Algèbre 1-GROUPES et GÉOMÉTRIE

David Harari

1. Le groupe linéaire $\mathrm{GL}(E)$

Dans toute cette section, on désignera par K un corps commutatif et par E un K-espace vectoriel de dimension finie n. Le *groupe linéaire* $(\mathrm{GL}(E), \circ)$ des endomorphismes bijectifs de E est isomorphe au groupe multiplicatif matriciel $\mathrm{GL}_n(K)$; nous travaillerons indifféremment avec l'un ou l'autre. On a déjà vu (dans le chapitre "Groupes") une décomposition en produit semi-direct $\mathrm{GL}_n(K) \simeq \mathrm{SL}_n(K) \rtimes K^*$.

1.1. Générateurs et centres de $\mathrm{GL}(E)$, $\mathrm{SL}(E)$

Deux types de transformations vont jouer un rôle privilegié : les *dilatations* et les *transvections*.

Proposition 1.1 *Soient H un hyperplan de E et $u \in \mathrm{GL}(E)$ dont la restriction à H est l'identité. Les assertions suivantes sont équivalentes :*

1. *Le déterminant λ de u est différent de 1.*
2. *u est diagonalisable et admet une valeur propre $\lambda \neq 1$.*
3. *$\mathrm{Im}\,(u - \mathrm{Id})$ n'est pas inclus dans H.*
4. *Il existe une base dans laquelle la matrice de u est $\mathrm{Diag}(1, 1, ..., 1, \lambda)$ avec $\lambda \neq 1$.*

On dit alors que u est une dilatation d'hyperplan H, de droite $D := \mathrm{Im}\,(u - \mathrm{Id})$, et de rapport λ.

Démonstration : 4. $\Rightarrow$ 1. est trivial.

1. $\Rightarrow$ 2.: on a déjà 1 comme valeur propre de multiplicité au moins $n-1$ puisque la restriction de u à H est l'identité. L'autre valeur propre est $\lambda \neq 1$, donc la dimension de chaque sous-espace propre est la multiplicité de la valeur propre associée, et u est diagonalisable.

2. $\Rightarrow$ 3.: Soit D le sous-espace propre (de dimension 1) associé à λ, alors la restriction de $(u - \mathrm{Id})$ à D est une homothétie de rapport non nul, donc $\mathrm{Im}\,(u - \mathrm{Id})$ contient D qui n'est pas incluse dans H (les sous-espaces propres sont en somme directe).

3. $\Rightarrow$ 4.: Comme $(u - \mathrm{Id})$ est de rang 1 (ce n'est pas 0 par 3., et son noyau contient un hyperplan), son image est une droite D et 3. dit que E est la somme directe de H et D. D'autre part D est stable par u car c'est l'image de $(u - \mathrm{Id})$; en recollant une base de H avec une base de D, on obtient une base de E dans laquelle la matrice de u est de la forme voulue.

□

Proposition 1.2 *Soient H un hyperplan de E et $u \in \mathrm{GL}(E)$ dont la restriction à H est l'identité. On suppose $u \neq \mathrm{Id}$ et on fixe une forme linéaire non nulle f sur E telle que $H = \ker f$. Alors les assertions suivantes sont équivalentes :*

1. *Le déterminant λ de u est 1.*
2. *u n'est pas diagonalisable*
3. *$\mathrm{Im}\,(u - \mathrm{Id})$ est incluse dans H.*
4. *L'endomorphisme induit $\bar{u} : E/H \to E/H$ est l'identité.*
5. *Il existe $a \neq 0$ dans H tel que $u(x) = x + f(x)a$ pour tout x de E.*
6. *Il existe une base dans laquelle la matrice de u s'écrit (par blocs)*

$$\begin{pmatrix} I_{n-2} & 0 \\ 0 & U \end{pmatrix}$$

où U est la matrice $(2,2)$ définie par

$$U = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

On dit alors que u est une transvection d'hyperplan H et de droite $D := \mathrm{Im}\,(u - \mathrm{Id})$.

Remarque : Attention u détermine complètement les sous-espaces H et D mais réciproquement n'est pas complètement déterminé par eux (en changeant dans la matrice ci-dessus le 1 qui n'est pas sur la diagonale par n'importe quel $\lambda \neq 0$, on obtient encore une transvection de même hyperplan et même droite). On fera également très attention au fait que par convention, l'identité n'est pas une transvection (en particulier dans la formule du 5., il est important que a soit non nul).

On notera $\tau(a, f)$ la transvection donnée par $\tau(a, f)(x) = x + f(x)a$ quand f est un élément non nul du dual E^*, et a un élément non nul de $\ker f$ (ceci est sans ambiguîté). En particulier l'inverse de $\tau(a, f)$ est la transvection $\tau(-a, f)$.

Preuve de la proposition : 6. $\Rightarrow$ 1., 1. $\Rightarrow$ 2., et 2. $\Rightarrow$ 3. s'obtiennent par contrapposée à partir de la proposition précédente.

3. $\Rightarrow$ 4.: si x est dans E, alors $u(x) - x$ est dans D, donc dans H donc sa classe dans E/H est nulle, i.e. $\bar{u}(\bar{x}) = \bar{x}$.

4. $\Rightarrow$ 5.: On choisit x_0 tel que $f(x_0) = 1$ et on pose $a = u(x_0) - x_0$, alors a est dans H car $\bar{a} = 0$ d'après 4. D'autre part $a \neq 0$ car x_0 n'est pas dans H, qui est le noyau de $(u - \mathrm{Id})$ vu que $u \neq \mathrm{Id}$. Maintenant l'égalité $u(x) = x + f(x)a$ a lieu pour tout x de H et pour x_0 donc pour tout x de E puisque E est la somme directe de H et de Kx_0.

5. $\Rightarrow$ 6.: on complète $e_{n-1} := a$ en une base $(e_1, ..., e_{n-1})$ de H, puis on complète encore avec un vecteur e_n tel que $f(e_n) = 1$ (donc qui n'est pas dans l'hyperplan H). Alors $u(e_i) = e_i$ pour $1 \leq i \leq n-1$, et $u(e_n) = e_n + f(e_n)a = e_n + e_{n-1}$.

□

Il est intéressant de savoir ce qui se passe quand on conjugue une transvection.

Proposition 1.3 *Soient $u \in \mathrm{GL}(E)$ et τ une transvection de droite D et d'hyperplan H. Alors $u\tau u^{-1}$ est une transvection de droite $u(D)$, d'hyperplan $u(H)$. Plus précisément, si $\tau = \tau(a, f)$, on a $u\tau u^{-1} = \tau(u(a), f \circ u^{-1})$.*

Démonstration : Si $\tau(x) = x + f(x)a$, alors $\tau(u^{-1}(x)) = u^{-1}(x) + f(u^{-1}(x))a$ et $u\tau(u^{-1}(x)) = x + f(u^{-1}(x))u(a)$, avec $u(a)$ non nul dans le noyau de la forme linéaire $(f \circ u^{-1})$.

□

On en déduit :

Proposition 1.4 *Le centre Z de $\mathrm{GL}(E)$ est constitué des $\lambda\mathrm{Id}$, $\lambda \in K^*$. Le centre de $\mathrm{SL}(E)$ est $\mathrm{SL}(E) \cap Z$, c'est-à-dire que c'est l'ensemble des $\lambda\mathrm{Id}$ avec $\lambda^n = 1$.*

Démonstration : Il suffit de montrer que si $u \in \mathrm{GL}(E)$ commute avec toutes les transvections, alors u est une homothétie. Mais d'après la proposition précédente, une telle u laisse stable toute droite de E. C'est un exercice élémentaire de vérifier alors que u est une homothétie.

□

Passons à l'engendrement :

Theorème 1.5 *Les transvections engendrent $\mathrm{SL}(E)$. Les transvections et les dilatations engendrent $\mathrm{GL}(E)$.*

Démonstration : Le deuxième point se déduit immédiatement du premier car si $u \in \mathrm{GL}(E)$ est de déterminant λ, on obtient un élément de $\mathrm{SL}(E)$ en composant u avec une dilatation de rapport λ^{-1}. Pour montrer que les transvections engendrent $\mathrm{SL}(E)$, on commence par un lemme (qui ne marche pas en dimension 1, attention...) :

Lemme 1.6 *Supposons $n \geq 2$ et soient x, y deux vecteurs non nuls de E. Alors il existe une transvection ou un produit de deux transvections envoyant x sur y.*

Démonstration : Supposons que x et y ne soient pas colinéaires. Posons $a = y - x$, il existe alors un hyperplan H contenant a et pas x, puis une forme linéaire f d'hyperplan H avec $f(x) = 1$. Alors la transvection u définie par $u(z) = z + f(z)a$ envoie x sur y. Si maintenant x et y sont colinéaires, on utilise l'hypothèse $n \geq 2$ pour trouver un troisième vecteur $w \neq 0$ qui n'est pas colináire à x et y. D'après ce qu'on a vu plus haut, il existe des transvections u, v avec $u(x) = w$ et $v(w) = y$ donc $vu(x) = y$.

□

Nous montrons maintenant le théorème par récurrence sur n. Pour $n = 1$, il est évident. Supposons donc $n \geq 2$, et le théorème démontré pour $n - 1$. Soit $u \in \mathrm{SL}(E)$ et $x_0 \neq 0$ dans E. D'après le lemme, on peut supposer $u(x_0) = x_0$ (quitte à composer u avec un produit de transvections qui ramène $u(x_0)$ sur x_0). Soit D la droite Kx_0, alors l'endomorphisme $\bar{u} : E/D \to E/D$ induit par u est encore de déterminant 1 : pour le voir il suffit de prendre une base $(\bar{e}_2, ..., \bar{e}_n)$ de E/D, alors $(x_0, e_2, ..., e_n)$ est une base de E dans laquelle la matrice de u s'écrit par blocs

$$\begin{pmatrix} 1 & * \\ 0 & \overline{M} \end{pmatrix}$$

où $\overline{M}$ est la matrice de $\bar{u}$ dans $(\bar{e}_2, ..., \bar{e}_n)$.

