Chapitre 3

Formes Sesquilinéaires

Sommaire

1 Généralités sur les formes sesquilinéaires

Dans ce chapitre, nous allons aborder une notion très intéressante qui est un "crochet" du type produit scalaire sur un espace vectoriel. L'enjeu est en fait double :

- d'une part, on va dégager tout un pan d'une théorie concernant ces "crochets",
- d'autre part, il va en ressortir des groupes assez particuliers qui seront atribués de noms bien spécifiques suivant le "crochet" et qui vont nous enrichir dans notre connaissance sur les groupes (on avait déjà signalé combien on était en fait démuni à ce niveau là...) !

MISE EN GARDE : Les notions qui seront abordées prêtent souvent à confusion sur pas mal de points, car il y aura parfois des identifications entre des espaces, certaines notions se ressemblent... Pour la peine on se rappelle que ce type de difficulté a déjà été rencontré pour la première fois lorsqu'on a rencontré la dualité. Il y a également de type de notion en analyse fonctionelle, avec les topologies normales, fortes et faibles, enfin bref vaut mieux faire attention à tout ceci... !

Etant donné k un corps commutatif et σ un automorphisme de k, pour ne pas surcharger les notations nous allons noter $\lambda^\sigma \overset{\triangle}{:=} \sigma(\lambda)$. Alors on a $(\lambda^\sigma)^\tau = \lambda^{\tau\sigma}$. On a à l'esprit l'exemple le plus classique qui est la conjugaison sur le corps des complexes...

Def. 3.1

Soit E un k-espace vectoriel, on désigne par forme « σ-semi-linéaire » sur E, toute application $g : E \to k$ qui vérifie :

- $g(x+y) = g(x) + g(y)$, c'est à dire " *additivité* ".
- $g(\lambda.x) = \lambda^\sigma.g(x)$.

Rmq

- Si $\sigma = id_k$, on retombe sur la notion de forme linéaire sur E.

Def. 3.2

Soit E un k-espace vectoriel, on désignera par forme « σ-sesquilinéaire » sur E toute application $f : E \times E \to k$ qui vérifie :

1. $\forall y$, l'application $E \to k$, $x \mapsto f(x,y)$ est linéaire (*linéaire par rapport à la* $1^{\text{ère}}$ *variable*).
2. $\forall x$, l'application $E \to k$, $y \mapsto f(x,y)$ est σ-semi-linéaire (σ-*semi-linéaire par rapport à la* $2^{\text{ème}}$ *variable*).

Dans la pratique, lorsqu'il n'y aura pas de confusion possible, on parlera tout simplement de forme sesquilinéaire.

Rmq

- Si $\sigma = id_k$, on retombe sur la notion de forme bilinéaire sur E.

• Nous sommes bien entendu en algèbre linéaire, donc toute notion théorique, a un pendant "pratique" qui se matérialise à travers le calcul matriciel et cela bien entendu lorsqu'on est en dimension finie :

Si E est de dimension finie, munie d'une certaine base $e_1, \cdots e_n$, notre forme σ-sesquilinéaire f est parfaitement déterminée par les n^2 nombres $a_{i,j} \overset{\Delta}{:=} f(e_i, e_j)$ déterminent parfaitement notre forme f : Si $x = \sum_{i=1}^{n} x_i.e_i$ et $y = \sum_{i=1}^{n} y_i.e_i$, alors on obtient la formule

$$f(x,y) = \sum_{i,j} a_{i,j} x_i y_j^{\sigma}.$$

Si on désigne par X et par Y les matrices colonnes de x et de y, on obtient alors la formule :

$$f(x,y) = {}^tXAY^{\sigma}$$

où $A = (a_{i,j}) \in \mathrm{Mat}(n;k)$ est ce que l'on appelle la matrice de f relativement à la base $\mathcal{B} = \{e_i\}$, parfois dans la littérature on l'appelle aussi la « matrice de Gram » de f relativement à la base $\mathcal{B}$.

Le déterminant de A est ce que l'on appelle le « discriminant » de f (ou parfois appelé le déterminant de Gram de f) et se note $\Delta_f(\mathcal{B})$. (Ce nom de "*discriminant*" ne vous évoque rien ? Nous reviendrons si nous y pensons, mais cela renvoie à la notion de "*résultant*" entre deux polynômes ; si jamais l'un des polynôme est la dérivée de l'autre, le résultant correspondant prendra le nom particulier de "*discriminant*").

•• Si P est une matrice de passage, la nouvelle matrice A' de f dans la nouvelle base est reliée par la formule :

$$A' = {}^tPAP^{\sigma}$$

Après une petite réflexion, on verra très bien que cette formule provient de celle qui est présentée ci-dessus (cf. Def.1.13 p.20).

On fera bien attention à ne pas confondre avec la formule classique $P^{-1}AP$ qui est valable lorsqu'on a affaire à une matrice d'une application linéaire, mais on y reviendra un peu plus loin (cf. p.20).

En particulier, on constate que $\det(A') = \delta\delta^\sigma \det(A)$, avec $\delta = \det(P)$.

Le discriminant n'est donc pas comme le déterminant d'une application linéaire un invariant, il n'est défini qu'à un élément $\delta\delta^\sigma$ près. Si $\sigma = \mathrm{id}_k$, il ne sera donc défini qu'à un carré près, i.e. modulo k^{*2}. On en revient encore à ce satané espace k^{*2} des "racines carrées" dans k^* , qui nous était déjà apparu par deux fois dans le chapitre du groupe linéaire,

- dans l'établissement des classes de conjugaison des transvections dans $SL(2;k)$ (cf. Prop.1.5. p.10 du chapitre du groupe linéaire), on verra un peu plus tard que cela nous servira aussi à classifier les classes d'équivalence des formes quadratiques non dégénérées lorsque $k = \mathbb{F}_q^*$.
- Lorsqu'on a voulu établir le Lemme 1.2. p.31 dans le chapitre du groupe linéaire.

La quantité $\delta\delta^\sigma$ sera appelée la « norme » de δ (cf. pour les nombres complexes qui nous donnerait en fait le module au carré). On en verra toute l'utilité de cette notion pour distinguer les classes d'équivalences de formes (cf. Thm.1.6 p.25 sur la classification des formes symétriques sur un corps fini), mais on verra plus exactement dans la preuve du Thm.1.8 p.30 pourquoi on va définir $N(x) = xx^\sigma$.

1.1 Le transfert $f \leftrightarrow \overline{f}$

Pour tout $y \in E$, notons par $f_y(x) \overset{\triangle}{:=} f(x,y)$ ou encore $f_y \overset{\triangle}{:=} f(\bullet, y)$ qui est alors une forme linéaire sur E [On fera attention à ne pas se tromper de variable, d'ailleurs dans tout ce chapitre lorsqu'on rencontra f_y ou même f_x, on s'accordera toujours à bien considérer y ou même x comme étant la deuxième variable]. On obtient ainsi une application "σ-semi-linéaire" :

$$\overline{f} : E \to E^*,\ y \mapsto f_y.$$

1 Il est classique de vouloir passer par $\overline{f}$ pour représenter la forme initiale f, et inversement si on a une telle application σ-semi-linéaire alors on se fabrique facilement une forme σ-sesquilinéaire. On a parfaitement une correspondance, on prendra seulement garde à ce que l'application $E \to E^*$ soit bien σ-semi-linéaire, le caracère "linéaire" qu'on a l'habitude d'employer pour $\overline{f}$ étant seulement réservée que lorsque f est bilinéaire. L'intérêt pour nous est de nous ramener à de l'algèbre linéaire classique au lieu de rester dans l'algèbre bilinéaire et de pouvoir ainsi transférer certaines notions ou propriétés [c'est d'ailleurs une des raisons qui explique l'ambiguïté, les confusions que l'on peut rencontrer dans ce chapitre].

2 On peut également retrouver la matrice de Gram de f dans la base $\{e_i\}$, en considérant seulement la matrice de $\overline{f}$ relativement aux bases $\{e_i\}$ et $\{e_i^*\}$.

Muni de notre nouvelle application $\overline{f}$, on donne cette nouvelle définition :

Def. 3.3

On dira que la forme σ-sesquilinéaire f est « non-dégénérée » si l'application $\overline{f}$ est injective.

Dans le cas contraire, on dira que la forme f est « dégénérée » (ou encore « singulière »).
Par abus de langage, lorsqu'on parlera de *noyau* et de *rang* d'une forme sesquilinéaire f, il s'agira du noyau et du rang de $\overline{f}$.

On prendra garde que cette définition n'est pas du tout "symétrique au niveau des variables", j'y reviendrai un peu plus tard si jamais j'y pense...

Si E est de dimension finie, cette définition devient :

$$f \text{ non-dégénérée} \quad \Leftrightarrow \quad \overline{f} \text{ est un isomorphisme} \quad \Leftrightarrow \quad \mathrm{Ker}(f) = \{0\} \quad \Leftrightarrow \quad \det(A) \neq 0.$$

1.2 Formes réflexives

Etant donné une forme sesquilinéaire f, on dira que deux vecteurs x et y sont «ortogonaux» si $f(x,y) = 0$, ue l'on désignera par la notation $x \perp_f y$ ou plus simplement par $x \perp y$ lorsqu'il n'y a pas de confusion possible.

Comme nous l'avions déjà signalé plus haut, il n'y a pas de raison que cela soit symétrique, ce n'est pas parce que x est orthogonal à y, que y sera orthogonal à x.

Def. 3.4

Nous dirons que la forme sesquilinéaire f est «réflexive», si justement la notion d'orthogonalité correspondante est symétrique, i.e. pour tout $x, y \in E$ on a

$$f(x,y) = 0 \quad \Leftrightarrow \quad f(y,x) = 0.$$

Si E est de dimension 1, alors toute forme est nécessairement réflexive.

Donnons immédiatement d'autres définitions sur les formes sesquilinéaires suivant qu'elles vérifient une autre propriété :

Def. 3.5

Une forme sesquilinéaire f relativement à σ.

- Si $\sigma = \mathrm{id}_k$, i.e. qu'on a affaire à une forme bilinéaire :
  a) On dira que f est «symétrique» si pour tout $x, y \in E$, on a $f(y,x) = f(x,y)$.
  b) On dira que f est «anti-symétrique» si pour tout $x, y \in E$, on a $f(y,x) = -f(x,y)$.
- Si $\sigma \neq \mathrm{id}_k$:
  c) On dira que f est «hermitienne» si pour tout $x, y \in E$, on a $f(y,x) = f(x,y)^\sigma$.

Rmq

1. La première observation triviale que nous pouvons faire immédiatement est que dans ces trois définitions, la forme f est nécessairement réflexive. Chose que l'on peut représenter par le schéma :
$$\{\text{symétrique, anti}-\text{symétrique, hermitienne}\} \Rightarrow \{\text{réflexive}\}.$$
2. On pourrait légitimement englober ces trois définitions à travers une nouvelle définition qui serait la suivante : Etant donné σ et τ deux automorphismes un corps k et soit f une forme σ-sesquilinéaire, on dira en plus que f est τ-symétrique si jamais on a $f(y,x) = f(x,y)^{\tau}$, alors les définitions a), b) et c) correspondent au couple d'automorphismes (id_k, id_k), $(id_k, -id_k)$ et (σ, σ).
Le seul intérêt de leur avoir attribué à chacun un nom spécifique c'est que nous allons être en mesure de les étudier au point même de les classifier... !
3. Tous les éléments de la droite engendrée par une forme symétrique ou anti-symétrique garde la même propriété. Mais c'est loin d'être le cas pour une forme hermitienne, néanmoins cela reste toujours vraie en multipliant par tout scalaire invariant par σ... ! Cette dernière observation (mine de rien) a tout son intérêt, je vais d'ailleurs l'utiliser à la fin de la preuve du Thm.1.1 p.6, mais ça sera aussi l'objet d'un résultat dans la Prop.1.1 p.9 qui suivra.... !

A ces trois définition vient se greffer une autre définition très liée à la définition b) :

Def. 3.6

Soit f une forme σ-sesquilinéaire,

b') on dira qu'elle est «alternée» si pour tout $x \in E$, on a $f(x,x) = 0$.

Alors le lien qui relie cette notion à celle d'une forme anti-symétrique est :

Lem. 3.1

1. Si f est une forme σ-sesquilinéaire alternée, non-nulle, alors on a :
   (i) $\sigma = id_k$,
   (ii) f est anti-symétrique.
2. La réciproque est vraie si le corps k n'est pas de caractéristique 2.

Preuve. 1) Pour tout $x, y \in E$, on a
$$f(x+y, x+y) \underset{\text{alt.}}{=} 0 = \underbrace{f(x,x)}_{=0} + f(x,y) + f(y,x) + \underbrace{f(y,y)}_{=0}$$
d'où $f(y,x) = -f(x,y)$.

