Algèbre et géométrie pour l'Agrégation de Mathématiques
Deuxième édition
Chapitres modifiés

Jean-Étienne ROMBALDI

12 septembre 2021

Table des matières

Chapitre 1

Représentations d'un groupe fini

[1]Pour ce chapitre, G est un groupe multiplicatif d'élément neutre noté e, $\mathbb{K}$ est un corps commutatif de caractéristique nulle et E un $\mathbb{K}$-espace vectoriel. Nous préciserons quand G est fini, $\mathbb{K} = \mathbb{C}$ ou E est de dimension finie.

1.1 Définitions et exemples

Définition 1.1. *Une représentation linéaire de G dans E est un morphisme de groupes $\rho : G \to GL(E)$. On dit aussi que E est un G-module.*

On note (ρ, E) une telle représentation linéaire du groupe G et on dira simplement représentation.

Si ρ est constante égale à Id_E (*i. e.* $\rho(g) = Id_E$ pour tout $g \in G$), on dit alors que la représentation est triviale.

Pour E est de dimension $n \in \mathbb{N}^*$, on dit que la représentation est de dimension finie et la dimension de E est le degré de cette représentation. En se fixant une base de E, une représentation de G dans E revient à se donner un morphisme de groupes de G dans $GL_n(\mathbb{K})$. Désignant, pour tout $g \in G$, par $R(g) = ((\rho_{ij}(g)))_{1 \leq i,j \leq n}$ la matrice de $\rho(g)$ dans une base $\mathcal{B}$ de E, on a pour g, h dans G, $R(g) \in GL_n(\mathbb{K})$ et $R(gh) = R(g) R(h)$, $R(g)^{-1} = R(g^{-1})$.

Se donner une représentation linéaire de G dans E revient à se donner une action à gauche de G sur E, $(g, x) \in G \times E \mapsto g * x \in E$, qui est $\mathbb{K}$-linéaire, c'est-à-dire telle que $g * (\lambda x + \mu y) = \lambda g * x + \mu g * y$ pour tous λ, μ dans $\mathbb{K}$ et x, y dans E (pour tout $g \in G$, l'application $x \mapsto g * x$ est linéaire). En effet, si $\rho : G \to GL(E)$ est une une telle représentation, on définit alors une action sur le groupe G par :

$$\forall (g, x) \in G \times E, \; g * x = \rho(g)(x)$$

puisque, pour tous g, g' dans G et tout x dans E, on a $e * x = \rho(e)(x) = Id_E(x) = x$ et $g * (g' * x) = \rho(g)(\rho(g')(x)) = \rho(g) \circ \rho(g')(x) = \rho(gg')(x) = (gg') * x$. Cette

1. Ce chapitre était présent dans la première édition et a été supprimé dans la deuxième.

action est bien $\mathbb{K}$-linéaire puisque :

$$g * (\lambda x + \mu y) = \rho(g)(\lambda x + \mu y) = \lambda \rho(g)(x) + \mu \rho(g)(y) = \lambda g * x + \mu g * y$$

Réciproquement, soit $(g, x) \mapsto g * x$ une action à gauche $\mathbb{K}$-linéaire de G sur E. Pour tout $g \in E$, l'application $\rho(g) : x \in E \mapsto g * x \in E$ est $\mathbb{K}$-linéaire et bijective. En effet, de la linéarité de l'action, on déduit que $\rho(g) \in \mathcal{L}(E)$, de $e * x = x$ pour tout $x \in E$, on déduit que $\rho(e) = Id_E$ et avec $g * \left(g^{-1} * x\right) = \left(gg^{-1}\right) * x = e * x = x$, $g^{-1} * (g * x) = x$, on déduit que $\rho(g) \circ \rho\left(g^{-1}\right) = \rho\left(g^{-1}\right) \circ \rho(g) = Id_E$, ce qui signifie que $\rho(g)$ est bijective d'inverse $\rho\left(g^{-1}\right)$. Enfin, avec $g * (g' * x) = (gg') * x$, pour tous g, g', x, on déduit que $\rho(gg') = \rho(g) \circ \rho(g')$, c'est-à-dire que l'application ρ est un morphisme de groupes de G dans $GL(E)$.

Exemples 1.1

1. *Un groupe diédral de type $\mathcal{D}_{2n}$ est isomorphe au sous-groupe de $GL_2(\mathbb{R})$ engendré par $R = \begin{pmatrix} \cos\left(\frac{2\pi}{n}\right) & -\sin\left(\frac{2\pi}{n}\right) \\ \sin\left(\frac{2\pi}{n}\right) & \cos\left(\frac{2\pi}{n}\right) \end{pmatrix}$ et $S = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$, ce qui nous donne une représentation de degré 2 de $\mathcal{D}_{2n}$ (voir le paragraphe* **??***).*
2. *Pour G fini et E de dimension* $\mathrm{card}(G)$, *en notant $\mathcal{B} = (e_k)_{k \in G}$ une base de E, on définit une représentation de G dans E en posant $\rho(g)(e_k) = e_{gk}$ pour tout $(g, k) \in G^2$. C'est la représentation régulière de G sur E.*
3. *Si $G = \mathfrak{S}_n$ est le groupe symétrique d'ordre n et E est de dimension n, en désignant par $\mathcal{B} = (e_k)_{1 \leq k \leq n}$ une base de E, on définit une représentation de $\mathfrak{S}_n$ dans E en posant $\rho(\sigma)(e_k) = e_{\sigma(k)}$ pour tous $\sigma \in \mathfrak{S}_n$ et $k \in \{1, \cdots, n\}$. C'est la représentation de G sur E par permutation. La trace de $\rho(\sigma)$ est le nombre de points fixes de la permutation σ.*
4. *Si $G = \langle a \rangle = \left\{e, a, \cdots, a^{m-1}\right\}$ est un groupe cyclique d'ordre m, on se donne un isomorphisme u_0 de E et on définit l'application $\rho : G \to GL(E)$ par :*

$$\forall k \in \{0, \cdots, m-1\},\ \rho\left(a^k\right) = u_0^k$$

C'est une représentation de G sur E si, et seulement si, on a $u_0^m = Id_E$. En effet si ρ est une représentation, on a alors $u_0^m = \rho(a^m) = \rho(e) = Id_E$. Réciproquement, si $u_0^m = Id_E$, on a alors pour tous j, k compris entre 0 et $m - 1$:

$$\rho\left(a^j a^k\right) = \rho\left(a^{j+k}\right) = u_0^{j+k} = u_0^j u_0^k = \rho\left(a^j\right) \rho\left(a^k\right) \text{ si } j + k \leq m - 1$$

$$\begin{aligned} \rho\left(a^j a^k\right) &= \rho\left(a^{j+k-m} a^m\right) = \rho\left(a^{j+k-m}\right) \\ &= u_0^{j+k-m} = u_0^{j+k} = u_0^j u_0^k = \rho\left(a^j\right) \rho\left(a^k\right) \text{ si } m \leq j + k \leq 2m - 2 \end{aligned}$$

et ρ est bien une représentation.

5. *Soient $((\rho_i, E_i))_{1 \leq i \leq p}$ une famille de représentations du groupe G dans des $\mathbb{K}$-espaces vectoriels $E_1, \cdots, E_p$ avec $p \geq 2$. On définit une représentation de G dans la somme directe $E = \bigoplus_{i=1}^{p} E_i$ en posant :*

$$\forall g \in G,\ \forall (x_1, \cdots, x_p) \in \prod_{i=1}^{p} E_i,\ \rho(g)\left(\sum_{i=1}^{p} x_i\right) = \sum_{i=1}^{p} \rho_i(g)(x_i)$$

En effet, comme $\rho_i(g) \in GL(E_i)$ pour $1 \leq i \leq p$, on a $\rho(g) \in \mathcal{L}(E)$. L'égalité $\sum_{i=1}^{p} \rho_i(g)(x_i) = 0$ équivaut à $\rho_i(g)(x_i) = 0$ pour tout i compris entre 1 et p, soit à $x_i = 0$ puisque chaque $\rho_i(g)$ est un isomorphisme, donc $\rho(g) \in GL(E)$ et on a bien une représentation. On dit que (ρ, E) est la somme directe des représentations (ρ_i, E_i), ce que l'on note $(\rho, E) = \left(\bigoplus_{i=1}^{p} \rho_i, \bigoplus_{i=1}^{p} E_i \right)$. Dans le cas où tous les E_i sont de dimension finie, en désignant par $\mathcal{B}_i$ une base de E_i, pour tout $g \in G$, la matrice de $\rho(g)$ dans la base $\bigcup_{i=1}^{p} \mathcal{B}_i$ de E est la matrice diagonale par blocs $R(g) = \operatorname{diag}(R_1(g), \cdots, R_p(g))$, où $R_i(g)$ est la matrice de $\rho_i(g)$ dans $\mathcal{B}_i$.

Définition 1.2. *Si (ρ, E) est une représentation de G, on dit alors que $\ker(\rho) = \{g \in G \mid \rho(g) = Id_E\}$ est le noyau de la représentation. La représentation est dite fidèle si son noyau est réduit à l'élément neutre de G.*

Exemples 1.2

1. *Si G est d'ordre n et E de dimension n, une représentation régulière de G dans E (exemple 2) est fidèle. En effet si g est dans le noyau de ρ, on alors $e_{gk} = \rho(g)(e_k) = e_k$ pour tout $k \in G$, donc $gk = k$ pour tout $k \in G$ et $g = e$.*
2. *Si G est le groupe $\mathcal{S}_n$ des permutations de $\{1, \cdots, n\}$, l'application ρ qui associe à toute permutation $\sigma \in \mathcal{S}_n$ la matrice de passage $R(\sigma)$ de la base canonique $\mathcal{B} = (e_j)_{1 \leq j \leq n}$ de $\mathbb{K}^n$ à la base $\mathcal{B}_\sigma = \left(e_{\sigma(j)}\right)_{1 \leq j \leq n}$ est un morphisme de groupes injectif de $\mathcal{S}_n$ dans $GL_n(\mathbb{K})$, c'est donc une représentation fidèle de $\mathcal{S}_n$.*
3. *Soient $G = \langle a \rangle$ un groupe cyclique d'ordre m, $u_0 \in GL(E)$ tel que $u_0^m = Id_E$ et $\rho : G \to GL(E)$ la représentation définie par $\rho\left(a^k\right) = u_0^k$ pour tout k compris entre 0 et $m-1$ (exemple 4). Si u_0 est d'ordre m, cette représentation est alors fidèle. En effet, si $a^k \in \ker(\rho)$ avec k compris entre 0 et $m-1$, on a alors $u_0^k = Id_E$ et nécessairement $k = 0$, donc $\ker(\rho) = \{e\}$ et ρ est fidèle. Sinon, l'ordre r de u_0 est un diviseur strict de m, soit $r \in \{1, \cdots, m-1\}$ et $a^r \in \ker(\rho)$ avec $a^r \neq e$, donc ρ n'est pas fidèle.*

1.2 Représentations irréductibles

Définition 1.3. *Si (ρ, E) est une représentation de G, on dit alors qu'un sous-espace vectoriel F de E est G-invariant (ou stable par G) si F est stable par tous les $\rho(g)$, ce qui signifie que $\rho(g)(F) \subset F$ pour tout $g \in G$.*

Les sous-espaces $\{0\}$ et E sont toujours G-invariant.

Si F est un sous-espace vectoriel de E de dimension finie qui est G-invariant, on a alors $\rho(g)(F) = F$ puisque $\rho(g)$ est un isomorphisme.

Si F est un sous-espace vectoriel de E qui est G-invariant, la représentation $\rho : G \to GL(E)$ induit alors une représentation $\rho' : G \to GL(F)$. On dit alors que ρ' est une sous-représentation de ρ.

Définition 1.4. *Si (ρ, E) est une représentation de G, on dit alors que $E^G = \{x \in E \mid \forall g \in G,\ \rho(g)(x) = x\}$ est l'ensemble des points fixes de G dans E.*

Pour tous x, y dans E^G, $\lambda \in \mathbb{K}$ et $g \in G$, on a :

$$\rho(g)(\lambda x + y) = \lambda \rho(g)(x) + \rho(g)(y) = \lambda x + y$$

donc E^G est un sous-espace vectoriel de E qui est G-invariant.

Définition 1.5. *On dit qu'une représentation de G dans E est irréductible (ou simple) si E n'est pas réduit au vecteur nul et les seuls sous-espaces G-invariants sont $\{0\}$ et E.*

Exemples 1.3

1. *Si G est d'ordre n et E de dimension n, une représentation régulière de G dans E (exemple 2) n'est pas irréductible. En effet, pour $x = \sum_{k \in G} e_k$ et tout $g \in G$, on a $\rho(g)(x) = \sum_{k \in G} \rho(g)(e_k) = \sum_{k \in G} e_{gk} = \sum_{h \in G} e_h = x$ puisque l'application $k \mapsto gk$ est une permutation de G. Donc la droite vectorielle $D = \mathbb{K}x$ dirigée par x est G-invariante non triviale. La restriction de ρ à D est la représentation triviale puisque $\rho(g)(x) = x$ pour tout $g \in G$.*
2. *Reprenant l'exemple 4 avec $\mathbb{K} = \mathbb{C}$ et E de dimension finie, comme $u_0^m = Id_E$, l'endomorphisme u_0 est annulé par le polynôme $X^m - 1$ qui est scindé à racines simples dans $\mathbb{C}[X]$, il est donc diagonalisable. Désignant par $\mathcal{B} = (e_i)_{1 \leq i \leq n}$ une base de E formé de vecteurs propres de u_0, $\mathcal{B}$ est aussi une base de vecteurs propres de $\rho(a^k) = u_0^k$ pour tout k compris entre 0 et $m-1$, donc chaque droite vectorielle $E_i = \mathbb{C}e_i$, pour i compris entre 1 et n, est G-invariante. Donc cette représentation de G n'est pas irréductible.*

Définition 1.6. *Soient E, F deux $\mathbb{K}$-espaces vectoriels et (ρ, E), (σ, F) deux représentations de G. On dit que $u \in \mathcal{L}(E, F)$ est G-linéaire (ou que c'est un G-morphisme, ou que c'est un opérateur d'entrelacement) si on a $u \circ \rho(g) = \sigma(g) \circ u$ pour tout $g \in G$. Si de plus u est bijective, on dit alors que c'est un G-isomorphisme ou que les représentations sont équivalentes.*

Dire que $u \in \mathcal{L}(E, F)$ est G-linéaire se traduit aussi par $\sigma(g) \circ u \circ \rho(g^{-1}) = u$ pour tout $g \in G$.

Dire les représentations (ρ, E) et (σ, F) sont équivalentes revient à dire qu'il existe un isomorphisme u de E sur F tel que $\rho(g) = u^{-1} \circ \sigma(g) \circ u$ pour tout $g \in G$ et dans le cas de la dimension finie, cela signifie qu'il existe une matrice inversible P telle que $R(g) = P^{-1}S(g)P$ pour tout $g \in G$ (les matrices $R(g)$, de $\rho(g)$ dans une base de E, et $S(g)$, de $\sigma(g)$ dans une base de F, sont semblables).

On note $\mathcal{L}_G(E,F)$ le sous-espace vectoriel de $\mathcal{L}(E,F)$ constitué des applications G-linéaires.

Lemme 1.1 *Soient (ρ, E) et (σ, F) deux représentations du même groupe G dans des $\mathbb{K}$-espaces vectoriels E et F.*

1. *L'application τ qui associe à $g \in G$ l'isomorphisme $\tau(g)$ de $\mathcal{L}(E,F)$ défini par $\tau(g)(u) = \sigma(g) \circ u \circ \rho(g^{-1})$ pour tout $u \in \mathcal{L}(E,F)$ est une représentation de G dans $\mathcal{L}(E,F)$. Si E et F sont de dimension finie, cette représentation est de degré* $\deg(\rho, E)\deg(\sigma, F)$.
2. *L'espace des points fixes de G dans cette représentation $(\tau, \mathcal{L}(E,F))$ est l'espace des applications G-linéaires, soit $\mathcal{L}(E,F)^G = \mathcal{L}_G(E,F)$.*

Preuve.

1. Pour tout $g \in G$ et tout $u \in \mathcal{L}(E,F)$, on a $\rho(g^{-1}) \in GL(E)$ et $\sigma(g) \in GL(F)$, donc $\tau(g)(u) = \sigma(g) \circ u \circ \rho(g^{-1}) \in \mathcal{L}(E,F)$. Pour tous scalaires λ, μ et tous morphismes u, v dans $\mathcal{L}(E,F)$, on a :
$$\tau(g)(\lambda u + \mu v) = \sigma(g) \circ (\lambda u + \mu v) \circ \rho(g^{-1}) = \lambda\tau(g)(u) + \mu\tau(g)(v)$$
donc $\tau(g)$ est bien une application linéaire de $\mathcal{L}(E,F)$ dans $\mathcal{L}(E,F)$. Pour u dans $\mathcal{L}(E,F)$, on a $\tau(e)(u) = \sigma(e) \circ u \circ \rho(e^{-1}) = u$, donc $\tau(e) = Id_{\mathcal{L}(E,F)}$. Pour g, h dans G et u dans $\mathcal{L}(E,F)$, on a :
$$\begin{aligned}\tau(gh)(u) &= \sigma(gh) \circ u \circ \rho(h^{-1}g^{-1}) = \sigma(g) \circ (\sigma(h) \circ u \circ \rho(h^{-1})) \circ \rho(g^{-1}) \\ &= \sigma(g) \circ (\tau(h)(u)) \circ \rho(g^{-1}) = \tau(g)(\tau(h)(u))\end{aligned}$$
donc $\tau(gh) = \tau(g) \circ \tau(h)$ et $\tau(g)$ est inversible d'inverse $\tau(g^{-1})$. En définitive, τ est bien un morphisme de groupes de G dans $GL(\mathcal{L}(E,F))$ et $(\tau, \mathcal{L}(E,F))$ est bien une représentation de G. Si E et F sont de dimension finie, cette représentation est alors de degré égal à $\dim(\mathcal{L}(E,F)) = \deg(\rho, E)\deg(\sigma, F)$.
2. On a :
$$\begin{aligned}\mathcal{L}(E,F)^G &= \{u \in \mathcal{L}(E,F) \mid \forall g \in G,\ \tau(g)(u) = u\} \\ &= \{u \in \mathcal{L}(E,F) \mid \forall g \in G,\ \sigma(g) \circ u \circ \rho(g^{-1}) = u\} = \mathcal{L}_G(E,F)\end{aligned}$$

□

Lemme 1.2 *Soient E, F deux $\mathbb{K}$-espaces vectoriels, (ρ, E), (σ, F) deux représentations de G et $u \in \mathcal{L}_G(E,F)$. Les espaces vectoriels* $\ker(u)$ *et* $\operatorname{Im}(u)$ *sont G-invariants dans E et F respectivement.*

Preuve. Pour tout $g \in G$ et tout $x \in \ker(u)$, on a :
$$u(\rho(g)(x)) = u \circ \rho(g)(x) = \sigma(g) \circ u(x) = \sigma(g)(u(x)) = \sigma(g)(0) = 0$$
donc $\rho(g)(x) \in \ker(u)$. Le sous-espace $\ker(u)$ est donc g-invariant dans E. Pour tout $g \in G$ et tout $y = u(x) \in \operatorname{Im}(u)$, on a :
$$\sigma(g)(y) = \sigma(g) \circ u(x) = u \circ \rho(g)(x) = u(\rho(g)(x)) \in \operatorname{Im}(u)$$
Le sous-espace $\operatorname{Im}(u)$ est donc G-invariant dans F. □

Lemme 1.3 (Schur) *Soient E, F deux $\mathbb{K}$-espaces vectoriels et (ρ, E), (σ, F) deux représentations irréductibles de G.*

1. *Si E et F ne sont pas isomorphes, on a alors $\mathcal{L}_G(E, F) = \{0\}$.*
2. *Si $\mathbb{K}$ est algébriquement clos et E, F sont isomorphes et de dimension finie, l'espace $\mathcal{L}_G(E, F)$ est alors de dimension 1 sur $\mathbb{K}$ (c'est donc un corps).*

Preuve.

1. Supposons qu'il existe une application G-linéaire $u \in \mathcal{L}_G(E, F)$ non nulle. Comme (ρ, E) est irréductible et $\ker(u)$ est G-invariant dans E distinct de E, on a $\ker(u) = \{0\}$, donc u est injective. Comme (σ, F) est irréductible et $\mathrm{Im}(u)$ est G-invariant dans F non réduit à $\{0\}$, on a $\mathrm{Im}(u) = F$, donc u est surjective et c'est un isomorphisme. On en déduit que $\mathcal{L}_G(E, F) = \{0\}$ si E et F ne sont pas isomorphes.
2. Le corps $\mathbb{K}$ étant algébriquement clos et $E \backsimeq F$ de dimension finie, toute application G-linéaire $u \in \mathcal{L}_G(E, F)$ identifiée à un endomorphisme de E admet une valeur propre $\lambda \in \mathbb{K}$ et le sous-espace propre $E_\lambda = \ker(u - \lambda Id_E)$ est G-invariant dans E (car $u - \lambda Id_E \in \mathcal{L}_G(E, E)$), non réduit au vecteur nul, c'est donc E tout entier si (ρ, E) est irréductible, ce qui signifie que u est une homothétie. On a donc $\mathcal{L}_G(E, F) \backsimeq \mathbb{K}$. On peut aussi vérifier que $\mathcal{L}_G(E, F)$ est un anneau dont tous les éléments non nuls sont inversibles, c'est donc un corps.

$\square$

Théorème 1.1.

Soient G fini, E, F deux $\mathbb{K}$-espaces vectoriels, (ρ, E), (σ, F) deux représentations de G, $u \in \mathcal{L}(E, F)$ et $\widehat{u} = \dfrac{1}{\mathrm{Card}(G)} \sum_{g \in G} \sigma(g) \circ u \circ \rho(g^{-1})$.

1. *$\widehat{u}$ est un G-morphisme de E dans F.*
2. *Si les représentations (ρ, E) et (σ, F) sont irréductibles et non G-isomorphes, on a alors $\widehat{u} = 0$.*
3. *Si $\mathbb{K}$ est algébriquement clos et $(\rho, E) = (\sigma, F)$ est irréductible de degré fini, on a alors $\widehat{u} = \dfrac{\mathrm{Tr}(u)}{\dim(E)} Id_E$ (homothétie de rapport $\dfrac{\mathrm{Tr}(u)}{\dim(E)}$).*

Preuve.

1. Pour tout $h \in G$, on a :

$$\begin{aligned}\sigma(h) \circ \widehat{u} \circ \rho(h^{-1}) &= \frac{1}{\mathrm{Card}(G)} \sum_{g \in G} \sigma(hg) \circ u \circ \rho\left((hg)^{-1}\right) \\ &= \frac{1}{\mathrm{Card}(G)} \sum_{k \in G} \sigma(k) \circ u \circ \rho(k^{-1}) = \widehat{u}\end{aligned}$$

du fait que l'application $k \mapsto g = h^{-1}k$ réalise une bijection de G sur lui même. Donc $\widehat{u} \in \mathcal{L}_G(E, F)$.

2. C'est une conséquence immédiate du lemme de Schur qui nous dit que E et F sont G-isomorphes si $\widehat{u} \in \mathcal{L}_G(E,F)$ est non nulle.
3. Pour $\mathbb{K}$ algébriquement clos, $E = F$ de dimension finie et ρ irréductible, le lemme précédent nous dit que $\mathcal{L}_G(E,E)$ est de dimension 1, il existe donc $\lambda \in \mathbb{K}$ tel que $\widehat{u} = \lambda Id_E$ et $\mathrm{Tr}(\widehat{u}) = \lambda \dim(E)$, avec :
$$\mathrm{Tr}(\widehat{u}) = \frac{1}{\mathrm{Card}(G)} \sum_{g\in G} \mathrm{Tr}\left(\sigma(g) \circ u \circ \rho\left(g^{-1}\right)\right) = \frac{1}{\mathrm{Card}(G)} \sum_{g\in G} \mathrm{Tr}(u) = \mathrm{Tr}(u)$$
ce qui donne bien $\lambda = \dfrac{\mathrm{Tr}(u)}{\dim(E)}$.

□

Le lemme qui suit nous dit que, sur un corps algébriquement clos, tout sous-espace G-invariant d'une représentation de degré fini d'un groupe fini admet un supplémentaire G-invariant.

Lemme 1.4 *On suppose que le groupe G est fini. On se donne une représentation (ρ, E) de G dans E, F un sous-espace vectoriel G-invariant de E et $\pi \in \mathcal{L}(E)$ un projecteur d'image F. L'application $\widehat{\pi} = \dfrac{1}{\mathrm{Card}(G)} \displaystyle\sum_{g\in G} \rho(g) \circ \pi \circ \rho\left(g^{-1}\right) \in \mathcal{L}_G(E,E)$ est un projecteur d'image F et $E = F \oplus H$, où H est un sous-espace vectoriel G-invariant de E.*

Preuve. On a vu avec le théorème précédent que $\widehat{\pi} \in \mathcal{L}_G(E,E)$. Pour tout $x \in E$ et $g \in G$, on a $\pi\left(\rho\left(g^{-1}\right)(x)\right) \in F$ et $\rho(g)\left(\pi\left(\rho\left(g^{-1}\right)(x)\right)\right) \in F$ puisque F est G-invariant, on a donc $\widehat{\pi}(x) \in F$, c'est-à-dire que $\widehat{\pi}$ est à valeurs dans F. D'autre part, pour tout $x \in F$, on a $\rho\left(g^{-1}\right)(x) \in F$ (F est G-invariant) donc $\pi\left(\rho\left(g^{-1}\right)(x)\right) = \rho\left(g^{-1}\right)(x) = \rho(g)^{-1}(x)$ et $\rho(g)\left(\pi\left(\rho\left(g^{-1}\right)(x)\right)\right) = x$. On a donc $\widehat{\pi}(x) = x$ pour tout $x \in F$ et $\mathrm{Im}(\widehat{\pi}) \subset F$, ce qui entraîne que $\widehat{\pi}$ est un projecteur d'image F. On a alors $E = \mathrm{Im}(\widehat{\pi}) \oplus \ker(\widehat{\pi}) = F \oplus H$, avec $H = \ker(\widehat{\pi})$ qui est G-invariant. □

Théorème 1.2. Maschke

On suppose que G est fini, que $\mathbb{K}$ est algébriquement clos et que (ρ, E) est une représentation de G dans E de degré fini. Dans ce cas, (ρ, E) est somme directe de sous-représentations irréductibles, c'est-à-dire qu'il existe des sous-espaces G-invariants et irréductibles $E_1, \cdots, E_p$ de E tels que $E = \displaystyle\bigoplus_{i=1}^{p} E_i$.

Preuve. On procède par récurrence sur le degré de la représentation. En dimension 1 toute représentation de G est irréductible. En supposant le résultat acquis pour les représentations de G de dimension $n \geq 1$, on se donne une représentation $\rho : G \to GL(E)$ de degré $n+1$. Si ρ est irréductible, c'est terminé, sinon il existe des sous-espaces vectoriels de E qui sont G-invariant et distincts de $\{0\}$ et de E. En désignant par F un tel sous-espace de dimension minimale, ce sous-espace G-invariant est irréductible et le lemme précédent nous dit qu'il admet un supplémentaire H qui est G-invariant et distincts de $\{0\}$ et de E. Il suffit alors d'appliquer l'hypothèse de récurrence à $\rho : G \to GL(H)$ pour conclure. $\square$

1.3 Caractères des groupes finis

Pour ce paragraphe et les suivants, le groupe G est fini et les espaces vectoriels considérés sont de dimension finie.

Si $\mathbb{K} = \mathbb{C}$ et G est commutatif, une représentation irréductible de G dans E est nécessairement de degré 1 (exercice 1.3), donc $GL(E)$ est isomorphe au groupe multiplicatif $\mathbb{K}^*$ et il existe une application $\chi : G \to \mathbb{K}^*$ telle que $\rho(g) = \chi(g) Id_E$ pour tout $g \in G$ et $\chi(g)$ est la trace de $\rho(g)$.

Pour G fini, non nécessairement commutatif, on donne la définition suivante.

Définition 1.7. *Si (ρ, E) est une représentation de G, son caractère est l'application $\chi_\rho : G \to \mathbb{K}$ définie par $\chi_\rho(g) = \mathrm{Tr}(\rho(g))$ pour tout $g \in G$.*

Définition 1.8. *On dit qu'une fonction $\chi : G \to \mathbb{K}$ est un caractère [resp. un caractère irréductible] de G s'il existe une représentation [resp. une représentation irréductible] (ρ, E) de G telle que $\chi = \chi_\rho$.*

On notera simplement χ pour χ_ρ quand la représentation est fixée.

Exemple 1.1 *Le caractère de la représentation triviale $\rho : g \mapsto Id_E$ est la fonction constante $\chi : g \mapsto n$.*

Des propriétés de la trace, on déduit facilement les résultats suivant.

Théorème 1.3.

Soient (ρ, E) une représentation de G et $\chi_\rho : G \to \mathbb{K}$ son caractère. On a :

1. *$\chi_\rho(e) = \deg(\rho, E) = \dim(E)$ (toutes les représentations de G de même caractère χ, ont le même degré $\chi(e)$) ;*
2. *pour tous g, h dans G, $\chi_\rho(gh) = \chi_\rho(hg)$ et $\chi_\rho\left(ghg^{-1}\right) = \chi_\rho(h)$ (χ_ρ est constant sur chaque classe de conjugaison de G) ;*
3. *si (ρ, E) est somme directe de $p \geq 2$ représentations de G, son caractère est alors la somme des caractères correspondants.*

Preuve.

1. On a $\chi_\rho(e) = \mathrm{Tr}(\rho(e)) = \mathrm{Tr}(Id_E) = \dim(E)$.
2. Pour g, h dans G, on a :

$$\chi_\rho(gh) = \mathrm{Tr}(\rho(gh)) = \mathrm{Tr}(\rho(g) \circ \rho(h)) = \mathrm{Tr}(\rho(h) \circ \rho(g)) = \chi_\rho(hg)$$

et $\chi_\rho\left(ghg^{-1}\right) = \chi\left(hg^{-1}g\right) = \chi(h)$.

3. Si $(\rho, E) = \left(\bigoplus_{i=1}^{p} \rho_i, \bigoplus_{i=1}^{p} E_i\right)$, en désignant par $\mathcal{B}_i$ une base de E_i, pour tout $g \in G$, la matrice de $\rho(g)$ dans la base $\bigcup_{i=1}^{p} \mathcal{B}_i$ de E est la matrice diagonale par blocs $R(g) = \operatorname{diag}(R_1(g), \cdots, R_p(g))$, où $R_i(g)$ est la matrice de $\rho_i(g)$ dans $\mathcal{B}_i$ et on a $\chi_\rho(g) = \operatorname{Tr}(R(g)) = \sum_{i=1}^{p} \operatorname{Tr}(R_i(g)) = \sum_{i=1}^{p} \chi_{\rho_i}(g)$.

□

Théorème 1.4.

Deux représentations de G équivalentes, ont même caractère.

Preuve. Si les représentations (ρ, E) et (σ, F) de G sont équivalentes, il existe alors un isomorphisme u de E sur F tel que $\rho(g) = u^{-1} \circ \sigma(g) \circ u$ pour tout $g \in G$ et on a $\operatorname{Tr}(\rho(g)) = \operatorname{Tr}(\sigma(g))$ pour tout $g \in G$. Donc $\chi_\rho = \chi_\sigma$. □

Nous verrons plus loin que la réciproque du théorème précédent est vraie.

Théorème 1.5.

On suppose que $\mathbb{K} = \mathbb{C}$ et on se donne une représentation (ρ, E) de G de caractère $\chi : G \to \mathbb{C}$.

1. *Si $g \in G$ est d'ordre r, $\chi(g)$ est alors somme de n racines r-ièmes de l'unité.*
2. *Pour tout $g \in G$, on a $\chi\left(g^{-1}\right) = \overline{\chi(g)}$ et $|\chi(g)| \leq \chi(e) = \dim(E)$.*
3. *Le noyau de (ρ, E) est $\ker(\rho) = \{g \in G \mid \chi(g) = \chi(e)\}$.*

Preuve.