En appliquant alors l'hypothèse de récurrence à $\bar{u}$, on peut écrire $\bar{u} = \bar{\tau}_1 ... \bar{\tau}_r$, où les $\bar{\tau}_i$ sont des transvections de E/D (rappelons que l'inverse d'une transvection est encore une transvection). Écrivons $\bar{\tau}_i = \tau(\bar{f}_i, \bar{a}_i)$ avec $a_i \in E$ et $\bar{f}_i \in (E/D)^*$. Définissons $f_i \in E^*$ par $f_i(x) = \bar{f}_i(\bar{x})$, alors $\tau_i := \tau(a_i, f_i)$ est une transvection de E. Soit alors $v = \tau_1 ... \tau_r$, on observe que u et v coïncident en x_0 car $u(x_0) = x_0$ et $\tau_i(x_0) = x_0$ vu que $f_i(x_0) = \bar{f}_i(\bar{x}_0) = \bar{f}_i(\bar{0}) = 0$. D'autre part $\bar{v} = \bar{u}$ donc $w := v^{-1}u$ est une transvection ou l'identité de E car $\mathrm{Im}\,(w - \mathrm{Id}) \subset D$, avec D incluse dans $\ker(w - \mathrm{Id})$ (qui est de dimension n ou $n-1$ puisque $\mathrm{Im}\,(w - \mathrm{Id})$ est de dimension au plus 1).

□

Remarque : Le procédé ci-dessus ne donne pas le nombre minimal de transvections dans la décomposition de u. Le nombre optimal est n si u n'est pas une homothétie, $n+1$ sinon, ce qu'on peut montrer en utilisant des méthodes matricielles plus algorithmiques (opérations élémentaires sur les lignes et les colonnes).

1.2. Conjugaison et commutateurs

On commence par la proposition évidente suivante, qui résulte par exemple de la forme matricielle d'une dilatation :

Proposition 1.7 *Deux dilatations sont conjuguées dans* $\mathrm{GL}(E)$ *si et seulement si elles ont même rapport.*

Pour les transvections, l'énoncé est un peu plus compliqué.

Proposition 1.8 *Deux transvections sont toujours conjuguées dans* $\mathrm{GL}(E)$. *Si* $n \geq 3$, *elles sont conjuguées dans* $\mathrm{SL}(E)$.

Démonstration : Soient u et v deux transvections. Comme elles ont même matrice dans deux bases différentes, il existe $g \in \mathrm{GL}(E)$ tel que $v = gug^{-1}$. Soit $\lambda = \det g$. Pour montrer que u et v sont conjuguées dans $\mathrm{SL}(E)$, il suffit de trouver $s \in \mathrm{GL}(E)$, de déterminant λ^{-1}, et qui commute avec v

car alors on aura $v = (sg)u(sg)^{-1}$ avec $sg \in \mathrm{SL}(E)$. Or si $n \geq 3$, la matrice de v dans une certaine base est

$$\begin{pmatrix} I_{n-2} & 0 \\ 0 & U \end{pmatrix}$$

et il suffit de prendre $s = \mathrm{Diag}(\lambda I_{n-2}, \lambda^{-1}, \lambda^{-1})$.

□

[Exercice : pour $n = 2$, les matrices $\begin{pmatrix} 1 & \lambda \\ 0 & 1 \end{pmatrix}$ et $\begin{pmatrix} 1 & \mu \\ 0 & 1 \end{pmatrix}$ sont conjuguées dans $\mathrm{SL}_2(K)$ si et seulement si λ/μ est un carré dans K^*. C'est toujours le cas si K est algébriquement clos, mais par exemple pour $K = \mathbf{R}$ il y a deux classes de conjugaison, et pour $K = \mathbf{Q}$ une infinité.]

Le théorème suivant sur les sous-groupes dérivés est un peu analogue aux résultats sur le groupe symétrique et le groupe alterné.

Theorème 1.9 *1. On a $D(\mathrm{GL}_n(K)) = \mathrm{SL}_n(K)$ sauf dans un cas : $n = 2$ et K est de cardinal 2 (i.e. isomorphe à $\mathbf{Z}/2\mathbf{Z}$).*

2. On a $D(\mathrm{SL}_n(K)) = \mathrm{SL}_n(K)$ sauf si les deux conditions suivantes sont remplies : $n = 2$, et K est de cardinal au plus 3 (i.e. isomorphe à $\mathbf{Z}/2\mathbf{Z}$ ou $\mathbf{Z}/3\mathbf{Z}$).

Démonstration : Déjà, tout commutateur est de déterminant 1 donc les sous-groupes dérivés sont inclus dans $\mathrm{SL}_n(K)$. Si d'autre part on montre qu'il existe une transvection u qui s'écrit comme un commutateur $[a, b] = aba^{-1}b^{-1}$ avec a, b dans $\mathrm{GL}_n(K)$ (resp. dans $\mathrm{SL}_n(K)$), alors toute transvection v s'écrit $v = gug^{-1}$ avec $g \in \mathrm{GL}_n(K)$ d'après la proposition précédente, donc $v = [gag^{-1}, gbg^{-1}]$ est dans $D(\mathrm{GL}_n(K))$ (resp. $D(\mathrm{SL}_n(K))$ puisque $\mathrm{SL}_n(K) \lhd \mathrm{GL}_n(K)$).

On pourrait être tenté, si $n \geq 3$, d'écrire que le carré u^2 d'une transvection u est une transvection, donc $u^2 = gug^{-1}$ avec $g \in \mathrm{SL}_n(K)$ soit $u = [u, g]$; mais ceci est incorrect si K est de caractéristique 2, [1] car alors u^2 est l'identité, et non pas une transvection. On va donc devoir calculer explicitement en distinguant différents cas :

i) Si K est de cardinal au moins 4, on peut choisir λ distinct de 0, 1, et -1 dans K. Posons $t = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ et $s = \begin{pmatrix} \lambda & 0 \\ 0 & \lambda^{-1} \end{pmatrix}$, alors $[s, t] = \begin{pmatrix} 1 & \lambda^2 - 1 \\ 0 & 1 \end{pmatrix}$

[1] Attention, un corps de caractéristique 2 peut être infini, ex. $K = \mathbf{Z}/2Z(T)$. On verra qu'il y a un et un seul corps fini de cardinal p^r pour tout nombre premier p et tout $r \geq 1$.

est une transvection, d'où le résultat si $n = 2$. Si maintenant n est quelconque, le commutateur de $s' = \begin{pmatrix} I_{n-2} & 0 \\ 0 & s \end{pmatrix}$ et $t' = \begin{pmatrix} I_{n-2} & 0 \\ 0 & t \end{pmatrix}$ est encore une transvection.

ii) Si K est de cardinal 2 ou 3 et $n \geq 3$, on prend $t = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ et $s = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, alors $[t, s] = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}$ est bien une transvection car $[t, s]$ est de déterminant 1 et le rang de $[t, s] - I_n$ est 1.

iii) Si enfin K est de cardinal 3 et $n = 2$, on prend $t = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ et $s = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$, alors $[s, t] = t$ est une transvection (mais noter qu'ici s n'est pas dans $\mathrm{SL}_2(K)$, donc on obtient juste le résultat pour le sous-groupe dérivé de $\mathrm{GL}_2(K)$).

□

[Exercice : Le groupe $\mathrm{GL}_2(\mathbf{Z}/2\mathbf{Z}) = \mathrm{SL}_2(\mathbf{Z}/2\mathbf{Z})$ est isomorphe à $\mathcal{S}_3$ (qui n'est pas égal à son sous-groupe dérivé). Le groupe $\mathrm{SL}_2(\mathbf{Z}/3\mathbf{Z})$ est d'ordre 24, et son sous-groupe dérivé d'ordre 8, plus précisément ce sous-groupe dérivé est le *groupe des quaternions* d'ordre 8, il est non-abélien mais pas isomorphe au groupe diédral.]

1.3. Le groupe $\mathrm{PSL}_n(K)$-cas général

Soit $\mathrm{PGL}_n(K) = \mathrm{GL}_n(K)/K^*$ le *groupe projectif linéaire* : c'est le quotient de $\mathrm{GL}_n(K)$ par son centre. De même on note $\mathrm{PSL}_n(K) = \mathrm{SL}_n(K)/\mu_n(K)$ le quotient de $\mathrm{SL}_n(K)$ par son centre. Notons que $\mathrm{PSL}_n(K)$ se plonge dans $\mathrm{PGL}_n(K)$ par l'application naturelle qui associe à $\bar{g}$ la classe de g dans $\mathrm{PGL}_n(K)$ pour tout $g \in \mathrm{SL}_n(K)$.

Le théorème principal de toute cette section est le suivant :

Theorème 1.10 *Le groupe* $\mathrm{PSL}_n(K)$ *est simple sauf dans deux cas exceptionnels :* $n = 2$*, avec* K *de cardinal 2 ou 3.*

Remarque : Si K est fini, cela fournit une nouvelle série de groupes finis simples. Notons aussi qu'à partir de ce résultat, il est facile de retrouver le théorème 1.9, mais la preuve du théorème 1.10 étant assez longue, il était utile d'en avoir une preuve directe.

