« Algèbre »
(cours de Magistère du second semestre 2014/15,
basé sur mes cours « Algèbre-2 » à l'ENSL en 2007 et
« Groupes classiques, algèbre géométrique et
applications » à l'Institut Fourier à 2010 et 2014).

Alexei PANTCHICHKINE (Institut Fourier)

PROGRAMME BREF

1. Algèbre tensorielle.
2. Exemples et constructions de corps.
Extensions algébriques, degré, polynôme minimal,
caractéristique. Corps de rupture d'un polynôme.
3. Corps finis. Applications aux codes correcteurs.
4. Exemples de groupes, groupes classiques.
5. Applications physiques*:
Espace-temps de Minkowski. Groupe de Lorentz.
6. Géométrie projective. Coniques, quadriques.
7. Variétés affines (exemples). Courbes planes, points singuliers

---

*Si le temps le permet

De quoi s'agit-il?

Les plus connus groupes classiques (sur les complexes) sont
GL(n), SL(n), O(n), SO(n), U(n), SU(n), Sp(2n).
De point de vue de la géométrie, ce sont certaines groupes de transformations du C-espace vectoriel V=C^n (à savoir, soit les transformations linéares inversibles, soit les transformations linéaires spéciales, orthogonales,
ainsi que les transformations unitaires ou symplectiques),
données par X->AX (A une matrice, X un vecteur colonne de V).

En gros, on définie ces sous-groupes de GL(n) (ou de GL(2n)) par une condition d'invariance de sorte B(AX, AY)=B(X,Y), où B soit une forme bilinéaire (symétrique ou alternée), soit une forme sésquilinéaire (hermitienne) connue des cours d'algébre linéaire comme des transformations orthogonales, unitaires etc.
B(X,Y) et X->AX sont des examples de la notion de tenseur. De plus, au lieu du corps k=C on considère un corps quelconque, y compris k=R, Q, un corps fini F_q (ou une extension de tels corps).

Programme

1. Algèbre tensorielle
2. ~~Polynômes et fractions rationnelles~~
3. Exemples et constructions de corps
4. Exemples de groupes. Groupes classiques
5. Géométrie projective. Coniques, quadriques

Programme bref

Algèbre tensorielle. ~~Polynômes et fractions rationnelles~~, Exemples et constructions de corps. Extensions algébriques, transcendance, degré, polynôme minimal, caractéristique. Corps de rupture d'un polynôme. Corps finis. Exemples de groupes, groupes classiques. Géométrie projective. Coniques, quadriques. Variétés affines (exemples). Courbes planes, points singuliers

Sommaire

Table des matières

J'ai signalé avec une * ce qui peut être sauté en premère lecture.

Cours N1. Lundi 19 janvier 2015

(disponible sur : `http ://www-fourier.ujf-grenoble.fr /~panchish`)

Motivations et contenu du cours

Le présent cours est centré sur les notions de l'algèbre tensorielle, ~~polynômes et fractions rationnelles~~, exemples et constructions de corps, exemples de groupes, groupes classiques, géométrie projective. Ce sont les outils algébriques pour la géométrie (en particulier la géométrie différentielle et la géométrie algébrique), pour l'arithmétique (equations diophantiennes, représentations galoisiennes), pour l'analyse (fonctions spéciales de variables matricielles, équations différentielles et groupes de monodromie), pour la physique mathématique (en particulier, la mécanique quantique, voir [Coq02], Partie IV de [KosMan]), ainsi que pour beaucoup d'applications (codes géométriques etc.)

Les corps finis donnent des exemples importants d'extensions de corps, et on étudie en détail les polynômes irréductibles sur les corps finis.

Le cours est considéré comme la suite du cours "Algèbre 1", et on utilise comme prérequis les notions de théorie des ensembles, groupes, anneaux et corps, les notions d'algèbre linéaire, y compris les formes quadratiques, géométrie affine (euclidienne).

Première partie
Algèbre tensorielle

Motivation : algèbre polylinéaire

0.1 Formalisme d'applications polylinéaires

Cette partie est consacrée à l'étude systématique des constructions polylinéaires pour les espaces vectoriels V sur un corps commutatif K, et pour les modules M sur un anneau commutatif A. Ici on introduit la notion du produit tensoriel qui sert de base aux constructions algébriques ; on l'étudie en détails.

Cependant, les applications principales de ce formalisme se trouve à l'extérieure de l'algèbre linéaire, notammement, dans la géométrie différentielle, théorie des représentations de groupes et dans la mécanique quantique, voir Partie IV de [KosMan] et [Coq02].

Soit V_1, ..., V_r une famille finie de K-espaces vectoriels, et soit $V_1 \times \cdots \times V_r$ leur produit cartesien.

DÉFINITION 0.1.1 *Une application*

$$f : V_1 \times \cdots \times V_r \to V$$

du produit cartesien $V_1 \times \cdots \times V_r$ dans un autre K-espace vectoriel V est dite polylinéaire, si elle est linéaire pour tous les arguments $v_i \in V_i$, $i = 1, \cdots r$. Notation *:*

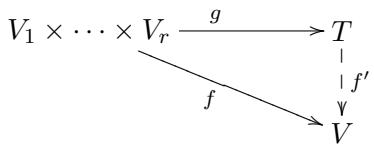
$$f \in \mathcal{L}(V_1 \times \cdots \times V_r; V).$$

On va constrire une application polylinéaire universelle

$$g : V_1 \times \cdots \times V_r \to V_1 \otimes \cdots \otimes V_r,$$

dont l'image $T = V_1 \otimes \cdots \otimes V_r$ est dit le *produit tensoriel* de $V_1, \ldots, V_r$.

DÉFINITION 0.1.2 (PROPRIÉTÉ UNIVERSELLE) *il existe une application K-polylinéaire $g : V_1 \times \cdots \times V_r \to T$, telle que toute application K-polylinéaire $f : V_1 \times \cdots \times V_r \to V$ se factorise par g : $f = f' \circ g$ pour une unique application K-linéaire $f' : T \to V$:*



C'est-à-dire, que $\mathcal{L}(V_1 \times \cdots \times V_r; V) = \mathcal{L}(V_1 \otimes \cdots \otimes V_r, V)$ (l'identification canonique).

Notation : $v_1 \otimes \cdots \otimes v_r = g(v_1, \ldots, v_r)$
tenseur décomposable

On appelle les vecteurs $v_1 \otimes \cdots \otimes v_r$ éléments décomposables de l'espace vectoriel des tenseurs générales $V_1 \otimes \cdots \otimes V_r$. Le produit tensoriel T est engendré par les tenseurs décomposables, et on le construire sur les corps et sur les anneaux commutatifs.

Idées de base du calcul tensoriel

- Dualité : on considère les vecteurs $v \in V$ comme des applications K-linéaires sur l'espace V^* des formes K-linéaires $\ell : V \to K$

$$v : \ell \mapsto \ell(v) \in K, \text{ de plus, } V \cong (V^*)^* \text{ si } \dim V < \infty$$

- Produit d'applications polylinéaires à valeurs scalaires (de variables différentes) est de nouveau polylinéaire

$$f : V_1 \times \cdots \times V_r \to K, \ h : V_{r+1} \times \cdots \times V_{r+s} \to K \Rightarrow$$
$$(f \otimes h)(v_1, \ldots, v_{r+s}) \stackrel{def}{=} f(v_1, \ldots, v_r) h(v_{r+1}, \ldots, v_{r+s})$$

- Identification d'applications K-linéaires $F \in \mathcal{L}(U, \mathcal{L}(V, W))$ avec certaines applications K-bilinéaires $\tilde{F} :\in \mathcal{L}(U, V; W)$

$$F \mapsto (\tilde{F} : (u, v) \mapsto F(u)(v)), \ \mathcal{L}(U, \mathcal{L}(V, W)) = \mathcal{L}(U, V; W) \text{ si } \dim U, \dim V, \dim W < \infty.$$

DÉFINITION 0.1.3 (TENSEURS CLASSIQUES) *Un tenseur F p-fois covariant et q-fois contravariant sur un K-espace vectoriel V est une application polylinéaire*

$$F : \underbrace{V \times \cdots \times V}_{p \text{ fois}} \times \underbrace{V^* \times \cdots \times V^*}_{q \text{ fois}} \to K, \quad i.e. F \in \underbrace{V^* \otimes \cdots \otimes V^*}_{p \text{ fois}} \otimes \underbrace{V \otimes \cdots \otimes V}_{q \text{ fois}}$$

On note $\mathbb{T}_p^q(V) = V^{*\otimes p} \otimes V^{\otimes q} = \mathcal{L}(V^{\otimes p} \otimes V^{*\otimes q}; K)$.

Soit $\dim(V) < \infty$. Les cas particuliers de la notion d'un tenseur F inclus :

- un vecteur $v \in \mathbb{T}_0^1(V) = (V^*)^*$.
- une forme K-linéaire $\ell : V \to K \in \mathbb{T}_1^0$
- une forme bilinéaire $f : V \times V \to K \in \mathbb{T}_2^0$
- un opérateur linéaire $u : V \to V : u \in \mathbb{T}_1^1(V) \cong \mathcal{L}(V^* \otimes V; K)$

Notation classique

Soit $V = \langle e_1, \ldots, e_n \rangle$ un K-espace vectoriel d'une base $\{e_1, \ldots, e_n\}$, donc $\dim(V) = n$. On note $e^1, \ldots, e^n$ la base duale de V^*, c'est-à-dire, $e^j(e_i) = \delta_{ij}$. Alors tout tenseur

$$F \in \mathbb{T}^p_q(V) = V^{*\otimes p} \otimes V^{\otimes q} = \mathcal{L}(V^{\otimes p} \otimes V^{*\otimes q}; K)$$

est détérminé par ces composantes $F^{j_1 \cdots, j_q}_{i_1 \cdots, i_p} = F(e_{i_1}, \cdots, e_{i_p}, e^{j_1} \cdots, e^{j_q})$. De plus

$$F = \sum_{\substack{i_1 \cdots, i_p \\ j_1 \cdots, j_q}} F^{j_1, \cdots, j_q}_{i_1 \cdots, i_p} e^{i_1} \otimes \cdots \otimes e^{i_p} \otimes e_{j_1} \otimes \cdots \otimes e_{j_q} = F^{j_1, \cdots, j_q}_{i_1 \cdots, i_p} e^{i_1} \otimes \cdots \otimes e^{i_p} \otimes e_{j_1} \otimes \cdots \otimes e_{j_q},$$

(selon la convention d'Einstein, on omet le symbole de sommation d'indices répétants).

1 ~~Rappels sur la notion d'anneau, exemples~~

1.1 Structure d'anneau et idéaux

Définition 1.1.1 *Un anneau est un groupe abélien A muni d'une loi interne*

$$A \times A \to A, \quad (x, y) \mapsto xy = x \cdot y$$

appelé produit ou multiplication, qui est associative

An1 $\forall x, y, z \in A, x(yz) = (xy)z$,

et distributive à droite et à gauche pour à l'addition :

An2 $\forall x, y, z \in A, x(y + z) = xy + xz$,

An3 $\forall x, y, z \in A, (y + z)x = yx + zx$,

On prendra également la convention que tout anneau est unifère, c'est-à-dire que la multiplication est munie d'un élément neutre 1 :

An4 $\forall x \in A, 1x = x1 = x$.

L'anneau est dit commutatif si la loi de multiplication est commutative :

Comm. $\forall x, y \in A, xy = yx$.

Définition 1.1.2 *Un morphisme d'anneau $\varphi : A \to B$ est une application telle que*

MorAn $\forall x, y, z \in A, \varphi(xy + z) = \varphi(x)\varphi(y) + \varphi(z) \in A, \varphi(1_A) = 1_B$

SAn *Une partie $A \subset B$ est dit un sous-anneau, si l'inclusion $A \hookrightarrow B$ est un morphisme d'anneau.*

Exemple On pose $B = \mathbb{Z} \times \mathbb{Z} = \{(x_1, x_2) \mid x_1, x_2 \in \mathbb{Z}\}$, alors $A = \{0\} \times \mathbb{Z}$ est un anneau, mais non un sous-anneau de B.

Définition 1.1.3 *Soit A un anneau commutatif. Une partie $I \subset A$ est dit un idéal si c'est un sous-groupe additif pour l'addition, stable par la multiplication externe (par un élément quelconque $a \in A$).*

Idéal $\forall x \in I, \forall a \in A, ax \in I$.

Opérations sur les idéaux

DÉFINITION 1.1.4 *(a) Soient I, J deux idéaux de A. Leur* somme

$$I + J = \{x + y \mid x \in I, y \in J\}$$

est le plus petit idéal de A contenant I et J.

La somme d'une famille d'idéaux $(I_\alpha)_{\alpha\in\Gamma}$ *est formée par toutes les sommes finies*

$$\sum_{\alpha\in\Gamma} I_\alpha = \left\{ \sum_{\alpha\in\Gamma} x_\alpha, x_\alpha \in I_\alpha \right\}$$

où $x_\alpha = 0$ sauf un nombre fini de $\alpha \in \Gamma$.

(b) L'intersection ensembliste

$$\bigcap_{\alpha\in\Gamma} I_\alpha$$

d'une famille d'idéaux $(I_\alpha)_{\alpha\in\Gamma}$ *est toujours un idéal de A.*

(c) Soit X une partie d'un anneau A. L'intersection de tous les idéaux de A, contenants X, est dit l'idéal engendré par X

(d) Le produit

$$I_1 \cdot I_2 \cdot \ldots \cdot I_n$$

d'un nombre fini d'idéaux est l'idéal engendré par

$$\{x_1 \cdot x_2 \cdot \ldots \cdot x_n \mid x_1 \in I_1, x_2 \in I_1, \cdots, x_n \in I_n\}$$

En particulier, l'idéal I^n est engendré par

$$\{x_1 \cdot x_2 \cdot \ldots \cdot x_n \mid x_1, x_2 \in I_1, \cdots, x_n \in I\}$$

EXEMPLE

a) Si $A = \mathbb{Z}$, $I = (m)$, $J = (n)$, alors

$$I + J = (\mathrm{pgcd}(m, n)), I \cap J = (\mathrm{ppcm}(m, n)), I \cdot J = (mn).$$

REMARQUE

L'idéal, engendré par une famille x_α, coïncide avec la somme

$$\sum_{\alpha\in\Gamma} (x_\alpha)$$

de tous les idéaux principaux $(x_\alpha) = x_\alpha A$.

REMARQUE Montrer en exercice que l'union d'une famille d'idéaux $(I_\alpha)_{\alpha\in\Gamma}$ n'est pas un idéal en général, mais c'est le cas si les idéaux I_α sont totalement ordonnés par l'inclusion :

$$\forall \alpha, \beta, \text{ soit } I_\alpha \subset I_\beta, \text{ soit } I_\beta \subset I_\alpha.$$

1.2 Anneau quotient

DÉFINITION 1.2.1 *Soit A un anneau commutatif, $I \subset A$ un idéal de A. Alors il existe sur le groupe quotient additif A/I une unique structure d'anneau telle que la projection canonique $\pi : A \to A/I$ est un morphisme d'anneaux.*

PROPOSITION 1.2.2 *(a) Soit A un anneau commutatif, $I \subset A$ un idéal de A. Alors il existe une bijection entre l'ensemble*

$$\{J \subset A \mid J \supset I\}$$

d'idéaux contenants I, et l'ensemble

$$\left\{\overline{J} \subset A/I\right\}$$

d'idéaux de A/I, donnée par $J = \pi^{-1}(\overline{J})$, où $\pi : A \to A/I$ est la projection canonique.

(b) Soit $\psi : A \to B$ un morphisme d'anneaux, alors $I = \operatorname{Ker}\psi := \psi^{-1}(0)$ est un idéal de A, $\psi(A) = C$ est un sous-anneau de B, et il y a un isomorphisme d'anneaux

$$\overline{\psi} : A/I \xrightarrow{\sim} C.$$

On écrit

$$x \equiv y \bmod I \iff x - y \in I.$$

EXERCICE Soit $A = \mathbb{Z}[X]$, $I = 5A = (5)$. Trouver tous les idéaux de A contenant I.

Diviseurs de zéro, éléments nilpotents et unités

DÉFINITION 1.2.3

(a) Un $x \in A\backslash\{0\}$ est dit diviseur de zéro, *s'il existe un $y \in A\backslash\{0\}$ tel que $xy = 0$. Un anneau $A \neq \{0\}$ sans diviseurs de zéro est dit* intègre.

(b) Un élément $x \in A\backslash\{0\}$ est dit nilpotent, *si $x^n = 0$ pour un $n \geq 1$.*

(c) Un élément $x \in A$ esi dit inversible *(ou une* unité*) de A s'il existe $y \in A, xy = 1$. On notera $x \in A^\times$.*

DÉFINITION 1.2.4 *Un corps est un anneau commutatif A, non réduit à $\{0\}$ dans lequel tout élément non-nul est inversible :*

Corps $\forall x \in A, x \neq 0, \exists y \in A, xy = 1$

PROPOSITION 1.2.5 *(a) Soit A un corps, alors A est un anneau intégre.*

(b) Soit A un corps, I un idéal de A. Alors soit $I = \{0\}$ soit $I = A$.

1.3 Idéaux premiers

DÉFINITION 1.3.1

(a) Un idéal $I \neq A$ est dit premier*, si*

$$\forall x, y \in A, x \cdot y \in I \Longleftrightarrow x \in I \text{ ou } y \in I,$$

i.e. l'anneau quotient A/I est intègre.

(b) Un idéal $I \neq A$ est dit maximal*, si*

$$\forall \text{ idéal } J \subset A, I \subset J \Rightarrow I = J, \text{ ou } J = A$$

PROPOSITION 1.3.2

(a) Un idéal $I \neq A$ est dit maximal*, si et seulement si A/I est un corps*

(b) Tout idéal maximal est premier.

PREUVE (a) On suppose I maximal. Si $x \notin I$, on considère l'idéal (x, I) engendré par x et I. Alors $(x, I) \neq I$ donc $(x, I) = A$; ceci dit, il existe $a \in A$ et $b \in I$ tels que $ax + b = 1$; ceci dit, $\overline{ax} = \overline{1}$ dans A/I.

Réciproquement, si A/I est un corps, les seuls idéaux de A/I sont $\{0\}$ et A/I. Par la proposition 1.2.2, a), il existe une bijection entre l'ensemble

$$\{J \subset A \mid J \supset I\}$$

d'idéaux contenants I, et l'ensemble

$$\left\{\overline{J} \subset A/I\right\}$$

d'idéaux de A/I. Donc il n'y a pas d'idéaux stricts intermédiaires entre I et A, i.e. I est maximal.

(b) Un corps est toujours un anneau intègre, donc I est premier.

EXEMPLE

Dans l'anneau $A = \mathbb{C}[X, Y]$ l'idéal $I = (X, Y)$ est maxmal, $A/I \xrightarrow{\sim} \mathbb{C}$.

L'idéal $J = (X)$ n'est pas maxmal, mais premier :$A/J \xrightarrow{\sim} \mathbb{C}[Y]$.

EXERCICE (à faire en TD) Montrer que tous les idéaux maximaux M de l'anneau $A = \mathbb{Z}[X]$ sont de la forme : $M = (p, f)$, où $f \in \mathbb{Z}[X]$ est un polynôme tel que $f \bmod p \in \mathbb{F}_p[X]$ est irréductible. Ici $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$ est le corps de p éléments.

L'idéal $J = (p)$ n'est pas maxmal, mais premier :$A/J \xrightarrow{\sim} \mathbb{F}_p[X]$.

EXERCICES

1.1 Trouver tous les diviseurs de zéro dans les anneaux

$$\mathbb{Z}/100\mathbb{Z}, \mathbb{Z}/72\mathbb{Z}, \mathbb{Z}/p^n\mathbb{Z}.$$

1.2 Trouver tous les éléments nilpotents dans les anneaux

$$\mathbb{Z}/100\mathbb{Z}, \mathbb{Z}/72\mathbb{Z}, \mathbb{Z}/p^n\mathbb{Z}.$$

1.3 Trouver tous les éléments inversibles dans les anneau

$$\mathbb{Z}/100\mathbb{Z}, \mathbb{Z}/72\mathbb{Z}, \mathbb{Z}/p^n\mathbb{Z}.$$

1.4 Montrer qu'un anneau fini A est intègre si et seulement s'il est un corps.

1.5 Trouver tous les éléments inversibles dans les anneau $\mathbb{Z}[i]$ et $\mathbb{Z}[j]$.

2 Modules et espaces vectoriels

2.1 Rappels sur les modules et espaces vectoriels

Définition 2.1.1 *Si A est un anneau commutatif. Un A-module est la donnée d'un groupe abélien M, muni d'une loi externe*

$$\times : A \times M \to M, (a, m) \mapsto am \in M \quad (\text{"}\textit{multiplication externe}\text{"})$$

satisfaisant les propriétés suivantes :

Mo1 $\forall a \in A,\ \forall x, y \in M,\ a(x + y) = ax + ay,$

Mo2 $\forall a, b \in A,\ \forall x \in M,\ (a + b)x = ax + bx,$

Mo3 $\forall a, b \in A,\ \forall x \in M,\ a(bx) = (ab)x.$

Si $A = K$ est un corps, *un* espace vectoriel *sur K est un K-module.*

Définition 2.1.2 *Soit A un anneau commutatif, et soient M, N deux A-modules. Une application $\phi : M \to N$ est dit un* morphisme *de A-modules (ou une application A-linéaire) si c'est un morphisme de groupe abéliens, et si elle vérifie la condition suivante :*

Mor. $\forall \lambda \in A, \forall m \in M, \phi(\lambda m) = \lambda \phi(m)$

Un isomorphisme de A-modules est un morphisme A-modules qui est bijectif. Son inverse est alors un morphisme A-modules.

2.2 Exemlples de modules

Exemple La notion de $\mathbb{Z}$-module coïncide avec celle de groupe abélien :

$$\times : \mathbb{Z} \times M \to M, (a, m) \mapsto am \in M$$

Exemple 2.2.1 *Si A est un anneau commutatif, et $n \in \mathbb{N}$, A^n est un A-module pour la lois externe*

$$\times : A \times A^n \to A^n, (a, (a_1, \cdots, a_n)) \mapsto (aa_1, \cdots, aa_n) \in A^n$$

Exemple 2.2.2 *Soit A un anneau commutatif, et soit M un A-module. Si X est un ensemble, alors pour tout $a \in A$ et pour toute application $f : X \to M$ on pose*

$$\forall x \in X, (af)(x) = a(f(x)) \in M.$$

L'ensemble M^X de toutes les applications $f : X \to M$ est un A-module avec la loi externe

$$\times : A \times M^X \to M^X, (a, f) \mapsto af \in M^X$$

2.3 Sous-A-modules, sous-espaces vectoriels

DÉFINITION 2.3.1 *Soit A un anneau commutatif, et soit M un A-module. Un sous-groupe abélien $N \subset M$ est dit un sous-A-module si il vérifie la condition suivante :*

$\mathsf{Sous-module}$ $\forall \lambda \in A, \forall x \in N, \lambda x \in N$.

Si K est un corps, un sous-K-module d'un espace vectoriel sur K est dit un sous-espace vectoriel sur K.

EXEMPLE 2.3.2 *Soit A un anneau commutatif, et soient M, N deux A-modules. L'ensemble $\mathcal{L}(M,N)$ des applications A-linéaires $\phi : M \to N$ est un sous-A-module du module N^M de toutes les applications de M vers N. En particulier, si K est un corps, et E un K-espace vectoriel, le dual de E, noté $E^\vee$ est l'espace vectoriel $\mathcal{L}(E,K)$.*

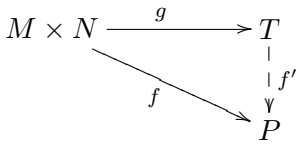
3 Produit tensoriel de modules

3.1 Existence et l'unicité du produit tensoriel

Soient M, N, P trois A-modules.

DÉFINITION 3.1.1 *a) Une application $f : M \times N \to P$ est dite A-bilinéaire, si pour tout $x \in M$, l'application $y \mapsto f(x,y)$ de N à P, et pour tout $y \in N$, l'application $x \mapsto f(x,y)$ de M à P, sont des homomorphismes de A-modules.*

b) Un A-module T est dit le produit tensoriel $M \otimes_A N$ s'il satisfait la propriété universelle suivante : il existe une application A-bilinéaire $g : M \times N \to T$, telle que toute application A-bilinéaire $f : M \times N \to P$ se factorise par g : $f = f' \circ g$ pour un unique homomorphisme $f' : T \to P$ de A-modules.



REMARQUE 3.1.2 *En général, g n'est pas surjective ! Mais $\mathrm{Im}(g)$ engendre T.*

PROPOSITION 3.1.3 (L'EXISTENCE ET L'UNICITÉ DU PRODUIT TENSORIEL) *a) Pour tous les A-modules M et N il existe un couple (T,g), où $g : M \times N \to T$ une application A-bilinéaire avec la propriété universelle de Definition 3.1.1, b).*

b) Le A-module T fourni avec l'application $g : M \times N \to T$ est unique à un isomorphisme près : pour tout autre tel couple (T', g') on a $g = j \circ g'$, pour un isomorphisme $j : T \xrightarrow{\sim} T'$ de A-modules.

PREUVE : Existence. Soit $C = A^{(M\times N)}$ le A-module libre de base

$$\{e_{x,y} \mid (x,y) \in M \times N\}\,, \text{ i.e. les éléments de } C \text{ sont}$$

$$C = \left\{ \sum_{i=1}^{n} \alpha_i e_{x_i,y_i} \mid \alpha_i \in A,\ x_i, \in M, y_i \in N \right\},$$

(toutes les combinaisons A-linéaires finies formelles ; on remarque que $e_{0,0} \neq 0$).

On identifie C avec l'ensemble des vecteurs infinis $(\alpha_{x,y})_{(x,y)\in M\times N}$ tels que $\alpha_{x,y} \in A$ sont présque tous nuls (sauf un nombre fini),

$$e_{x,y} \longleftrightarrow (\cdots, 0, \underbrace{1}_{(x,y)-\text{place}}, 0\cdots).$$

On considère le sous-A-module D de C engendré par tous les éléments de type

$$e_{x+x',y} - e_{x,y} - e_{x',y}, e_{x,y+y'} - e_{x,y} - e_{x,y'},$$
$$e_{ax,y} - ae_{x,y}, e_{x,ay} - ae_{x,y}.$$

Puis on pose $T = C/D$ et on note $x \otimes y$ la classe $e_{x,y} + D$ dans T.

Le module T est engendré par

$$x \otimes y = e_{x,y} + D,$$

de plus

$$(x + x') \otimes y = x \otimes y + x' \otimes y, x \otimes (y + y') = x \otimes y + x \otimes y',$$
$$(ax) \otimes y = a(x \otimes y), x \otimes (ay) = a(x \otimes y).$$

On pose $g(x,y) = x \otimes y$, alors $f'(x \otimes y) = f(x,y)$ est bien définie comme une unique application A-linéaire avec la propriété $f = f' \circ g$.

REMARQUE a) On a $e_{0,0} - 0e_{0,0} \in D$, donc $0 \otimes 0 = 0$ dans T.

b) Le produit $x \otimes y$ dépend de choix de modules $M \ni x$, $N \ni y$.

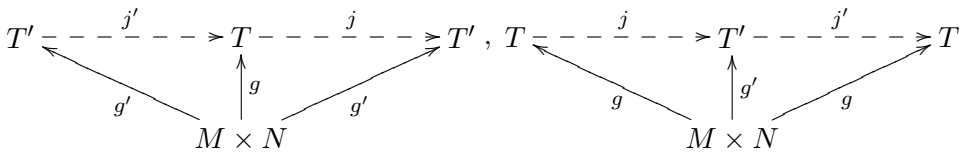
Il se peut que $x \in M' \subset M$, $y \in N' \subset N$, $x \otimes y = 0$ dans $M \otimes N$ mais

$$x \otimes y \neq 0 \text{ dans } M' \otimes N'.$$

EXEMPLE Soient $A = \mathbb{Z}$, $M = \mathbb{Z}$, $M' = 2\mathbb{Z}$, $N = N' = \mathbb{Z}/2\mathbb{Z} = \langle y \rangle$, alors $2 \otimes y = 1 \otimes 2y$ dans $M \otimes N$ mais $2 \otimes y \neq 0$ dans $M' \otimes N'$, puisque $\{2\}$ est une base du $\mathbb{Z}$-module libre, $M' \cong \mathbb{Z}$, donc $\mathbb{Z} \otimes N \cong N$ par la propriété universelle.

EXERCICE Montrer que $A^n \otimes M \cong M^n$ pour tout $n \in \mathbb{N}$.

PREUVE de b) (l'unicité de T). Par la propriété universelle 3.1.1, b), il existe j, j', $j : T \to T'$, $j' : T' \to T$ tels que $j \circ j' = id_{T'}$, $j' \circ j = id_T$:



puisque j et j' sont détérminées par la propriété universelle 3.1.1, b).

Dans notre construction de $T = M \otimes_A N$ on pose $g(x,y) = x \otimes y \in T$, mais en pratique on n'utilise que l'existence de $M \otimes_A N$, et non la construction ci-dessus.

De plus, pour les espaces vectoriels de dimension finie sur un corps $k = A$ il existe une autre construction de $M \otimes_k N$ comme $Bil(M^* \times N^*, k) \cong k^{mn}$ où $M \cong k^m$, $N \cong k^n$.