Puisque $f \neq 0$, il existe un couple de vecteurs $x, y \in E$ tels que $f(x,y) \neq 0$. Pour tout $\lambda \in k$, on a alors
$$f(\lambda.x, y) = \lambda.f(x,y) = -\lambda.f(y,x)$$
d'après ce que l'on vient de montrer, et réappliquant à nouveau ceci on a $f(\lambda.x, y) = -f(y, \lambda.x)$, or f est σ-sesquilinéaire si bien que l'on a
$$f(\lambda.x, y) = -f(y, \lambda.x) = -\lambda^{\sigma} f(y,x).$$
Si bien que pour tout $\lambda \in k$, on a $\lambda.f(y,x) = \lambda^{\sigma}.f(y,x)$, avec $f(y,x) = -f(x,y) \neq 0$ d'où $\lambda = \lambda^{\sigma}$ et donc on a bien $\sigma = id_k$ et f est bien anti-symétrique.

2) Si $\mathrm{car}(k) \neq 2$ et f une forme anti-symétrique, on a trivialement pour $y = x$, $f(x,x) = -f(x,x)$ d'où $2.f(x,x) = 0$ ❑

• Si E est de dimension finie, et si A est la matrice représentative de la forme f σ-sesquilinéaire, alors les notions de forme symétrique, anti-symétrique ou hermitienne se traduisent maticiellement par $A^t = A$, $A^t = -A$ et $A^t = A^\sigma$.

• Je ne sais pas vous, mais le premier résultat de ce Lemme me fascine assez, en effet l'hypothèse "alternée" d'une forme σ-sesquilinéaire nous permet de déterminer complètement l'automorphisme σ du corps k. Je persiste sur ce point, en effet on obtient également une information en supposant cette fois-ci le caractère "hermitien" : En effet si $f \neq 0$, il existe $x, y \in E$ tel que $f(x,y) \neq 0$, alors comme au dessus pour tout $\lambda \in k$, on a $f(\lambda.x, y) = \lambda.f(x,y)$, mais nous avons aussi $f(\lambda.x,y) \underset{\text{herm.}}{=} f(y,\lambda.x)^\sigma = [\lambda^\sigma.f(y,x)]^\sigma = \lambda^{\sigma^2}.f(y,x)^\sigma \underset{\text{herm.}}{=} \lambda^{\sigma^2}.f(x,y)$, grâce au fait que $f(x,y) \neq 0$, on peut simplifier pour finalement trouver $\lambda = \lambda^{\sigma^2}$, i.e. $\sigma^2 = \mathrm{id}_k$; on obtient alors :

Cor. 3.1

si f est une forme σ-sesquilinéaire hermitienne, alors σ est une involution sur le corps k, i.e. $\sigma^2 = \mathrm{id}_k$.

Nous avons vu un peu plus haut, dans la Rmq.1.2 p.5, que les formes a), b) et c) étaient nécessairement réflexives, ici nous allons voir que, modulo quelques hypothèses supplémentaires, on va avoir la réciproque :

Thm. 3.1 [Formes Réflexives]

Soit f une forme σ-sesquilinéaire sur un k-espace vectoriel E

(i) de dimension finie n, avec $n \geqslant 2$,

(ii) réflexive,

(iii) non-dégénérée.

Alors ,

1 L'automorphisme σ est une involution sur k, i.e. $\sigma^2 = \mathrm{id}_k$.

2 a) Si $\sigma = \mathrm{id}_k$, alors f est une forme bilinéaire symétrique (ou) anti-symétrique.

b) Si $\sigma \neq \mathrm{id}_k$, alors il existe $\alpha \in k^*$ tel que αf soit hermitienne.

Preuve. Puisque f est non-dégénérée , pour tout $x \in E \setminus \{0\}$, la forme linéaire f_x est non nulle, et comme nous somme en en dimenson finie le noyau de f_x est un hyperplan de E.

Comme de plus f est réflexive , on obtient les équivalences suivantes :

$$f(y,x) = \underset{\text{réflexive}}{\Leftrightarrow} f(x,y) = 0 \quad \Leftrightarrow \quad f(x,y)^{\sigma^{-1}} = 0. \tag{1.1}$$

Posons $g_x \overset{\triangle}{:=} f(x,\bullet)^{\sigma^{-1}}$, alors on a

$$g_x(\lambda.y) = f(x,\lambda.y)^{\sigma^{-1}} \underset{\text{sesq.}}{=} [\lambda^\sigma.f(x,y)]^{\sigma^{-1}} = \lambda.f(x,y)^{\sigma^{-1}} = \lambda.g_x(y),$$

i.e. que g_x est une forme linéaire et d'après les équivalences de (1.1) on trouve que $\mathrm{Ker}(f_x) = \mathrm{Ker}(g_x)$, on en déduit que f_x et g_x sont proportionnels ; on en déduit que pour tout $x \in E \setminus \{0\}$, il existe $\alpha(x) \in k^*$ tel que

$$g_x = \alpha(x).f_x \tag{1.2}$$

Rappelons l'application définie au paragraphe précédent $\overline{f} : E \to E^*$ qui est σ-semi-linéaire et en plus un isomorphisme car E est de dimension finie et f non-dégénérée.

Si on considère l'application $\overline{g} : E \to E^*,\ x \mapsto g_x$, on a

$$\begin{aligned} \overline{g}(\lambda.x) &= g_{\lambda.x} = f(\lambda.x,\bullet)^{\sigma^{-1}} = [\lambda.f(x,\bullet)]^{\sigma^{-1}} = \lambda^{\sigma^{-1}}.f(x,\bullet)^{\sigma^{-1}} \\ &= \lambda^{\sigma^{-1}}.\overline{g}(x), \end{aligned}$$

i.e. que $\overline{g}$ est quant à elle σ^{-1}-semi-linéaire.

1. Bien que $\overline{f}$ soit un isomorphisme, pour le moment on ne peut pas se prononcer pour dire que $\overline{g}$ est aussi un isomorphisme, en effet on avait déjà attiré l'attention sur le fait que cette notion n'est nullement "symétrique au niveau des variables" (cf. Attention p.4) : En effet, si $\overline{g}(x) = 0$, cela signifie que $f(x,\bullet)^{\sigma^{-1}} \equiv 0$, cela veut certes dire que $f(x,\bullet) \equiv 0$, mais on n'a pas nécessairement $f(\bullet,x) \equiv 0$, car on n'a pas d'hypothèse faite sur f (du genre sym., anti-sym. ou herm.) pour pouvoir intervertir les variables.
2. De la relation (1.2) ci-dessus, il serait tentant de penser que α est une forme linéaire sur E, mais en fait on n'en sait absolument rien.

Puisque $\overline{f}$ est un isomorphisme σ-semi-linéaire, on en déduit facilement que $\overline{f}^{-1}$ que σ^{-1}-semi-linéaire. Alors $\overline{f}^{-1} \circ \overline{g} : E \to E$ sera quant à elle σ^{-2}-semi-linéaire et on a

$$\begin{aligned} \overline{f}^{-1} \circ \overline{g}(x) &= \overline{f}^{-1}(g_x) \underset{(1.2)}{=} \overline{f}^{-1}(\alpha(x).f_x) \\ &= \alpha(x)^{\sigma^{-1}}.\overline{f}^{-1}(\underbrace{f_x}_{=\overline{f}(x)}) = \alpha(x)^{\sigma^{-1}}.\overline{f}^{-1} \circ \overline{f}(x) \\ &= \alpha(x)^{\sigma^{-1}}.x \end{aligned} \tag{1.3}$$

Alors on a besoin du résultat suivant :

Lem. 3.2

Soit $u : E \to E$ une application τ-semi-linéaire telle que

(i) $\dim(E) \geqslant 2$,

(ii) Pour tout $x \in E$, il existe $\lambda_x \in k$ tel que $u(x) = \lambda_x.x$,

Alors u est une homothétie et on a $\tau = \mathrm{id}_k$.

Preuve. Soit $x, y \in E$, il nous suffit de montrer que l'on a $\lambda_x = \lambda_y$:

• On peut effectivement reprendre la preuve du Lemme 1.2 p.11 du chapitre sur le groupe linéaire : Soient $x, y \in E$ deux vecteurs non-colinéaires qui existent bien car $\dim(E) \geqslant 2$: Nous avons

$$u(x+y) \underset{(ii)}{=} \lambda_{x+y}(x+y).$$

D'autre part, nous avons aussi

$$u(x+y) \underset{\text{add.}}{=} u(x) + u(y) \underset{(ii)}{=} \lambda_x x + \lambda_y y$$

Donc on a $\lambda_{x+y}(x+y) = \lambda_x x + \lambda_y y$, soit $(\lambda_{x+y} - \lambda_x)x = (\lambda_y - \lambda_{x+y})y$; mais x et y ne sont pas colinéaires, si bien que l'égalité précédente ne peut se produire que si $\lambda_{x+y} = \lambda_x = \lambda_y$.

• Si x et y sont colinéaires, on passe par un troisième point, chose qui est réalisable car $\dim(E) \geqslant 2$.

• Donc on peut poser $\lambda_0 = \lambda_x$, toujours à cause de l'hypothèse (ii) on en déduit que $\lambda_0 \neq 0$ et pour tout $\lambda \in k$, on obtient $u(\lambda.x) = \lambda_0.(\lambda.x)$, d'autre part on a :

$$u(\lambda.x) \underset{\tau-\text{semi}-\text{lin.}}{=} \lambda^\tau.u(x) = \lambda^\tau \lambda_0.x,$$

si bien que l'on a $\lambda_0\lambda = \lambda^\tau\lambda_0$, et comme $\lambda_0 \neq 0$ on trouve que $\lambda = \lambda^\tau$, i.e. $\tau = \mathrm{id}_k$. ❑

Rmq

1. Une petite nuance avec le Lemme 1.2. p.11 du chapitre sur le groupe linéaire, ici l'hypothèse (ii) ne stipule nullement que u doit laisser stable toute droite vectorielle et encore moins agirait comme une homothétie sur chaque droite vectorielle.
2. Ce Lemme est faux si $\dim(E) = 1$, en effet si $E = k.e$ alors $u(\lambda.x) = \lambda^\tau.x$ est une forme τ-semi-linéaire qui n'est pourtant pas une homothétie.

Revenons à notre théorème : Puisque $\dim(E) \geqslant 2$ et d'après la relation (1.3), on en déduit qu'il existe $\lambda \neq 0$ tel que

$$\overline{f}^{-1} \circ \overline{g} = \lambda.\mathrm{id}_E, \tag{1.4}$$

et $\sigma^{-2} = \mathrm{id}_k$, soit encore $\sigma^2 = \mathrm{id}_k$ ce qui montre bien que σ est une involution sur k. Comme $\overline{f}$ est σ-semi-linéaire, on obtient alors

$$\overline{g} = \lambda^\sigma.\overline{f}.$$

Donc pour tout $x, y \in E$, on trouve que

$$[\overline{g}(x)](y) \underset{\text{def.}}{=} \underbrace{f(x,y)^{\sigma^{-1}}} = [\lambda^\sigma.\overline{f}(x)](y) \underset{\text{def.}}{=} \underbrace{\lambda^\sigma.f(y,x)}$$

soit encore

$$f(x,y) = \lambda.f(y,x)^\sigma. \tag{1.5}$$

[Remarquons que pour l'instant σ reste de manière abstraite dans la relation précédente, cela ressemble presque à a notion de σ-symétrie que j'ai introduite dans la Rmq.(2) à la page 5, donc quelque part on n'est plus très loin de notre résultat.]

- Si $\sigma = \mathrm{id}_k$, alors f est bilinéaire. Avec le caractère non-dégénérée, on peut trouver $x, y \in E$ tel que $f(x,y) \neq 0$ et on a
$$f(x,y) = \lambda.f(y,x) = \lambda^2.f(x,y).$$
Si bien que $\lambda^2 = 1$ et donc $\lambda = \pm 1$ ce qui prouve bien 2-a).
- Si $\sigma \neq \mathrm{id}_k$, d'après Lemme 1.1 p.5 notre forme f ne saurait être alternée, donc il existe $x_0 \in E$ tel que $f(x_0,x_0) \neq 0$ et la relation (1.5) se transforme alors en
$$f(x_0,x_0) = \lambda.f(x_0,x_0)^\sigma. \tag{1.6}$$
L'idée provient de la Remarque (3) p.5 : En se basant sur la relation (1.5) il faudrait que l'on puisse multiplier f par un scalaire μ tel que $\mu.f$ puisse devenir σ-hermitienne, i.e. doit vérifier la relation
$$[\mu.f(x,y)] = [\mu.f(y,x)]^\sigma$$
Or $[\mu.f(y,x)]^\sigma = \mu^\sigma.f(y,x)^\sigma$ et d'après la relation (1.5), on trouve que $[\mu.f(x,y)] = \mu.\lambda.f(y,x)^\sigma$, si bien que μ devrait vérifier la relation $\mu\lambda = \mu^\sigma$, soit encore
$$\lambda = \frac{\mu^\sigma}{\mu}.$$

D'autre part la relation (1.6) nous donne

$$\lambda = \frac{f(x_0,x_0)}{f(x_0,x_0)^\sigma}.$$

Si bien que l'on obtiendrait la relation

$$\frac{\mu^\sigma}{\mu} = \frac{f(x_0,x_0)}{f(x_0,x_0)^\sigma} \quad \Leftrightarrow \quad \mu^\sigma.f(x_0,x_0)^\sigma = f(x_0,x_0).\mu \quad \Leftrightarrow \quad (\mu.f(x_0,x_0))^\sigma = \mu.f(x_0,x_0).$$

Or une telle "équation" nous dit que l'on doit rechercher parmi les les points fixes de l'automorphismes σ ; on en connait deux en particulier qui sont 0 et 1 (cela ne veut pas dire que ce sont les seuls, mais au moins on dispose déjà de ceux là), on peut aussi exclure 0 et donc il nous faudrait trouver un μ tel que $\mu.f(x_0,x_0) = 1$, on va bien entendu prendre $\mu = f(x_0,x_0)^{-1}$ et l'affaire est dans le sac....