Démonstration : On va démontrer le théorème à l'exception du cas $n = 2$, $\#K = 4$ ou 5, qui sera vu en 1.4. comme conséquence des isomorphismes exceptionnels. La méthode est assez similaire à celle que l'on avait faite pour le groupe alterné, en utilisant des commutateurs; mais il y a des complications en dimension 2 qui obligent à faire des calculs directs dans ce cas.

Preuve dans le cas $n \geq 3$. Soit E un K-espace vectoriel de dimension $n \geq 3$. On considère un sous groupe distingué $\overline{N}$ non trivial de $\mathrm{PSL}(E)$ (quotient de $\mathrm{SL}(E)$ par son centre Z), alors son image réciproque N par la surjection canonique $\pi : \mathrm{SL}(E) \to \mathrm{PSL}(E)$ est un sous-groupe distingué de $\mathrm{SL}(E)$ contenant strictement Z, et tout revient à montrer que $N = \mathrm{SL}(E)$. Il suffit pour cela de montrer que N contient une transvection, d'après le théorème 1.5 et la proposition 1.8.

Choissisons $\sigma \in N$ qui n'est pas dans Z, alors comme Z est constitué des homothéties de $\mathrm{SL}(E)$ (proposition 1.4), il existe $a \in E$ tel que $b := \sigma(a)$ ne soit pas colinéaire à a. Soit τ une transvection de droite Ka, on pose $\rho = [\sigma, \tau] = \sigma\tau\sigma^{-1}\tau^{-1}$, alors $\rho \in N$, et $\rho \neq \mathrm{Id}$ car les droites des transvections τ et $\sigma\tau\sigma^{-1}$ sont respectivement Ka et Kb. Comme $n \geq 3$, il existe un hyperplan H de E qui contient $Ka \oplus Kb$. D'autre part, on a $\rho(x) - \tau^{-1}(x) \in Kb$ car $\sigma\tau\sigma^{-1}$ est une transvection de droite Kb, donc comme $\tau^{-1}(x) - x \in Ka$ vu que τ^{-1} est une transvection de droite Ka, on obtient

$$\mathrm{Im}\,(\rho - \mathrm{Id}) \subset H$$

Ainsi H est stable par ρ. L'idée est alors de reprendre un commutateur $v = [\rho, u]$, où u est une transvection d'hyperplan H. Alors $v \in N$, et comme v est le produit de deux transvections ($\rho u \rho^{-1}$ et u^{-1}) de même hyperplan (rappelons que $\rho(H) = H$), c'est l'identité ou une transvection d'après la formule $\tau(c, f)\tau(d, f) = \tau(c + d, f)$, où f est une forme linéaire d'hyperplan H. On a donc deux cas :

- S'il existe une transvection u d'hyperplan H qui ne commute pas à ρ, alors v (construit plus haut) est une transvection dans N, ce qui termine la preuve.

- Sinon montrons que la restriction de ρ à H est l'identité. Soit c dans H, prenons pour u la transvection $\tau(c, f)$. Alors $\rho u = u\rho$ donne pour tout x de E :

  $$\rho(x) + f(x)\rho(c) = \rho(x) + f(\rho(x))c$$

  mais si x n'est pas dans H, alors $x = \rho(x) + y$ avec y dans H (on a vu que $\mathrm{Im}\,(\rho - \mathrm{Id}) \subset H$), d'où $f(x)\rho(c) = f(x)c$ avec $f(x) \neq 0$, i.e. $\rho(c) = c$. Comme $\det \rho = 1$ et $\rho_{|H} = \mathrm{Id}_{|H}$, ρ était déjà une transvection.

□

La preuve dans le cas $n \geq 3$ est ainsi achevée. Pour le cas $n = 2$, il faut malheureusement faire un certain nombre de calculs explicites.

Le cas $n = 2$, $\#K \geq 7$. On commence par trois lemmes.

Lemme 1.11 *Supposons K de cardinal au moins 7. Soit $s = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ dans $\mathrm{SL}_2(K)$ avec $c \neq 0$. Alors il existe $g \in \mathrm{SL}_2(K)$ tel que $g^{-1}s^{-1}gs$ ait une valeur propre λ dans $K - \{0, 1, -1\}$.*

Démonstration : On cherche $g = \begin{pmatrix} \alpha & \beta \\ 0 & \delta \end{pmatrix}$ tel que $g^{-1}s^{-1}gs(e_1) = \lambda e_1$ où e_1 est le premier vecteur de la base canonique et $\lambda \neq 0, 1, -1$. Cela s'écrit $gs(e_1) = \lambda sg(e_1)$, soit

$$a\alpha + c\beta = \lambda a\alpha \quad c\delta = \lambda c\alpha \quad \alpha\delta = 1$$

Les deux dernières égalités sont équivalentes à $\delta = \lambda\alpha$, $\delta^2 = \lambda$, et si elles sont satisfaites on a un β qui convient car $c \neq 0$. On choisit pour λ un carré de K^* autre que $-1, 0, 1$, ce qui est possible car K^{*^2} est un sous-groupe de K^* avec $K^*/\{\pm 1\} \simeq K^{*^2}$ (ainsi K^{*^2} est infini, ou bien d'indice 1 ou 2 dans K^*) et K est de cardinal au moins 7. Puis on pose $\alpha = \delta/\lambda$.

□

Lemme 1.12 *Soit $s \in \mathrm{SL}_2(K)$ possédant $\lambda \neq 0, 1, -1$ comme valeur propre. Alors s est conjugué de $t := \begin{pmatrix} \lambda & 0 \\ 0 & \lambda^{-1} \end{pmatrix}$ dans $\mathrm{SL}_2(K)$.*

Démonstration : Déjà s est conjuguée de t dans $\mathrm{GL}_2(K)$, car ce sont deux matrices diagonalisables à valeurs propres distinctes λ et $1/\lambda$. On écrit $s = utu^{-1}$ avec $\det u = d$, alors $v = \begin{pmatrix} d^{-1} & 0 \\ 0 & 1 \end{pmatrix}$ commute avec u d'où $t = (uv)^{-1}s(uv)$ avec $\det(uv) = 1$.

□

Lemme 1.13 *Soit $\lambda \neq 0, 1, -1$, $t := \begin{pmatrix} \lambda & 0 \\ 0 & \lambda^{-1} \end{pmatrix}$, et $u = \begin{pmatrix} 1 & \mu \\ 0 & 1 \end{pmatrix}$ Alors il existe g dans $\mathrm{SL}_2(K)$ tel que $g^{-1}t^{-1}gt = u$.*

Démonstration : On cherche g comme dans le lemme 1.11. On obtient $\alpha\delta = 1$, $\beta/\lambda = \alpha\lambda\mu + \beta\lambda$. On prend alors $\beta = \lambda\mu$, $\delta = 1/\alpha$, ce qui donne $\alpha = 1/\lambda - \lambda$ qui est bien non nul.

□

Fin de la preuve du cas $n = 2$, $\#K \geq 7$. Soit N un sous-groupe distingué de $\mathrm{SL}_2(K)$ contenant un élément s qui n'est pas le centre $Z = \{\pm\mathrm{Id}\}$. On distingue trois cas :

- Si s possède une valeur propre $\lambda \neq 0, 1, -1$ dans K, alors $t = \begin{pmatrix} \lambda & 0 \\ 0 & \lambda^{-1} \end{pmatrix}$ est dans N d'après le lemme 1.12. Par le lemme 1.13, toute $\begin{pmatrix} 1 & \mu \\ 0 & 1 \end{pmatrix}$ avec $\mu \neq 0$ est dans N; mais si τ est une transvection, alors il existe $g \in \mathrm{GL}_2(K)$ tel que $g\tau g^{-1} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, et pour $u = \begin{pmatrix} 1 & 0 \\ 0 & \mu^{-1} \end{pmatrix}$, on a alors $(ug)\tau(ug)^{-1} = \begin{pmatrix} 1 & \mu \\ 0 & 1 \end{pmatrix}$ donc τ est conjuguée dans $\mathrm{SL}_2(K)$ de $\begin{pmatrix} 1 & \mu \\ 0 & 1 \end{pmatrix}$ avec $\mu = \det g$. Finalement toute transvection est dans N, donc $N = \mathrm{SL}_2(K)$.
- Si $s = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ avec $c \neq 0$, on est ramené au cas précédent avec le lemme 1.11.
- Si enfin $s = \begin{pmatrix} a & b \\ 0 & d \end{pmatrix}$ et n'est pas comme dans le premier cas, alors comme s n'est pas $\pm\mathrm{Id}$, s n'est pas diagonalisable donc $b \neq 0$ et $a = d = \pm 1$. Alors $u = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ est dans $\mathrm{SL}_2(K)$ et $usu^{-1} = \begin{pmatrix} a & 0 \\ -b & a \end{pmatrix}$ ce qui ramène au deuxième cas.

□

La preuve du théorème 1.10 est ainsi achevée, à l'exception du cas $n = 2$, K de cardinal 4 ou 5, qu'on traitera en 1.4. [2]

[2] Rappelons que le cardinal d'un corps fini est une puissance de sa caractéristique, donc il n'y a pas de corps de cardinal 6.

1.4. Le groupe $\mathrm{PSL}_2(K)$: les cas exceptionnels.

Dans ce paragraphe, on va s'intéresser plus particulièrement à $\mathrm{PSL}_2(K)$ quand K est fini. On commence par une proposition donnant le cardinal des groupes $\mathrm{PSL}_n(K)$.