1. Si $g \in G$ est d'ordre $r \geq 1$, l'isomorphisme $\rho(g)$ est diagonalisable et ses valeurs propres $\lambda_1, \cdots, \lambda_n$ sont des racines r-ièmes de l'unité (exercice 1.2), donc $\chi(g) = \operatorname{Tr}(\rho(g)) = \sum_{i=1}^{n} \lambda_i$ est somme de n racines r-ièmes de l'unité.
2. On a $|\lambda_i| = 1$ pour tout i compris entre 1 et n et les $\overline{\lambda_i} = \lambda_i^{-1}$ sont les valeurs propres de $\rho\left(g^{-1}\right) = \rho(g)^{-1}$, donc $|\chi(g)| \leq \sum_{i=1}^{n} |\lambda_i| = n = \dim(E) = \chi(e)$ et $\chi\left(g^{-1}\right) = \operatorname{Tr}\left(\rho\left(g^{-1}\right)\right) = \sum_{i=1}^{n} \lambda_i^{-1} = \sum_{i=1}^{n} \overline{\lambda_i} = \overline{\chi(g)}$.
3. Pour tout $g \in \ker(\rho)$, on a $\rho(g) = Id_E = \rho(e)$ et $\chi(g) = \operatorname{Tr}(\rho(g)) = n = \chi(e)$. Réciproquement, soit $g \in G$ tel $\chi(g) = \chi(e)$. En désignant par $\lambda_1, \cdots, \lambda_n$ les valeurs propres de $\rho(g)$, on a $\sum_{k=1}^{n} \lambda_k = n$, les $\lambda_k = e^{i\theta_k}$ étant des racines de l'unité. On a alors $\sum_{k=1}^{n} \cos(\theta_k) = n$ et nécessairement $|\cos(\theta_k)| = 1$ pour tout k

(sinon $n \leq \sum_{k=1}^{n} |\cos(\theta_k)| < n$, ce qui est impossible), soit $\lambda_k = \pm 1$ pour tout k et l'égalité $\sum_{k=1}^{n} \lambda_k = n$ impose $\lambda_k = 1$ pour tout k compris entre 1 et n (sinon, en séparant les -1 des 1 dans la somme, on aboutit à $p = n + q$ avec $0 \leq p \leq n-1$ et $1 \leq q \leq n$, ce qui est impossible). L'endomorphisme diagonalisable $\rho(g)$ a donc toutes ses valeurs propres égales à 1, c'est donc Id_E et $g \in \ker(\rho)$.

□

Définition 1.9. *Si (ρ, E) est une représentation de G, sa moyenne est l'application linéaire définie par* $\mu = \dfrac{1}{\mathrm{Card}(G)} \sum_{g \in G} \rho(g)$.

Lemme 1.5 *Soient (ρ, E) une représentation de G et μ sa moyenne.*

1. μ *est un projecteur.*

2. $E^G = \mathrm{Im}(\mu) = \ker(\mu - Id_E)$.

3. $\mathrm{Tr}(\mu) = \dim\left(E^G\right) = \dfrac{1}{\mathrm{Card}(G)} \sum_{g \in G} \chi(g)$.

Preuve.

1. Pour tout $g \in G$, on a :

$$\rho(g) \circ \mu = \frac{1}{\mathrm{Card}(G)} \sum_{h \in G} \rho(g) \circ \rho(h) = \frac{1}{\mathrm{Card}(G)} \sum_{h \in G} \rho(gh)$$

et du fait que l'application $k \mapsto h = g^{-1}k$ est une permutation de G, on déduit que $\rho(g) \circ \mu = \dfrac{1}{\mathrm{Card}(G)} \sum_{k \in G} \rho(k) = \mu$. Il en résulte que $\mu \circ \mu = \mu$, c'est-à-dire que μ est un projecteur.

2. Avec $(\mu - Id_E) \circ \mu = 0$, on déduit que $\mathrm{Im}(\mu) \subset \ker(\mu - Id_E)$ et du fait que pour tout $x \in \ker(\mu - Id_E)$, on a $x = \mu(x) \in \mathrm{Im}(\mu)$, on déduit l'égalité $\mathrm{Im}(\mu) = \ker(\mu - Id_E)$. Pour $x \in E^G$, on a :

$$\mu(x) = \frac{1}{\mathrm{Card}(G)} \sum_{g \in G} \rho(g)(x) = \frac{1}{\mathrm{Card}(G)} \sum_{g \in G} x = x$$

c'est-à-dire que $E^G \subset \ker(\mu - Id_E)$. Réciproquement si $x = \mu(x)$, on a alors $\rho(g)(x) = (\rho(g) \circ \mu)(x) = \mu(x) = x$ pour tout $g \in G$ et $x \in E^G$. On a donc bien $E^G = \mathrm{Im}(\mu) = \ker(\mu - Id_E)$.

3. Comme μ est un projecteur, on a $\mathrm{Tr}(\mu) = \dim(\mathrm{Im}(\mu)) = \dim\left(E^G\right)$ avec :

$$\mathrm{Tr}(\mu) = \frac{1}{\mathrm{Card}(G)} \sum_{g \in G} \mathrm{Tr}(\rho(g)) = \frac{1}{\mathrm{Card}(G)} \sum_{g \in G} \chi(g)$$

□

Lemme 1.6 *Si E, F sont des espaces vectoriels (de dimension finie), $f \in \mathcal{L}(E)$, $g \in \mathcal{L}(F)$ et $\varphi : \mathcal{L}(E,F) \to \mathcal{L}(E,F)$ est l'application linéaire définie par :*

$$\forall u \in \mathcal{L}(E,F), \ \varphi(u) = g \circ u \circ f$$

on a alors $\operatorname{Tr}(\varphi) = \operatorname{Tr}(f)\operatorname{Tr}(g)$.

Preuve. On se donne une base $\mathcal{B} = (e_i)_{1 \le i \le n}$ de E, une base $\mathcal{B}' = (e'_i)_{1 \le i \le n}$ une base de F et on note $\mathcal{B}'' = (u_{ij})_{\substack{1 \le i \le n \\ 1 \le j \le m}}$ la famille d'applications linéaires de E dans F définie par :

$$\forall k \in \{1, \cdots, n\}, \ u_{ij}(e_k) = \begin{cases} 0 \text{ si } k \neq i \\ e'_j \text{ si } k = i \end{cases}$$

(soit en désignant par $(e_i^*)_{1 \le i \le n}$ la base duale de $\mathcal{B}$, $u_{ij}(x) = e_i^*(x)\, e'_j$). On vérifie facilement que cette famille est une base de $\mathcal{L}(E,F)$ (il y a $nm = \dim(\mathcal{L}(E,F))$ éléments et $\sum_{\substack{1 \le i \le n \\ 1 \le j \le m}} \alpha_{ij} u_{ij} = 0$ appliqué à e_k donne $\sum_{j=1}^{m} \alpha_{kj} e'_j = 0$ qui entraîne la nullité de tous les coefficients α_{kj}), toute application linéaire $u \in \mathcal{L}(E,F)$ s'écrivant $u = \sum_{\substack{1 \le i \le n \\ 1 \le j \le m}} \alpha_{ij} u_{ij}$, où les coefficients α_{ij} sont définis par $u(e_k) = \sum_{j=1}^{m} \alpha_{kj} e'_j$ pour tout $k \in \{1, \cdots, n\}$ (c'est-à-dire que $((\alpha_{ij})) = {}^t A$, où A est la matrice de u dans les bases $\mathcal{B}$ et $\mathcal{B}'$). Pour $1 \le r \le n$ et $1 \le s \le m$, on a $\varphi(u_{rs}) = \sum_{\substack{1 \le i \le n \\ 1 \le j \le m}} \alpha_{ij}^{(r,s)} u_{ij}$ et

$\operatorname{Tr}(\varphi) = \sum_{\substack{1 \le r \le n \\ 1 \le s \le m}} \alpha_{rs}^{(r,s)}$. Avec $\varphi(u_{rs})(e_k) = \sum_{j=1}^{m} \alpha_{kj}^{(r,s)} e'_j$ et :

$$\begin{aligned} \varphi(u_{rs})(e_k) &= (g \circ (u_{rs}) \circ f)(e_k) = (g \circ (u_{rs})) \left(\sum_{i=1}^{n} a_{ik} e_i \right) \\ &= g(a_{rk} e'_s) = a_{rk} \sum_{j=1}^{m} b_{js} e'_j \end{aligned}$$

on déduit que $\alpha_{kj}^{(r,s)} = a_{rk} b_{js}$, ce qui donne $\operatorname{Tr}(\varphi) = \sum_{\substack{1 \le r \le n \\ 1 \le s \le m}} a_{rr} b_{ss} = \operatorname{Tr}(f)\operatorname{Tr}(g)$.

□

On note $\mathbb{K}^G$ l'espace vectoriel des applications de G dans $\mathbb{K}$. Cet espace vectoriel est de dimension $\operatorname{card}(G)$.

Pour toute fonction $v \in \mathbb{K}^G$, on note v^* la fonction définie par $v^*(g) = v(g^{-1})$ pour tout $g \in G$ et on définit sur $\mathbb{K}^G$ la forme bilinéaire $\langle \cdot \mid \cdot \rangle$ par :

$$\forall (u,v) \in \mathbb{K}^G \times \mathbb{K}^G, \ \langle u \mid v \rangle = \frac{1}{\operatorname{Card}(G)} \sum_{g \in G} u(g)\, v^*(g)$$

Lemme 1.7 *La forme bilinéaire $\langle \cdot \mid \cdot \rangle$ est symétrique et non dégénérée sur $\mathbb{K}^G$.*

Preuve. L'application $g \mapsto g^{-1}$ étant une permutation de G et le corps $\mathbb{K}$ commutatif, on déduit que la forme bilinéaire $\langle \cdot \mid \cdot \rangle$ est symétrique. Si $u \in \mathbb{K}^G$ est telle que $\langle u \mid v \rangle = 0$ pour tout $v \in \mathbb{K}^G$, en prenant pour fonction v la fonction caractéristique du singleton $\left\{h^{-1}\right\}$, où h est quelconque dans G, on a, $0 = \langle u \mid v \rangle = \dfrac{1}{\text{Card}\,(G)} u\,(h)$ et $u\,(h) = 0$, donc u est l'application nulle. La forme bilinéaire $\langle \cdot \mid \cdot \rangle$ est donc non dégénérée. □

Si $\mathbb{K} = \mathbb{C}$ et v est un caractère, on a alors $v^*\,(g) = v\left(g^{-1}\right) = \overline{v\,(g)}$ et :

$$\langle u \mid v \rangle = \frac{1}{\text{Card}\,(G)} \sum_{g \in G} u\,(g)\,\overline{v\,(g)}$$

L'application $\langle \cdot \mid \cdot \rangle$ ainsi définie est un produit scalaire hermitien sur $\mathbb{K}^G$.

Théorème 1.6.

Soient $(\rho, E)\,,\,(\sigma, F)$ deux représentations de G et χ_ρ, χ_σ leurs caractères respectifs.

1. *Le caractère de la représentation associée $(\tau, \mathcal{L}\,(E, F))$ (voir le lemme 1.1) est $\chi_\tau = \chi_\rho^* \chi_\sigma$.*
2. *On a $\langle \chi_\rho \mid \chi_\sigma \rangle = \dim\left(\mathcal{L}_G\,(E, F)\right)$, où $\mathcal{L}_G\,(E, F)$ est l'espace des applications G-linéaires de E dans F ($\langle \chi_\rho \mid \chi_\sigma \rangle$ est donc un entier).*
3. *Si E n'est pas réduit au vecteur nul, on a alors $\langle \chi_\rho \mid \chi_\rho \rangle \in \mathbb{N}^*$.*
4. *Si les représentations (ρ, E) et (σ, F) sont irréductibles, on a alors :*

$$\left(\langle \chi_\rho \mid \chi_\sigma \rangle \geq 1\right) \Leftrightarrow \left((\rho, E) \text{ et } (\sigma, F) \text{ sont équivalentes}\right) \Leftrightarrow \left(\chi_\rho = \chi_\sigma\right)$$

Preuve.

1. Pour tous g dans G et u dans $\mathcal{L}\,(E, F)$, on a $\tau\,(g)\,(u) = \sigma\,(g) \circ u \circ \rho\left(g^{-1}\right)$ et, en utilisant le lemme précédent :

$$\chi_\tau\,(g) = \text{Tr}\,(\tau\,(g)) = \text{Tr}\left(\rho\left(g^{-1}\right)\right) \text{Tr}\,(\sigma\,(g)) = \chi_\rho\left(g^{-1}\right) \chi_\sigma\,(g)$$

   On a donc bien $\chi_\tau = \chi_\rho^* \chi_\sigma$.

2. On a :

$$\begin{aligned}\langle \chi_\rho \mid \chi_\sigma \rangle = \langle \chi_\sigma \mid \chi_\rho \rangle &= \frac{1}{\text{Card}\,(G)} \sum_{g \in G} \chi_\rho\left(g^{-1}\right) \chi_\sigma\,(g) \\ &= \frac{1}{\text{Card}\,(G)} \sum_{g \in G} \chi_\tau\,(g) = \dim\left(\mathcal{L}\,(E, F)^G\right)\end{aligned}$$

   et avec $\mathcal{L}\,(E, F)^G = \mathcal{L}_G\,(E, F)$, on déduit que $\langle \chi_\rho \mid \chi_\sigma \rangle = \dim\left(\mathcal{L}_G\,(E, F)\right)$.

3. En particulier, on a $\langle \chi_\rho \mid \chi_\rho \rangle = \dim\left(\mathcal{L}_G\,(E, E)\right) \in \mathbb{N}^*$ si E n'est pas réduit au vecteur nul puisque $\mathcal{L}_G\,(E, E)$ contient Id_E qui est non nulle.

4. Si $\langle \chi_\rho \mid \chi_\sigma \rangle \geq 1$, on a alors $\dim(\mathcal{L}_G(E,F)) \geq 1$ et le lemme de Schur, pour (ρ, E) et (σ, F) irréductibles, nous dit que les représentations sont équivalentes et en conséquence, elles ont même caractère (théorème 1.4). Si $\chi_\rho = \chi_\sigma$, on a alors $\langle \chi_\rho \mid \chi_\sigma \rangle = \langle \chi_\rho \mid \chi_\rho \rangle \geq 1$ puisque E n'est pas réduit au vecteur nul ((ρ, E) étant irréductible, on a $E \neq \{0\}$).

□

Avec les notations du théorème précédent, pour $\mathbb{K} = \mathbb{C}$, on a $\chi_\tau = \overline{\chi_\rho}\chi_\sigma$.

Pour $\mathbb{K}$ algébriquement clos et (ρ, E) irréductible on a vu que $\mathcal{L}_G(E,E)$ est de dimension 1 (lemme 1.3), donc $\langle \chi_\rho \mid \chi_\rho \rangle = 1$.

Le théorème précédent nous dit que si χ_ρ, χ_σ sont les caractères de deux représentations irréductibles, ils sont orthogonaux dans $\mathbb{K}^G$ pour la forme bilinéaire $\langle \cdot \mid \cdot \rangle$ si, et seulement si, $\chi_\rho \neq \chi_\sigma$.

Corollaire 1.1. *Si $\chi_1, \cdots, \chi_p$ sont des caractères irréductibles distincts de G, ils sont alors linéairement indépendants dans $\mathbb{K}^G$ et $p \leq \mathrm{Card}(G)$.*

Preuve. Si $\sum_{j=1}^{p} \lambda_j \chi_j = 0$, on a alors $0 = \sum_{j=1}^{p} \lambda_j \langle \chi_j \mid \chi_i \rangle = \lambda_i \langle \chi_i \mid \chi_i \rangle$ pour tout i compris entre 1 et p, compte tenu de l'orthogonalité des χ_j avec $\langle \chi_i \mid \chi_i \rangle > 0$, ce qui donne $\lambda_i = 0$. Les χ_i sont donc linéairement indépendants dans $\mathbb{K}^G$ qui est de dimension $\mathrm{Card}(G)$ et nécessairement $p \leq \mathrm{Card}(G)$. □

On a donc au plus $\mathrm{Card}(G)$ caractères irréductibles sur G.

1.4 Fonctions centrales

On suppose que le groupe G est fini et que $\mathbb{K} = \mathbb{C}$.

Définition 1.10. *On appelle fonction centrale de G dans $\mathbb{C}$ toute application $\varphi : G \to \mathbb{C}$ constante sur les classes de conjugaison de G, c'est-à-dire telle que $\varphi(ghg^{-1}) = \varphi(h)$ pour tous g, h dans G.*

Le point **2.** du théorème 1.3 nous dit qu'un caractère est une fonction centrale.

Une fonction $\varphi : G \to \mathbb{C}$ est centrale si, et seulement si, $\varphi(gk) = \varphi(kg)$ pour tous g, k dans G.

On désigne par $\mathcal{H}$ le sous-espace de $\mathbb{C}^G$ formé des fonctions centrales sur G et on définit sur $\mathcal{H}$ un produit hermitien en posant :

$$\forall (\varphi, \psi) \in \mathcal{H}^2,\ \langle \varphi, \psi \rangle = \frac{1}{|G|} \sum_{g \in G} \overline{\varphi(g)} \psi(g)$$

Lemme 1.8 *Si $\varphi \in \mathcal{H}$ et (ρ, E) est une représentation de G de caractère χ, l'application $\rho_\varphi \in \mathcal{L}(E)$ définie par $\rho_\varphi = \dfrac{1}{\mathrm{Card}(G)} \sum_{g \in G} \varphi(g^{-1}) \rho(g)$ est un G-morphisme de E et, si de plus la représentation (ρ, E) est irréductible, ρ_φ est alors une homothétie de rapport $\dfrac{\langle \varphi \mid \chi \rangle}{\dim(E)}$.*

Preuve. Pour tout $h \in G$, on a $\rho_\varphi \circ \rho(h) = \dfrac{1}{\mathrm{Card}(G)} \sum_{g \in G} \varphi(g^{-1}) \rho(gh)$ et comme l'application $k \mapsto g = hkh^{-1}$ est une permutation de G, cela s'écrit :

$$\begin{aligned}\rho_\varphi \circ \rho(h) &= \frac{1}{\mathrm{Card}(G)} \sum_{k \in G} \varphi(hk^{-1}h^{-1}) \rho(hk) \\ &= \rho(h) \circ \frac{1}{\mathrm{Card}(G)} \sum_{k \in G} \varphi(k^{-1}) \rho(k) = \rho(h) \circ \rho_\varphi\end{aligned}$$

si φ est une fonction centrale. On a donc $\rho_\varphi \in \mathcal{L}_G(E, E)$. Si de plus (ρ, E) est irréductible, le lemme de Schur nous dit alors que $\mathcal{L}_G(E, E)$ est de dimension 1 et ρ_φ est une homothétie, soit $\rho_\varphi = \lambda Id_E$ et :

$$\begin{aligned}\lambda = \frac{\mathrm{Tr}(\rho_\varphi)}{\dim(E)} &= \frac{1}{\dim(E)} \frac{1}{\mathrm{Card}(G)} \sum_{g \in G} \varphi(g^{-1}) \chi(g) \\ &= \frac{1}{\dim(E)} \frac{1}{\mathrm{Card}(G)} \sum_{g \in G} \overline{\varphi(g)} \chi(g) = \frac{\langle \varphi \mid \chi \rangle}{\dim(E)}\end{aligned}$$

□

On a vu que les caractères irréductibles de G forment un système orthonormé de l'espace des fonctions centrales (théorème 1.6) et donc, il y a au plus $\mathrm{Card}(G)$ caractères irréductibles sur G. En fait ils en forment une base.

Théorème 1.7.

Les caractères irréductibles de G forment une base orthonormée de l'espace $\mathcal{H}$ des fonctions centrales sur G et le nombre de ces caractères irréductibles (qui est la dimension de $\mathcal{H}$) est égal au nombre des classes de conjugaison de G.

Preuve. Soit $\mathcal{F}$ le sous-espace vectoriel de $\mathcal{H}$ engendré par tous les caractères irréductibles de G, $\chi_1, \cdots, \chi_p$. Comme ces caractères forment un système orthonormé, ils sont linéairement indépendants et pour montrer qu'ils forment une base de $\mathcal{H}$, il nous suffit de prouver que l'orthogonal de $\mathcal{F}$ est réduit à $\{0\}$. Si $\varphi \in \mathcal{H}$ est orthogonal à $\mathcal{F}$, on a alors $\langle \varphi \mid \chi \rangle = 0$ pour tout caractère irréductible et la fonction ρ_φ associée par le procédé du lemme précédent à un tel caractère est nulle. Si (ρ, E) est une représentation de G, elle s'écrit alors $(\rho, E) = \bigoplus_{j=1}^{r} (\rho_j, E_j)$, où les (ρ_j, E_j) sont des représentions irréductibles de G. En notant $\rho_{j,\varphi}$ l'application associée à φ et (ρ_j, E_j) par le procédé du lemme précédent, on a pour tout $x = \sum_{j=1}^{r} x_j$, avec $x_j \in E_j$, $\rho_\varphi(x) = \sum_{j=1}^{r} \rho_\varphi(x_j) = \sum_{j=1}^{r} \rho_{j,\varphi}(x_j) = 0$. On a donc $\rho_\varphi = 0$ pour toute représentation de G. En utilisant la représentation régulière de G, on a $\sum_{g \in G} \varphi(g^{-1}) \rho(g) = 0$, les $\rho(g)$, pour $g \in G$ étant linéairement indépendants dans $GL(\mathbb{C}^G)$ (en effet si $\sum_{g \in G} \lambda_g \rho(g) = 0$, on a alors $\sum_{g \in G} \lambda_g \rho(g)(e_1) = \sum_{g \in G} \lambda_g e_g = 0$ et

$\lambda_g = 0$ pour tout $g \in G$), ce qui nous donne $\varphi\left(g^{-1}\right) = 0$ pour $g \in G$, c'est-à-dire que $\varphi = 0$. Comme $\mathcal{H}$ est l'espace des fonctions constantes sur les classes de conjugaison de G, sa dimension p est égale au nombre de ces classes. □

Pour la suite de ce paragraphe, on note :

— p la dimension de l'espace $\mathcal{H}$ des fonctions centrales sur G ;

— $(\chi_1, \cdots, \chi_p)$ la base de $\mathcal{H}$ formée de tous les caractères irréductibles de G ;

— $\overline{g_1}, \cdots, \overline{g_p}$ toutes les classes de conjugaison de G.

On rappelle qu'une fonction centrale φ est constante sur chaque classe de conjugaison $\overline{g_i}$, cette valeur constante étant $\varphi(g_i)$.

Théorème 1.8.

Pour $1 \leq i, j \leq p$, *on a :*

$$\sum_{g \in G} \overline{\chi_i(g)} \chi_j(g) = \begin{cases} \operatorname{Card}(G) & \textit{si } i = j \\ 0 \ \textit{si } i \neq j \end{cases}$$

$$\sum_{k=1}^{p} \overline{\chi_k(g_i)} \chi_k(g_j) = \begin{cases} \dfrac{\operatorname{Card}(G)}{\operatorname{Card}(\overline{g_i})} & \textit{si } i = j \\ 0 \ \textit{si } i \neq j \end{cases}$$

Preuve. Les premières formules traduisent simplement le fait que la famille $(\chi_i)_{1 \leq i \leq p}$ est orthogonale. Pour tout entier i compris entre 1 et p, on désigne par δ_i la fonction caractéristique de $\overline{g_i}$. Ces fonctions δ_i étant des fonctions centrales, elles s'écrivent $\delta_i = \sum\limits_{k=1}^{p} \lambda_k \chi_k$, avec :

$$\lambda_k = \langle \delta_i \mid \chi_k \rangle = \frac{1}{\operatorname{Card}(G)} \sum_{g \in G} \delta_i(g) \chi_k\left(g^{-1}\right) = \frac{1}{\operatorname{Card}(G)} \sum_{g \in \overline{g_i}} \chi_k\left(g^{-1}\right)$$

et pour tout $g = h g_i h^{-1}$ dans $\overline{g_i}$, on a $\chi_k\left(g^{-1}\right) = \chi_k\left(g_i^{-1}\right) = \overline{\chi_k(g_i)}$, donc $\lambda_k = \dfrac{1}{\operatorname{Card}(G)} \operatorname{Card}(\overline{g_i}) \overline{\chi_k(g_i)}$, soit $\delta_i = \dfrac{\operatorname{Card}(\overline{g_i})}{\operatorname{Card}(G)} \sum\limits_{k=1}^{p} \overline{\chi_k(g_i)} \chi_k$. On en déduit que $1 = \delta_i(g_i) = \dfrac{\operatorname{Card}(\overline{g_i})}{\operatorname{Card}(G)} \sum\limits_{k=1}^{p} \overline{\chi_k(g_i)} \chi_k(g_i)$, soit $\sum\limits_{k=1}^{p} |\chi_k(g_i)|^2 = \dfrac{\operatorname{Card}(G)}{\operatorname{Card}(\overline{g_i})}$. Et pour $j \neq i$, on a $0 = \delta_i(g_j) = \frac{\operatorname{Card}(\overline{g_i})}{\operatorname{Card}(G)} \sum\limits_{k=1}^{p} \overline{\chi_k(g_i)} \chi_k(g_j)$, soit $\sum\limits_{k=1}^{p} \overline{\chi_k(g_i)} \chi_k(g_j) = 0$.

□

Corollaire 1.2. *En désignant, pour tout i compris entre 1 et p, par (ρ_i, E_i) une représentation irréductible de G de caractère χ_i, on a :*

$$\sum_{k=1}^{p} (\dim(E_k))^2 = \operatorname{Card}(G)$$

Preuve. Prenant $g_i = g_j = e$, on a $\overline{g_i} = \{e\}$ et :

$$\sum_{k=1}^{p} (\dim (E_k))^2 = \sum_{k=1}^{p} |\chi_k (g_i)|^2 = \frac{\mathrm{Card}\,(G)}{\mathrm{Card}\,(\overline{g_i})} = \mathrm{Card}\,(G)$$

□

Corollaire 1.3. *Il y a p représentations irréductibles de G, à isomorphisme près. Précisément, en désignant, pour tout i compris entre 1 et p, par (ρ_i, E_i) une représentation irréductible de G de caractère χ_i, ces représentations sont non G-isomorphes et toute représentation irréductible de G est G-isomorphe à l'une des représentations (ρ_i, E_i).*

Preuve. Soit (ρ, E) une représentation irréductible de G. Son caractère χ_ρ est l'un des caractères de base χ_k, où k est compris entre 1 et r. On a donc $\langle \chi_\rho \mid \chi_\rho \rangle = 1$ et les représentations (ρ, E) et (ρ_k, E_k) sont équivalentes (théorème 1.6). Pour $1 \leq j \neq k \leq r$, on a $\chi_j \neq \chi_k$ et les les représentations (ρ_j, E) et (ρ_k, E_k) ne sont pas équivalentes. □

Théorème 1.9.

Soient (ρ, E), (ρ', E') deux représentations de G décomposées en sommes directes de représentations irréductibles :

$$(\rho, E) = \bigoplus_{j=1}^{m} (\tau_j, F_j) \ \text{ et } \ (\rho', E') = \bigoplus_{j=1}^{m'} \left(\tau'_j, F'_j\right)$$

Les trois assertions suivantes sont équivalentes :

1. *les représentations (ρ, E) et (ρ', E') sont équivalentes ;*
2. *$\chi_\rho = \chi_{\rho'}$ (les représentations ont même caractère) ;*
3. *$m = m'$ et il existe une permutation $\sigma \in \mathcal{S}_n$ telle que, pour tout j compris entre 1 et m, $\left(\tau'_j, F'_j\right)$ est équivalente à $\left(\tau_{\sigma(j)}, F_{\sigma(j)}\right)$.*

Preuve. Pour tout i compris entre 1 et p, (ρ_i, E_i) est une représentation irréductible de G de caractère χ_i. En désignant, pour i compris entre 1 et p, par m_i [resp. m'_i] le nombre de composantes irréductibles (τ_j, F_j) de (ρ, E) [resp. $\left(\tau'_j, F'_j\right)$ de (ρ', E')] équivalentes à (ρ_i, E_i), on a $\chi_\rho = \sum_{i=1}^{p} m_i \chi_i$ et $\chi_{\rho'} = \sum_{i=1}^{p} m'_i \chi_i$ et tenant compte de l'orthonormalité des χ_i, on a :

$$m_i = \langle \chi_\rho \mid \chi_i \rangle, \ m'_i = \langle \chi_{\rho'} \mid \chi_i \rangle, \ (1 \leq i \leq p)$$

On sait déjà que si les représentations (ρ, E) et (ρ', E') sont équivalentes, elles ont alors même caractère. Si $\chi_\rho = \chi_{\rho'}$, on a alors $m_i = m'_i$ pour tout i compris entre 1 et p, donc $m = \sum_{i=1}^{p} m_i = m'$ et les représentations (ρ, E) et (ρ', E') sont équivalentes à une même représentation $\bigoplus_{i=1}^{p}\bigoplus_{k=1}^{m_i}(\rho_k, E_k)$ $(\bigoplus_{k=1}^{m_i}(\rho_k, E_k) = \{0\}$ si $m_i = 0)$ et en conséquence elles sont équivalentes. □

Avec les notations de la démonstration précédente, on dit que, pour i compris entre 1 et p, m_i est la multiplicité de (ρ_i, E_i) dans (ρ, E).

Corollaire 1.4. *Les caractères de G sont les fonctions centrales de la forme $\chi = \sum_{i=1}^{p} m_i\chi_i$, où les m_i sont des entiers naturel.*
Si χ est un caractère non nul de G, $\langle \chi \mid \chi \rangle$ est alors un entier naturel non nul qui vaut 1 si, et seulement si, χ et irréductible.

Preuve. Soit χ un caractère de G. On a vu que les m_i sont les composantes de $\chi \in \mathcal{H}$ dans la base orthonormée $(\chi_i)_{1\leq i\leq p}$, donc $\chi = \sum_{i=1}^{p} m_i\chi_i$ et $\langle \chi \mid \chi \rangle = \sum_{i=1}^{p} m_i^2$, c'est donc un entier naturel non nul si $\chi \neq 0$. Si χ et irréductible, c'est l'un des χ_i et $\langle \chi \mid \chi \rangle = 1$. Réciproquement si $\langle \chi \mid \chi \rangle = 1$, on a alors $\sum_{i=1}^{p} m_i^2 = 1$, donc tous les m_i sont nuls nuls, sauf un qui vaut 1 et χ et irréductible. □

1.5 Caractères des groupes abéliens finis

On suppose toujours que le groupe G est fini et que $\mathbb{K} = \mathbb{C}$.

Théorème 1.10.

Le groupe G est abélien si, et seulement si, tous ses caractères irréductibles sont de degré 1.

Preuve. Pour $\mathbb{K} = \mathbb{C}$ et G abélien fini, toute représentation irréductible de G est nécessairement de degré 1 (exercice 1.3). On a vu que le nombre de représentations irréductibles de G est égal au nombre p de classes de conjugaisons et notant $((\rho_i, E_i))_{1\leq i\leq p}$ la famille de ces représentations, on a $\sum_{k=1}^{p} (\dim(E_k))^2 = \mathrm{Card}(G)$. Dans le cas où G est commutatif, les classes de conjugaison sont réduites à un élément, donc $p = \mathrm{Card}(G)$ et nécessairement $\dim(E_k) = 1$ pour tout k compris entre 1 et p. □

On se propose de démontrer que tout groupe abélien fini est produit de groupes cycliques. Au paragraphe **??** on donne une autre démonstration de ce résultat.