Remarque

a) On définit le produit $M_1 \otimes_A M_2 \otimes_A \cdots \otimes_A M_n$ à partir d'applications A-multilinéaire avec une propriété universelle analogue à 3.1.1, b) (formuler en exercice).

b) Si $M \cong M'$, $N \cong N'$, alors $M \otimes_A N \cong M' \otimes_A N'$,

3.2 Exemples et propriétés du produit tensoriel

Exemple 3.2.1 *Soit* $A = \mathbb{Z}$, $M = \mathbb{Z}/m\mathbb{Z}$, $N = \mathbb{Z}/n\mathbb{Z}$. *Alors* $(\mathbb{Z}/m\mathbb{Z}) \otimes (\mathbb{Z}/n\mathbb{Z}) \cong \mathbb{Z}/d\mathbb{Z}$ *avec* $d = \text{pgcd}(m,n)$.

En effet, il suffit de vérifier la propriété universelle pour l'application

$$(\mathbb{Z}/m\mathbb{Z}) \otimes (\mathbb{Z}/n\mathbb{Z}) \to \mathbb{Z}/d\mathbb{Z},\ g(a,b) = ab \bmod d.$$

pour toute application $\mathbb{Z}$-bilinéaire $f : (\mathbb{Z}/m\mathbb{Z}) \times (\mathbb{Z}/n\mathbb{Z}) \to P$. On pose $f(\overline{1},\overline{1}) = p \in P$, alors $f(\overline{a},\overline{b}) = abp$.

De plus $f(\overline{m},\overline{1}) = f(\overline{1},\overline{n}) = 0 = mp = np$. Ceci dit, $dp = 0$ puisque $d = mu + nv$ $(u,v \in \mathbb{Z})$, et f se factorise par

$$\begin{array}{ccc} (\mathbb{Z}/m\mathbb{Z}) \times (\mathbb{Z}/n\mathbb{Z}) & \xrightarrow{\ g\ } & \mathbb{Z}/d\mathbb{Z} \\ & \searrow^{f} & \downarrow f' \\ & & P \end{array} \quad , \quad (\overline{1},\overline{1}) \mapsto \overline{1} \bmod d$$

$\overline{1}(\text{mod}\, d) \mapsto p$ est unique

Proposition 3.2.2 (propriétés du produit tensoriel)

Soient M, N, P *trois* A*-modules. Il existe les isomorphismes canoniques*

a)

$M \otimes_A N \cong N \otimes_A M,\ x \otimes y \longmapsto y \otimes x,$

b)

$(M \otimes_A N) \otimes_A P \cong M \otimes_A (N \otimes_A P) \xrightarrow{\sim} M \otimes_A N \otimes_A P,$
$(x \otimes y) \otimes z \longmapsto x \otimes (y \otimes z) \longmapsto x \otimes y \otimes z$

c)

$(M \oplus N) \otimes_A P \cong (M \otimes_A P) \oplus (N \otimes_A P),$
$(x \oplus y) \otimes z \longmapsto (x \otimes z) \oplus (y \otimes z),$

d)

$A \otimes_A M \cong M, a \otimes x \longmapsto ax,$

Preuve. Dans tous les cas on vérifie que les applications existent (elles sont bien définies), et qu'elles sont détérminées.

Cours N 2. Lundi 26 janvier 2015

REMARQUE . Le produit tensoriel $T = M \otimes_A N \otimes_A P$ est donné par la propriété universelle d'applications A-trilinéaires $f : M \times N \times P \to Q$ de A-modules.

a) On construit d'abord les A-homomorphismes
$f : M \otimes_A N \cong N \otimes_A M,\ x \otimes y \longmapsto y \otimes x,$
$g : N \otimes_A M \cong M \otimes_A N,\ y \otimes x \longmapsto x \otimes x.$

En effet, l'application $\phi : (x, y) \mapsto y \otimes x$ est A-bilinéaire puisque $\phi(ax, y) = \phi(x, ay) = a(y \otimes x)$ donc il existe un seul A-homomorphisme $f(x \otimes y) = y \otimes x$; la même chose pour g.

La composée $f \circ g \ : N \otimes_A M \to N \otimes_A M$, $x \otimes y \longmapsto x \otimes y$ est donc $\mathrm{Id}_{N \otimes_A M}$.
Puis, $g \circ f \ : M \otimes_A N \to M \otimes_A N$ est $\mathrm{Id}_{M \otimes_A N}$.

On construit les A-homomorphismes
$h : (M \otimes_A N) \otimes_A P \to M \otimes_A N \otimes_A P$, $k : M \otimes_A N \otimes_A P \otimes_A P \to (M \otimes_A N) \otimes_A P$, en utilisant de nouveaux les propriétés universelles : l'application A-bilinéaire $\psi : (x, y) \mapsto x \otimes y \otimes z$ (pour tout $z \in P$, induit

$$h_z : (M \otimes_A N \to M \otimes_A N \otimes_A P, h_z(x \otimes y) = x \otimes y \otimes z.$$

Puis $\chi : (t, z) \mapsto h_z(t)$, $(M \otimes_A N) \otimes_A P \to M \otimes_A N \otimes_A P$ est A-linéaire.
A partir de $\chi : (M \otimes_A N) \otimes_A P \to M \otimes_A N \otimes_A P$ on obtient une application cherchée

$$h : (M \otimes_A N) \otimes_A P \to M \otimes_A N \otimes_A P.$$

Pour construire $k : M \otimes_A N \otimes_A P \to (M \otimes_A N) \otimes_A P$, on considère l'application A-trilinéaire

$$\varkappa : M \otimes_A N \otimes_A P \to (M \otimes_A N) \otimes_A P, \varkappa(x, y, z) = (x \otimes y) \otimes z,$$

qui induit $k : M \otimes_A N \otimes_A P \otimes_A P \to (M \otimes_A N) \otimes_A P$ ci-dessus La construction implique : $h \circ k = \mathrm{Id}_{M \otimes_A N \otimes_A P}$, $k \circ h = \mathrm{Id}_{(M \otimes_A N) \otimes_A P}$ (sur les générateurs !).
Ceci dit, h et k sont des isomorphismes de A-modules.
PREUVE de c) et d) de Proposition 3.2.2 est en exercice à faire.

4 Algèbre symétrique d'un module

4.1 Applications multilinéaires symétriques

DÉFINITION 4.1.1 (APPLICATION MULTILINÉAIRE SYMÉTRIQUE) *Une application r-multilinéaire $\varphi : M \times \cdots \times M \to N$ est dit symétrique, si pour toute permutation $\sigma = \begin{pmatrix} 1 & 2 & \cdots & r \\ i_1 & i_2 & \cdots & i_r \end{pmatrix}$ et pour tous $x_1, \cdots, x_r \in M$ on a*

$$\varphi(x_1, \cdots, x_r) = \varphi(x_{i_1}, \cdots, x_{i_r})$$

Exemple .

a) Soit $M = N = A$, on pose

$$\varphi(x_1, \cdots, x_r) = x_1 \cdot \ldots \cdot x_r$$

(le produit dans l'anneau A). Alors φ est multilinéaire et symétrique puisque A est supposé commutatif.

b) Soit $M = A[X, Y]$, $N = A[X]$, $x_i = f_i(X, Y) \in M$. On pose

$$\varphi(f_1, \cdots, f_r) = (f_1 \cdot \ldots \cdot f_r)(X, 0) \in N.$$

c) Soient $l_1, l_2, \cdots, l_r : M \to A, N = A$. On pose

$$\varphi(x_1, \cdots, x_r) = \sum_{\sigma = \begin{pmatrix} 1 & 2 & \cdots & r \\ i_1 & i_2 & \cdots & i_r \end{pmatrix} \in S_r} l_1(x_{i_1}) \cdot \ldots \cdot l_r(x_{i_r})$$

Alors l'application φ est r-multilinéaire et symétrique

Proposition 4.1.2 (l'existence et l'unicité du produit symétrique multiple)

a) Il existe un A-module $T = S^r(M)$ et un A-homomorphisme symétrique $\mu : M \times \cdots \times M \to T$ telle que toute application r-multilinéaire symétrique $\varphi : M \times \cdots \times M \to N$ se factorise de façon unique par μ : il existe une unique application A-linéaire $f : T \to N$ telle que $\varphi = f \circ \mu$:

$$\begin{array}{ccc} M \times \cdots \times M & \xrightarrow{\mu} & T \\ & \searrow^{\varphi} & \swarrow_{f} \\ & N & \end{array} \tag{4.1}$$

b) L'unicité de $T = S^r(M)$ (à un A-isomorphisme près).

Preuve de b) découle directement de la propriété universelle (4.2) comme dans Proposition 3.1.3 (en exercice à faire).

Preuve de a) *L'existence de* $(S^r(M), \mu)$: on pose d'abord

$$M^{\otimes r} = M \otimes_A \cdots \otimes_A M \quad (r \text{ fois}), \quad m(x_1, \cdots, x_r) = x_1 \otimes \cdots \otimes x_r,$$

et on définit

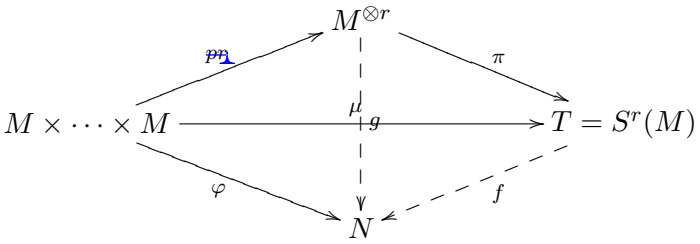
$$S^r(M) = M^{\otimes r} / \left\langle x_1 \otimes \cdots \otimes x_r - x_{i_1} \otimes \cdots \otimes x_{i_r} \mid \sigma = \begin{pmatrix} 1 & 2 & \cdots & r \\ i_1 & i_2 & \cdots & i_r \end{pmatrix} \in S_r, x_i \in M \right\rangle,$$

(on factorise pas le sous-A-module R_{sym} engendré par les relations de symétrie). Puis on pose

$$\mu(x_1, \cdots, x_r) = \text{ la classe } (x_1 \otimes \cdots \otimes x_r + R_{sym}) \in S^r(M).$$

Vérification de la propriété universelle (4.2) : pour toute application A-multilinéaire symétrique $\varphi : M \times \cdots \times M \to N$ il existe $g : M^{\otimes r} \to N$ (un unique homomorphisme de A-modules) tel que $\varphi(x_1, \cdots, x_r) = g(x_1 \otimes \cdots \otimes x_r)$, $\varphi = g \circ m$.

Par la symétrie de φ, $\operatorname{Ker} g \supset R_{sym}$ puisque $g(x_1 \otimes \cdots \otimes x_r - x_{i_1} \otimes \cdots \otimes x_{i_r}) = \varphi(x_1, \cdots, x_r) - \varphi(x_{i_1}, \cdots, x_{i_r}) = 0$, alors il existe un seul homomorphisme de A-modules $f : S^r(M) = M^{\otimes r}/R_{sym} \longrightarrow N$ tel que $g = f \circ \pi$, où $\pi : M^{\otimes r} \to S^r(M)$ la projection naturelle.



(4.2)

De plus, $\mu = \pi \circ m$, donc

$$f \circ \mu = f \circ \pi \circ m = g \circ m = \varphi,$$

puisque $g = f \circ \pi$. Ceci dit, f est un unique homomorphisme de A-modules tel que $f \circ \mu = \varphi$ ∎

4.2 Définition et propriétés de l'algèbre symétrique d'un module

NOTATION

$$S(M) = \bigoplus_{r \geq 0} S^r(M), \ S^0(M) = A, \tag{4.3}$$

$$T(M) = \bigoplus_{r \geq 0} T^r(M), \text{ où } T^r(M) = M^{\otimes r} = \overbrace{M \otimes_A \cdots \otimes_A M}^{r \text{ fois}}. \tag{4.4}$$

On montre que $S(M)$ est une A-algèbre commutative et que $T(M)$ est une A-algèbre non-commutative.

Rappels

DÉFINITION 4.2.1 *Soit A un anneau commutatif. Une A-algèbre B est un anneau fourni d'un homomorphisme d'anneau $\varphi : A \to B$ tel que pour tout $a \in A$ et pour tout $b \in B$ on a $\varphi(a)b = b\varphi(a)(= ab)$.*

EXEMPLE

a) Soit $A = \mathbb{R}$, $B = \mathbb{C}$, $\varphi : \mathbb{R} \to \mathbb{C}$ l'inclusion d'anneau. Alors $\mathbb{C}$ est une $\mathbb{R}$-algèbre.

b) *L'algèbre des quaternions (de Cayley)*

$$\mathbb{H} = \Big\{ a + bi + cj + dk \;\Big|\; a,b,c,d \in \mathbb{R}, ij = -ji = k, jk = -kj = i, ki = -ik = j, \\ i^2 = j^2 = k^2 = -1 \Big\}$$

Alors $\varphi : a \mapsto a + 0i + 0j + 0k$ est un homomorphisme, $\mathbb{H}$ est une $\mathbb{R}$.

Attention :

$\mathbb{H}$ n'est pas une $\mathbb{C}$-algèbre (pour $\psi : \mathbb{C} \to \mathbb{H}, a + bi \mapsto a + bi + 0j + 0k$).

PROPOSITION 4.2.2 (DESCRIPTION DU PRODUIT) *Pour tout r_1, $r_2 \in \mathbb{N}$, il existe une unique application A-bilinéaire*

$$t : T(M)^{r_1} \times T^{r_2}(M) \to T^{r_1+r_2},$$

telle que

$$t(x_1 \otimes \cdots \otimes x_{r_1}, y_1 \otimes \cdots \otimes y_{r_2}) = x_1 \otimes \cdots \otimes x_{r_1} \otimes y_1 \otimes \cdots \otimes y_{r_2},$$

et il existe un unique application A-bilinéaire

$$s : S(M)^{r_1} \times S^{r_2}(M) \to S^{r_1+r_2},$$

telle que

$$s(x_1 \cdot \ldots \cdot x_{r_1}, y_1 \cdot \ldots \cdot y_{r_2}) = x_1 \cdot \ldots \cdot x_{r_1} \cdot y_1 \cdot \ldots \cdot y_{r_2},$$

où $x_1 \cdot \ldots \cdot x_r = \mu(x_1, \cdots, x_r)$.

b) Les applications $t = t_{r_1,r_2}, s = s_{r_1,r_2}$ fournissent

$$T(M) = \bigoplus_{r \geq 0} T^r(M), S(M) = \bigoplus_{r \geq 0} S^r(M)$$

des structures de A-algèbres telles que $T^0(M) = S^0(M) = A$ donnent les morphismes de structure

$$A \to T(M), A \to S(M).$$

4.3 Exemples de l'algèbre symétrique

PROPOSITION 4.3.1 (STRUCTURE DE L'ALGÈBRE SYMÉTRIQUE) *Soit $M = \langle x_\alpha \rangle_{\alpha\in\Gamma}$ un A-module engendré par une famille d'éléments $x_\alpha \in M$, avec un ensemble Γ totalement ordonné, par exemple $\Gamma = \{1, 2, \cdots, n\}$, alors*

$$S^r(M) = \langle x_{\alpha_1} \cdot \ldots \cdot x_{\alpha_r} \mid (\alpha_1, \cdots, \alpha_r) \in \Gamma^r \rangle$$

(avec un système de générateurs redondant), de plus

$$S^r(M) = \langle x_{\beta_1}^{n_1} \cdot \ldots \cdot x_{\beta_s}^{n_s} \mid , n_1 + \cdots + n_s = r, \beta_1, \cdots, \beta_s \in \Gamma \text{ distincts avec } \beta_1 < \cdots < \beta_s \rangle$$

(un système réduit de générateurs).

b) Si $M = A^{(\Gamma)}$ est libre alors $S(M) = \bigoplus_{r\geq 0} S^r(M)$ est une A-algèbre commutative isomorphe à l'anneau des polynômes (commutatifs) sur A :

$$S(M) \cong A[X_\alpha]_{\alpha\in\Gamma}.$$

EXERCICE Montrer que si $M = A^m$ est libre de rang m, alors

$$S^r(M) \cong A^{\binom{m+r-1}{r}}.$$

5 Algèbre extérieure d'un module

5.1 Application multilinéaires alternées

DÉFINITION 5.1.1 (APPLICATION MULTILINÉAIRE ALTERNÉES) *Une application r-multilinéaire $\varphi : M \times \cdots \times M \to N$ est dite alternée, si $\varphi(x_1, \cdots, x_r) = 0$ pour tous les $x_i \in M$ avec la propriété qu'un élément paraît plus qu'une fois : dans ce cas*

$$\varphi(x_1, \cdots, x_{i-1}, x_i + x_j, x_{i+1}, \cdots, x_{j-1}, x_i + x_j, x_{j+1}, \cdots) = 0$$

donc

$$\begin{aligned} &\varphi(x_1, \cdots, x_{i-1}, x_i, x_{i+1}, \cdots, x_{j-1}, x_j, x_{j+1}, \cdots) + \\ &\varphi(x_1, \cdots, x_{i-1}, x_j, x_{i+1}, \cdots, x_{j-1}, x_i, x_{j+1}, \cdots) = 0. \end{aligned}$$

Ceci implique que pour toute permutation $\sigma = \begin{pmatrix} 1 & 2 & \cdots & r \\ i_1 & i_2 & \cdots & i_r \end{pmatrix}$ et pour tous $x_1, \cdots, x_r \in M$ on a

$$\varphi(x_1, \cdots, x_r) = \varepsilon(\sigma)\varphi(x_{i_1}, \cdots, x_{i_r}),$$

c'est-à-dire, que φ est antisymétrique.

~~REMARQUE Si $A = \mathbb{F}_2$, $M = N$, alors $\varphi(x_1, \cdots, x_r) = x_1 + \cdots + x_r$ n'est pas alternée, mais φ est antisymétrique!~~

PROPOSITION 5.1.2 (L'EXISTENCE ET L'UNICITÉ DU PRODUIT ALTERNÉE MULTIPLE)
a) Il existe un A-module $T = \wedge^r(M)$ et un A-homomorphisme alterné $\lambda_r : M \times \cdots \times M \to T$ tel que toute application r-multilinéaire alternée $\varphi : M \times \cdots \times M \to N$ se factorise de façon unique par λ : il existe une unique application A-linéaire $f : T \to N$ telle que $\varphi = f \circ \lambda$:

$$\begin{array}{ccc} M \times \cdots \times M & \xrightarrow{\lambda_r} & T \\ & \searrow^{\varphi} \quad \swarrow_{f} & \\ & N & \end{array} \tag{5.1}$$

b) L'unicité de $T = \wedge^r(M)$ (à un A-isomorphisme près).

PREUVE de a) *L'existence de* $(\wedge^r(M), \lambda_r)$: on utilise de nouveaux

$$M^{\otimes r} = M \otimes_A \cdots \otimes_A M \quad (r \text{ fois}), \quad m(x_1, \cdots, x_r) = x_1 \otimes \cdots \otimes x_r,$$

et on définit

$$\wedge^r(M) = M^{\otimes r} / \left\langle x_1 \otimes \cdots \otimes x_{i-1} \otimes x \otimes x_{i+1} \otimes \cdots \otimes x_{j-1} \otimes x \otimes x_{j+1} \otimes \cdots \right| x_l, x \in M \rangle,$$

(on factorise pas le sous-A-module R_{alt} engendré par les relations d'alternance). Puis on pose

$$\lambda_r(x_1, \cdots, x_r) = \text{ la classe } (x_1 \otimes \cdots \otimes x_r + R_{alt}) \in \wedge^r(M).$$

Vérification de la propriété universelle (4.2) : pour toute application A-multilinéaire alternée $\varphi : M \times \cdots \times M \to N$ il existe $g : M^{\otimes r} \to N$ (un unique homomorphisme de A-modules) tel que $\varphi(x_1, \cdots, x_r) = g(x_1 \otimes \cdots \otimes x_r)$, $\varphi = g \circ m$. ∎

5.2 Définition et propriétés de l'algèbre extérieure d'un module

Pour définir l'algèbre extérieure, on pose

$$\wedge(M) = \bigoplus_{r \geq 0} \wedge^r(M), \text{ où } \wedge^0(M) = A, \wedge^1(M) = M. \tag{5.2}$$

On montre que $\wedge(M)$ est une A-algèbre non-commutative (en général). On définit pour tout r, $s \in \mathbb{N}$, une unique application A-bilinéaire

$$\lambda_{r,s} : \wedge(M)^r \times \wedge^s(M) \to \wedge^{r+s},$$

telle que

$$\lambda_{r,s}(x_1 \wedge \cdots \wedge x_r, y_1 \wedge \cdots \wedge y_s) = x_1 \wedge \cdots \wedge x_r \wedge y_1 \wedge \cdots \wedge y_s,$$

b) Les applications $\lambda = \lambda_{r,s}$ fournissent

$$\wedge(M) = \bigoplus_{r \geq 0} \wedge^r(M)$$

de structure d'une A-algèbre telles que $\wedge^0(M) = A$ donnent les morphismes de structure

$$A \to \wedge(M).$$

NOTATION . Pour $x \in \wedge^r(M)$ et $y \in \wedge^s(M)$ on pose $x \wedge y = \lambda_{r+s}(x, y) \in \wedge^{r+s}(M)$.

PROPOSITION 5.2.1 (PROPRIÉTÉ D'ALGÈBRE EXTÉRIEURE) *Soit* $x \in \wedge^r(M)$, $y \in \wedge^s(M)$, $z \in \wedge^t(M)$. *Alors*

a) $x \wedge y = (-1)^{r \cdot s} y \wedge x$ *(la symétrie gauche du produit extérieure)*

b) $(x \wedge y) \wedge z = x \wedge (y \wedge z)$ *(l'associativité du produit extérieure)* ∎

PROPOSITION 5.2.2 (PUISSANCE EXTÉRIEURE) *Soit* $M \cong A^r$ *libre de base* $e_1, \cdots, e_r \in M$.

Alors $\wedge^r(M) = \langle e_{i_1} \wedge \cdots \wedge e_{i_r} \mid 1 \leq i_1 < \cdots < i_r \leq n \rangle$ *est libre de rang* $\binom{n}{r}$. ∎

EXERCICE Soit $A = k[x, y]$, k un corps, $I = (x, y) = Ax + Ay$. Alors $\wedge^2 I \neq 0$: il existe une application $\varphi : I \times I \to A/I$ alternée non nulle : $\varphi(x, y) = \frac{\partial(f,g)}{\partial(x,y)} = \frac{\partial f}{\partial x}\frac{\partial g}{\partial y} - \frac{\partial f}{\partial y}\frac{\partial g}{\partial x}$.

RAPPELS : **5 Algèbre extérieure d'un module**

Deuxième partie
Polynômes et fractions rationnelles

6 Polynômes à une variable

6.1 Anneau de polynômes

Sur un corps fini, il convient de distinguer les polynômes et les fonctions polynômes. Nous revenons donc sur la définition des polynômes.

DÉFINITION 6.1.1 *Si A un anneau commutatif, l'anneau des polynômes à une variable X sur A est l'anneau $A[X]$ formé des suites $(a_i)_{i\in\mathbb{N}}$ d'élément de A telles que $a_i = 0$ sauf un nombre fini d'entiers i. Cet ensemble est muni de la somme*

$$(a_i)_{i\in\mathbb{N}} + (b_i)_{i\in\mathbb{N}} = (a_i + b_i)_{i\in\mathbb{N}}$$

et du produit

$$(a_i)_{i\in\mathbb{N}} \times (b_i)_{i\in\mathbb{N}} = \left(\sum_{i=j+k} a_j b_k\right)_{i\in\mathbb{N}}$$

On a une application injective $A \to A[X]$ qui envoie a sur $(a, 0, \cdots)$, et on identifie A avec son image. Tout élément de $A[X]$ s'écrit de façon unique $\sum_{i\in\mathbb{N}} a_i X^i$, où on note par X la suite :

$$(0, 1, 0, 0, \cdots).$$

Soit A un anneau. Il est commode de voir un polynôme $f(X)$ à coefficients dans A une expression formelle du type

$$f(X) = a_n X^n + a_{n-1} X^{n-1} + \cdots + a_0$$

qui est donnée par la suite de ces coefficients $a_0, a_1, \ldots, a_n \in A$, $n \in \mathbb{N}$ telle que presque tout a_n (sauf un nombre fini) est nul.

Si tous les coefficients a_i sont nuls on appelle $f(X)$ un polynôme nul : $f = 0$. Si $f(X)$ est non-nul, alors $a_n \neq 0$ pour un n.

DÉFINITION 6.1.2 *Le plus grand indice n avec cette propriété est appele le* degré *de $f(X)$ et il est noté* $\deg f$. *Le degré du polynîme nul n'est pas défini, mais parfois on pose* $\deg 0 = -\infty$.

L'anneau $A[X]$ est défini donc comme l'ensemble des expressions $f(X)$ ci-dessus avec des opérations données par les règles suivantes : si

$$f(X) = a_n X^n + a_{n-1} X^{n-1} + \cdots + a_0, \quad a_n \neq 0$$

$$g(X) = b_s X^s + b_{s-1} X^{s-1} + \cdots + b_0, \quad b_s \neq 0$$

deux polynômes et si par exemple $n \geq s$, on appelle leur somme le polynôme

$$(f+g)(X) = f(X) + g(X) = c_n X^n + c_{n-1} X^{n-1} + \cdots + c_0,$$

dont les coefficients sont obtenus par l'addition des coefficients correspondants de X dans f et dans g, c'est à dire $c_i = a_i + b_i$, $i = 0, 1, \ldots, n$, où pour $i > s$ les coefficients b_i sont considérés comme nuls.

Le produit des polynômes $f(X)$ et $g(X)$ est le polynôme

$$(fg)(X) = f(X) \cdot g(X) = d_{n+s} X^{n+s} + d_{n+s-1} X^{n+s-1} + \cdots + d_0,$$

où

$$d_i = \sum_{k+l=i} a_k b_l,$$

et les coefficients a_i, b_j sont considérés comme zéros pour $i > n$, et $j > s$, $d_0 = a_0 b_0$, $d_1 = a_0 b_1 + a_1 b_0, \ldots, d_{n+s} = a_n b_s$.

THÉORÈME 6.1.3 *Si l'anneau A n'a pas de diviseurs de zéro, l'anneau des polynômes $A[X]$ aussi n'a pas de diviseurs de zéro. Le degré du produit des polynômes non nuls est égal à la somme de ces degrés.*

La démonstration est directement impliquée par les formules ci-dessus, en particulièr $d_{n+s} = a_n b_s$, où $n = \deg f$, $s = \deg g$.

L'anneau A peut être identifié à un sous-anneau de $A[X]$ formé par des constantes (polynômes de degré nul et polynôme nul). Ceci implique que la multiplication des élémets $a \in A$ par $f(X) \in A[X]$ est aussi définie. En particulier, si $A = K$ est un corps, l'anneau $K[X]$ est aussi un éspace vectoriel. Du point de vue de la structure algébrique, l'anneau $K[X]$ devient un algèbre de dimension infinie sur K, c'est à dire un anneau et un espace vectoriel en même temps dans lequel la multiplication d'éléments commute avec la multiplication par des constantes.

6.2 Division euclidienne sur les anneaux

Division des polynômes avec reste

PROPOSITION 6.2.1 *Soit A un anneau commutatif intègre. On se donne un polynôme*

$$P(X) = \sum_{i=0}^{d} a_i X^i$$

à coefficients dans A tel que a_d soit un élément inversible de A. *Alors pour tout polynôme $f(X)$ de $A[X]$ il existe une unique paire $(Q, R) \in A[X]^2$ telle que*

$$f = PQ + R \text{ avec } R = 0 \text{ ou } \deg R < \deg P$$

PREUVE de l'existence. Nous allons procéder par récurrence sur le degré de f. Si $\deg f < d$, alors $(0, f)$ convient. Sinon, on écrit

$$f = \sum_{i=0}^{n} b_i X^i \text{ avec } b_n \neq 0 \text{ et } n \geq d.$$

Alors

$$f - b_n a_d^{-1} X^{n-d} P = PQ_1 + R_1 \text{ avec } R_1 = 0 \text{ ou } \deg R_1 < d.$$

La paire $(Q_1 + b_n a_d^{-1} X^{n-d}, R_1)$ convient.

Unicité. Si

$$PQ_0 + R_0 = PQ_1 + R_1$$

avec $\deg R_0 < \deg P$ et $\deg R_1 < \deg P$ alors $P(Q_0 - Q_1) = (R_1 - R_0)$. Comme le coefficient dominant de P est inversible, on a $\deg(P(Q_0 - Q_1)) = \deg P + \deg(Q_0 - Q_1)$ mais ce degré est strictement inférieur à celui de P si et seulement si $Q_0 - Q_1 = 0$, c'est-à-dire, $Q_0 = Q_1$, ce qui entraîne que $R_0 = R_1$.

EXEMPLE Illustrons sur un exemple la façon d'effectuer une telle division dans $\mathbb{Z}[X]$:

$$\begin{array}{r|l} 3X^4 + 7X^3 - 7X^2 + 16X - 5 & X^2 + 3X - 2 \\ \cline{2-2} -3X^4 - 9X^3 + 6X^2 & 3X^2 - 2X + 5 \\ \cline{1-1} \quad -2X^3 - X^2 + 16X - 5 & \\ \quad 2X^3 + 6X^2 - 4X & \\ \cline{1-1} \quad 5X^2 + 12X - 5 & \\ \quad -5X^2 - 15X + 10 & \\ \cline{1-1} \quad -3X + 5 & \end{array}$$

Division euclidienne dans $K[x]$ sur un corps K

THÉORÈME 6.2.2 *Pour tous les polynômes $f(X)$ et $P(X)$ tels que $P(X)$ soit non nul à coefficients dans un corps K, il existe une unique paire $(Q, R) \in A[X]^2$ telle que*

$$f = PQ + R \text{ avec } R = 0 \text{ ou } \deg R < \deg P$$

Les polynômes $Q(X)$ et $R(X)$ sont uniquement déterminés par cette condition.