Proposition 1.14 *Soient K un corps fini [3] de cardinal q et $n \in \mathbf{N}^*$. Alors :*

$$\#\mathrm{GL}_n(K) = (q^n - 1)(q^n - q)...(q^n - q^{n-1})$$
$$\#\mathrm{SL}_n(K) = \#\mathrm{PGL}_n(K) = (q^n - 1)(q^n - q)...(q^n - q^{n-2})q^{n-1}$$
$$\#\mathrm{PSL}_n(K) = \#\mathrm{SL}_n(K)/d$$

où $d = (n, q-1)$ est le pgcd de n et $q-1$.

Démonstration : On a déjà vu le résultat pour $\mathrm{GL}_n(K)$ (dans la preuve du premier théorème de Sylow), qu'on montre en comptant les bases de K^n. Le deuxième résultat vient de ce que $\mathrm{PGL}_n(K) = \mathrm{GL}_n(K)/K^*$ et $\mathrm{SL}_n(K)$ est le noyau du morphisme surjectif det : $\mathrm{GL}_n(K) \to K^*$. Pour montrer le troisième point, il suffit de vérifier que le cardinal de l'ensemble $\mu_n(K)$ des racines n-ièmes de l'unité de K est d, vu que $\mathrm{PSL}_n(K) = \mathrm{SL}_n(K)/\mu_n(K)$. Déjà on a $\mu_n(K) = \mu_d(K)$ car si $x \in K^*$, on a $x^{q-1} = 1$ vu que le groupe multiplicatif K^* est d'ordre $q-1$, donc l'ordre d'une racine n-ième de l'unité divise n et $q-1$, donc divise d.

Il y a au plus d racines de l'unité dans K car le polynôme $X^d - 1$ a au plus d racines. D'autre part le polynôme $X^{q-1} - 1$ est scindé sur K (il a $q-1$ racines distinctes qui sont les éléments de K^*), donc aussi $X^d - 1$ qui le divise puisque d divise $q-1$ (si $q-1 = md$, on a $X^{q-1} - 1 = (X^d - 1)(1 + X^d + ... + X^{d(m-1)})$. Finalement il y a bien exactement d racines d-ièmes de l'unité dans K.

□

Theorème 1.15 *Soit $\mathbf{F}_q$ le corps fini à q éléments. On a les isomorphismes (dits exceptionnels) :*

$$\mathrm{GL}_2(\mathbf{F}_2) = \mathrm{SL}_2(\mathbf{F}_2) = \mathrm{PGL}_2(\mathbf{F}_2) = \mathrm{PSL}_2(\mathbf{F}_2) \simeq \mathcal{S}_3$$
$$\mathrm{PGL}_2(\mathbf{F}_3) \simeq \mathcal{S}_4 \quad \mathrm{PSL}_2(\mathbf{F}_3) \simeq \mathcal{A}_4$$
$$\mathrm{PGL}_2(\mathbf{F}_4) = \mathrm{PSL}_2(\mathbf{F}_4) \simeq \mathcal{A}_5$$
$$\mathrm{PGL}_2(\mathbf{F}_5) \simeq \mathcal{S}_5 \quad \mathrm{PSL}_2(\mathbf{F}_5) \simeq \mathcal{A}_5$$

En particulier $\mathrm{PSL}_2(\mathbf{F}_2)$ et $\mathrm{PSL}_2(\mathbf{F}_3)$ ne sont pas simples, tandis que $\mathrm{PSL}_2(\mathbf{F}_4)$ et $\mathrm{PSL}_2(\mathbf{F}_5)$ le sont.

[3] On note souvent $\mathbf{F}_q$ *le* corps fini de cardinal q, car on verra plus tard que si q est la puissance d'un nombre premier p, il existe un et un seul corps fini de cardinal q à isomorphisme près.

Remarque : On a $\mathbf{F}_2 = \mathbf{Z}/2\mathbf{Z}$, $\mathbf{F}_3 = \mathbf{Z}/3\mathbf{Z}$, $\mathbf{F}_5 = \mathbf{Z}/5\mathbf{Z}$, mais $\mathbf{F}_4$ n'est ni l'anneau $\mathbf{Z}/4\mathbf{Z}$, ni $\mathbf{Z}/2\mathbf{Z} \times \mathbf{Z}/2\mathbf{Z}$ (qui ne sont pas des corps !). En fait comme on le verra plus tard $\mathbf{F}_4$ est l'anneau quotient de $\mathbf{Z}/2\mathbf{Z}[X]$ par l'idéal engendré par $X^2 + X + 1$.

Démonstration : Soient K un corps fini, E le K-espace vectoriel K^n. On note $\mathbf{P}(E) = \mathbf{P}_K^{n-1}$ (noter le décalage d'indice) l'ensemble des droites vectorielles de K^n, appelé *espace projectif de dimension* $n-1$. C'est aussi l'ensemble quotient de $K^n - \{0\}$ par la relation d'équivalence $x \sim y$ si et seulement si x et y sont colinéaires. Le groupe $\mathrm{GL}(E)$ opère sur $\mathbf{P}(E)$ via $g.D = g(D)$, et comme le centre opère trivialement on obtient un morphisme $\Phi : \mathrm{PGL}(E) \to \mathcal{S}(\mathbf{P}(E))$ qui est injectif, vu que les seuls $g \in \mathrm{GL}(E)$ qui stabilisent toutes les droites sont les homothéties.

Prenons maintenant $n = 2$ et K de cardinal q, alors comme le cardinal de $\mathbf{P}_K^1$ est $q+1$ (il y a $q+1$ droites dans le plan sur $\mathbf{F}_q$), on obtient un plongement

$$\Phi : \mathrm{PGL}_2(\mathbf{F}_q) \to \mathcal{S}_{q+1}$$

On passe alors en revue tous les cas en calculant les cardinaux avec la proposition 1.14.

a) Si $q = 2$, tous les groupes considérés sont de cardinal 6 car $\mathbf{F}_q^*$ est trivial. En particulier Φ est un isomorphisme de but $\mathcal{S}_3$.

b) Ici $\mathrm{PGL}_2(\mathbf{F}_3)$ est de cardinal 24, donc Φ est un isomorphisme. Comme $\mathrm{PSL}_2(\mathbf{F}_3)$ est de cardinal 12, il est d'indice 2 dans $\mathrm{PGL}_2(\mathbf{F}_3) \simeq \mathcal{S}_4$, donc isomorphe à $\mathcal{A}_4$.

c) D'après la proposition 1.14, on a $\mathrm{PGL}_2(\mathbf{F}_4) = \mathrm{PSL}_2(\mathbf{F}_4)$ et ces groupes sont de cardinal 60, donc d'indice 2 (via Φ) dans $\mathcal{S}_5$, donc isomorphes à $\mathcal{A}_5$.

d) Le groupe $\mathrm{PGL}_2(\mathbf{F}_5)$ est d'ordre 120, il peut être vu via Φ comme un sous-groupe d'indice 6 de $\mathcal{S}_6$, donc il est isomorphe à $\mathcal{S}_5$ d'après un corollaire de la simplicité des groupes alternés vu dans le chapitre "Groupes". Comme $\mathrm{PSL}_2(\mathbf{F}_5)$ est de cardinal 60, il est d'indice 2 dans $\mathcal{S}_5$, donc isomorphe à $\mathcal{A}_5$. □

Remarques : a) On obtient que le groupe $\mathrm{GL}_2(\mathbf{F}_2) = \mathrm{SL}_2(\mathbf{F}_2) \simeq \mathcal{S}_3$ n'est pas parfait. Le groupe $\mathrm{SL}_2(\mathbf{F}_3)$ ne l'est pas non plus car $\mathrm{PSL}_2(\mathbf{F}_3) \simeq \mathcal{A}_4$ a un quotient isomorphe à $\mathbf{Z}/3\mathbf{Z}$ (le quotient par le groupe V_4 constitué de l'identité et des doubles transpositions), donc également $\mathrm{SL}_2(\mathbf{F}_3)$ (en prenant l'image réciproque de V_4 par la surjection canonique). Au passage on voit que $\mathrm{SL}_2(\mathbf{F}_3)$ n'est pas isomorphe à $\mathrm{PGL}_2(\mathbf{F}_3)$, bien que ces deux groupes aient même cardinal (le premier a un quotient d'ordre 3 et pas le second).

b) Le plongement $\Phi : \mathrm{PGL}_2(\mathbf{F}_5) \to \mathcal{S}_6$ donne un sous-groupe d'indice 6 de $\mathcal{S}_6$ qui opère transitivement sur $\{1, ..., 6\}$, donc n'est pas conjugué du stabilisateur d'un point. Cela permet de construire un automorphisme de $\mathcal{S}_6$ qui n'est pas intérieur (ce phénomène ne se produit pour $\mathcal{S}_n$ que quand $n = 6$).

c) Il y a des isomorphismes exceptionnels plus compliqués, par exemple $\mathrm{PSL}_2(\mathbf{F}_7) \simeq \mathrm{PSL}_3(\mathbf{F}_2)$ (l'unique groupe simple d'ordre 168, voir TD...).

2. Formes quadratiques

Dans toute cette section, K est un corps *de caractéristique différente de 2* et E un K-espace vectoriel de dimension finie n.

2.1. Généralités

Définition 2.1 Une *forme bilinéaire* sur E est une application $\varphi : E \times E \to K$ telle que pour tout x_0 de E, les applications $x \mapsto \varphi(x_0, x)$ et $x \mapsto \varphi(x, x_0)$ soient linéaires. La forme φ est dite *symétrique* (resp. *antisymétrique*) si $\varphi(x, y) = \varphi(y, x)$ (resp. $\varphi(x, y) = -\varphi(y, x)$) pour tous x, y de E.