Lemme 1.9 *Tout groupe abélien d'ordre p^α où p est un nombre premier est produit de groupes cycliques.*

Preuve. Soit G un groupe abélien d'ordre p^α. Si $x \neq 1 \in G$, x est d'ordre p^k avec $k \leq \alpha$. Si G admet un élément d'ordre p^α, alors il est cyclique. Sinon soit x_0 un élément d'ordre maximal $\beta < \alpha$. On a $x_1 = x_0^{p^{\beta-1}} \neq 1$ de sorte qu'il existe un caractère χ de G tel que $\chi(x_1) \neq 1$ en effet, d'après ce qui précède, si $\chi(x_1) = 1$ pour tout caractère irréductible de G on a pour toute fonction f sur G :

$$f(1) = \sum_{\chi \in \hat{G}} < f, \chi > \chi(1) = \sum_{\chi \in \hat{G}} < f, \chi > \chi(x_1) = f(x_1)$$

il suffit alors de prendre $f = \delta_{x_1}$ pour obtenir une contradiction. Le nombre complexe $\chi(x_0)$ est une racine p^β-ième de l'unité et c'est une racine primitive puisque $\chi(x_1) \neq 1$. L'homomorphisme χ est surjectif et envoie x_0 (qui est d'ordre p^β) sur un générateur du groupe des racines p^β-ième de l'unité : le groupe H engendré par x_0 est cyclique d'ordre p^β. Si $y \in G$ est d'ordre p^γ avec $\gamma \leq \beta$, $\chi(y)$ est une racine de l'unité d'ordre p^γ c'est donc aussi une racine p^β-ième de l'unité : il existe donc $h \in H$ tel que $\chi(h) = \chi(y)$. Il en résulte que $\chi(h^{-1}y) = 1$ *i.e.* que $h^{-1}y \in N = \ker \chi$. Pour tout $y \in G$ il existe donc $h \in H$ et $k \in K = \ker \chi$ tels que $y = hk$. Comme $H \cap K = \{1\}$ on en déduit que G est produit direct de H et K(utiliser $(h, k) \in H \times K \mapsto hk$). Or K est un groupe abélien d'ordre strictement inférieur à celui de G. Une récurrence simple montre alors que G est produit de groupes cycliques (d'ordres p^{α_i}). □

On a $\prod p^{\alpha_i} = p^\alpha$ d'où $\alpha = \sum \alpha_i$. De plus deux groupes $\prod \frac{\mathbb{Z}}{p^{\alpha_i}\mathbb{Z}}$ et $\prod \frac{\mathbb{Z}}{p^{\alpha'_i}\mathbb{Z}}$ ne sont isomorphes que si (α_i) et (α'_i) donnent la même partition de l'entier α : le nombre de groupes abéliens d'ordre p^α deux à deux non isomorphes est donc égal à $p(\alpha)$. (décompositions de α en somme d'entiers $0 < n_1 \leq \cdots \leq n_p$). On a $p(4) = 5$, $p(5) = 7$, $p(6) = 11$, *etc* ... Ce nombre croît très vite et peut se calculer par récurrence (on a $P(n) \sim \dfrac{\exp(\pi\sqrt{\frac{2n}{3}})}{4n\sqrt{3}}$ d'après Hardy et Ramanujan). Il permet de calculer le nombre de groupes abéliens d'ordre p^n donc des groupes abéliens finis. Du point de vue des séries formelles on a $1 + \sum p(n)X^n = \prod_k \frac{1}{1 - X^k}$.

Théorème 1.11.

Tout groupe abélien fini est produit de groupes cycliques.

Preuve. Soit G un groupe abélien d'ordre $m \cdot n$ où m et n sont deux entiers premiers entre eux. Soit H l'ensemble des éléments de G dont l'ordre divise m et K celui des éléments dont l'ordre divise n. H et K sont des sous-groupes de G puisque si $x \in H$ est d'ordre k et $y \in H$ d'ordre ℓ leur produit est d'ordre ppcm(k, ℓ) (car G est abélien) ordre qui divise m. Le même raisonnement s'applique à K. Les entiers m et n étant premiers entre eux, Tout élément de $H \cap K$ est d'ordre 1 donc $H \cap K = \{1\}$. De plus, $\varphi : (H, K) \to G$ défini par $\varphi(h, k) = hk$ est un morphisme injectif de groupes. Montrons qu'il est surjectif : m et n étant premiers entre eux, il existe des entiers u et v tels que $um + vn = 1$ de sorte que si $g \in G$ on a $g^{um+vn} = (g^n)^u (g^m)^v$. Mais l'ordre de $(g^n)^u$ divise m d'où $g^{un} \in H$ et l'ordre de $(g^m)^v$ divise n et $g^{vm} \in K$. On a donc $G = H \times K$. Il suffit alors de raisonner par récurrence pour voir que tout groupe abélien fini s'écrit comme produit de groupes cycliques. □

1.6 Exercices

Exercice 1.1. *Soit (ρ, E) une représentation de G de degré fini. Montrer que l'application $\rho^* : G \to GL(E^*)$ définie par $\rho^*(g) = {}^t\left(\rho\left(g^{-1}\right)\right)$ pour tout $g \in G$ est une représentation de G sur E^* de même degré.*

Solution. Pour tout $g \in G$, on a $\rho\left(g^{-1}\right) \in GL(E)$ et en dimension finie, sa transposée est dans $GL(E^*)$. Pour tout $g \in G$ et tout $\ell \in E^*$, on a :

$$\rho^*(g)(\ell) = {}^t\rho\left(g^{-1}\right)(\ell) = \ell \circ \rho\left(g^{-1}\right) = \ell \circ \rho(g)^{-1}$$

et pour $h \in G$:

$$\begin{aligned}\rho^*(gh)(\ell) &= \ell \circ \rho(gh)^{-1} = \ell \circ (\rho(g) \circ \rho(h))^{-1} \\ &= \ell \circ \rho(h)^{-1} \circ \rho(g)^{-1} = (\rho^*(g) \circ \rho^*(h))(\ell)\end{aligned}$$

donc ρ^* est bien un morphisme de groupes de G dans $GL(E^*)$.

Exercice 1.2. *Soit (ρ, E) une représentation de G. Dans le cas où le groupe G est fini et E est un $\mathbb{C}$-espace vectoriel de dimension finie, montrer que chaque isomorphisme $\rho(g)$ est diagonalisable et que ses valeurs propres sont des racines de l'unité.*

Solution. Comme G est fini, tout ses élément sont d'ordre fini, donc pour $g \in G$ d'ordre $r \geq 1$, on a $g^r = e$ et $\rho(g)^r = \rho(g^r) = \rho(e) = Id_E$, c'est-à-dire que l'endomorphisme $\rho(g)$ est annulé par le polynôme $X^r - 1$ qui est scindé à racines simples sur $\mathbb{C}$, en conséquence $\rho(g)$ est diagonalisable et ses valeurs propres sont des racines r-ièmes de l'unité.

Exercice 1.3. *On suppose que $\mathbb{K} = \mathbb{C}$ et que le groupe G est fini et commutatif. Montrer qu'une représentation linéaire irréductible de G dans E est nécessairement de degré 1.*

Solution. Si G est fini et commutatif, l'ensemble $\{\rho(g) \mid g \in G\}$ est alors une famille commutative d'endomorphismes de E diagonalisables (voir l'exercice 1.2) et en conséquence, il existe une base commune de diagonalisation. Si $x \in E\backslash\{0\}$ est un vecteur propre commun à tous les $\rho(g)$, où g décrit G, la droite vectorielle $\mathbb{C}x$ est alors G-invariante non réduite au vecteur nul, donc $E = \mathbb{C}x$ si la représentation est irréductible.

Exercice 1.4. *On propose ici une autre démonstration du théorème 1.2 dans le cas où $\mathbb{K}$ est le corps $\mathbb{R}$ des nombre réels [resp. le corps $\mathbb{C}$ des nombre complexes]. On se donne un groupe fini G et (ρ, E) une représentation de G dans un espace euclidien [resp. hermitien] E de dimension n, le produit* *scalaire euclidien [resp. hermitien] étant noté $\langle \cdot \mid \cdot \rangle$. On notera $\|\cdot\|$ la norme associée.*

1. *Montrer que l'application :*

$$(x, y) \in E^2 \mapsto \langle x \mid y \rangle_\rho = \sum_{g \in G} \langle \rho(g)(x) \mid \rho(g)(y) \rangle$$

définit un produit scalaire euclidien [resp. hermitien] sur E et que pour tout $g \in G$, l'isomorphisme $\rho(g)$ est orthogonal [resp. unitaire] pour ce produit scalaire (on dit que la représentation (ρ, E) est orthogonale [resp. unitaire]). On notera $\|\cdot\|_\rho$ la norme associée.

2. *Montrer que si F est un sous-espace vectoriel G-invariant de E, son orthogonal $F^{\perp_\rho}$ relativement à $\langle \cdot \mid \cdot \rangle_\rho$ est aussi G-invariant.*
3. *Montrer que (ρ, E) est somme directe de sous-représentations irréductibles.*

Solution.

1. Pour x, y dans E, on a :

$$\langle x \mid y \rangle_\rho = \sum_{g \in G} \langle \rho(g)(x) \mid \rho(g)(y) \rangle = \sum_{g \in G} \langle \rho(g)(y) \mid \rho(g)(x) \rangle = \langle y \mid x \rangle_\rho$$

dans le cas euclidien et :

$$\langle x \mid y \rangle_\rho = \sum_{g \in G} \langle \rho(g)(x) \mid \rho(g)(y) \rangle = \sum_{g \in G} \overline{\langle \rho(g)(y) \mid \rho(g)(x) \rangle} = \overline{\langle y \mid x \rangle_\rho}$$

dans le cas hermitien. Chaque application $\rho(g)$ étant linéaire, l'application $\langle \cdot \mid \cdot \rangle_\rho$ est bilinéaire. Pour tout $x \in E$, on a $\langle x \mid x \rangle_\rho = \sum\limits_{g \in G} \|\rho(g)(x)\|^2 \geq 0$ et l'égalité est réalisée si, et seulement si, $\rho(g)(x) = 0$ pour tout $g \in G$, ce qui équivaut à $x = 0$ puisque $\rho(G) \in GL(E)$. On a donc bien un produit scalaire euclidien [resp. hermitien] sur E. Pour tout $g \in G$ et tout $x \in E$, on a :

$$\begin{aligned}\|\rho(g)(x)\|_\rho^2 &= \sum_{h \in G} \|\rho(h) \circ \rho(g)(x)\|^2 = \sum_{h \in G} \|\rho(hg)(x)\|^2 \\ &= \sum_{k \in G} \|\rho(k)(x)\|^2 = \|x\|_\rho^2\end{aligned}$$

puisque l'application $h \mapsto hg$ est une permutation de G. Donc $\rho(g)$ est orthogonal [resp. unitaire].

2. Soit F un sous-espace vectoriel G-invariant de E. Comme, pour tout $g \in G$ l'endomorphisme $\rho(g)$ est orthogonal [resp. unitaire] relativement à $\langle \cdot \mid \cdot \rangle_\rho$, on a $\langle \rho(g)(y) \mid x \rangle_\rho = \langle y \mid \rho(g)(x) \rangle_\rho = 0$ pour tout $y \in F^{\perp_\rho}$ et tout $x \in F$ puisque $\rho(g)(x) \in G$ (F est G-invariant), ce qui signifie que $F^{\perp_\rho}$ est G-invariant.
3. Il suffit d'écrire, pour (ρ, E) non irréductible, $E = F \oplus F^{\perp_\rho}$ et de raisonner par récurrence sur la dimension.

Exercice 1.5. *On suppose que E est de dimension n et que G est d'ordre n. Déterminer le caractère de la représentation régulière de G.*

Solution. En désignant par $\mathcal{B} = (e_k)_{k \in G}$ une base de $E = \mathbb{C}^G$ indexée par G, on rappelle que la représentation régulière de G est définie par $\rho(g)(e_k) = e_{gk}$ pour tout $(g, k) \in G^2$. Pour $g \in G$, la matrice $R(g)$ de $\rho(g)$ dans la base $\mathcal{B}$ est donc une matrice de permutation. Pour $g = e$, on a $R(e) = I_n$ et $\chi(g) = \mathrm{Tr}(\rho(e)) = n$, pour $g \neq e$, on a $\rho(g)(e_k) = e_{gk} \neq e_k$ pour tout $k \in G$, c'est-à-dire que $R(g)$ est la matrice d'une permutation sans points fixes, donc ses termes diagonaux sont nuls et sa trace $\chi(g)$ est nulle.

Exercice 1.6. *Calculer le caractère de la représentation par permutation.*

Solution. On rappelle que la représentation par permutation. est définie comme suit. $G = \mathfrak{S}_n$ étant le groupe symétrique d'ordre n, E étant de dimension n et $\mathcal{B} = (e_k)_{1 \leq k \leq n}$ une base de E, cette représentation est définie par :

$$\forall \sigma \in \mathfrak{S}_n,\ \forall k \in \{1, \cdots, n\},\ \rho(\sigma)(e_k) = e_{\sigma(k)}$$

Pour $\sigma \in \mathfrak{S}_n$, la trace de $\rho(\sigma)$ est égale au nombre d'éléments de $\mathcal{B}$ invariants par σ, c'est à dire au nombre de points fixes de la permutation σ.

Exercice 1.7. *Montrer que tout caractère $\chi : \mathcal{S}_n \to \mathbb{C}$ du groupe symétrique $\mathcal{S}_n$ est à valeurs réelles.*

Solution. Une permutation est conjuguée à son inverse, donc $\chi(\sigma) = \chi\left(\sigma^{-1}\right) = \overline{\chi(\sigma)}$ pour tout $\sigma \in \mathcal{S}_n$ et χ est à valeurs réelles.

Exercice 1.8. *On suppose que $\mathbb{K} = \mathbb{C}$. Montrer que si (ρ, E) est une représentation de G et χ_ρ son caractère, le caractère de la représentation (ρ^*, E^*) (voir l'exercice 1.1) est alors $\overline{\chi_\rho}$.*

Solution. On rappelle que la représentation $\rho^* : G \to GL(E^*)$ est définie par $\rho^*(g) = {}^t\rho\left(g^{-1}\right)$ pour tout $g \in G$ et on a, pour tout $g \in G$, en désignant par $R(g)$ la matrice de $\rho(g)$ dans une base de E :

$$\begin{aligned}\chi_{\rho^*}(g) &= \mathrm{Tr}\left({}^t\rho\left(g^{-1}\right)\right) = \mathrm{Tr}\left({}^tR\left(g^{-1}\right)\right) = \mathrm{Tr}\left(R\left(g^{-1}\right)\right) = \mathrm{Tr}\left(\rho\left(g^{-1}\right)\right) \\ &= \chi_\rho\left(g^{-1}\right) = \overline{\chi_\rho(g)}\end{aligned}$$

Exercice 1.9. *Déterminer tous les caractères d'un groupe cyclique G d'ordre n.*

Solution. Si $G = \langle g_0 \rangle$ est cyclique d'ordre n et g_0 un générateur, un caractère χ de G est uniquement déterminé par sa valeur en g_0. Avec $\chi(g_0)^n = \chi(g_0^n) = \chi(e) = 1$, on déduit que $\chi(g_0)$ est une racine n-ième de l'unité, c'est-à-dire qu'il existe un entier r compris entre 0 et $n-1$ tel que $\chi(g_0) = \theta_r = e^{\frac{2ir\pi}{n}}$. Réciproquement pour tout entier r compris entre 0 et $n-1$, l'application :

$$\chi_r : g = g_0^k \in G \mapsto \chi_r(g) = \theta_r^k = e^{\frac{2ikr\pi}{n}}$$

définit un caractère de G. On a donc ainsi tous les caractères de G et il y en a n.

Exercice 1.10. *Soient p un nombre premier impair et χ un caractère du groupe multiplicatif $H = \left(\frac{\mathbb{Z}}{p\mathbb{Z}}\right)^*$ que l'on prolonge à $\frac{\mathbb{Z}}{p\mathbb{Z}}$ en posant $\chi(\overline{0}) = 0$. On note 1 le caractère constant égal à 1 sur H et si χ est un caractère de H, $\overline{\chi}$ est le caractère défini par $\overline{\chi}(\overline{k}) = \overline{\chi(\overline{k})}$, où $\overline{k}$ est la classe de l'entier k modulo p. On note $\omega = e^{\frac{2i\pi}{p}}$ et α est le morphisme du groupe additif $\frac{\mathbb{Z}}{p\mathbb{Z}}$ dans le groupe multiplicatif $\mathbb{C}^*$ défini par $\alpha(\overline{k}) = \omega^k$ pour tout $k \in \mathbb{Z}$. Si χ est un caractère de H, on définit sa somme de Gauss par $G(\chi) = \sum_{g\in H} \alpha(g)\chi(g)$. Si χ et χ' sont deux caractères de G, on définit leur somme de Jacobi par $J(\chi,\chi') = \sum_{g\in H} \chi(g)\chi'(\overline{1}-g)$.*

1. *Déterminer tous les caractères de H.*
2. *Soit χ un caractère de H. Calculer $J(\chi,1)$ et $J(1,\chi)$.*
3. *Soit χ un caractère non constant de H. Calculer $J(\chi,\overline{\chi})$.*
4. *Soient χ et χ' deux caractères de H avec $\chi' \neq \overline{\chi}$. Montrer que $G(\chi)G(\chi') = J(\chi,\chi')G(\chi\chi')$.*
5. *Soit χ un caractère de H. Montrer que $G(\chi)G(\overline{\chi}) = p\chi(-\overline{1})$ et $|G(\chi)| = \sqrt{p}$.*

Solution.

1. Pour $p \geq 3$ premier, le groupe multiplicatif $H = \left(\frac{\mathbb{Z}}{p\mathbb{Z}}\right)^*$ est cyclique d'ordre $p-1$ (théorème **??**) et l'exercice précédent nous dit que ses caractères sont les applications définies par $\chi_r : g = g_0^k \in H \mapsto \chi_r(g) = e^{\frac{2ikr\pi}{p-1}}$, où g_0 est un générateur de H et r, k sont des entiers compris entre 0 et $p-2$.
2. En gardant les notations qui précèdent, on a pour r compris entre 0 et $p-2$ (en faisant attention au fait que $1(\overline{0}) = 0$) ;

$$J(\chi_r, 1) = \sum_{g\in H-\{\overline{1}\}} \chi_r(g) = \sum_{k=0}^{p-2} e^{\frac{2ikr\pi}{p-1}} - 1 = \begin{cases} \frac{1-e^{\frac{2i(p-1)r\pi}{p-1}}}{1-e^{\frac{2ir\pi}{p-1}}} - 1 = -1 \text{ si } r \neq 0 \\ p-2 \text{ si } r = 0 \end{cases}$$

En tenant compte du fait que $\chi_r\left(\overline{0}\right)=0$ et que l'application $g\mapsto\overline{1}-g$ réalise une bijection de $H-\left\{\overline{1}\right\}$ sur lui même, on a :

$$J(1,\chi_r)=\sum_{g\in H-\left\{\overline{1}\right\}}\chi_r\left(\overline{1}-g\right)=\sum_{h\in H-\left\{\overline{1}\right\}}\chi_r(h)=J(\chi_r,1)$$

3. On a pour r compris entre 0 et $p-2$:

$$J\left(\chi_r,\overline{\chi_r}\right)=\sum_{g\in H}\chi_r(g)\,\overline{\chi_r}\left(\overline{1}-g\right)=\sum_{g\in H-\left\{\overline{1}\right\}}\chi(g)\,\overline{\chi}\left(\overline{1}-g\right)$$

En se donnant un générateur g_0 du groupe cyclique H, pour tout $g\in H-\left\{\overline{1}\right\}$ il existe un unique entier k compris entre 1 et $p-2$ tel que $g=g_0^k$ et un unique entier $\nu(k)$ compris entre 0 et $p-2$ tel que $\overline{1}-g=g_0^{\nu(k)}$. L'application ν ainsi définie est alors une injection de $\{1,\cdots,p-2\}$ dans $\{0,\cdots,p-2\}$. On a donc :

$$J\left(\chi_r,\overline{\chi_r}\right)=\sum_{k=1}^{p-2}\chi_r\left(g_0^k\right)\overline{\chi_r}\left(g_0^{\nu(k)}\right)$$

avec $\chi_r(g_0)=e^{\frac{2ir\pi}{p-1}}$, $\chi_r\left(g_0^k\right)=e^{\frac{2ikr\pi}{p-1}}$, $\overline{\chi_r}\left(g_0^{\nu(k)}\right)=\overline{e^{\frac{2i\nu(k)r\pi}{p-1}}}=e^{-\frac{2i\nu(k)r\pi}{p-1}}$ et $\chi_r\left(g_0^k\right)\overline{\chi_r}\left(g_0^{\nu(k)}\right)=e^{-\frac{2i(k-\nu(k))r\pi}{p-1}}=\chi_r\left(g_0^{k-\nu(k)}\right)$, ce qui nous donne :

$$J\left(\chi_r,\overline{\chi_r}\right)=\sum_{k=1}^{p-2}\chi_r\left(g_0^{k-\nu(k)}\right)$$

Des égalités $g=g_0^k$ et $\overline{1}-g=g_0^{\nu(k)}$, on déduit que $g_0^k=\overline{1}-g_0^{\nu(k)}$ et :

$$g_0^{k-\nu(k)}=\left(\overline{1}-g_0^{\nu(k)}\right)g_0^{-\nu(k)}=g_0^{-\nu(k)}-\overline{1}$$

donc $J\left(\chi_r,\overline{\chi_r}\right)=\sum_{k=1}^{p-2}\chi_r\left(g_0^{-\nu(k)}-\overline{1}\right)$. L'application $k\mapsto g_0^{-\nu(k)}-\overline{1}$ réalisant une injection de $\{1,\cdots,p-2\}$ dans H, le seul élément non atteint de H étant $-\overline{1}$ ($g_0^{-\nu(p)}\neq\overline{0}$ car $\overline{0}\notin H$), on en déduit que :

$$J\left(\chi_r,\overline{\chi_r}\right)=\sum_{g\in H-\left\{-\overline{1}\right\}}\chi_r(g)=\sum_{g\in H}\chi_r(g)-\chi\left(-\overline{1}\right)=-\chi\left(-\overline{1}\right)$$

4. On a :

$$G(\chi)\,G(\chi')=\sum_{g\in H}\alpha(g)\,\chi(g)\sum_{h\in H}\alpha(h)\,\chi'(h)=\sum_{g\in H}\chi(g)\sum_{h\in H}\alpha(g+h)\,\chi'(h)$$

(α est un morphisme du groupe additif $\dfrac{\mathbb{Z}}{p\mathbb{Z}}$ dans le groupe multiplicatif $\mathbb{C}^*$).

Pour g fixé dans H l'application $h\mapsto g+h$ étant une permutation de $\dfrac{\mathbb{Z}}{p\mathbb{Z}}$, on peut écrire que :

$$G(\chi)\,G(\chi')=\sum_{g\in H}\chi(g)\sum_{k\in\frac{\mathbb{Z}}{p\mathbb{Z}}}\alpha(k)\,\chi'(k-g)=\sum_{k\in\frac{\mathbb{Z}}{p\mathbb{Z}}}\alpha(k)\sum_{g\in H}\chi(g)\,\chi'(k-g)$$

Pour $k \in \frac{\mathbb{Z}}{p\mathbb{Z}} \setminus \{\overline{0}\} = H$, l'application $h \mapsto kh$ est une permutation de H, donc :

$$\begin{aligned}\sum_{g\in H} \chi(g)\chi'(k-g) &= \sum_{h\in H} \chi(kh)\chi'(k-kh) \\ &= \chi(k)\chi'(k)\sum_{h\in H}\chi(h)\chi'(e-h) = \chi(k)\chi'(k)J(\chi,\chi')\end{aligned}$$

Pour $k = \overline{0}$, on a :

$$\begin{aligned}\sum_{g\in H} \chi(g)\chi'(k-g) = \sum_{g\in H}\chi(g)\chi'(-g) &= \chi'\left(-\overline{1}\right)\sum_{g\in H}\chi(g)\chi'(g) \\ &= \chi'\left(-\overline{1}\right)\sum_{g\in H}(\chi\chi')(g) = 0\end{aligned}$$

si $\chi\chi' \neq 1$, ce qui est réalisé si $\chi' \neq \overline{\chi}$. Il reste donc :

$$G(\chi)G(\chi') = \sum_{k\in H}\alpha(k)\chi(k)\chi'(k)J(\chi,\chi') = J(\chi,\chi')G(\chi\chi')$$

5. On reprend les calculs précédents, avec $\chi' = \overline{\chi}$. Pour $k \in \frac{\mathbb{Z}}{p\mathbb{Z}} - \{\overline{0}\}$ il n'y a pas de changement. Pour $k = \overline{0}$, on a :

$$\begin{aligned}\sum_{g\in H}\chi(g)\overline{\chi}(k-g) = \sum_{g\in H}\chi(g)\overline{\chi}(-g) &= \overline{\chi}\left(-\overline{1}\right)\sum_{g\in G}|\chi(g)|^2 \\ &= (p-1)\overline{\chi}\left(-\overline{1}\right) = (p-1)\chi\left(-\overline{1}\right)\end{aligned}$$

ce qui nous donne $G(\chi)G(\overline{\chi}) = \alpha\left(\overline{0}\right)(p-1)\chi\left(-\overline{1}\right) + J(\chi,\overline{\chi})\sum_{k\in H}\alpha(k)$ avec $\alpha\left(\overline{0}\right) = 1$, $J(\chi,\overline{\chi}) = -\chi\left(-\overline{1}\right)$ et $\sum_{k\in H}\alpha(k) = \sum_{k\in\frac{\mathbb{Z}}{p\mathbb{Z}}}\alpha(k) - \alpha\left(\overline{0}\right) = -1$, donc $G(\chi)G(\overline{\chi}) = p\chi\left(-\overline{1}\right)$. Un calcul analogue donne $G(\chi)\overline{G(\chi)} = p$.

Chapitre 2

Déterminants (nouvelle version du 12/06/2021)

$\mathbb{K}$ est un corps commutatif de caractéristique différente de 2, E un $\mathbb{K}$-espace vectoriel de dimension finie $n \geq 1$, $\mathcal{L}(E)$ est l'algèbre des endomorphismes de E et $GL(E)$ est le groupe des automorphismes de E. $E^* = \mathcal{L}(E, \mathbb{K})$ est l'espace des formes linéaires sur E (voir le chapitre **??**). Pour n, m entiers naturels non nuls, $\mathcal{M}_{m,n}(\mathbb{K})$ est l'espace des matrices à m lignes, n colonnes et à coefficients dans $\mathbb{K}$. Pour $m = n$, on note $\mathcal{M}_n(\mathbb{K})$ l'algèbre des matrices carrées d'ordre n à coefficients dans $\mathbb{K}$, $GL_n(\mathbb{K})$ est le groupe des matrices inversibles dans $\mathcal{M}_n(\mathbb{K})$.

Pour tout entier $n \geq 1$, $\mathcal{S}_n$ est le groupe symétrique et pour toute permutation $\sigma \in \mathcal{S}_n$, $\varepsilon(\sigma)$ est la signature de σ (voir le chapitre **??**).

2.1 Formes multilinéaires alternées

Sauf précision contraire, p est un entier naturel supérieur ou égal à 2.

Définition 2.1. *Une forme p-linéaire sur l'espace E est une application $\varphi : E^p \to \mathbb{K}$ telle que pour tout k compris entre 1 et p et $(x_i)_{\substack{1\leq i\leq p\\ i\neq k}}$ fixé dans E^{p-1} l'application partielle :*

$$\varphi_i : x \in E \mapsto \begin{cases} \varphi(x, x_2, \cdots, x_p) \text{ pour } k = 1 \\ \varphi(x_1, \cdots, x_{k-1}, x, x_{k+1}, \cdots, x_p) \text{ pour } 1 \leq k \leq p-1 \\ \varphi(x_1, \cdots, x_{p-1}, x) \text{ pour } k = p \end{cases}$$

est une forme linéaire sur E. On dit que φ est alternée, si de plus on a $\varphi(x_1, \cdots, x_p) = 0$ pour tout $(x_i)_{1\leq i\leq p} \in E^p$ pour lequel il existe $j \neq k$ compris entre 1 et p tels que $x_i = x_j$.

Pour tout entier $p \geq 1$, on note $\mathcal{L}_p(E, \mathbb{K})$ l'ensemble des formes p-linéaires sur E (pour $p = 1$, $\mathcal{L}_1(E, \mathbb{K}) = E^*$).

Pour tous $\varphi \in \mathcal{L}_p(E, \mathbb{K})$, $(x_i)_{1\leq i\leq p} \in E^p$ et $(\lambda_i)_{1\leq i\leq p} \in \mathbb{K}^p$, on a en utilisant la multilinéarité, $\varphi(\lambda_1 x_1, \cdots, \lambda_p x_p) = \left(\prod_{i=1}^{p} \lambda_i\right) \varphi(x_1, \cdots, x_p)$. Il en résulte que

$\varphi(0, \cdots, 0) = 0$. Dans le cas où tous les λ_i sont égaux à un même scalaire λ, on obtient $\varphi(\lambda x_1, \cdots, \lambda x_p) = \lambda^p \varphi(x_1, \cdots, x_p)$.

Théorème 2.1.

Pour tout $p \in \mathbb{N}^*$, $\mathcal{L}_p(E, \mathbb{K})$ est un $\mathbb{K}$-espace vectoriel de dimension n^p.

Preuve. Il est facile de vérifier que $\mathcal{L}_p(E, \mathbb{K})$ est un sous-espace vectoriel de l'espace $\mathbb{K}^{E^p}$ des applications de E^p dans $\mathbb{K}$.

Soit $\mathcal{B} = (e_i)_{1 \leq i \leq n}$ une base de E. Pour tout $\varphi \in \mathcal{L}_p(E, \mathbb{K})$ et tout $(x_i)_{1 \leq i \leq p}$ dans E^p, en notant pour j compris entre 1 et p, $x_j = \sum_{i=1}^{n} x_{i,j} e_i$ et en utilisant le caractère p-linéaire de φ, on a :

$$\varphi(x_1, \cdots, x_p) = \varphi\left(\sum_{i_1=1}^{p} x_{i_1,1} e_{i_1}, x_2, \cdots, x_p\right) = \sum_{i_1=1}^{p} x_{i_1,1} \varphi(e_{i_1}, x_2 \cdots, x_p)$$

$$= \sum_{i_1=1}^{p}\sum_{i_2=1}^{p} x_{i_1,1} x_{i_2,2} \varphi(e_{i_1}, e_{i_2}, x_3 \cdots, x_p) = \sum_{i_1=1}^{n} \cdots \sum_{i_p=1}^{n} x_{i_1,1} \cdots x_{i_p,p} \varphi\left(e_{i_1}, \cdots, e_{i_p}\right)$$

$$= \sum_{1 \leq i_1, \cdots, i_p \leq n} x_{i_1,1} \cdots x_{i_p,p} \varphi\left(e_{i_1}, \cdots, e_{i_p}\right)$$

Il est clair que l'application :

$$\begin{array}{rccc} \Phi : & \mathcal{L}_p(E, \mathbb{K}) & \to & \mathbb{K}^{n^p} \\ & \varphi & \mapsto & \left(\varphi\left(e_{i_1}, \cdots, e_{i_p}\right)\right)_{1 \leq i_1, \cdots, i_p \leq n} \end{array}$$

est linéaire est injective. Pour tout vecteur $\alpha = \left(\alpha_{i_1, \cdots, i_p}\right)_{1 \leq i_1, \cdots, i_p \leq n} \in \mathbb{K}^{n^p}$, l'application $\varphi : (x_1, \cdots, x_p) \mapsto \sum_{1 \leq i_1, \cdots, i_p \leq n} \alpha_{i_1, \cdots, i_p} x_{i_1,1} \cdots x_{i_p,p}$ est p-linéaire, donc Φ est surjective et c'est un isomorphisme de $\mathcal{L}_p(E, \mathbb{K})$ sur $\mathbb{K}^{n^p}$. Il en résulte que $\dim(\mathcal{L}_p(E, \mathbb{K})) = n^p$. □

Si $\varphi \in \mathcal{L}_p(E, \mathbb{K})$ est alternée, on a alors $\varphi(x_1, \cdots, x_p) = 0$ pour toute famille liée $(x_i)_{1 \leq i \leq p}$ de vecteurs de E. Il en résulte que $\varphi(x_1, \cdots, x_p)$ est inchangé si on ajoute à l'un des vecteurs x_k une combinaison linéaire des autres vecteurs x_j (avec $j \neq k$).

Théorème 2.2.

Une forme p-linéaire φ sur E est alternée si, et seulement si, on a $\varphi\left(x_{\sigma(1)}, \cdots, x_{\sigma(p)}\right) = \varepsilon(\sigma) \varphi(x_1, \cdots, x_p)$ pour toute permutation $\sigma \in \mathcal{S}_p$ et tout $(x_1, \cdots, x_p) \in E^p$.

Preuve. La signature étant un morphisme de groupes de $\mathcal{S}_p$ sur $\{-1, 1\}$ et le groupe $\mathcal{S}_p$ étant engendré par les transpositions, il revient au même de montrer que $\varphi \in \mathcal{L}_p(E, \mathbb{K})$ est alternée si, et seulement si, on a $\varphi\left(x_{\tau(1)}, \cdots, x_{\tau(p)}\right) = -\varphi(x_1, \cdots, x_p)$ pour toute transposition τ.