❑

Rmq

1. Si $\sigma \neq id_k$, la forme f n'est pas hermitienne en général : Considérons $E = \mathbb{C}^2$ et σ la conjugaison des nombres complexes et g la forme $x\overline{x'} + y\overline{y'}$ (qui est déjà hermitienne), alors $f = \lambda.g$ est hermitienne (si et seulement si) λ est réel (ceci rejoint bien la Remarque (3) p.5, car les réels sont les points fixes de la conjugaison). On a donc une floppée de formes non hermitiennes.

2. Observons tout d'abord que si on a $F = E \oplus E'$ et qu'au départ on avait une forme σ-sesquilinéaire f définie sur E, on pourra toujours prolonger f sur F de manière nulle sur E'. Si de plus f est symétrique, anti-symétrique ou même hermitienne, alors ce prolongement gardera le même caractère. Maintenant, en imaginant qu'on avait affaire à une forme dégénérée f, alors on pourra toujours étudier la forme induite $\tilde{f}$ sur $E/Ker(f)$ qui sera bien non-dégénérée et notre observation nous permet d'affirmer que l'on obtiendrait la même conclusion du moment que $\dim(E/Ker(f)) \geqslant 2$, i.e. si on n'est pas dans la situation $\dim(E) = \dim(Ker(f)) + 1$. En effet, tout se base esentiellement sur le Lemme 1.2 p.7, en particulier si on est dans ce mauvais contexte, alors on ne pourrait pas non plus obtenir le caractère involutif de σ.

Prop. 3.1 [Les involutions sur un corps]

Soit k un corps et $\sigma \in Aut(k)$ tel que $\sigma^2 = id_k$ et $\sigma \neq id_k$. Alors il existe un sous-corps k_0 de k et un élément $a \in k$ tels que :

1. k est une extension de corps de degré 2 de k_0 (i.e. $[k : k_0] = 2$) et en plus on obtient
$$k = k_0[a] = \{\lambda + \mu.a \quad | \quad \lambda, \mu \in k_0\}$$

2.
- Si $car(k) \neq 2$, alors
  a) l'élément a vérifie l'équation $a^2 = \alpha$ pour un certain $\alpha \in k_0$,
  b) l'involution σ vérifie $\sigma_{|k_0} = id_{k_0}$ et $\sigma(a) = -a$.
- Si $car(k) = 2$, alors
  a) l'élément a vérifie l'équation $a^2 + a = \alpha$ pour un certain $\alpha \in k_0$,
  b) l'involution σ vérifie $\sigma_{|k_0} = id_{k_0}$ et $\sigma(a) = 1 + a$.

Preuve. Définissons

$$k_0 \stackrel{\triangle}{:=} \{x \in k \quad | \quad x^\sigma = x\}.$$

Alors k_0 est bien un sous-corps de k [cf. Théorie de Galois], et comme $\sigma \neq \mathrm{id}_k$, on a bien k_0 est strictement plus petit que k, si bien que l'on peut extirper un élément $a \in k \setminus k_0$.

- Si $\mathrm{car}(k) \neq 2$: L'élément $a - a^\sigma$ vérifie :

$$(a - a^\sigma)^\sigma = a^\sigma - a = -(a - a^\sigma),$$

et est un élément non-nul, anti-invariant. [Remarquons que ça serait également invariant si on avait été en caractéristique 2] C'est en particulier un élément en dehors de k_0. Il y a donc des éléments non-triviaux, anti-invariants dans k. Quitte à considérer ce dernier élément, on peut supposer que $a^\sigma = -a$. Posons alors $\alpha \overset{\Delta}{:=} -aa^\sigma$, cet élément vérifie $\alpha^\sigma = -a^\sigma a = \alpha$, c'est donc un élément non-nul dans k_0, et cette définition de α se traduit encore par

$$\alpha = -a \underbrace{a^\sigma}_{-a} = (-a)(-a) = a^2.$$

Qui n'est rien d'autre que l'équation annoncée que doit vérifier a. [Notons que l'on aurait pu aussi poser $\alpha = aa^\sigma$, cela n'affecte pas le reste de la démonstration, cela affecte seulement l'équation que doit vérifier notre élément a].

D'autre part, on sait bien écrire tout élément x sous la forme :

$$x = \underbrace{\frac{x + x^\sigma}{2}}_{\text{inv.}} + \underbrace{\frac{x - x^\sigma}{2}}_{\text{anti-inv.}}, \tag{1.7}$$

comme la somme d'un élément invariant et d'un élément anti-invariant, possible car $\mathrm{car}(k) \neq 2$ [c'est assez classique, non ?], i.e. d'un élément de k_0 et d'un élément anti-invariant. Mais si b est anti-invariant, alors b/a sera quant à lui invariant, donc tout élément anti-invariant pourra toujours s'écrire sous la forme $\mu.a$ avec $\mu \in k_0$. On a donc montré que

$$k = \{\lambda + \mu.a \quad | \quad \lambda, \mu \in k_0\}.$$

D'autre part $k_0[a]$ est en général un espace plus gros que celui qui est présenté au dessus, mais comme il est contenu dans k, il lui est alors égal et on en a fini.

- Si $\mathrm{car}(k) = 2$: D'après ce que l'on a pu raconter au dessus, en caractéristique 2 les notions d'invariant et d'anti-invariant se confondent, il nous faut donc paser par autre chose. On ne va donc pas rechercher un élément a qui soit anti-invariant, on peut même rajouter en disant que la décomposition (1.7) perd tout son sens pour deux raisons : la division par 2 est impossible, mais d'après ce que l'on vient de dire, invariant et anti-invariant n'a plus de sens ici. Il nous faut développer une autre technique qui puisse se substituer à ceci.

Autre enseignement : On peut toujours écrire par $x = x_1 + x_2$ avec x_1 invariant, mais cela n'implique pas nécessairement que l'on puisse arriver à dire quelque chose sur x_2 même si on est en caractéristique différent de 2 ; dans la situation précédente, si on a pu dire que x_2 était anti-symétrique c'est parce qu'on avait "explicitement" $x_1 = \dfrac{x + x^\sigma}{2}$, si bien que l'on avait également explicitement x_2 et ainsi vérifier s'il pouvait vérifier certaines propriétés, mais si on se cantonne juste à un seule propriété sur x_1 cela n'est pas suffisant et on devine facilement pourquoi car en bougeant par exemple x_1 il n'y a pas de raison que x_2 qui est égal à $x - x_1$ puisse garder une propriété qui nous sauterait aux yeux, ça tiendrait presque à du miracle si cela se faisait.

Pour essayer de coller au plus proche à la technique précédente, on va essayer de fabriquer deux fonctions $f_1(x)$ et $f_2(x)$ telles que :

1. l'on puisse écrire $x = f_1(x) + f_2(x)$,
2. les fonctions f_1 et f_2 vérifient respectivement une certaine propriété "intéressante".

On ne va pas mener simultanément la construction de f_1 et de f_2, ça sera une débauche d'énergie inutile. En effet, il nous suffit de commencer par construire f_1 et de vérifier s'il s'en dégage une certaine propriété, alors par la relation 1. on aura f_2 et on pourra vérifier s'il s'en dégage à son tour une autre propriété. Pour la construction de $f_1(x)$, on reste surtout "modeste" en essayant d'être le plus élémentaire possible : On a à notre disposition x, x^σ, on va alors s'autoriser à faire des opérations les plus élémentaires possibles qui nous sont autorisées à faire dans cette structure qui ici est un corps, donc on va s'autoriser ici à faire $+, -, \times, \div$, et l'inverse.

On ne va pas tout essayer, je vais transcrire maintenant les résultats : Puique $a \in k \setminus k_0$, on a $a^\sigma \neq \pm a$ (j'ai mis $\pm$ car car(k) = 2 et donc $a = -a$), donc $a + a^\sigma \neq 0$, et la fraction $\frac{a}{a+a^\sigma}$ a bien un sens et vérifie

$$\begin{aligned}\left(\frac{a}{a+a^\sigma}\right)^\sigma &= \frac{a^\sigma}{a+a^\sigma} = \frac{a^\sigma + a}{a+a^\sigma} - \frac{a}{a+a^\sigma} = \frac{a^\sigma}{a+a^\sigma} + \frac{a}{a+a^\sigma} \\ &= 1 + \left(\frac{a}{a+a^\sigma}\right).\end{aligned}$$

Donc quitte à remplacer a par ce dernier élément, on peut supposer que a vérifie l'équation $a^\sigma = 1 + a$. Comme tout à l'heure, définissons l'élément $\alpha \overset{\Delta}{:=} a a^\sigma$, c'est bien un élément de k_0 et cette définition se traduit encore par

$$\alpha = a \underbrace{a^\sigma}_{1+a} = a + a^2$$

qui est là aussi l'équation annoncée que doit vérifier notre élément a.

D'autre part, si $x \in k \setminus k_0$, alors comme ci-dessus la fraction $\overline{x} := \frac{x}{x+x^\sigma}$ a bien un sens et vérifie la relation $\overline{x}^\sigma = 1 + \overline{x}$. Et on a

$$(a + \overline{x})^\sigma = a^\sigma + \overline{x}^\sigma = (1+a) + (1+\overline{x}) = a + \overline{x},$$

i.e. que $a + \overline{x} \in k_0$, et donc $\overline{x} \in k_0 + a$, soit encore

$$\frac{x}{x+x^\sigma} \in k_0 + a \quad \Leftrightarrow \quad x \in \underbrace{\underbrace{(x+x^\sigma)}_{\in k_0}.k_0}_{\in k_0} + \underbrace{(x+x^\sigma)}_{\in k_0}.a \quad ,$$

là aussi cela nous dit bien que l'on a effectivement $k = k_0[a]$.

❑

Rmq

Ce résultat nous montre qu'on n'a finalement pas trop de choix au niveau des involutions sur un corps. C'est au moins pratique en ce qui nous concerne, car cela va au moins nous limiter dans les possibilités des formes sesquilinéaires... ! On poussera un peu plus loin cette analyse des involutions lorsque k sera un corps fini (cf. Prop.1.3 p.28).

Signalons enfin, que σ est l'unique involution laissant fixe le corps k_0, c'est en effet une conséquence d'un résultat en théorie de Galois (cf. Cours de maîtrise, Chap. des extensions galoisiennes, p.7 : "*Si* L/K *est une extension finie, alors* $\mathcal{G}(L/K)$ *est fini et son ordre divise* [L : K]". Ici l'extension $[k : k_0]$ valant 2, je vous laisse conclure.

1.3 Notions et objets "géométriques" rattachés aux formes sesquilinéaires

1.3.1 Les SETIM

Jusqu'à présent, nous n'avons présenté que les définitions générales des formes sesquilinéaires, et notamment les formes symétriques, anti-symétriques et hermitiennes (qui à quelques hypothèses près, se résumaient aux formes réflexives, cf. Thm.1.1 p.6), d'ailleurs dans la preuve de ce Thm.1.1, on s'est seulement contenté de l'utilisation de leurs définitions "primaires".

Pour aller plus loin, en repensant à l'algèbre linéaire classique, lorsqu'on a affaire à une application linéaire, il y apparaît des objets géométriques qui lui sont intimement rattachés, comme le noyau, son image, son rang. Le rang permet notamment de classifier les applications linéaires selon une relation d'équivalence près qui est la relation d'être semblable, et en plus ces trois notions sont reliées par le théorème du rang. Pourquoi ne pas adopter cette stratégie pour justement pousser un peu plus loin notre analyse des formes sesquilinéaires. De plus comme cela nous a été suggéré dans la Remarque (2) p.9, autant étudier directement les formes qui sont non-dégénérées.... ! C'est d'ailleurs dans ce cadre là que nous allons tout de suite nous placer. Et puisque nous voulons classifier les

choses, il vaut mieux se placer dans un contexte où dans un premier temps les choses pourraient être plus abordables, plus classifiables comme par exemple en dimension finie.

Hyp. : Dans cette partie, nous allons considérer f une forme sesquilinéaire qui est "**réflexive**".