DÉFINITION 6.2.3 *Le polynôme $Q(X)$ est appellé le* quotient *de la division $f(X)$ par $Q(X)$, et $R(X)$ le* reste.

Divisibilité des polynômes

Soit K un corps. Soient $f(X)$, $\phi(X) \in K[X]$. Si le reste de la division de $f(X)$ par $\phi(X)$ est nul, on dit que $f(X)$ est divisible par $\phi(X)$ ou aussi que $\phi(X)$ divise $f(X)$, la notation : $\phi \mid f$. La condition $\phi \mid f$ est équivalente au fait qu'il existe un polynôme $\psi(X)$ tel que $f(X) = \phi(X) \cdot \psi(X)$.

La définition implique directement les propriétés suivantes de la divisibilité :

1) Si f est divisible par g, et g est divisible par h, f est divisible par h.

2) Si f et g sont divisibles par ϕ, leur somme et leur différence sont divisibles par ϕ.

3) Tout polynôme est divisible par n'importe quel polynôme de degré zéro.

4) $f(X)$ divise $g(X)$ et $g(X)$ divise $f(X)$ en même temps si et seulement si $g(X) = cf(X)$, où $c \in K^*$ est un élément inversible.

5) Les ensembles de diviseurs de $f(X)$ et $cf(X)$ coïncident.

On appelle le pgcd (plus grand diviseur commun) de $f(X)$ et $g(X)$ le polynôme $d(X)$ tel que d divise f et g, et d est divisible par tout autre diviseur commun de ces polynômes.

THÉORÈME 6.2.4 *Pour tous les polynômes f et g dans $K[X]$ sur un corps K il existe un pgcd correspondant uniquement déterminé à une constante multiplicative près.*

C'est une propriété générale dans les anneaux euclidiens.

6.3 Valeurs et racines d'un polynôme

PROPOSITION 6.3.1 *Soit A un anneau commutatif. On se donne un polynôme*

$$f(X) = \sum_{i=0}^{n} a_i X^i$$

à coefficients dans A, et soit c un élément de A. Alors la valeur de f en c est définie comme $f(c) = a_n c^n + a_{n-1} c^{n-1} + \cdots + a_0 \in A$.

(a) L'application

$$\psi_c : A[X] \to A, f \longmapsto f(c) = \sum_{i=0}^{n} a_i c^i \in A$$

est un seul morphisme d'anneaux tel que la réstriction de ψ_c sur le sous-anneau $A \subset A[X]$ soit triviale, et que $\psi_c(X) = c$.

(b) Réciproquement, pour tout morphisme d'anneaux $\psi : A[X] \to A$ tel que la réstriction de ψ sur le sous-anneau $A \subset A[X]$ est triviale, il existe un seul $c \in A$ tel que $\psi = \psi_c$.

PREUVE de la proposition 6.3.1 découle de la définition 1.1.2 de morphisme d'anneaux.

Fonctions polynômes.

DÉFINITION 6.3.2 *(a) L'application*

$$f : A \to A, c \longmapsto f(c)$$

notée par la même lettre f, est dite la fonction polynôme.

(b) Si $f(c) = 0$ (c'est à dire, que f s'annule en c), on appelle c racine *de f.*

REMARQUE . La fonction polynôme en général ne définit pas le polynôme de manière unique. Par exemple, sur un corps fini K, il convient de distinguer les polynômes et les fonctions polynômes.

Soit par exemple $A = K = \mathbb{F}_2 = \{0, 1\}$ (le corps de deux éléments). Considérons le polynôme $f(X) = X^2 + X + 1$, alors $f(0) = 1$ et $f(1) = 1$ c'est à dire f definie une fonction constante sur K mais $f(X)$ n'est pas une constante (polynôme de degré zéro) comme un polynôme.

THÉORÈME 6.3.3 *Soit $A = K$ un corps. Si $f(X) \in K[X]$, le reste de la division de $f(X)$ par $(X - c)$ est égal à $f(c)$. En particulier, il s'agit d'une racine c de $f(X)$ si et seulement si $f(X)$ est divisible par $X - c$.*

Démonstration impliquée par l'unicité de la division avec reste :

$$f(X) = q(X)(X - c) + r,$$

où $r = f(c)$.

COROLLAIRE 6.3.4 *Si $f(X) \in K[X]$, le nombre des racines $c \in K$ est majoré par le degré* $\deg(f)$.

Démonstration est impliquée par l'unicité de la décomposition en facteurs irréductibles dans un anneau euclidien.

Méthode d'Hörner

permet de trouver facilement le quotient $q(X)$ de la division par $X - c$

$$f(X) = q(X)(X - c) + r,$$

et la valeur $r = f(c)$.

On considère le polynôme
$f(X) = \sum_{i=0}^{n} a_i X^i$, et sa valeur en $X = c$; on calcule $f(c)$ de façon suivante : soient

$f(X) = a_n X^n + a_{n-1} X^{n-1} + \cdots + a_0,\ a_n \neq 0$ un polynôme, et on cherche un autre polynôme $q(X) = b_{n-1} X^{n-1} + b_{n-2} X^{n-2} + \cdots + b_0,\ \ b_{n-1} \neq 0$ tel que

$$f(X) = (X - c)q(X) + r,$$

En comparant les coefficients des puissances de X on obtient

$$a_n = b_{n-1},\ a_{n-1} = b_{n-2} - cb_{n-1},\ \cdots,$$

$$a_1 = b_0 - cb_1,\ a_0 = r - cb_0,\ r = f(c).$$

Ceci implique

$$b_{n-1} = a_n,\ b_{n-k} = cb_{n-k+1} + a_{n-k} (k = 2,\ \cdots, n)$$

On peut faire le tableau suivant (le schéma de Hörner)

	a_n	a_{n-1}	$\cdots$	a_1	a_0
c	$b_{n-1} = a_n$	$b_{n-2} = cb_{n-1} + a_{n-1}$	$\cdots$	$b_0 = cb_1 + a_1$	$f(c) = cb_0 + a_0$

La formule de Taylor

Soit K un corps. On se donne un polynôme

$$f(X) = \sum_{i=0}^{n} a_i X^i$$

à coefficients dans K, et soit c un élément de K. Alors il existe $b_1, \cdots, b_n \in K$ tels que

$$f(X) = f(c) + b_1(X-c) + b_2(X-c)^2 + \cdots + b_n(X-c)^n,$$

avec la propriété

$$k! b_k = f^{(k)}(c), \quad k = 1, \cdots, n, \tag{6.1}$$

où

$$f'(X) = n a_n X^{n-1} + (n-1) a_{n-1} X^{n-2} + \cdots + a_1$$

est la dérivée formelle de $f(X)$.

PREUVE. L'existence des $b_1, \cdots, b_n \in K$ découle de la division euclidienne dans $K[X]$, par récurrence à partir de

$$f(X) = (X-c)q(X) + f(c).$$

Ensuite, on deduit la formule (6.1) par récurrence à partir de l'identité :

$$f(X) = f(c) + b_1(X-c) + b_2(X-c)^2 + \cdots + b_n(X-c)^n,$$

en utilisant l'égalité formelle $((X-c)^k)' = k(X-c)^{k-1}$.

6.4 Formule d'interpolation de Lagrange

La formule d'interpolation de Lagrange donne un polynôme sur un corps K de degré inférieur ou égal à n qui prend pour les valeurs distinctes de la variable X en α_0, α_1, ..., $\alpha_n \in K$, les valeurs β_0, β_1, ..., $\beta_n \in K$. La solution est donnée par le polynôme de Lagrange

$$f(X) = \sum_{i=0}^{n} \beta_i \frac{(X-\alpha_0)\ldots(X-\alpha_{i-1})(X-\alpha_{i+1})\ldots(X-\alpha_n)}{(\alpha_i-\alpha_0)\ldots(\alpha_i-\alpha_{i-1})(\alpha_i-\alpha_{i+1})\ldots(\alpha_i-\alpha_n)} \tag{6.2}$$

En effet son degré et inferieur ou égal à n, et $f(\alpha_i) = \beta_i$ $(i = 0, 1, \ldots, n)$.

PROPOSITION 6.4.1 *Le polynome de Lagrange (6.2) est l'unique polynôme*

$$f(X) = a_n X^n + a_{n-1} X^{n-1} + \cdots + a_0 \in K[X]$$

de degré $\leq n$ *tel que* $f(\alpha_i) = \beta_i (i = 0, 1, \ldots, n)$ *pour arbitraires* distincts *éléments* α_i $(i = 0, 1, \ldots, n)$ *du corps* K *et pour arbitraires* $\beta_i \in K$ $(i = 0, 1, \ldots, n)$.

En effet si l'on considère les coefficients $a_j (j = 0, 1, \ldots, n)$ comme des inconnues, on obtiendrait le système des équations linéaires suivant :

$$\begin{cases} f(\alpha_0) = a_n \alpha_0^n + a_{n-1} \alpha_0^{n-1} + \cdots + a_0 = \beta_0 \\ f(\alpha_1) = a_n \alpha_1^n + a_{n-1} \alpha_1^{n-1} + \cdots + a_0 = \beta_1 \\ \ldots\ldots\ldots\ldots\ldots \\ f(\alpha_n) = a_n \alpha_n^n + a_{n-1} \alpha_n^{n-1} + \cdots + a_0 = \beta_n \end{cases}$$

Le déterminant de ce système carrée coïncide (à signe près) avec le déterminant de Vandermonde

$$\begin{vmatrix} 1 & 1 & \ldots & 1 \\ \alpha_0 & \alpha_1 & \ldots & \alpha_n \\ \alpha_0^2 & \alpha_1^2 & \ldots & \alpha_n^2 \\ \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots \\ \alpha_0^n & \alpha_1^n & \ldots & \alpha_n^n \end{vmatrix} = \prod_{0 \leq i < j \leq n} (\alpha_j - \alpha_i)$$

qui est non-nul, d'où l'existence et l'unicité des coefficients cherchés $a_j (j = 0, 1, \ldots, n)$ du polynôme $f(X)$ (dans le corps K).

PROPOSITION 6.4.2 *Soit* K *un corps* infini. *Alors un polynôme* $f(X)$ *est uniquement déterminé par la fonction polynôme correspondente* $c \mapsto f(c) : K \to K$.

REMARQUE . Au contraire, sur un corps fini K, toute fonction $f : K \to K$ est polynômiale, mais le polynôme correspondant f n'est pas uniquement déterminé par cette fonction, car il existe des polynômes non-nuls représentant la fonction identiquement nulle. Par exemple, le polynôme non-nul de degré p

$$f(X) = X^p - X = \prod_{a \in \mathbb{F}_p} (X - a)$$

s'annule en tous les points du corps $K = \mathbb{Z}/p\mathbb{Z}$.

6.5 Polynômes irréductibles.

Rappelons que si K est un corps, l'anneau de polynômes $K[X]$ est euclidien, donc principal et factoriel, en particulier tout polynôme non nul s'écrit de manière unique comme produit de son coefficient dominant et de polynômes irréductibles unitaires.

Et on obtient la définition suivante pour les éléments irréductibles de $K[X]$:

Définition 6.5.1 *Soit K un corps. Un polynôme $P \in K[X]$ est dit* irréductible*, s'il vérifie les deux conditions suivantes :*

Irr1. $P \notin K$

Irr2. *Si $P = QR$ avec $Q, R \in K[X]$ alors $Q \in K^\times$ ou $R \in K^\times$.*

Exemples 6.5.2 *Soit K un corps.*

(i) Un polynôme $P \in K[X]$ de degré un est irréductible.

(ii) Un polynôme $P \in K[X]$ de degré 2 ou 3 est irréductible si et seulement si il n'admet pas de racine dans K.

(iii) Tout polynôme irréductible sur $\mathbb{C}$ est de degré un.

(iv) Tout polynôme irréductible sur $\mathbb{R}$ est de degré ≤ 2.

Proposition 6.5.3 *Soit K un corps. Un polynôme $f \in K[X]$ est irréductible si et seulement si l'anneau quotient $K[X]/(f)$ est un corps.*

Exercices

6.1 Soit P un polynôme dans $A[X]$, A anneau commutatif. Montrer que $P(P(X)) - X$ est divisible par $P(X) - X$

6.2 Soit P un polynôme dans $\mathbb{R}[X]$ tel que $P(1) = 1$ et $P(2) = 4$. On pose $B(X) = X^2 - 3X + 2$. Déterminer le reste de la division de P par B.

6.3 Soit P un polynôme dans $\mathbb{Q}[X]$ tel que $P(X) = (X-a)^2(X-b)$ où $a, b \in \mathbb{C}$. En considérant $P'(X)$, montrer que $a, b \in \mathbb{Q}$.

6.4 Soit
$$f(X) = a_n X^n + a_{n-1} X^{n-1} + \cdots + a_0 \in \mathbb{Z}[X],$$
et p/q fraction irréductible telle que $P(p/q) = 0$. Montrer que q divise a_n, p divise a_0, et pour tout m, $p - mq$ divise $P(m)$.

6.5 Dans $\mathbb{R}[X]$ montrer qu'il existe un unique polynôme P de degré ≤ 7 tel que $P + 1$ soit divisible par $(X-1)^4$, et $P - 1$ soit divisible par $(X+1)^4$. Déterminer P

6.6 Soit $a \in \mathbb{Z}$, n un entier ≥ 2 tel que $\mathrm{pgcd}(a, n) = 1$ Montrer que n est premier si et seulement si les polynômes $(X + a)^n$, et $X^n + a$ sont congrus modulo n.

7 Fractions rationnelles

Outre les polynômes, on étudie encore en analyse les fonctions dites *fractions rationnelles*; ce sont les quotients de deux polynômes $\frac{f(x)}{g(x)}$, où $g(x) \neq 0$. On effectue les opérations algébriques sur ces fonctions d'après les mêmes règles qu'en arithmetique sur les nombres rationnels, c'est-à-dire d'après les règles d'opérations sur les fractions dont les numérateurs et dénominateurs sont entiers. L'identité de deux *fractions rationelles* a le même sens que l'identité des fractions en arithmétique.

7.1 Corps des fractions

PROPOSITION 7.1.1 *Si A est un anneau intègre, alors il existe un corps K appelé* corps des fractions *de A et noté $Frac(A)$ tel que*

(i) $A \subset K$;

(ii) Pour tout corps L et tout morphisme d'anneaux injectif $\phi : A \to L$, il existe un unique morphisme de corps $\psi : K \to L$ tel que $\psi|_A = \phi$.

PREUVE. Construction. On définit sur $A \times (A\backslash\{0\})$ la relation $\mathcal{R}$ par

$$(a,b)\mathcal{R}(c,d) \Longleftrightarrow ad = bc.$$

On vérifie en utilisant l'intégrité de A que $\mathcal{R}$ est une relation d'équivalence. On note K l'ensemble quotient $A \times A\backslash\{0\}/\mathcal{R}$, et $\frac{a}{b}$ l'image de (a,b) dans ce quotient. L'application de A dans K qui envoie a sur $(a,1) = \frac{a}{1}$ est injective, et on identifie A avec son image. On muni alors K des lois

$$+ : K \times K \to K, (\frac{a}{b}, \frac{c}{d}) \mapsto \frac{ad+bc}{bd} \in K \quad (\text{"addition"});$$
$$\times : K \times K \to K, (\frac{a}{b}, \frac{c}{d},) \mapsto \frac{ac}{bd} \in K \quad (\text{"multiplication"});$$

On vérifie que ces lois sont bien définies et munissent K d'une structure de corps, l'élément neutre pour l'addition étant $0/1$, l'élément neutre pour la multiplication étant $1/1$, et l'inverse d'un élément non nul $\frac{a}{b}$ étant $\frac{b}{a}$.

Proprieté universelle. Soient L un corps et $\phi : A \to L$ un morphisme d'anneaux injectif il existe un unique morphisme de corps $\psi : K \to L$ tel que $\psi|_A = \phi$. L'application

$$A \times A\backslash(\{0\}) \to L, \quad (a,b) \mapsto \frac{\phi(a)}{\phi(b)} \in L$$

passe au quotient $(a,b) \mapsto \frac{a}{b}$ et définit un morphisme de corps $K \to L$ qui convient.

D'autre part, si ψ est un tel morphisme de corps, alors on a $\psi(a/b) = \frac{\psi(a)}{\psi(b)}$, ce qui montre l'unicité de ψ.

7.2 Rappel : caractéristique d'un corps, sous-corps premier

DÉFINITION 7.2.1 *Si A est un anneau commutatif, il existe un unique morphisme d'anneaux $\phi : \mathbb{Z} \to A$, donné par $\phi(n) = n \cdot 1$. Son noyau,* $\operatorname{Ker}\phi$*, est un sous-groupe de $\mathbb{Z}$. Le générateur positif de ce sous-groupe est appelé la* caractéristique *de A et est noté $car(A)$.*

PROPOSITION 7.2.2 *La caractéristique d'un corps K est nul ou un nombre premier $p = car(K)$. Si la caractéristique du corps est nulle, celui-ci contient un sous-corps, isomorphe à $\mathbb{Q}$. Sinon, il contient un sous-corps, isomorphe à $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$, où p est sa caractéristique. Le corps ainsi obtenu est le plus petit corps contenu dans K, on l'appelle sous-corps premier de K.*

PREUVE. D'après le théorème d'isomorphisme, on a $\mathbb{Z}/\operatorname{Ker}\phi \simeq \operatorname{Im}\phi$, ce quotient est donc un anneau intègre, c'est-à-dire $\operatorname{Ker}\phi$ est un idéal premier de $\mathbb{Z}$, $\{0\}$ ou $p\mathbb{Z}$, p premier. Si c'est $\{0\}$, la propriété 7.1.1(ii) montre que $\mathbb{Q}$ est isomorphe à un sous-corps K' de K ; K' est alors le sous-corps engendré par 1 donc le plus petit sous-corps de K. Si $\operatorname{Ker}\phi = p\mathbb{Z}$, alors le sous-anneau $\operatorname{Im}\phi$ de K engendré par 1 est isomorphe à $\mathbb{Z}/p\mathbb{Z}$, donc c'est un corps et le plus petit sous-corps de K.

7.3 Décomposition des fractions rationelles

Pour fixer les idées, nous considérons les fractions rationelles à coefficients *réels* ; le lecteur n'aura aucune peine à remarquer que tous les résultats de ce paragraphe se généralisent presque mot à mot dans le cas des fractions rationelles à coefficients dans un corps K.

Une fraction est dite *irréductible* si son numérateur et son dénominateur sont des polynômes premiers entre eux.

Toute fraction rationelle est égale à une fraction irréductible, cette dernière étant bien définie à un facteur numerique près, ce dernier étant commun pour le dénominateur et le numérateur.

En effet, toute fraction rationelle peut être simplifiée en divisant ses deux polynômes par leur plus grand commun diviseur, après quoi cette fraction devient irréductible. Soient deux fractions irréductibles égales $\dfrac{f(x)}{g(x)}$ et $\dfrac{\varphi(x)}{\psi(x)}$, c'est-à-dire

$$f(x)\psi(x) = g(x)\varphi(x); \tag{7.1}$$

alors, $f(x)$ et $g(x)$ étant premiers entre eux, il en résulte que $f(x)$ est un diviseur de $\varphi(x)$, tandis que, en raison de la même propriété pour $\varphi(x)$ et $\psi(x)$ (qui sont également premiers entre eux) il s'ensuit que $f(x)$ est divisible par $\varphi(x)$. Ainsi, $f(x) = c\varphi(x)$, et de (1) il résulte que $g(x) = c\varphi(x)$.

Une fraction rationelle est dite *régulière* si le degré du numérateur est inférieur à celui du dénominateur. Ajoutant à l'ensemble des fractions régulières le polynôme nul, le théorème suivant est vrai :

THÉORÈME 7.3.1 *Toute fraction rationelle peut être mise d'une façon unique sous la forme de la somme d'un polynôme et d'une fraction régulière.*

En effet, soit une fraction rationelle $\dfrac{f(x)}{g(x)}$ et supposons qu'en divisant $f(x)$ par $g(x)$ on obtienne l'égalité

$$f(x) = g(x)q(x) + r(x),$$

où le degré de $r(x)$ est inférieur à celui de $g(x)$. Alors, il est facile de vérifier que

$$\frac{f(x)}{g(x)} = q(x) + \frac{r(x)}{g(x)}.$$

Si on a en même temps une autre égalité pour $\dfrac{f(x)}{g(x)}$:

$$\frac{f(x)}{g(x)} = \bar{q}(x) + \frac{\varphi(x)}{\psi(x)},$$

avec le degré de $\varphi(x)$ inférieur à celui de $\psi(x)$, alors on obtient l'égalité

$$q(x) - \bar{q}(x) = \frac{\varphi(x)}{\psi(x)} - \frac{r(x)}{g(x)} = \frac{\varphi(x)g(x) - \psi(x)r(x)}{\psi(x)g(x)}.$$

Le premier membre étant un polynôme et le second une fraction régulière, il en resulte que $q(x) - \bar{q}(x) = 0$ et que

$$\frac{\varphi(x)}{\psi(x)} - \frac{r(x)}{g(x)} = 0.$$

Les fractions rationelles régulières peuvent être l'objet d'une étude plus détaillée. Pour cela rappelons l'irréductibilité sur $K = \mathbb{R}$ des polynômes de la forme $x - \alpha$ avec α réel et des polynômes de la forme $x^2 - (\beta + \bar{\beta})x + \beta\bar{\beta}$, où β et $\bar{\beta}$ sont deux nombres complexes conjugués. Il est facile de vérifier que dans le cas complexe les polynômes de la forme $x - \alpha$ avec a complexe jouent le même rôle.

Une fraction rationelle régulière $\dfrac{f(x)}{g(x)}$ est dite *simple* si son dénominateur $g(x)$ est une puissance d'un polynôme irréductible $p(x)$,

$$g(x) = p^k(x),\ k \geq 1,$$

et le numérateur $f(x)$ de degré inférieur à celui de $p(x)$.

THÉORÈME 7.3.2 *Toute fraction rationelle régulière se décompose en une somme de fractions simples.*

Démonstration. Soit d'abord une fraction rationelle régulière $\dfrac{f(x)}{g(x)h(x)}$ avec les polynômes $g(x)$ et $h(x)$ premiers entre eux,

$$(g(x), h(x)) = 1.$$

Par conséquent il existe des polynômes $\bar{u}(x)$ et $\bar{v}(x)$ tels que l'on a

$$g(x)\bar{u}(x) + h(x)\bar{v}(x) = 1.$$

Il en résulte

$$g(x)\Big[\bar{u}(x)f(x)\Big] + h(x)\Big[\bar{v}(x)f(x)\Big] = f(x). \tag{7.2}$$

Soit $u(x)$ le reste de la division du produit $\bar{u}(x)f(x)$ par $h(x)$, le degré de $u(x)$ étant inférieur à celui de $h(x)$. Alors l'égalité (7.2) peut être réecrite sous la forme

$$g(x)u(x) + h(x)v(x) = f(x), \tag{7.3}$$

où $v(x)$ est un polynôme qui peut être facilement calculé. Les degrés du produit $g(x)u(x)$ et du polynôme $f(x)$ étant inférieurs au degré du produit $g(x)h(x)$, le degré du produit $h(x)v(x)$ est également inférieur à celui de $g(x)h(x)$, de sorte que le degré de $v(x)$ est inférieur à celui de $g(x)$. De (7.3) il résulte l'égalité

$$\frac{f(x)}{g(x)h(x)} = \frac{v(x)}{g(x)} + \frac{u(x)}{h(x)},$$

dont le second membre est une somme de fractions régulières.

Si au moins un des dénominateurs $g(x)$ et $h(x)$ peut être représenté sous la forme d'un produit de polynômes premiers entre eux, alors on peut réaliser encore une décomposition. Continuant ce processus nous obtiendrons *une décompsition de toute fraction régulière en une somme d'un certain nombre de fractions régulieres dont chacune a pour dénominateur une puissance d'un polynôme irréductible.* Plus precisement, soit une fraction régulière $\dfrac{f(x)}{g(x)}$ dont le dénominateur $g(x)$ se décompose en un produit de facteurs irréductibles :

$$g(x) = p_1^{k_1}(x)p_2^{k_2}(x)\cdots p_l^{k_l}(x)$$

(on peut toujours supposer que le coefficient du terme principal du dénominateur d'une fraction rationelle est égal à l'unité) ; en outre, $p_i(x) \neq p_j(x)$ pour $i \neq j$. Alors on a

$$\frac{f(x)}{g(x)} = \frac{u_1(x)}{p_1^{k_1}(x)} + \frac{u_2(x)}{p_2^{k_2}(x)} + \cdots + \frac{u_l(x)}{p_l^{k_l}(x)};$$

les termes du second membre de cette égalité sont des fractions régulières.

Il reste à considérer le cas d'une fraction régulière de la forme $\dfrac{u(x)}{p^k(x)}$, où $p(x)$ est irréductible. Appliquant l'algorithme de la division avec reste, divisons $u(x)$ par p^{k-1}, puis le reste de la division par $p^{k-2}(x)$, etc.

Nous sommes conduits aux égalités suivantes :

$u(x) = p^{k-1}(x)s_1(x) + u_1(x)$, $u_1(x) = p^{k-2}(x)s_2(x) + u_2(x)$, $u_{k-2}(x) = p(x)s_{k-1}(x) + u_{k-1}(x)$.

Le degré de $u(x)$ étant, en vertu de notre hypothèse, inférieur au degré de $p^k(x)$ et les degrés des restes $u_i(x)$, $i = 1, 2, \cdots, k-1$, inférieurs aux degrés des diviseurs correspondants $p^{k-i}(x)$, les degrés de tous les quotients $s_1(x)$, $s_2(x), \cdots$, $s_{k-1}(x)$ sont

strictement inférieurs au degré du polynôme $p(x)$. Le degré du dernier reste $u_{k-1}(x)$ également inférieur à celui de $p(x)$.

Il résulte des égalités obtenues que

$$u(x) = p^{k-1}(x)s_1(x) + p^{k-2}(x)s_2(x) + \cdots + p(x)s_{k-1}(x) + u_{k-1}(x).$$

Cela nous conduit à la représentation cherchée de la fonction rationelle $\dfrac{u(x)}{p^k(x)}$ sous la forme d'une somme de fractions simples :

$$\frac{u(x)}{p^k(x)} = \frac{u_{k-1}(x)}{p^k(x)} + \frac{s_{k-1}(x)}{p^{k-1}(x)} + \cdots + \frac{s_2(x)}{p^2(x)} + \frac{s_1(x)}{p(x)}.$$

Le théorème est démontré. On peut le compléter par le *théorème d'unicité* :

THÉORÈME 7.3.3 *Toute fraction rationelle régulière se décompose d'une façon unique en une somme de fractions simples.*

En effet, soient deux représentations différentes d'une fraction régulière sous la forme d'une somme de fractions simples. Retranchant l'une des décomposition de l'autre et regroupant les termes semblables, nous obtenons une somme de fractions simples identiquement nulle. Les dénominateurs des fractions simples formant cette somme sont certaines puissances des polynômes irréductibles distincts $p_1(x),\ p_2(x), \ldots,\ p_s(x)$; soit $p_i^{k_i}(x)$, $i = 1, 2, \ldots, s$, la plus grande puissance du polynôme $p_i(x)$ intervenant aux dénominateurs des fractions sipmles. Multiplions les deux membres de l'égalité en question par le produit $p_1^{k_1-1}(x)p_2^{k_2}(x)\cdots p_s^{k_s}(x)$. Tous les termes de la somme, excepté un, deviennent, après cette multiplication, des polynômes. En ce qui concerne le terme $\dfrac{u(x)}{p^{k-1}(x)}$ il se transforme en une fraction dont le dénominateur est le polynôme $p_1(x)$ et le numérateur le produit $u(x)p_2^{k_2}(x)\cdots p_s^{k_s}(x)$. Le polynôme $p_1(x)$ étant irréductible et tout facteur du numerateur formant avec $p_1(x)$ un couple de polynômes premiers entre eux, le numéroteur n'est pas divisible par le dénominateur. Effectuant la division avec reste, nous obtenons que la somme d'un polynôme et d'une fraction régulière non nulle est nulle, ce qui est impossible.

EXEMPLE 7.3.4 *Décomposer en une somme de fractions simples la fraction régulière réele* $\dfrac{f(x)}{g(x)}$ *avec* $f(x) = 2x^4 - 10x^3 + 7x^2 + 4x + 3$, $g(x) = x^5 - 2x^3 + 2x^2 - 3x + 2$.

Il est facile de vérifier que

$g(x) = (x + 2)(x - 1)^2(x^2 + 1)$;

en outre, chacun des polynômes $x + 2$, $x - 1$, $x^2 + 1$ est irréductible. Il découle de la théorie exposée ci-dessus que la décomposition cherchée doit être de la forme

$$\frac{f(x)}{g(x)} = \frac{A}{x+2} + \frac{B}{(x-1)^2} + \frac{C}{x-1} + \frac{Dx+E}{x^2+1}, \tag{7.4}$$

où les nombres A, B, C, D et E sont à déterminer.

De (4) résulte l'égalité

$$f(x) = A(x-1)^2(x^2+1) + B(x+2)(x^2+1) + C(x+2)(x-1)(x^2+1) + \\ + Dx(x+2)(x-1)^2 + E(x+2)(x-1)^2. \tag{7.5}$$

Identifiant les coefficients des mêmes puissances de x dans les deux membres de l'égalité (7.5), nous obtenons un système de cinq equations linéaires à cinq inconnues A, B, C, D, E ; en outre, il découle du théorème démontré ci-dessus que ce système possède une solution unique. Néanmoins, nous allons choisir une autre méthode.