Supposons que $\varphi \in \mathcal{L}_p(E, \mathbb{K})$ soit alternée et soit $\tau = (j, k)$ une transposition avec $1 \leq j < k \leq p$. En écrivant que :

$$\begin{aligned}
0 &= \varphi(x_1, \cdots, x_j + x_k, \cdots, x_j + x_k, \cdots, x_p) \\
&= \varphi(x_1, \cdots, x_j, \cdots, x_j, \cdots, x_p) + \varphi(x_1, \cdots, x_p) \\
&+ \varphi\left(x_{\tau(1)}, \cdots, x_{\tau(p)}\right) + \varphi(x_1, \cdots, x_k, \cdots, x_k, \cdots, x_p) \\
&= \varphi(x_1, \cdots, x_p) + \varphi\left(x_{\tau(1)}, \cdots, x_{\tau(p)}\right)
\end{aligned}$$

on déduit que $\varphi\left(x_{\tau(1)}, \cdots, x_{\tau(p)}\right) = -\varphi(x_1, \cdots, x_p)$. Réciproquement, supposons cette condition vérifiée. Si $x_j = x_k$ pour deux indices $j < k$ compris entre 1 et p, on a alors, pour $\tau = (j, k)$:

$$\varphi(x_1, \cdots, x_p) = \varphi\left(x_{\tau(1)}, \cdots, x_{\tau(p)}\right) = -\varphi(x_1, \cdots, x_p)$$

et $\varphi(x_1, \cdots, x_p) = 0$ pour $\mathbb{K}$ de caractéristique différente de 2. □

Le résultat qui suit nous montre comment transformer une forme p-linéaire en forme p-linéaire alternée.

Théorème 2.3.

Si $\varphi \in \mathcal{L}_p(E, \mathbb{K})$ est une forme p-linéaire, l'application $\widetilde{\varphi}$ définie sur E^p par :

$$\widetilde{\varphi}(x_1, \cdots, x_p) = \sum_{\sigma \in \mathcal{S}_p} \varepsilon(\sigma) \varphi\left(x_{\sigma(1)}, \cdots, x_{\sigma(p)}\right)$$

est alors une forme p-linéaire alternée sur E.

Preuve. Pour tout $\sigma \in \mathcal{S}_p$, l'application $(x_1, \cdots, x_n) \mapsto \varphi\left(x_{\sigma(1)}, \cdots, x_{\sigma(p)}\right)$ est une forme p-linéaire, donc il en est de même de $\widetilde{\varphi}$. Pour toute permutation $\tau \in \mathcal{S}_p$ et tout $(x_1, \cdots, x_p) \in E^p$, on a :

$$\begin{aligned}
\widetilde{\varphi}\left(x_{\tau(1)}, \cdots, x_{\tau(p)}\right) &= \sum_{\sigma \in \mathcal{S}_p} \varepsilon(\sigma) \varphi\left(x_{\sigma\circ\tau(1)}, \cdots, x_{\sigma\circ\tau(p)}\right) \\
&= \sum_{\sigma' \in \mathcal{S}_p} \varepsilon\left(\sigma' \circ \tau^{-1}\right) \varphi\left(x_{\sigma'(1)}, \cdots, x_{\sigma'(p)}\right) \\
&= \sum_{\sigma' \in \mathcal{S}_p} \varepsilon(\sigma') \varepsilon\left(\tau^{-1}\right) \varphi\left(x_{\sigma'(1)}, \cdots, x_{\sigma'(p)}\right) = \varepsilon(\tau) \widetilde{\varphi}(x_1, \cdots, x_p)
\end{aligned}$$

(l'application $\sigma \mapsto \sigma \circ \tau$ réalise une permutation de $\mathcal{S}_p$ et $\varepsilon\left(\tau^{-1}\right) = \varepsilon(\tau)$). Il en résulte que $\widetilde{\varphi}$ est une forme p-linéaire alternée sur E. □

2.2 Déterminants

On se donne une base $\mathcal{B} = (e_i)_{1 \leq i \leq n}$ de E et pour tout $x = \sum_{i=1}^{n} x_i e_i \in E$, on note $X = (x_i)_{1 \leq i \leq n}$ dans $\mathbb{K}^n$.

Théorème 2.4.

L'espace vectoriel $\mathcal{A}_n(E,\mathbb{K})$ des formes n-linéaires alternées sur E est de dimension 1 engendré par l'application $\det_{\mathcal{B}} : E^n \to \mathbb{K}$ définie par :

$$\det_{\mathcal{B}}(x_1,\cdots,x_n) = \sum_{\sigma\in\mathcal{S}_n} \varepsilon(\sigma) \prod_{i=1}^{n} x_{\sigma(i),i}$$

où $x_j = \sum_{i=1}^{n} x_{ij} e_i$ pour tout j compris entre 1 et n.

Preuve. Vérifions tout d'abord que l'application $\det_{\mathcal{B}}$ est n-linéaire alternée. En notant, pour tout j compris entre 1 et n et tout $x = \sum_{i=1}^{n} x_i e_i \in E$, $\pi_j(x) = x_j$, on a $\det_{\mathcal{B}}(x_1,\cdots,x_n) = \sum_{\sigma\in\mathcal{S}_n} \varepsilon(\sigma) \prod_{i=1}^{n} \pi_{\sigma(i)}(x_i)$. Chaque application $\pi_{\sigma(i)}$ étant linéaire, l'application $(x_i)_{1\le i\le n} \mapsto \prod_{i=1}^{n} \pi_{\sigma(i)}(x_i)$ est n-linéaire et il en est de même de $\det_{\mathcal{B}}$ comme combinaison linéaire d'applications n-linéaires. Pour tout permutation τ, en effectuant le changement d'indice $k = \tau(i)$, on a :

$$\det_{\mathcal{B}}\left(x_{\tau(1)},\cdots,x_{\tau(n)}\right) = \sum_{\sigma\in\mathcal{S}_n} \varepsilon(\sigma) \prod_{i=1}^{n} \pi_{\sigma(i)}\left(x_{\tau(i)}\right) = \sum_{\sigma\in\mathcal{S}_n} \varepsilon(\sigma) \prod_{k=1}^{n} \pi_{\sigma\circ\tau^{-1}(k)}(x_k)$$

et en utilisant le fait que l'application $\sigma' \mapsto \sigma = \sigma'\circ\tau$ est une bijection de $\mathcal{S}_n$ sur lui même, on en déduit que :

$$\begin{aligned}
\det_{\mathcal{B}}\left(x_{\tau(1)},\cdots,x_{\tau(n)}\right) &= \sum_{\sigma'\in\mathcal{S}_n} \varepsilon(\sigma'\circ\tau) \prod_{k=1}^{n} \pi_{\sigma'(k)}(x_k) \\
&= \varepsilon(\tau) \sum_{\sigma'\in\mathcal{S}_n} \varepsilon(\sigma') \prod_{k=1}^{n} \pi_{\sigma'(k)}(x_k) = \varepsilon(\tau) \det_{\mathcal{B}}(x_1,\cdots,x_n)
\end{aligned}$$

ce qui signifie que $\det_{\mathcal{B}}$ est alternée. Pour $\varphi \in \mathcal{A}_n(E,\mathbb{K})$ et $(x_1,\cdots,x_n) \in E^n$, on a :

$$\begin{aligned}
\varphi(x_1,\cdots,x_n) &= \sum_{1\le i_1,\cdots,i_n\le n} x_{i_1,1}\cdots x_{i_n,n}\varphi(e_{i_1},\cdots,e_{i_n}) \\
&= \sum_{\sigma\in\mathcal{F}_n} \prod_{i=1}^{n} x_{\sigma(i),i}\varphi\left(e_{\sigma(1)},\cdots,e_{\sigma(n)}\right)
\end{aligned}$$

où $\mathcal{F}_n$ est l'ensemble des applications de $\{1, \cdots, n\}$ dans $\{1, \cdots, n\}$. Comme φ est alternée, on a $\varphi\left(x_{\sigma(1)}, \cdots, x_{\sigma(n)}\right) = 0$ pour σ non bijective et :

$$\begin{aligned}\varphi(x_1, \cdots, x_n) &= \sum_{\sigma \in \mathcal{S}_n} \prod_{i=1}^{n} x_{\sigma(i),i} \varphi\left(e_{\sigma(1)}, \cdots, e_{\sigma(n)}\right) \\ &= \left(\sum_{\sigma \in \mathcal{S}_n} \varepsilon(\sigma) \prod_{i=1}^{n} x_{\sigma(i),i}\right) \varphi(e_1, \cdots, e_n)\end{aligned}$$

soit $\varphi = \lambda \det_{\mathcal{B}}$ avec $\lambda = \varphi(e_1, \cdots, e_n) \in \mathbb{K}$ et $\det_{\mathcal{B}} \in \mathcal{A}_n(E, \mathbb{K}) \setminus \{0\}$. En conclusion, $\mathcal{A}_n(E, \mathbb{K})$ est de dimension 1 engendré par $\det_{\mathcal{B}}$. □

Avec la démonstration précédente, on constate que $\det_{\mathcal{B}}$ est l'unique forme n-linéaire alternée sur E telle que $\varphi(e_1, \cdots, e_n) = 1$.

Définition 2.2. *Avec les notations qui précèdent, on dit que* $\det_{\mathcal{B}}(x_1, \cdots, x_n)$ *est le déterminant dans la base* $\mathcal{B}$ *du* n*-uplet de vecteurs* $(x_i)_{1 \leq i \leq n}$.

Pour $\varphi \in \mathcal{A}_n(E, \mathbb{K})$, on a $\varphi = \lambda \det_{\mathcal{B}}$ avec $\lambda = \varphi(e_1, \cdots, e_n)$ et en conséquence, $\varphi(x_1, \cdots, x_n) = \varphi(e_1, \cdots, e_n) \det_{\mathcal{B}}(x_1, \cdots, x_n)$ pour tout $(x_i)_{1 \leq i \leq n} \in E^n$. On en déduit la formule de changement de base qui suit.

Théorème 2.5. Relation de Chasles

Si $\mathcal{B}' = (e'_i)_{1 \leq i \leq n}$ *est une autre base de* E, *on a alors pour pour tout* $(x_i)_{1 \leq i \leq n} \in E^n$ *:*

$$\begin{aligned}\det_{\mathcal{B}'}(x_1, \cdots, x_n) &= \det_{\mathcal{B}'}(e_1, \cdots, e_n) \det_{\mathcal{B}}(x_1, \cdots, x_n) \\ &= \det_{\mathcal{B}'}(\mathcal{B}) \det_{\mathcal{B}}(x_1, \cdots, x_n)\end{aligned}$$

On déduit du résultat précédent que $\det_{\mathcal{B}'}(\mathcal{B}) \det_{\mathcal{B}}(\mathcal{B}') = \det_{\mathcal{B}'}(\mathcal{B}') = 1$.

Théorème 2.6.

Soit $\mathcal{B}' = (x_i)_{1 \leq i \leq n}$ *un* n*-uplet de vecteurs de* E. *Les assertions suivantes sont équivalentes :*

1. *la famille* $\mathcal{B}'$ *est liée ;*
2. *pour toute base* $\mathcal{B}$ *de* E, $\det_{\mathcal{B}}(\mathcal{B}') = 0$ *;*
3. *il existe une base* $\mathcal{B}$ *de* E *telle que* $\det_{\mathcal{B}}(\mathcal{B}') = 0$.

Preuve. (1) $\Rightarrow$ (2) Si la famille $\mathcal{B}'$ est liée, on a alors $\varphi(x_1, \cdots, x_n) = 0$ pour toute forme n-linéaire alternée et c'est en particulier vrai pour $\det_{\mathcal{B}}$, quelle que soit la base $\mathcal{B}$ de E.

(2) $\Rightarrow$ (3) est évident.

(3) $\Rightarrow$ (1) Soit $\mathcal{B}$ une base de E telle que $\det_{\mathcal{B}}(\mathcal{B}') = 0$. Si la famille $\mathcal{B}'$ est libre, c'est alors une base de E et on a $1 = \det_{\mathcal{B}'}(\mathcal{B}') = \det_{\mathcal{B}'}(\mathcal{B}) \det_{\mathcal{B}}(\mathcal{B}') = 0$, ce qui est impossible. □

Corollaire 2.1. *Soit $\mathcal{B}' = (x_i)_{1 \leq i \leq n}$ un n-uplet de vecteurs de E. Cette famille est une base de E si, et seulement si, il existe une base $\mathcal{B}$ de E telle que* $\det_{\mathcal{B}}(\mathcal{B}') \neq 0$.

Le résultat qui suit nous permet de définir le déterminant d'un endomorphisme de E.

Théorème 2.7.

Pour tout endomorphisme $u \in \mathcal{L}(E)$, il existe un unique scalaire λ_u tel que pour toute forme $\varphi \in \mathcal{A}_n(E, \mathbb{K}) \setminus \{0\}$ et tout $(x_i)_{1 \leq i \leq n} \in E^n$, on a :

$$\varphi(u(x_1), \cdots, u(x_n)) = \lambda_u \varphi(x_1, \cdots, x_n)$$

On a $\lambda_u = \det_{\mathcal{B}}(u(e_1), \cdots, u(e_n))$, où $\mathcal{B} = (e_i)_{1 \leq i \leq n}$ est une base quelconque de E.

Preuve. Pour tout endomorphisme $u \in \mathcal{L}(E)$ et toute forme n-linéaire alternée $\varphi \in \mathcal{A}_n(E, \mathbb{K}) \setminus \{0\}$, l'application $\varphi \circ u : (x_i)_{1 \leq i \leq n} \in E^n \mapsto \varphi(u(x_i))_{1 \leq i \leq n}$ est aussi une forme n-linéaire alternée, donc il existe un scalaire λ_u tel que $\varphi \circ u = \lambda_u \varphi$ ($\dim(\mathcal{A}_n(E, \mathbb{K})) = 1$ et φ est non nulle). Si $\psi \in \mathcal{A}_n(E, \mathbb{K}) \setminus \{0\}$ est une autre forme n-linéaire alternée non nulle, on a alors $\psi = \rho\varphi$ et $\psi \circ u = \rho\varphi \circ u = \rho\lambda_u\varphi = \lambda_u\psi$, c'est-à-dire que le scalaire λ_u ne dépend pas de la forme n-linéaire alternée non nulle choisie. Prenant $\varphi = \det_{\mathcal{B}}$ puis $\varphi = \det_{\mathcal{B}'}$, où $\mathcal{B} = (e_i)_{1 \leq i \leq n}$ et $\mathcal{B}' = (e'_i)_{1 \leq i \leq n}$ sont des bases de E, on a $\det_{\mathcal{B}}(u(e_1), \cdots, u(e_n)) = \lambda_u = \det_{\mathcal{B}'}(u(e'_1), \cdots, u(e'_n))$. □

Définition 2.3. *Avec les notations du théorème précédent, on dit que le scalaire λ_u est le déterminant de u et on le note* $\det(u)$.

Le scalaire $\det(u)$ ne dépend que de u et pas du choix d'une base de E.

Si $(x_i)_{1 \leq i \leq n} \in E^n$ et $u \in \mathcal{L}(E)$ est défini par $u(e_i) = x_i$ pour i compris entre 1 et n, on a alors $\det(u) = \det_{\mathcal{B}}(u(e_1), \cdots, u(e_n)) = \det_{\mathcal{B}}(x_1, \cdots, x_n)$, où $\mathcal{B} = (e_i)_{1 \leq i \leq n}$ est une base quelconque de E.

Théorème 2.8.

On a $\det(Id) = 1$ et pour u, v dans $\mathcal{L}(E)$, $\lambda \in \mathbb{K}$, on a :

$$\det(\lambda u) = \lambda^n \det(u), \ \det(u \circ v) = \det(v \circ u) = \det(u)\det(v)$$

Un endomorphisme $u \in \mathcal{L}(E)$ est inversible si, et seulement si, $\det(u) \neq 0$ et dans ce cas, on a $\det(u^{-1}) = \dfrac{1}{\det(u)}$.

Preuve. Pour toute base $\mathcal{B} = (e_i)_{1 \leq i \leq n}$ de E, on a :

$$\det(Id) = \det_{\mathcal{B}}(Id(e_i))_{1 \leq i \leq n} = \det_{\mathcal{B}}(\mathcal{B}) = 1$$

Avec la n-linéarité de $\det_{\mathcal{B}}$, on obtient :

$$\det(\lambda u) = \det_{\mathcal{B}}(\lambda u(e_i))_{1\leq i\leq n} = \lambda^n \det_{\mathcal{B}}(u(e_i))_{1\leq i\leq n} = \lambda^n \det(u)$$

Pour toute forme $\varphi \in \mathcal{A}_n(E,\mathbb{K}) \setminus \{0\}$ et tout $(x_i)_{1\leq i\leq n} \in E^n$, on a :

$$\varphi(u\circ v(x_1), \cdots, u\circ v(x_n)) = \det(v\circ u)\,\varphi(x_1,\cdots,x_n)$$

et aussi :

$$\begin{aligned}\varphi(u\circ v(x_1), \cdots, u\circ v(x_n)) &= \det(u)\,\varphi(v(x_1),\cdots,v(x_n))\\ &= \det(u)\det(v)\,\varphi(x_1,\cdots,x_n)\end{aligned}$$

donc $\det(u\circ v) = \det(u)\det(v)$. Comme $\det(u)\det(v) = \det(v)\det(u)$, on a aussi $\det(u\circ v) = \det(v\circ u)$.

Si $u \in GL(E)$, on a alors $1 = \det(Id) = \det\left(u\circ u^{-1}\right) = \det(u)\det\left(u^{-1}\right)$ et $\det(u) \neq 0$. Si $u \notin GL(E)$, on a alors $\det(u) = \det_{\mathcal{B}}(u(e_i))_{1<i<n} = 0$ puisque la famille $(u(e_i))_{1<i<n}$ est liée. □

L'application det est un morphisme de groupes surjectif de $GL(E)$ sur $\mathbb{K}^*$ et le noyau de ce morphisme est le groupe spécial linéaire $SL(E)$ qui est distingué de $GL(E)$ (le groupe linéaire $GL(E)$ est étudié en détail au chapitre **??**).

Si $A = ((a_{i,j}))_{1\leq i,j\leq n}$ est la matrice de $u \in \mathcal{L}(E)$ dans une base $\mathcal{B} = (e_i)_{1\leq i\leq n}$ de E, on a alors $u(e_j) = \displaystyle\sum_{i=1}^{n} a_{ij}e_i$ $(1 \leq i \leq n)$ et :

$$\det(u) = \det_{\mathcal{B}}(u(e_1),\cdots,u(e_n)) = \sum_{\sigma\in\mathcal{S}_n}\varepsilon(\sigma)\prod_{i=1}^{n} a_{\sigma(i),i}$$

Définition 2.4. *Si $A = ((a_{i,j}))_{1\leq i,j\leq n} \in \mathcal{M}_n(\mathbb{K})$, le déterminant de A est le scalaire :*

$$\det(A) = \sum_{\sigma\in\mathcal{S}_n}\varepsilon(\sigma)\prod_{i=1}^{n} a_{\sigma(i),i} \tag{2.1}$$

Ce déterminant est le déterminant de la famille $(C_j)_{1\leq j\leq n}$ des vecteurs colonnes de A dans la base canonique de $\mathbb{K}^n$. C'est aussi le déterminant de l'endomorphisme de $\mathbb{K}^n$ canoniquement associé à la matrice A.

Si $u \in \mathcal{L}(E)$ a pour matrice $A \in \mathcal{M}_n(\mathbb{K})$ dans une base $\mathcal{B}$ de E, on a alors $\det(u) = \det(A)$. Comme deux matrices semblables dans $\mathcal{M}_n(\mathbb{K})$ définissent le même endomorphisme de $\mathbb{K}^n$ dans deux bases différentes, elles ont le même déterminant.

En identifiant $\mathcal{M}_n(\mathbb{K})$ à $\mathbb{K}^{n^2}$, l'application $\det : \mathcal{M}_n(\mathbb{K}) \to \mathbb{K}$ est une fonction polynomiale homogène de degré n.

Théorème 2.9.

On a $\det(I_n) = 1$ *et pour* A, B *dans* $\mathcal{M}_n(\mathbb{K})$, $\lambda \in \mathbb{K}$, *on a :*

$$\det(\lambda A) = \lambda^n \det(A), \ \det\left({}^tA\right) = \det(A)$$

$$\det(AB) = \det(BA) = \det(A)\det(B)$$

Une matrice $A \in \mathcal{M}_n(\mathbb{K})$ *est inversible si, et seulement si,* $\det(A) \neq 0$ *et dans ce cas, on a* $\det\left(A^{-1}\right) = \dfrac{1}{\det(A)}$.

Preuve. Toutes les propriétés, exceptée celle sur la transposée se déduisent du théorème 2.8. Pour $A = ((a_{i,j}))_{1\leq i,j\leq n} \in \mathcal{M}_n(\mathbb{K})$, on a :

$$\begin{aligned}\det(A) &= \sum_{\sigma\in\mathcal{S}_n} \varepsilon(\sigma)\prod_{i=1}^{n} a_{\sigma(i),i} = \sum_{\sigma\in\mathcal{S}_n} \varepsilon(\sigma)\prod_{i=1}^{n} a_{\sigma(i),\sigma^{-1}(\sigma(i))} \\ &= \sum_{\sigma\in\mathcal{S}_n} \varepsilon\left(\sigma^{-1}\right)\prod_{k=1}^{n} a_{k,\sigma^{-1}(k)} = \sum_{\tau\in\mathcal{S}_n} \varepsilon(\tau)\prod_{i=1}^{n} a_{k,\tau(k)} = \det\left({}^tA\right)\end{aligned}$$

puisque l'application $\sigma \mapsto \sigma^{-1}$ est une permutation de $\mathcal{S}_n$ et $\varepsilon\left(\sigma^{-1}\right) = \varepsilon(\sigma)$ pour tout $\sigma \in \mathcal{S}_n$. □

Avec le théorème précédent, on retrouve le fait que si A et B sont deux matrices semblables dans $\mathcal{M}_n(\mathbb{K})$, on a alors $\det(A) = \det(B)$.

2.3 Méthodes de calcul du déterminant d'une matrice

Le calcul du déterminant d'une matrice A d'ordre n par la formule (2.1) nécessite un nombre d'opérations de l'ordre de $n!$ ce qui est beaucoup trop pour la capacité d'un ordinateur actuel sachant que $n! \geq \left(\dfrac{n}{e}\right)^n$. Par exemple, pour $n = 30$, cela donne plus de 10^{30} opérations.

Des propriétés des formes multilinéaires alternées, on déduit les propriétés suivantes du déterminant, où $A \in \mathcal{M}_n(\mathbb{K})$ et $C_j = (a_{ij})_{1\leq i\leq n} \in \mathbb{K}^n$ est la colonne numéro j de A pour tout j compris entre 1 et n.

— En désignant pour toute permutation σ de $\{1, \cdots, n\}$ par A_σ la matrice $A_\sigma = \left(C_{\sigma(1)}, \cdots, C_{\sigma(n)}\right)$ déduite de A en faisant agir σ sur ses colonnes, on a $\det(A_\sigma) = \varepsilon(\sigma)\det(A)$.

— Si l'une des colonnes de A est combinaison linéaire des autres, on a alors $\det(A) = 0$. En particulier, $\det(A) = 0$ si l'une des colonnes de A est nulle.

— $\det(A)$ est linéaire par rapport à chacune des colonnes.

— $\det(A)$ est inchangé si on ajoute à une de ses colonnes une combinaison linéaire des autres.

— Les propriétés analogues sur les lignes de A sont vérifiées.

Exemple 2.1 *De la première propriété, on déduit que si P_σ est une matrice de permutation (matrice déduite de I_n en faisant agir σ sur les colonnes de I_n), on a alors* $\det(P_\sigma) = \varepsilon(\sigma)$.

On s'intéresse d'abord au calcul du déterminant d'une matrice triangulaire par blocs.

Lemme 2.1 *Pour toute matrice* $A = \begin{pmatrix} a_{11} & \alpha \\ 0 & B \end{pmatrix} \in \mathcal{M}_n(\mathbb{K})$, *où* $B \in \mathcal{M}_{n-1}(\mathbb{K})$, $a_{11} \in \mathbb{K}$ *et* $\alpha \in \mathcal{M}_{1,n-1}(\mathbb{K})$, *on a* $\det(A) = a_{11}\det(B)$.

Preuve. Si $a_{11} = 0$, la première colonne de la matrice A est alors nulle et on a $\det(A) = 0 = a_{11}\det(B)$. Si $a_{11} \neq 0$, on a alors $\det(A) = a_{11}^n \det(A')$ où $A' = \begin{pmatrix} 1 & \alpha' \\ 0 & B' \end{pmatrix}$ avec $\alpha' = \dfrac{1}{a_{11}}\alpha$ et $B' = \dfrac{1}{a_{11}}B$. En notant C'_j la colonne numéro j de A', pour $1 \leq j \leq n$, les opérations qui consistent à remplacer C'_j par $C'_j - \alpha'_j C'_1$ pour j compris entre 2 et n, ne changent pas la valeur de $\det(A')$, ce qui nous donne $\det(A') = \det\begin{pmatrix} 1 & 0 \\ 0 & B' \end{pmatrix}$. L'application $\varphi : B' \in \mathcal{M}_{n-1}(\mathbb{K}) \mapsto \det\begin{pmatrix} 1 & 0 \\ 0 & B' \end{pmatrix}$ est $(n-1)$-linéaire alternée sur les colonnes (en identifiant l'espace vectoriel $\mathcal{M}_{n-1}(\mathbb{K})$ à $\left(\mathbb{K}^{n-1}\right)^{n-1}$) avec $\varphi(I_{n-1}) = \det(I_n) = 1$, donc $\det(A') = \varphi(B') = \det(B') = \dfrac{1}{a_{11}^{n-1}}\det(B)$ et $\det(A) = a_{11}\det(B)$. □

Lemme 2.2 *Pour toute matrice triangulaire supérieure ou inférieure* $A = ((a_{i,j}))_{1\leq i,j\leq n}$ *dans* $\mathcal{M}_n(\mathbb{K})$, *on a* $\det(A) = \prod_{i=1}^{n} a_{i,i}$.

Preuve. Se déduit par récurrence du lemme précédent et de $\det({}^tA) = \det(A)$. □

Lemme 2.3 *Pour toute matrice* $M = \begin{pmatrix} I_p & B \\ 0 & A \end{pmatrix} \in \mathcal{M}_n(\mathbb{K})$, *où* $p + q = n$, $A \in \mathcal{M}_q(\mathbb{K})$, $B \in \mathcal{M}_{p,q}(\mathbb{K})$, *on a* $\det(M) = \det(A)$.

Preuve. Pour $B \in \mathcal{M}_{p,q}(\mathbb{K})$ fixée, l'application $\varphi : A \in \mathcal{M}_q(\mathbb{K}) \mapsto \det\begin{pmatrix} I_p & B \\ 0 & A \end{pmatrix}$ est p-linéaire alternée sur les colonnes (en identifiant l'espace vectoriel $\mathcal{M}_q(\mathbb{K})$ à $(\mathbb{K}^q)^q$) avec $\varphi(I_q) = \det\begin{pmatrix} I_p & B \\ 0 & I_q \end{pmatrix} = 1$. On a donc, par définition du déterminant, $\varphi(A) = \det(A)$ pour toute matrice $A \in M_q(\mathbb{K})$. De manière analogue on montre que $\det\begin{pmatrix} A & B \\ 0 & I_q \end{pmatrix} = \det(A)$. □

Lemme 2.4 *Pour toute matrice* $M = \begin{pmatrix} A & B \\ 0 & D \end{pmatrix} \in \mathcal{M}_n(\mathbb{K})$, *où* $A \in \mathcal{M}_p(\mathbb{K})$, $B \in \mathcal{M}_{p,q}(\mathbb{K})$, $D \in \mathcal{M}_q(\mathbb{K})$, $p + q = n$, *on a* $\det(M) = \det(A)\det(D)$.

Preuve. Pour $(B, D) \in \mathcal{M}_{p,q}(\mathbb{K}) \times \mathcal{M}_q(\mathbb{K})$ fixé, l'application $\varphi : A \in \mathcal{M}_p(\mathbb{K}) \mapsto \det \begin{pmatrix} A & B \\ 0 & D \end{pmatrix}$ est p-linéaire alternée telle que $\varphi(I_p) = \det \begin{pmatrix} I_p & B \\ 0 & D \end{pmatrix} = \det(D)$, donc $\det(M) = \varphi(A) = \det(A)\det(D)$. □

Théorème 2.10.

Si $A = \begin{pmatrix} A_1 & B_{12} & \cdots & B_{1p} \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & B_{p-1,p-1} \\ 0 & \cdots & 0 & A_p \end{pmatrix}$ *est une matrice carrée où les* A_k *sont des matrices carrées, on a alors* $\det(A) = \prod_{k=1}^{p} \det(A_k)$.

Preuve. Se déduit de ce qui précède par récurrence sur p. □

Pour toute matrice $A = ((a_{i,j}))_{1 \leq i,j \leq n} \in \mathcal{M}_n(\mathbb{K})$ et tous i, j compris entre 1 et n, on note $A_{i,j}$ la matrice carrée d'ordre $n-1$ déduite de A en supprimant la ligne i et la colonne j. Le scalaire $\det(A_{i,j})$ est le mineur d'indice (i,j) et le scalaire $(-1)^{i+j}\det(A_{i,j})$ est le cofacteur d'indice (i,j).

Théorème 2.11.

Pour toute matrice $A \in \mathcal{M}_n(\mathbb{K})$, *on a les développements de* $\det(A)$ *suivant la colonne* $j \in \{1, \cdots, n\}$ *ou la ligne* $i \in \{1, \cdots, n\}$ *:*

$$\det(A) = \sum_{i=1}^{n} (-1)^{i+j} a_{i,j} \det(A_{i,j}) = \sum_{j=1}^{n} (-1)^{i+j} a_{i,j} \det(A_{i,j})$$

Preuve. Fixons la colonne j et notons $(e_i)_{1 \leq i \leq n}$ la base canonique de $\mathbb{K}^n$. La colonne C_j s'écrit $C_j = \sum_{i=1}^{n} a_{ij} e_i$ et en utilisant la linéarité du déterminant par rapport à la j-ième colonne, on a $\det(A) = \sum_{i=1}^{n} a_{ij} \det(B_{i,j})$, où B_{ij} est la matrice déduite de A en remplaçant C_j par e_i. En permutant la colonne j avec la colonne $j-1$, puis $j-1$ avec $j-2$, $\cdots$, 2 avec 1 et ensuite la ligne i avec la ligne $i-1$,

$i-1$ avec $i-2, \cdots, 2$ avec 1 (on ne fait rien pour $i=1$), on aboutit à :

$$\begin{aligned}
\det\left(B_{i,j}\right) &= (-1)^{i+j} \begin{vmatrix}
1 & a_{i1} & \cdots & a_{i,j-1} & a_{i,j+1} & \cdots & a_{in} \\
0 & a_{11} & \cdots & a_{1,j-1} & a_{1,j+1} & \cdots & a_{1n} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
0 & a_{i-1,1} & \cdots & a_{i-1,j-1} & a_{i-1,j+1} & \cdots & a_{i-1,n} \\
0 & a_{i+1,1} & \cdots & a_{i+1,j-1} & a_{i+1,j+1} & \cdots & a_{i+1,n} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
0 & a_{n,1} & \cdots & a_{n,j-1} & a_{n,j+1} & \cdots & a_{n,n}
\end{vmatrix} \\
&= (-1)^{i+j} \det\left(A_{i,j}\right)
\end{aligned}$$

et on a le résultat annoncé. On procède de manière analogue pour la deuxième formule. □

Définition 2.5. *La comatrice d'une matrice $A \in \mathcal{M}_n(\mathbb{K})$ est la matrice :*

$$C(A) = \left(\left((-1)^{i+j} \det\left(A_{i,j}\right)\right)\right)_{1 \leq i,j \leq n}$$

des cofacteurs de A.

Théorème 2.12.

Pour toute matrice $A \in \mathcal{M}_n(\mathbb{K})$, on a :

$$A \cdot {}^tC(A) = {}^tC(A) \cdot A = \det(A)\, I_n$$

Pour $\det(A) \neq 0$, *A est inversible d'inverse* $A^{-1} = \dfrac{1}{\det(A)}\, {}^tC(A)$.