Grâce à cette hypothèse de "réflexivité", nous avons vu que la notion d'orthogonalité devient bien symétrique au niveau des variables (cf. Attention 1.1 p.4 et 1.2 p.4), on va tout de suite généraliser cette définition aux ensembles :

Def. 3.7

1. Soit A et B deux sous-ensembles de E, on dira que A et B sont « orthogonaux » , si pour tout $x \in A$ et $y \in B$, on a $f(x,y) = 0$. On notera cela par $A \perp_f B$ ou encore $A \perp B$ lorsqu'il n'y a pas de confusion.
2. Si A est un sous-ensemble de E, on désigne « l'orthogonal » de A, le sous-ensemble $A^\perp$ défini par
$$A^\perp \overset{\triangle}{:=} \{x \in E \quad | \quad \forall a \in A,\ a \perp x\},$$
qui est un sous-espace vectoriel de E.

Présentons tout de suite des propriétés "intrinsèquement" rattachées à la deuxième définition :

Prop. 3.2

Soit V et W des sous-espaces vectoriels E. On va faire les hypothèses supplémentaires suivantes :

(i) f est non-dégénérée.

(ii) E est de dimension finie n.

Alors on obtient :

1. Si V est de dimension p, alors $\dim(V^\perp) = n - p$,
2. $V^{\perp\perp} = V$,
3. $(V+W)^\perp = V^\perp \cap W^\perp$, *(résultat valable sans hyp. de dimension sur* E*)*
4. $(V \cap W)^\perp = V^\perp + W^\perp$.

Preuve. Soit $\overline{f} : E \to E^*$ la forme σ-semi-linéaire, définie au §1.1 p.3, puisque $\dim(E) < +\infty$ et f est non-dégénérée f est un "isomorphisme" (au sens de la semi-linéarité et non pas au sens de la linéarité, cf. Rmq.(1) p.3).

(1) Considérons $e_1, \cdots, e_p$ une base de V que l'on complète en une base de E par $e_{p+1}, \cdots, e_n$. On a la base duale $\{e_i\}_{1 \leqslant i \leqslant n}$ de E^*, toute forme linéaire u sur E peut donc sécrire comme $u = \sum_{i=1}^{n} \lambda_i . e_i^*$. Puisque $\overline{f}$ est un isomorphisme semi-linéaire, il nous suffit de voir que $\overline{f}(V^\perp)$ est bien un sous-espace vectoriel de E' de dimension $n-p$. Par définition même, on a

$$V^\perp = \text{Vect}(e_1, \cdots, e_p)^\perp = \{x \in E \quad | \quad \forall v \in V, \quad v \perp x\} = \{x \in E \quad | \quad \forall 1 \leqslant i \leqslant p, \quad e_i \perp x\}.$$

De plus on a,

$$\begin{array}{rcl} \overline{f}(V^\perp) & = & \{u \in E^* \quad | \quad \exists x \in V^\perp,\ u = \overline{f}(x) = f_x = f(\bullet, x)\}, \\ & = & \{f(\bullet, x) \quad | \quad x \in V^\perp\}, \\ & \underbrace{=}_{\text{subtil}} & \{u \in E^* \quad | \quad u_{|V} \equiv 0\}. \end{array}$$

De cette analyse, on en déduit que $u = \sum_{i=1}^{n} \lambda_i . e_i^* \in \overline{f}(V^\perp)$ (si et seulement si) pour tout $1 \leqslant i \leqslant p$ on a $u(e_i) = 0$ (si et seulement si) pour tout $1 \leqslant i \leqslant p$ on a $\lambda_i = 0$, et donc $\{e_i^*\}_{p+1 \leqslant i \leqslant n}$ est une base de $\overline{f}(V^\perp)$ et le résultat est démontré.

(2) On a déjà $V \subset V^{\perp\perp}$, et le résultat découle pour une raison de dimension.

Je laisse le reste.... ❑

L'isomorphisme $\overline{f}$ fait un transfert de la notion d'orthogonalité en celle de forme s'annulant, cela ne s'arrête pas là, en fait la forme sesquilinéaire se retrouve à travers le crochet de dualité entre E et E^*.

Les notions que nous avons données dans Def.1.7 p.12, sont certes intéressantes, mais le seul reproche que je pourrais donner serait que cela n'introduit pas pour autant des objets intrinsèquement liés à la forme f. Nous allons pousser un peu plus loin les choses grâce aux notions que voici :

Def. 3.8

1. Un sous-espace vectoriel V de E sera qualifié de sous-espace « isotrope » (ou encore de « singulier ») si on a une des (équivalences) suivantes :
   (i) La restriction $f_{|V}$ est dégénérée,
   (ii) Il existe $v \in V$ tel que pour tout $x \in V$, on a $f(v, x) = 0$,
   (iii) $V \cap V^\perp \neq \{0\}$.
2. Un sous-espace vectoriel V de E sera qualifié de sous-espace « totalement isotrope » (ou encore appelé « SETI ») si on a une des (équivalences) suivantes :
   (i) La restriction de f à V est nulle,
   (ii) $V \subset V^\perp$.

Rmq

1. Un vecteur x est qualifié de vecteur « isotrope », si le sous-espace vectoriel qu'il engendre est isotrope, c'est donc (équivalent) au fait que $f(x, x) = 0$. (On voit aussi que dans ce cas particulier, $\mathrm{Vect}(x)$ est également totalement isotrope).
2. D'après la Prop.1.2 p.12, si $\dim(E) = n$ et f non-dégénérée, alors $\dim(V^\perp) = n - \dim(V)$, par conséquent si de plus V est totalement isotrope on obtient $V \subset V^\perp$, et en particulier on aura $\dim(V) \leqslant n - \dim(V)$, soit encore

$$\dim(V) \leqslant \frac{n}{2}.$$

Def. 3.9

Etant donné une forme sesquilinéaire, réflexive, non-dégénéré f sur un espace vectoriel E de dimension finie n. On appelle «indice» de f, le maximum des dimensions des sous-espaces totalement isotropes. On le notera par ν_f ; en particulier, on a $\nu_f \leqslant n/2$. Les SETI maximaux sont qualifiés à leur tour les «SETIM» de E.

Cette définition d'indice ne nous garantit pas pour autant que tous les SETIM soient de même dimension, c'est pourtant le cas comme l'affirme e résultat suivant :

Thm. 3.2 [SETIM]

Les SETIM d'un espace vectoriel de dimension finie, sont tous de même dimension.

Preuve. Remarquons que la forme f induit sur $\overline{E} := E/\mathrm{Ker}(f)$ une forme $\tilde{f}$, qui elle est non-dégénérée, reliée par

$$\tilde{f}(\overline{x},\overline{y}) = f(x,y)$$

où $\overline{x},\overline{y} \in \overline{E} = E/\mathrm{Ker}(f)$, le caractère réflexif de f nous permet de bien définir $\tilde{f}$. Notons

$$\pi : E \to \overline{E}$$

la projection canonique. Alors on pourra (sans aucune difficulté) vérifier les quelques propriétées suivantes :

- Soit $\overline{F}$ un sous-espace vectoriel de $\overline{E}$, alors $\overline{F}$ est un SETI de $\overline{E}$ (si et seulement si) $\pi^{-1}(\overline{F})$ est un SETI de E.
- Soit F un sous-espace-vectoriel de E, alors on a les (équivalences) suivantes :
  (i) F est un SETI de E,
  (ii) $F + \mathrm{Ker}(f)$ est un SETI de E,
  (iii) $\pi(F)$ est un SETI de $\overline{E}$.

  On obtient ainsi des caractérisations des SETIM dans chacun des espaces E et $\overline{E}$, à travers une traduction dans l'autre espace :

  - ($\overline{F}$ est un SETIM de $\overline{E}$) $\Leftrightarrow$ ($\pi^{-1}(\overline{F})$ est un SETIM dans E).
  - (F est un SETIM de E) $\Leftrightarrow$ ($\mathrm{Ker}(f) \subset F$ et $\pi(F)$ un SETIM dans $\overline{E}$).

La dernière équivalence nous permet d'avoir la relation suivante :

$$\nu_f = \nu_{\tilde{f}} + \dim(\mathrm{Ker}(f))$$

Hyp. : La relation précédente entre les deux indices va nous permettre, pour l'établissement de notre résultat, de nous limiter à considérer la forme f non-dégénérée.
[Toutes les observations faites ci-dessus vont bien dans le sens de Remarque (2) p.9, c'est sensiblement pour cette même raison que nous voulions seulement considérer que l'étude des formes non-dégénérées.]

Considérons deux SETIM V_1 et V_2 de E, nous allons montrer que V_1 et V_2 sont de même dimension. Il est déjà clair que V_1 et V_2 sont deux SETIM de $V_1 + V_2$. Deux situations se présentent à nous :

- Si $V_1 + V_2 \subsetneq E$: Alors le résultat s'obtient par induction sur la dimension du sous-espace $V_1 + V_2$.

- Si $V_1 + V_2 = E$: Or on a la formule classique

$$\dim(V_1 + V_2) = \dim(V_1) + \dim(V_2) - \dim(V_1 \cap V_2). \tag{1.8}$$

Puisque $\dim(E) < +\infty$ et f non-dégénérée d'après Remarque (2) p.13 on a $\dim(V_i) \leqslant \dim(E)/2$ pour $i = 1, 2$, si bien que la relation (1.8) devient :

$$\dim(V_1 + V_2) = \dim(E) - \dim(V_1 \cap V_2),$$

on en déduit alors que nécessairement on a $V_1 \cap V_2 = \{0\}$, i.e. qu'on a une somme directe entre V_1 et V_2 et qu'en plus $\dim(V_1) = \dim(V_2) = \dim(E)/2$.

❏

1.3.2 Les formes quadratiques

Dans la section précédente, nous avons réussi à dégager des objets intrinsèquement liés à une forme sesquilinéaire qui sont les SETIM. Néanmoins leur manipulation n'est pas du tout pratique, on peut très bien se demander comment est-ce que l'on pourrait en extirper un exemplaire à partir de la donnée de la forme f ?.

On a vu que son noyau était en soit bien gentille, mais est-ce qu'il peut nous être vraiment utile ? La question se pose car, on a vu par exemple que dans les preuves (cf. Thm.1.1 p.6, Thm1.2 p.14 ou encore Remarque (2) p.9) il nous suffisait de nous ramener au cas non-dégénérée... !

C'est presque une question philosophique, mais la plus part du temps on s'emploie à définir des objets à travers des contraintes, des équations, c'est par exemple le cas d'un noyau, alors pourquoi ne pas essayer de comprendre justement l'ensemble des x tel $f(x,x) = 0$, c'est à mon avis l'objet le plus naturel qui se présente à nous après le noyau de la forme. Mais alors, ne serait-ce que cette observation on voit bien qu'il est naturel d'essayer de comprendre la fonction $x \mapsto f(x,x)$, d'où l'introduction de :

Def. 3.10

1. L'ensemble des x tels que $f(x,x) = 0$ s'appelle le « cône isotrope » de la forme sesquilinénaire f et sera noté par $\mathfrak{C}_f$.

2. Si f est une forme sesquilinéaire sur un k-espace vectoriel E, l'application
$$q_f : E \to k, \quad x \mapsto f(x,x).$$
est appelée la « forme quadratique » associée à f.

   Signalons par la même occasion que f est appelée la « forme polaire » de q_f.

Rmq

Cette notion n'a pas d'intérêt pour les formes alternées, car leurs formes quadratiques associées sont triviales, on ne va donc pas s'étendre dans leur étude, on ne s'intéressera seulement qu'aux formes symétiques et hermitiennes lorsqu'on évoquera la notion de formes quadratiques.

Comme le suggère cette dernière remarque, la forme quadratique n'a vraiment d'intérêt que si elle n'est pas nulle, d'où la définition qui suit :

Nous pouvons présenter les premières propriétés élémentaires associées à ces formes quadratiques ; leur établissement ne pose aucun problème :

Cor. 3.2

Soit f une forme σ-sesquilinéaire et q_f sa forme quadratique associée, alors on a :

1. $q_f(\lambda.x) = \lambda\lambda^\sigma.q_f(x)$,
2. $q_f(x+y) = q_f(x) + q_f(y) + f(x,y) + f(y,x)$,
3. $q_f(x-y) = q_f(x) + q_f(y) - f(x,y) - f(y,x)$,
4. $q_f(x+y) - q_f(x-y) = 2(f(x,y) + f(y,x))$.

En joutant 2. et 3., on obtient une formule communément appelée

Identité du parallélogramme

$$q_f(x+y) + q_f(x-y) = 2\left[q_f(x) + q_f(y)\right]$$

Dans le cas symétrique : les formules deviennent

1. $q_f(\lambda.x) = \lambda^2.q_f(x)$,
2. $q_f(x+y) = q_f(x) + q_f(y) + 2.f(x,y)$,
3. $q_f(x-y) = q_f(x) + q_f(y) - 2.fx,y)$,
4. $q_f(x+y) - q_f(x-y) = 4.f(x,y)$.

