. Soient E un espace affine de dimension finie et f une application affine de E dans E. Démontrer que si 1 n'est pas valeur propre de $\vec{f}$, alors f admet un unique point fixe.

10.2 Exercice. Soit E un espace affine de dimension $\geqslant 2$. Démontrer que toute bijection de E dans lui-même qui envoie toute droite sur une droite qui lui est parallèle est une homothétie-translation.

10.3 Exercice. Soient F, G des sous-espaces affines de E, A un point de F et B un point de G. Démontrer que l'on a $F \cap G \neq \emptyset$ si et seulement si $\overrightarrow{AB} \in \overrightarrow{F} + \overrightarrow{G}$.

10.4 Exercice. Dans un espace affine de dimension supérieure ou égale à 3, on considère un ensemble $\mathcal{D}$ de droites (affines). On suppose que deux droites de $\mathcal{D}$ ont un point commun. Démontrer que, soit toutes ces droites ont un point commun, soit elles sont coplanaires.

10.5 Exercice. Soit E un espace affine euclidien. Soient $A_1, \ldots, A_n$ des points de E et $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$. Pour $M \in E$ on pose $\varphi(M) = \sum_{i=1}^{n} \lambda_i \|A_i M\|^2$. Quels sont les extrema locaux de φ ?

10.6 Exercice. Soit C une partie convexe de E et $f : C \to \mathbb{R}$ une application convexe. Soit $a \in \mathbb{R}$. Démontrer que les ensembles $\{M \in E;\ f(M) < a\}$ et $\{M \in E;\ f(M) \leqslant a\}$ sont des parties convexes de E.

10.7 Exercice. *Coordonnées barycentriques des points remarquables du triangle.* Soit E un plan affine euclidien. Le but de cet exercice est de retrouver les coordonnées barycentriques des points remarquables d'un triangle non aplati ABC dans le repère affine (A, B, C). On note $\hat{A}, \hat{B}, \hat{C}$ les (mesures dans $]0, \pi[$ des) angles $\widehat{CAB}, \widehat{ABC}, \widehat{BCA}$ (on peut soit prendre des angles non orientés, ou mieux, on choisit l'orientation du plan de telle sorte que $\hat{A}, \hat{B}, \hat{C} \in]0, \pi[$).

1. Quelles sont les coordonnées barycentriques du centre de gravité G ?
2. a) Comment choisir un repère orthonormé de E pour lequel les affixes de A, B, C aient même module ?

   On fixe un tel repère.

   b) Démontrer qu'il existe $z \in \mathbb{C}^*$ et $s, t \in \mathbb{R}$ tels que s, t et $s+t$ ne sont pas multiples entiers de 2π tels que les affixes de A, B, C soient $z_A = z,\ z_B = ze^{is}$ et $z_C = ze^{-it}$. Exprimer s et t en fonction de $\hat{B}$ et $\hat{C}$.

   c) En remarquant que $(\sin t)e^{is} + (\sin s)e^{-it} = \sin(s+t)$, trouver des coordonnées barycentriques du centre du cercle circonscrit du triangle ABC.
3. a) Notons A' la projection orthogonale de A sur (BC). Calculer BA' et CA' en fonction de AA', $\hat{B}$ et $\hat{C}$. En déduire les coordonnées barycentriques de A' dans le repère affine (B, C) (discuter suivant le cas où les angles $\hat{B}$ et $\hat{C}$ sont aigus ou non).

   b) Démontrer qu'un système de coordonnées barycentriques de l'orthocentre H dans le repère (A, B, C) est $(\tan\hat{A}, \tan\hat{B}, \tan\hat{C})$.
4. On note I le centre du cercle inscrit à ABC.

   a) Démontrer que les vecteurs $\overrightarrow{AI}$ et $\dfrac{\overrightarrow{AB}}{AB} + \dfrac{\overrightarrow{AC}}{AC}$ sont colinéaires.

   b) Démontrer que (BC, AC, AB) est un système de coordonnées barycentriques de I dans le repère (A, B, C).

   c) Démontrer que $(\sin\hat{A}, \sin\hat{B}, \sin\hat{C})$ est un système de coordonnées barycentriques de I dans le repère (A, B, C).

10.8 Exercice. Soient E un espace vectoriel euclidien et $f : E \to E$ une isométrie.

1. On suppose que l'ensemble $F = \{x \in E;\ f(x) = x\}$ des point fixes de f n'est pas vide. Démontrer que f est produit de $\dim E - \dim F$ réflexions et qu'on ne peut pas faire mieux.
2. On suppose que f n'a pas de points fixes et on pose $\overrightarrow{F} = \{\vec{v} \in \overrightarrow{E};\ \vec{f}(\vec{v}) = \vec{v}\}$ des point fixes de l'application linéaire tangente à f. Démontrer que f est produit de $\dim \overrightarrow{E} - \dim \overrightarrow{F} + 2$ réflexions et qu'on ne peut pas faire mieux.

10.9 Exercice. Soient E un espace affine et $f : E \to E$ une application affine. Soit $A \in E$. Démontrer qu'il existe $\vec{v}$ tel que l'on ait $T_{\vec{v}} \circ f = f \circ T_{\vec{v}}$ et $f \circ T_{\vec{v}}$ possède des points fixes si et seulement si $\overrightarrow{Af(A)} \in \ker(\mathrm{id}_{\vec{E}} - \vec{f}) + \mathrm{im}(\mathrm{id}_{\vec{E}} - \vec{f})$ et qu'on a unicité de ce vecteur si et seulement si $\ker(\mathrm{id}_{\vec{E}} - \vec{f}) \cap \mathrm{im}(\mathrm{id}_{\vec{E}} - \vec{f}) = \{0\}$.

10.10 Exercice. Soient E un espace vectoriel euclidien et $f : E \to E$ une isométrie. Quels sont les points M de E qui minimisent la distance $Mf(M)$?

10.11 Exercice. Groupe de cube.

1. Décrire les 24 isométries directes du cube et leur action sur les diagonales, *i.e.* leur image dans le groupe $\mathfrak{S}_4$.
2. Construire un homomorphisme du groupe du cube à valeurs dans $\mathfrak{S}_3$. Quel est son noyau ?

Index