Preuve. Pour $1 \leq i,j \leq n$, on note $c_{i,j} = (-1)^{i+j} \det\left(A_{i,j}\right)$ le cofacteur d'indice (i,j). Le terme d'indice (i,j) du produit $A \cdot {}^tC(A)$ est $d_{i,j} = \sum_{k=1}^{n} a_{i,k} c_{j,k}$. Pour $i=j$, on a $d_{i,i} = \sum_{k=1}^{n} (-1)^{i+k} a_{i,k} \det\left(A_{i,k}\right) = \det(A)$ et pour $i \neq j$, on a $d_{i,j} = \sum_{k=1}^{n} (-1)^{j+k} a_{i,k} \det\left(A_{j,k}\right) = 0$ puisque c'est le développement du déterminant de la matrice ayant pour ligne $k \neq j$, celle de A et pour ligne j, la ligne i de A, donc elle a deux lignes identiques. On a donc $A \cdot {}^tC(A) = \det(A)\, I_n$, l'autre égalité se vérifiant de manière analogue. □

En identifiant $\mathcal{M}_n(\mathbb{K})$ à $\mathbb{K}^{n^2}$, on déduit du théorème précédent que pour A inversible, les coefficients de A^{-1} sont des fonctions rationnelles des coefficients a_{ij}.

Exemple 2.2 *Pour* $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in GL_2(\mathbb{K})$, *on a :*

$$A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

Pour calculer un déterminant, on dispose aussi de la méthode des pivots de Gauss qui est basée sur les résultats suivants où une matrice de transvection est une matrice de la forme $I_n + \lambda E_{ij}$ en notant $(E_{i,j})_{1 \leq i,j \leq n}$ la base canonique de $\mathcal{M}_n(\mathbb{K})$ (voir le paragraphe **??**).

Lemme 2.5 *Soit* $A \in \mathcal{M}_n(\mathbb{K})$ *de coefficient* a_{11} *non nul. Il existe des matrices de transvection* $P_1, \cdots, P_r$ *telles que :*

$$P_r \cdots P_1 A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22}^{(1)} & \cdots & a_{2n}^{(1)} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n2}^{(1)} & \cdots & a_{nn}^{(1)} \end{pmatrix}$$

Preuve. Voir [**?**]. □

Si le coefficient a_{11} est nul et qu'il existe un indice i compris entre 2 et n tel que $a_{i1} \neq 0$, on se ramène à l'hypothèse du lemme en permuttant les lignes 1 et i. Si tous les a_{i1} pour $1 \leq i \leq n$ sont nuls, le déterminant de A est alors nul.

Théorème 2.13.

Toute matrice $A \in GL_n(\mathbb{K})$ *peut être réduite à la forme triangulaire supérieure en la multipliant à gauche par des matrices de transvection ou de permutation.*

Preuve. Voir [**?**]. □

Le théorème précédent ramène au signe près le calcul du déterminant de A au déterminant d'une matrice triangulaire supérieure, le « signe près » étant $(-1)^p$, où p est le nombre de permutations effectuées dans l'algorithme de Gauss.

2.4 Quelques déterminants classiques

2.4.1 Déterminants de Vandermonde

À tout entier $n \geq 2$ et toute suite $(\alpha_k)_{1 \leq k \leq n}$ d'éléments de $\mathbb{K}$, on associe la matrice de Vandermonde :

$$V(\alpha_1, \cdots, \alpha_n) = \left(\left(\alpha_j^{i-1}\right)\right)_{1 \leq i,j \leq n} = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ \alpha_1 & \alpha_2 & \cdots & \alpha_n \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_1^{n-1} & \alpha_2^{n-1} & \cdots & \alpha_n^{n-1} \end{pmatrix}$$

et on note $\Delta(\alpha_1, \cdots, \alpha_n)$, le déterminant de cette matrice.

Théorème 2.14.

La matrice de Vandermonde $V(\alpha_1, \cdots, \alpha_n)$ est inversible si, et seulement si, les α_k pour k compris entre 1 et n sont deux à deux distincts.

Preuve. S'il existe $i \neq j$ tels que $\alpha_i = \alpha_j$, la matrice $V(\alpha_1, \cdots, \alpha_n)$ a alors deux colonnes identiques, ce qui implique qu'elle est non inversible.

Réciproquement, si $V(\alpha_1, \cdots, \alpha_n)$ est non inversible, il en est alors de même de ${}^tV(\alpha_1, \cdots, \alpha_n) = \begin{pmatrix} 1 & \alpha_1 & \cdots & \alpha_1^{n-1} \\ 1 & \alpha_2 & \cdots & \alpha_2^{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \alpha_n & \cdots & \alpha_n^{n-1} \end{pmatrix}$ et le système linéaire :

$$\sum_{j=1}^{n} \alpha_i^{j-1} x_j = 0 \ \ (1 \leq i \leq n)$$

a une solution $(x_j)_{1 \leq j \leq n}$ non nulle, ce qui revient à dire que $\alpha_1, \cdots, \alpha_n$ sont des racines du polynôme $P(X) = \sum_{j=1}^{n} x_j X^{j-1} \in \mathbb{K}_{n-1}[X] \setminus \{0\}$ et nécessairement il existe $i \neq j$ tels que $\alpha_i = \alpha_j$, sans quoi P aurait n racines distinctes en étant non nul de degré au plus égal à $n-1$, ce qui n'est pas possible. □

Le théorème précédent peut être utilisé pour prouver l'existence et l'unicité des polynômes d'interpolation de Lagrange.

Théorème 2.15.

Soit $(\alpha_k)_{1 \leq i \leq n}$ une famille de $n \geq 2$ scalaires deux à deux distincts. Pour toute famille de scalaires $(\beta_k)_{1 \leq i \leq n}$, il existe un unique polynôme $P \in \mathbb{K}_{n-1}[x]$ tel que $P(\alpha_k) = \beta_k$ pour tout k compris entre 0 et n.

Preuve. L'application $\varphi_n : \mathbb{K}_{n-1}[X] \to \mathbb{K}^n$ définie par $\varphi_n(P) = (P(\alpha_k))_{1 \leq k \leq n}$ pour tout $P \in \mathbb{K}_{n-1}[X]$ est linéaire et sa matrice dans les bases canoniques de $\mathbb{K}_{n-1}[X]$ et $\mathbb{K}^n$ est la matrice inversible :

$$\begin{pmatrix} 1 & \alpha_1 & \cdots & \alpha_1^{n-1} \\ 1 & \alpha_2 & \cdots & \alpha_2^{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \alpha_n & \cdots & \alpha_n^{n-1} \end{pmatrix} = {}^tV(\alpha_1, \cdots, \alpha_n)$$

c'est donc un isomorphisme de $\mathbb{K}_{n-1}[X]$ sur $\mathbb{K}^n$ et tout vecteur $(\beta_k)_{1 \leq i \leq n} \in \mathbb{K}^n$ a un unique antécédent $P \in \mathbb{K}_{n-1}[X]$. □

Théorème 2.16.

On a $\Delta(\alpha_1, \cdots, \alpha_n) = \prod_{1 \leq i < j \leq n} (\alpha_j - \alpha_i)$.

Preuve. Pour $n = 2$, on a $\Delta(\alpha_1, \alpha_2) = \begin{vmatrix} 1 & 1 \\ \alpha_1 & \alpha_2 \end{vmatrix} = \alpha_2 - \alpha_1$. En supposant le résultat acquis pour $n - 1 \geq 2$, on propose trois méthodes de démonstration par récurrence.

1. En retranchant, pour $i = n, n-1, \cdots, 2$, à la ligne i de $V(\alpha_1, \cdots, \alpha_n)$ sa ligne $i - 1$ multipliée par α_1, on obtient :

$$\begin{aligned}\Delta(\alpha_1, \cdots, \alpha_n) &= \begin{vmatrix} 1 & 1 & \cdots & 1 \\ 0 & \alpha_2 - \alpha_1 & \cdots & \alpha_n - \alpha_1 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \alpha_2^{n-2}(\alpha_2 - \alpha_1) & \cdots & \alpha_n^{n-2}(\alpha_n - \alpha_1) \end{vmatrix} \\ &= \begin{vmatrix} \alpha_2 - \alpha_1 & \alpha_3 - \alpha_1 & \cdots & \alpha_n - \alpha_1 \\ \alpha_2(\alpha_2 - \alpha_1) & \alpha_3(\alpha_3 - \alpha_1) & \cdots & \alpha_n(\alpha_n - \alpha_1) \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_2^{n-2}(\alpha_2 - \alpha_1) & \alpha_3^{n-2}(\alpha_3 - \alpha_1) & \cdots & \alpha_n^{n-2}(\alpha_n - \alpha_1) \end{vmatrix}\end{aligned}$$

soit par n-linéarité du déterminant :

$$\begin{aligned}\Delta(\alpha_1, \cdots, \alpha_n) &= \left(\prod_{k=2}^{n}(\alpha_k - \alpha_1)\right) \begin{vmatrix} 1 & \cdots & 1 \\ \alpha_2 & \cdots & \alpha_n \\ \vdots & \ddots & \vdots \\ \alpha_2^{n-2} & \cdots & \alpha_n^{n-2} \end{vmatrix} \\ &= \left(\prod_{k=2}^{n}(\alpha_k - \alpha_1)\right) \Delta(\alpha_2, \cdots, \alpha_n)\end{aligned}$$

et avec l'hypothèse de récurrence, cela donne :

$$\begin{aligned}\Delta(\alpha_1, \cdots, \alpha_n) &= \left(\prod_{k=2}^{n}(\alpha_k - \alpha_1)\right) \prod_{2 \leq i < j \leq n}^{n}(\alpha_j - \alpha_i) \\ &= \prod_{1 \leq i < j \leq n}^{n}(\alpha_j - \alpha_i)\end{aligned}$$

2. Le déterminant étant une forme n-linéaire alternée, on a pour tout polynôme P unitaire de degré $n - 1$:

$$\begin{vmatrix} 1 & 1 & \cdots & 1 \\ \alpha_1 & \alpha_2 & \cdots & \alpha_n \\ \vdots & \vdots & \ddots & \vdots \\ P(\alpha_1) & P(\alpha_2) & \cdots & P(\alpha_n) \end{vmatrix} = \Delta(\alpha_1, \cdots, \alpha_n)$$

Prenant $P(X) = \prod_{k=1}^{n-1} (X - \alpha_k)$, on en déduit que :

$$\begin{aligned}
\Delta(\alpha_1, \cdots, \alpha_n) &= \begin{vmatrix} 1 & 1 & \cdots & 1 \\ \alpha_1 & \alpha_2 & \cdots & \alpha_n \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & P(\alpha_n) \end{vmatrix} = P(\alpha_n)\,\Delta(\alpha_1, \cdots, \alpha_{n-1}) \\
&= \left(\prod_{k=1}^{n-1} (\alpha_n - \alpha_k) \right) \Delta(\alpha_1, \cdots, \alpha_{n-1})
\end{aligned}$$

et avec l'hypothèse de récurrence, cela donne :

$$\begin{aligned}
\Delta(\alpha_1, \cdots, \alpha_n) &= \left(\prod_{k=1}^{n-1} (\alpha_n - \alpha_k) \right) \prod_{1 \leq i < j \leq n-1}^{n} (\alpha_j - \alpha_i) \\
&= \prod_{1 \leq i < j \leq n}^{n} (\alpha_j - \alpha_i)
\end{aligned}$$

3. Dans le cas où deux des α_k sont égaux, la formule est triviale. Dans le cas contraire, le polynôme $V(X) = V(\alpha_1, \cdots, \alpha_{n-1}, X)$ qui est de degré $n-1$ (développement par rapport à la dernière colonne) s'annule en $\alpha_1, \cdots, \alpha_{n-1}$ deux à deux distincts, donc il existe $\lambda \in \mathbb{K}$ telle que $V(X) = \lambda \prod_{k=1}^{n-1} (X - \alpha_k)$.
   En développant le déterminant $V(X)$ par rapport à la dernière colonne, on voit que λ qui est le coefficient de X^{n-1} est égal à $\Delta(\alpha_1, \cdots, \alpha_{n-1})$, ce qui nous donne par évaluation en α_n, $\Delta(\alpha_1, \cdots, \alpha_n) = \Delta(\alpha_1, \cdots, \alpha_{n-1}) \prod_{k=1}^{n-1} (\alpha_n - \alpha_k)$ et on conclut encore par récurrence.

□

Avec le théorème précédent, on retrouve le fait que $V(\alpha_1, \cdots, \alpha_n)$ est inversible si, et seulement si, les α_k pour k compris entre 1 et n sont deux à deux distincts.

Exemple 2.3 *Pour $n \geq 2$, on a :*

$$\begin{vmatrix} 1 & 1 & \cdots & 1 \\ 1 & 2 & \cdots & n \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 2^{n-1} & \cdots & n^{n-1} \end{vmatrix} = \prod_{1 \leq 1 < j \leq n} (j - i) = \prod_{j=2}^{n} \prod_{i=1}^{j-1} (j - i) = \prod_{j=2}^{n} (j-1)! = \prod_{j=1}^{n-1} j!$$

Dans le cas où les α_k sont deux à deux distincts, on peut donner une expression de l'inverse d'une matrice de Vandermonde, en utilisant les polynômes d'interpolation de Lagrange. Pour ce faire, on introduit l'application linéaire $\varphi_n : \mathbb{K}_{n-1}[X] \to \mathbb{K}^n$ définie par $\varphi_n(P) = (P(\alpha_i))_{1 \leq i \leq n}$ pour tout $P \in \mathbb{K}_{n-1}[X]$. La matrice de φ_n

dans les bases canoniques $\mathcal{B} = (e_i)_{1\leq i\leq n}$ de $\mathbb{K}^n$ et $\mathcal{B}' = \left(X^{j-1}\right)_{1\leq j\leq n}$ de $\mathbb{K}_{n-1}[X]$ et $\mathbb{K}^n$ est donnée par :

$$\varphi_n\left(X^{j-1}\right) = \begin{pmatrix} \alpha_1^{j-1} \\ \alpha_2^{j-1} \\ \vdots \\ \alpha_n^{j-1} \end{pmatrix} \ (1 \leq j \leq n)$$

ce qui nous dit que c'est la matrice transposée de Vandermonde ${}^tV(\alpha_1, \cdots, \alpha_n)$. En notant $\mathcal{L} = (L_i)_{1\leq i\leq n}$ la base de Lagrange définie par $L_i(X) = \prod_{\substack{j=1\\ j\neq i}}^{n} \frac{X-\alpha_j}{\alpha_i-\alpha_j}$, on a $L_i(\alpha_j) = \delta_{i,j}$ pour tous i, j compris entre 1 et n, soit $\varphi_n(L_i) = e_i$ et $\varphi_n^{-1}(e_i) = L_i$ pour tout i compris entre entre 1 et n, ce qui nous dit que ${}^tV(\alpha_1, \cdots, \alpha_n)^{-1} = P_{\mathcal{B}',\mathcal{L}}$ est la matrice de passage de la base canonique $\mathcal{B}' = \left(X^{j-1}\right)_{1\leq j\leq n}$ à la base de Lagrange $\mathcal{L} = (L_i)_{1\leq i\leq n}$ et $V(\alpha_1, \cdots, \alpha_n)^{-1} = {}^tP_{\mathcal{B}',\mathcal{L}}$. Par exemple pour $n = 2$, on a $V(\alpha_1, \alpha_2)^{-1} = \frac{1}{\alpha_2-\alpha_1}\begin{pmatrix} \alpha_2 & -1 \\ -\alpha_1 & 1 \end{pmatrix}$ et pour $n = 3$:

$$\begin{aligned} &L_1(X) = \frac{(X-\alpha_2)(X-\alpha_3)}{(\alpha_1-\alpha_2)(\alpha_1-\alpha_3)},\ L_2(X) = \frac{(X-\alpha_1)(X-\alpha_3)}{(\alpha_2-\alpha_1)(\alpha_2-\alpha_3)} \\ &L_3(X) = \frac{(X-\alpha_1)(X-\alpha_2)}{(\alpha_3-\alpha_1)(\alpha_3-\alpha_2)} \end{aligned}$$

$$\begin{aligned} V(\alpha_1, \alpha_2, \alpha_3)^{-1} &= \begin{pmatrix} \frac{\alpha_2\alpha_3}{(\alpha_1-\alpha_2)(\alpha_1-\alpha_3)} & -\frac{\alpha_2+\alpha_3}{(\alpha_1-\alpha_2)(\alpha_1-\alpha_3)} & \frac{1}{(\alpha_1-\alpha_2)(\alpha_1-\alpha_3)} \\ \frac{\alpha_1\alpha_3}{(\alpha_2-\alpha_1)(\alpha_2-\alpha_3)} & -\frac{\alpha_1+\alpha_3}{(\alpha_2-\alpha_1)(\alpha_2-\alpha_3)} & \frac{1}{(\alpha_2-\alpha_1)(\alpha_2-\alpha_3)} \\ \frac{\alpha_1\alpha_2}{(\alpha_3-\alpha_1)(\alpha_3-\alpha_2)} & -\frac{\alpha_1+\alpha_2}{(\alpha_3-\alpha_1)(\alpha_3-\alpha_2)} & \frac{1}{(\alpha_3-\alpha_1)(\alpha_3-\alpha_2)} \end{pmatrix} \\ &= \frac{1}{\Delta(\alpha_1, \alpha_2, \alpha_3)}\begin{pmatrix} \alpha_2\alpha_3(\alpha_3-\alpha_2) & -\left(\alpha_3^2-\alpha_2^2\right) & (\alpha_3-\alpha_2) \\ -\alpha_1\alpha_3(\alpha_3-\alpha_1) & \alpha_3^2-\alpha_1^2 & -(\alpha_3-\alpha_1) \\ \alpha_1\alpha_2(\alpha_2-\alpha_1) & -\left(\alpha_2^2-\alpha_1^2\right) & (\alpha_2-\alpha_1) \end{pmatrix} \end{aligned}$$

Les matrices de Vandermonde peuvent être généralisées en considérant les matrices de la forme :

$$V(\alpha_1, \cdots, \alpha_n, P_1, \cdots, P_n) = ((P_i(\alpha_j)))_{1\leq i,j\leq n} = \begin{pmatrix} P_1(\alpha_1) & \cdots & P_1(\alpha_n) \\ \vdots & \ddots & \vdots \\ P_n(\alpha_1) & \cdots & P_n(\alpha_n) \end{pmatrix}$$

où $(P_j)_{1\leq j\leq n}$ est une suite de polynômes dans $\mathbb{K}_{n-1}[X]$. Pour simplifier, on notera V_n une telle matrice. Cette matrice est une matrice de Vandermonde pour $P_j(X) = X^{j-1}$.

Théorème 2.17.

En notant $Q(P_1, \cdots, P_n)$ *la matrice dont les colonnes sont formées des composantes de chaque polynôme* P_k *dans la base canonique de* $\mathbb{K}_{n-1}[X]$, *on a :*

$$\det \begin{pmatrix} P_1(\alpha_1) & \cdots & P_1(\alpha_n) \\ \vdots & \ddots & \vdots \\ P_n(\alpha_1) & \cdots & P_n(\alpha_n) \end{pmatrix} = \Delta(\alpha_1, \cdots, \alpha_n) \det(Q(P_1, \cdots, P_n))$$

Dans le cas particulier où chaque polynôme P_k, *pour* k *compris entre* 1 *et* $n-1$ *est unitaire de degré* $k-1$, *on a* $\det(A_n) = \Delta(\alpha_1, \cdots, \alpha_n)$.

Preuve. En notant $P_j(X) = \sum_{i=0}^{n-1} a_{ij} X^j$ pour tout entier j compris entre 1 et n, on a :

$$\begin{aligned} V_n &= \begin{pmatrix} a_{0,1} & \cdots & a_{n-1,1} \\ \vdots & \ddots & \vdots \\ a_{n,0} & \cdots & a_{n-1,n} \end{pmatrix} \begin{pmatrix} 1 & 1 & \cdots & 1 \\ \alpha_1 & \alpha_2 & \cdots & \alpha_n \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_1^{n-1} & \alpha_2^{n-1} & \cdots & \alpha_n^{n-1} \end{pmatrix} \\ &= {}^tQ(P_1, \cdots, P_n) V(\alpha_1, \cdots, \alpha_n) \end{aligned}$$

où $Q(P_1, \cdots, P_n)$ est la matrice dont les colonnes sont formées des composantes de chaque polynôme P_j dans la base canonique de $\mathbb{K}_{n-1}[X]$ (dans le où la famille $(P_j)_{1 \leq j \leq n}$ est libre, cette matrice est la matrice de passage de $\left(X^{j-1}\right)_{1 \leq j \leq n}$ à $(P_j)_{1 \leq j \leq n}$), donc $\det(V_n) = \Delta(\alpha_1, \cdots, \alpha_n) \det(Q(P_1, \cdots, P_n))$.

Dans le cas particulier où chaque polynôme P_k, pour k compris entre 1 et $n-1$ est unitaire de degré $k-1$, la matrice $Q(P_1, \cdots, P_n)$ est triangulaire supérieure avec tous ses termes diagonaux égaux à 1, donc $\det(Q(P_1, \cdots, P_n)) = 1$ et $\det(A_n) = \Delta(\alpha_1, \cdots, \alpha_n)$. □

Dans le cas où les α_k sont deux à deux distincts, en prenant pour polynômes P_j les polynômes de Lagrange $L_j(X) = \prod_{\substack{i=1 \\ i \neq j}}^{n} \frac{X - \alpha_i}{\alpha_j - \alpha_i}$, on a :

$$V_n = I_n = {}^tQ(L_1, \cdots, L_n) V(\alpha_1, \cdots, \alpha_n)$$

ce qui nous permet de retrouver l'inverse $V(\alpha_1, \cdots, \alpha_n)^{-1} = {}^tQ(L_1, \cdots, L_n)$.

Exemple 2.4 *Prenant* $\alpha_k = k$ *et* $P_k(X) = (X + k - 1)^{n-1}$, *on obtient :*

$$\det \begin{pmatrix} 1^{n-1} & 2^{n-1} & \cdots & n^{n-1} \\ 2^{n-1} & 3^{n-1} & \cdots & (n+1)^{n-1} \\ \vdots & \vdots & \cdots & \vdots \\ n^{n-1} & (n+1)^{n-1} & \cdots & (2n-1)^{n-1} \end{pmatrix} = (-1)^{\frac{n(n-1)}{2}} ((n-1)!)^{n-1}$$

(exercice 2.7).

Les déterminants de Vandermonde peuvent aussi être utilisés pour prouver que les vecteurs propres non nuls d'une matrice $A \in \mathcal{M}_n(\mathbb{K})$ associés à des valeurs propres deux à deux distinctes sont linéairement indépendants (exercice 2.4) ou encore pour prouver le théorème d'analyse réelle qui suit.

Théorème 2.18.

Si f est une fonction de classe $\mathcal{C}^{n+1}$ de $\mathbb{R}$ dans $\mathbb{R}$, avec $n \geq 1$, telle que f et $f^{(n+1)}$ soient bornées sur $\mathbb{R}$, alors toutes les dérivées $f^{(k)}$, pour k compris entre 1 et n, sont également bornées sur $\mathbb{R}$.

Preuve. Pour tout réel x et tout entier p compris entre 1 et n, la formule de Taylor à l'ordre n sur l'intervalle $[x, x+p]$ s'écrit :

$$f(x+p) - f(x) - \frac{f^{(n+1)}(x+p\theta_p)}{(n+1)!}p^{n+1} = \sum_{k=1}^{n} \frac{f^{(k)}(x)}{k!}p^k$$

avec $0 < \theta_p < 1$, ce qui peut s'écrire matriciellement $V(x) = AU(x)$, où on a noté :

$$V(x) = \begin{pmatrix} f(x+1) - f(x) - \dfrac{f^{(n+1)}(x+\theta_1)}{(n+1)!} \\ \vdots \\ f(x+n) - f(x) - \dfrac{f^{(n+1)}(x+n\theta_n)}{(n+1)!}n^{n+1} \end{pmatrix}, \ U(x) = \begin{pmatrix} \dfrac{f^{(1)}(x)}{1!} \\ \vdots \\ \dfrac{f^{(n)}(x)}{n!} \end{pmatrix}$$

et $A = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 2 & 2^2 & \cdots & 2^n \\ \vdots & \vdots & \ddots & \vdots \\ n & n^2 & \cdots & n^n \end{pmatrix}$. On a alors :

$$\det(A) = n! \begin{vmatrix} 1 & 1 & \cdots & 1 \\ 1 & 2 & \cdots & 2^{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & n & \cdots & n^{n-1} \end{vmatrix} = n! \prod_{j=1}^{n-1} j! \neq 0$$

(exemple 2.3), c'est-à-dire que A est inversible. On peut donc écrire que $U(x) = A^{-1}V(x)$ et $\|U(x)\|_\infty \leq \left\|A^{-1}\right\|_\infty \|V(x)\|_\infty$, soit :

$$\begin{aligned} \max_{1\leq k\leq n} \left| \frac{f^{(k)}(x)}{k!} \right| &\leq \left\|A^{-1}\right\|_\infty \max_{1\leq k\leq n} \left| f(x+p) - f(x) - \frac{f^{(n+1)}(x+p\theta_p)}{(n+1)!}p^{n+1} \right| \\ &\leq \left\|A^{-1}\right\|_\infty \left(2\|f\|_\infty + \frac{n^{n+1}}{(n+1)!} \left\|f^{(n+1)}\right\|_\infty \right) \end{aligned}$$

le réel x étant quelconque. On a donc pour tout entier k compris entre 1 et n :

$$\left\|f^{(k)}\right\|_\infty \leq k! \left\|A^{-1}\right\|_\infty \left(2\|f\|_\infty + \frac{n^{n+1}}{(n+1)!} \left\|f^{(n+1)}\right\|_\infty \right)$$

$\square$

2.4.2 Déterminants circulants

On se donne un entier $n \geq 2$, des nombres complexes $a_0, \cdots, a_{n-1}$ et on leur associe la matrice circulante :

$$C(a_0, \cdots, a_{n-1}) = \begin{pmatrix} a_0 & a_1 & \cdots & a_{n-2} & a_{n-1} \\ a_{n-1} & a_0 & \ddots & a_{n-3} & a_{n-2} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ a_2 & a_3 & \ddots & a_0 & a_1 \\ a_1 & a_2 & \cdots & a_{n-1} & a_0 \end{pmatrix} = \left(\left(a_{j-i \bmod n}\right)\right)_{0 \leq i,j \leq n-1}$$

où $j - i \bmod n$ désigne le reste appartenant à $\{0, 1, \cdots, n-1\}$ dans la division euclidienne de $j - i$ par n, ce qui signifie que $a_{ij} = a_{j-i}$ pour $0 \leq i \leq j \leq n-1$ et $a_{ij} = a_{j+n-i}$ pour $j + 1 \leq i \leq n - 1$. Pour toute racine n-ième de l'unité $z \in \mathbb{C}$, on a $z^n = 1$ et en notant $Z = \left(z^j\right)_{0 \leq j \leq n-1}$ dans $\mathbb{C}^n$:

$$C(a_0, \cdots, a_{n-1}) Z = \begin{pmatrix} a_0 & a_1 & \cdots & a_{n-2} & a_{n-1} \\ a_{n-1} & a_0 & \ddots & a_{n-3} & a_{n-2} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ a_2 & a_3 & \ddots & a_0 & a_1 \\ a_1 & a_2 & \cdots & a_{n-1} & a_0 \end{pmatrix} \begin{pmatrix} 1 \\ z \\ \vdots \\ z^{n-2} \\ z^{n-1} \end{pmatrix}$$

$$= \begin{pmatrix} a_0 + a_1 z + \cdots + a_{n-2} z^{n-2} + a_{n-1} z^{n-1} \\ z\left(a_0 + a_1 z + \cdots + a_{n-2} z^{n-2} + a_{n-1} z^{n-1}\right) \\ \vdots \\ z^{n-2}\left(a_0 + a_1 z + \cdots + a_{n-2} z^{n-2} + a_{n-1} z^{n-1}\right) \\ z^{n-1}\left(a_0 + a_1 z + \cdots + a_{n-2} z^{n-2} + a_{n-1} z^{n-1}\right) \end{pmatrix} = P(z) \begin{pmatrix} 1 \\ z \\ \vdots \\ z^{n-2} \\ z^{n-1} \end{pmatrix}$$

où $P(X) = \sum\limits_{k=0}^{n-1} a_k X^k$, ce qui nous dit que $P(z)$ est vecteur propre de la matrice $C(a_0, \cdots, a_{n-1})$, le vecteur Z étant un vecteur propre associé. Ce résultat peut aussi se justifier en disant que, pour $0 \leq i \leq n-1$, la composante numéro de i de $C(a_0, \cdots, a_{n-1}) Z$ est :

$$\sum_{j=0}^{n-1} a_{ij} z^j = \sum_{j=0}^{i-1} a_{j+n-i} z^j + \sum_{j=i}^{n-1} a_{j-i} z^j = z^i \left(\sum_{j=0}^{i-1} a_{j+n-i} z^{n+j-i} + \sum_{j=i}^{n-1} a_{j-i} z^{j-i} \right)$$

$$= z^i \left(\sum_{k=n-i}^{n-1} a_k z^k + \sum_{k=0}^{n-i-1} a_k z^k \right) = z^i \sum_{k=0}^{n-1} a_k z^k = z^i P(z)$$

En notant $\omega_n = e^{\frac{2i\pi}{n}}$, chaque racine n-ième de l'unité ω_n^k pour $0 \leq k \leq n-1$, nous donne la valeur propre $P\left(\omega_n^k\right)$ de $C(a_0, \cdots, a_{n-1})$ avec $Z_k = \left(\omega_n^{kj}\right)_{0 \leq j \leq n-1}$ comme vecteur propre associé.

Comme $\det(Z_0, \cdots, Z_{n-1}) = \Delta\left(1, \omega_n, \omega_n^2, \cdots, \omega_n^{n-1}\right) \neq 0$ (ces racines de l'unité sont deux à deux distinctes), la famille de vecteur propres $(Z_k)_{0 \leq k \leq n-1}$ est une base de $\mathbb{C}^n$, ce qui nous dit que $C(a_0, \cdots, a_{n-1})$ est diagonalisable de valeurs propres $P(1), P(\omega_n), \cdots, P\left(\omega_n^{n-1}\right)$. Il en résulte que $\det C(a_0, \cdots, a_{n-1}) = \prod_{k=0}^{n-1} P\left(\omega_n^k\right)$.

2.4.3 Déterminants de Gram

$(E, \langle \cdot \mid \cdot \rangle)$ désigne un espace vectoriel réel préhilbertien et $(x_i)_{1 \leq i \leq n}$ est une famille de $n \geq 2$ vecteurs de E.

Définition 2.6. *On appelle matrice de Gram de $(x_i)_{1 \leq i \leq n}$, la matrice :*

$$\begin{aligned} G(x_1, \cdots, x_n) &= ((\langle x_i \mid x_j \rangle))_{1 \leq i,j \leq n} \\ &= \begin{pmatrix} \langle x_1 \mid x_1 \rangle & \langle x_1 \mid x_2 \rangle & \cdots & \langle x_1 \mid x_n \rangle \\ \langle x_2 \mid x_1 \rangle & \langle x_2 \mid x_2 \rangle & \cdots & \langle x_2 \mid x_n \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle x_n \mid x_1 \rangle & \langle x_n \mid x_2 \rangle & \cdots & \langle x_n \mid x_n \rangle \end{pmatrix} \end{aligned}$$

et le déterminant $g(x_1, \cdots, x_n)$ de cette matrice est appelé déterminant de Gram de la famille $(x_i)_{1 \leq i \leq n}$.

Théorème 2.19.

On a, $\operatorname{rg}(G(x_1, \cdots, x_n)) = \operatorname{rg}(x_1, \cdots, x_n)$.