Faisant dans (7.5) $x = -2$, nous avons l'égalité $45A = 135$, d'où l'on a

$$A = 3. \tag{7.6}$$

Faisons ensuite $x = 1$ dans (7.5), il vient $6B = 6$, c'est-à-dire

$$B = 1. \tag{7.7}$$

Maintenant, faisons dans (7.5) successivement $x = 0$ et $x = -1$. Utilisant (7.6) et (7.7) nous obtenons les equations

$$\begin{cases} & - & 2C & + & 2E & = & -2, \\ -4C & - & 4D & + & 4E & = & -8. \end{cases} \tag{7.8}$$

Il en résulte que

$$D = 1. \tag{7.9}$$

Enfin, faisons $x = 2$ dans (7.5). Utilisant (7.6), (7.7) et (7.9), nous trouvons l'équation

$$20C + 4E = -52,$$

qui, avec la première équation (7.8), donne

$$C = -2,\ E = -3$$

Ainsi,

$$\frac{f(x)}{g(x)} = \frac{3}{x+2} + \frac{1}{(x-1)^2} - \frac{2}{x-1} + \frac{x-3}{x^2+1}.$$

8 Polynômes à plusieurs variables

8.1 Anneau de polynômes à plusieurs variables

DÉFINITION 8.1.1 *Soit A un anneau commutatif. L'anneau des polynômes à n variables peut être défini par récurrence comme*

$$A[X_1, \cdots, X_n] = A[X_1, \cdots, X_{n-1}][X_n].$$

Si $\boldsymbol{\alpha} = (\alpha_1, \cdots, \alpha_n)$ appartient à $\mathbb{N}^n$, on note $X^{\boldsymbol{\alpha}}$ pour le produit $\prod_{i=1}^n X_i^{\alpha_i}$. Tout polynôme s'écrit alors de manière unique

$$P = \sum_{\boldsymbol{\alpha} \in \mathbb{N}^n} a_{\boldsymbol{\alpha}} X^{\boldsymbol{\alpha}}$$

avec $(a_{\boldsymbol{\alpha}})_{\boldsymbol{\alpha} \in \mathbb{N}^n}$ une famille de $A^{\mathbb{N}^n}$ telle que $a_{\boldsymbol{\alpha}} = 0$ sauf pour un nombre fini d'éléments $\boldsymbol{\alpha} \in \mathbb{N}^n$.

Pour un élément $\boldsymbol{\alpha} = (\alpha_1, \cdots, \alpha_n) \in \mathbb{N}^n$, on note $|\boldsymbol{\alpha}| = \sum_{i=1}^n \alpha_i$. On définit alors le degré total d'un polynôme non nul $P = \sum_{\boldsymbol{\alpha} \in \mathbb{N}^n} a_{\boldsymbol{\alpha}} X^{\boldsymbol{\alpha}}$ comme

$$\deg(P) = \sup\{|\boldsymbol{\alpha}| \mid a_{\boldsymbol{\alpha}} \neq 0\}$$

PROPOSITION 8.1.2 *Soit A un anneau commutatif. Pour tout polynôme non nul*

$$P, Q \in A[X_1, \cdots, X_n],$$

on a

(i) si $P + Q \neq 0$, alors $\deg(P+Q) \leq \sup(\deg P, \deg Q)$
avec l'égalité si $\deg P \neq \deg Q$ *;*

(ii) $\deg(PQ) \leq \deg P + \deg Q$
avec l'égalité si A est intègre.

DÉFINITION 8.1.3 *Soit A un anneau commutatif. Un polynôme non nul*

$$P \in A[X_1, \cdots, X_n],$$

est dit homogène *de degré d si et seulement si*

$$\boldsymbol{\alpha} \in \mathbb{N}^n, |\boldsymbol{\alpha}| \neq d \Rightarrow a_{\boldsymbol{\alpha}} = 0.$$

PROPOSITION 8.1.4 *Soient A et B deux anneaux commutatifs, $\phi : A \to B$ un morphisme d'anneaux, et $b_1, \ldots b_n$ des éléments de B, il existe un unique morphisme d'anneaux*

$$\psi : A[X_1, \cdots, X_n] \to B,$$
$$P = \sum_{\boldsymbol{\alpha} \in \mathbb{N}^n} a_{\boldsymbol{\alpha}} X^{\boldsymbol{\alpha}} \longmapsto P(b_1, \ldots b_n) = \sum_{\boldsymbol{\alpha} \in \mathbb{N}^n} a_{\boldsymbol{\alpha}} b_1^{\alpha_1} \cdot \ldots \cdot b_n^{\alpha_n},$$

tel que la réstriction de ψ sur le sous-anneau $A \subset A[X_1, \cdots, X_n]$ coïncide avec ϕ.

DÉFINITION 8.1.5 (INDÉPENDANCE ALGÉBRIQUE) *Soit A' un anneau contenant A. on dit que $x_1, \cdots, x_n \in A'$ sont algébriquement indépendants sur A s'il existe un isomorphisme d'anneaux*

$$\Phi \ : A[X_1, \cdots, X_n] \to A[x_1, \cdots, x_n] \subset A'$$

tel que $\Phi(X_i) = x_i$, c'est-à-dire, que $F(x_1, \cdots, x_n) = 0$ dans A' implique $F = 0$ (tous les coefficients de F sont nuls).

8.2 Polynômes symétriques

Le groupe S_n opère sur les polynômes à n variables $F \in A[X_1, \cdots, X_n]$ par

$$\pi F(X_1, \cdots, X_n) = F(X_{\pi(1)}, \cdots, X_{\pi(n)}).$$

Définition 8.2.1 (polynôme symétrique) *Un polynôme $F \in A[X_1, \cdots, X_n]$ est dit symétrique si $\forall \pi \in S_n$, on a $\pi F = F$.*
Notation $F \in A[X_1, \cdots, X_n]^{S_n}$.

Exemples 8.2.2
a) $f_k = X_1^k + \cdots + X_n^k \in A[X_1, \cdots, X_n]^{S_n}$;
b) $s_0 = 1$, $s_1 = X_1 + \cdots + X_n, s_2 = X_1X_2 + X_1X_3 + \cdots + X_{n-1}X_n$,

$$s_r = \sum_{1 \leq i_1 < \cdots < i_r \leq n} X_{i_1} \cdot \ldots \cdot X_{i_r} \in A[X_1, \cdots, X_n]^{S_n}.$$

Remarque On a

$$\prod_{i=1}^{n} (T - X_i) = \sum_{j=0}^{n} s_j (-1)^j T^{n-j}. \tag{8.1}$$

Théorème 8.2.3 (Théorème fondamental sur les polynômes symétriques) *Soit A un anneau intègre.*
a) $A[X_1, \cdots, X_n]^{S_n} = A[s_1, \cdots, s_n]$.
b) *Les polynômes $s_1, \cdots, s_n$ sont algébriquement indépendants sur A.*
c) *Soit $N \subset \mathbb{N}^n$ l'ensemble de tous les n-multiplets $(\nu_1, \ldots, \nu_n)$ avec $0 \leq \nu_i \leq i - 1$ pour $1 \leq i \leq n$. Alors pour tout $F \in A[X_1, \cdots, X_n]$ il existe $g_\nu \in A[s_1, \cdots, s_n]$ tels que $F = \sum_{\nu \in N} g_\nu X_1^{\nu_1} \cdots X_n^{\nu_n}$, et g_ν sont détérminés, donc $F = g_{0,\cdots,0}$ pour tout $F \in A[X_1, \cdots, X_n]^{S_n}$.*

Rappels : Théorème fondamental sur les polynômes symétriques

Théorème 8.2.3 *Soit A un anneau intègre.*
a) $A[X_1, \cdots, X_n]^{S_n} = A[s_1, \cdots, s_n]$.
b) *Les polynômes $s_1, \cdots, s_n$ sont algébriquement indépendants sur A.*
c) *Soit $N \subset \mathbb{N}^n$ l'ensemble de tous les n-multiplets $(\nu_1, \ldots, \nu_n)$ avec $0 \leq \nu_i \leq i-1$ pour $1 \leq i \leq n$. Alors pour tout $F \in A[X_1, \cdots, X_n]$ il existe $g_\nu \in A[s_1, \cdots, s_n]$ tels que $F = \sum_{\nu \in N} g_\nu X_1^{\nu_1} \cdots X_n^{\nu_n}$, et g_ν sont détérminés, donc $F = g_{0,\cdots,0}$ pour tout $F \in A[X_1, \cdots, X_n]^{S_n}$.*

8.3 Calculs avec des polynômes symétriques. Résultant et discriminant

On considère d'abord le polynôme générique

$$\prod_{i=1}^{n}(T - X_i) = \sum_{j=0}^{n} s_j(-1)^j T^{n-j} \in A[X_1, \ldots, X_n][T]$$

Définition 8.3.1
(a) Le discriminant du polynôme générique *est un polynôme unique $\Delta = \Delta(s_1, \ldots, s_n)$ de n variables sur A tel que*

$$\prod_{1 \leq i < j \leq n} (X_i - X_j)^2 = \Delta(s_1, \ldots, s_n) \in A[X_1, \ldots, X_n]$$

L'existence et l'unicité de $\Delta(s_1, \ldots, s_n)$ est démontré par le théorème 8.2.3
(b) Le discriminant Δ_f d'un polynôme normalisé *$f(X) = X^n + c_{n-1}X^{n-1} + \cdots + c_0$ est défini par la substitution dans $\Delta(s_1, \ldots, s_n)$*

$$s_1 = -c_1, s_2 = c_2, \cdots, s_n = (-1)^n c_n,$$

Exemples 8.3.2 *(1) $n = 2$, $\Delta(s_1, s_2) = (X_1 - X_2)^2 = (X_1 + X_2)^2 - 4X_1X_2 = s_1^2 - 4s_2$*
(2) $n = 3$,

$$\begin{aligned}\Delta(s_1, s_2, s_3) &= (X_1 - X_2)^2(X_1 - X_3)^2(X_2 - X_3)^2 \\ &= s_1^2 s_2^2 - 4s_1^3 s_3 - 4s_2^3 + 18 s_1 s_2 s_3 - 27 s_3^2\end{aligned}$$

On cherche Δ sous la forme $s_1^2 s_2^2 + a s_1^3 s_3 + b s_2^3 + c s_1 s_2 s_3 + d s_3^2$ (suivant les termes principaux $X_1^{\nu_1} X_2^{\nu_2} X_3^{\nu_3}$ pour l'ordre lexicographique, homogènes de degré total 6, avec

$$(\nu_1, \nu_2, \nu_3) = (4,2,0), (4,1,1), (3,3,0), (3,2,1), (2,2,2)).$$

(3) En particulier, pour $f = T^3 + pT + q$, $n = 3$, on a

$$\Delta_f = -4p^3 - 27q^2$$

Soient

$$f(T) = a_n T^n + \cdots + a_0, \quad g(T) = b_m T^m + \cdots + b_0$$

deux polynômes sur un anneau intègre A.

DÉFINITION 8.3.3 *Le résultant* $\operatorname{res}(f,g)$ *est le déterminant* $\det S$,

$$S = \begin{pmatrix} a_n & a_{n-1} & \cdots & a_0 & & & & \\ & a_n & a_{n-1} & \cdots & a_0 & & & \\ & & \cdots & \cdots & \cdots & \cdots & & \\ & & & a_n & a_{n-1} & \cdots & \cdots & a_0 \\ b_m & b_{m-1} & \cdots & b_0 & & & & \\ & b_m & b_{m-1} & \cdots & b_0 & & & \\ & & \cdots & \cdots & \cdots & \cdots & & \\ & & & & b_m & b_{m-1} & \cdots & b_0 \end{pmatrix} \begin{matrix} \left.\vphantom{\begin{matrix}a\\a\\a\\a\end{matrix}}\right\} m \text{ lignes} \\ \left.\vphantom{\begin{matrix}a\\a\\a\\a\end{matrix}}\right\} n \text{ lignes} \end{matrix}$$

PROPOSITION 8.3.4 (PROPRIÉTÉS DU RÉSULTANT) *Soient*

$$f(T) = a_n T^n + \cdots + a_0, \quad g(T) = b_m T^m + \cdots + b_0$$

deux polynômes sur un anneau intègre A. On a
a) $\operatorname{res}(f,g) = (-1)^{mn}\operatorname{res}(g,f)$ *;*
b) Pour tous $a, b \in A$, *on a* $\operatorname{res}(af, bg) = a^m b^n \operatorname{res}(f,g)$ *;*
c) Pour tout morphisme d'anneau $\varphi : A \to A'$ *on a*

$$\operatorname{res}(\varphi(f), \varphi(g)) = \varphi(\operatorname{res}(f,g)) \in A'.$$

THÉORÈME 8.3.5 (SUR L'ÉLIMINATION) *Soient*

$$f(T) = a_n T^n + \cdots + a_0, \quad g(T) = b_m T^m + \cdots + b_0$$

deux polynômes sur un anneau intègre A, tels que $d^\circ(f) + d^\circ(g) \geq 1$. *Il existe* $p, q \in A[T]$ *tels que* $\deg p < m$, $\deg q < n$, *et* $pf + qg = res(f,g)$.

Préparation : résultant comme déterminant d'une application linéaire

Soit $A = K$ un corps, $K_m[T] = \{P \in K[T] \mid d^\circ P \leq m-1 \text{ ou } P = 0\}$, $f(T), g(T) \in K[T]$ deux polynômes $d^\circ f = n$, $d^\circ g = m$, $\mathcal{R} : K_m[T] \times K_n[T] \to K_{m+n}[T]$ l'application K-linéaire telle que $(u,v) \mapsto fu + gv$. Alors le déterminant de $\mathcal{R}$ dans les bases usuelles coincide avec le *résultant* des polynômes f, g.

Soient $f(T) = a_n T^n + \cdots + a_0, g(T) = b_m T^m + \cdots + b_0$.
On trouve la matrice $\mathcal{A} = \mathcal{A}_{e,f}$ de $\mathcal{R}$ dans les bases

$$e = \{(T^{m-1}, 0), (T^{m-2}, 0), \cdots, (T, 0), (1, 0), (0, T^{n-1}), (0, T^{n-2}), \cdots, (0, T), (0, 1)\}$$
de $K_m[T] \times K_n[T]$, et
$$e' = \{T^{m+n-1}, T^{m+n-2}, \cdots, T, 1\} \text{ de } K_{m+n}[T].$$

La matrice $S = \mathcal{A}^t$ est dit la *matrice de Sylvester* des polynômes f et g.
Donc $\operatorname{res}(f,g) = \det_{e,e'} \mathcal{R} = \det \mathcal{A} = \det S$.

Théorème 8.3.6 (Résultant comme la norme algébrique) *Soit f est normalisé par $a_n = 1$, et $B = A[T]/(f)$ un A-module libre avec une base $T^n, \cdots, 1$. Alors*

$$res(f, g) = \mathrm{N}_{B/A}(g(T))$$

où $\mathrm{N}_{B/A}(g(T))$ est le détérminant d'une application A-linéaire

$$g(T) : B \to B, \quad x \mapsto \overline{g(T)}x, \ \text{où } \overline{g(T)} = g(T) \bmod f(T).$$

Corollaire 8.3.7
1) $\mathrm{res}(f, g_1 g_2) = \mathrm{res}(f, g_1)\mathrm{res}(f, g_2)$;
2) $\mathrm{res}(f_1 f_2, g) = \mathrm{res}(f_1, g)\mathrm{res}(f_2, g)$;
3) *Soient*

$$f(T) = a_n T^n + \cdots + a_0, g(T) = b_m T^m + \cdots + b_0$$

deux polynômes sur un anneau intègre A, tels que

$$f = a_n \prod_{i=1}^{n}(T - \alpha_i) \ et \ g = b_m \prod_{j=1}^{m}(T - \beta_j).$$

Alors le résultant de f et g par rapport à T coincide avec

$$a_n^m b_m^n \prod_{i=1}^{n}\prod_{j=1}^{m}(\alpha_i - \beta_j) = a_n^m \prod_{i=1}^{n} g(\alpha_i) = (-1)^{mn} b_m^n \prod_{j=1}^{m} f(\beta_j).$$

4)

$$\Delta_f = (-1)^{n(n-1)/2}\mathrm{res}(f, f')$$

pour tout polynôme normalisé f.

Preuve du théorème 8.3.6 utilise une base auxiliaire

$$e'' = \{f \cdot T^{m-1}, f \cdot T^{m-2}, \cdots, f \cdot 1, T^{n-1}, \cdots, 1\} \text{ de } K_{m+n}[T],$$

et une application linéaire auxiliaire

$\mathcal{R}' : K_{m+n}[T] \to K_{m+n}[T]$, $\mathcal{R}'(f \cdot T^j) = f \cdot T^j, j = m-1, \cdots, 0, \mathcal{R}'(T^i) = g \cdot T^i, i = n-1, \cdots, 0.$

On montre que ([Bosch], p. 172) :

$$\det \mathcal{R}' = \mathrm{res}(f, g) = \mathrm{N}_{B/A}(g(T)). \quad \blacksquare$$

Exemples de calul du résultant avec "Maple"

"resultant - *compute the resultant of two polynomials*

Calling Sequence

resultant(a, b, x)

Parameters

a,b - polynomials in x

x - a name

1. Description

- The function resultant computes the resultant of the two polynomials a and b with respect to the indeterminate x.

- If a and b are polynomials over an integral domain, where

  a = an * product(x - alpha[i], i=1..n)

  and

  b = bm * product(x - beta[i], i=1..m)

  then the resultant of the two polynomials a and b with respect to x is defined to be the polynomial

  an^m*bm^n*product(product(alpha[i] - beta[j], j=1..m), i=1..n)

- The resultant can be computed from the Euclidean algorithm, or computed as the determinant of Sylvester's matrix or Bezout's matrix. For univariate and bivariate resultants over the rationals, modular methods are used for polynomials of high degree and the subresultant algorithm is used for polynomials of low degree. Otherwise Bezout's determinant is computed using minor expansion.

- For efficient computation, resultant takes advantage of any factorization of a and b that is present, although no explicit factorization is attempted.

Reference : "Computer Algebra : Symbolic & Algebraic Computation" Edited by B. Buchberger, G. E. Collins, and R. Loos, Springer-Verlag, Wien, 1982, pp. 115-138 . "

2. Exemples

```
> resultant(a*x+b, c*x+d, x);
```

$$-b\,c + d\,a$$

```
> resultant((x+a)^5,(x+b)^5,x);
```

$$(-a+b)^{25}$$

```
> with(linalg):
> p := a+b*x:
> q := c+d*x+e*x^2:
> sylvester(p,q,x);
```

$$\begin{bmatrix} b & a & 0 \\ 0 & b & a \\ e & d & c \end{bmatrix}$$

```
> resultant(p,q,x);
```

$$b^2\,c - a\,b\,d + a^2\,e$$

```
> sylvester((x-a)^3,(x-b)^3,x);
```

$$\begin{bmatrix} 1 & -3\,a & 3\,a^2 & -a^3 & 0 & 0 \\ 0 & 1 & -3\,a & 3\,a^2 & -a^3 & 0 \\ 0 & 0 & 1 & -3\,a & 3\,a^2 & -a^3 \\ 1 & -3\,b & 3\,b^2 & -b^3 & 0 & 0 \\ 0 & 1 & -3\,b & 3\,b^2 & -b^3 & 0 \\ 0 & 0 & 1 & -3\,b & 3\,b^2 & -b^3 \end{bmatrix}$$

```
> resultant((x-a)^3,(x-b)^3,x);
```

$$(a-b)^9$$

Cours N 3. Lundi 2 février 2015

Extensions des corps commutatifs

9 Extensions et algébricité. Exemples et constructions de corps

9.1 Extensions, degré.

DÉFINITION 9.1.1

(i) Soit K un corps et L un autre corps, contenant K. On dit que L est une extension *de K. C'est un espace vectoriel sur K.*

(ii) Soit L une extension d'un corps K. On appelle degré de L sur K la dimension $\dim_K L$ de L considéré comme espace vectoriel sur K. On le note $[L : K]$, le degré est éventuellement infini. Si le degré $[L : K]$ est fini on dit que L est une extension finie *de K.*

(iii) Si L est une extension de K et $A = (\alpha_i)_{i\in I}$ une partie de L, on appelle extension de K engendrée par A *le sous-corps minimal $K(A)$ de L contenant K et A. Les α_i s'appellent les* générateurs *de $K(A)$ sur K. Tout élément de $K(A)$ s'écrit comme une fraction rationnelle à coefficients dans K d'élément α_i.*

Par exemple, $\mathbb{C}$ est une extension finie de $\mathbb{R}$ de degré 2.
De plus $\mathbb{C} = \mathbb{R}(i)$

THÉORÈME 9.1.2 *Si K, L et E sont trois corps emboîtés tels que $K \subset L \subset E$, alors*

$$[E : K] = [E : L] \cdot [L : K]$$

PREUVE : On note $(a_i)_{i\in I}$ une base de E sur L, et $(b_j)_{j\in J}$ une base de L sur K.

Pour tout $x \in E$, il existe une famille finie $(\alpha_i)_{i\in I_1}$, $I_1 \subset I$, d'éléments de L tels que $x = \sum_{i\in I_1} \alpha_i a_i$. Mais chaque α_i est une combinaison linéaire à coefficients dans K d'éléments b_j : $\alpha_i = \sum_{j\in J_1} \beta_{i,j} b_j$, pour une famille finie $\beta_{i,j} \in K$. Ceci implique que $x = \sum_{(i,j)\in I_1\times J_1} \beta_{i,j} a_i b_j$, et donc la famille $(a_i b_j)_{i\in I, j\in J}$ est génératrice pour le K-espace vectoriel E.

C'est une famille libre : si $(\beta_{i,j})_{(i,j)\in X}$, $X \subset I_1 \times J_1 \subset I \times J$ est une famille finie d'éléments de K telle que $\sum_{(i,j)\in X} \beta_{i,j} a_i b_j = 0$, alors les images I_1 et J_1 de X par projection sur I et J sont finies et on a :

$$\sum_{(i,j)\in X} \beta_{i,j} a_i b_j = \sum_{i\in I_1} \left(\sum_{j\in J_1} \beta_{i,j} b_j \right) a_i = 0$$

et comme pour tout i $\sum_{j\in J_1} \beta_{i,j} b_j$ appartiennent à L, on trouve $\sum_{j\in J_1} \beta_{i,j} b_j = 0$ pour tout $i \in I_1$, puis $\beta_{i,j} = 0$ pour tout $(i, j) \in X$. ∎

COROLLAIRE 9.1.3 *Pour n corps emboîtés*

$$K \subset K_1 \subset \cdots \subset K_n,$$

on a l'égalité

$$[K_n : K] = [K_1 : K] \cdot [K_2 : K_1] \cdot \ldots \cdot [K_n : K_{n-1}].$$

EXEMPLE 9.1.4 *On considère le sous-corps $K = \mathbb{Q}(\sqrt[3]{2}, i)$ de $\mathbb{C}$, il contient $\mathbb{Q}(\sqrt[3]{2})$, et $K = \mathbb{Q}(\sqrt[3]{2})(i)$, donc*

$$[K : \mathbb{Q}] = [\mathbb{Q}(\sqrt[3]{2})(i) : \mathbb{Q}(\sqrt[3]{2})] \cdot [\mathbb{Q}(\sqrt[3]{2}) : \mathbb{Q}].$$

Le polynôme X^3-2 est irréductible dans $\mathbb{Q}[X]$ et le polynôme X^2+1 l'est dans $\mathbb{Q}(\sqrt[3]{2})[X]$. Donc,

$$[\mathbb{Q}(\sqrt[3]{2})(i) : \mathbb{Q}(\sqrt[3]{2})] = 2, [\mathbb{Q}(\sqrt[3]{2}) : \mathbb{Q}] = 3, \text{ et } [K : \mathbb{Q}] = 6.$$

9.2 Éléments algébriques

Soit E une extension d'un corps K.

DÉFINITION 9.2.1

(i) Un élément α de E est dit algébrique *sur K s'il existe un polynôme non nul P de $K[X]$ tel que $P(\alpha) = 0$.*

(ii) Une extension E de K est dite algébrique *si tout élément α de E est algébrique sur K.*

(iii) Si $\alpha \in E$ est un élément algébrique sur K l'ensemble des polynômes $P \in K[X]$ tels que $P(\alpha) = 0$, forme un idéal de $K[X]$, non réduit à (0). Cet idéal est principal, et son générateur unitaire s'appelle le polynôme minimal *de α sur K.*

PROPOSITION 9.2.2 *Soit E une extension d'un corps K, et soit α un élément de E algébrique sur K de polynôme minimal P.*

i) Si $Q \in K[X]$ admet α comme racine, alors P divise Q dans $K[X]$.

(ii) Le polynôme P est irréductible dans $K[X]$.

(iii) Le sous-anneau $K[\alpha]$ de E est un corps $K(\alpha)$ et on a $[K(\alpha) : K] = \deg P$. La famille $(1, \alpha, \cdots, \alpha^{n-1})$, où $n = \deg P$, est une base de $K(\alpha)$ sur K.

PREUVE : L'assertion (i) traduit le fait que P engendre l'idéal des éléments de $K[X]$, ayant α comme racine.

(ii) Si P se factorise en QR dans $K[X]$, alors on a $P(\alpha) = Q(\alpha)R(\alpha) = 0$ dans le corps E, donc $Q(\alpha) = 0$ ou $R(\alpha) = 0$, et P divise Q ou R. ∎

(iii) Le sous-anneau $K[\alpha]$ est l'image de l'homomorphisme d'évaluation $Q \mapsto Q(\alpha)$ de $K[X]$ dans E, dont le noyau est l'idéal (P), maximal d'après (ii) Par le théorème d'isomorphisme, l'anneau $K[\alpha]$ est donc isomorphe au quotient $K[X]/(P)$, qui est un corps. Si $n = \deg P$, la famille $(1, \alpha, \cdots, \alpha^{n-1})$ est libre sur K car P est le polynôme non nul de plus petit degré qui annule α. Elle est génératrice car pour tout $Q \in K[X]$ on a $Q(\alpha) = R(\alpha)$, où R est le reste de la division euclidienne de Q par P, et donc $R(\alpha) = \sum_{i=0}^{n-1} r_i \alpha^i$, $r_i \in K$. La famille est donc une base de $K(\alpha)$ sur K. En particulier, on a $[K(\alpha) : K] = n = \deg P$. ∎

PROPOSITION 9.2.3 *Soit E une extension finie d'un corps K ($[E : K] = n \in \mathbb{N}$), alors E est algébrique sur K.*

PREUVE. Soit α un élément de E. Si $n = [E : K]$, la famille de $n + 1$ éléments $(1, \alpha, \cdots, \alpha^n)$ n'est pas libre, donc il existe $P(X) = \sum_{i=0}^{n} a_i X^i$ non nul dans $K[X]$ tel que $\sum_{i=0}^{n} a_i \alpha^i = P(\alpha) = 0$. ∎

REMARQUE . L'assertion réciproque est fausse : l'extension $E = \overline{\mathbb{Q}} \subset \mathbb{C}$ de $\mathbb{Q}$, formée par tous les nombres complexes algébriques sur $\mathbb{Q}$ (E est bien un corps, exercice), est algébrique par définition, mais $[E : \mathbb{Q}]$ n'est pas finie. En effet, on vérifie qu'il existe des polynômes irréductibles sur $\mathbb{Q}$ de tout degré $n \geq 1$, par exemple $X^n - 2$ (exercice).

9.3 Corps de rupture, corps de décomposition

Remarquons que pour un corps arbitraire K et pour tout polynôme P non constant de $K[X]$ on peut construire une extension L de K dans laquelle P possède une racine : quitte à factoriser P, on peut le supposer irréductible, auquel cas on a vu que l'anneau quotient $K[X]/(P)$ est un corps. La classe $X + (P)$ est alors une racine de P dans $K[X]/(P)$.

THÉORÈME 9.3.1 (SUR L'ISOMORPHISME) *Soient L une extension de K et $\alpha \in L$ une racine d'un polynôme irréductible P de $K[X]$. Alors l'anneau $K[\alpha]$ est un corps isomorphe à $K[X]/(P)$.*

On peut déduire d'après le paragraphe précédent la proposition 9.2.2 et sa preuve, puisque P est, à un facteur constant près, le polynôme minimal de α sur K.

DÉFINITION 9.3.2 *Soit K un corps, et P un polynôme irréductible. On dit qu'un corps L contenant K est un* corps de rupture *de P sur K s'il existe un $\alpha \in L$ tel que $L = K(\alpha)$ et $P(\alpha) = 0$.*

REMARQUE 9.3.3 *Avec les notations de la proposition, on peut donner une construction matricielle de l'anneau $K[X]/(P)$, corps de rupture de P sur K.*

En effet, on écrit $P = \sum_{j=0}^{n} a_j X^j$, et on suppose que $a_n = 1$, alors dans la base

$$(\overline{1}, \overline{X}, \cdots, \overline{X^{n-1}}) \bmod (P)$$

de $K[X]/(P)$ la multiplication $\mu : Q \mapsto \overline{X}Q \bmod (P)$ a pour matrice

$$A = \begin{pmatrix} 0 & 0 & \dots & 0 & -a_0 \\ 1 & 0 & \dots & 0 & -a_1 \\ 0 & 1 & \dots & 0 & -a_2 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 & -a_{n-1} \end{pmatrix},$$

dont le polynôme minimal est P. En résulte que $K[A] \simeq K[X]/\operatorname{Ker}(\mu)$ est isomorphe à $K[X]/(P)$.