Nous voyons grâce à ces formules qu'elles peuvent s'inverser du moment où l'on n'est pas en caractéristique 2, c'est bien dans ce cadre là où nous nous placerons lorsqu'on parlera des formes quadratiques, on a alors les formules :

$\mathrm{car}(k) \neq 2$

Symétrique

$$f(x,y) = \frac{1}{2}\left[q_f(x+y) - q_f(x) - q_f(y)\right]$$

$$f(x,y) = \frac{1}{4}\left[q_f(x+y) - q_f(x-y)\right]$$

Dans le cas hermitien : En nous plaçant également en caractéristique différent de 2, avec Prop.1.1 p.9, on a $a^\sigma = -a$. Les formules 2. et 3. de Cor.1.2 deviennent

2'. $q_f(x+y) = q_f(x) + q_f(y) + f(x,y) + f(x,y)^\sigma$,

3'. $q_f(x-y) = q_f(x) + q_f(y) - f(x,y) - f(x,y)^\sigma$.

en particulier on constate que $q_f(x+y) - q_f(x-y) = 2.(f(x,y) + f(x,y)^\sigma)$, soit encore

$$\frac{1}{2}\left[q_f(x+y) - q_f(x-y)\right] = f(x,y) + f(x,y)^\sigma. \tag{1.9}$$

Et on va trouver en plus les relations suivantes :

$$\begin{aligned} q_f(x+a.y) &= q_f(x) + aa^\sigma.q_f(y) + a^\sigma.f(x,y) + a.f(y,x) \\ &= q_f(x) + aa^\sigma.q_f(y) - a.f(x,y) + a.f(x,y)^\sigma, \end{aligned}$$

ainsi que

$$\begin{aligned} q_f(x-a.y) &= q_f(x) + aa^\sigma.q_f(y) - a^\sigma.f(x,y) - a.f(y,x) \\ &= q_f(x) + aa^\sigma.q_f(y) + a.f(x,y) - a.f(x,y)^\sigma. \end{aligned}$$

qui nous donne $q_f(x+a.y)-q_f(x-a.y)=-2a(f(x,y)-f(x,y)^\sigma)$, soit encore

$$-\frac{1}{2a}\Big[q_f(x+a.y)-q_f(x-a.y)\Big]=f(x,y)-f(x,y)^\sigma. \tag{1.10}$$

Maintenant, en ajoutant les relations (1.9) et (1.10), on va trouver que

Hermitienne

$$f(x,y)=\frac{1}{4}\Big[q_f(x+y)-q_f(x-y)\Big]-\frac{1}{4a}\Big[q_f(x+a.y)-q_f(x-a.y)\Big]$$

(cf. [Per] p.124, [Gou] Prop.2. p.230).
Les deux formules ainsi obtenues, qui nous permettent de retrouver la forme polaire est ce que l'on appelle "*l'identité de polarisation*". Dans la suite, grâce à cette formule de transition, les qualificatifs attribués à une forme sesquilinéaire pourra être également employée pour sa forme quadratique, par exemple on pourra parler de *forme quadratique non-dégénée*.

Pour finir j'aimerai juste rajouter une dernière formule qui sera très utile lorsqu'on abordera les espaces préhilbertiens et qu'on appelle aussi par le nom de "*d'identité du parallélogramme*" à juste titre (cf. [Wag] p.394 à 399) : Ici $k=\mathbb{C}$ et σ est la conjugaison des nombres complexes, en particulier on a $f(y,x)=\overline{f(x,y)}$ et la formule que l'on va donner n'est rien d'autre que la retranscription de 2. et de 3. du Cor.1.2 p.16 ci-dessus :

Identité du parallélogramme

$$q_f(x\pm y)=q_f(x)+q_f(y)\pm 2.\underbrace{\mathrm{Re}(f(x,y))}_{\text{partie réelle}}$$

Nous avons vu dans la Remarque 15 précédente que la notion de forme quadratique n'avait aucun intérêt pour une forme alternée, car celle-ci était tout simplement nulle ; mais plus génénéralement, cela n'a pas du tout d'intérêt lorsque qu'on étudie quelque chose de trivial ; néanmoins une forme quadratique s'annulle toujours en 0, si jamais on a affaire à une forme quadratique non-nulle, on peut aller dans le sens inverse et lui demander de ne pas être nulle sur les vecteurs non-nuls, ce qui nous amène à la définition :

Def. 3.11

Etant donné une forme sesquilinéaire f sur E, on dira que f ou encore que q_f est « définie » si $f(x,x)=q_f(x)=0$ entraine $x=0$.

Rmq

Quelques conséquences de ces définitions :

1. Relation tautologique entre le noyau et le cône isotrope d'une forme sesquilinéaire :

$$\underbrace{\mathrm{Ker}(f)}_{\text{noyau}} \quad \subset \quad \underbrace{\mathfrak{C}_f}_{\text{cône isotrope}}$$

2. Lien entre indice, définie et non-dégénérée :

$$\nu_f = 0 \quad \Leftrightarrow \quad f \text{ est définie} \quad \Rightarrow \quad f \text{ est non} - \text{dégénérée.}$$

L'implication de droite provient de l'inclusion précédente, et la première équivalence provient du fait que tout vecteur isotrope se trouve nécessairement dans un SETIM et que ce dernier n'est constitué que de vecteurs isotropes.

2 Quelques classification des formes sesquiliéaires

Le but d'une classification, c'est de rassembler des éléments selon certains critères. On peut rappeler par exemple pour les transvections, que pour $n \geqslant 3$, il n'y avait qu'une seule classe de conjugaison dans $SL(n;k)$ (cf. Prop.1.4. p.9 dans le chapitre du groupe linéaire), tandis que pour $n = 2$, on avait vu qu'il y en avait autant que k^*/k^{*2} ; l'intérêt de passer par une classification est de toujours donner un "représentant" privilégié, et cela dans le but de pousser plus loin l'analyse sur autre chose (cela a été par exemple le cas dans la recherche des sous-groupes distingués de $PSL(E)$ pour ensuite permettre l'établissement de sa simplicité, cf. Thm.1.5. p.26 dans le chapitre du groupe linéaire).

En dimension finie, nous avons rappelé que le fait d'avoir un représentant priviliégié pour une classification donnée était une chose fort intéressante. En dimension finie, cela se faisait toujours à travers un représentant de nature matricielle (cf. pour les transvections) et qui dit matrice dit automatiquement "*base*". C'est sur cet axe que nous allons faire le travail qui va suivre :

Def. 3.12

Etant donné une forme sesquilinéaire f, réflexive sur E. On dira que $e_1, \cdots, e_n$ est une base «orthogonale» pour f si on a :

$$\forall i \neq j \quad \Rightarrow \quad f(e_i, e_j) = 0.$$

(le caractère "réflexif" nous permet de nous passer de la discussion sur la nullité ou non de $f(e_j, e_i)$).

Dans une telle base, la matrice A de notre forme sesquilinéaire, sera diagonale

$$A = \begin{pmatrix} \gamma_1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \gamma_n \end{pmatrix}$$

avec $\gamma_i := f(e_i, e_i)$.

Voici le résultat qui va nous permettre d'arranger pas mal de chose :

Thm. 3.3 [Base adaptée à une forme sesquilinéaire]

Soit f une forme σ-sesquilinéaire sur un k-espace vectoriel E

(i) de dimension finie n,

(ii) $\mathrm{car}(k) \neq 2$,

(iii) f est symétrique (ou) hermitienne.

Alors il existe une base orthogonale pour f. En particulier, on a $\gamma_i = \gamma_i^{\sigma}$.

Preuve. • Tout comme l'observation faite dans la preuve du Thm.1.3 précédent (voir aussi Remarque (2) p.9), on se ramène au cas où f est non-dégénérée, en considérant une base de $\mathrm{Ker}(f) = E^{\perp}$, puis un supplémentaire de ce dernier, alors on est sûr d'obtenir ainsi deux sous-espaces orthogonaux.

• On raisonne par récurrence sur n. Pour $n = 1$, il n'y a rien à faire. Par l'hypothèse (iii) et surtout l'hypothese (ii) nous assure que f n'est pas anti-symétrique (en effet en caractéristique 2, les notions de forme symétique et anti-symétrique se confondent) et donc la forme n'est pas du tout alternée (cf. Lemme 1.1 (2) p.5), et donc il existe un vecteur e non-isotrope, i.e. $f(e,e) \neq 0$.

Considérons $H \overset{\triangle}{:=} \mathrm{Vect}(e)^{\perp}$, Nous allons montrer que $E = \mathrm{Vect}(e) \perp H$. La "projection" de tout vecteur x sur la droite $\mathrm{Vect}(e)$ est le vecteur $f(x,e).e$, alors le vecteur $x - \dfrac{f(x,e)}{f(e,e)}.e$ est orthogonal à $\mathrm{Vect}(e)$: en effet

$$f(x - \frac{f(x,e)}{f(e,e)}.e, e) = f(x,e) - \frac{f(x,e)}{f(e,e)}.f(e,e) = 0$$

Ceci montre que $x - \dfrac{f(x,e)}{f(e,e)}.e \in H = \mathrm{Vect}(e)^{\perp}$ et on a bien

$$x = \underbrace{\left(x - \frac{f(x,e)}{f(e,e)}.e\right)}_{\in H} + \underbrace{\frac{f(x,e)}{f(e,e)}.e}_{\in \mathrm{Vect}(e)},$$

ceci montre que l'on a $E = \mathrm{Vect}(e) \oplus H$ et c'est même une somme directe car on a bien choisi e non-isotrope

[Signalons à titre d'information que la décomposition obtenue ci-dessus $E = \mathrm{Vect}(e) \oplus \mathrm{Vect}(e)^{\perp}$ ne demande pas que l'on soit en dimension finie, mais seulement d'avoir une forme non-alternée].

Alors par hypothèse de récurrence on va pouvoir trouver une base orthogonale dans H que l'on joint au vecteur e pour avoir une base de E. L'hypothèse $\dim(E) < +\infty$ nous permet seulement de nous assurer du fonctionnement du raisonnement par récurrence. ❑

J'ai pris le soin d'avoir encadré le mot "*orthognale*" pour insister là dessus, car par racouci on aurait envie de remplacer ce mot par le terme "*orthonormale*", cela serait tout à fait inexact, on verra par exemple dans les preuves de Thm.1.4 p.22, Thm.1.5 p.23, Thm.1.6, Thm.1.7 p.27 ou encore Thm.1.8 p.30, que la construction d'une base orthonormale partira toujours d'une base orthogonale (en l'occurrence celle qu'on a dans le Thm.1.3 ci-dessus) et tout se focalisera ensuite sur les coefficients qui apparaissent sur cette diagonale, on va les "*normaliser*", en fait la discussion portera à chaque fois sur le fait de savoir si chacun de ces coefficients est "*la norme*" ou non d'un autre scalaire (c'est surtout dans la preuve de Thm.1.8 p.30 que ma remarque sera criante de vérité, voir aussi Rmq. p.23). Donc le passage

orthogonale $\dashrightarrow$ **orthonormale**

pourra essentiellement se faire par la discussion de la "*norme*" si elle existe sur le corps de base k avec lequel on évolue.

Venons-en à la relation d'équivalence entre formes sesquilinéaires

Un petit rappel d'algèbre linéaire ne fera pas de mal : Etant donné deux endomorphismes f et g sur E, on dit que f et g sont "*semblables*" s'il existe existe un automorphisme $u \in GL(E)$ tel que $g = u \circ f \circ u^{-1}$; d'un point de vue matriciel, cela signifie qu'il existe deux bases $\mathcal{B}$ et $\mathcal{B}'$ tel que $Mat_{\mathcal{B}}(f) = Mat_{\mathcal{B}'}(g)$, chose que l'on peut encore dire de la manière suivante : si on exprime les matrices de f et de g dans une même base $\mathcal{B}$, alors il va exister une matrice inversible P telle que $Mat_{\mathcal{B}}(g) = PMat_{\mathcal{B}}(f)P^{-1}$. La matrice P est la matrice de u qui va transformer la base $\mathcal{B}$ en la base $\mathcal{B}'$.

En pensant à ce rappel, il sera alors logique de dire que deux formes f et g sont équivalentes si : soit A la matrice de f dans une base $\mathcal{B}$ (cf. page 2) il devrait exister un automorphisme $u \in GL(E)$ tel que la matrice représentative de g dans la base $u(\mathcal{B})$ soit également A.

Def. 3.13

Etant donné deux formes σ-sesquilinéaires f et g sur E. On dira que f et g sont «équivalentes» , s'il existe $u \in GL(E)$ tel que pour tout $x, y \in E$, on a

$$g(x,y) = f(u(x), u(y)).$$

(cf. p.3, d'un point de vue matriciel, on trouvera $A_g = {}^tPA_fP^{\sigma}$).