Preuve. Si tous les vecteurs x_i sont nuls, la matrice de Gram correspondante est alors la matrice nulle et le résultat est évident. On suppose donc que les x_i ne sont pas tous nuls et on désigne par F le sous-espace vectoriel de E engendré par $\{x_1, \cdots, x_n\}$. Le procédé de Gram-Schmidt nous permet de construire une base orthonormée $(e_i)_{1 \leq i \leq p}$ de F, avec $1 \leq p \leq n$, et dans cette base on écrit $x_i = \sum_{k=1}^{p} a_{ki} e_k$. Pour i, j compris entre 1 et n, on a alors :

$$\langle x_i \mid x_j \rangle = \sum_{k=1}^{p} a_{ki} a_{kj} = (a_{1i}, \cdots, a_{pi}) \begin{pmatrix} a_{1j} \\ \vdots \\ a_{pj} \end{pmatrix}$$

soit, $G(x_1, \cdots, x_n) = {}^t AA$ où $A = ((a_{ij}))_{\substack{1 \leq i \leq p \\ 1 \leq j \leq n}} \in \mathcal{M}_{p,n}(\mathbb{R})$. Munissant $\mathbb{R}^n$ et $\mathbb{R}^p$ de leurs structures euclidiennes canoniques, on a $\langle {}^t AAX \mid X \rangle_{\mathbb{R}^n} = \langle AX \mid AX \rangle_{\mathbb{R}^p} = \|AX\|_{\mathbb{R}^p}^2$ pour tout $X \in \mathbb{R}^n$ et on en déduit que $\ker({}^t AA) = \ker(A)$, puis avec le théorème du rang que les matrices ${}^t AA$ et A ont même rang. On a donc $\operatorname{rg}(G(x_1, \cdots, x_n)) = \operatorname{rg}(A) = \operatorname{rg}(x_1, \cdots, x_n)$ puisque A est la matrice du système de vecteurs $(x_i)_{1 \leq i \leq n}$ dans la base $(e_i)_{1 \leq i \leq p}$. □

Corollaire 2.2. *On a $g(x_1, \cdots, x_n) \geq 0$ et la famille $(x_i)_{1\leq i\leq n}$ est libre si, et seulement si, on a $g(x_1, \cdots, x_n) > 0$.*

Preuve. On reprend les notations de la démonstration du théorème 2.19. Si le système $(x_i)_{1\leq i\leq n}$ est lié, on a alors $\mathrm{rg}(G(x_1, \cdots, x_n)) = \mathrm{rg}(A) < n$ et la matrice $G(x_1, \cdots, x_n)$ est non inversible dans $\mathcal{M}_n(\mathbb{R})$, son déterminant est donc nul. Si le système $(x_i)_{1\leq i\leq n}$ est libre, on a alors $\mathrm{rg}(G(x_1, \cdots, x_n)) = \mathrm{rg}(A) = n$ et la matrice $G(x_1, \cdots, x_n)$ est inversible dans $\mathcal{M}_n(\mathbb{R})$, son déterminant est donc non nul. En considérant que la matrice $G(x_1, \cdots, x_n) = {}^tAA$ est symétrique positive ($\langle {}^tAAX \mid X\rangle_{\mathbb{R}^n} = \|AX\|^2_{\mathbb{R}^n} \geq 0$ pour tout $X \in \mathbb{R}^n$), on déduit que ce déterminant est strictement positif. □

Les déterminants de Gram peuvent être utilisés pour calculer la distance d'un point à un sous-espace vectoriel de dimension finie de E.

Théorème 2.20.

Soient F un sous-espace vectoriel de dimension finie de E et $(x_i)_{1\leq i\leq n}$ une base de F. Pour tout x dans E, on a $d(x, F) = \sqrt{\dfrac{g(x_1, \cdots, x_n, x)}{g(x_1, \cdots, x_n)}}$ et la meilleure approximation de x par des éléments de F est le vecteur $y = \sum\limits_{j=1}^{n} \dfrac{g_{j,x}(x_1, \cdots, x_n)}{g(x_1, \cdots, x_n)} x_j$, où $g_{j,x}(x_1, \cdots, x_n)$ est le déterminant de la matrice déduite de la matrice de Gram $G(x_1, \cdots, x_n)$ en remplaçant sa colonne numéro j par $\begin{pmatrix} \langle x_1 \mid x\rangle \\ \vdots \\ \langle x_n \mid x\rangle \end{pmatrix}$.

Preuve. En notant y la projection orthogonale de x sur F, on a :

$$d(x, F)^2 = \|x - y\|^2 = \|x\|^2 - \|y\|^2$$

et pour tout i compris entre 1 et n, $\langle x_i \mid x\rangle = \langle x_i \mid x - y\rangle + \langle x_i \mid y\rangle = \langle x_i \mid y\rangle$ du fait que $x - y \in F^\perp$ et $x_i \in F$. En utilisant la linéarité du déterminant par rapport à la dernière colonne, on en déduit que :

$$\begin{aligned} g(x_1, \cdots, x_n, x) &= \begin{pmatrix} \langle x_1 \mid x_1\rangle & \cdots & \langle x_1 \mid x_n\rangle & \langle x_1 \mid y\rangle \\ \vdots & \ddots & \cdots & \vdots \\ \langle x_n \mid x_1\rangle & \cdots & \langle x_n \mid x_n\rangle & \langle x_n \mid y\rangle \\ \langle x \mid x_1\rangle & \cdots & \langle x \mid x_n\rangle & d(x, F)^2 + \|y\|^2 \end{pmatrix} \\ &= d(x, F)^2 g(x_1, \cdots, x_n) + g(x_1, \cdots, x_n, y) \\ &= d(x, F)^2 g(x_1, \cdots, x_n) \end{aligned}$$

(le système $\{x_1, \cdots, x_n, y\}$ étant lié, puisque $y \in F$, on a $g(x_1, \cdots, x_n, y) = 0$), soit $d(x, F) = \sqrt{\dfrac{g(x_1, \cdots, x_n, x)}{g(x_1, \cdots, x_n)}}$. D'autre part, la projection orthogonale y de x

sur F s'écrit $y = \sum_{j=1}^{n} a_j x_j$, les coefficients a_j étant solutions du système d'équations normales :

$$G(x_1, \cdots, x_n) \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} \langle x_1 \mid x \rangle \\ \vdots \\ \langle x_n \mid x \rangle \end{pmatrix}$$

et en utilisant les formules de Cramer, on obtient $a_j = \dfrac{g_{j,x}(x_1, \cdots, x_n)}{g(x_1, \cdots, x_n)}$ pour j compris entre 1 et n. □

En utilisant le procédé d'orthogonalisationn de Gram-Schmidt, on peut donner une expression du déterminant de Gram d'une famille libre. On rappelle que la mesure principale de l'angle géométrique de deux vecteurs x, y non nuls dans E est le réel $\theta \in [0, \pi]$ définie par $\langle x \mid y \rangle = \cos(\theta) \|x\| \|y\|$.

Théorème 2.21.

Soient $(x_i)_{1 \leq i \leq n}$ une famille libre de vecteurs dans $(E, \langle \cdot \mid \cdot \rangle)$, $F = \mathrm{Vect}\{x_1, \cdots, x_n\}$ et $(e_i)_{1 \leq i \leq n}$ la base orthonormée de F déduite par le procédé de Gram-Schmidt avec la condition $\langle x_k \mid e_k \rangle > 0$ pour tout k compris entre 1 et n. On a :

$$\begin{aligned} g(x_1, \cdots, x_n) &= \prod_{k=2}^{n} \cos^2(x_k, e_k) \prod_{k=1}^{n} \|x_k\|^2 \\ &= \sin^2(x_1, x_2) \prod_{k=3}^{n} \cos^2(x_k, e_k) \prod_{k=1}^{n} \|x_k\|^2 \quad \textit{(pour } n \geq 3\textit{)} \end{aligned}$$

Preuve. La base orthonormée de F déduite de la base $(x_i)_{1 \leq i \leq n}$ par le procédé de Gram-Schmidt est définie par $e_k = \dfrac{1}{\|y_k\|} y_k$ pour $1 \leq k \leq n$, où $y_1 = x_1$ et $y_k = x_k - \sum_{j=1}^{k-1} \langle x_k \mid e_j \rangle e_j$ pour $2 \leq k \leq n$. La matrice de passage de $(e_i)_{1 \leq i \leq n}$ à $(x_i)_{1 \leq i \leq n}$ est alors donnée par :

$$A = \begin{pmatrix} \|y_1\| & \langle x_2 \mid e_1 \rangle & \cdots & \langle x_n \mid e_1 \rangle \\ 0 & \|y_2\| & \ddots & \vdots \\ \vdots & \ddots & \ddots & \langle x_n \mid e_{n-1} \rangle \\ 0 & \cdots & 0 & \|y_n\| \end{pmatrix}$$

et $g(x_1, \cdots, x_n) = \det(A)^2 = \prod_{k=1}^{n} \|y_k\|^2$ (voir ladémonstration du théorème 2.19).

De $x_k = \|y_k\| e_k + \sum_{j=1}^{k-1} \langle x_k \mid e_j \rangle e_j$, pour k compris entre 2 et n, et de l'orthogonalité des e_j, on déduit que $\|y_k\| = \langle x_k \mid e_k \rangle = \|x_k\| \cos(x_k, e_k)$, ce qui donne en tenant compte de $\|y_1\| = \|x_1\|$:

$$g(x_1, \cdots, x_n) = \prod_{k=2}^{n} \cos^2(x_k, e_k) \prod_{k=1}^{n} \|x_k\|^2$$

Pour $n \geq 3$, en écrivant que :

$$\|y_2\|^2 = \|x_2 - \langle x_2 \mid e_1 \rangle e_1\|^2 = \|x_2\|^2 - \langle x_2 \mid e_1 \rangle^2$$

avec $\langle x_2 \mid e_1 \rangle = \|x_2\| \cos(x_2, e_1) = \|x_2\| \cos(x_2, x_1)$, on obtient :

$$\|y_2\|^2 = \|x_2\|^2 \left(1 - \cos^2(x_2, x_1)\right) = \|x_2\|^2 \sin^2(x_1, x_2)$$

et $g(x_1, \cdots, x_n) = \sin^2(x_1, x_2) \prod_{k=3}^{n} \cos^2(x_k, e_k) \prod_{k=1}^{n} \|x_k\|^2$. □

Théorème 2.22. Inégalités de Hadamard

On a $0 \leq g(x_1, \cdots, x_n) \leq \prod_{k=1}^{n} \|x_k\|^2$, la borne inférieure [resp. supérieure] étant atteinte si, et seulement si, la famille $(x_i)_{1 \leq i \leq n}$ est liée [resp. orthogonale].

Preuve. Si la famille $(x_i)_{1 \leq i \leq n}$ est liée, on a alors $g(x_1, \cdots, x_n) = 0$. On suppose donc cette famille libre et dans ce cas, on a $g(x_1, \cdots, x_n) > 0$ et le théorème précédent nous dit que $g(x_1, \cdots, x_n) \leq \prod_{k=1}^{n} \|x_k\|^2$, l'égalité étant réalisée si, et seulement si, on a $\cos^2(x_k, e_k) = 1$ pour tout k compris entre 2 et n, ce qui compte tenu de $\|y_k\|^2 = \cos^2(x_k, e_k) \|x_k\|^2 = \|x_k\|^2 - \sum_{j=1}^{k-1} \langle x_k \mid e_j \rangle^2$ équivaut à $\langle x_k \mid e_j \rangle = 0$ pour tout j compris entre 1 et $k-1$, soit à $x_k = y_k = \|y_k\| e_k$ pour tout k compris entre 1 et n et la famille $(x_i)_{1 \leq i \leq n}$ est orthogonale. □

En notant $C_j = (a_{ij})_{1 \leq i \leq n} \in \mathbb{R}^n$ la colonne numéro j d'une matrice A dans $\mathcal{M}_n(\mathbb{R})$ et en munissant $\mathbb{R}^n$ de son produit scalaire canonique, on a :

$$\begin{aligned} {}^tAA &= \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ a_{12} & a_{22} & \cdots & a_{n2} \\ \vdots & \vdots & \vdots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{n,n} \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{n,n} \end{pmatrix} \\ &= \begin{pmatrix} \langle C_1 \mid C_1 \rangle & \langle C_1 \mid C_2 \rangle & \cdots & \langle C_1 \mid C_n \rangle \\ \langle C_2 \mid C_1 \rangle & \langle C_2 \mid C_2 \rangle & \cdots & \langle C_2 \mid C_n \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle C_n \mid C_1 \rangle & \langle C_n \mid C_2 \rangle & \cdots & \langle C_n \mid C_n \rangle \end{pmatrix} \end{aligned}$$

et l'inégalité de Hadamard prend la forme :

$$(\det(A))^2 = \det\left({}^tAA\right) = g(C_1, \cdots, x_n) \leq \prod_{k=1}^{n} \|C_k\|^2$$

soit $|\det(A)| \leq \prod_{k=1}^{n} \|C_k\|$.

Exemple 2.5 *Prenant pour matrice A une matrice de Vandermonde $V(\alpha_1, \cdots, \alpha_n)$, où $(\alpha_k)_{1\leq k\leq n} \in \mathbb{R}^n$, on aboutit à :*

$$\prod_{1\leq i<j\leq n}^{n} |\alpha_j - \alpha_i| \leq \prod_{k=1}^{n} \sqrt{1 + \alpha_k^2 + \cdots + \alpha_k^{2(n-1)}}$$

prenant les α_k dans $[-1, 1]$, on obtient $\prod_{1\leq i<j\leq n}^{n} |\alpha_j - \alpha_i| \leq \prod_{k=1}^{n} \sqrt{n} = n^{\frac{n}{2}}$.

2.4.4 Déterminants de Cauchy

À tout entier $n \geq 2$ et toutes suites $(\alpha_k)_{1\leq k\leq n}$ et $(\beta_k)_{1\leq k\leq n}$ d'éléments de $\mathbb{K}$ telles $\alpha_i + \beta_j$ pour tous i, j compris entre 1 et n, on associe la matrice de Cauchy :

$$\left(\left(\frac{1}{\alpha_i + \beta_j}\right)\right)_{1\leq i,j\leq n} = \begin{pmatrix} \frac{1}{\alpha_1+\beta_1} & \cdots & \frac{1}{\alpha_1+\beta_{n-1}} & \frac{1}{\alpha_1+\beta_n} \\ \vdots & \ddots & \vdots & \vdots \\ \frac{1}{\alpha_{n-1}+\beta_1} & \cdots & \frac{1}{\alpha_{n-1}+\beta_{n-1}} & \frac{1}{\alpha_{n-1}+\beta_n} \\ \frac{1}{\alpha_n+\beta_1} & \cdots & \frac{1}{\alpha_n+\beta_{n-1}} & \frac{1}{\alpha_n+\beta_n} \end{pmatrix}$$

Théorème 2.23.

On a :

$$\det\left(\left(\frac{1}{\alpha_i + \beta_j}\right)\right)_{1\leq i,j\leq n} = \frac{\prod\limits_{1\leq i<j\leq n} (\beta_j - \beta_i)(\alpha_j - \alpha_i)}{\prod\limits_{1\leq i,j\leq n} (\alpha_i + \beta_j)}$$

(déterminant de Cauchy).

Preuve. S'il existe deux indices $i \neq j$ tels que $\alpha_j = \alpha_i$ [resp. $\beta_j = \beta_i$], la matrice de Cauchy $A_n = \left(\left(\frac{1}{\alpha_i + \beta_j}\right)\right)_{1\leq i,j\leq n}$ a alors deux lignes [resp. deux colonnes] identiques et son déterminant est nul. L'égalité annoncée est donc trivialement vérifiée dans ce cas. On suppose que les α_i, ainsi que les β_j sont deux à deux distincts et on désigne par F_n la fraction rationnelle $F_n(X) = \frac{\prod\limits_{j=1}^{n-1} (X - \beta_j)}{\prod\limits_{i=1}^{n} (X + \alpha_i)}$. Le numérateur de F_n est de degré $n-1$ et son dénominateur de degré n, on a donc une décomposition en éléments simples de la forme $F_n(X) = \sum_{k=1}^{n} \frac{\lambda_k}{X + \alpha_k}$, les coefficients λ_k étant donnés par $\lambda_k = \dfrac{\prod\limits_{j=1}^{n-1}(\alpha_k+\beta_j)}{\prod\limits_{\substack{i=1\\ i\neq k}}^{n}(\alpha_k-\alpha_i)}$. En développant le déterminant suivant la dernière ligne, on a compte tenu de $F_n(\beta_j)=0$ pour tout j compris entre 0 et $n-1$:

$$D_n = \begin{vmatrix} \frac{1}{\alpha_1+\beta_1} & \cdots & \frac{1}{\alpha_1+\beta_{n-1}} & \frac{1}{\alpha_1+\beta_n} \\ \vdots & \ddots & \vdots & \vdots \\ \frac{1}{\alpha_{n-1}+\beta_1} & \cdots & \frac{1}{\alpha_{n-1}+\beta_{n-1}} & \frac{1}{\alpha_{n-1}+\beta_n} \\ F_n(\beta_1) & \cdots & F_n(\beta_{n-1}) & F_n(\beta_n) \end{vmatrix} = F_n(\beta_n)\det(A_{n-1})$$

D'autre part, en écrivant que $F_n(\beta_j) = \sum\limits_{k=1}^{n}\dfrac{\lambda_k}{\beta_j+\alpha_k}$ et en utilisant le fait que det est une forme multilinéaire alternée, on a aussi $D_n = \lambda_n\det(A_n)$. On a donc $F_n(\beta_n)\det(A_{n-1}) = \lambda_n\det(A_n)$ et :

$$\det(A_n) = \frac{F_n(\beta_n)}{\lambda_n}\det(A_{n-1}) = \frac{\prod\limits_{j=1}^{n-1}(\beta_n-\beta_j)(\alpha_n-\alpha_i)}{\prod\limits_{i=1}^{n}(\beta_n+\alpha_i)\prod\limits_{j=1}^{n-1}(\alpha_n+\beta_j)}\det(A_{n-1})$$

On conclut alors par récurrence sur $n\geq 2$. Pour $n=2$, on a :

$$\det(A_2) = \begin{vmatrix} \frac{1}{\alpha_1+\beta_1} & \frac{1}{\alpha_1+\beta_2} \\ \frac{1}{\alpha_2+\beta_1} & \frac{1}{\alpha_2+\beta_2} \end{vmatrix} = \frac{(\beta_2-\beta_1)(\alpha_2-\alpha_1)}{(\alpha_1+\beta_1)(\alpha_1+\beta_2)(\alpha_2+\beta_1)(\alpha_2+\beta_2)}$$

Et supposant le résultat acquis pour $n-1$, on a :

$$\begin{aligned}\det(A_n) &= \frac{\prod\limits_{j=1}^{n-1}(\beta_n-\beta_j)(\alpha_n-\alpha_i)}{\prod\limits_{i=1}^{n}(\beta_n+\alpha_i)\prod\limits_{j=1}^{n-1}(\alpha_n+\beta_j)}\,\frac{\prod\limits_{1\leq i<j\leq n-1}(\beta_j-\beta_i)(\alpha_j-\alpha_i)}{\prod\limits_{1\leq i,j\leq n-1}(\alpha_i+\beta_j)} \\ &= \frac{\prod\limits_{1\leq i<j\leq n}(\beta_j-\beta_i)(\alpha_j-\alpha_i)}{\prod\limits_{i=1}^{n}(\beta_n+\alpha_i)\prod\limits_{i=1}^{n}(\beta_{n-1}+\alpha_i)\cdots\prod\limits_{i=1}^{n}(\beta_1+\alpha_i)} \\ &= \frac{\prod\limits_{1\leq i<j\leq n}(\beta_j-\beta_i)(\alpha_j-\alpha_i)}{\prod\limits_{1\leq i,j\leq n}(\alpha_i+\beta_j)}\end{aligned}$$

□

Exemple 2.6 *On se place sur l'espace préhilbertien $\mathcal{C}^0([0,1])$ muni du produit scalaire $(f,g)\mapsto\langle f\mid g\rangle = \int_0^1 f(t)\,g(t)\,dt$ et on se donne une suite strictement*

strictement croissante $(\lambda_k)_{k\in\mathbb{N}}$ *de réels positif à laquelle on associe la suite de fonctions* $(\theta_k)_{k\in\mathbb{N}}$ *définie par* $\theta_k(t) = t^{\lambda_k}$ *pour tout* $t \in [0,1]$. *Pour tout* $n \in \mathbb{N}^*$, *le déterminant de Gram de* $(\theta_0, \cdots, \theta_n)$ *est alors donné par :*

$$\begin{aligned} g(\theta_0, \cdots, \theta_n) &= \det\left((\langle \theta_i \mid \theta_j \rangle)\right)_{0\le i,j\le n} = \det\left(\left(\int_0^1 t^{\lambda_i+\lambda_j} dt\right)\right)_{0\le i,j\le n} \\ &= \det\left(\left(\frac{1}{\lambda_i+\lambda_j+1}\right)\right)_{0\le i,j\le n} = \frac{\prod\limits_{0\le i<j\le n} (\lambda_j-\lambda_i)^2}{\prod\limits_{0\le i,j\le n} (\lambda_i+\lambda_j+1)} \\ &= \prod_{j=0}^{n} \frac{\prod\limits_{i=0}^{j-1} (\lambda_j-\lambda_i)^2}{\prod\limits_{i=0}^{n} (\lambda_j+\lambda_i+1)} \end{aligned}$$

Les matrices de Gram correspondantes au choix de $\lambda_k = k$ *pour tout* $k \in \mathbb{N}$ *(*$(\theta_k)_{k\in\mathbb{N}}$ *est la base canonique de l'espace des fonctions polynomiales sur* $[0,1]$*) sont les matrices de Hilbert* $H_n = \left(\left(\int_0^1 t^{i+j} dt\right)\right)_{0\le i,j\le n} = \left(\left(\frac{1}{i+j+1}\right)\right)_{0\le i,j\le n}$ *de déterminants :*

$$\det(H_n) = \prod_{j=0}^{n} \frac{\prod\limits_{i=0}^{j-1} (j-i)^2}{\prod\limits_{i=0}^{n} (i+j+1)} = \prod_{j=0}^{n} \frac{(j!)^3}{(n+j+1)!} = \frac{\left(\prod\limits_{j=2}^{n} j!\right)^4}{\prod\limits_{j=2}^{2n+1} j!}$$

Les déterminants de l'exemple précédent apparaissent dans la démonstration du résultat de densité suivant.

Théorème 2.24. Müntz

L'espace vectoriel $\mathrm{Vect}\{\theta_n \mid n \in \mathbb{N}\}$ *est dense dans l'espace préhilbertien* $\left(\mathcal{C}^0([0,1]), \langle \cdot \mid \cdot \rangle\right)$ *(convergence quadratique) si, et seulement si, on a* $\sum\limits_{n=1}^{+\infty} \frac{1}{\lambda_n} = +\infty$. *Si on suppose de plus que* $\lambda_0 = 0$ *et* $\lambda_1 \ge 1$, *alors l'espace vectoriel* $\mathrm{Vect}\{\theta_n \mid n \in \mathbb{N}\}$ *est dense dans* $\left(\mathcal{C}^0([0,1]), \|\cdot\|_\infty\right)$ *(convergence uniforme) si, et seulement si, on a* $\sum\limits_{n=1}^{+\infty} \frac{1}{\lambda_n} = +\infty$.

Exemple 2.7 *En notant* $(p_n)_{n\ge 1}$ *la suite des nombres premiers rangée dans l'ordre croissant,* $\mathrm{Vect}\{1, x^{p_n} \mid n \ge 1\}$ *est dense dans* $\left(\mathcal{C}^0([0,1]), \|\cdot\|_\infty\right)$.

2.5 Exemples d'utilisation du déterminant

2.5.1 Rang d'une matrice ou d'un endomorphisme

Pour tout matrice $A \in \mathcal{M}_{m,n}(\mathbb{K})$ et tout entier j compris entre 1 et n, on note $C_j \in \mathbb{K}^m$ la colonne numéro j de A.

Définition 2.7. *Le rang d'une matrice $A \in \mathcal{M}_{m,n}(\mathbb{K})$ est la dimension du sous-espace vectoriel de $\mathbb{K}^m$ engendré par ses colonnes :*

$$\mathrm{rg}(A) = \dim(\mathrm{Vect}(C_1, \cdots, C_n))$$

Il est équivalent de dire que le rang de A est égal au nombre maximum de vecteurs colonnes de A linéairement indépendants dans $\mathbb{K}^n$.

Définition 2.8. *Soient E, F deux espaces vectoriels de dimension finie et u une application linéaire de E dans F. Le rang de u est la dimension de* $\mathrm{Im}(u)$. *On le note* $\mathrm{rg}(u)$.

Théorème 2.25. Du rang

Soient E, F deux espaces vectoriels de dimension finie et u une application linéaire de E dans F. On a $\dim(E) = \dim(\ker(u)) + \mathrm{rg}(u)$.

Preuve. Soit H un supplémentaire de $\ker(u)$ dans E et v la restriction de u à H, c'est-à-dire l'application v définie sur H par $v(x) = u(x)$ pour tout $x \in H$. Le noyau de cette application est $\ker(v) = H \cap \ker(u) = \{0\}$, ce qui signifie que v est injective de H dans F et réalise une bijection de H sur $\mathrm{Im}(v)$. En écrivant tout vecteur y de $\mathrm{Im}(u)$ sous la forme $y = u(x)$ avec $x \in E$ qui s'écrit $x = x_1 + x_2$ où $x_1 \in \ker(u)$ et $x_2 \in H$, on a $y = u(x_1) + u(x_2) = v(x_2)$, c'est-à-dire que y est dans $\mathrm{Im}(v)$. On a donc $\mathrm{Im}(u) \subset \mathrm{Im}(v)$ et comme $\mathrm{Im}(v) \subset \mathrm{Im}(u)$, on a en fait $\mathrm{Im}(v) = \mathrm{Im}(u)$ et v réalise un isomorphisme de H sur $\mathrm{Im}(u)$. Il en résulte que $\mathrm{rg}(u) = \dim(\mathrm{Im}(u)) = \dim(H) = \dim(E) - \dim(\ker(u))$. □

Théorème 2.26.

Le rang d'une matrice $A \in \mathcal{M}_{m,n}(\mathbb{K})$ est égal au rang de l'application linéaire $u \in \mathcal{L}(\mathbb{K}^n, \mathbb{K}^m)$ canoniquement associé à la matrice A.

Preuve. En notant $\mathcal{B} = (e_k)_{1 \le k \le n}$ la base canonique de $\mathbb{K}^n$ et $\mathcal{B}' = (f_k)_{1 \le k \le m}$ celle de $\mathbb{K}^m$, pour tout j compris entre 1 et n, la colonne numéro j de A est $C_j = u(e_j)$ et on a :

$$\mathrm{rg}(A) = \mathrm{rg}(C_1, C_2, \cdots, C_n) = \mathrm{rg}(u(e_1), u(e_2), \cdots, u(e_n)) = \mathrm{rg}(u)$$

□

Théorème 2.27.

Si $A \in \mathcal{M}_{m,n}(\mathbb{K})$ est la matrice de $u \in \mathcal{L}(E,F)$ dans des bases $\mathcal{B}$ de E et $\mathcal{B}'$ de F, on a alors $\mathrm{rg}(u) = \mathrm{rg}(A)$.

Preuve. Soient $\mathcal{B} = (e_k)_{1\leq k\leq n}$ une base de E, $\mathcal{B}' = (f_k)_{1\leq k\leq m}$ une base de F et $u \in \mathcal{L}(E,F)$ de matrice A dans ces bases. Dire que $x = \sum_{j=1}^{n} x_j e_j$ est dans le noyau de u équivaut à dire que $u(x) = 0$, ce qui se traduit par $AX = 0$, où $X = (x_j)_{1\leq j\leq n}$ est le vecteur de $\mathbb{K}^n$ formé des composantes de x dans la base $\mathcal{B}$. En désignant par v l'application linéaire canoniquement associée à A, on a $v(X) = AX = 0$, c'est-à-dire que X est dans le noyau de v. Réciproquement si $X \in \ker(v)$, on a $AX = 0$, ce qui équivaut à $u(x) = 0$, soit $x \in \ker(u)$. L'application $x \mapsto X$ réalise donc un isomorphisme de $\ker(u)$ sur $\ker(v)$ et $\dim(\ker(u)) = \dim(\ker(v))$, ce qui équivaut à $\mathrm{rg}(u) = \mathrm{rg}(v) = rg(A)$ en utilisant le théorème du rang. □

Définition 2.9. *Une matrice extraite de $A = ((a_{i,j}))_{\substack{1\leq i\leq m \\ 1\leq j\leq n}} \in \mathcal{M}_{m,n}(\mathbb{K})$ est une matrice $A_{I,J} = ((a_{i,j}))_{(i,j)\in I\times J}$, où :*

$$I = \{1 \leq i_1 < \cdots < i_q \leq m\} \text{ et } J = \{1 \leq j_1 < \cdots < j_p \leq n\}$$

Un déterminant extrait de A (ou de $\det(A)$) est le déterminant d'une matrice carrée extraite de A. Si δ_p est un déterminant extrait de A d'ordre p, on appelle bordant de δ tout déterminant δ_{p+1} extrait d'ordre $p+1$ de A tel que δ_p soit extrait de δ_{p+1}.

Théorème 2.28.

Une matrice $A \in \mathcal{M}_{m,n}(\mathbb{K})$ est de rang r si, et seulement si, il existe un déterminant extrait de A, δ_r d'ordre r et non nul tel que tous les bordants de δ_r sont nuls (si $r = \min(n,m)$, la deuxième condition n'est pas à prendre en compte).

Preuve. Voir [**?**], volume 1. □

Corollaire 2.3. *Le rang d'une matrice $A \in \mathcal{M}_{m,n}(\mathbb{K})$ est l'ordre du plus grand déterminant extrait de A qui est non nul.*

2.5.2 Systèmes de Cramer

On appelle système de Cramer tout système linéaire $Ax = b$ d'inconnue $x \in \mathbb{K}^n$, où $A \in GL_n(\mathbb{K})$ et $b \in \mathbb{K}^n$. L'unique solution d'un tel système est $x = A^{-1}b$. En notant $x = (x_j)_{1\leq j\leq n} \in \mathbb{K}^n$ cette solution, on a $b = \sum_{j=1}^{n} x_j C_j$ et en utilisant le caractère multilinéaire alterné du déterminant, on a pour tout entier k compris entre 1 et n :

$$\begin{aligned}\det\left(C_1,\cdots,C_{k-1},b,C_{k+1},\cdots,C_n\right) &= \sum_{j=1}^{n} x_j \det\left(C_1,\cdots,C_{k-1},C_j,C_{k+1},\cdots,C_n\right)\\ &= x_k \det(A)\end{aligned}$$

soit $x_k = \dfrac{\det\left(C_1,\cdots,C_{k-1},b,C_{k+1},\cdots,C_n\right)}{\det(A)}$ pour $1 \leq k \leq n$. En calculant un déterminant par la formule (2.1), cela nécessite $n!\,(n-1)$ multiplications et $(n!-1)$ additions, donc environ $nn!$ opérations élémentaires. Comme il y a $n+1$ déterminants à calculer puis n divisions à faire pour les formules de Cramer, on aura un total d'environ $n^2 n!$ opérations élémentaires à effectuer, ce qui peut être beaucoup trop important pour de grandes valeurs de n.

Les formules de Cramer peuvent être utilisées pour déterminer l'ensemble des solutions d'un système linéaire de m équations à n inconnues $Ax = b$ d'inconnue $x \in \mathbb{K}^n$, où $A \in \mathcal{M}_{m,n}(\mathbb{K}) \setminus \{0\}$ est de rang $r \in \{1, \min(n,m)\}$ et $b \in \mathbb{K}^m$.

Il existe des matrices carrées d'ordre r extraites de A de déterminant non nul (un tel déterminant est appelé déterminant principal de A) et toute matrice carrée extraite d'ordre plus grand ou égal à $r+1$ (s'il en existe) est de déterminant nul.

En effectuant au besoin des permutations de lignes ou de colonnes du système linéaire, on se ramène à une matrice A tel que la matrice $A_r = ((a_{ij}))_{1 \leq i,j \leq r}$ soit inversible. Le système linéaire :

$$\sum_{j=1}^{n} a_{ij} x_j = b_i \ \ (1 \leq i \leq r)$$

est alors appelé système d'équations principales du système $Ax = b$.

Considérons d'abord le cas d'un système homogène $Ax = 0$. Le système d'équations principales étant de rang r, l'ensemble de ses solutions est un sous-espace vectoriel de $\mathbb{K}^n$ de dimension $n-r$ et comme il contient l'ensemble $S(A,0)$ des solutions de $Ax = 0$, qui est aussi de dimension $n-r$, ces deux espaces sont égaux. La résolution du système d'équations principales peut se faire en utilisant les formules de Cramer pour le système $A_r x^{(r)} = b^{(r)}$, où $x^{(r)} \in \mathbb{K}^r$ a pour composantes les inconnues principales $x_1, \cdots, x_r$ et $b^{(r)} = \left(-\displaystyle\sum_{j=r+1}^{n} a_{ij} x_j\right)_{1 \leq i \leq r} \in \mathbb{K}^r$ (les sommes valant 0 pour $r = n$).