Remarques : a) Comme $\mathrm{Car}\, K \neq 2$, φ antisymétrique est équivalent à φ alternée, i.e. $\varphi(x, x) = 0$ pour tout x de E.

b) Contrairement à un endomorphisme, on peut toujours restreindre une forme bilinéaire à un sous-espace.

[Exercice : toute forme bilinéaire s'écrit de manière unique comme la somme d'une forme symétrique et d'une forme antisymétrique.]

La *matrice* d'une forme bilinéaire φ dans une base $\mathcal{B} = (e_1, ..., e_n)$ est par définition la matrice $(\varphi(e_i, e_j))_{1 \leq i,j \leq n}$. Elle est symétrique (resp. antisymétrique) si et seulement si φ l'est. On a la formule classique de changement de base : si M, M' sont les matrices respectives de φ dans $\mathcal{B}, \mathcal{B}'$, et si P est la matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$, alors :

$$M' = {}^tPMP$$

Ceci justifie la définition suivante :

Définition 2.2 Soit φ une forme bilinéaire, M sa matrice dans une base. On appelle *discriminant* de φ la classe de $\det M$ dans K^*/K^{*2} si $\det M \neq 0$. On convient que le discriminant est nul si $\det M = 0$.

En effet changer de base multiplie $\det M$ par un carré de K^*.

Définition 2.3 On dit qu'une application $q : E \to K$ est une *forme quadratique* s'il existe une forme bilinéaire symétrique φ telle que $q(x) = \varphi(x, x)$ pour tout x de E. La forme φ s'appelle la *forme polaire* de q.

Remarques : a) La forme bilinéaire φ est bien déterminée par q via les formules $\varphi(x, y) = \frac{1}{4}(q(x + y) - q(x - y)) = \frac{1}{2}(q(x + y) - q(x) - q(y))$ (noter l'importance de l'hypothèse $\operatorname{Car} K \neq 2$).

b) Si (a_{ij}) est la matrice de q (i.e. la matrice de φ) dans une base, alors q est donnée par la formule $q(x) = \sum_{1 \leq i,j \leq n} a_{ij} x_i x_j$, où $x_1, ..., x_n$ sont les coordonnées du vecteur x dans cette base. Autrement dit une forme quadratique sur K^n n'est pas autre chose qu'une *fonction polynôme homogène de degré 2*.

c) Si $q : E \to K$ est donnée par $q(x) = \psi(x, x)$, où ψ est une forme bilinéaire quelconque, alors q est encore une forme quadratique car si on écrit $\psi = \psi_1 + \psi_2$ avec ψ_1 symétrique et ψ_2 antisymétrique, alors $q(x) = \psi_1(x, x)$.

Définition 2.4 Soit q une forme quadratique sur E de forme polaire φ. On dit que deux vecteurs x et y de E sont *orthogonaux* (sous-entendu : relativement à q) si $\varphi(x, y) = 0$. L'*orthogonal* d'une partie F de E est le sous-espace vectoriel de E constitué des vecteurs orthogonaux à tout vecteur de F. On le note $F^\perp$.

Définition 2.5 On appelle *noyau* d'une forme quadratique q le sous-espace $E^\perp$ de E. On dit que q est *non dégénérée* si son noyau est réduit à $\{0\}$. On appelle *cône isotrope* [4] de q l'ensemble des x de E tels que $q(x) = 0$.

Remarque : Le cône isotrope contient le noyau, mais la réciproque est fausse; en général le cône isotrope n'est même pas un espace vectoriel. Par exemple la forme $q(x_1, x_2) = x_1^2 - x_2^2$ sur K^2 est non dégénérée, mais son cône isotrope est non trivial.

Proposition 2.6 *Soit q une forme quadratique sur E de forme polaire φ. On a l'équivalence entre :*

i) q est non dégénérée.

ii) L'application linéaire $\Phi : E \to E^*$ qui envoie x sur la forme linéaire $y \mapsto \varphi(x, y)$ est un isomorphisme.

iii) La matrice de q dans une base de E est inversible (i.e. le discriminant de q est non nul).

[4] On appellera également *vecteur isotrope* tout vecteur x *non nul* de E tel que $q(x) = 0$.

Démonstration : i) signifie précisément que Φ est injective, donc i) et ii) sont équivalents puisque $\dim E = \dim E^*$ est finie. ii) équivaut à iii) car si M est la matrice de q dans une base $\mathcal{B}$, alors M est exactement la matrice de Φ dans les bases $\mathcal{B}$, $\mathcal{B}^*$, où $\mathcal{B}^*$ est la base duale de $\mathcal{B}$.

□

La proposition suivante est importante; il faut faire très attention à ses hypothèses (différentes pour les deux parties de l'énoncé) qui sont souvent à l'origine de confusions.

Proposition 2.7 *Soient q une forme quadratique sur E et F un sous-espace de E. Alors :*

i) Si q est non dégénérée, on a $\dim F + \dim F^\perp = \dim E$.

ii) Si la restriction de q à F est non dégénérée, on a $E = F \oplus F^\perp$.

Remarque : q non dégénérée n'implique pas que $q_{|F}$ soit non dégénérée : prendre $q(x_1, x_2) = x_1^2 - x_2^2$ sur K^2, et pour F la droite engendrée par $(1, 1)$. Bien que la conclusion de ii) soit plus forte que celle de i), $q_{|F}$ non dégénérée n'implique pas non plus q non dégénérée : prendre $q(x_1, x_2, x_3) = x_1^2 - x_2^2$ sur K^3, et pour F le sous-espace engendré par les deux premiers vecteurs de la base canonique. Le cas d'une forme définie positive avec $K = \mathbf{R}$ est trompeur, car la propriété "définie positive" se transmet aux sous-espaces, contrairement à la propriété "non dégénérée".

Démonstration : i) L'application $u : E^* \to F^*$, qui associe à une forme linéaire sur E sa restriction à F, est surjective (prendre une base de F et la compléter en une base de E). Comme $\Phi : E \to E^*, x \mapsto (y \mapsto \varphi(x, y))$ est bijective vu que φ est non dégénérée, la composée $\Phi_{F^*} := u \circ \Phi : E \to F^*$ qui associe à tout x de E la forme linaire $y \mapsto \varphi(x, y)$ sur F est surjective. Le noyau de Φ_{F^*} est par définition $F^\perp$. La formule des dimensions donne alors $\dim E = \dim F^\perp + \dim F^* = \dim F^\perp + \dim F$.

ii) $q_{|F}$ non dégénérée signifie $F \cap F^\perp = \{0\}$. Il suffit donc de montrer que $\dim F + \dim F^\perp = \dim E$, ou encore que Φ_{F^*} est surjective (on termine alors comme en i)). On ne sait pas ici que Φ est surjective, mais Φ_{F^*} a pour restriction à F une application linéaire ψ dont le noyau est $F \cap F^\perp = \{0\}$, donc ψ est injective. Comme $\dim F = \dim F^*$ est finie, ψ est aussi surjective et comme c'est la restriction à F de Φ_{F^*}, Φ_{F^*} est a fortiori surjective ce qui termine la preuve.

□

Corollaire 2.8 *Soit q une forme quadratique sur E. Alors q admet une base orthogonale (i.e. une base dans laquelle sa matrice est diagonale).*

Démonstration : On procède par récurrence sur $n = \dim E$. Pour $n = 1$, c'est clair. Supposons le résultat vrai en dimension $< n$. Alors si q est nulle le résultat est évident. Sinon il existe e_1 dans E avec $q(e_1) \neq 0$. La restriction de q à la droite Ke_1 est alors non dégénérée, donc d'après la proposition précédente on a $E = Ke_1 \oplus (Ke_1)^\perp$; par hypothèse de récurrence $(Ke_1)^\perp$ admet une base orthogonale $(e_2, ..., e_n)$ et on obtient la base voulue en prenant $(e_1, e_2, ..., e_n)$.

□

Remarque : Il n'y a pas toujours de base orthonormée, même pour une forme non dégénérée. Par exemple une forme dont le discriminant n'est pas 1 (modulo K^{*^2}) ne peut admettre de base orthonormée.

2.2. Plans hyperboliques

Définition 2.9 *Soit P un plan muni d'une forme quadratique q. On dit que (P, q) est un plan hyperbolique s'il existe une base (e_1, e_2) de P dans laquelle la matrice de q est* $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$.

Remarque : En particulier ceci implique que q est non dégénérée, de discriminant -1. On peut aussi trouver une base dans laquelle la matrice de q est $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ en considérant $(e_1 + \frac{1}{2}e_2, e_1 - \frac{1}{2}e_2)$. Le terme "hyperbolique" vient de ce que si $K = \mathbf{R}$, l'image réciproque par q d'un réel non nul est une hyperbole.

Proposition 2.10 *Soit P un plan muni d'une forme quadratique q. On a équivalence entre :*

i) (P, q) est hyperbolique.

ii) q est non dégénérée et il existe $x \neq 0$ dans P tel que $q(x) = 0$.

iii) Le discriminant de q est -1 (modulo K^{*^2}).

Démonstration : i) implique ii) et iii) : clair à partir de la définition.

Supposons ii) vrai et montrons i). On prend comme premier vecteur de base e_1 tel que $e_1 \neq 0$ et $q(e_1) = 0$. Comme q est non dégénérée, il existe y dans P avec $\varphi(e_1, y) \neq 0$, où φ est la forme polaire de q. Quitte à multiplier y par un scalaire, on peut supposer $\varphi(e_1, y) = 1$. On pose alors $e_2 = \lambda e_1 + y$ avec $\lambda \in K$, alors $\varphi(e_1, e_2) = 1$ et pour avoir $q(e_2) = 0$, il suffit de poser $\lambda = -q(y)/2$. Alors (e_1, e_2) est une base (e_2 n'est pas proportionnel à e_1 car $\varphi(e_1, e_2) \neq 0$) dans laquelle la matrice de q a la forme voulue.