Rmq

De cette définition, si f et g sont équivalentes, alors il est facile de voir que l'on a :

1. Comme nous l'avions déjà signalé à la suite des identités de polarisation, il convenait d'attribuer les mêmes qualificatifs à une forme sesquilinéaire et sa forme quadratique, ici nous pouvons également dire que deux formes quadratiques q_f et q_g sont « équivalentes » s'il existe $u \in GL(E)$ tel que $q_g(x) = q_f(u(x))$ et grâce aux formules de polarisation on voit clairement qu'on a bien fait de souligner ce point là puisque cette définition est bien (équivalente) au fait que les formes f et g sont équivalentes. De cette observation, on voit très bien que la classification des formes sesquilinéaires revient à celle des formes quadratiques (pour les formes symétriques et des formes hermitiennes en caractéristique différente de 2).

2. Quelques invariants numériques :
   (i) $\text{rang}(f) = \text{rang}(g)$, elles ont même rang,
   (ii) $\nu_f = \nu_g$, elles ont même indice.

3. Rappelons le lien qu'il y a entre le discriminant de f et de g (voir p.1) :

$$\Delta_g = \det(u)\det(u)^{\sigma}.\Delta_f,$$

si bien que le discriminant n'est pas un véritable invariant pour la classe d'équivalence ; les discriminants de diffèrentes formes équivalentes diffèrent d'une "*norme*" près, c'est bien là une première différence majeure avec la relation d'équivalence entre applications linéaires. Néanmoins Si $\sigma = \text{id}_k$, on constate que le discriminant apparaît alors comme un élément bien spécifique de k^*/k^{*2}, chose que l'on peut encore de manière équivalente exprimer (en référence aux classes d'équivalence des transvections dans $SL(2,;k)$) par le fait que le quotient des discriminants est un carré. [On ne pouvait pas tellement faire cela lorsque $\sigma \neq \text{id}_k$, car on ne sait pas quel type de structure pourrait avoir "*l'espace des normes*", tandis que k^{*2} est l'espace des racines carrées. On pourrait (je pense) que si on avait explicitement σ... !

Rmq

Enfin, j'aimerai également rajouter, une deuxième différence notable : En revenant à la relation d'équivalence entre application linéaire, on sait que le rang suffisait à caractériser la classe d'équivalence, alors qu'ici cela n'est pas le cas, en effet si on considère sur $\mathbb{R}^2$ avec $\sigma = \text{id}_{\mathbb{R}}$, et les formes $f((x_1,y_1),(x_2,y_2)) = x_1x_2 + y_1y_2$ et $g((x_1,y_1),(x_2,y_2)) = -x_1x_2 - y_1y_2$; il est facile de voir qu'elles ont même rang, même indice et leurs matrices de Gram dans la base canonique sont $A_f = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ et $A_g = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$ tous les deux de discriminant 1, donc vérifient bien la contrainte d'avoir leur quotient égal à un carré. Et pourtant on va voir qu'elles ne sont pas équivalentes.

Présentons à présent la première classification :

2.1 Le cas symétrique

Thm. 3.4 [Formes symétriques sur un corps algébriquement clos]

Soit E un k-espace vectoriel de dimension finie n sur un corps k algébriquement clos avec $\mathrm{car}(k) \neq 2$.

Alors Il n'y a qu'une unique classe d'équivalence parmi les formes quadratiques, symétrique, non-dégénérées.

- Dans une base convenable sa matrice sera l'identité et si on exprime tout vecteur dans une telle base $x = (x_1, \cdots, x_n)$ on aura :
$$q(x) = x_1^2 + \cdots + x_n^2.$$
- De plus, l'indice est égal à $[n/2]$ (*partie entière*).

Preuve. Soit q une forme quadratique non-dégénérée, c'est en particulier non-alternée et comme $\dim(E) < +\infty$ et $\mathrm{car}(k) \neq 2$, on sait l'existence d'une base orthogonale (cf. Thm.1.3 p.19). Soit donc $e_1, \cdots e_n$ une telle base, et $\gamma_i := q(e_i)$, on a $\gamma_i = \gamma_i^\sigma$ (cf. Thm.1.3) (mais cela ne pose pas de problème pour ici car $\sigma = \mathrm{id}_k$) et en plus on a $\gamma_i \neq 0$ car q est non-dégénérée.

- Soit $x \in E$, on peut écrire $x = \sum_{i=1}^{n} x_i.e_i$, alors grâce au Cor.1.2 p.16 on obtient
$$q(x) = \sum_{i=1}^{n} \gamma_i.x_i x_i^\sigma = \sum_{i=1}^{n} \gamma_i.x_i^2.$$

Puisque k est algébriquement clos, on peut écrire $\gamma_i = \delta_i^2$, posons alors $x_i' \overset{\Delta}{=} \delta_i x_i$, on a alors
$$q(x) = \sum_{i=1}^{n} x_i'^2.$$

Dans la base $\mathcal{B} = \{b_1 := \delta_i^{-1} e_1, \cdots, b_n := \delta_n^{-1} e_n\}$, la matrice de q est la matrice identié I_n. Cette nouvelle base est mieux qu'une base orthogonale puisque c'est une base « orthonormale ».
Notons que pour avoir une telle base orthonormale, il a bien fallu que q soit non-dégénérée
Dans cette construction, on a bien l'impression d'avoir défini $\mathcal{B}$ à partir du vecteur x, mais il n'en est rien, et relativement à cette base on a effectivement $q(y) = \sum_{i=1}^{n} y_i^2$ si jamais $y = \sum_{i=1}^{n} y_i.b_i$.

- Les hypothèses $\dim(E) < +\infty$ et q non-dégénérée nous permettent de dire que l'indice ν n'excède pas $n/2$ (cf. Remarque (2) p.13), et donc on a encore
$$\nu \leqslant [n/2].$$

Enfin, dénotons par $\mathbf{i}$ une racine de -1 qui existe bien puisque k est algébriquement clos : Considérons la famille de vecteurs
$$b_1 + \mathbf{i}b_2, b_3 + \mathbf{i}b_4, \cdots, b_{[n/2]} + \mathbf{i}b_{[n/2]+1}.$$
Faisons alors les observations suivantes :

1. On a une famille de $[n/2]$ vecteurs linéairement indépendants,
2. Le sous-espace engendré par ces vecteurs est totalement isotrope : Désignons par f la forme polaire de q, alors on a
$$f(b_k + \mathbf{i}b_{k+1}, b_l + \mathbf{i}b_{l+1}) = \underbrace{f(b_k, b_l)}_{0 \text{ ou } 1} + \mathbf{i}\underbrace{f(b_k, b_{l+1})}_{0 \text{ ou } 1} + \mathbf{i}\underbrace{f(b_{k+1}, b_l)}_{0 \text{ ou } 1} + \mathbf{i}^2 \underbrace{f(b_{k+1}, b_{l+1})}_{0 \text{ ou } 1}.$$
Cela nous amène à des discussions : Si $k = l$ on aura $f(b_k, b_l) = 1$, mais on aura aussi $f(b_{k+1}, b_{l+1}) = 1$ et puisque $\mathbf{i}^2 = -1$ le premier et le dernier terme se compensent pour donner 0, et on aura aussi $f(b_k, b_{l+1}) =$

$f(b_{k+1}, b_l) = 0$. On fera exactement la même discussion pour les autres situations. En fin de compte on va bien trouver que ceci vaut 0.

3. Avec les points 1. et 2. précédents, on a montré en fin de compte que $\nu = [n/2]$.
Mais il est aussi légitime de se demander pourquoi ne pas avoir rajouté les autres vecteurs : $e_2 + \mathbf{i}e_3, \cdots$, i.e. les vecteurs de la forme $e_{2k} + \mathbf{i}e_{2k+1}$ avec $k \geqslant 1$.
En effet il nous suffit de considérer le vecteur $e_{2k-1} + \mathbf{i}e_{2k}$ déjà comptabilisé, on a alors

$$f(e_{2k-1} + \mathbf{i}e_{2k}, e_{2k} + \mathbf{i}e_{2k+1}) = \underbrace{f(e_{2k-1}, e_{2k})}_{0} + \mathbf{i}\underbrace{f(e_{2k-1}, e_{2k+1})}_{0} + \mathbf{i}\underbrace{f(e_{2k}, e_{2k})}_{1} + \mathbf{i}\underbrace{f(e_{2k}, e_{2k+1})}_{0}.$$

Cela n'étant pas égal à 0, et donc on n'aura plus une famille totalement isotrope.

❏

Rmq

Je voudrai revenir sur l'unicité, car une étant donné deux bases orthogonales pour deux formes (dégénérées ou pas), il existe toujours $u \in GL(E)$ qui va envoyer l'une sur l'autre, on a alors l'impression que les deux formes vont être équivalentes. C'est un racourci qui est légitime, mais faux pour la raison suivante : Soit $\{e_i\}$ (resp. $\{f_i\}$) une base orthogonale pour q_1 (resp. pour q_2) comme nous donne le Thm.1.3 p.19, i.e. on aura $q_1(e_i) = \zeta_i$ et $q_2(f_i) = \xi_i$, alors avec la donnée de u qui enverrait e_i sur f_i il n'est pas garanti qu'on puisse avoir $q_1 = q_2 \circ u$. De toute façon, pour que cela puisse marcher il est déjà impératif d'avoir le même rang (cf. Remarque (2) p.21), ensuite s'ils ont le même rang, alors comme dans la preuve ci-dessus, on va pouvoir normaliser les bases respectives afin d'avoir $q_1(\bullet) = 0, 1$ (on normalise bien entendu que les vecteurs e_i pour lesquels $\zeta_i \neq 0$), on fait la même chose pour q_2 avec sa base orthogonale, alors une fois que cela a été fait, on pourra effectivement définir u qui pourra envoyer certains vecteurs sur ceux que l'on souhaite pour que la partie puisse être remportée. Donc on voit qu'il y a exactement $(n+1)$ classes d'équivalence de formes quadratiques symétrique sur E.

Notons que dans cette preuve, on utilise que parcimonie le caractère "*algébriquement clos*" du corps k, ce résultat est encore valable pour un corps sur lequel tout élément est un carré (cf. Attention p.20).

Thm. 3.5 [Formes symétriques sur le corps $\mathbb{R}$ - Loi d'inertie de Sylvester]

Soit E un $\mathbb{R}$-espace vectoriel de dimension finie n.

Alors il y a exactement $(n+1)$ classes de formes quadratiques symétriques, non-dégégérées.

• Dans une base convenable, la matrice de chacune de ces classes sera de la forme

$$\begin{pmatrix} 1 & & & & \\ & \ddots & & 0 & \\ & & 1 & & \\ & & & -1 & \\ & 0 & & & \ddots \\ & & & & & -1 \end{pmatrix}$$

où 1 y apparaît p fois et -1 y apparaît $(n-p)$ fois (avec p parfaitement déterminé), et si on on exprime tout vecteur dans une telle base $x = (x_1, \cdots, x_n)$ on aura :

$$q(x) = (x_1^2 + \cdots x_p^2) - (x_{p+1}^2 + \cdots + x_n^2).$$

• De plus, l'indice est égal à $\min(p, n-p)$.

Preuve. On reprend le début de la preuve précédente avec cette fois-ci $\gamma_i \in \mathbb{R}\setminus\{0\}$; on peut supposer que les p premiers nombres $\gamma_i > 0$ et les autres < 0. Considérons alors α_i les racines carrées de γ_i des p premiers et les racines carrées de $-\gamma_i$ pour les derniers.
Comme tout à l'heure on définit la base $\mathcal{B} = \{b_1 := \alpha_1^{-1}e_1, \cdots, \alpha_n^{-1}e_n\}$ et dans cette base on aura la matrice annoncée.

• De plus, pour une telle base on peut décomposer $E = E^+ \oplus E^-$ où E^+ (resp. E^-) est le sous-espace vectoriel engendré par les p premiers (resp. les derniers) vecteurs. Sur E^+ (resp. E^-) q sera définie positive (resp. négative). Si maintenant on a une autre base adaptée comme ci-dessus, on pourra de même écrire $E = F^+ \oplus F^-$. Alors il est facile de voir que l'on a la somme directe $E^+ \oplus F^-$ et donc ce dernier est de dimension $p + (n-p') \leqslant n$, ce qui montre que $p \leqslant p'$, on refait la même chose dans l'autre sens et pour finalement montrer que $p = p'$.

• L'ensemble $\{1, \cdots, n\}$ se découpe ainsi en 2 paquets ; le nombre total de découpes possibles est égal à $(n+1)$ et un argument comme dans la Remarque p.23 nous permettrait de voir que l'on a exactement $(n+1)$ classes d'équivalence de formes quadratiques symétriques, non-dégénérée sur un $\mathbb{R}$-espace vectoriel de dimension n.