En itérant la construction du corps de rupture ci–dessus (on choisit à chaque étape un facteur irréductible de P de degré > 1 sur le corps obtenu), on peut construire pour tout polynôme $P \in K[X]$ une extension finie L' de K dans laquelle P s'écrit comme produit de facteurs du premier degré. La construction implique que l'extension L' peut être choisie de telle façon que $[L' : K] \leq n!$.

9.3. Corps de rupture, corps de décomposition (rappels et suite)

THÉORÈME 9.3.1 (SUR L'ISOMORPHISME) *Soient L une extension de K et $\alpha \in L$ une racine d'un polynôme irréductible P de $K[X]$. Alors l'anneau $K[\alpha]$ est un corps isomorphe à $K[X]/(P)$.*

DÉFINITION 9.3.2 Soit K un corps, et P un polynôme irréductible. On dit qu'un corps L contenant K est un corps de rupture de P sur K s'il existe un $\alpha \in L$ tel que $L = K(\alpha)$ et $P(\alpha) = 0$.

COROLLAIRE A (DÉCOMPOSITION EN FACTEURS LINÉAIRES) *Soient K un corps, P un polynôme de $K[X]$ de degré ≥ 1. Alors il existe une extension E de K dans laquelle le polynôme P se décompose en produit de facteurs linéaires.*

COROLLAIRE B (SUR L'EXISTENCE DES RACINES) *Soient K un corps, P_i un ensemble fini de polynômes de $K[X], i = 1, \cdots, n$, de degré ≥ 1. Alors il existe une extension L de K dans laquelle tout polynôme P_i possède une racine $\beta_i \in L$.*

DÉFINITION 9.3.4 (CORPS DE DÉCOMPOSITION) *On considère un polynôme P de $K[X]$ de degré supérieur ou égal à 1, et une extension E de K, dans laquelle P s'écrit $P(X) = (X-\alpha_1)\cdot\ldots\cdot(X-\alpha_n)$. Alors le corps $L = K(\alpha_1, \cdots, \alpha_n)$ s'appelle* corps de décomposition *de P dans E. C'est l'extension minimale de K dans E, dans laquelle P se décompose en produit de facteurs linéaires.*

On va vérifier que ce corps est uniquement déterminé à un isomorphisme près :

THÉORÈME 9.3.5 (SUR UN PROLONGEMENT D'ISOMORPHISME) *On considère un isomorphisme de corps $\sigma : K \to K'$, et un polynôme irréductible $P(X) = \sum_{j=0}^{n} a_j X^j \in K[X]$. On note $P^\sigma = \sum_{j=0}^{n} \sigma(a_j) X^j \in K'[X]$. On choisit une extension de K contenant une racine β de P, et une extension de K', contenant une racine β' de $P^\sigma(X)$. Alors il existe un isomorphisme de corps*

$$\tilde{\sigma} : K(\beta) \to K'(\beta'),$$

qui prolonge σ et tel que $\tilde{\sigma}(\beta) = \beta'$.

PREUVE. Comme σ est un isomorphisme de corps de K dans K', l'application

$$\varphi : K[X] \to K'[X], \quad Q(X) = \sum_{j=0}^{n} b_j X^j \mapsto Q^\sigma(X) = \sum_{j=0}^{n} \sigma(b_j) X^j$$

est un isomorphisme d'anneaux. On en déduit

$$\begin{array}{rccc} \tilde{\sigma} : & K(\beta) & \to & K'(\beta') \\ & \uparrow & & \uparrow \\ \overline{\varphi} : & K[X]/(P) & \to & K'[X]/(P^\sigma), \\ & Q + (P) & \mapsto & Q^\sigma + (P^\sigma) \end{array}$$

Cela montre à la fois que P^σ est un élément irréductible de $K'[X]$ (par le théorème sur l'isomorphisme 9.3.1, $K'[X]/(P^\sigma)$ est un corps), et que si $\theta = Q(\beta) \in K(\beta)$, son image par $\tilde{\sigma}$ est bien définie par l'égalité $\tilde{\sigma}(\theta) = Q^\sigma(\beta')$. ∎

THÉORÈME 9.3.6 (DE L'UNICITÉ) *Soient $\sigma : K \xrightarrow{\sim} K'$ un isomorphisme de corps, $P(X) = \sum_{j=0}^n a_j X^j$ un polynôme de $K[X]$, et notons $P^\sigma = \sum_{j=0}^n \sigma(a_j) X^j \in K'[X]$.*

À P, on associe une extension E de K, dans laquelle il se factorise en produit des termes de premier degré, $P(X) = a(X-\alpha_1)\cdot\ldots\cdot(X-\alpha_n)$, et on note $B = K(\alpha_1, \cdots, \alpha_n)$ le corps de décomposition de P dans E.

On définit de même E' et B' pour le polynôme P^σ. Il existe alors un isomorphisme de corps

$$\tau : B \xrightarrow{\sim} B',$$

dont la restriction à K est égale à σ.

PREUVE. Le polynôme P s'écrit $P(X) = a(X-\alpha_1)\cdot\ldots\cdot(X-\alpha_n)$, les α_i étant des éléments de E. Raisonnons par récurrence sur le nombre N d'éléments α_i, qui n'appartiennent pas à K.

Si $N = 0$, tous les α_i appartiennent à K, donc $B = K$ est isomorphe à $B' = K'$, et $\tau = \sigma$.

Pour $N \geq 1$, supposons que α_1 n'appartienne pas à K.

C'est un élément algébrique sur K, de polynôme minimal S, et il existe $Q \in K[X]$ tel que $P = SQ$.

Dans $E'[X]$, on a les égalités

$$P^\sigma = S^\sigma Q^\sigma = a'(X-\alpha'_1)\cdot\ldots\cdot(X-\alpha'_n).$$

Si β est une racine de S^σ dans une extension de E', il s'ensuit que

$$P^\sigma(\beta) = 0 = a'(\beta-\alpha'_1)\cdot\ldots\cdot(\beta-\alpha'_n),$$

donc il existe un indice i, qu'on peut supposer égal à 1, tel que $\beta = \alpha'_1 \in E'$. On utilise le théorème 9.3.5 pour le polynôme irréductible S : il existe un isomorphisme de corps

$$\tau : K(\alpha_1) \xrightarrow{\sim} K'(\alpha'_1),$$

qui prolonge σ.

On considère maintenant P comme un polynôme à coefficients dans $L = K(\alpha_1)$.

Le nombre de racines de P qui n'appartiennent pas à L est strictement inférieur à N, et l'hypothèse de récurrence nous donne l'existence d'un prolongement de τ,

$$\pi : L(\alpha_2, \cdots, \alpha_n) \xrightarrow{\sim} L'(\alpha'_2, \cdots, \alpha'_n),$$

avec $L' = K'(\alpha'_1)$. On termine en remarquant que $L(\alpha_2, \cdots, \alpha_n) = K(\alpha_1, \alpha_2, \cdots, \alpha_n)$, et que $L'(\alpha'_2, \cdots, \alpha'_n) = K'(\alpha'_1, \alpha'_2, \cdots, \alpha'_n)$. ∎

COROLLAIRE 9.3.7 (DE L'UNICITÉ) *Soient K un corps, P un polynôme de $K[X]$, et L, L' deux corps de décomposition de P sur K. Alors il existe un isomorphisme de corps de L sur L' dont la restriction à K est l'identité.*

En effet, il suffit d'utiliser le résultat précédent pour $\sigma = \mathrm{id} : K \to K$.

10 Clôture algébrique (voir [Lang], Ch.VII, §2)

10.1 Prolongement d'isomorphismes sur les extensions algébriques

On considère un isomorphisme de corps $\sigma : K \to K'$, et un polynôme irréductible $P(X) = \sum_{j=0}^{n} a_j X^j \in K[X]$. On note $P^\sigma = \sum_{j=0}^{n} \sigma(a_j) X^j \in K'[X]$. On choisit une extension L de K contenant une racine β de P, et une extension L' de K', contenant une racine β' de $P^\sigma(X)$. Rappelons

THÉORÈME 9.3.5 (SUR UN PROLONGEMENT D'ISOMORPHISME) *Il existe un isomorphisme de corps*

$$\tilde{\sigma} : K(\beta) \to K'(\beta'),$$

qui prolonge σ et tel que $\tilde{\sigma}(\beta) = \beta'$.

On écrira

$$\begin{array}{ccc} K(\beta) & \xrightarrow{\tilde{\sigma}} & K'(\beta') \\ {\scriptstyle \text{incl}}\uparrow & & \uparrow{\scriptstyle \text{incl}} \\ K & \xrightarrow{\sigma} & K' \end{array}$$

LEMME 10.1.1 *On considère une extension algébrique E de K, et un morphisme injectif $\sigma : E \to E$ tel que $\sigma(x) = x$ pour tout $x \in K$.*

Alors σ est un automorphisme de E.

PREUVE. Il suffit de vérifier que σ est surjectif. Pour un α, soit P son polynôme minimal sur K, et soit E' le sous-corps de E engendré par toutes les racines de P dans E. Alors $\sigma(E') \subset E'$, et $[E' : K] < \infty$, donc $\sigma(E') = E'$.

10.2 Extensions algébriquement clôses

DÉFINITION 10.2.1 (CORPS ALGÉBRIQUEMENT CLÔS) *Un corps L est dit algébriquement clôs, si tout polynôme de $L[X]$ de degré ≥ 1 posséde une racine dans L.*

Rapellons : THÉORÈME DE D'ALEMBERT-GAUSS : *Le corps $\mathbb{C}$ est algébriquement clos (c'est-à-dire que tout polynôme à coefficients dans $\mathbb{C}$, de degré au moins un, a une racine dans $\mathbb{C}$).*

THÉORÈME 10.2.2 (SUR L'EXISTENCE D'UNE EXTENSION ALGÉBRIQUEMENT CLÔSE) *Pour tout corps K il existe un corps L algébriquement clôs contenant K.*

PREUVE. 1) On construit d'abord une extension E_1 dans laquelle tout polynôme de $K[X]$ de degré ≥ 1 posséde une racine (Artin).

On associe à tout polynôme $f \in K[X]$ de degré ≥ 1 une variable formelle X_f, et on considère l'anneau $A = K[S]$, où

$$S = \{X_f\}_{f \in K[X] \text{ de degré } \geq 1}$$

l'ensemble infini de variables indépendantes.

L'idéal $I = \langle f(X_f)\rangle_{f\in K[X]}$ de degré ≥ 1 est propre $I \neq A$. Dans le cas contraire, on aurait

$$1 = \sum_{i=1}^{n} \alpha_i f(X_{f_i}).$$

D'après le Corollaire B du Théorème 9.3.1, il existe une extension L de K dans laquelle tout polynôme f_i possède une racine $\beta_i, i = 1, \cdots, n$. En substituant $X_{f_i} = \beta_i$, on obtient $1 = 0$ (contradiction).

On utilise le fait qu'il existe un idéal maximal $\mathfrak{m} \supset I$ (avec Lemme de Zorn), alors l'anneau quotient $E_1 = A/\mathfrak{m} \supset K$ est un corps, et tout polynôme de $f \in K[X]$ de degré ≥ 1 posséde une racine $X_f \bmod \mathfrak{m} \in E_1 = A/\mathfrak{m}$.

2) On construit par récurrence une tour infinie

$$K \subset E_1 \subset \cdots \subset E_n \subset \ldots$$

telle que tout polynôme de $E_n[X]$ de degré ≥ 1 posséde une racine dans l'extension E_{n+1}.

On pose $E = \bigcup_{n\in\mathbb{N}} E_n$, alors il existe n tel que $f \in E_n[X]$. Tout polynôme de $f \in E_n[X]$ de degré ≥ 1 posséde une racine $X \in E_{n+1} \subset E$. ∎

Corollaire 10.2.3 (sur l'existence d'une clôture algébrique) *Pour tout corps K il existe une extension $\overline{K}$ algébrique sur K telle que $\overline{K}$ est algébriquement clôs. On apelle $\overline{K}$ une clôture algébrique de K.*

Preuve. Pour tout corps K il existe un corps L algébriquement clôs contenant K (par Théorème 10.2.4).

On construit $\overline{K}$ comme l'ensemble de tous les éléments de L algébriques sur K. On utilise le fait qu'un élément $\alpha \in L$ est algébrique sur K si et seulement si α engendre une extension finie de K : α est une racine d'un polynôme irréductible Q de $K[X]$, et l'anneau $K[\alpha]$ est un corps isomorphe à $K[X]/(Q)$ (par Théorème 9.3.1 sur l'isomorphisme).

Il s'agit d'une extension algébrique de K, et tout polynôme $P(X) = \sum_{j=0}^{n} a_j X^j \in \overline{K}[X] \subset L[X]$ possède une racine β dans L.

Comme β est algébrique sur $\overline{K}$, et $\overline{K}$ algébrique sur K, on en déduit que β est algébrique sur K, donc $\beta \in \overline{K}$. En effet, β est algébrique sur l'extension finie $K(a_0, \cdots, a_{n-1})$ de K par la multiplicativité de degré (Corollaire 9.1.3). ∎

Théorème 10.2.4 *Soit K un corps, E son extension algébrique, et $\sigma : K \to L$ une inclusion de K dans un corps algébriquement clôs.*
Alors il existe un prolongement de σ à une inclusion $\tilde{\sigma} : E \to L$.
Si E est algébriquement clôs, et L est algébrique sur $\sigma(K)$, alors toute telle extension $\tilde{\sigma}$ est un isomorphisme de E dans L.

Preuve. Soit S l'ensemble de couples (F, τ), où F est un sous-corps de E, contenant K, et τ est un prolongement de σ à une inclusion de F dans L.

On écrira $(F, \tau) \leq (F', \tau')$, si $F \subset F'$, $\tau'|_F = \tau$. Alors S est non-vide et on vérifie qu'il est inductivement ordonné (toute chaîne de S possède un majorant).

Lemme de Zorn (voir Halmos P.R., Naive Set Theory, *sans démonstration)* Soit S un ensemble ordonné telle que toute chaîne T de S possède un majorant. Alors S possède au moins un élément maximal.

Soit (K', λ) un élément maximal de S .

On affirme que $K' = E$: sinon il existe $\alpha \in E$, $\alpha \notin K'$, et il existe un prolongement de λ sur $K'(\alpha)$ ce qui contrairie à la maximalité de (K', λ).

Ceci dit, il existe un prolongement $\tilde{\sigma}$ de σ sur E.

De plus, si E est algébriquement clôs, et L est algébrique sur $\sigma(K)$, toute telle extension $\tilde{\sigma}$ est un isomorphisme de E dans L. Il suffit de voir que $\tilde{\sigma}$ est surjectif. Mais pour tout $\gamma \in L$ de polynôme minimal (noté R^σ) sur $\sigma(K)$, l'ensemble des racines de R est envoyé par $\tilde{\sigma}$ sur l'ensemble des racines de R^σ. ∎

Corollaire 10.2.5 (sur l'unicité d'une clôture algébrique) *Soit K un corps, $\overline{K}$ une clôture algébrique de K. Alors $\overline{K}$ est unique à un isomorphisme près.*

11 Morphisme de Frobenius, structure des corps finis

Pour un plus large développement sur les corps finis, le lecteur est renvoyé au Lidl-Niederreiter, [Li-Ni].

Jusqu'ici nous avons rencontré l'exemple fondamental des corps finis $\mathbb{Z}/p\mathbb{Z}$ (p premier), quotients de $\mathbb{Z}$ par un idéal maximal. Il s'agit maintenant de décrire tous les corps finis.

11.1 Structure

Proposition-Définition 11.1.1 *Soit A un anneau commutatif de caractéristique p un nombre premier. L'application*

$$\mathrm{Fr}_p : x \mapsto x^p, x \in A$$

est un morphisme d'anneau appelé morphisme de Frobenius. Plus généralement, si A est un anneau commutatif de caractéristique p premier et si q est une puissance de p, on note $\mathrm{Fr}_q : x \mapsto x^q$.

Preuve. Le point à montrer est l'additivité. Or si $x, y \in A$ on développe $(x+y)^p$ par la formule du binôme, et on conclut par le fait que si $1 \leq i \leq p-1$, p divise l'entier C_p^i.

Théorème 11.1.2 *: Soit $\mathcal{K}$ un corps fini. Alors $\mathcal{K}$ est de caractéristique p un nombre premier, $\mathcal{K}$ est de cardinal $q = p^d$, avec $d = [\mathcal{K} : \mathbb{Z}/p\mathbb{Z}]$, et Fr_p est un automorphisme du corps $\mathcal{K}$.*

Inversement, si p est premier et d est un entier strictement positif, il existe à isomorphisme près un unique corps à $q = p^d$ éléments, qui est le corps de décomposition de $X^q - X$ sur $\mathbb{Z}/p\mathbb{Z}$. On note ce corps $\mathbb{F}_q$. De plus, le groupe $(\mathbb{F}_q, +)$ est isomorphe au groupe $((\mathbb{Z}/p\mathbb{Z})^d, +)$ et le groupe multiplicatif $\mathbb{F}_q^*$ est isomorphe à $\mathbb{Z}/(q-1)\mathbb{Z}$ (groupe cyclique d'ordre $q-1$).

PREUVE. Vérifions ensuite qu'un corps $\mathcal{K}$ à q éléments est un corps de décomposition pour le polynôme $X^q - X$. Comme $\mathcal{K}$ a q éléments, $\mathcal{K}^*$ est un groupe d'ordre $q-1$. Par conséquent,

$$\forall x \in \mathcal{K}^*,\ x^{q-1} = 1.$$

Autrement dit, les q éléments de $\mathcal{K}$ sont racines de $X^q - X$. Du fait du degré, on obtient

$$X^q - X = \prod_{\alpha \in \mathcal{K}} (X - \alpha).$$

En particulier, $\mathcal{K}$ est un corps de décomposition pour le polynôme $X^q - X$.

Soit inversement $\mathcal{K}$ un corps de décomposition pour le polynôme $X^q - X$ sur $\mathbb{Z}/p\mathbb{Z}$, où p est premier et q est une puissance de p. Comme Fr_q est un morphisme de corps de $\mathcal{K}$, l'ensemble de ses points fixes les racines de $X^q - X$ est un sous-corps de $\mathcal{K}$. Comme $\mathcal{K}$ est engendré par ces racines sur $\mathbb{F}_p$, $\mathcal{K}$ est l'ensemble des racines de $P(X) = X^q - X$. Comme $P' = -1$, toutes les racines de P sont simples (par les propriétés de la dérivé formelle P' d'un polynôme), P a donc ses q racines distinctes dans $\mathcal{K}$ et $\mathcal{K}$ a exactement q éléments. D'après 9.3.7 ceci établit l'assertion de l'existence et de l'unicité à isomorphisme près du corps $\mathbb{F}_q$.

Enfin la cyclicité du groupe $\mathbb{F}_q^*$ est un cas particulier du théorème de cyclicité de tout sous-groupe fini du groupe multiplicatif d'un corps fini (rappel). ∎

THÉORÈME 11.1.3 *: Soit $q = p^n$ où p est premier. Tout sous-corps du corps $\mathbb{F}_q$ est de cardinal p^m, où m est un diviseur de n. Et pour tout diviseur m de n, $\mathbb{F}_q$ possède un unique sous-corps de cardinal p^m, qui est l'ensemble des racines du polynôme $X^{p^m} - X$ dans $\mathbb{F}_q$.*

PREUVE. Si $\mathcal{K}$ est un sous-corps de $\mathbb{F}_q$, alors il contient le sous-corps premier F_p, donc c'en est une extension finie et $\mathbb{F}_q$ est une extension finie de $\mathcal{K}$. De là $\mathcal{K}$ a pour cardinal p^m, et le cardinal de $\mathbb{F}_q$ est une puissance de $(p^m)^d$ de celui de $\mathcal{K}$. Ainsi $n = md$. Inversement, si m divise n, alors $p^m - 1$ divise $p^n - 1$, donc le polynôme $X^{p^m-1} - 1$ divise le polynôme $X^{p^n-1} - 1$, donc le polynôme $X^{p^m} - X$ divise $X^{p^n} - X$, ce qui entraîne que $X^{p^m} - X$ a p^m racines distinctes dans $\mathbb{F}_q$. Ces racines forment l'unique sous-corps de cardinal p^m de $\mathbb{F}_q$. ∎

Dans la pratique, pour pouvoir faire les calculs, le corps $\mathbb{F}_{p^n}$ $(n > 1)$ sera construit comme anneau quotient de type $\mathbb{F}_p[X]/(Q)$, en choisissant un polynôme irréductible Q de degré n. (voir le théorème 11.2.1)

11.2 Polynômes sur les corps finis. Nombre de polynômes irréductibles

THÉORÈME 11.2.1 *Soit p un nombre premier, et q une puissance de p. Pour tout entier $m \geq 1$, il existe $\theta \in \mathbb{F}_{q^m}$ tel que $\mathbb{F}_{q^m} = \mathbb{F}_q[\theta]$ et il existe P polynôme irréductible de degré m sur $\mathbb{F}_q$.*

PREUVE. Soient θ un générateur du groupe cyclique $\mathbb{F}_{q^m}^*$, et P son polynôme minimal sur $\mathbb{F}_q$. Alors on a $\mathbb{F}_{q^m} = \mathbb{F}_q[\theta]$ et P est un polynôme irréductible sur $\mathbb{F}_q$ dont $\mathbb{F}_{q^m}$ est un corps de rupture. Autrement dit, on a un isomorphisme

$$\mathbb{F}_q[X]/(P) \xrightarrow{\sim} \mathbb{F}_{q^m}$$

qui envoie la classe de X sur θ. En particulier, on a $\deg P = \dim_{\mathbb{F}_q}(\mathbb{F}_{q^m}) = m$.

REMARQUE 11.2.2 *Si on a un tel polynôme et α une racine de P dans $\mathbb{F}_{q^m}$, la famille $\{1, \alpha, \cdots, \alpha^{m-1}\}$ est une base du $\mathbb{F}_q$-espace vectoriel $\mathbb{F}_{q^m}$.*

PROPOSITION 11.2.3 *Soient P un polynôme irréductible sur $\mathbb{F}_q$ et α une racine de P dans une extension de $\mathbb{F}_q$. Alors, pour tout polynôme Q sur $\mathbb{F}_q$, $Q(\alpha) = 0$ si et seulement si P divise Q.*

En effet, P est alors le polynôme minimal de α sur $\mathbb{F}_q$, voir la proposition 9.2.2.

LEMME 11.2.4 *Soit P un polynôme irréductible de degré m sur $\mathbb{F}_q$. Alors P divise $X^{q^n} - X$ si et seulement si m divise n.*

PREUVE. Le corps de rupture de P sur $\mathbb{F}_q$ est de cardinal q^m, donc tout élément y vérifie $x^{q^m} = x$, donc aussi en itérant si m divise n, $x^{q^n} = x$. On conclut que P divise $X^{q^n} - X$ en appliquant la proposition précédente avec $Q = X^{q^n} - X$. Inversement, si P divise $X^{q^n} - X$ alors le corps $\mathbb{F}_{q^n}$ contient un corps de rupture de P, de cardinal q^m, donc par le théorème 11.1.3 m divise n.

THÉORÈME 11.2.5 *Soit P un polynôme irréductible sur $\mathbb{F}_q$ de degré m. Alors P est scindé sur le corps $\mathbb{F}_{q^m}$ et a toutes ses racines simples. Si α est l'une d'elles, ces m racines sont $\alpha, \alpha^q, \cdots, \alpha^{q^{m-1}}$. En particulier si $P \neq X$ toutes les racines de P ont le même ordre multiplicatif dans $\mathbb{F}_{q^m}^*$.*

PREUVE. On a vu que le corps de rupture de P, $\mathbb{F}_q[X]/(P)$, de cardinal q^m, est formé de l'ensemble des racines du polynôme $X^{q^m} - X$. Par suite $X^{q^m} - X$ s'annule en une racine de P, donc par la proposition il est divisible par P sur $\mathbb{F}_q$ et a fortiori sur $\mathbb{F}_{q^m}$. Ainsi puisque $X^{q^m} - X$ est scindé à racines simples sur $\mathbb{F}_{q^m}$, il en est de même pour P.

Ensuite, on écrit $P = \sum_{i=0}^m a_i X^i$ avec $a_i \in \mathbb{F}_q$. Si α est une racine de P, alors

$$\mathrm{Fr}_q(P(\alpha)) = \sum_{i=0}^m \mathrm{Fr}_q(a_i)\mathrm{Fr}_q(\alpha)^i = P(\mathrm{Fr}_q(\alpha)) = 0$$

où l'avant-dernière égalité vient du fait que $\mathrm{Fr}_q(x) = x$ pour tout $x \in \mathbb{F}_q$. Par conséquent $\alpha, \alpha^q, \cdots, \alpha^{q^{m-1}}$ sont des racines de P. Montrons par l'absurde que ces m racines sont distinctes. En effet, dans le cas contraire, il existe i, j avec $0 \leq i < j \leq m-1$ tels que $\alpha^{q^i} = \alpha^{q^j}$ et donc $\alpha^{q^j - q^i} = 1$. Par conséquent

$$\mathrm{ord}(\alpha) | q^j - q^i = q^i(q^{j-i} - 1).$$

Mais comme $\alpha \in \mathbb{F}_{q^m}^*$, l'ordre de α est premier à q donc, par le lemme de Gauss, $\mathrm{ord}(\alpha) | q^{i-j} - 1$, et $\alpha^{q^{j-i}} = \alpha$, donc α appartient au corps $\mathbb{F}_{q^{j-i}}$, ce qui est en contradiction avec le fait que $[\mathbb{F}_q(\alpha) : \mathbb{F}_q] = \deg P = m$. Ainsi on a

$$P(X) = (X - \alpha)(X - \alpha^q) \cdot \ldots \cdot (X - \alpha^{q^{m-1}}).$$

La dernière assertion résulte de ce que Fr_q est un automorphisme du corps $\mathbb{F}_{q^m}$, donc il conserve l'ordre multiplicatif des éléments.

COROLLAIRE 11.2.6 *Le corps de décomposition de tout polynôme de degré m irréductible sur $\mathbb{F}_q$ est $\mathbb{F}_{q^m}$.*

Tout élément $t \in \mathbb{F}_{q^n}$ est racine d'un unique polynôme irréductible unitaire $P = P_t$ de $\mathbb{F}_q[X]$ de degré d divisant n, ainsi

$$X^{q^n} - X = \prod_{d|n} \prod_{\substack{P \text{ unitaire} \\ \text{irréductible sur } \mathbb{F}_q \\ \deg P = d}} P(X)$$

(dans cette formule comme dans la suite du paragraphe, on convient que la notation $d|n$ signifie que d est un diviseur POSITIF de n).
PREUVE. Puisque $\mathbb{F}_q[X]$ est factoriel, on applique le lemme 11.2.4 en utilisant que $X^{q^n} - X$ est premier avec sa dérivée, donc sa factorisation est sans multiplicité.

Soit $\nu_n(q)$ le nombre de polynômes unitaires irréductibles de degré n sur $\mathbb{F}_q$. Alors l'identité ci–dessus montre que $q^n = \sum_{d|n} d\nu_d(q)$, et pour récupérer $\nu_d(q)$ de cette formule on utilise la formule d'inversion de Möbius.

Formule d'inversion de Möbius

DÉFINITION 11.2.7 *On appelle fonction de Möbius la fonction définie sur $\mathbb{N}$ par :*

$$\mu(n) = \begin{cases} 1 & \text{si } n = 1, \\ (-1)^k & \text{si } n \text{ est le produit de } k \text{ nombres premiers distincts,} \\ 0 & \text{si } n \text{ est divisible par le carré d'un nombre premier.} \end{cases}$$

On voit que $\mu(nm) = \mu(n)\mu(m)$, si m et n sont premiers entre eux.

REMARQUE On peut aussi définir la fonction de Möbius $\mu(n)$ par l'égalité formelle

$$\sum_{n=1}^{\infty} \mu(n) n^{-s} = \prod_{p \text{ premier}} (1 - p^{-s}) = \zeta(s)^{-1}.$$

PROPOSITION 11.2.8 (FORMULE D'INVERSION) *Soient (a_n), (b_n) $(n \geq 1)$ deux suites d'entiers liées par*

$$b_n = \sum_{d|n} a_d \quad (n \geq 1).$$

Alors on a

$$a_n = \sum_{d|n} \mu(n/d) b_d \quad (n \geq 1).$$

En effet on a

$$\sum_{d|n} \mu(n/d) b_d = \sum_{d|n} \mu(n/d) \sum_{d'|d} a_{d'} = \sum_{d'|n} a_{d'} \sum_{\delta|(n/d')} \mu(\delta),$$

où $\delta = n/d$ divise n/d'. Pour $m > 1$, si s désigne le nombre de diviseurs premiers distincts positifs de m, on a

$$\sum_{\delta|m} \mu(\delta) = \sum_{t=0}^{s} C_t^s (-1)^t = (1-1)^s = 0.$$

REMARQUE La formule d'inversion $a_n = \sum_{d|n} \mu(n/d) b_d$ résulte aussi facilement de l'identité formelle

$$\sum_{n=1}^{\infty} b_n n^{-s} = \sum_{n=1}^{\infty} a_n n^{-s} \zeta(s), \quad \zeta(s) = \sum_{n=1}^{\infty} n^{-s}$$

c'est-à-dire

$$\sum_{n=1}^{\infty} a_n n^{-s} = \sum_{n=1}^{\infty} b_n n^{-s} \zeta(s)^{-1}.$$

Une application directe de la formule d'inversion nous permet d'énoncer :

THÉORÈME 11.2.9 *Pour tout corps fini $\mathbb{F}_q$ et tout entier $n \geq 1$ on a :*

(i) $X^{q^n} - X = \prod_{d|n} \prod_{\substack{P \text{ unitaire} \\ \text{irréductible sur } \mathbb{F}_q \\ \deg P = d}} P(X),$

(ii) $q^n = \sum_{d|n} d\nu_d(q).$

(iii) Le nombre de polynômes irréductibles unitaires de degré n sur $\mathbb{F}_q$ est

$$\nu_n(q) = \frac{1}{n} \sum_{d|n} \mu(n/d) q^d.$$

REMARQUE On voit facilement par l'absurde que l'expression à droite est non nulle, car il y a unicité de l'écriture (éventuelle) d'un entier comme somme de puissances différentes de q. Comme $\nu_n \geq 0$, on obtient $\nu_n > 0$; on a ainsi une nouvelle preuve du fait qu'il existe un polynôme irréductible de degré n sur $\mathbb{F}_q$ (théorème 11.2.1).