Il ne reste plus qu'à montrer que iii) implique ii). Si iii) est vraie, alors dans une certaine base la matrice de q est $\mathrm{Diag}(a,b)$ (d'après le corollaire 2.8), avec $-ab \in K^{*2}$, donc $-\frac{b}{a}$ est aussi dans K^{*2}. Écrivons $-\frac{b}{a} = \lambda^2$; Alors le vecteur x de coordonnées $(\lambda, 1)$ vérifie $q(x) = 0$, d'où ii).

□

Corollaire 2.11 *Soit q une forme quadratique non dégénérée sur E. On suppose qu'il existe x non nul tel que $q(x) = 0$.*[5] *Alors :*

1. *Il existe un plan P, contenant x, et tel que $(P, q_{|P})$ soit hyperbolique.*
2. *L'image de E par q est K tout entier.*

Démonstration : 1. Soit y dans E tel que $\varphi(x,y) \neq 0$, où φ est la forme polaire de q; alors (x,y) est libre et la matrice dans (x,y) de la restriction de q au plan $P := (x,y)$ est de la forme $\begin{pmatrix} 0 & \varphi(x,y) \\ \varphi(x,y) & * \end{pmatrix}$. Le déterminant de cette matrice est non nul, donc $q_{|P}$ est non dégénérée et admet un vecteur isotrope non nul, à savoir x. D'après la proposition précédente, $(P, q_{|P})$ est hyperbolique.

2. D'après 1., on peut supposer que (E,q) est un plan hyperbolique. Soit alors $\mathcal{B}$ une base dans laquelle la matrice de q est $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. Le résultat résulte de l'expression $q(v) = 2v_1v_2$, où (v_1, v_2) sont les coordonnées d'un vecteur v de E dans la base $\mathcal{B}$.

□

Corollaire 2.12 *Soit q une forme non dégénérée sur E. Alors l'espace E se décompose en somme directe orthogonale $E = (\bigoplus_{i=1}^{r} P_i) \oplus F$, où les (P_i, q) sont des plans hyperboliques, et la restriction de q à F est anisotrope (i.e. son cône isotrope est nul).*

Démonstration : On note que si q est anisotrope, c'est fini. Sinon d'après le corollaire 2.11, il existe un plan P tel que la restriction de q à P soit hyperbolique (en particulier non dégénérée). Alors E est somme directe orthogonale de P et $P^\perp$, et on en déduit le résultat par récurrence sur n en considérant la restriction de q à $P^\perp$, qui est encore non dégénérée puisque E est la somme directe orthogonale de P et $P^\perp$ (donc le discriminant de q est le produit des discriminants de $q_{|P}$ et $q_{|P^\perp}$).

□

[5] On dit parfois que q *représente* 0.

Remarque : On verra plus tard que r ne dépend que de q, ainsi que $q_{|F}$ (à équivalence près, voir la definition 2.13 plus bas pour cette notion). Le problème de la classification se ramène donc aux formes anisotropes, mais ce n'est en réalité souvent pas une bien grande simplification. Voir le livre de Serre pour le cas $K = \mathbf{Q}$.

2.3. Classification des formes quadratiques sur C, R, $\mathbf{F}_q$

On commence par deux définitions.

Définition 2.13 Soit q une forme quadratique sur E. Le *rang* de q est la codimension du noyau de q. C'est aussi le rang de la matrice de q dans une base de E.

Définition 2.14 Soient q, q' deux formes quadratiques sur des K-espaces vectoriels E, E'. On dit que q et q' sont *équivalentes* s'il existe un isomorphisme $u : E \to E'$ tel que $q' = q \circ u$. Il revient au même de dire qu'il existe des bases respectives $\mathcal{B}, \mathcal{B}'$ de E, E' telles que la matrice de q dans $\mathcal{B}$ coïncide avec la matrice de q' dans E'. On note alors $q \sim q'$.

Pour $E = E'$, $\sim$ est une relation d'équivalence sur l'ensemble des formes quadratiques de E, mais la définition ci-dessus nous sera parfois utile dans le cas général.[6] Notons aussi que deux formes quadratiques sur E sont équivalentes ssi leurs matrices respectives M, M' dans une base (fixée) de E sont *congruentes*, i.e. s'il existe une matrice inversible P telle que $M' = {}^tPMP$.

Le problème de classifier les formes quadratiques sur un K-espace vectoriel de dimension finie E est en général très compliqué. On va se borner ici à considérer trois cas particuliers pour K. Notons tout de suite qu'on peut toujours se limiter aux formes *non dégénérées* : en effet si N est le noyau d'une forme q, la restriction de q à un supplémentaire F de N est non dégénérée, ce qui fait que la classification des formes de rang r sur E est ramenée à la classification des formes non dégénérées sur un K-espace vectoriel de dimension r.

Theorème 2.15 *Soit E un K-espace vectoriel de dimension n.*

[6] On ne peut pas à proprement parler de relation d'équivalence si E n'est pas fixé, car les "K-espaces vectoriels munis d'une forme quadratique" ne forment pas un ensemble...

1. *Si $K = \mathbf{C}$ (et plus généralement si tout élément de K est un carré dans K), il n'y a qu'une seule classe de forme quadratique non dégénérée sur E, donnée par la matrice I_n.*

2. *Si $K = \mathbf{R}$, il y a $n+1$ classes de formes quadratiques non dégénérées sur E, correspondant aux matrices-blocs :*

$$M_s := \begin{pmatrix} I_s & 0 \\ 0 & -I_{n-s} \end{pmatrix}$$

avec $0 \leq s \leq n$.

3. *Si $K = \mathbf{F}_m$ est un corps fini (de caractéristique différente de 2), il y a deux classes de formes quadratiques non dégénérées sur E, correspondant aux matrices I_n et*

$$J_n := \begin{pmatrix} I_{n-1} & 0 \\ 0 & \alpha \end{pmatrix}$$

où α est un élément de K^* qui n'est pas un carré dans K^*.

Démonstration : 1. Dans une certaine base $(e_1, ..., e_n)$, la matrice de q est diagonale, de la forme $\mathrm{Diag}(a_1, ..., a_n)$ avec tous les a_i non nuls. Chaque a_i s'écrit $a_i = b_i^2$ avec $b_i \in K^*$. Alors la matrice de q dans la base $(e_1/b_1, ..., e_n/b_n)$ est I_n.

2. De même, comme tout réel non nul est un carré ou l'opposé d'un carré, la matrice de q dans une certaine base est bien de la forme voulue (en multipliant les vecteurs d'une base de diagonalisation par des scalaires, on obtient une base $\mathcal{B}$ dans laquelle la matrice de q est de la forme $\mathrm{Diag}(\varepsilon_1, ..., \varepsilon_n)$ avec les ε_i dans $\{\pm 1\}$; il suffit ensuite de permuter le cas échéant les vecteurs de $\mathcal{B}$).

Il reste à montrer que si $s \neq s'$, les matrices symétriques M_s et $M_{s'}$ ne sont pas congruentes. On remarque que si E_+ désigne l'espace engendré par $(e_1, ..., e_s)$, la restriction de q à E_+ est définie positive. De même sa restriction à l'espace E_- engendré par $(e_{s+1}, ..., e_n)$ est définie négative. Si dans une autre base $(e'_1, ..., e'_n)$, la matrice de q était $M_{s'}$, la restriction de q au sous-espace E'_+ engendré par $(e'_1, ..., e'_{s'})$ serait définie positive, d'où $E_- \cap E'_+ = \{0\}$, ce qui implique $\dim E_- + \dim E'_+ \leq n$, i.e. $(n-s) + s' \leq n$ soit $s' \leq s$. Par symétrie $s' = s$.

3. Comme K est de caractéristique différente de 2, le groupe $K^*/{K^*}^2$ est d'ordre 2 (le noyau du morphisme multiplicatif $K^* \to K^*$, $x \mapsto x^2$ est $\{\pm 1\}$). Soit donc α dans K^* qui n'est pas un carré. Déjà I_n et J_n ne sont

pas congruentes car $\det J_n = \alpha$ n'est pas un carré. Il suffit donc de montrer que pour toute forme quadratique non dégénérée q sur E, il existe une base dans laquelle la matrice de q est I_n ou J_n. On procède par récurrence sur n. Pour $n = 1$, le résultat résulte de ce que tout élément de K^* est égal à 1 ou à α modulo K^{*2}.

Supposons le résultat vrai en dimension $< n$, et supposons $n \geq 2$. Alors il existe $x \in E$ tel que $q(x) = 1$: en effet il suffit de montrer que l'équation $ax_1^2 + bx_2^2 = 1$ a une solution dans K si a, b sont dans K^* (diagonaliser q sous la forme $\mathrm{Diag}(a, b, ...)$, et prendre x de la forme $(x_1, x_2, 0, ...0)$ dans la base correspondante). Or, si on note m le cardinal de K, il y a $(m+1)/2$ carrés dans K (en comptant 0), donc $(m+1)/2$ éléments de la forme ax_1^2, et aussi $(m+1)/2$ éléments de la forme $1 - bx_2^2$, ce qui fait qu'au moins un élément de K (qui est de cardinal $m < (m+1)/2 + (m+1)/2$) s'écrit à la fois ax_1^2 et $1 - bx_2^2$.

Maintenant la restriction de q à Kx est non dégénérée, donc $E = Kx \oplus (Kx)^\perp$, et on conclut en appliquant l'hypothèse de récurrence à $(Kx)^\perp$.