• Un peu comme tout à l'heure, on considère la famille de vecteurs $b_1 + b_{p+1}, b_2 + b_{p+2}, \cdots$, on prend la somme entre un élément sur qui q vaut 1 et un élément sur qui q vaut -1, on continue ainsi jusqu'à l'épuisement du stock, c'est à dire qu'on aura formé en tout $\min(p, n-p)$ vecteurs. On vérifiera bien comme tout à l'heure que le sous-espace vectoriel F engendré par ces vecteurs est totalement isotrope, ce qui montrera que l'indice est plus grand que $\min(p, n-p)$. Montrons l'égalité :
D'après ce qui précède, il existe un base $\{b_i\}$ dans laquelle la matrice de notre forme f aura l'aspect suivant :

$$\begin{pmatrix} I_p & 0 \\ 0 & I_{n-p} \end{pmatrix}$$

Considérons G un SETI contenant F. Soit $x \in G$, on va écrire $x = \sum_i \lambda_i . b_i$, alors on a

$$0 \underbrace{=}_{G \text{ SETI}} q(x) = f(x,x) = \sum_{i=1}^{p} \lambda_i - \sum_{j=p+1}^{n} \lambda_j. \tag{2.1}$$

D'autre part, pour tout $1 \leqslant k \leqslant \min(p, n-p)$, on a $x \perp (b_k + b_{p+k})$ car G est un SETI qui contient $b_k + b_{p+k}$, ce qui s'exprime par :

$$\begin{aligned} f(x, b_k + b_{p+k}) &= 0 = \underbrace{=}_{\text{identité paral.}} \frac{1}{2}\Big(\underbrace{q(x + b_k + b_{p+k}) - q(x)}_{[(\lambda_k+1)^2 - (\lambda_{p+k}+1)^2] - [\lambda_k^2 - \lambda_{p+k}^2]} - \underbrace{q(b_k + b_{p+k})}_{=0 \text{ car } \in F} \\ &= \frac{1}{2}\big(2\lambda_k - 2\lambda_{p+k}\big) = \lambda_k - \lambda_{p+k}. \end{aligned}$$

Ceci montre que pour tout $1 \leqslant k \leqslant \min(p, n-p)$, on a $\lambda_k = \lambda_{p+k}$, de plus en tenant compte de (2.1) on a soit $\lambda_k = 0$ pour $\min(p, n-p) + 1 \leqslant k \leqslant p$ si jamais $\min(p, n-p) = n-p$, soit pour $p + \min(p, n-p) + 1 \leqslant k \leqslant n$ si jamais $\min(p, n-p) = p$ (j'avoue que c'est un peu rapide ici mais je le laisse en réflexion), de cette analyse on en déduit que

$$x = \sum_{i=1}^{\min(p, n-p)} \lambda_i . (b_i + b_{p+i}),$$

ce qui montre que l'on a $x \in F$, i.e. que F est bien un SETIM et que l'indice $\nu_q = \min(p, n-p)$.

❏

Le couple d'entiers p et $n-p$ apparu dans ce résultat, porte un nom bien particulier :

Def. 3.14

Avec les notations du Thm.1.5, le couple d'entiers $(p, n-p)$ s'appelle la « signature » de la forme q.

Rmq

1. Comme dans la Remarque 23, il est facile de voir qu'en considérant cette fois-ci les formes dégénérées, on pourra généraliser la signature comme étant un triplet de la forme (p,q,r) avec $p+q+r=n$, alors le nombre de classes d'équivalence sera égal à $(n+1)+n+(n-1)+\cdots+1=\frac{(n+1)(n+2)}{2}$.
2. Si $p=n$, on aura $q(x)>0$ pour tout $x\neq 0$, une telle forme sera dite «définie positive» . En particulier, dans une base adaptée comme dans le Thm.1.5 précédent, la matrice de q ne sera rien d'autre que l'identité.

Thm. 3.6 [Formes symétriques sur un corps fini]

Soit E un $\mathbb{F}_q$-espace vectoriel de dimension finie n, où $\mathbb{F}_q$ n'est pas de caractéristique 2. Soit $\alpha\in\mathbb{F}_q^*\setminus\mathbb{F}_q^{*2}$.

Alors il y a exactement deux classes d'équivalence de formes quadratiques symétriques, non-dégénérées sur E, de matrices représentatrices :

$$q_1=\begin{pmatrix}1&&0\\&\ddots&\\&&\ddots\\0&&1\end{pmatrix}\ (\text{ou})\ q_2=\begin{pmatrix}1&&0\\&\ddots&\\&&1\\0&&\alpha\end{pmatrix}.$$

Une forme q est dans l'une ou l'autre classe suivant que son discriminant Δ_q est (ou) non, un carré de $\mathbb{F}_q^*$.

Preuve. Soit q une forme quadratique, symétrique, non-dégénérée. Par le Thm.1.3, on peut lui trouver une base orthogonale $e_1,\cdots,e_n$.

- Pour $n=2$: Matriciellement la matrice correspondante sera de la forme $Q=\begin{pmatrix}a&0\\0&b\end{pmatrix}$. Si X représente la matrice colonne du vecteur (x,y), en utilisant la formule $q(x,y)={}^tXQX$ (cf. p.2), on obtient alors

$$q(x,y)=ax^2+by^2.$$

Alors on va pouvoir trouver un vecteur $b_1=(x,y)$ pour lequel on a $q(b_1)=1$, l'existence de b_1 provient du Lemme suivant :

Lem. 3.3

L'équation ax^2+by^2 avec $a,b\in\mathbb{F}_q\setminus\{0\}$ a des solutions en x,y dans $\mathbb{F}_q$.

Preuve. Par hypothèse, $b\neq 0$, on peut alors définir l'application

$$\rho:\mathbb{F}_q\to\mathbb{F}_q,\ x\mapsto\frac{1-ax^2}{b}.$$

D'autre part, on a

$$\rho(x_1)=\rho(x_2)\quad\Leftrightarrow\quad x_1^2=x_2^2.$$

Cela signifie en paticulier que si on se limite aux "*carrés*", ρ devient injective. Or on pourra voir [Per] Prop.2.10.(2) p.74 que l'on a $\text{card}(\mathbb{F}_q^2)=\frac{q+1}{2}$ pour $q\neq 2$. Si bien que ρ va prendre au moins $\frac{q+1}{2}$ valeurs possibles. Si on considère les autres éléments dans $\mathbb{F}_q$ qui ne sont pas des carrés, on constate qu'ils sont au nombre de $\frac{q-1}{2}$. Par conséquent ρ aura au moins un carré parmi ses images et l'affaire est entendue... ! ❑

En revenant à notre théorème : Considérons alors b_2 un vecteur orthogonal à b_1, pour une question de rang, on a alors $q(b_2) \neq 0$. D'autre part, nous avons la décomposition

$$\mathbb{F}_q = \mathbb{F}_q^2 \cup \alpha\mathbb{F}_q^2.$$

Cette décomposition provient de la suite exacte courte

$$1 \to \{1,-1\} \to \mathbb{F}_q^* \xrightarrow{x^2} {\mathbb{F}_q^*}^2 \to 1.$$

⋆ Si $q(b_2) = \lambda^2 \in {\mathbb{F}_q^*}^2$, on va pouvoir considérer le vecteur normalisé $\frac{1}{\lambda}b_2$ et obtenir ainsi dans la base $\{b_1; \frac{1}{\lambda}b_2\}$ la matrice de q qui ne sera rien d'autre que l'identité qui est la matrice de q_1.

⋆ Sinon, $q(b_2) = \alpha\lambda^2$, on va considérer le vecteur normalisé $\frac{1}{\lambda}b_2$ et obtenir ainsi dans la base $\{b_1; \frac{1}{\lambda}b_2\}$ la matrice de q qui ne sera rien d'autre que celle de q_2.

• • Pour $n > 2$: Suivant la base orthogonale considérée $e_1, \cdots, e_n$, on peut regarder la restriction de q au plan $\text{Vect}(e_1, e_2)$, l'intérêt d'avoir considéré un tel plan issu en quelque sorte du choix d'une base orthogonale, c'est d'avoir une restriction qui reste non-dégénérée, par le Lemme 1.3 précédent, on va pouvoir trouver un vecteur $b_1 \in \text{Vect}(e_1, e_2)$ tel que $q(b_1) = 1$, On applique alors la récurrence sur l'hyperplan $\text{Vect}(b_1)^\perp$ sur lequel la restriction de q est encore non-dégénérée. [On constate en fin de compte que le choix initial de la base orthogonale est double.... !].

• • • D'autre part, il est facile de calculer les déterminants de q_1 et de q_2, i.e. leurs discriminants, le premier valant 1 et le second valant α, avec α qui n'est pas un carré, or on a vu que les discriminants entre deux formes équivalentes sont reliée par un carré (cf. Remarque (3) 2 p.21). Elle ne sont donc pas équivalentes. ❑

Dans les démonstrations de ces trois théorèmes, j'aimerai revenir sur la discussion de ou des classes d'équivalence. A chaque fois, on a fabriqué une base dans laquelle la matrice de q avait l'aspect souhaité. En algèbre linéaire classique, c'est aussi de cette manière que l'on fait pour montrer par exemple que les matrices sont "*semblables*". Pourtant ce n'est pas ainsi que l'on passe d'une matrice à l'autre, où est donc la nuance, la subtilité, à déceler ? Il faut se convaincre que ce sont bien deux notions d'équivalence bien distinctes pour une même vision d'approche : Dire que deux endomorphismes f et g sont équivalents signifie qu'il existe $u \in GL(E)$ tel que $g = u^{-1} \circ f \circ u$, tandis que deux formes sesquilinéaires ou quadratiques) sont "*équivalentes*" si $q_g = q_f \circ u$; le point qui diverge, est que pour la première notion, la composition entre applications linéaires se retranscrit parfaitement par le produit entre matrices, tandis que pour la deuxième cela n'est pas le cas, car ce que l'on peut observer c'est que les entités que représentent les matrices ne sont pas de même "nature". Pourtant on y retrouve quand même une petite partie car dans l'expression ${}^tPAP$ on retrouve AP le début de $q \circ u$, ceci s'explique par le caracère "linéaire de la deuxième variable de f. Je laisse la fin de cette réflexion au lecteur.... !

Si maintenant on considère toutes les formes quadratiques dégénérées ou pas, on aura exactement $2n+1$ classes d'équivalence de formes quadratiques symétriques.

2.2 Le cas hermitien

Nous allons présenter ici tout d'abord présenter sur le corps des nombres complexes puis sur un corps fini.

Thm. 3.7 [Formes hermitiennes sur le corps $\mathbb{C}$]

Soit E un $\mathbb{C}$-espace vectoriel de dimension fini n.

Alors il y a exactement $(n+1)$ classes de formes hermitiennes, non-dégénérées, relativement à la conjugaison.

• Dans une base convenable, la matrice de chacune de ces classes sera de la forme

$$\begin{pmatrix} 1 & & & & & \\ & \ddots & & 0 & & \\ & & 1 & & & \\ & & & -1 & & \\ & 0 & & & \ddots & \\ & & & & & -1 \end{pmatrix}$$

où 1 y apparaît p fois et -1 y apparaît $(n-p)$ fois (avec p parfaitement déterminé), et si on on exprime tout vecteur dans cette base $z=(z_1,\cdots,z_n)$ on aura :

$$q(z)=(z_1\overline{z_1}+\cdots z_p\overline{z_p})-(z_{p+1}\overline{z_{p+1}}+\cdots+z_n\overline{z_n}).$$

Preuve. • 1ère méthode : Bien que le corps sur lequel on évolue soit algébriquement clos, la preuve va plus ressembler à la preuve du Thm.1.5 p.23 : Soit $e_1,\cdots,e_n$ une $\mathbb{C}$-base orthogonale adaptée à q (cette base est assurée par le Thm.1.3 p.19). Notons $\gamma_i=q(e_i)=f(e_i,e_i)$ les coefficients sur cette diagonale, alors on a $\gamma_i=\overline{\gamma_i}$ car ici $f(e_i,e_i)=\overline{f(e_i,e_i)}$ on a une forme hermitienne associée à la conjugaison. Alors on a $\gamma_i\in\mathbb{R}\setminus\{0\}$, c'est la raison pour laquelle cela ressemble au cas symétrique sur $\mathbb{R}$; on recopie cette preuve pour avoir la matrice et le reste du résultat annoncés.

• 2ème méthode : Considérons toujours $e_1,\cdots,e_n$ une $\mathbb{C}$-base orthogonale adaptée à q. Considérons E_1 le $\mathbb{R}$-espace vectoriel engendré par ces derniers, ce dernier espace est de dimension n et l'observation intéressante que nous devons faire est que q_1, la restriction de q à ce dernier espace, devient alors une forme "symétrique" non-dégénérée. En appliquant le Thm.1.5 p.23, après normalisation, on aura sur la diagonale p termes en 1 et $(n-p)$ termes en -1. Considérons le sous-espace $E_2 \overset{\triangle}{:=} \mathbf{i}.E_1$, la même observation peut se faire concernant q_2, la restriction de q sur cet espace, de plus l'observation supplémentaire que nous pouvons faire est que après la normalisation (qui sera la même), la matrice de q sur cet espace sera exactement la même, c'est en effet à cause du fait que $q(e)=q(\mathbf{i}e)$.