Exemple 2.8 *Considérons un système linéaire homogène de m équations à $m+1$ inconnues, $Ax = 0$ avec $A = ((a_{ij}))_{\substack{1 \leq i \leq m \\ 1 \leq j \leq m+1}} \in \mathcal{M}_{m,m+1}(\mathbb{K})$ de rang m, la matrice $((a_{ij}))_{1 \leq i,j \leq m}$ étant inversible. Ce système est équivalent au système :*

$$\sum_{j=1}^{m} a_{ij} x_j = -a_{1,m+1} x_{m+1} \ \ (1 \leq i \leq m)$$

où x_{m+1} est l'inconnue non principale. Sa solution est donnée par :

$$\begin{aligned}
x_k &= \frac{\det\left(C_1, \cdots, C_{k-1}, b_m, C_{k+1}, \cdots, C_m\right)}{\det\left(A_m\right)} \\
&= -x_{m+1} \frac{\det\left(C_1, \cdots, C_{k-1}, C_{m+1}, C_{k+1}, \cdots, C_m\right)}{\det\left(A_m\right)} \\
&= (-1)^{m+1-k} x_{m+1} \frac{\det\left(C_1, \cdots, C_{k-1}, C_{k+1}, \cdots, C_m, C_{m+1}\right)}{\det\left(A_m\right)} \\
&= \frac{(-1)^{m+1}}{\det\left(A_m\right)} x_{m+1} (-1)^k \det\left(C_1, \cdots, C_{k-1}, C_{k+1}, \cdots, C_m, C_{m+1}\right) \ (1 \leq k \leq m)
\end{aligned}$$

L'ensemble des solutions $S(A, 0)$ est donc la droite vectorielle dirigée par le vecteur $v = (\alpha_k)_{1 \leq k \leq m+1}$, où :

$$\alpha_k = (-1)^k \det\left(C_1, \cdots, C_{k-1}, C_{k+1}, \cdots, C_m, C_{m+1}\right) \ (1 \leq k \leq m)$$

et $\alpha_{m+1} = (-1)^{m+1} \det(A_m)$. Considérons, par exemple, le système :

$$\left\{\begin{array}{l} x+y+z-t=0 \\ 2x+y+z-2t=0 \\ x-y+z-3t=0 \end{array}\right.$$

Il équivaut au système d'inconnue non principale t :

$$\left\{\begin{array}{l} x+y+z=t \\ 2x+y+z=2t \\ x-y+z=3t \end{array}\right.$$

et de solution $x = t,\ y = -t,\ z = t$.

Revenons au système linéaire $Ax = b$ d'inconnue $x \in \mathbb{K}^n$, où $A \in \mathcal{M}_{m,n}(\mathbb{K})$ est de rang $r \geq 1$ et $b \in \mathbb{K}^m$. Ce système aura des solutions si, et seulement si, $b \in F = \operatorname{Vect}\{C_1, \cdots, C_n\}$, ce qui est équivalent à dire que $F = \operatorname{Vect}\{C_1, \cdots, C_n, b\}$ et revient à dire que $\operatorname{rg}(A, b) = \operatorname{rg}(A) = r$ et comme $\det(A_r) \neq 0$, cela équivaut à dire que :

$$\left|\begin{array}{cccc} a_{11} & \cdots & a_{1r} & b_1 \\ \vdots & \ddots & \vdots & \vdots \\ a_{r1} & & a_{rr} & b_r \\ a_{k1} & & a_{kr} & b_k \end{array}\right| = 0 \ (r+1 \leq k \leq n)$$

Si le système est compatible (*i.e.* l'ensemble de ses solutions est non vide), l'ensemble $S(A, b)$ des solutions de $Ax = b$ est alors un sous-espace affine de $\mathbb{K}^n$ de dimension $n - r$ dirigé par $S(A, 0)$. L'ensemble des solutions du système d'équations principales est aussi un sous-espace affine de $\mathbb{K}^n$ de dimension $n - r$ et comme il contient $S(A, b)$, ces deux sous-espaces sont égaux. En définitive, quand il est compatible, le système $Ax = b$ est équivalent au système d'équations principales $A'_r x = b$, où $A'_r = ((a_{ij}))_{\substack{1 \leq i \leq r \\ 1 \leq j \leq n}}$ (deux systèmes linéaires sont équivalents s'ils ont même ensemble de solutions). La résolution de ce dernier système se faisant en utilisant les formules de Cramer pour le système $A_r x^{(r)} = b^{(r)}$, où $x^{(r)} \in \mathbb{K}^r$ a pour composantes les inconnues principales $x_1, \cdots, x_r$ et $b^{(r)} = \left(b_i - \sum_{j=r+1}^{n} a_{ij} x_j \right)_{1 \leq i \leq r} \in \mathbb{K}^r$ (les sommes valant 0 pour $r = n$). Nous avons donc montré le résultat suivant.

Théorème 2.29. Rouché-Fontené

Soient $A = ((a_{ij}))_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}} \in \mathcal{M}_{m,n}(\mathbb{K})$ de rang $r \geq 1$ et $b \in \mathbb{K}^m$. On se donne une matrice $A_r = ((a_{ij}))_{(i,j) \in I \times J}$ dans $GL_r(\mathbb{K})$ extraite de A telle que $\det(A_r) \neq 0$.

1. *Le système $Ax = b$ est compatible si, et seulement si, $r = n$ ou $r \leq n-1$ et :*
$$\det \begin{pmatrix} A_r & (b_i)_{i \in I} \\ (a_{k,j})_{j \in J} & b_k \end{pmatrix} = 0 \ (k \in \{1, \cdots, n\} \setminus I)$$
2. *Si le système $Ax = b$ est compatible, il est alors équivalent au système d'équations principales $A'_r x = b$, où $A_r = ((a_{ij}))_{\substack{i \in I \\ 1 \leq j \leq n}}$ et les inconnues principales $(x_j)_{j \in J}$ s'obtienne comme solutions du système de Cramer $A_r x_r = b^{(r)}$, où $b^{(r)} \in \mathbb{K}^r$ est fonction de b et des inconnues non principales $(x_j)_{j \in \{1, \cdots, m\} \setminus J}$.*

2.5.3 Orientation d'un espace euclidien

E est un espace euclidien de dimension $n \geq 2$.

Théorème 2.30.

Si $\mathcal{B} = (e_i)_{1 \leq i \leq n}$ et $\mathcal{B}' = (e'_i)_{1 \leq i \leq n}$ sont deux bases orthonormées de E, alors la matrice de passage P de $\mathcal{B}$ à $\mathcal{B}'$ est une matrice orthogonale.

Preuve. L'application linéaire u définie par $u(e_j) = e'_j = \sum_{i=1}^{n} p_{ij} e_i$ pour tout j compris entre 1 et n est une isométrie puisqu'elle transforme une base orthonormée en base orthonormée et en conséquence sa matrice dans la base $\mathcal{B}$, qui n'est autre que la matrice $P = ((p_{ij}))_{1 \leq i,j \leq n}$, est orthogonale. □

Avec les notations du théorème précédent, on a $\det(P) = \det_{\mathcal{B}}(\mathcal{B}') = \pm 1$.

On définit une relation sur l'ensemble des bases orthonormées de E en disant qu'une base orthonormée $\mathcal{B}$ est en relation avec une base orthonormée $\mathcal{B}'$ si, et seulement si, la matrice de passage P de $\mathcal{B}$ à $\mathcal{B}'$ est dans $\mathcal{O}_n^+(\mathbb{R})$. On notera $\sim$ cette relation.

Théorème 2.31.

La relation $\sim$ ainsi définie est une relation d'équivalence et il y a exactement deux classes d'équivalence pour cette relation.

Preuve. Cette relation est réflexive puisque $I_n \in \mathcal{O}_n^+(\mathbb{R})$ et $\mathcal{B} = I_d(\mathcal{B})$. Elle est symétrique puisque $P \in \mathcal{O}_n^+(\mathbb{R})$ entraîne $P^{-1} \in \mathcal{O}_n^+(\mathbb{R})$ et si P est la matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$, alors P^{-1} est la matrice de passage de $\mathcal{B}'$ à $\mathcal{B}$. Elle est transitive puisque le produit de deux matrices de $\mathcal{O}_n^+(\mathbb{R})$ est dans $\mathcal{O}_n^+(\mathbb{R})$ ($\mathcal{O}_n^+(\mathbb{R})$ est un groupe).

Soit $\mathcal{B} = (e_i)_{1\leq i\leq n}$ une base orthonormée de E fixée. Pour toute autre base orthonormée $\mathcal{B}' = (e'_i)_{1\leq i\leq n}$, en désignant par $P = ((p_{ij}))_{1\leq i,j\leq n}$ la matrice de passage P de $\mathcal{B}$ à $\mathcal{B}'$, on a soit $P \in \mathcal{O}_n^+(\mathbb{R})$ et $\mathcal{B}' \sim \mathcal{B}$, soit $P \in \mathcal{O}_n^-(\mathbb{R})$ et en désignant par $\mathcal{B}^-$ la base orthonormée définie par $\mathcal{B}^- = (e_1, \cdots, e_{n-1}, -e_n)$, la matrice de passage P^- de $\mathcal{B}^-$ à $\mathcal{B}'$ est :

$$P^- = \begin{pmatrix} p_{11} & \cdots & \cdots & p_{1n} \\ \vdots & \ddots & & \vdots \\ & & \ddots & \\ p_{n-1,1} & & & p_{n-1,n} \\ -p_{nn} & \cdots & \cdots & -p_{nn} \end{pmatrix}$$

et $\det(P^-) = -\det(P) = 1$, donc $P^- \in \mathcal{O}_n^+(\mathbb{R})$ et $\mathcal{B}' \sim \mathcal{B}^-$. Donc $\mathcal{B}'$ est soit dans la classe de $\mathcal{B}$, soit dans celle de $\mathcal{B}^-$ et ces deux classes sont distinctes puisque la matrice de passage de $\mathcal{B}$ à $\mathcal{B}^-$ est $\begin{pmatrix} I_{n-1} & 0 \\ 0 & -1 \end{pmatrix} \in \mathcal{O}_n^-(\mathbb{R})$. On a donc deux classes distinctes. □

Définition 2.10. *Orienter l'espace euclidien E revient à choisir une des deux classes d'équivalence pour la relation $\sim$. Si l'espace E est orienté par le choix d'une classe d'équivalence $\overline{\mathcal{B}_0}$, où $\mathcal{B}_0$ est base orthonormée de E, on dit qu'une base orthonormée $\mathcal{B}$ est directe (ou qu'elle définit la même orientation que $\mathcal{B}_0$) si $\mathcal{B}$ est dans la classe d'équivalence de $\mathcal{B}_0$ et on dit que cette base $\mathcal{B}$ est indirecte (ou rétrograde) dans le cas contraire.*

Orienter l'espace euclidien E revient donc à choisir une base orthonormée $\mathcal{B}_0$.

L'espace $\mathbb{R}^n$, pour $n \geq 2$, est en général orienté par le choix de la base canonique.

Les isométries directes sont celles qui respectent l'orientation. Précisément, on a le résultat suivant.

Théorème 2.32.

Soient $\mathcal{B} = (e_i)_{1\leq i\leq n}$ une base orthonormée de E et u une application linéaire de E dans E. L'application u est une isométrie positive si, et seulement si, elle transforme $\mathcal{B}$ en une base orthonormée de E définissant la même orientation.

Preuve. Soient $u \in \mathcal{L}(E)$ et $\mathcal{B}' = (u(e_i))_{1\leq i\leq n}$. On a :

$$\big(u \in \mathcal{O}^+(E)\big) \Leftrightarrow (u \in \mathcal{O}(E) \text{ et } \det(u) = 1)$$

avec $\det(u) = \det_{\mathcal{B}}(\mathcal{B}')$, donc :

$$\begin{aligned}\left(u \in \mathcal{O}^+(E)\right) &\Leftrightarrow \left(\mathcal{B}' = (u(e_i))_{1\leq i\leq n} \text{ base orthonormée et } \det_{\mathcal{B}}(\mathcal{B}') = 1\right)\\ &\Leftrightarrow (\mathcal{B}' \sim \mathcal{B})\end{aligned}$$

□

2.5.4 Produit mixte et produit vectoriel

E est un espace euclidien de dimension $n \geq 3$ orienté par le choix d'une base orthonormée $\mathcal{B}_0 = (e_i)_{1\leq i\leq n}$.

Si $\mathcal{B}$ est une autre base de E, alors pour tout n-uplet $(x_1, x_2, \cdots, x_n)$ de vecteurs de E, on a $\det_{\mathcal{B}_0}(x_1, x_2, \cdots, x_n) = \det_{\mathcal{B}_0}(\mathcal{B})\det_{\mathcal{B}}(x_1, x_2, \cdots, x_n)$, donc le réel $\det_{\mathcal{B}}(x_1, x_2, \cdots, x_n)$ est indépendante du choix d'une base orthonormée directe $\mathcal{B}$ de E. On le note $[x_1, x_2, \cdots, x_n]$ et on dit que c'est le produit mixte des vecteurs ordonnés $x_1, x_2, \cdots, x_n$.

En remarquant que, pour tout $(n-1)$-uplet $x_1, x_2, \cdots, x_{n-1}$ de vecteurs de E, l'application $x \mapsto \det(x_1, x_2, \cdots, x_{n-1}, x)$ est une forme linéaire, on déduit qu'il existe un unique vecteur $a \in E$ tel que :

$$\forall x \in E,\ \det(x_1, \cdots, x_{n-1}, x) = \langle a \mid x\rangle \tag{2.2}$$

ce vecteur a étant fonction des vecteurs $x_1, x_2, \cdots, x_{n-1}$. On peut donc donner la définition suivante.

Définition 2.11. *Le produit vectoriel des $n-1$ vecteurs $x_1, x_2, \cdots, x_{n-1}$ de E est le vecteur a défini par* (2.2)*. On le note $x_1 \wedge x_2 \wedge ... \wedge x_{n-1}$.*

Dans la base orthonormée $\mathcal{B}_0$, en notant $x_j = \sum_{i=1}^{n} x_{ij}e_i$ pour tout j compris entre 1 et n, les réels $\det(x_1, x_2, \cdots, x_{n-1}, e_i) = \langle x_1 \wedge x_2 \wedge ... \wedge x_{n-1} \mid e_i\rangle$ sont les composantes du vecteur $x_1 \wedge x_2 \wedge ... \wedge x_{n-1}$ dans la base $\mathcal{B}_0$. On a donc :

$$x_1 \wedge ... \wedge x_{n-1} = \sum_{i=1}^{n} (-1)^{i+n}\,\delta_i e_i$$

en désignant par δ_i le déterminant de la matrice d'ordre $n-1$ déduite de la matrice $(X_1, \cdots, X_{n-1})$ en supprimant de cette matrice la ligne numéro i (X_i étant le vecteur de $\mathbb{R}^n$ formé des composantes de x_i dans la base $\mathcal{B}_0$).

On peut remarquer que $(-1)^{i+n}\,\delta_i$ est aussi le cofacteur $C_{i,n}(x_1, x_2, \cdots, x_{n-1})$ d'indice (i, n) de la matrice $(X_1, \cdots, X_{n-1}, 0)$ (*i.e.* celui en ligne i et colonne n).

En utilisant les propriétés du déterminant, on obtient le résultat suivant.

Théorème 2.33.

— *Le produit vectoriel est une application $(n-1)$-linéaire alternée de E^{n-1} dans E ;*

— *le vecteur* $x_1 \wedge x_2 \wedge ... \wedge x_{n-1}$ *est orthogonal à tous les vecteurs* x_i *pour* $1 \leq i \leq n-1$ *;*

— $x_1 \wedge x_2 \wedge ... \wedge x_{n-1} = 0$ *si, et seulement si, la famille* $(x_1, x_2, ..., x_{n-1})$ *est liée ;*

— *si la famille* $(x_1, x_2, ..., x_{n-1})$ *est libre, on a alors :*

$$\det(x_1, \cdots, x_{n-1}, x_1 \wedge ... \wedge x_{n-1}) = \|x_1 \wedge ... \wedge x_{n-1}\|^2 > 0$$

et la famille $(x_1, ..., x_{n-1}, x_1 \wedge ... \wedge x_{n-1})$ *est une base directe de* E *;*

— *si la famille* $(x_1, ..., x_{n-1})$ *est orthonormée, la famille* $(x_1, \cdots, x_{n-1}, x_1 \wedge ... \wedge x_{n-1})$ *est alors une base orthonormée directe de* E.

Preuve.

— Chacune des applications $(x_1, ..., x_{n-1}) \mapsto \det(x_1, \cdots, x_{n-1}, e_i)$ étant $(n-1)$-linéaire alternée, il en est de même de l'application

$$(x_1, ..., x_{n-1}) \mapsto x_1 \wedge ... \wedge x_{n-1} = \sum_{i=1}^{n} \det(x_1, \cdots, x_{n-1}, e_i)\, e_i$$

— Avec $\langle x_1 \wedge ... \wedge x_{n-1} \mid x_i \rangle = \det(x_1, \cdots, x_{n-1}, x_i) = 0$ pour $1 \leq i \leq n-1$, on déduit que $x_1 \wedge ... \wedge x_{n-1}$ est orthogonal à x_i.

— Si la famille $(x_i)_{1 \leq i \leq n-1}$ est liée, il en est alors de même de $(x_1, ..., x_{n-1}, x)$ pour tout $x \in E$ et $\langle x_1 \wedge ... \wedge x_{n-1} \mid x \rangle = \det(x_1, \cdots, x_{n-1}, x) = 0$, ce qui signifie que $x_1 \wedge ... \wedge x_{n-1} \in E^\perp = \{0\}$. Si la famille $(x_i)_{1 \leq i \leq n-1}$ est libre, elle se prolonge alors en une base $(x_1, \cdots, x_{n-1}, x)$ et on a $\langle x_1 \wedge ... \wedge x_{n-1} \mid x \rangle = \det(x_1, \cdots, x_{n-1}, x) \neq 0$, ce qui entraîne $x_1 \wedge ... \wedge x_{n-1} \neq 0$.

— Si la famille $(x_i)_{1 \leq i \leq n-1}$ est libre, on a alors $x_1 \wedge ... \wedge x_{n-1} \neq 0$, donc :

$$\det(x_1, \cdots, x_{n-1}, x_1 \wedge ... \wedge x_{n-1}) = \|x_1 \wedge ... \wedge x_{n-1}\|^2 \neq 0$$

et $(x_1, \cdots, x_{n-1}, x_1 \wedge ... \wedge x_{n-1})$ est une base directe de E.

— Si la famille $(x_i)_{1 \leq i \leq n-1}$ est orthonormée, elle est alors libre et la famille $(x_1, \cdots, x_{n-1}, x_1 \wedge ... \wedge x_{n-1})$ est une base directe de E. De plus le vecteur $x_1 \wedge ... \wedge x_{n-1}$ est orthogonal à l'hyperplan H engendré par $x_1, ..., x_{n-1}$. En prolongeant la famille $(x_i)_{1 \leq i \leq n-1}$ en une base orthonormée directe de E, $(x_1, ..., x_{n-1}, x_n)$, on a $x_1 \wedge ... \wedge x_{n-1} = \lambda x_n$ (ces deux vecteurs sont dans la droite $H^\perp$) et $\lambda = \langle x_1 \wedge ... \wedge x_{n-1} \mid x_n \rangle = \det(x_1, \cdots, x_{n-1}, x_n) = 1$, donc $x_1 \wedge ... \wedge x_{n-1} = x_n$ est de norme 1.

□

Le produit vectoriel peut être utilisé pour donner une expression relativement simple de la distance d'un point à un hyperplan.

Théorème 2.34.

Si H est un hyperplan de E et $(x_1, ..., x_{n-1})$ une base de H, alors la droite $D = H^{\perp}$ est dirigée par le vecteur $x_1 \wedge ... \wedge x_{n-1}$ et pour tout vecteur x de E, la projection orthogonale de x sur H est :

$$p_H(x) = x - \frac{\langle x_1 \wedge ... \wedge x_{n-1} \mid x\rangle}{\|x_1 \wedge ... \wedge x_{n-1}\|^2}(x_1 \wedge ... \wedge x_{n-1})$$

et la distance de x à H est donnée par :

$$d(x, H) = \frac{|\langle x_1 \wedge ... \wedge x_{n-1} \mid x\rangle|}{\|x_1 \wedge ... \wedge x_{n-1}\|} = \frac{|\det(x_1, \cdots, x_{n-1}, x)|}{\|x_1 \wedge ... \wedge x_{n-1}\|}$$

Preuve. Le vecteur $x_1 \wedge ... \wedge x_{n-1}$ étant orthogonal à tous les x_i qui engendrent H, est nécessairement dans $H^{\perp}$. Comme $H^{\perp}$ est une droite et $x_1 \wedge ... \wedge x_{n-1}$ non nul, la droite $D = H^{\perp}$ est dirigée par $x_1 \wedge ... \wedge x_{n-1}$. On a $d(x, H) = \|x - y\|$ où $y = p_H(x)$ est la projection orthogonale de x sur H. Comme $x - y \in H^{\perp}$, il existe un réel λ tel que $x - y = \lambda(x_1 \wedge ... \wedge x_{n-1})$ et avec $\lambda\|x_1 \wedge ... \wedge x_{n-1}\|^2 = \langle x - y \mid x_1 \wedge ... \wedge x_{n-1}\rangle = \langle x \mid x_1 \wedge ... \wedge x_{n-1}\rangle$ (puisque le vecteur y est dans H et $x_1 \wedge ... \wedge x_{n-1}$ dans $H^{\perp}$), on déduit que $\lambda = \dfrac{\langle x_1 \wedge ... \wedge x_{n-1} \mid x\rangle}{\|x_1 \wedge ... \wedge x_{n-1}\|^2}$ et :

$$y = x - \frac{\langle x_1 \wedge ... \wedge x_{n-1} \mid x\rangle}{\|x_1 \wedge ... \wedge x_{n-1}\|^2}(x_1 \wedge ... \wedge x_{n-1})$$

$$d(x, H) = \|x - y\| = \frac{|\langle x_1 \wedge ... \wedge x_{n-1} \mid x\rangle|}{\|x_1 \wedge ... \wedge x_{n-1}\|}$$

□

Le théorème précédent nous dit aussi qu'une équation de l'hyperplan H de base $(x_1, ..., x_{n-1})$ est donnée par :

$$(x \in H) \Leftrightarrow (d(x, H) = 0) \Leftrightarrow (\langle x_1 \wedge ... \wedge x_{n-1} \mid x\rangle = 0)$$

En prenant pour $(x_1, ..., x_{n-1})$ une base orthonormée de H, on a :

$$p_H(x) = x - \langle x_1 \wedge ... \wedge x_{n-1} \mid x\rangle(x_1 \wedge ... \wedge x_{n-1})$$

et $d(x, H) = |\langle x_1 \wedge ... \wedge x_{n-1} \mid x\rangle|$.

2.5.5 Le produit vectoriel en dimension 3

On suppose ici que E est un espace euclidien de dimension 3 orienté par le choix d'une base orthonormée $\mathcal{B}_0 = (e_1, e_2, e_3)$.

Le produit vectoriel de $x = x_1e_1 + x_2e_2 + x_3e_3$ et $y = y_1e_1 + y_2e_2 + y_3e_3$ est le vecteur $z = z_1e_1 + z_2e_2 + z_3e_3$ défini par :

$$z_1 = \langle x \wedge y \mid e_1 \rangle = \det(x, y, e_1) = \begin{vmatrix} x_2 & y_2 \\ x_3 & y_3 \end{vmatrix} = x_2y_3 - x_3y_2$$
$$z_2 = \langle x \wedge y \mid e_2 \rangle = \det(x, y, e_2) = -\begin{vmatrix} x_1 & y_1 \\ x_3 & y_3 \end{vmatrix} = x_3y_1 - x_1y_3$$
$$z_3 = \langle x \wedge y \mid e_3 \rangle = \det(x, y, e_3) = \begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix} = x_1y_2 - x_2y_1$$

Lemme 2.6 *Si (f_1, f_2, f_3) est une base orthonormée directe, on a alors :*

$$f_1 \wedge f_2 = f_3,\ f_2 \wedge f_3 = f_1,\ f_3 \wedge f_1 = f_2$$

Preuve. Pour $1 \leq i \neq j \leq 3$, le vecteur $f_i \wedge f_j$ est orthogonal au plan engendré par f_i, f_j, donc colinéaire à f_k où $\{k\} = \{1, 2, 3\} \setminus \{i, j\}$ et il existe un réel λ tel que $f_i \wedge f_j = \lambda f_k$. Ce réel λ est déterminé par $\lambda = \langle f_i \wedge f_j \mid f_k \rangle = \det(f_i, f_j, f_k) = \pm 1$. On a donc $f_i \wedge f_j = \det(f_i, f_j, f_k) f_k$, ce qui nous donne :

$$f_1 \wedge f_2 = \det(f_1, f_2, f_3) f_3 = f_3$$

$$f_1 \wedge f_3 = \det(f_1, f_3, f_2) f_2 = -f_2$$

$$f_2 \wedge f_3 = \det(f_2, f_3, f_1) f_1 = f_1$$

□

Le caractère 2-linéaire alterné du produit vectoriel se traduit par :

$$\begin{cases} (x + y) \wedge z = x \wedge z + y \wedge z \\ x \wedge (y + z) = x \wedge y + x \wedge z \\ (\lambda x) \wedge y = x \wedge (\lambda y) = \lambda (x \wedge y) \\ x \wedge y = -(y \wedge x) \end{cases}$$

pour tous vecteurs x, y, z et tout réel λ.

Avec la dernière propriété, on déduit que le produit vectoriel est non commutatif ($x \wedge y = -(y \wedge x) \neq y \wedge x$ pour y, x libre).

On a vu que $x \wedge y = 0$ équivaut à dire que x et y sont liés. On en déduit que le produit vectoriel est non associatif. Par exemple, on a :

$$e_1 \wedge (e_1 \wedge e_2) = e_1 \wedge e_3 = -e_2 \neq (e_1 \wedge e_1) \wedge e_2 = 0$$

De manière plus précise, on a les formules du double produit vectoriel suivantes.

Théorème 2.35.

Pour x, y, z dans E, on a :

$$x \wedge (y \wedge z) = \langle x \mid z \rangle y - \langle x \mid y \rangle z$$
$$(x \wedge y) \wedge z = \langle x \mid z \rangle y - \langle y \mid z \rangle x$$

Preuve. Pour tout $x \in E$, on désigne par φ_x l'application linéaire définie par :

$$\forall x \in E,\ \varphi_x(t) = x \wedge t$$

On a alors, pour x, y, z dans E $x \wedge (y \wedge z) = \varphi_x \circ \varphi_y(z)$ et la matrice dans $\mathcal{B}_0$ de $\varphi_x \circ \varphi_y$ est :

$$\begin{aligned} A_x A_y &= \begin{pmatrix} 0 & x_3 & -x_2 \\ -x_3 & 0 & x_1 \\ x_2 & -x_1 & 0 \end{pmatrix} \begin{pmatrix} 0 & y_3 & -y_2 \\ -y_3 & 0 & y_1 \\ y_2 & -y_1 & 0 \end{pmatrix} \\ &= \begin{pmatrix} -x_2y_2 - x_3y_3 & x_2y_1 & y_1x_3 \\ x_1y_2 & -x_1y_1 - x_3y_3 & x_3y_2 \\ x_1y_3 & x_2y_3 & -x_1y_1 - x_2y_2 \end{pmatrix} \\ &= \begin{pmatrix} x_1y_1 & x_2y_1 & y_1x_3 \\ x_1y_2 & x_2y_2 & x_3y_2 \\ x_1y_3 & x_2y_3 & x_3y_3 \end{pmatrix} - (x_1y_1 + x_2y_2 + x_3y_3) \begin{pmatrix} 1 & 0 & 0_3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} \begin{pmatrix} x_1 & x_2 & x_3 \end{pmatrix} - \langle x \mid y \rangle I_3 \end{aligned}$$

et on reconnaît la matrice dans $\mathcal{B}_0$ de l'application linéaire $t \mapsto \langle x \mid t \rangle y - \langle x \mid y \rangle t$. Il en résulte que :

$$(x \wedge y) \wedge z = -z \wedge (x \wedge y) = -(\langle z \mid y \rangle x - \langle z \mid x \rangle y) = \langle x \mid z \rangle y - \langle y \mid z \rangle x$$

On peut aussi se placer dans une base orthonormée directe adaptée. Pour y, z liés, on a $x \wedge (y \wedge z) = 0$. Si $y = 0$, on a alors $\langle x \mid z \rangle y - \langle x \mid y \rangle z = 0$, sinon on a $z = \lambda y$ et $\langle x \mid z \rangle y - \langle x \mid y \rangle z = \lambda (\langle x \mid y \rangle y - \langle x \mid y \rangle y) = 0$. Pour (y, z) libre, on utilise une base orthonormée directe $\mathcal{B} = (f_1, f_2, f_3)$ où $f_1 = \dfrac{1}{\|y\|} y$, $\operatorname{Vect}\{y, z\} = \operatorname{Vect}\{f_1, f_2\}$ et dans cette base $\mathcal{B}$, on a :

$$\left\{ \begin{array}{l} x = x_1 f_1 + x_2 f_2 + x_3 f_3 \\ y = \|y\| f_1 = y_1 f_1 \\ z = z_1 f_1 + z_2 f_2 \end{array} \right. , \ y \wedge z = y_1 z_2 f_3$$

$$\begin{aligned} x \wedge (y \wedge z) &= y_1 z_2 (x_2 f_1 - x_1 f_2) = x_2 z_2 y - x_1 y_1 \left(z - \frac{z_1}{y_1} y \right) \\ &= (x_1 z_1 + x_2 z_2) y - \langle x \mid y \rangle z = \langle x \mid z \rangle y - \langle x \mid y \rangle z \end{aligned}$$

et la relation $x \wedge (y \wedge z) = \langle x \mid z \rangle y - \langle x \mid y \rangle z$ s'en déduit par anti-symétrie du produit vectoriel. □

Corollaire 2.4. (Formule de Lagrange). *Pour x, y dans E, on a $\langle x \mid y \rangle^2 + \|x \wedge y\|^2 = \|x\|^2 \|y\|^2$.*

Preuve. On a :

$$\begin{aligned} \|x \wedge y\|^2 &= \langle x \wedge y \mid x \wedge y \rangle = \det(x, y, x \wedge y) = \det(y, x \wedge y, x) = \langle y \wedge (x \wedge y) \mid x \rangle \\ &= \left\langle \|y\|^2 x - \langle x \mid y \rangle y \mid x \right\rangle = \|x\|^2 \|y\|^2 - \langle x \mid y \rangle^2 \end{aligned}$$

$\square$

On rappelle que si x et y sont des vecteurs non nuls dans E, la mesure principale de leur angle géométrique est le réel $\theta \in [0, \pi]$ définie par $\langle x \mid y \rangle = \cos(\theta) \|x\| \|y\|$. L'identité de Lagrange nous dit alors que :

$$\begin{aligned} \|x \wedge y\|^2 &= \|x\|^2 \|y\|^2 - \langle x \mid y \rangle^2 \\ &= \|x\|^2 \|y\|^2 \left(1 - \cos^2(\theta)\right) = \|x\|^2 \|y\|^2 \sin^2(\theta) \end{aligned}$$

et pour x, y linéairement indépendants, on a $x \wedge y \neq 0$ et $\|x \wedge y\| = \|x\| \|y\| \sin(\theta)$, cette dernière identité étant encore valable pour x, y liés puisque dans ce cas on a, $|\langle x \mid y \rangle| = \|x\| \|y\|$ et θ vaut 0 ou π.

On peut donner une définition géométrique du produit vectoriel en dimension 3 comme suit.

Pour x, y liés, on a $x \wedge y = 0$.

Pour x, y linéairement indépendants, on définit une base orthonormée (f_1, f_2) du plan P engendré par x, y avec $f_1 = \frac{1}{\|x\|} x$, puis on complète cette base en une base orthonormée directe $\mathcal{B} = (f_1, f_2, f_3)$ de E. On dit que le plan P est orienté par le choix de f_3.

La mesure principale $\theta \in \left]-\pi, \pi\right]$ de l'angle orienté des vecteurs x et y dans le plan P ainsi orienté est telle que la rotation ρ_θ d'angle θ transforme $f_1 = \frac{1}{\|x\|} x$ en $\frac{1}{\|y\|} y$.

Dans la base (f_1, f_2) de P, la matrice de ρ_θ est $R_\theta = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$ et l'égalité $\rho_\theta\left(\frac{1}{\|x\|} x\right) = \frac{1}{\|y\|} y$ avec $x = \|x\| f_1$ se traduit par :

$$y = \|y\| \rho_\theta(f_1) = \|y\| (\cos(\theta) f_1 + \sin(\theta) f_2)$$

Comme $\mathcal{B} = (f_1, f_2, f_3)$ est une base orthonormée directe de E, les coordonnées de $x \wedge y$ dans cette base sont données par :

$$\begin{pmatrix} \|x\| \\ 0 \\ 0 \end{pmatrix} \wedge \begin{pmatrix} \|y\| \cos(\theta) \\ \|y\| \sin(\theta) \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \|x\| \|y\| \sin(\theta) \end{pmatrix}$$

c'est-à-dire que $x \wedge y = \|x\| \|y\| \sin(\theta) f_3$.