□

Remarque : Le 2. est le classique théorème d'inertie de Sylvester. Si q a pour matrice M_s dans une certaine base, on retrouve s (resp. $n - s$) comme la dimension maximale d'un sous-espace F tel que $q_{|F}$ soit définie positive (resp. définie négative).

2.4. Le théorème de Witt

Définition 2.16 Soit q une forme quadratique sur le K-ev E. Le *groupe des isométries* ou *groupe orthogonal* de q, noté $O(q)$, est le sous-groupe de $\mathrm{GL}(E)$ constitué des automorphismes u qui conservent q, i.e. tels que $q(u(x)) = q(x)$ pour tout x de E.

Notons que pour $u \in \mathrm{GL}(E)$, la propriété $u \in O(q)$ est équivalente à $\varphi(u(x), u(y)) = \varphi(x, y)$ pour tous x, y de E (où φ est la forme polaire de q) via les formules redonnant φ en fonction de q.

Dans le cas d'une forme q définie positive sur un espace vectoriel réel, le groupe $O(q)$ opère transitivement sur les sous-espaces de dimension d fixée (considérer des bases orthonormées de deux tels sous-espaces, et les compléter en des bases orthonormées de E). Pour une forme q (non dégénérée) quelconque, cela n'est pas vrai en général car $q_{|F}$ et $q_{|G}$ n'ont plus de raison d'être équivalentes dès que $\dim F = \dim G$. Le théorème de Witt dit précisément que la condition $q_{|F} \sim q_{|G}$ est suffisante pour qu'il existe $u \in O(q)$ tel que

$u(F) = G$. C'est sans doute le théorème de base le plus important sur les formes quadratiques.

Theorème 2.17 (Witt) *Soit q une forme quadratique non dégénérée sur E. Soient F un sous-espace de E et $s : F \to E$ une application linéaire injective qui conserve q (i.e. telle que $q(s(x)) = q(x)$ pour tout x de F). Alors il existe $u \in O(q)$ dont la restriction à F est s.*

Notons que l'injectivité de s n'est pas automatique, on pourrait par exemple avoir $q_{|F} = 0$.

On déduit immédiatement du théorème de Witt le résultat annoncé plus haut :

Corollaire 2.18 *Soient F et G deux sous-espaces de E tels que $q_{|F} \sim q_{|G}$. Alors il existe $u \in O(q)$ tel que $u(F) = G$.*

Démonstration : Il existe $s_1 : F \to G$ bijective linéaire telle que $q(s_1(x)) = q(x)$ pour tout x de F, et on applique le théorème de Witt à $s : F \to E$ définie par $s(x) = s_1(x)$.

□

Avant de faire la preuve du théorème de Witt, voici d'autres corollaires :

Corollaire 2.19 *Soit q une forme non dégénérée sur E, et F, G deux sous-espaces. Si $q_{|F} \sim q_{|G}$, alors $q_{|F^\perp} \sim q_{|G^\perp}$*

Démonstration : D'après le corollaire 2.18, il existe $u \in O(q)$ telle que $u(F) = G$. Comme u est une isométrie, on a alors $u(F^\perp) \subset G^\perp$ donc $u(F^\perp) = G^\perp$ par égalité des dimensions, et a fortiori $q_{|F^\perp} \sim q_{|G^\perp}$.

□

Corollaire 2.20 (Théorème de simplification) *Soient q, q' deux formes quadratiques non dégénérées sur des K-espaces vectoriels E, E'. On suppose que $E = E_1 \oplus E_2$ et $E' = E'_1 \oplus E'_2$ avec E_1, E_2 (resp. E'_1, E'_2) orthogonaux pour q (resp. q'). On suppose également que $q \sim q'$ et $q_{|E_1} \sim q'_{|E'_1}$. Alors $q_{|E_2} \sim q'_{|E'_2}$.*

Démonstration : Via $q \sim q'$, on se ramène immédiatement au cas où $E = E'$, $q = q'$; le corollaire 2.19 dit alors que $q_{|E_1^\perp} \sim q'_{|(E_1')^\perp}$. Mais $E_1^\perp = E_2$ car $E_2 \subset E_1^\perp$ et comme q est non dégénérée, il y a égalité des dimensions. De même $(E_1')^\perp = E_2'$.

□

On peut maintenant préciser le corollaire 2.12

Corollaire 2.21 *Soit q une forme non dégénérée sur E. Alors l'espace E se décompose en somme directe orthogonale $E = (\bigoplus_{i=1}^r P_i) \oplus F$, où les P_i sont des plans hyperboliques, et la restriction de q à F est anisotrope. De plus r ne dépend que de q, ainsi que $q_{|F}$ à équivalence près.*

Démonstration : L'existence a déjà été vue (corollaire 2.12). L'unicité est une conséquence immédiate du théorème de simplification.

□

Preuve du théorème de Witt. Il y a principalement deux étapes : réduction au cas où F est non dégénéré, et preuve (par récurrence sur $\dim F$) dans ce cas.

a) Réduction à F non dégénéré.

Pour cette réduction, il suffit de montrer le lemme suivant :

Lemme 2.22 *Si F est dégénéré, il existe un sous-espace $F_1 \supset F$, de dimension $\dim F + 1$, et tel que $s : F \to E$ se prolonge en $s_1 : F_1 \to E$ linéaire, injectif, conservant q.*

En effet si on connaît le lemme et le cas non dégénéré, le théorème de Witt en résulte par récurrence sur la codimension de F.

Preuve du lemme : Comme F est dégénéré, il existe x non nul dans $F \cap F^\perp$. Soit φ la forme polaire de q, alors comme q est non dégénérée sur E, il existe $y \in E$ tel que $\varphi(x, y) = 1$, et on peut également supposer $q(y) = 0$, quitte à remplacer y par $y - \frac{q(y)}{2}x$. Posons $F_1 = F \oplus Ky$; on "transporte" tout par s en posant $F' = s(F)$, $x' = s(x)$. Soit l' la forme linéaire définie sur F' par $l'(z) = \varphi(s^{-1}(z), y)$ (ceci a un sens car s, qui est injective, induit un isomorphisme de F sur F'); alors comme q est non dégénérée, il existe y' dans E tel que $l'(z) = \varphi(z, y')$ pour tout z de F' (pour le voir, prolonger l' en une forme linéaire de E). On peut encore supposer $q(y') = 0$ en changeant y' en $y' - \frac{q(y')}{2}x'$ car $q(x') = 0$ et $\varphi(x', y') = l'(x') = \varphi(x, y) = 1$. Alors

l'application linéaire s_1 définie par $s_1(y) = y'$ et $(s_1)_{|F} = s$ convient : en effet pour tout $f \in F$, on a $q(s(f)) = q(f)$, $q(y') = q(y)$, et $\varphi(s(f), y') = \varphi(f, y)$ par construction, ce qui assure que s_1 conserve encore la forme q. D'autre part s_1 reste injectif car $y' \notin F'$ (sinon $\varphi(x', y')$ serait nul puisque $x \in F^\perp$, d'où $x' \in (F')^\perp$).

□

b) Preuve dans le cas F non dégénéré.

On procède par récurrence sur $\dim F$. En réalité, le cas significatif est $\dim F = 1$. Supposons donc $F = Kx$ avec $q(x) \neq 0$. Soit $y = s(x)$, on a donc $q(y) = q(x)$. L'idée est alors de prendre pour u la symétrie orthogonale par rapport à l'hyperplan orthogonal à $(x - y)$, mais il y a une petite difficulté : $x - y$ peut être isotrope. Ceci dit l'un des deux scalaires $q(x - y)$, $q(x + y)$ est non nul, sinon leur somme $2(q(x) + q(y)) = 4q(x)$ le serait. Soit donc $\varepsilon \in \pm 1$ tel que $q(x + \varepsilon y) \neq 0$. Alors E est somme directe de $K(x + \varepsilon y)$ et de son hyperplan orthogonal H. La symétrie v par rapport à H parallèlement à $K(x+\varepsilon y)$ est alors dans $O(q)$, et $v(x) = -\varepsilon y$ car $x - \varepsilon y \in H$ (il est orthogonal à $x + \varepsilon y$) d'où $v(x - \varepsilon y) = x - \varepsilon y$ et $v(x + \varepsilon y) = -x - \varepsilon y$. On prend alors $u = -\varepsilon v$, alors $u \in O(q)$, et u prolonge $s : Kx \to E$.

Le cas $\dim F = 1$ étant fait, supposons $\dim F \geq 2$, et le résultat vrai pour les dimensions $< \dim F$. On écrit alors F comme somme directe orthogonale de F_1 et F_2, avec F_1 et F_2 de dimension $< \dim F$ (c'est possible, par exemple en prenant une base orthogonale de $q_{|F_1}$). En particulier la restriction de q à F_1 et F_2 est encore non dégénérée. L'hypothèse de récurrence permet d'étendre $s_{|F_1}$ en une isométrie $v \in O(q)$, et quitte à composer par v^{-1}, on peut supposer $s_{|F_1} = \mathrm{Id}$. Alors $s(F_2) \subset F_1^\perp$, d'où une application linéire injective $s_{|F_2} : F_2 \to F_1^\perp$, qui par hypothèse de récurrence appliquée à $F_2 \subset F_1^\perp$, s'étend en un automorphisme linéaire s_2 de $F_1^\perp$ conservant q. Il suffit alors de définir u par $u_{|F_1} = \mathrm{Id}$ et $u_{|F_1^\perp} = s_2$, en notant que $E = F_1 \oplus F_1^\perp$.