EXEMPLE Soit $q = 3$, $n = 2$, alors $\nu_2(3) = \frac{1}{2}(3^2 - 3) = 3$.

```
Factor(T^9-T) mod 3;
```

$$T\,(T+2)\,(T^2+T+2)\,(T^2+2\,T+2)\,(T+1)\,(T^2+1)$$

Ordre d'un polynôme, polynômes primitifs

La notion d'ordre est présentée ici comme complément, les démonstrations sont laissées en exercice (voir [Li-Ni]).

DÉFINITION 11.2.10 *Soit P un polynôme non nul sur $\mathbb{F}_q$. Si $P(0) \neq 0$, l'*ordre *de P est le plus petit entier strictement positif e tel que P divise $X^e - 1$. Si P(0)=0, alors il existe Q dans $\mathbb{F}_q[X]$ non nul en 0 et h entier positif tels que $P = X^h Q$, et dans ce cas on pose* $\mathrm{ord}(P) = \mathrm{ord}(Q)$.

EXERCICE Montrer l'existence d'un tel nombre e, avec $e \leq q^m - 1$ si $m = \deg P \geq 1$ [Indication : raisonner dans l'anneau fini $\mathbb{F}_q[X]/(P)$].

REMARQUE 11.2.11 *Si P est irréductible de degré m sur $\mathbb{F}_q$, alors l'ordre e de P divise $q^m - 1$. De plus d'après le lemme 11.2.4 si $e > 1$ (donc $P(X) \neq X$), m est minimal > 0 pour cette propriété, donc le degré m de P est l'ordre multiplicatif de q modulo e.*

THÉORÈME 11.2.12 *Soient $m \geq 1$ et $e > 1$. Le nombre de polynômes irréductibles unitaires sur $\mathbb{F}_q$ de degré m et d'ordre e est*

$$N_{q,m,e} = \begin{cases} \varphi(e)/m & , \textit{si } m \textit{ est l'ordre multiplicatif de } q \bmod e \\ 0 & , \textit{sinon,} \end{cases}$$

où $\varphi(e)$ est l'indicateur d'Euler de e.

PREUVE (en exercice).

On considère le groupe cyclique $\mathbb{F}_{2^{11}}^*$ d'ordre $2^{11} - 1 = 23 \cdot 89$. Soit $\alpha \in \mathbb{F}_{2^{11}}^*$ un élément d'ordre 23. La factorisation de $X^{23} - 1$ en irréductibles sur $\mathbb{F}_2$ est :

$$X^{23} - 1 = X^{23} + 1 = (X+1)P_0(X)P_1(X) =$$
$$(X+1)\left(X^{11} + X^{10} + X^6 + X^5 + X^4 + X^2 + 1\right)\left(X^{11} + X^9 + X^7 + X^6 + X^5 + X + 1\right)$$

où

$$P_0(X) = X^{11} + X^{10} + X^6 + X^5 + X^4 + X^2 + 1 = \prod_{i \in I}(X - \alpha^i),$$
$$I = \{1, 2, 3, 4, 6, 8, 9, 12, 13, 16, 18\}$$
$$P_1(X) = X^{11} + X^9 + X^7 + X^6 + X^5 + X + 1 = \prod_{j \in J}(X - \alpha^j),$$
$$J = \{5, 7, 10, 11, 14, 15, 17, 19, 20, 21, 22\}.$$

(Notons que toute racine $\neq 1$ de $X^{23} - 1$ est d'ordre 23. Pour écrire les racines de P_0 et P_1 ci-dessus, on a noté α une racine de P_0 et appliqué le théorème 11.2.5.)

Pour $e = 23$ et $q = 2$, on a bien $\mathrm{ord}(2 \bmod 23) = 11$: les polynômes irréductibles d'ordre 23 sur $\mathbb{F}_2$ sont de degré 11, il y en a $\varphi(23)/11 = 2$.

REMARQUE L'ensemble

$$I = \{1, 2, 3, 4, 6, 8, 9, 12, 13, 16, 18\}$$

coïncide avec l'ensemble des résidus quadratiques modulo 23, et l'ensemble complémentaire

$$J = \{5, 7, 10, 11, 14, 15, 17, 19, 20, 21, 22\}$$

coïncide avec l'ensemble des non-résidus quadratiques modulo 23.

Le morphisme de Frobenius $\alpha^k \mapsto \alpha^{2k}$ laisse bien sûr les ensembles d'exposants I et J stables, et en effet on a $\left(\frac{2}{23}\right) = 1$ (voir la loi complémentaire de la loi de réciprocité quadratique. L'application $\alpha^k \mapsto \alpha^{-k}$ échange les ensembles I et J puisque $\left(\frac{-1}{23}\right) = -1$ par le critère d'Euler ; en particulier P_1 est le polynôme minimal de α^{-1} sur $\mathbb{F}_2$.

REMARQUE 11.2.13 *: Le degré de P irréductible sur $\mathbb{F}_q$ d'ordre e est l'ordre multiplicatif de q modulo e.*

En effet, on a $e|q^m - 1$, avec un m minimal. Par exemple, si $e = 23$, $q = 2$, alors $\mathrm{ord}(2) \bmod 23 = 11$.

DÉFINITION 11.2.14 *Un polynôme P de degré m sur $\mathbb{F}_q$ est dit* primitif *sur $\mathbb{F}_q$ s'il est le polynôme minimal sur $\mathbb{F}_q$ d'une racine primitive de $\mathbb{F}_{q^m}$ (un générateur du groupe cyclique $\mathbb{F}_{q^m}^*$).*

Une étude de l'ordre des produits de polynômes fournit la caractérisation suivante, où on voit qu'à degré m fixé ce sont les polynômes primitifs qui atteignent l'ordre maximum $q^m - 1$ (voir l'exercice 11.2.10) :

THÉORÈME 11.2.15 *Un polynôme P de degré m est primitif sur $\mathbb{F}_q$ si et seulement si P est unitaire, non nul en 0 et d'ordre $q^m - 1$.*

Résumé des propriétés des polynômes irréductibles sur $\mathbb{F}_q$

THÉORÈME 11.2.16 *Soit α un élément de $\mathbb{F}_{q^m}$, une extension de $\mathbb{F}_q$. Soient d le degré, et P le polynôme minimal de α sur $\mathbb{F}_q$. Alors,*

(i) P est irréductible sur $\mathbb{F}_q$ et son degré d divise m.

(ii) un polynôme Q sur $\mathbb{F}_q$ s'annule α si et seulement si P divise Q.

(iii) tout polynôme irréductible unitaire sur $\mathbb{F}_q$ nul en α est égal à P.

(iv) P divise $X^{q^d} - X$ et $X^{q^m} - X$.

(v) Les racines de P sont $\alpha, \alpha^q, \cdots, \alpha^{q^{d-1}}$ et P est le polynôme minimal sur $\mathbb{F}_q$ de toutes ces racines.

Si de plus $\alpha \neq 0$, on a :

(vi) l'ordre de P est égal à celui de α dans le groupe multiplicatif $\mathbb{F}_{q^m}^*$.

(vii) P est un polynôme primitif sur $\mathbb{F}_q$ si et seulement si α est d'ordre $q^d - 1$ dans $\mathbb{F}_{q^m}^*$.

EXERCICES

11.1 Ecrire les tables d'addition et de multiplication des corps $\mathbb{F}_4$, $\mathbb{F}_8$ et $\mathbb{F}_9$.

11.2 Ecrire la factorisation de $T^9 - T$ (resp. $T^8 - T$) en irréductibles sur $\mathbb{F}_3$ (resp. sur $\mathbb{F}_2$). Quels sont les facteurs primitifs ?

1.3 Généralités sur les codes

DÉFINITION 1.1 (a) *Soit F un ensemble de cardinal q, M, n deux entiers strictement positifs. On appelle* code C sur l'alphabet F de longueur n *bloc-par-bloc toute partie $C \subset F^n$ de cardinal* $\mathrm{Card}\,(C) = M$.

(b) *Un code $C \subset F^n$ est dit q-aire si $C = \mathrm{Im}E$ pour une application injective*

$$E : F^k \longrightarrow F^n$$

L'élément $E(u)$, pour un u de F^k est appelé un mot code, k est dit la dimension du code, n est sa longueur. Dans ce cas $\mathrm{Card}\,(C) = q^k$

DÉFINITION 1.2 (a) *Soit F un ensemble fini non vide et n entier strictement positif. L'application $d : F^n \times F^n \longrightarrow \mathbb{N}$*

$$(a, b) \mapsto \mathrm{Card}\,\{i \in \{1, \ldots, n\} \mid a_i \neq b_i\}$$

avec $a = (a_1, \cdots, a_n)$ et $b = (b_1, \cdots, b_n)$ est la distance de Hamming sur F^n.

(b) *Soit F un corps fini. L'application $w : F^n \to \mathbb{N}$*

$$a \mapsto d(a, \underline{0}) = \mathrm{Card}\,\{i \in \{1, \ldots, n\} \mid a_i \neq 0\}$$

est le poids de Hamming.

REMARQUE 1.3 *: La distance de Hamming sur F^n est bien une distance sur F^n. En effet, on a :*

$$d(a, b) = 0 \Longleftrightarrow (a_i = b_i; 1 \leq i \leq n) \Longleftrightarrow a = b$$

$$d(a, b) = d(b, a) \text{ pour tous } a, b \text{ pour tous } E^n$$

$$d(a, c) \leq d(a, b) + d(b, c) \text{ pour tous } a, b, c \text{ pour tous } E^n.$$

DÉFINITION 1.4 *Soit $C = \mathrm{Im}E$ un code q-aire, i.e., l'image d'une application injective $E : F^k \to F^n$, où $\#F = q$.*

(a) *On appelle écart ou "distance" de C le nombre*

$$d = d(C) = \min_{\substack{x,y \in F^k \\ x \neq y}} d(E(x), E(y))$$

Soit $C = \mathrm{Im}E$ un code q-aire, $E : F^k \to F^n$ de distance d. Dans ce cas on dit que C est un $[n, k, d]_q$-code.

(b) *On appelle vitesse de transmission (le "rendement" ou "information rate" en anglais) de C le rapport $R = k/n$*

(c) *$1/R$ est le coefficient de redondance de C.*

(d) *$\delta = d/n$ est la distance relative (ou le "taux de correction") de C*

Soit $x = (x_1, \ldots, x_n)$ le mot transmis par le canal. Le *mot reçu* $y = (y_1, y_2, \ldots, y_n)$, eventuellement entaché d'erreurs, diffère de x en $d(x, y)$ positions.

Lorsque on suppose qu'il n'y a pas plus de t erreurs commises, $d(x, y) \leq t$, on pourra retrouver x à la condition que chaque mot erroné reçu ne puisse provenir que d'un seul mot du code.

DÉFINITION 1.5 *Un code de longueur n sur l'alphabet F vérifie la* condition de decodage *d'ordre t si pour tout $y \in F^n$ il existe au plus un mot $x \in C \subset F^n$ tel que $d(x, y) \leq t$. Dans ce cas les boules*

$$B(x, t) = \{y \in F^n \mid d(x, y) \leq t\} \subset F^n$$

(pour la distance de Hamming) sont deux-à-deux disjointes.

THÉORÈME 1.6 *Un code C peut corriger t erreurs si son écart d est tel que $d \geq 2t + 1$.*

Preuve Si c est envoyé et y reçu, tels que $d(y, c) \leq t$, tout mot code c' de C est tel que $d(c, c') \geq 2t + 1$. Or, d est une distance, donc

$$d(y, c') \geq d(c, c') - d(c, y)$$

$$d(y, c') \geq t + 1$$

C peut donc corriger t erreurs. Les boules $B(c, t)$ sont donc deux-à-deux disjointes.

DÉFINITION 1.7 (a) *Un décodage de E est une application*

$$D : F^n \to F^k$$

telle que $D \circ E = Id_{F^k}$.

(b) *On dit que D est standard ("de vraisemblence maximale") si*

$$\forall y \in F^n \forall u \in F^k, \quad d(E(u), y) \geq d(E(D(y)), y)$$

c'est-à-dire que $E(D(y))$ se trouve parmi les mots de codes $E(u)$ les plus proches de y.

REMARQUE 1.8 *L'existence d'un décodage est garantie par l'injectivité de E.*

1.4 Codage et decodage optimal sur un canal bruité

On considère les problèmes mathématiques de transmission d'information (vue comme une longue suite

$$F^{kN} \ni a = (a_1, \ldots, a_k, a_{k+1}, \ldots, a_{kN}),$$

où

$$u_1 = (a_1, \ldots, a_k), u_2 = (a_{k+1}, \cdots, a_{2k}), \cdots, u_N = (a_{k(N-1)+1}, \cdots, a_{kN-1}, a_{kN}),$$

avec les *blocs d'information* $u_1, \ldots u_N$. On transmet l'information bloc-par-bloc a l'aide du code

$$a \mapsto E_N(a) = (E(u_1), \ldots, E(u_N)) \in F^{nN}$$

avec les hypothèses 1.2 :

$$E_N(a) \mapsto \tilde{E}_N(a) \in F^{nN}$$

Il est possible de diminuer considerablement la proportion d'erreures de transmission d'information avec des bons codes.

3 Codes linéaires et codes cycliques. Matrice génératrice et calcul du syndrome

d'erreur

3.1 Codes linéaires

Une classe de codes très importante est celle des codes linéaires, notamment en raison des outils dont nous disposons pour manipuler et représenter les applications linéaires à l'aide de l'écriture matricielle.

En général, pour un alphabet fini F, étant donné une énumération de F^k, la donnée d'un code $C = \mathrm{Im}(E)$, $E : F^k \to F^n$ est la donnée de $n \times q^k$ éléments de F, ce qui représente un gros volume d'information.

Si l'on munit F^k et F^n de structures supplémentaires et si l'on prend une application E qui respecte ces structures, on peut économiser sur le volume d'information représentant E au prix de calculs des valeurs non mémorisées de E.

En ce sens le plus simple est de prendre pour q un nombre primaire, $q = p^r$, pour F le corps $\mathbb{F}_q$ et pour E une application linéaire injective de $\mathbb{F}_q^k$ dans $\mathbb{F}_q^n$. Rapportant C aux bases canoniques adéquates, on caractérise cette application par $n \times k$ éléments de $\mathbb{F}_q$.

Rapellons
DÉFINITION 1.2 (a) Soit F un ensemble fini non vide et n entier strictement positif. L'application $d : F^n \times F^n \longrightarrow \mathbb{N}$

$$(a, b) \mapsto \mathrm{Card}\,\{i \in \{1, \ldots, n\} \mid a_i \neq b_i\}$$

avec $a = (a_1, \cdots, a_n)$ et $b = (b_1, \cdots, b_n)$ est la *distance de Hamming* sur F^n.
(b) Soit F un corps fini. L'application $w : F^n \to \mathbb{N}$

$$a \mapsto d(a, \underline{0}) = \mathrm{Card}\,\{i \in \{1, \ldots, n\} \mid a_i \neq 0\}$$

est le *poids de Hamming*.

DÉFINITION 3.1 *Soit $F = \mathbb{F}_q$ un corps fini. Un code linéaire C est un sous-espace vectoriel de dimension k de l'espace vectoriel F^n (vu comme l'image d'une application $E : F^k \to F^n$ linéaire injective). Les vecteurs lignes $a = (a_1, \cdots, a_k) \in F^k$ sont les* mots d'information, *et les vecteurs lignes $c = E(a) = (c_1, \cdots, c_n) \in F^n$ sont les* mots de code. *La* matrice génératrice *G du code E est la matrice attachée à l'application linéaire $E : F^k \to F^n$ (dans les bases standards de F^k et F^n), de telle façon que*

$$c = E(a) = aG.$$

REMARQUE 3.2 *Un code linéaire C est l'image d'une application linéaire injective, donc on peut considérer C comme un sous-espace vectoriel de dimension k de l'espace vectoriel F^n. On peut ainsi caractériser les codes linéaires à partir de matrices à coefficients dans F comme* noyau *d'une autre application linéaire $S : F^n \to F^{n-k}$. La matrice H de S est appelée* matrice de contrôle *de C :*

$$S(c) = Hc^t.$$

EXEMPLE. Soit H une matrice $(n-k) \times n$ à coefficients dans $\mathbb{F}_q$ de rang $n-k$. Le noyau de l'application représentée par H est un sous-espace vectoriel de $\mathbb{F}_q^n$. On peut donc définir un code linéaire par un système d'équations linéaires :

$$C = \Big\{ c = (c_1, \cdots, c_n) \;\Big|\; Hc^t = 0 \Big\}.$$

Posons

$$H = (A, I_{n-k}) = \begin{pmatrix} 1 & 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 1 \end{pmatrix}$$

et soit $q = 2$ (C code binaire) On désire transmettre le message

$$a = (a_1 a_2 a_3 a_4).$$

On le code en $c = (a_1 a_2 a_3 a_4 c_5 c_6 c_7)$, avec c_5, c_6, c_7 tels que $Hc^t = 0$. Or,

$$Hc^t = 0 \iff \begin{array}{l} a_1 + a_3 + a_4 + c_5 = 0 \\ a_1 + a_2 + a_4 + c_6 = 0 \\ a_1 + a_2 + a_3 + c_7 = 0 \end{array} \iff \begin{array}{l} c_5 = a_1 + a_3 + a_4 \\ c_6 = a_1 + a_2 + a_4 \\ c_7 = a_1 + a_2 + a_3 \end{array}$$

On obtient l'application linéaire injective $E : \mathbb{F}_2^4 \to \mathbb{F}_2^7$

$$(a_1, a_2, a_3, a_4) \longmapsto (a_1, a_2, a_3, a_4, a_1 + a_3 + a_4, a_1 + a_2 + a_4, a_1 + a_2 + a_3)$$

La matrice H est appelée *matrice de contrôle* de C. On a $Hc^t = 0$ pour tous les mots de code $c \in C$.

REMARQUE 3.3 *Si $H = (A, I_{n-k})$, alors un message $a = a_1 \cdots a_k$ est codé en $c = a_1 \cdots a_k c_{k+1} \cdots c_n$ le code est alors dit systématique. Ici $A \in \mathrm{Mat}_{n-k,k}(F)$*

De plus, on a

$$\{Hc^t = 0\} \Longrightarrow c^t = \begin{pmatrix} I_k \\ -A \end{pmatrix} a^t = [a(I_k, -A^t)]^t$$

c'est-à-dire, que

$$\begin{pmatrix} c_{k+1} \\ c_{k+2} \\ \cdots \\ c_n \end{pmatrix} = -A \begin{pmatrix} a_1 \\ \cdots \\ a_k \end{pmatrix}, \begin{pmatrix} c_{k+1} \\ c_{k+2} \\ \cdots \\ c_n \end{pmatrix} + A \begin{pmatrix} a_1 \\ \cdots \\ a_k \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \cdots \\ 0 \end{pmatrix}.$$

Il vient la définition suivante :

DÉFINITION 3.4 *$G = (I_k, -A^t)$ est la matrice génératrice canonique du code linéaire C de matrice de contrôle $H = (A, I_{n-k})$. D'une manière plus générale, toute matrice G engendrant un code C est une matrice génératrice de C.*

REMARQUE 3.5 *Pour tout mot code c, on a $Hc^t = 0$ et $c = aG$. Donc*

$$GH^t = 0 \in \mathrm{Mat}_{k,n-k}(F),\ HG^t = 0 \in \mathrm{Mat}_{n-k,k}(F),$$

puisque $Hc^t = HG^t a^t = 0$ pour tous les $a \in F^k$.

3.2 Détection et correction d'erreurs, décodage

Dans ce qui suit, nous noterons c un mot code émis, y le message reçu, et $e = y - c$ le vecteur erreur.

La distance de Hamming $d(y, c)$ est alors le nombre d'erreurs survenues au cours de la transmission. Pour décoder y reçu, on peut supposer que le nombre d'erreurs est minimal, c'est à dire que l'on va chercher le mot code c le plus proche de y au sens de la distance de Hamming. C'est la règle du décodage *par plus proche voisin.*

DÉFINITION 3.6 *Soit t un entier naturel. C un code linéaire de dimension r et de longueur n est dit t-correcteur d'erreurs si*

$$\forall y \in \mathbb{F}_q^n, |\{c \in C : d(y, c) \leq t\}| \leq 1$$

Si alors $c \in C$ est transmis et qu'au plus t erreurs surviennent, on a $d(y, c) \leq t$ et $d(y, c') > t$ pour tout autre élément de C. Ainsi, la méthode du décodage par plus proche voisin donne le bon résultat.

Il apparaît qu'un des objectifs de la théorie du codage consiste à élaborer des codes dont les mots sont très éloignés les uns des autres au sens de la distance de Hamming. Toutefois, un autre est de transmettre un maximum d'information et donc de garder des vitesses de transmission acceptables, et réunir les deux est épineux.

THÉORÈME 3.7 *Un code C peut corriger t erreurs si son écart d est tel que $d \geq 2t + 1$*

PREUVE : On a déja vu ce resultat (Théorème 1.6) Si c est envoyé et y reçu, tels que $d(y, c) \leq t$, tout mot code c' de C est tel que $d(c, c') \geq 2t+1$. Or, d est une distance, donc

$$d(y, c') \geq d(c, c') - d(c, y)$$

$$d(y, c') \geq t + 1$$

C peut donc corriger t erreurs.

EXEMPLE. Reprenons le code déjà vu plus haut

$$\begin{aligned} E : \quad & \mathbb{F}_2^4 \to \mathbb{F}_2^7 \\ & (a_1, a_2, a_3, a_4) \mapsto (a_1, a_2, a_3, a_4, a_1 + a_3 + a_4, a_1 + a_2 + a_4, a_1 + a_2 + a_3) \end{aligned} \tag{3.1}$$

Soient a, b des éléments de $\mathbb{F}_2^4$
Si $d(a, b) = 1$, alors $d(E(a), E(b)) = 3$ ou 4.
Si $d(a, b) = 2$, alors $d(E(a), E(b)) = 3$ ou 4.
Si $d(a, b) = 3$, alors $d(E(a), E(b)) = 3$ ou 4.
Si $d(a, b) = 4$, alors $d(E(a), E(b)) = 7$.
Ceci dit, l'application E écarte vraiment les mots d'information.

En effet, on peut toujours supposer $b = (0, 0, 0, 0)$, et on utilise directement la formule (3.1) pour le mot $E(a)$.

Donc $d = 3$, et le code corrige 1 erreur.

LEMME 3.8 *Soit un code linéaire C de matrice de correction H et d'écart d. Alors $d \geq s+1$ si et seulement si s colonnes de H sont linéairement indépendantes.*

PREUVE : Supposons que s colonnes de H soient linéairement dépendantes. Alors il existe $c \in C$ non nul tel que $Hc^t = 0$ et $w(c) \leq s$. Ainsi, $d \leq s$. Inversement, si s colonnes de H sont toujours indépendantes, $c \in C$ non nul est toujours tel que $w(c) \geq s+1$ et donc $d \geq s+1$.

Ce qui suit est un algorithme simple de décodage des codes linéaires : le décodage par *leader de classe*.

Soit C un code linéaire de longueur n et de dimension k sur $\mathbb{F}_q$. L'espace vectoriel $\mathbb{F}_q^n/C$ est formé de toutes les classes $a+C$, $a \in \mathbb{F}_q^n$. Pour tout a, $|a+C| = q^k$, et

$$\mathbb{F}_q^n = (a^{(0)} + C) \cup \cdots \cup (a^{(s)} + C),$$

avec $a^{(0)} = 0$, $s = q^{n-k} - 1$.

Alors, quel que soit le message y reçu, il existe i tel que $y \in a^{(i)}+C$, et si c est le message envoyé, $e = y - c = a^{(i)} + z \in a^{(i)} + C$. On peut ainsi construire une méthode de décodage des codes linéaires. En effet, quel que soit $y \in a^{(i)} + C$ reçu, tous les vecteurs erreur possibles pour y sont également dans $a^{(i)} + C$. La règle de décodage par plus proche voisin nous conduit à choisir pour vecteur erreur le vecteur $e \in a^{(i)} + C$ de poids de Hamming minimum, et on décode y en $x = y - e$. Nous allons voir maintenant l'algorithme, à proprement parler, plus en détail.

Ecrivons H dont la $j^{\text{ème}}$ colonne est α^{j-1} exprimé dans la base $\{1, \alpha, \alpha^2, \alpha^3\}$, $0 \leq j \leq 14$. Il vient

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

Si alors $a(x) = a_0 + \cdots + a_{10}x^{10}$ est le message à transmettre, il sera codé en $w(x) = a(x)(x^4+x+1)$. Supposons qu'une erreur survienne au cours de la transmission, le message reçu est alors $v(x) = w(x) + x^{e-1}$. Son syndrome est $S(v) = v(\alpha) = w(\alpha) + \alpha^{e-1} = \alpha^{e-1}$, et on sait qu'une erreur est survenue en $e^{\text{ème}}$ position.

4.2 Exemples : codes de Golay

Les calculs suivants sont disponibles à l'adresse cachée :
`http ://www-fourier.ujf-grenoble.fr/~panchish/04mag-maple` dans le fichier 4mag-7cycl-gol.mws :

4.2.1. Code G_{23}

On considère le groupe cyclique $\mathbb{F}_{2^{11}}^*$ d'ordre $2^{11} - 1 = 23 \cdot 89$. Soit $\alpha \in \mathbb{F}_{2^{11}}^*$ une racine primitive de degré 23. On pose

$$G_{23} = \left\{ x = (x_0, \cdots, x_{22}) \in \mathbb{F}_2^{23} \mid \sum_{i=0}^{22} x_i \alpha^i = 0 \right\} \subset \mathbb{F}_2[x]/(x^{23} - 1).$$

On a n=23, $q = 2$,

$$x^{23} - 1 = x^{23} + 1 = (x+1)g_0(x)g_1(x) =$$
$$(x+1)\left(x^{11} + x^{10} + x^6 + x^5 + x^4 + x^2 + 1\right)\left(x^{11} + x^9 + x^7 + x^6 + x^5 + x + 1\right)$$

où

$$\begin{aligned} g_0(x) &= x^{11}+x^{10}+x^6+x^5+x^4+x^2+1 \\ &= \prod_{i\in I}(x-\alpha^i), I=\{1,2,3,4,6,8,9,12,13,16,18\} \\ g_1(x) &= x^{11}+x^9+x^7+x^6+x^5+x+1 \\ &= \prod_{j\in J}(x-\alpha^j), J=\{5,7,10,11,14,15,17,19,20,21,22\} \end{aligned}$$

Remarque. L'ensemble

$$I=\{1,2,3,4,6,8,9,12,13,16,18\}$$

coïncide avec l'ensemble des résidues quadratiques modulo 23, et l'ensemble complémentaire

$$J=\{5,7,10,11,14,15,17,19,20,21,22\}$$

coïncide avec l'ensemble des non-résidues quadratiques modulo 23.