Pour finir observons enfin que si on note $z=(z_1,\cdots,z_n)$ relativement à cette base que l'on vient de normaliser, notons $z_k=a_k+\mathbf{i}b_k$, alors z s'écrit aussi $z=\underbrace{(a_1,\cdots,a_n)}_{z_1}+\mathbf{i}\underbrace{(b_1,\cdots,b_n)}_{z_2}$. Je ne détaille pas les calculs mais on va bien trouver

$$\begin{aligned} q(z) &= q(z_1+\mathbf{i}z_2)=q(z_1)+\underbrace{\mathbf{i}\bar{\mathbf{i}}}_{1}q(z_2)=q_1(z_1)+q_2(z_2)\\ &= \left[(a_1^2+\cdots+a_p^2)-(a_{p+1}^2+\cdots+a_n)\right]+\left[(b_1^2+\cdots+b_p^2)-(b_{p+1}^2+\cdots+b_n)\right]\\ &= \Big[\underbrace{(a_1^2+b_1^2)}_{z_1\overline{z_1}}+\cdots+\underbrace{(a_p^2+b_p^2)}_{z_p\overline{z_p}}\Big]-\Big[\underbrace{(a_{p+1}^2+b_{p+1}^2)}_{z_{p+1}\overline{z_{p+1}}}+\cdots+\underbrace{(a_n^2+b_n^2)}_{z_n\overline{z_n}}\Big]. \end{aligned}$$

❑

Pour ce qui a été du corps des nombres complexes, nous avions explicitement une involution non triviale qui est la conjugaison, mais pour ce qui est des autres corps, nous n'avons pas assez de recul pour exiber une involution qui ne soit pas triviale ; nous allons avoir besoin pour cela d'étudier un petit peu comme cela avait déjà annoncé dans la Remarque p.11 (voir aussi Remarque (3) p.21).

Prop. 3.3 [Les involutions sur un corps fini]

Soit k un corps fini avec $\mathrm{car}(k) \neq 2$. Alors on a :

1. Si il existe une involution σ de k, non-triviale, alors le cardinal de k est un carré. En particulier, il existe un nombre q (pas nécesairement premier) tel que $k = \mathbb{F}_{q^2}$.
2. Réciproquement, il existe une unique involution σ, non-triviale, sur $\mathbb{F}_{q^2}$, donnée par $\sigma(x) = x^q$. De plus, le sous-corps des invariants de σ est égal à $\mathbb{F}_q$.

Preuve. (1) Nous avons vu dans la Prop.1.1 p.9, que le corps des invariants d'une involution non-triviale noté k_0 vérifie la relation $[k : k_0] = 2$ (i.e. le degré de l'extension k/k_0 vaut 2). En tant que corps fini k_0 sera de la forme $\mathbb{F}_q$, et par conséquent le cardinal de k sera égal à q^2, mais on sait par la théorie de Galois qu'on n'a qu'un seul corps ayant ce cardinal, ça ne peut être que $k = \mathbb{F}_{q^2}$. D'autre part, d'après Remarque p.11, l'involution σ étant non-triviale est l'unique involution de k qui fixe les points de k_0.

(2) A cause du premier point il nous suffirait de voir que l'on n'a qu'un seul sous-corps de cardinal q dans $\mathbb{F}_{q^2}$, mais cela provient du :

Lem. 3.4

Soit p un nombre premier et soit $\mathbb{F}_{p^m}$ un corps fini à p^m éléments. Alors il existe une (correspondence) entre les sous-corps de $\mathbb{F}_{p^m}$ et les diviseurs de m. Pour tout diviseur d de m, le sous-corps correspondant est l'ensemble des racines de $X^{p^d} - X$.

Preuve. Si K est un sous-corps de $\mathbb{F}_{p^m}$, alors $\mathbb{F}_{p^m}$ est un K-espace vectoriel. De plus, $\dim_K(\mathbb{F}_{p^m})$ est fini car $\mathbb{F}_{p^m}$ est fini, on a alors $\mathbb{F}_{p^m} \simeq K^n$ et donc $p^m = \mathrm{card}(\mathbb{F}_{p^m}) = \big(\mathrm{card}(K)\big)^n$, par l'unicité de la décomposition des nombres (en sachant que p est premier), on en déduit que $\mathrm{card}(K)$ est de la forme p^d avec d qui vérifie $dn = m$, d est donc un diviseur de m.

Inversement, si d divise m, on peut écrire $dn = m$ et posons $q := p^d$. Alors on a $q^n = (p^d)^n = p^{dn} = p^m$. Considérons le morphisme de Frobenius

$$F : \mathbb{F}_{p^m} \to \mathbb{F}_{p^m},\ x \mapsto x^p,$$

On sait très bien que le résultat fondamental relatif à ce morphisme, affirme que *l'ensemble de ses points fixes $\mathbb{F}_{p^m}$ coïncide exactement avec son corps premier $\mathbb{F}_p$* (cf. [Per] Prop.2.4. p.73). Ce résulat trouve en fait une généralisation très naturelle : Considérons plutôt son itérée

$$\tilde{F} : \mathbb{F}_{p^m} \to \mathbb{F}_{p^m},\ x \mapsto \underbrace{(x^p)^d}_{x^{p^d}} = x^q,$$

alors on obtient le même résultat :

Cor. 3.3

Notons $q := p^d$, où p est un nombre premier. Et soit $n \in \mathbb{N} \setminus \{0\}$, alors on peut considérer le morphisme de Frobenius

$$\tilde{F} : \mathbb{F}_{q^n} \to \mathbb{F}_{q^n},\ x \mapsto x^q.$$

Alors l'ensemble des points fixes de $\tilde{F}$ dans $\mathbb{F}_{q^n}$ est isomorphe à $\mathbb{F}_q$.

Preuve. Les points fixes de $\tilde{F}$ est bien un sous-corps qui vérifient $x^q = x$, ce sont donc des racines du polynôme $X^q - X = 0$ (cf. Chapitre sur le groupe linéaire p.24, lorsqu'on est passé d'une équation algébrique que doivent vérifier des éléments, à une équation polynômiale). Ici on est sur un corps, qui est en particulier intègre, et tout polynôme n'a donc pas plus de racines que son degré (cf. cours de maîtrise sur les anneaux), en particulier l'ensemble des points fixes est donc un sous-corps de cardinal au plus q.
D'autre part, le groupe multiplicatif $\mathbb{F}_{q^n}^*$ est cyclique, isomorphe à $\mathbb{Z}/(q^n-1)\mathbb{Z}$ et d'ordre $q^n - 1$ (cf. [Per] Thm.2.7. p.74). Or $q - 1$ est un diviseur de $q^n - 1$, on a alors à l'intérieur du groupe $\mathbb{F}_{q^n}^* \simeq \mathbb{Z}/(q^n-1)\mathbb{Z}$ un sous-groupe G d'ordre $q - 1$ (cf. [Per] p.74, on a là aussi une correspondence). Le Théorème de Lagrange nous permet de dire que les éléments de G vérifient donc $x^{q-1} = 1$ et donc $x^q = x$. On rajoute alors 0, pour pouvoir affirmer que les points fixes de $\tilde{F}$ sont au nombre de q, et par unicité des corps finis, on en déduit que cet ensemble est isomorphe à $\mathbb{F}_q$. ❑

Poursuivons la preuve du Lemme : Par le Cor.1.3 précédent, il existe bien un sous-corps de $\mathbb{F}_{p^m}$ à $q := p^d$ éléments qui correspond aux points fixes de l'automorphisme de Frobenius $x \mapsto x^q$ sur $\mathbb{F}_{p^m}$, (ou) encore aux racines du polynôme $X^q - X$, mais ce dernier est un polynôme de degré q, et il a au plus q racines, par conséquent, il n'y a qu'un seul sous-corps de $\mathbb{F}_{p^m}$ qui comprend q éléments. ❑

Revenons à la fin de la preuve de Prop.1.3 : On voit très bien que c'est un cas particulier de ce qui a été développé dans le Lemme 1.4 et dans le Cor.1.3 précédent, en effet il suffit de prendre ici $m = 2$, $d = 1$, et $\sigma(x) = x^q$ l'automorphisme de Frobenius. ❑

Rmq

Avant d'entamer la preuve du Lemme 1.4 p.28 précédent, ce dernier me faisait penser à cette correspondence entre les sous-groupes d'un groupe cyclique d'ordre n et les diviseurs de n (cf. [Per] p.74) ; cela se reposait sur une forme de "*rigidité*" que devait vérifier globalement tous les éléments du groupe, on partitionnait le groupe en paquets d'éléments ayant même ordre d, notons par $N(d)$ le cardinal de chacun de ces paquets ; ce partage est à mettre en parallèle avec la relation $n = \sum_{d|n} \varphi(d)$, où φ est la fonction d'Euler, mais on parvenait à montrer que $N(d) \leqslant \varphi(d)$ (en fait dans [Per] p.74, il montre que $N(d)$ valait 0 ou $\varphi(d)$), mais comme la somme des $N(d)$ et des $\varphi(d)$ donnaient le même entier n, on en déduisait l'égalité.
Ici c'est un peu la même histoire, ce qu'il nous sauve dans la preuve c'est l'argument "*tout polynôme* P *admet au plus* $\deg(P)$ *racines sur un anneau intègre, ici on a un corps c'est encore mieux*", c'est ce qui nous a permis de ne pas avoir d'autre sous-corps que celui qui nous intéressait. On pourra consulter le problème 6 p.39 ou encore l'exercice 10 p.26/27 dans le livre [Gou] qui relie bien ce que j'évoque ici.

Enfin, nous pouvons maintenant présenter :

Thm. 3.8 [Formes hermitiennes sur un corps fini]

Soit E un k-espace vectoriel de dimension finie n, où k est un corps fini, avec $\mathrm{car}(k) \neq 2$, muni d'une involution σ non-triviale.

Alors il n'y a qu'une unique classe d'équivalence parmi les formes quadratiques, hermitiennes, relatives à σ.

- Dans une base convenable sa matrice sera l'identité.

Preuve. D'après la Prop.1.3 p.28, l'existence d'une involution σ non-triviale sur un corps fini, nous permet de supposer que $k = \mathbb{F}_{q^2}$ et σ l'automorphisme de Frobenius $x \mapsto x^q$. Soit $e_1, \cdots, e_n$ une base orthogonale pour la forme hermitinne f, nous donne $\gamma_i = f(e_i, e_i) = f(e_i, e_i)^\sigma$, si bien que l'on a $\gamma_i \in \mathbb{F}_q^*$.

Considérons l'application :

$$N : \mathbb{F}_{q^2}^* \to \mathbb{F}_{q^2}^*,\ x \mapsto N(x) \overset{\triangle}{:=} x x^\sigma = x^{q+1}\ .$$

[Remarquons que ce n'est pas un hasard si on a posé xx^σ, cela ne vous fait penser à rien ? Il suffit de considérer le cas symétrique ou encore dans le cas de la conjugaison des nombres complexes... ! Dans chacun de ces cas, on a trouvé sur la diagonale des matrices des coefficients de la formes $N(x)$... !]

Commençons par remarquer que $N(x)^\sigma = (x^{q+1})^q = (x^q x)^q = \underbrace{x^{q^2}}_{x} x^q = x^{q+1} = N(x)$, i.e. que l'image de l'application N est contenue dans $\mathbb{F}_q^*$. D'autre part, on remarque facilement que N est un morphisme de groupes. On obtient ainsi un morphisme de groupes $N : \mathbb{F}_{q^2}^* \to \mathbb{F}_q^*$. Et on a

$$\mathrm{Ker}(N) = \left\{ x \in \mathbb{F}_{q^2}^* \quad | \quad x^{q+1} = 1 \right\}.$$

Comme dans l'argument dans la preuve du Cor.1.3 p.29 ci-dessus, l'équation $X^{q+1} - 1 = 0$ a au plus $(q+1)$ racines, et si $\varphi : G \to G'$ est un morphisme de groupes (finis), on a $\mathrm{card}(G) = \mathrm{card}(\mathrm{Ker}(\varphi)) \times \mathrm{card}(\mathrm{Im}(\varphi))$, si bien que

$$\mathrm{card}(\mathrm{Im}(N)) \geqslant \frac{q^2 - 1}{q+1} = q - 1.$$

Cette relation, nous permet de dire que le morphisme N est nécessairement surjectif, on a donc $\mathrm{Im}(N) = \mathbb{F}_q^*$.

De ce résultat, on en déduit que chaque γ_i admet un antécédent λ_i par N, i.e. $\gamma_i = N(\lambda_i) = \lambda_i \lambda_i^\sigma$ pour un certain $\lambda_i \in \mathbb{F}_{q^2}^*$. Posons alors $b_i \overset{\triangle}{:=} \dfrac{1}{\lambda_i} e_i$, on va trouver que $f(b_i, b_i) = \dfrac{\gamma_i}{\lambda_i \lambda_i^\sigma} = 1$ et notre résultat est démontré.

❏

Références

[Gou] Gourdon X. : *Algèbre*, 2^{e} édition, (les maths en tête), Ellipses 2008.

[Per] Perrin D. : *Cours d'Algèbre*, Ellipses, 1996.

[Wag] Wagschal C. : *Topologie et analyse fonctionnelle*, Hermann 1995.

Vocabulaire