□

Voici une autre application importante du théorème de Witt.

Définition 2.23 Soit q une forme quadratique non dégénérée sur E. Un *sous-espace totalement isotrope* ("seti" en abrégé) F est un sous-espace de E tel que la restriction de q à F soit nulle. Il est dit maximal si aucun seti ne le contient strictement (on abrègera sous-espace totalement isotrope maximal en "setim").

Theorème 2.24 *1. Tout seti est contenu dans un setim*

2. Si F est un seti et F' un setim, alors $\dim F \leq \dim F'$.

3. *Tous les setim ont la même dimension, notée $\nu(q)$, et appelée indice de q.*

Ainsi q anisotrope équivaut à $\nu(q) = 0$.

Démonstration : 1. est évident (si F est un seti, prendre un seti de dimension maximale qui le contient). 3. découle immédiatement de 2., et il suffit donc de prouver 2. On raisonne par l'absurde. Si $\dim F > \dim F'$, alors on choisit $F_1 \subset F$ avec $\dim F_1 = \dim F'$. Alors les restrictions de q à F_1, F' sont équivalentes (car nulles), donc par le corollaire 2.18 du théorème de Witt, il existe $u \in O(q)$ tel que $u(F_1) = F'$. alors $u(F)$ contient strictement F' et est un seti, ce qui contredit la maximalité de F'.

□

[Exercice : dans la décomposition $E = (\bigoplus_{i=1}^{r} P_i) \oplus F$ avec (P_i, q) hyperbolique et $q_{|F}$ anisotrope, l'entier r est égal à $\nu(q)$; en particulier $\nu(q) \leq n/2$).]

2.5. Quelques résultats sur $O(q)$

On désigne toujours par E un K-espace vectoriel de dimension finie n, muni d'une forme quadratique non dégénérée q. On note $O(q)$ le groupe orthogonal[7] de q et $O^+(q)$ le sous-groupe de $O(q)$ constitué des éléments de déterminant 1. Notons que si $u \in O(q)$, alors $\det u \in \{\pm 1\}$ car si P, M désignent les matrices respectives de u, q dans une base, on a ${}^tPMP = M$ d'où $(\det P)^2 \det M = \det M$.

Si D est une droite *non dégénérée* de E, on notera s_D la réflexion d'hyperplan $D^\perp$, i.e. la symétrie par rapport à $D^\perp$ parallèlement à D. Si P est un plan non dégénéré de E, on notera r_P le *renversement* (ou retournement, ou demi-tour) par rapport à $P^\perp$, i.e. la symétrie par rapport à $P^\perp$ parallèlement à P. Notons que le déterminant d'une réflexion est -1, celui d'un renversement est 1.

On commence par les centres de $O(q)$, $O^+(q)$:

Theorème 2.25 *Supposons $n \geq 3$. Alors*

1. *Le centre de $O(q)$ est $\{\pm \mathrm{Id}\}$.*
2. *Le centre de $O^+(q)$ est $O^+(q) \cap \{\pm \mathrm{Id}\}$, c'est-à-dire $\{\pm \mathrm{Id}\}$ si n est pair, $\{\mathrm{Id}\}$ si n est impair.*

[7] Pour $K = \mathbf{C}$, on fera attention de ne pas confondre ce groupe avec le groupe unitaire d'une forme hermitienne définie positive.

Démonstration : Notons déjà que l'homothétie $\lambda\mathrm{Id}$ est dans $O(q)$ si et seulement si $\lambda \in \{\pm 1\}$. L'idée est d'écrire qu'un élément u du centre de $O(q)$ (resp. $O^+(q)$) commute avec toute réflexion (resp. tout renversement), et d'en déduire que u laisse stable toute droite, donc est une homothétie. Mais par rapport au cas d'une forme définie positive sur un espace vectoriel réel, il y a une difficulté supplémentaire liée au fait que certaines droites peuvent être dégénérées. On y remédie grâce au lemme suivant :

Lemme 2.26 *Pour $n \geq 3$, toute droite D de E est intersection de deux plans non dégénérés.*

Démonstration : Si $D = Kx$ est non dégénérée, on a $E = D \oplus D^\perp$; soit y_1, y_2 deux vecteurs faisant partie d'une base orthogonale de $D^\perp$, alors on a $D = (x, y_1) \cap (x, y_2)$ et les plans (x, y_1), (x, y_2) sont non dégénérés.

Si maintenant D est dégénérée, i.e. $q(x) = 0$, alors on inclut x dans un plan hyperbolique $P_1 = (x, y)$ avec $\varphi(x, y) \neq 0$. Alors P_1 est non dégénéré et $E = P_1 \oplus P_1^\perp$ avec $\dim P_1^\perp \geq 1$. Choisissons alors $z \neq 0$ dans le sous-espace (non dégénéré) $P_1^\perp$, alors $P_2 = (x, y+z)$ est hyperbolique car $\varphi(x, y+z) \neq 0$, et $D = P_1 \cap P_2$ ($y + z$ n'est pas colinéaire à x car $\varphi(x, y + z) \neq 0$, ni à y car y, z sont respectivement dans P_1, $P_1^\perp$).

□

On peut maintenant démontrer les deux assertions du théorème :

1. Si D est une droite non dégénérée, alors un élément u du centre de $O(q)$ commute avec s_D, donc laisse stable D (en effet $us_Du^{-1} = s_{u(D)}$). Comme tout plan P non dégénéré est engendré par deux droites non dégénérées (considérer une base orthogonale de P), P est stable par u. D'après le lemme précédent, toute droite de E est stable par u, donc u est une homothétie.

2. Si P est un plan non dégénéré, alors un élément u du centre de $O^+(q)$ commute avec r_P, donc laisse stable P. Le lemme permet alors encore de conclure.

□

[Exercice : si $n = 2$, alors le centre de $O(q)$ est encore $\{\pm id\}$, sauf si E est un plan hyperbolique sur le corps $\mathbf{Z}/3\mathbf{Z}$, auquel cas $O(q)$ est isomorphe à $\mathbf{Z}/2\mathbf{Z} \times \mathbf{Z}/2\mathbf{Z}$. Pour $n = 2$, $O^+(q)$ est commutatif (distinguer les cas q hyperbolique et q anisotrope).]

Passons à l'engendrement.

Theorème 2.27 *Le groupe $O(q)$ est engendré par les réflexions (i.e. par les s_D, où D est une droite non dégénérée).*

Démonstration : On procède par récurrence sur n. Le résultat est clair si $n = 1$, supposons le vrai en dimension $< n$. Soit $u \in O(q)$. On distingue deux cas :

i) S'il existe x avec $u(x) = x$ et $q(x) \neq 0$, alors E est somme directe de la droite $D = Kx$ et de l'hyperplan $H = (Kx)^\perp$. On applique l'hypothèse de récurrence à H, qui est stable par u. Alors $u_{|H} = s_1...s_m$, où les s_i sont des réflexions de H. En prolongeant les s_i par Id sur D, on obtient des réflexions $s'_1, ..., s'_m$ de E telles que $u = s'_1...s'_m$.

ii) Dans le cas général, soit $x \in E$ tel que $q(x) \neq 0$, posons $y = s(x)$. Alors l'un des deux vecteurs $(x - y)$, $(x + y)$ est non isotrope. Si c'est $x - y$, on compose u avec la réflexion $s_{K(x-y)}$, pour se ramener à i). Si c'est $x + y$, on compose u avec $s_{K(x+y)}$, puis avec s_{Kx} pour se ramener encore à i).

□

Remarque : Le procédé ci-dessus majore le nombre de réflexions par $2n$. En fait la meilleure borne possible est n, mais ce résultat (théorème de Cartan-Dieudonné) est plus subtil.

Theorème 2.28 *Pour $n \geq 3$, $O^+(q)$ est engendré par les renversements.*

Démonstration : D'après le théorème précédent il suffit de voir que la composée $s_1 s_2$ de deux réflexions s'écrit $r_1 r_2$, où r_1 et r_2 sont des renversements. Si $n = 3$, il suffit d'écrire $s_1 s_2 = (-s_1)(-s_2)$. Dans le cas général, soient H_1 et H_2 les hyperplans de s_1, s_2 (qu'on peut supposer distincts). Alors $(H_1 \cap H_2)^\perp$ est de dimension 2. Le sous-espace $(H_1 \cap H_2)$ peut être dégénéré, mais il ne contient pas $(H_1 \cap H_2)^\perp$ car un élément non nul de $(H_1)^\perp$ est dans $(H_1 \cap H_2)^\perp$ et pas dans H_1. En particulier, le noyau de la restriction de q à $H_1 \cap H_2$ est de dimension au plus 1, et $H_1 \cap H_2$ contient un sous-espace V non dégénéré de dimension $n-3$. Alors $E = V \oplus V^\perp$, avec s_1 et s_2 coïncidant avec l'identité sur V. Les restrictions de s_1, s_2 à $V^\perp$ (qui est de dimension 3) sont donc des réflexions. On définit alors des renversements r_1, r_2 de E par $(r_i)_V = \mathrm{Id}$ et $(r_i)_{V^\perp} = (-s_i)_{V^\perp}$ pour $i = 1, 2$. Alors $s_1 s_2 = r_1 r_2$ comme on voulait.

□

Il y a beaucoup d'autres résultats sur $O(q)$ (théorèmes de simplicité, sous-groupes dérivés...), mais contrairement aux théorèmes précédents on n'obtient plus la même chose dans le cas euclidien et dans le cas général (le cas le plus difficile étant celui d'une forme anisotrope sur un corps quelconque). On renvoie au livre de Dieudonné sur les groupes classiques pour plus de détails.