L'application de Frobenius $\alpha^k \mapsto \alpha^{2k}$ laisse I et J stable puisque $\left(\frac{2}{23}\right)=1$, et l'application $\alpha^k \mapsto \alpha^{-k}$ échange les ensembles I et J puisque $\left(\frac{-1}{23}\right)=-1$, grâce à la loi de

réciprocité quadratique de Gauss : pour les nombres premiers positifs impairs p, q on a

$$\left(\frac{p}{q}\right)\left(\frac{q}{p}\right)=(-1)^{\frac{(p-1)}{2}\frac{(q-1)}{2}},$$

et on a les deux compléments suivants de cette loi :

$$\left(\frac{2}{p}\right)=(-1)^{(p^2-1)/8}, \qquad \left(\frac{-1}{p}\right)=(-1)^{(p-1)/2},$$

DÉFINITION 4.3 *Code de Golay G_{23} est un sous-espace vectoriel (g_0) de dimension 12 dans le quotient*

$$\mathbb{F}_2[x]/(x^{23}-1)$$

vu comme un espace vectoriel de dimension 23 sur $\mathbb{F}_2$ avec le polynôme générateur

$$g_0(x) = x^{11}+x^{10}+x^6+x^5+x^4+x^2+1$$

et avec le polynome de contrôle

$$h(x) = (x+1)\left(x^{11}+x^9+x^7+x^6+x^5+x+1\right) = (x+1)g_1(x)$$

C'est un $[23,12,7]_2$-code.

```
> restart ;
> with(linalg) :
```

```
Warning, the protected names norm and trace have been redefined and
unprotected
```

```
> Factor(x^23+1) mod 2;
```

$$(x+1)\left(x^{11}+x^{10}+x^6+x^5+x^4+x^2+1\right)\left(x^{11}+x^9+x^7+x^6+x^5+x+1\right)$$

```
> g:=x^11+x^10+x^6+x^5+x^4+x^2+1;irreduc(g) mod 2;
```

$$g := x^{11}+x^{10}+x^6+x^5+x^4+x^2+1$$

true

```
> alias(alpha = RootOf(g)) ;
```

$$\alpha$$

```
> Factor(g,alpha) mod 2;
```

$$(x+\alpha^9)\left(x+\alpha^6\right)\left(x+\alpha^9+\alpha^8+\alpha^6+\alpha^5+\alpha^2+\alpha\right)\left(x+\alpha^4\right)$$
$$\left(x+\alpha^8+\alpha^7+\alpha^6+\alpha^5+\alpha^3+\alpha^2+1\right)\left(x+\alpha^{10}+\alpha^8+\alpha^6+\alpha^3+\alpha+1\right)$$
$$\left(x+\alpha^{10}+\alpha^7+\alpha^4+\alpha^3+\alpha^2+\alpha+1\right)(x+\alpha)\left(x+\alpha^2\right)\left(x+\alpha^8\right)\left(x+\alpha^3\right)$$

```
> for i from 0 to 23 do
> if Eval(g, x=alpha^i) mod 2 = 0 then Expand (alpha^i) mod 2;
> print('i'=i, alpha^i=Expand (alpha^i) mod 2, 'g'(alpha^i)=0) fi
> od ;
```

$$i = 1,\ \alpha = \alpha,\ \mathrm{g}(\alpha) = 0$$
$$i = 2,\ \alpha^2 = \alpha^2,\ \mathrm{g}(\alpha^2) = 0$$
$$i = 3,\ \alpha^3 = \alpha^3,\ \mathrm{g}(\alpha^3) = 0$$
$$i = 4,\ \alpha^4 = \alpha^4,\ \mathrm{g}(\alpha^4) = 0$$
$$i = 6,\ \alpha^6 = \alpha^6,\ \mathrm{g}(\alpha^6) = 0$$
$$i = 8,\ \alpha^8 = \alpha^8,\ \mathrm{g}(\alpha^8) = 0$$

$$i = 9,\ \alpha^9 = \alpha^9,\ \mathrm{g}(\alpha^9) = 0$$

$$i = 12,\ \alpha^{12} = \alpha^{10} + \alpha^7 + \alpha^4 + \alpha^3 + \alpha^2 + \alpha + 1,\ \mathrm{g}(\alpha^{12}) = 0$$

$$i = 13,\ \alpha^{13} = \alpha^{10} + \alpha^8 + \alpha^6 + \alpha^3 + \alpha + 1,\ \mathrm{g}(\alpha^{13}) = 0$$

$$i = 16,\ \alpha^{16} = \alpha^9 + \alpha^8 + \alpha^6 + \alpha^5 + \alpha^2 + \alpha,\ \mathrm{g}(\alpha^{16}) = 0$$

$$i = 18,\ \alpha^{18} = \alpha^8 + \alpha^7 + \alpha^6 + \alpha^5 + \alpha^3 + \alpha^2 + 1,\ \mathrm{g}(\alpha^{18}) = 0$$

```
>  for i from 0 to 23 do
>  if Eval(g, x=alpha^(-i)) mod 2 = 0 then Expand (alpha^(-i)) mod 2;
>  print('i'=i, alpha^(i)=Expand (alpha^i) mod 2, 'g'(alpha^(-i))=0) fi
>  od ;
```

$$i = 5,\ \alpha^5 = \alpha^5,\ \mathrm{g}(\frac{1}{\alpha^5}) = 0$$

$$i = 7,\ \alpha^7 = \alpha^7,\ \mathrm{g}(\frac{1}{\alpha^7}) = 0$$

$$i = 10,\ \alpha^{10} = \alpha^{10},\ \mathrm{g}(\frac{1}{\alpha^{10}}) = 0$$

$$i = 11,\ \alpha^{11} = \alpha^{10} + \alpha^6 + \alpha^5 + \alpha^4 + \alpha^2 + 1,\ \mathrm{g}(\frac{1}{\alpha^{11}}) = 0$$

$$i = 14,\ \alpha^{14} = \alpha^{10} + \alpha^9 + \alpha^7 + \alpha^6 + \alpha^5 + \alpha + 1,\ \mathrm{g}(\frac{1}{\alpha^{14}}) = 0$$

$$i = 15,\ \alpha^{15} = \alpha^8 + \alpha^7 + \alpha^5 + \alpha^4 + \alpha + 1,\ \mathrm{g}(\frac{1}{\alpha^{15}}) = 0$$

$$i = 17,\ \alpha^{17} = \alpha^{10} + \alpha^9 + \alpha^7 + \alpha^6 + \alpha^3 + \alpha^2,\ \mathrm{g}(\frac{1}{\alpha^{17}}) = 0$$

$$i = 19,\ \alpha^{19} = \alpha^9 + \alpha^8 + \alpha^7 + \alpha^6 + \alpha^4 + \alpha^3 + \alpha,\ \mathrm{g}(\frac{1}{\alpha^{19}}) = 0$$

$$i = 20,\ \alpha^{20} = \alpha^{10} + \alpha^9 + \alpha^8 + \alpha^7 + \alpha^5 + \alpha^4 + \alpha^2,\ \mathrm{g}(\frac{1}{\alpha^{20}}) = 0$$

$$i = 21,\ \alpha^{21} = \alpha^9 + \alpha^8 + \alpha^4 + \alpha^3 + \alpha^2 + 1,\ \mathrm{g}(\frac{1}{\alpha^{21}}) = 0$$

$$i = 22,\ \alpha^{22} = \alpha^{10} + \alpha^9 + \alpha^5 + \alpha^4 + \alpha^3 + \alpha,\ \mathrm{g}(\frac{1}{\alpha^{22}}) = 0$$

```
>  g[1]:=x^11+x^9+x^7+x^6+x^5+x+1;Factor(g[1],alpha) mod 2;
```

$$g_1 := x^{11} + x^9 + x^7 + x^6 + x^5 + x + 1$$

$$(x+\alpha^{10}+\alpha^9+\alpha^7+\alpha^6+\alpha^3+\alpha^2)\,(x+\alpha^{10}+\alpha^6+\alpha^5+\alpha^4+\alpha^2+1)\,(x+\alpha^7)$$
$$(x+\alpha^8+\alpha^7+\alpha^5+\alpha^4+\alpha+1)\,(x+\alpha^{10})\,(x+\alpha^5)\,(x+\alpha^{10}+\alpha^9+\alpha^5+\alpha^4+\alpha^3+\alpha)$$
$$(x+\alpha^9+\alpha^8+\alpha^7+\alpha^6+\alpha^4+\alpha^3+\alpha)\,(x+\alpha^9+\alpha^8+\alpha^4+\alpha^3+\alpha^2+1)$$
$$(x+\alpha^{10}+\alpha^9+\alpha^7+\alpha^6+\alpha^5+\alpha+1)\,(x+\alpha^{10}+\alpha^9+\alpha^8+\alpha^7+\alpha^5+\alpha^4+\alpha^2)$$

Code de Golay $G_{23}=(g)$ est un sous-espace vectoriel de dimension 12 dans le quotient

$$\mathbb{F}_2[x]/(x^{23}-1)$$

(vu comme un espace vectoriel de dimension 23 sur $\mathbb{F}_2$) avec le polynôme générateur

$g := x^{11}+x^{10}+x^6+x^5+x^4+x^2+1$
et avec le polynome de contrôle $h=(x+1)\,(x^{11}+x^9+x^7+x^6+x^5+x+1)$,
$h=x^{12}+x^{11}+x^{10}+x^9+x^8+x^5+x^2+1$.

C'est un $[23,12,7]_2$-code.

```
> for i from 0 to 23 do
> if Eval(g[1], x=alpha^i) mod 2 = 0 then Expand (alpha^i) mod 2;
> print('i'=i, alpha^i=Expand (alpha^i) mod 2, 'g[1]'(alpha^i)=0) fi
> od ;
```


$$i=5,\ \alpha^5=\alpha^5,\ g_1(\alpha^5)=0$$
$$i=7,\ \alpha^7=\alpha^7,\ g_1(\alpha^7)=0$$
$$i=10,\ \alpha^{10}=\alpha^{10},\ g_1(\alpha^{10})=0$$
$$i=11,\ \alpha^{11}=\alpha^{10}+\alpha^6+\alpha^5+\alpha^4+\alpha^2+1,\ g_1(\alpha^{11})=0$$
$$i=14,\ \alpha^{14}=\alpha^{10}+\alpha^9+\alpha^7+\alpha^6+\alpha^5+\alpha+1,\ g_1(\alpha^{14})=0$$
$$i=15,\ \alpha^{15}=\alpha^8+\alpha^7+\alpha^5+\alpha^4+\alpha+1,\ g_1(\alpha^{15})=0$$
$$i=17,\ \alpha^{17}=\alpha^{10}+\alpha^9+\alpha^7+\alpha^6+\alpha^3+\alpha^2,\ g_1(\alpha^{17})=0$$
$$i=19,\ \alpha^{19}=\alpha^9+\alpha^8+\alpha^7+\alpha^6+\alpha^4+\alpha^3+\alpha,\ g_1(\alpha^{19})=0$$
$$i=20,\ \alpha^{20}=\alpha^{10}+\alpha^9+\alpha^8+\alpha^7+\alpha^5+\alpha^4+\alpha^2,\ g_1(\alpha^{20})=0$$
$$i=21,\ \alpha^{21}=\alpha^9+\alpha^8+\alpha^4+\alpha^3+\alpha^2+1,\ g_1(\alpha^{21})=0$$
$$i=22,\ \alpha^{22}=\alpha^{10}+\alpha^9+\alpha^5+\alpha^4+\alpha^3+\alpha,\ g_1(\alpha^{22})=0$$

```
> (x^11+x^10+x^6+x^5+x^4+x^2+1)*(x+1)*(x^11+x^9+x^7+x^6+x^5+x+1);g:=x^1
> 1+x^10+x^6+x^5+x^4+x^2+1;
```

$$(x^{11}+x^{10}+x^6+x^5+x^4+x^2+1)\,(x+1)\,(x^{11}+x^9+x^7+x^6+x^5+x+1)$$
$$g := x^{11}+x^{10}+x^6+x^5+x^4+x^2+1$$

```
> G:= matrix(12, 23,
> [[1,0,1,0,1,1,1,0,0,0,1,1,0,0,0,0,0,0,0,0,0,0,0],
> [0,1,0,1,0,1,1,1,0,0,0,1,1,0,0,0,0,0,0,0,0,0,0],
> [0,0,1,0,1,0,1,1,1,0,0,0,1,1,0,0,0,0,0,0,0,0,0],
> [0,0,0,1,0,1,0,1,1,1,0,0,0,1,1,0,0,0,0,0,0,0,0],
> [0,0,0,0,1,0,1,0,1,1,1,0,0,0,1,1,0,0,0,0,0,0,0],
> [0,0,0,0,0,1,0,1,0,1,1,1,0,0,0,1,1,0,0,0,0,0,0],
> [0,0,0,0,0,0,1,0,1,0,1,1,1,0,0,0,1,1,0,0,0,0,0],
> [0,0,0,0,0,0,0,1,0,1,0,1,1,1,0,0,0,1,1,0,0,0,0],
> [0,0,0,0,0,0,0,0,1,0,1,0,1,1,1,0,0,0,1,1,0,0,0],
> [0,0,0,0,0,0,0,0,0,1,0,1,0,1,1,1,0,0,0,1,1,0,0],
> [0,0,0,0,0,0,0,0,0,0,1,0,1,0,1,1,1,0,0,0,1,1,0],
> [0,0,0,0,0,0,0,0,0,0,0,1,0,1,0,1,1,1,0,0,0,1,1]]) ;
```


$$G := \begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 \end{bmatrix}$$

```
> transpose(G);
```

$$\begin{bmatrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 \\
1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 0 & 1 & 0 \\
1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 0 & 1 \\
0 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 0 \\
0 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 \\
0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \\
0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 \\
0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{bmatrix}$$

123

```
> (x^11+x^10+x^6+x^5+x^4+x^2+1)*(x+1)*(x^11+x^9+x^7+x^6+x^5+x+1);
```

$$(x^{11}+x^{10}+x^6+x^5+x^4+x^2+1)\,(x+1)\,(x^{11}+x^9+x^7+x^6+x^5+x+1)$$

```
> 'h'=(x+1)*(x^11+x^9+x^7+x^6+x^5+x+1);
> h:=Expand((x+1)*(x^11+x^9+x^7+x^6+x^5+x+1)) mod 2;
```

$$h = (x+1)\,(x^{11}+x^9+x^7+x^6+x^5+x+1)$$

$$h := x^{12}+x^{11}+x^{10}+x^9+x^8+x^5+x^2+1$$

```
> 'h'=(x+1)*(x^11+x^9+x^7+x^6+x^5+x+1);'h'=sort(h,x);
```

$$h = (x+1)\,(x^{11}+x^9+x^7+x^6+x^5+x+1)$$

$$h = x^{12}+x^{11}+x^{10}+x^9+x^8+x^5+x^2+1$$

```
> Expand(h*g) mod 2;
```

$$1+x^{23}$$

```
> H:= matrix(11, 23,
> [[0,0,0,0,0,0,0,0,0,0,1,1,1,1,1,0,0,1,0,0,1,0,1],
> [0,0,0,0,0,0,0,0,0,0,1,1,1,1,1,0,0,1,0,0,1,0,1,0],
> [0,0,0,0,0,0,0,0,0,1,1,1,1,1,0,0,1,0,0,1,0,1,0,0],
> [0,0,0,0,0,0,0,0,1,1,1,1,1,0,0,1,0,0,1,0,1,0,0,0],
> [0,0,0,0,0,0,0,1,1,1,1,1,0,0,1,0,0,1,0,1,0,0,0,0],
> [0,0,0,0,0,0,1,1,1,1,1,0,0,1,0,0,1,0,1,0,0,0,0,0],
> [0,0,0,0,0,1,1,1,1,1,0,0,1,0,0,1,0,1,0,0,0,0,0,0],
> [0,0,0,0,1,1,1,1,1,0,0,1,0,0,1,0,1,0,0,0,0,0,0,0],
> [0,0,0,1,1,1,1,1,0,0,1,0,0,1,0,1,0,0,0,0,0,0,0,0],
> [0,0,1,1,1,1,1,0,0,1,0,0,1,0,1,0,0,0,0,0,0,0,0,0],
> [0,1,1,1,1,1,0,0,1,0,0,1,0,1,0,0,0,0,0,0,0,0,0,0],
> [1,1,1,1,1,0,0,1,0,0,1,0,1,0,0,0,0,0,0,0,0,0,0,0]]) ;
```

$$H := \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$


```
> K:=multiply(H,transpose(G)) ;
```

$$K := \begin{bmatrix} 2 & 2 & 2 & 2 & 2 & 2 & 4 & 4 & 4 & 4 & 4 & 4 \\ 2 & 2 & 2 & 2 & 2 & 4 & 4 & 4 & 4 & 4 & 4 & 4 \\ 2 & 2 & 2 & 2 & 4 & 4 & 4 & 4 & 4 & 4 & 4 & 2 \\ 2 & 2 & 2 & 4 & 4 & 4 & 4 & 4 & 4 & 4 & 2 & 2 \\ 2 & 2 & 4 & 4 & 4 & 4 & 4 & 4 & 4 & 2 & 2 & 2 \\ 2 & 4 & 4 & 4 & 4 & 4 & 4 & 4 & 2 & 2 & 2 & 2 \\ 4 & 4 & 4 & 4 & 4 & 4 & 4 & 2 & 2 & 2 & 2 & 2 \\ 4 & 4 & 4 & 4 & 4 & 4 & 2 & 2 & 2 & 2 & 2 & 2 \\ 4 & 4 & 4 & 4 & 4 & 2 & 2 & 2 & 2 & 2 & 2 & 0 \\ 4 & 4 & 4 & 4 & 2 & 2 & 2 & 2 & 2 & 2 & 0 & 2 \\ 4 & 4 & 4 & 2 & 2 & 2 & 2 & 2 & 2 & 0 & 2 & 0 \end{bmatrix}$$

Les termes de la matrice obtenue ne sont pas « réduits » à leur forme canonique dans $\mathrm{GF}(2^{11}) = F_2[\ \alpha]$. Pour obtenir la réduction, on utilise :

```
> map(item -> Expand(item) mod 2, K) ;
```

$$\begin{bmatrix} 0&0&0&0&0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0&0&0&0&0 \end{bmatrix}$$

$\%1 := [0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0]$

Une autre matrice de contrôle :
on considère la matrice dont la j-me colonne est formée par les coordonnées de $\alpha^{j-1}(j = 1, 2, ..., 23)$ dans la base

$$\langle 1, \alpha, \alpha^2, \ldots, \alpha^{10} \rangle$$

$GF(2^{11}) = \mathbb{F}_2[\alpha]$.


```
> for j from 1 to 23 do print(alpha^(j-1)=Expand(alpha^(j-1)) mod 2) ;
> od;
```

$$1 = 1$$
$$\alpha = \alpha$$
$$\alpha^2 = \alpha^2$$
$$\alpha^3 = \alpha^3$$
$$\alpha^4 = \alpha^4$$
$$\alpha^5 = \alpha^5$$
$$\alpha^6 = \alpha^6$$
$$\alpha^7 = \alpha^7$$
$$\alpha^8 = \alpha^8$$
$$\alpha^9 = \alpha^9$$
$$\alpha^{10} = \alpha^{10}$$
$$\alpha^{11} = \alpha^{10} + \alpha^6 + \alpha^5 + \alpha^4 + \alpha^2 + 1$$
$$\alpha^{12} = \alpha^{10} + \alpha^7 + \alpha^4 + \alpha^3 + \alpha^2 + \alpha + 1$$
$$\alpha^{13} = \alpha^{10} + \alpha^8 + \alpha^6 + \alpha^3 + \alpha + 1$$
$$\alpha^{14} = \alpha^{10} + \alpha^9 + \alpha^7 + \alpha^6 + \alpha^5 + \alpha + 1$$
$$\alpha^{15} = \alpha^8 + \alpha^7 + \alpha^5 + \alpha^4 + \alpha + 1$$

$$\alpha^{16} = \alpha^9 + \alpha^8 + \alpha^6 + \alpha^5 + \alpha^2 + \alpha$$
$$\alpha^{17} = \alpha^{10} + \alpha^9 + \alpha^7 + \alpha^6 + \alpha^3 + \alpha^2$$
$$\alpha^{18} = \alpha^8 + \alpha^7 + \alpha^6 + \alpha^5 + \alpha^3 + \alpha^2 + 1$$
$$\alpha^{19} = \alpha^9 + \alpha^8 + \alpha^7 + \alpha^6 + \alpha^4 + \alpha^3 + \alpha$$
$$\alpha^{20} = \alpha^{10} + \alpha^9 + \alpha^8 + \alpha^7 + \alpha^5 + \alpha^4 + \alpha^2$$
$$\alpha^{21} = \alpha^9 + \alpha^8 + \alpha^4 + \alpha^3 + \alpha^2 + 1$$
$$\alpha^{22} = \alpha^{10} + \alpha^9 + \alpha^5 + \alpha^4 + \alpha^3 + \alpha$$

```
> H1:= matrix(11, 23,
> [[1,0,0,0,0,0,0,0,0,0,0,1,1,1,1,1,0,0,1,0,0,1,0],
> [0,1,0,0,0,0,0,0,0,0,0,0,1,1,1,1,1,0,0,1,0,0,1],
> [0,0,1,0,0,0,0,0,0,0,0,1,1,0,0,0,1,1,1,0,1,1,0],
> [0,0,0,1,0,0,0,0,0,0,0,0,1,1,0,0,0,1,1,1,0,1,1],
> [0,0,0,0,1,0,0,0,0,0,0,1,1,0,0,1,0,0,0,1,1,1,1],
> [0,0,0,0,0,1,0,0,0,0,0,1,0,0,1,1,1,0,1,0,1,0,1],
> [0,0,0,0,0,0,1,0,0,0,0,1,0,1,1,0,1,1,1,1,0,0,0],
> [0,0,0,0,0,0,0,1,0,0,0,0,1,0,1,1,0,1,1,1,1,0,0],
> [0,0,0,0,0,0,0,0,1,0,0,0,0,1,0,1,1,0,1,1,1,1,0],
> [0,0,0,0,0,0,0,0,0,1,0,0,0,0,1,0,1,1,0,1,1,1,1],
> [0,0,0,0,0,0,0,0,0,0,1,1,1,1,1,0,0,1,0,0,1,0,1]]) ;
```


$$H1 := \begin{bmatrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 1 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 1 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1
\end{bmatrix}$$

```
> K1:=multiply(H1,transpose(G)) ;
```

$$K1 := \begin{bmatrix} 2 & 2 & 2 & 2 & 2 & 2 & 2 & 4 & 4 & 4 & 4 & 4 \\ 0 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 4 & 4 & 4 & 4 \\ 2 & 2 & 2 & 0 & 0 & 2 & 4 & 4 & 2 & 2 & 4 & 4 \\ 0 & 2 & 2 & 2 & 0 & 0 & 2 & 4 & 4 & 2 & 2 & 4 \\ 2 & 2 & 2 & 0 & 2 & 2 & 2 & 2 & 2 & 4 & 4 & 4 \\ 2 & 2 & 0 & 2 & 2 & 4 & 2 & 2 & 2 & 4 & 4 & 4 \\ 2 & 2 & 2 & 2 & 2 & 2 & 4 & 4 & 4 & 4 & 2 & 4 \\ 0 & 2 & 2 & 2 & 2 & 2 & 2 & 4 & 4 & 4 & 4 & 2 \\ 0 & 0 & 2 & 2 & 2 & 2 & 2 & 2 & 4 & 4 & 4 & 4 \\ 0 & 0 & 0 & 2 & 2 & 2 & 2 & 2 & 2 & 4 & 4 & 4 \\ 2 & 2 & 2 & 2 & 2 & 2 & 4 & 4 & 4 & 4 & 4 & 4 \end{bmatrix}$$

Les termes de la matrice obtenue ne sont pas « réduits » à leur forme canonique dans $\mathrm{GF}(2^{11}) = F_2[\,\alpha]$. Pour obtenir la réduction.

```
>  map(item -> Expand(item) mod 2, K1) ;
```


$$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\%1 := [0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0]$$

4.2.2. Code G_{24}

de type $[24, 12, 8]_2$ est un 3-correctur ; il est obtenu en ajoutant un contrôle total de parité à la matrice H de code G_{23}.

Ce code est bien adapté à la transmission de 4096 nuances de couleur.

4.2.3. Code G_{11} 

On considère le groupe cyclique $\mathbb{F}_{3^5}^*$ d'ordre $3^5-1=11\cdot 22$. Soit $\alpha\in\mathbb{F}_{3^5}^*$ une racine primitive de degré 11. On pose

$$G_{11}=\left\{x=(x_0,\cdots,x_{10})\in\mathbb{F}_3^{11}\mid \sum_{i=0}^{10}x_i\alpha^i=0\right\}\subset\mathbb{F}_3[x]/(x^{11}-1).$$

On a n=11, $q=3$,

$$X^{11}-1=X^{11}+2=(x+2)g_0(x)g_1(x)=(x+2)\,(x^5+2\,x^3+x^2+2\,x+2)\,(x^5+x^4+2\,x^3+x^2+2)$$

où

$$g_0(x)=x^5+2\,x^3+x^2+2\,x+2=\prod_{i\in I}(x-\alpha^i),I=\{1,3,4,5,9\}$$
$$g_1(x)=x^5+x^4+2\,x^3+x^2+2=\prod_{j\in J}(x-\alpha^j),J=\{2,6,7,8,10\}$$

Remarque. L'ensemble

$$I=\{1,3,4,5,9\}$$

coïncide avec l'ensemble des résidues quadratiques modulo 11, et l'ensemble complémentaire

$$J=\{2,6,7,8,10\}$$

coïncide avec l'ensemble des non-résidues quadratiques modulo 11.

L'application de Frobenius $\alpha^k\mapsto\alpha^{3k}$ laisse I et J stable puisque $\left(\frac{3}{11}\right)=1$, et l'application $\alpha^k\mapsto\alpha^{-k}$ échange les ensembles I et J puisque $\left(\frac{-1}{11}\right)=-1$, grâce à la loi de réciprocité quadratique de Gauss : pour les nombres premiers positifs impairs p, q on a

$$\left(\frac{p}{q}\right)\left(\frac{q}{p}\right)=(-1)^{\frac{(p-1)}{2}\frac{(q-1)}{2}},$$

et on a les deux compléments suivants de cette loi :

$$\left(\frac{2}{p}\right)=(-1)^{(p^2-1)/8},\qquad\left(\frac{-1}{p}\right)=(-1)^{(p-1)/2},$$

DÉFINITION 4.4 *Code de Golay G_{11} est un sous-espace vectoriel (g_0) de dimension 6 dans le quotient*

$$\mathbb{F}_3[x]/(x^{11}-1)$$

vu comme un espace vectoriel de dimension 11 sur $\mathbb{F}_3$ avec le polynôme générateur

$$g_0(x) = g_0(x) = x^5 + 2\,x^3 + x^2 + 2\,x + 2$$

et avec le polynome de contrôle

$$h(x) = (x+2)\,(x^5 + x^4 + 2\,x^3 + x^2 + 2) = (x+1)g_1(x)$$

C'est un $[11, 6, 5]_3$-code.

```
> q:=3;
```

$$q := 3$$

```
> restart ;
```

```
> with(linalg) :
```

```
Warning, the protected names norm and trace have been redefined and
unprotected
```



```
> Factor(x^11-1) mod 3;
```

$$(x+2)\,(x^5 + 2\,x^3 + x^2 + 2\,x + 2)\,(x^5 + x^4 + 2\,x^3 + x^2 + 2)$$

```
> g:=x^5+2*x^3+x^2+2*x+2;irreduc(g) mod 3;
```

$$g := x^5 + 2\,x^3 + x^2 + 2\,x + 2$$

$$true$$

```
> alias(alpha = RootOf(g)) ;
```

$$\alpha$$

```
> Factor(g,alpha) mod 3;
```

$$(x+2\,\alpha)\,(x+2\,\alpha^3+\alpha^2+2\,\alpha+2)\,(x+2\,\alpha^4)\,(x+\alpha^4+2\,\alpha^3+2\,\alpha^2+2\,\alpha+1)\,(x+2\,\alpha^3)$$

```
> for i from 0 to 11 do
> if Eval(g, x=alpha^i) mod 3 = 0 then Expand (alpha^i) mod 3;
> print('i'=i, alpha^i=Expand (alpha^i) mod 3, 'g'(alpha^i)=0) fi
> od ;
```

$$i = 1,\ \alpha = \alpha,\ \mathrm{g}(\alpha) = 0$$

$$i = 3,\ \alpha^3 = \alpha^3,\ \mathrm{g}(\alpha^3) = 0$$
$$i = 4,\ \alpha^4 = \alpha^4,\ \mathrm{g}(\alpha^4) = 0$$
$$i = 5,\ \alpha^5 = \alpha^3 + 2\,\alpha^2 + \alpha + 1,\ \mathrm{g}(\alpha^5) = 0$$
$$i = 9,\ \alpha^9 = 2\,\alpha^4 + \alpha^3 + \alpha^2 + \alpha + 2,\ \mathrm{g}(\alpha^9) = 0$$

```
> for i from 0 to 11 do
> if Eval(g, x=alpha^(-i)) mod 3 = 0 then Expand (alpha^(-i)) mod 3;
> print('i'=i, alpha^(i)=Expand (alpha^i) mod 3, 'g'(alpha^(-i))=0) fi
> od ;
```

$$i = 2,\ \alpha^2 = \alpha^2,\ \mathrm{g}(\frac{1}{\alpha^2}) = 0$$
$$i = 6,\ \alpha^6 = \alpha^4 + 2\,\alpha^3 + \alpha^2 + \alpha,\ \mathrm{g}(\frac{1}{\alpha^6}) = 0$$
$$i = 7,\ \alpha^7 = 2\,\alpha^4 + 2\,\alpha^3 + \alpha + 1,\ \mathrm{g}(\frac{1}{\alpha^7}) = 0$$
$$i = 8,\ \alpha^8 = 2\,\alpha^4 + 2\,\alpha^3 + 2\,\alpha^2 + 2,\ \mathrm{g}(\frac{1}{\alpha^8}) = 0$$
$$i = 10,\ \alpha^{10} = \alpha^4 + 2\,\alpha^2 + \alpha + 2,\ \mathrm{g}(\frac{1}{\alpha^{10}}) = 0$$


```
> (x+2)*(x^5+2*x^3+x^2+2*x+2)*(x^5+x^4+2*x^3+x^2+2);
> 'h'=(x+2)*(x^5+x^4+2*x^3+x^2+2);
> h:= sort(Expand((x+2)*(x^5+x^4+2*x^3+x^2+2)) mod 3,x);
```

$$(x + 2)\,(x^5 + 2\,x^3 + x^2 + 2\,x + 2)\,(x^5 + x^4 + 2\,x^3 + x^2 + 2)$$
$$h = (x + 2)\,(x^5 + x^4 + 2\,x^3 + x^2 + 2)$$
$$h := x^6 + x^4 + 2\,x^3 + 2\,x^2 + 2\,x + 1$$

Code de Golay C_{11} = (g) est un sous-espace vectoriel de dimension 6 dans le quotient $Z_3[\,X]/(\,x^{11} - 1)$ vu comme un espace vectoriel de dimension 11 sur Z_3 avec le polynôme générateur

$g := x^5 + 2\,x^3 + x^2 + 2\,x + 2$
et avec le polynome de contrôle $h = (x + 2)\,(x^5 + x^4 + 2\,x^3 + x^2 + 2)$,
$h := x^6 + x^4 + 2\,x^3 + 2\,x^2 + 2\,x + 1$.