On a aussi $\|x \wedge y\| = \|x\| \|y\| |\sin(\theta)|$ et $\langle x \mid y \rangle = \|x\| \|y\| \cos(\theta)$ ($|\theta| \in [0, \pi]$ est la mesure principale de l'angle géométrique des vecteurs x, y). On retrouve ainsi l'identité de Lagrange :

$$\langle x \mid y \rangle^2 + \|x \wedge y\|^2 = \|x\|^2 \|y\|^2 \left(\cos^2(\theta) + \sin^2(\theta)\right) = \|x\|^2 \|y\|^2$$

2.6 Exercices

Exercice 2.1. *On suppose que le corps $\mathbb{K}$ est infini et on se donne une matrice* $M = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$ *dans* $\mathcal{M}_{2n}(\mathbb{K})$, *avec* A, B, C, D *dans* $\mathcal{M}_n(\mathbb{K})$ *telles que C et D commutent.*

1. *On suppose que D inversible et on pose* $T = \begin{pmatrix} D & 0 \\ -C & D^{-1} \end{pmatrix}$. *Calculer le produit MT et en déduire que* $\det(M) = \det(AD - BC)$.
2. *Désignant par P le polynôme défini par* $P(X) = \det \begin{pmatrix} A & B \\ C & D - XI_n \end{pmatrix}$ *(il s'agit ici du déterminant d'une matrice à coefficients dans le corps $\mathbb{K}(X)$), montrer que le polynôme Q défini par* $Q(X) = \det(D - XI_n)$ *n'a qu'un nombre fini de racines dans le corps $\mathbb{K}$, puis en déduire que* $\det(M) = \det(AD - BC)$.
3. *Montrer que pour A, B dans $\mathcal{M}_n(\mathbb{R})$ on a* $\det \begin{pmatrix} A & B \\ -B & A \end{pmatrix} \geq 0$.

Solution.

1. En utilisant le fait que D et C commutent on a :

$$MT = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} D & 0 \\ -C & D^{-1} \end{pmatrix} = \begin{pmatrix} AD - BC & BD^{-1} \\ 0 & I_q \end{pmatrix}$$

et $\det \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \det(AD - BC)$.

2. $Q(X) = \det(D - XI_n)$ est un polynôme de degré q, il n'a donc qu'un nombre fini de racines dans $\mathbb{K}$. En conséquence la matrice $D - XI_n$ est inversible pour une infinité de valeurs de X. Pour ces valeurs on a :

$$P(X) = \det \begin{pmatrix} A & B \\ C & D - XI_n \end{pmatrix} = \det(A(D - XI_n) - BC) = R(X)$$

Les polynômes P et R prenant la même valeur pour une infinité de valeurs de X sont donc égaux. Prenant $X = 0$ on obtient $\det \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \det(AD - BC)$.

3. On a :

$$\begin{aligned} \det \begin{pmatrix} A & B \\ -B & A \end{pmatrix} &= \det\left(A^2 + B^2\right) = \det((A + iB)(A - iB)) \\ &= \det(A + iB)\det(A - iB) = |\det(A + iB)|^2 \geq 0 \end{aligned}$$

Exercice 2.2. *Soit $A \in \mathcal{M}_n(\mathbb{K})$. Montrer que :*

$$\operatorname{rg}(C(A)) = \begin{cases} n \text{ si } \operatorname{rg}(A) = n \\ 1 \text{ si } \operatorname{rg}(A) = n-1 \\ 0 \text{ si } \operatorname{rg}(A) \leq n-2 \end{cases}$$

Solution. Si $\operatorname{rg}(A) = n$, la matrice A est alors inversible, donc aussi la matrice $C(A) = \det(A)\ {}^t(A^{-1})$. Si $\operatorname{rg}(A) = n-1$, il existe au moins un déterminant d'ordre $n-1$ extrait de A qui est non nul, donc $\operatorname{rg}(C(A)) \geq 1$. Comme ${}^tC(A) \cdot A = \det(A) I_n = 0$, on a $\operatorname{Im}(A) \subset \ker({}^tC(A))$, donc $\operatorname{rg}(A) \leq n - \operatorname{rg}(C(A))$ et $\operatorname{rg}(C(A)) \leq n - \operatorname{rg}(A) = 1$, ce qui nous donne $\operatorname{rg}(C(A)) = 1$. Si $\operatorname{rg}(A) = n-2$, tous les mineurs de A sont nuls et $C(A) = 0$.

Exercice 2.3. *Calculer le déterminant de l'endomorphisme τ de transposition défini par $\tau(A) = {}^tA$ pour tout $A \in \mathcal{M}_n(\mathbb{K})$.*

Solution. En utilisant le fait que $\mathcal{M}_n(\mathbb{K})$ est somme directe de l'espace $\mathcal{S}_n(\mathbb{K})$ des matrices symétriques et de l'espace $\mathcal{A}_n(\mathbb{K})$ des matrices anti-symétriques (tout $A \in \mathcal{M}_n(\mathbb{K})$ s'écrit $A = \frac{1}{2}(A + {}^tA) + \frac{1}{2}(A - {}^tA)$ et la matrice nulle est l'unique matrice à la fois symétrique et anti-symétrique), on se donne une base $\mathcal{B}$ de $\mathcal{M}_n(\mathbb{K})$ formée de la réunion d'une base $\mathcal{B}_1$ de $\mathcal{S}_n(\mathbb{K})$ et d'une base $\mathcal{B}_2$ de $\mathcal{A}_n(\mathbb{K})$. Pour tout $A \in \mathcal{B}_1$, on a $\tau(A) = A$ et pour tout $A \in \mathcal{B}_2$, on a $\tau(A) = -A$, l'espace $\mathcal{S}_n(\mathbb{K})$ étant de dimension $p = \frac{n(n+1)}{2}$ et $\mathcal{A}_n(\mathbb{K})$ de dimension $q = \frac{n(n-1)}{2}$. La matrice de τ dans la base $\mathcal{B}$ est donc $T = \begin{pmatrix} I_p & 0 \\ 0 & -I_q \end{pmatrix}$ et son déterminant est $\det(\tau) = \det(T) = (-1)^q$.

Exercice 2.4. *En utilisant un déterminant de Vandermonde, montrer que les vecteurs propres $x_1, \cdots, x_p$ non nuls d'une matrice $A \in \mathcal{M}_n(\mathbb{K})$ associés à des valeurs propres $\alpha_1, \cdots, \alpha_p$ deux à deux distinctes sont linéairement indépendants.*

Solution. On a $Ax_i = \alpha_i x_i$ pour $1 \leq i \leq p$ et pour tout entier naturel j, $A^j x_i = \alpha_i^j x_i$ pour $1 \leq i \leq p$. Si $\sum_{i=1}^{p} \lambda_i x_i = 0$, en appliquant A^j, pour $j \in \mathbb{N}$, on a $\sum_{i=1}^{p} \lambda_i \alpha_i^j x_i = 0$. On notant $x_i = (x_{i,k})_{1 \leq k \leq n}$, on a $\sum_{i=1}^{p} \lambda_i \alpha_i^j x_{i,k} = 0$ pour tout $j \in \mathbb{N}$, donc chaque vecteur $X_k = (\lambda_i x_{i,k})_{1 \leq i \leq p}$ est solution de $V(\alpha_1, \cdots, \alpha_p) X = 0$, ce qui donne $X_k = 0$ pour tout k compris entre 1 et n, soit $\lambda_i x_{i,k} = 0$ pour tout i compris entre 1 et p et tout k compris entre 1 et n. On a donc, pour tout i compris entre 1 et p, $\lambda_i x_i = 0$ et $\lambda_i = 0$ puisque $v_i \neq 0$.

Exercice 2.5. *On suppose que $\mathbb{K}$ est de caractéristique nulle et on se donne des scalaires $\alpha_0, \alpha_1, \cdots, \alpha_n$ deux à deux distincts. Montrer que pour tout polynôme $P \in \mathbb{K}[X]$ de degré n, la famille de polynômes $(P(X + \alpha_k))_{0 \leq k \leq n}$ est une base de $\mathbb{K}_n[X]$.*

Solution. Notons $P_k(X) = P(X + \alpha_k)$ pour $0 \leq k \leq n$. Comme P est de degré n, la famille $\left(P^{(j)}\right)_{0 \leq j \leq n}$, qui est échelonnée en degrés, est une base de $\mathbb{K}_n[X]$ (si $P(X) = \sum\limits_{k=0}^{n} a_k X^k$, la matrice de ce système dans la base canonique de $\mathbb{K}_n[x]$ est alors triangulaire supérieure avec les $\dfrac{n!}{(n-j)!} a_n \neq 0$ pour éléments diagonaux). En utilisant la formule de Taylor-Lagrange, on a :

$$P_k(x) = P(x + \alpha_k) = \sum_{j=0}^{n} \frac{\alpha_k^j}{j!} P^{(j)}(x)$$

C'est à dire que la matrice du système $(P_k)_{0 \leq k \leq n}$ dans la base $\left(\dfrac{1}{j!} P^{(j)}\right)_{0 \leq j \leq n}$ est la matrice de Vandermonde $V(\alpha_0, \cdots, \alpha_n)$. Dans le cas particulier où les α_i sont deux à deux distincts cette matrice est inversible et en conséquence $(P_k)_{0 \leq k \leq n}$ est une base de $\mathbb{K}_n[x]$.

Exercice 2.6. *On désigne par E l'espace vectoriel des applications de $\mathbb{R}$ dans $\mathbb{R}$. Montrer, pour tous réels $0 < a_1 < \cdots < a_n$, la famille :*

$$\mathcal{L} = \{f_{a_k} : x \mapsto \sin(a_k x) \mid 1 \leq k \leq n\}$$

est libre dans E.

Solution. Soient $0 < a_1 < \cdots < a_n$ et $\lambda_1, \cdots, \lambda_n$ des réels tels que :

$$\forall x \in \mathbb{R}, \ \sum_{k=1}^{n} \lambda_k \sin(a_k x) = 0$$

Un développement limité en 0 nous donne $\sum\limits_{k=1}^{n} \lambda_k a_k^{2p+1} = 0$ pour tout entier p compris en 0 et $n-1$. C'est–à–dire que $(\lambda_1 a_1, \cdots, \lambda_n a_n)$ est solution d'un système de Vandermonde :

$$\sum_{k=1}^{n} \lambda_k x_k^p = 0 \ \ (0 \leq p \leq n-1)$$

Les a_k étant strictement positifs et deux à deux distincts, on en déduit que les λ_k sont tous nuls. Ce qui prouve que le système $\mathcal{L}$ est libre dans E.

Exercice 2.7. *Soit* $A_n = \begin{pmatrix} 1^{n-1} & 2^{n-1} & \cdots & n^{n-1} \\ 2^{n-1} & 3^{n-1} & \cdots & (n+1)^{n-1} \\ \vdots & \vdots & \cdots & \vdots \\ n^{n-1} & (n+1)^{n-1} & \cdots & (2n-1)^{n-1} \end{pmatrix}$.

Montrer que :

$$\det(A_n) = (-1)^{\frac{n(n-1)}{2}} ((n-1)!)^n$$

Solution. On utilise le théorème 2.17 en prenant $\alpha_k = k$ et $P_k(X) = (X+k-1)^{n-1}$ pour $1 \leq k \leq n$. Dans ce cas, on a $\Delta(1, \cdots, n) = \prod_{j=1}^{n-1} j!$ (exemple 2.3), $P_1(X) = X^{n-1}$ et $P_k(X) = \sum_{i=0}^{n-1} \binom{n-1}{i} (k-1)^{n-1-i} X^i$ pour $2 \leq k \leq n$, donc :

$$\begin{aligned} \det(Q(P_1, \cdots, P_n)) &= \begin{vmatrix} 0 & \binom{n-1}{0} & \binom{n-1}{0} 2^{n-1} & \cdots & \binom{n-1}{0} (n-1)^{n-1} \\ 0 & \binom{n-1}{1} & \binom{n-1}{1} 2^{n-2} & \cdots & \binom{n-1}{1} (n-1)^{n-2} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & \binom{n-1}{n-2} & \binom{n-1}{n-2} 2 & \cdots & \binom{n-1}{n-2} n-1 \\ 1 & \binom{n-1}{n-1} & \binom{n-1}{n-1} & \cdots & \binom{n-1}{n-1} \end{vmatrix} \\ &= \prod_{i=0}^{n-1} \binom{n-1}{i} \begin{vmatrix} 0 & 1 & 2^{n-1} & \cdots & (n-1)^{n-1} \\ 0 & 1 & 2^{n-2} & \cdots & (n-1)^{n-2} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 1 & 2 & \cdots & n-1 \\ 1 & 1 & 1 & \cdots & 1 \end{vmatrix} \end{aligned}$$

soit :

$$\begin{aligned} \det(Q(P_1, \cdots, P_n)) &= (-1)^{n+1} \prod_{i=1}^{n-2} \binom{n-1}{i} \begin{vmatrix} 1 & 2^{n-1} & \cdots & (n-1)^{n-1} \\ 1 & 2^{n-2} & \cdots & (n-1)^{n-2} \\ \vdots & \vdots & \cdots & \vdots \\ 1 & 2 & \cdots & n-1 \end{vmatrix} \\ &= (-1)^{n+1} \prod_{i=1}^{n-2} \binom{n-1}{i} (n-1)! \begin{vmatrix} 1 & 2^{n-2} & \cdots & (n-1)^{n-2} \\ 1 & 2^{n-3} & \cdots & (n-1)^{n-3} \\ \vdots & \vdots & \cdots & \vdots \\ 1 & 1 & \cdots & 1 \end{vmatrix} \end{aligned}$$

avec :

$$\begin{pmatrix} 1 & 2^{n-2} & \cdots & (n-1)^{n-2} \\ 1 & 2^{n-3} & \cdots & (n-1)^{n-3} \\ \vdots & \vdots & \cdots & \vdots \\ 1 & 1 & \cdots & 1 \end{pmatrix} = P_\sigma \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 1 & 2 & \cdots & n-1 \\ \vdots & \vdots & \cdots & \vdots \\ 1 & 2^{n-2} & \cdots & (n-1)^{n-2} \end{pmatrix}$$

où P_σ est la matrice de pemutation associée à $\sigma = \begin{pmatrix} 1 & 2 & \cdots & n-1 \\ n-1 & n-2 & \cdots & 1 \end{pmatrix}$,

ce qui nous donne en notant $\lambda_n = (-1)^{n+1} \prod_{i=1}^{n-2} \binom{n-1}{i} (n-1)!$ et $\varepsilon(\sigma)$ la signature de la parmutation σ :

$$\begin{aligned}
\det(Q(P_1, \cdots, P_n)) &= \lambda_n \varepsilon(\sigma) \begin{vmatrix} 1 & 1 & \cdots & 1 \\ 1 & 2 & \cdots & n-1 \\ \vdots & \vdots & \cdots & \vdots \\ 1 & 2^{n-2} & \cdots & (n-1)^{n-2} \end{vmatrix} \\
&= \lambda_n \varepsilon(\sigma) \Delta(1, \cdots, n-1)
\end{aligned}$$

avec :

$$\begin{aligned}
\lambda_n \Delta(1, \cdots, n-1) &= (-1)^{n+1} ((n-1)!)^{n-1} \prod_{i=1}^{n-2} \frac{1}{i!(n-1-i)!} \prod_{j=1}^{n-2} j \\
&= (-1)^{n+1} ((n-1)!)^{n-1} \prod_{j=1}^{n-2} \frac{1}{j!}
\end{aligned}$$

et :

$$\begin{aligned}
\varepsilon(\sigma) &= \prod_{1 \le i < j \le n-1} \frac{\sigma(j) - \sigma(i)}{j-i} = \prod_{1 \le i < j \le n-1} \frac{(n-j) - (n-i)}{j-i} \\
&= \prod_{1 \le i < j \le n-1} (-1) = (-1)^{\frac{(n-1)(n-2)}{2}}
\end{aligned}$$

donc :

$$\begin{aligned}
\det(A_n) &= \Delta(1, \cdots, n) \det(Q(P_1, \cdots, P_n)) \\
&= \left(\prod_{j=1}^{n-1} j! \right) (-1)^{n+1} \varepsilon(\sigma) ((n-1)!)^{n-1} \prod_{j=1}^{n-2} \frac{1}{j!} \\
&= (-1)^{\frac{n(n-1)}{2}} ((n-1)!)^n
\end{aligned}$$

Exercice 2.8. *Soient E un espace euclidien orienté par le choix d'une base orthonormée $\mathcal{B}_0 = (e_i)_{1 \le i \le n}$ et σ une permutation de $\{1, 2, \cdots, n\}$. À quelle condition portant sur σ la base $\mathcal{B}_\sigma = \left(e_{\sigma(i)}\right)_{1 \le i \le n}$ est-elle directe ?*

Solution. Notant $\varepsilon(\sigma)$ la signature de σ, on a $\det_{\mathcal{B}_0}(\mathcal{B}_\sigma) = \varepsilon(\sigma) \det(I_n) = \varepsilon(\sigma)$ et $\mathcal{B}_\sigma$ est directe si, et seulement si, σ est une permutation paire.

Exercice 2.9. *Donner une équation du plan vectoriel P de $\mathbb{R}^3$ engendré par $u = (1,1,1)$ et $v = (1,2,3)$. Calculer la distance de $x = (1-1,1)$ à P.*

Solution. Ce plan est orthogonal au vecteur $u \wedge v = (1,-2,1)$, une équation est donc $x_1 - 2x_2 + x_3 = 0$ et la distance de $x = (1-1,1)$ à P est donnée par :

$$d(x,P) = \frac{|\langle u \wedge v \mid x\rangle|}{\|x_1 \wedge v\|} = \frac{4}{\sqrt{6}}$$

Exercice 2.10. *E est un espace euclidien orienté de dimension 3.*

1. *Montrer que, pour tous vecteurs x,y,z,t dans E, on a :*

$$\begin{aligned}\langle x \wedge y \mid z \wedge t\rangle &= \begin{vmatrix} \langle x \mid z\rangle & \langle y \mid z\rangle \\ \langle x \mid t\rangle & \langle y \mid t\rangle \end{vmatrix} \\ &= \langle x \mid z\rangle \langle y \mid t\rangle - \langle x \mid t\rangle \langle y \mid z\rangle\end{aligned}$$

   (pour $(z,t) = (x,y)$, on retrouve la formule de Lagrange).

2. *Soient x,y,z dans $E \setminus \{0\}$. À quelle condition a-t-on :*

$$x \wedge (y \wedge z) = (x \wedge y) \wedge z$$

3. *Soient a,b dans E avec $a \neq 0$. Résoudre l'équation $a \wedge x = b$ (division vectorielle).*
4. *Montrer que pour x,y,z,t dans E, on a :*

$$(x \wedge y) \wedge (z \wedge t) = -\det(x,y,z)\,t + \det(x,y,t)\,z$$

   et $\det(t,y,z)\,x + \det(x,t,z)\,y + \det(x,y,t)\,z = \det(x,y,z)\,t$.

Solution.

1. On a :

$$\begin{aligned}\langle x \wedge y \mid z \wedge t\rangle &= \det(x,y,z \wedge t) = \det(y, z \wedge t, x) \\ &= \langle y \wedge (z \wedge t) \mid x\rangle = \langle \langle y \mid t\rangle z - \langle y \mid z\rangle t \mid x\rangle \\ &= \langle x \mid z\rangle \langle y \mid t\rangle - \langle x \mid t\rangle \langle y \mid z\rangle = \begin{vmatrix} \langle x \mid z\rangle & \langle y \mid z\rangle \\ \langle x \mid t\rangle & \langle y \mid t\rangle \end{vmatrix}\end{aligned}$$

2. En utilisant les formules de double produit vectoriel, on a :

$$x \wedge (y \wedge z) - (x \wedge y) \wedge z = \langle y \mid z\rangle x - \langle x \mid y\rangle z$$

   Si x et z sont liés, il existe alors un réel $\lambda \neq 0$ tel que $z = \lambda x$ et :

$$x \wedge (y \wedge z) - (x \wedge y) \wedge z = \lambda (\langle y \mid x\rangle x - \langle x \mid y\rangle x) = 0$$

Si x et z sont linéairement indépendants, on a alors :

$$\begin{aligned}(x \wedge (y \wedge z) - (x \wedge y) \wedge z = 0) &\Leftrightarrow (\langle y \mid z \rangle = \langle x \mid y \rangle = 0) \\ &\Leftrightarrow \left(y \in (\mathrm{vect}\,(x, z))^{\perp} \right)\end{aligned}$$

avec $(\mathrm{vect}\,(x, z))^{\perp} = \mathbb{R}\,(x \wedge z)$. On en déduit que :

$$(x \wedge (y \wedge z) = (x \wedge y) \wedge z) \Leftrightarrow (y \text{ et } x \wedge z \text{ sont liés})$$

En définitive, $x \wedge (y \wedge z) = (x \wedge y) \wedge z$ si, et seulement si, x et z sont liés ou x et z sont linéairement indépendants et y est orthogonal au plan engendré par x et z.

3. Notons $S = \{x \in E \mid a \wedge x = b\}$. Comme $a \wedge x$ est orthogonal au vecteur non nul a, on a $S = \emptyset$ si b n'est pas orthogonal à a. Supposons b orthogonal à a, soit que $\langle a \mid b \rangle = 0$. En utilisant la formule du double produit vectoriel, on a $a \wedge (a \wedge b) = \langle a \mid b \rangle\, a - \langle a \mid a \rangle\, b = -\left\| a \right\|^2 b$, donc le vecteur $x_0 = -\dfrac{1}{\left\| a \right\|^2} a \wedge b$ est une solution particulière. Si $x \in E$ est une autre solution, on a alors $a \wedge (x - x_0) = 0$, ce qui équivaut à dire que $x - x_0$ est colinéaire au vecteur a. L'ensemble des solutions est donc $S = \{x_0 + \lambda a \mid \lambda \in \mathbb{R}\}$, à savoir la droite affine passant par $x_0 = -\dfrac{1}{\left\| a \right\|^2} a \wedge b$ et dirigée par a.
4. En utilisant la formule du double produit vectoriel, on a :

$$(x \wedge y) \wedge (z \wedge t) = \langle x \wedge y \mid t \rangle\, z - \langle x \wedge y \mid z \rangle\, t = \det(x, y, t)\, z - \det(x, y, z)\, t$$

On a aussi :

$$(x \wedge y) \wedge (z \wedge t) = -\,(z \wedge t) \wedge (x \wedge y) = \det(z, t, x)\, y - \det(z, t, y)\, x$$

donc :

$$\begin{aligned}\det(x, y, t)\, z - \det(x, y, z)\, t &= \det(z, t, x)\, y - \det(z, t, y)\, x \\ &= -\det(x, t, z)\, y - \det(t, y, z)\, x\end{aligned}$$

soit $\det(t, y, z)\, x + \det(x, t, z)\, y + \det(x, y, t)\, z = \det(x, y, z)\, t$.

Exercice 2.11. *E est un espace euclidien de dimension 3 orienté par le choix d'une base orthonormée $\mathcal{B}_0 = (e_1, e_2, e_3)$ et pour y dans E on désigne par φ_y l'endomorphisme de E défini par :*

$$\forall x \in E,\ \varphi_y(x) = x \wedge y$$

1. *Déterminer la matrice de φ_y dans la base $\mathcal{B}_0$.*
2. *Montrer que $\varphi_y^3 = -\left\| y \right\|^2 \varphi_y$.*
3. *En déduire une expression simplifiée de $e^{\varphi_y}(x)$ pour x, y dans E avec $y \neq 0$.*

Solution.

1. Les coordonnées dans la base $\mathcal{B}_0$ de $\varphi_y(x)$ sont :
$$\begin{pmatrix} x_2y_3 - x_3y_2 \\ x_3y_1 - x_1y_3 \\ x_1y_2 - x_2y_1 \end{pmatrix} = \begin{pmatrix} 0 & y_3 & -y_2 \\ -y_3 & 0 & y_1 \\ y_2 & -y_1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$
ce qui nous donne la matrice A_y (anti-symétrique) de φ_y dans la base $\mathcal{B}_0$.
2. On a :
$$\begin{aligned} A_y^3 &= \begin{pmatrix} 0 & -y_2^2y_3 - y_1^2y_3 - y_3^3 & y_2y_3^2 + y_1^2y_2 + y_2^3 \\ y_1^2y_3 + y_2^2y_3 + y_3^3 & 0 & -y_1y_3^2 - y_1^3 - y_1y_2^2 \\ -y_1^2y_2 - y_2^3 - y_2y_3^2 & y_1y_2^2 + y_1y_1^3 + y_1y_3^2 & 0 \end{pmatrix} \\ &= -\left(y_1^2 + y_2^2 + y_3^2\right) \begin{pmatrix} 0 & y_3 & -y_2 \\ -y_3 & 0 & y_1 \\ y_2 & -y_1 & 0 \end{pmatrix} \end{aligned}$$
On peut aussi écrire $\varphi_y^2(x) = (x \wedge y) \wedge y = \langle x \mid y \rangle y - \|y\|^2 x$ et :
$$\varphi_y^3(x) = \langle x \mid y \rangle (y \wedge y) - \|y\|^2 (x \wedge y) = -\|y\|^2 \varphi_y(x)$$
3. On rappelle que l'exponentielle de φ_y est définie par $e^{\varphi_y} = \sum_{n=0}^{+\infty} \frac{1}{n!}\varphi_y^n$. De la question précédente, on déduit que pour tout $p \geq 0$, on a $\varphi_y^{2(p+1)} = \left(-\|y\|^2\right)^p \varphi_y^2$ et $\varphi_y^{2p+1} = \left(-\|y\|^2\right)^p \varphi_y$. Pour $p = 0$ c'est vrai et supposant le résultat acquis pour $p \geq 0$, on a :
$$\varphi_y^{2(p+2)} = \left(-\|y\|^2\right)^p \varphi_y^4 = \left(-\|y\|^2\right)^p \left(-\|y\|^2 \varphi_y^2\right) = \left(-\|y\|^2\right)^{p+1} \varphi_y^2$$
et $\varphi_y^{2p+3} = \left(-\|y\|^2\right)^p \varphi_y^3 = \left(-\|y\|^2\right)^{p+1} \varphi_y$. Il en résulte que :
$$\begin{aligned} e^{\varphi_y} &= Id + \left(\sum_{p=0}^{+\infty} \frac{1}{(2(p+1))!} \left(-\|y\|^2\right)^p\right) \varphi_y^2 + \left(\sum_{p=0}^{+\infty} \frac{1}{(2p+1)!} \left(-\|y\|^2\right)^p\right) \varphi_y \\ &= Id - \frac{1}{\|y\|^2} \left(\sum_{p=0}^{+\infty} \frac{(-1)^p \|y\|^{2p}}{(2p)!} - 1\right) \varphi_y^2 + \frac{1}{\|y\|} \left(\sum_{p=0}^{+\infty} \frac{(-1)^p \|y\|^{2p+1}}{(2p+1)!}\right) \varphi_y \\ &= Id + \frac{1}{\|y\|^2} (1 - \cos(\|y\|)) \varphi_y^2 + \frac{1}{\|y\|} (\sin(\|y\|)) \varphi_y \end{aligned}$$
soit pour tout $x \in E$:
$$\begin{aligned} e^{\varphi_y}(x) &= x - \frac{1}{\|y\|^2} (\cos(\|y\|) - 1) \left(\langle x \mid y \rangle y - \|y\|^2 x\right) + \frac{\sin(\|y\|)}{\|y\|} x \wedge y \\ &= x + (1 - \cos(\|y\|)) \left(\frac{\langle x \mid y \rangle}{\|y\|^2} y - x\right) + \frac{\sin(\|y\|)}{\|y\|} x \wedge y \\ &= \cos(\|y\|)\, x + (1 - \cos(\|y\|)) \frac{\langle x \mid y \rangle}{\|y\|^2} y + \frac{\sin(\|y\|)}{\|y\|} x \wedge y \end{aligned}$$

Exercice 2.12. *E est un espace euclidien de dimension 3 orienté. Montrer que $u \in \mathcal{L}(E)$ est anti-symétrique si, et seulement si, il existe un unique vecteur $a \in E$ tel que $u(x) = a \wedge x$ pour tout $x \in E$.*

Solution. Pour tout $a \in E$, l'application linéaire $u : x \mapsto a \wedge x$ est bien anti-symétrique. En effet, pour x, y dans E, on a :

$$\begin{aligned} \langle u(x) \mid y \rangle &= \langle a \wedge x \mid y \rangle = \det(a, x, y) = -\det(a, y, x) = -\langle a \wedge y \mid x \rangle \\ &= -\langle u(y) \mid x \rangle = -\langle x \mid u(y) \rangle \end{aligned}$$

donc $u^* = -u$.

Soit $u \in \mathcal{L}(E)$ de matrice A dans la base orthonormée $\mathcal{B}_0$. La matrice de u^* dans la base $\mathcal{B}_0$ est ${}^tA$ et pour u anti-symétrique, on a $u^* = -u$, donc ${}^tA = -A$. Étant en dimension impaire, on a $\det(A) = \det({}^tA) = -\det(A)$ et $\det(A) = 0$. Il en résulte que $\ker(u)$ est de dimension $1, 2$ ou 3. Si $u = 0$, on a alors $u(x) = a \wedge x$ pour tout $x \in E$ avec $a = 0$. Pour $u \neq 0$, $\ker(u)$ est de dimension 1 ou 2. Pour tout $x \in \ker(u)$ et tout $y \in E$, on a $\langle x \mid u(y) \rangle = -\langle u(x) \mid y \rangle = 0$, donc $x \in (\mathrm{Im}(u))^\perp$. On a donc $\ker(u) \subset (\mathrm{Im}(u))^\perp$ et l'égalité avec les dimensions (le théorème du rang nous dit que $\dim(\ker(u)) = 3 - \dim(\mathrm{Im}(u)) = \dim\left((\mathrm{Im}(u))^\perp\right)$). On a donc $E = \mathrm{Im}(u) \oplus (\mathrm{Im}(u))^\perp = \mathrm{Im}(u) \oplus \ker(u)$. Si $\ker(u)$ est de dimension 2, $\mathrm{Im}(u)$ est alors de dimension 1, soit $\mathrm{Im}(u) = \mathbb{R} \cdot e$ avec $e \neq 0$. Il en résulte qu'il existe un réel non nul λ tel que $u(e) = \lambda e$ ($\lambda = 0$ donne $e \in \ker(u) \cap \mathrm{Im}(u)$ avec $e \neq 0$, ce qui n'est pas possible) et on a :

$$\lambda \|e\|^2 = \langle e \mid u(e) \rangle = -\langle u(e) \mid e \rangle = -\lambda \|e\|^2$$

donc $\lambda \|e\| = 0$ avec λ et $\|e\|$ non nul, ce qui n'est pas possible. En définitive $\ker(u)$ est de dimension 1 et $\mathrm{Im}(u)$ de dimension 2. En désignant par (f_1, f_2, f_3) une base orthonormée directe telle que f_1 dirige $\ker(u)$ et (f_2, f_3) engendre $\mathrm{Im}(u)$, la matrice de u dans cette base est anti-symétrique de la forme $A' = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & \alpha \\ 0 & -\alpha & 0 \end{pmatrix}$ ($\mathrm{Im}(u)$ est stable par u) avec $\alpha \neq 0$. Pour $x = x_1 f_1 + x_2 f_2 + x_3 f_3$ dans E, on a alors :

$$u(x) = \alpha(x_3 f_2 - x_2 f_3) = (-\alpha f_1) \wedge x = a \wedge x$$

S'il existe un autre vecteur $b \in E$ tel que $u(x) = b \wedge x$ pour tout $x \in E$, on a alors $u(a) = a \wedge a = 0 = b \wedge a$, les vecteur a et b sont liés. Si $a = 0$, on a alors $u(x) = b \wedge x = 0$ pour tout $x \in E$ et b est nul (sinon en prenant x linéairement indépendant de b, on a $b \wedge x \neq 0$). Si $a \neq 0$, il existe alors un réel λ tel que $b = \lambda a$ et $u(x) = a \wedge x = \lambda a \wedge x$ pour tout $x \in E$. Pour x indépendant de a, on a $a \wedge x \neq 0$ et nécessairement $\lambda = 1$, soit $b = a$.