C'est un $[11, 6, 5]_3$ -code.

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% %%%%%%

```
> G:= matrix(6, 11,
> [[2,2,1,2,0,1,0,0,0,0,0],
> [0,2,2,1,2,0,1,0,0,0,0],
> [0,0,2,2,1,2,0,1,0,0,0],
> [0,0,0,2,2,1,2,0,1,0,0],
> [0,0,0,0,2,2,1,2,0,1,0],
> [0,0,0,0,0,2,2,1,2,0,1]]);
```

$$G := \begin{bmatrix} 2 & 2 & 1 & 2 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 2 & 1 & 2 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 2 & 1 & 2 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 2 & 1 & 2 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 2 & 1 & 2 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 & 2 & 1 & 2 & 0 & 1 \end{bmatrix}$$

```
> 'h'= x^6+x^4+2*x^3+2*x^2+2*x+1;
```

$$h = x^6 + x^4 + 2\,x^3 + 2\,x^2 + 2\,x + 1$$


```
> H:= matrix(5, 11,
> [[0,0,0,0,1,0,1,2,2,2,1],
> [0,0,0,1,0,1,2,2,2,1,0],
> [0,0,1,0,1,2,2,2,1,0,0],
> [0,1,0,1,2,2,2,1,0,0,0],
> [1,0,1,2,2,2,1,0,0,0,0]]);
```

$$H := \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 0 & 1 & 2 & 2 & 2 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 & 2 & 2 & 2 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 2 & 2 & 2 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 2 & 2 & 2 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 2 & 2 & 2 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}$$

```
> sort(Expand(h*g) mod 3,x);
```

$$x^{11} + 2$$

```
> K:=multiply(H,transpose(G)) ;
```

$$K := \begin{bmatrix} 0 & 3 & 3 & 6 & 9 & 9 \\ 3 & 3 & 6 & 9 & 9 & 12 \\ 3 & 6 & 9 & 9 & 12 & 12 \\ 6 & 9 & 9 & 12 & 12 & 9 \\ 9 & 9 & 12 & 12 & 9 & 6 \end{bmatrix}$$

Les termes de la matrice obtenue ne sont pas « réduits » à leur forme canonique dans $\mathrm{GF}(3^5) = F_3[\,\alpha]$. Pour obtenir la réduction.

```
> map(item -> Expand(item) mod 3, K) ;
```

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

4.2.4. Code G_{12}

de type $[12, 6, 6]_3$ est un 2-correctur ; il est obtenu en ajoutant une ligne de contrôle total à la matrice H de code G_{11}.

EXERCICE. (a) Calculer le volume de la boule de Hamming $V_q(n, t)$, $t = 1, 2, 3$.
(b) Montrer que la borne de Hamming est atteinte pour les codes G_{23} et G_{11} donc on a un emplilement parfait de sphères.

REMARQUE 4.5 *On peut montrer que les codes G_{23} , G_{11} et $C(m, q)$ (voir Section 3.16) sont tous les codes parfaits.*

4.3 Locateurs d'erreurs

Pour détecter et corriger les erreurs, nous avons vu qu'il fallait déterminer le syndrome du message reçu. Dans le cas de certains codes cycliques, ce vecteur de longueur $n - k$ peut être remplacé par un objet plus léger ayant les mêmes possibilités. En effet, soit α une racine primitive $n^{\text{ème}}$ de l'unité, contenue dans $\mathbb{F}_{2^m}$, *) et considérons le code généré par $g(x)$, le polynôme minimal de α sur $\mathbb{F}_2$. Soit par exemple

Ici $q = 2$

$$H = (1\, \alpha\, \alpha^2 \cdots \alpha^{n-1})$$

et

$$S(v) = Hv^t = v(\alpha).$$

Soit w le message émis, posons $e^{(j)}(x) = x^{j-1}$, $1 \leq j \leq n$, et plaçons-nous dans le cas d'une erreur simple. Il existe donc $j, 1 \leq j \leq n$, tel que $v = w + e^{(j)}$ donc

$$S(v) = v(\alpha) = w(\alpha) + e^{(j)}(\alpha) = \alpha^{j-1}$$

$e^{(j)}(\alpha)$ est appelé locateur d'erreur. En effet, comme on a $e^{(j)}(\alpha) \neq e^{(i)}(\alpha)$ pour $i \neq j$, $1 \leq i, j \leq n$, α^{j-1} détermine la position de l'erreur. *)

Faire en e
le cas q

11.3 Soit p premier. Calculer "directement" le nombre de polynômes irréductibles de degré 5 sur $\mathbb{F}_p$.

11.4 Soit $P \in \mathbb{F}_{q^m}[X]$. Montrer $P \in \mathbb{F}_q[X]$ si et seulement si $P(X)^q = P(X^q)$.

11.5 Ecrire tous les polynômes irréductibles unitaires de degré 4 sur $\mathbb{F}_2$ et trouver leur ordre.

11.6 Montrer que $X^4 + 2$ est irréductible sur $\mathbb{F}_5$ et trouver son ordre.

11.7 Un polynôme P de degré m sur $\mathbb{F}_q$ est dit primitif sur $\mathbb{F}_q$ s'il est le polynôme minimal sur $\mathbb{F}_q$ d'une racine primitive de $\mathbb{F}_{q^m}$ (un générateur du groupe cyclique $\mathbb{F}_{q^m}^*$). Trouver le nombre des polynômes primitifs de degré m sur $\mathbb{F}_q$.

11.8 Soit p premier impair. Montrer que $\mathbb{F}_{p^2}$ est un corps de rupture de $X^4 + 1$ sur $\mathbb{F}_p$. Si $\alpha^4 + 1 = 0$ dans $\mathbb{F}_{p^2}$, montrer que $y = \alpha + \alpha^{-1}$ vérifie $y^2 = 2$. En déduire que 2 est un carré dans $\mathbb{F}_p$ si et seulement si $p \equiv \pm 1 \bmod 8$.

Cours N 5. Lundi 23 février 2015

Exemples de groupes. Groupes classiques

12 Structure de groupe

12.1 Compléments sur les groupes

La notion de groupe est fondamentale pour l'étude d'autres structures algébriques. Les groupes possèdent, d'autre part, des liens avec d'autres domaines des mathématiques. Un point de vue naïf s'était assez répandu, consistant à dire qu'il est possible de traîter les groupes finis à l'aide de moyens assez directs comme l'est le théorème de Cayley [W] affirmant que tout groupe se réalise comme sous-groupe d'un groupe de permutations d'un ensemble :

> Tout groupe fini G est isomorphe à un sous-groupe du groupe symétrique $\mathfrak{S}(G)$ des permutations de G. En particulier, si G est un groupe d'ordre n, il est isomorphe à un sous groupe fini de $\mathfrak{S}_n$.

Ce théorème fut démontré en 1854 dans un article de Arthur Cayley (16 août 1821 - 26 janvier 1895).

Cependant, ce théorème n'est pas toujours aussi puissant qu'il n'y paraît. Par exemple, dans le cas d'un cardinal fini, il plonge un groupe dans un autre groupe de cardinal supérieur et égal à $n!$.

12.2 Rappels sur l'action d'un groupe sur un ensemble

DÉFINITION 12.2.1 *Soit G un groupe, Ω un ensemble.*
a) Une action de G sur Ω est un morphisme de groupes
$\Phi : G \to S(\Omega) =$ *le groupe des permutations de* Ω, $g \mapsto (\Phi_g : \Omega \to \Omega)$
Notation : $gx := \Phi_g(x)$ *pour* $g \in G$, $x \in \Omega$, *et on a* $g_1(g_2x) = (g_1g_2)x$.
b) L'orbite d'un élément $x \in \Omega$ *est* $Gx := \{gx \mid g \in G\}$.
c) Le stabilisateur d'un élément $x \in \Omega$ *est* $\mathrm{St}_x := \{g \in G \mid gx = x\} \subset G$.

REMARQUE 12.2.2
a) $x_2 \in Gx_1 \iff Gx_1 = Gx_2$.
b) $g_1x = g_2x \iff g_2^{-1}g_1 \in \mathrm{St}_x$
c) $\Omega = \bigcup_i Gx_i$ *(l'union disjointe,* x_i *parcourt les représentants d'orbites).*

12.2.1 Formule des classes, formule de Burnside, voir [W]

À travers les notions d'orbite et de stabilisateur, les actions d'un groupe sont un outil commode en combinatoire. D'autre part, un certain nombre de propriétés concernant la structure de certains groupes peuvent être démontrées par des arguments de dénombrement.

Deux identités reviennent fréquemment. La formule des classes

$$\mathrm{Card}\,\Omega = \sum_x \mathrm{Card}\,Gx = \sum_x \frac{\mathrm{Card}\,G}{\mathrm{Card}\,St_x}$$

où x prend exactement une valeur dans chaque orbite, relie le cardinal de l'ensemble à la structure du groupe. La formule de Burnside affirme pour sa part que le nombre d'orbites est

$$\sum_{x\in\Omega} \frac{1}{\mathrm{Card}\,Gx} = \frac{1}{\mathrm{Card}\,G} \sum_{g\in G} \mathrm{Card}\,Fix_g$$

12.3 Groupes résolubles

Rappelons :

DÉFINITION 12.3.1 (a) *Un sous-groupe $H \subset G$ d'un groupe G est dit* distingué, *si pour tout $g \in G$ on a $gH = Hg$, notation $H \triangleleft G$. Dans ce cas on définit le groupe quotient*

$$G/H = \{gH \mid g_1H \cdot g_2H = g_1g_2H\}$$

(b) *G est dit* résoluble *s'il existe une série*

$$G = G_0 \triangleright G_1 \triangleright \cdots \triangleright G_n = \{e\}$$

telle que G_{k-1}/G_k sont tous abéliens ($k = 1, 2, \cdots, n$)

EXEMPLES 12.3.2 *Les groupes*

$$G = S_2, S_3, A_4, S_4$$

et

$$G = \left\{ \begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix} \mid a \in \mathbb{F}_5^*, b \in \mathbb{F}_5 \right\}$$

sont résolubles, mais les groupes

$$G = S_n \text{ et } A_n \subset S_n (n \geq 5)$$

ne sont pas résolubles.

12.3.1 Résolubilité et groupes de Galois

On associe à un polynôme $f(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_0 = (x - x_1) \cdot \ldots \cdot (x - x_n) \in \mathbb{Q}[x]$ le groupe G_f, dit "le groupe de Galois", un sous-groupe du groupe des permutations $S(x_1, \cdots x_n)$ des racines $x_1, \cdots, x_n$:

$$G_f \subset S(x_1, \cdots, x_n),$$

induites par certains automorphismes d'un corps commutatif contenant $x_1, \cdots, x_n$.

Pour trouver $x_1, \cdots x_n$, on cherche les expressions fixées par toutes les permutations dans G_f. En général, le sous-groupe $G_f \subset S(x_1, \cdots, x_n)$ peut être plus petit que S_n (pour une équation particulière).

On peut montrer que la résolubilité de l'équation $f(x) = 0$ en radicaux est équivanente la résolubilité du groupe de Galois G_f du polynôme f.

12.4 Groupes simples

DÉFINITION 12.4.1 *Un groupe G est dit simple si pour tout sous-groupe distingué $H \subset G$ on aura soit $H = \{1\}$ soit $H = G$.*

Le terme « simple » signifie que de tels groupes ne sont pas, en quelque sorte, « réductibles » à un groupe plus maniable. L'intérêt d'un sous-groupe distingué non trivial H d'un groupe G est souvent de permettre la construction du groupe quotient G/H. L'étude de G se ramène alors à celle de H et de G/H. Cette construction n'est pas possible pour un groupe simple et on ne peut donc pas ramener son étude à celle d'un groupe quotient de cardinal plus petit que lui.

Les groupes simples finis sont importants car il peuvent être perçus comme des briques de base de tous les groupes finis, de la même façon que tous les nombres entiers peuvent être décomposés en produits de nombres premiers.

Les seuls groupes simples abéliens sont les groupes cycliques d'ordre premier. La classification des groupes simples finis fut achevée en 1982.

12.4.1 Simplicité du groupe A_5

Pour démontrer que le groupe A_5 est simple, c'est à dire que pour tout sous-groupe distingué $H \subset A_5$ on aura soit $H = \{1\}$ soit $H = A_5$, il faut utiliser les classes d'éléments conjugués de ce groupe. En effet un tel sous-groupe doit contenir avec tout élément x sa classe d'éléments conjugués. Les éléments de A_5 sont de type suivant :

1) Les cycles d'ordre 3 $\tau = (\alpha, \beta, \gamma)$ qui sont tous conjugés (il en a 20, et $St_\tau = \{1, \tau, \tau^2\}$) ;

2) Les produits $\tau = (\alpha, \beta)(\gamma, \delta)$ des deux transpositions disjointes (il y en a 15, et

$$St_\tau \cong V_4 = \{1, \tau, (\alpha, \gamma)(\beta, \delta), (\alpha, \delta)(\beta, \gamma)\}$$

le groupe isomorphe à groupe de Klein V_4) ;

3) Les cycles d'ordre 5 (il en a 24, il ne peuvent pas tous être conjugués, car l'ordre d'un orbite doit diviser l'ordre du groupe, mais 24 ne divise pas 60) ; cependent si $\tau = (\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5)$ on voit facilement que $St_\tau = \langle\tau\rangle$ a l'ordre 5 donc l'orbite (la classe conjugué correspondant a l'ordre 12), et on obtient la classe de tel type suivant : celle de $(1, 2, 3, 4, 5)$ et celle de $(1, 2, 3, 5, 4)$.

4) L'élément 1 forme une classe conjuguée tout seule.

On voit donc qu'il y a que 5 classes (des ordres correspondants 1, 12, 12, 15, 20). Ces classes peuvent être facilement décrites géometriquement si l'on utilise l'identification ci-dessus du groupe A_5 avec les groupes d'octaèdre (et de dodecaèdre).

THÉORÈME 12.4.2 *Le groupe A_5 est simple, i.e. pour tout sous-groupe distingué $H \subset A_5$ soit $H = A_5$ soit $H = \{1\}$.*

Démonstration. On ne peut pas obtenir un sous-groupe H comme une réunion disjointe des classes si-dessus (l'ordre de H doit diviser 60 et il doit être de la forme 1 + la somme d'une partie des nombres 12, 12, 15, 20 n'est possible que dans les cas $|H| = 1$ où $|H| = 60$).

12.5 Groupe orthogonal $G = SO(3)$ et les angles de Euler

On considère le groupe orthogonal réel $G = SO(3) = SO(3, \mathbb{R})$.

Tout d'abord, on montre que G est compact (comme une partie bornée et fermée de $Mat(3,\mathbb{R}) \cong \mathbb{R}^9$).

Soient ϕ, ψ, θ trois angles tels que

$$0 \leq \phi, \psi < 2\pi, \quad 0 \leq \theta \leq \pi,$$

et on pose

$$B_\phi := \begin{pmatrix} \cos(\phi) & -\sin(\phi) & 0 \\ \sin(\phi) & \cos(\phi) & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad B_\psi := \begin{pmatrix} \cos(\psi) & -\sin(\psi) & 0 \\ \sin(\psi) & \cos(\psi) & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

$$C_\theta := \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta) & -\sin(\theta) \\ 0 & \sin(\theta) & \cos(\theta) \end{pmatrix} \in G.$$

On montre par la multiplication directe que

$$B_\phi C_\theta B_\psi = A, \text{ où}$$

$$A = \begin{pmatrix} \cos(\phi)\cos(\psi) - \sin(\phi)\cos(\theta)\sin(\psi) & -\cos(\phi)\sin(\psi) - \sin(\phi)\cos(\theta)\cos(\psi) & \sin(\phi)\sin(\theta) \\ \sin(\phi)\cos(\psi) + \cos(\phi)\cos(\theta)\sin(\psi) & -\sin(\phi)\sin(\psi) + \cos(\phi)\cos(\theta)\cos(\psi) & -\cos(\phi)\sin(\theta) \\ \sin(\theta)\sin(\psi) & \sin(\theta)\cos(\psi) & \cos(\theta) \end{pmatrix}.$$

En effet,

```
> restart;with(LinearAlgebra):B[phi]:=Matrix([[cos(phi), -sin(phi), 0],
> [sin(phi), cos(phi), 0],
> [0,0,1]]);
```

$$B_\phi := \begin{bmatrix} \cos(\phi) & -\sin(\phi) & 0 \\ \sin(\phi) & \cos(\phi) & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

```
> B[psi]:=Matrix([[cos(psi), -sin(psi), 0],
> [sin(psi), cos(psi), 0],
> [0,0,1]]);
```

$$B_\psi := \begin{bmatrix} \cos(\psi) & -\sin(\psi) & 0 \\ \sin(\psi) & \cos(\psi) & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

```
> C[theta]:=Matrix([[1,0,0],[0, cos(theta), -sin(theta)],
> [0, sin(theta), cos(theta)]]);
```

$$C_\theta := \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta) & -\sin(\theta) \\ 0 & \sin(\theta) & \cos(\theta) \end{bmatrix}$$

```
> A:=Multiply(B[phi], Multiply(C[theta], B[psi]));
```

$A :=$
$[\cos(\phi)\cos(\psi) - \sin(\phi)\cos(\theta)\sin(\psi)\,,\ -\cos(\phi)\sin(\psi) - \sin(\phi)\cos(\theta)\cos(\psi)\,,$
$\sin(\phi)\sin(\theta)]$
$[\sin(\phi)\cos(\psi) + \cos(\phi)\cos(\theta)\sin(\psi)\,,\ -\sin(\phi)\sin(\psi) + \cos(\phi)\cos(\theta)\cos(\psi)\,,$
$-\cos(\phi)\sin(\theta)]$
$[\sin(\theta)\sin(\psi)\,,\ \sin(\theta)\cos(\psi)\,,\ \cos(\theta)]$

On va montrer que pour tout $A \in SO(3)$, il existe une décomposition de Euler

$$A = B_\phi C_\theta B_\psi$$

où ϕ, ψ, θ sont trois angles tels que

$$0 \leq \phi, \psi < 2\pi, \quad 0 \leq \theta \leq \pi.$$

On peut présenter graphiquement la signification géométrique des angles de Euler, voir
`http ://www.sciences.univ-nantes.fr/physique/perso/gtulloue/Meca/Cinematique/euler1.html`

12.6 Homomorphisme remarquable de $SU(2)$ dans $SO(3)$

Soit

$$g = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \in SU(2) \text{ alors } g^{-1} = \begin{pmatrix} \delta & -\beta \\ -\gamma & \alpha \end{pmatrix} \in SU(2)$$

On montre que

$$g = \begin{pmatrix} \alpha & \beta \\ -\overline{\beta} & \overline{\alpha} \end{pmatrix}, |\alpha|^2 + |\beta|^2 = 1.$$

On en déduit que $SU(2)$ est homéomorphe à $S^3 \subset \mathbb{R}^4$.

On pose

$$b_\varphi = \begin{pmatrix} e^{i\frac{\varphi}{2}} & 0 \\ 0 & e^{-i\frac{\varphi}{2}} \end{pmatrix} \in SU(2), \quad c_\theta = \begin{pmatrix} \cos\frac{\theta}{2} & i\sin\frac{\theta}{2} \\ i\sin\frac{\theta}{2} & \cos\frac{\theta}{2} \end{pmatrix} \in SU(2)$$

On vérifie que

$$a(\varphi,\theta,\psi) = b_\varphi c_\theta b_\psi = \begin{pmatrix} \cos\frac{\theta}{2}e^{i\frac{\varphi+\psi}{2}} & i\sin\frac{\theta}{2}e^{i\frac{\varphi-\psi}{2}} \\ i\sin\frac{\theta}{2}e^{-i\frac{\varphi-\psi}{2}} & \cos\frac{\theta}{2}e^{-i\frac{\varphi+\psi}{2}} \end{pmatrix}$$

On en déduit que toute matrice

$$g = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \in SU(2)$$

se décompose sous la forme

$$g = a(\varphi,\theta,\psi) = b_\varphi c_\theta b_\psi$$

(INDICATION :Il suffit de poser

$$|\alpha| = \cos\frac{\theta}{2}, \mathrm{Arg}\alpha = \frac{\varphi+\psi}{2}, \quad |\beta| = \sin\frac{\theta}{2}, \mathrm{Arg}\beta = \frac{\varphi-\psi+\pi}{2}).$$

Ceci implique l'existence d'un homomorphisme

$$\Phi : SU(2) \to SO(3)$$

tel que

$$\Phi(b_\varphi) = B_\varphi, \quad \Phi(c_\theta) = C_\theta,$$

(utiliser l'action du groupe $SU(2)$ par conjugaison sur les matrices hermitiennes de trace nulle

$$H_x = \begin{pmatrix} x_3 & x_1 + ix_2 \\ x_1 - ix_2 & -x_3 \end{pmatrix}$$

en base de matrices de Pauli

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \ \sigma_2 = \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}, \ \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix})$$

On trouve le noyau de Φ, et on en déduit que $SO(3)$ est homéomorphe à l'espace projectif réel $\mathbb{RP}^3$.

12.7 Groupes finis des rotations

On va donner la liste des groupes finis de rotation, c'est à dire, des sous-groupes finis du groupe

$$\mathrm{SO}(3) = \{A \in M_3(\mathbb{R}) | A \cdot {}^tA = 1_3\},$$

oú ${}^tA$ est la matrice transposée à A.

Nous savons que tout élément $A \in \mathrm{SO}(3)$, $A \neq 1_3$ est une rotation de $\mathbb{R}^3$ d'une axe (x). C'est à dire, il y a exactement deux points sur la sphére $S^2 \subset \mathbb{R}^3$ fixés par rapport á A (qui sont apellés les pôles de la rotation).

Soit G un sous-groupe fini de SO(3), S l'ensemble des pôles des éléments non-triviaux de G.

Pour $x \in S$ et $B \in G$ on a $(BAB^{-1})Bx = B \cdot Ax = Bx$, c'est à dire $Bx \in S$. Désignons par Ω l'ensemble des pairs (A, x) où $A \in G$, $A \neq 1_3$ x le pôle de A. Soit G_x le stabilisateur de x. Si

$$G = \cup G_x \cup g_2 G_x \cup g_3 G_x \cdots \cup g_{m_x} G_x$$

la décomposition en classes d'équivalence, l'orbite de x s'identifie à

$$x, g_2 x, \cdots, g_{m_x} x$$

de cardinalité $G(x) = m_x$, et on a $N = m_x n_x$ avec $N = |G|, n_x = |G_x|$. On dit que n_x est la multiplicité du pôle x.

On a $|\Omega| = 2(N-1)$ puisque à tout $A \in G$, $A \neq 1_3$ correspondent exactement deux pôles. De l'autre côté,

$$|\Omega| = \sum_{x \in S} (n_x - 1).$$

Soit $\{x_1, \cdots, x_k\}$ l'ensemble des réprésentants d'orbites. Posons $n_i = n_{x_i}$, $m_i = m_{x_i}$, alors

$$|\Omega| = \sum_{x \in S}(n_x - 1) = \sum_{i=1}^{k} m_i(n_i - 1) = \sum_{i=1}^{k}(N - m_i).$$

Donc,

$$2N - 2 = \sum_{i=1}^{k}(N - m_i)$$

d'où

$$2 - \frac{2}{N} = \sum_{i=1}^{k}\left(1 - \frac{1}{n_i}\right). \tag{1}$$

Supposons que $N > 1$ alors $1 \leq 2 - \frac{2}{N} < 2$. On a $n_i \geq 2 \Rightarrow \frac{1}{2} \leq 1 - \frac{1}{n_i} < 1$, donc k doit être égale à 2 ou 3.

Consiérons deux cas :

Cas 1. $k = 2$.

On a

$$2 - \frac{2}{N} = \left(1 - \frac{1}{n_1}\right) + \left(1 - \frac{1}{n_2}\right)$$

où

$$2 = \frac{N}{n_1} + \frac{N}{n_2} = m_1 + m_2.$$

Ceci implique $m_1 = m_2 = 1$, $n_1 = n_2 = N$, c'est à dire il y un seul axe de rotation et $G = C_N$ est le groupe cyclique d'ordre N.

Cas 2. $k = 3$. On peut supposer que $n_1 \leq n_2 \leq n_3$. Si $n_1 \geq 3$ on aura

$$\sum_{i=1}^{3}\left(1 - \frac{1}{n_i}\right) \geq \left(1 - \frac{1}{3}\right) = 2,$$

ce qui est impossible. Alors $n_1 = 2$ et l'équation (1) s'écrit sous la forme

$$\frac{1}{2} + \frac{2}{N} = \frac{1}{n_2} + \frac{1}{n_3}.$$

Il est clair que $n_2 \geq 3 \Rightarrow \frac{1}{n_2} + \frac{1}{n_3} \leq \frac{1}{2}$. Alors $n_2 = 2$ ou 3.

Si $n_2 = 2$ alors $n_3 = \frac{N}{2} = m$ (N doit être impair) et $m_1 = m_2 = m$, $m_3 = 2$. Ces données correspondent au groupe du dyèdre D_m.

Si $n_2 = 3$ on a

$$\frac{1}{6} + \frac{2}{N} = \frac{1}{n_3},$$

et il y a les trois possibilités suivantes :

$$2') \quad n_3 = 3,\ N = 12,\ m_1 = 6,\ m_2 = 4,\ m_3 = 4;$$
$$2'') \quad n_3 = 4,\ N = 24,\ m_1 = 12,\ m_2 = 8,\ m_3 = 6;$$
$$2''') \quad n_3 = 5,\ N = 60,\ m_1 = 30,\ m_2 = 20,\ m_3 = 12;$$

On a donc

THÉORÈME 12.7.1 *Soit G un sous-groupe fini de* SO(3) *qui n'est pas cyclique ou dyèdral. Alors* $|G| =$ 12, 24 *ou 60.*

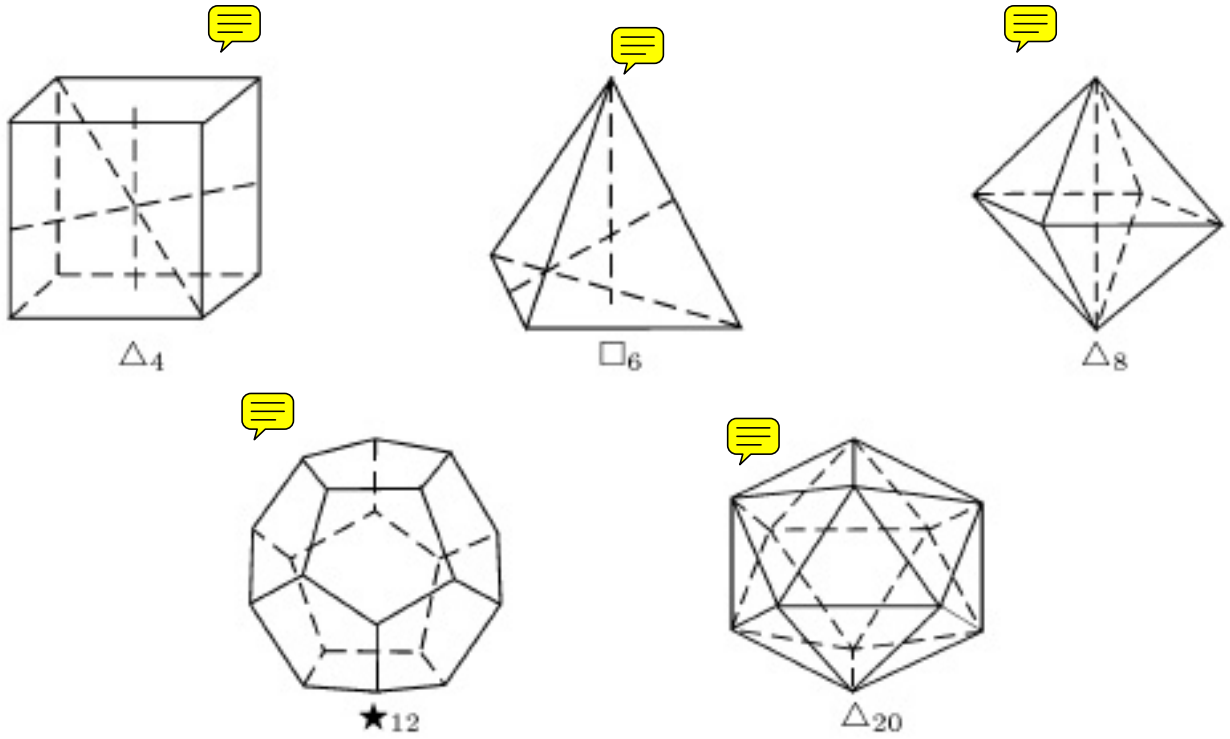


FIG. 1 – Polyèdres reguliers

12.8 Groupes de polyèdres réguliers.

L'existence de groupes d'ordre 12, 24, ou 60 dans SO(3) est facile de voir : il est connu depuis longtemps qu'il existe exactement cinq polyèdres réguliers convexes dans $\mathbb{R}^3$:

1) thétraèdre $\triangle_4$
2) cube $\square_6$
3) octaèdre $\triangle_8$
4) dodecaèdre $\star_{12}$
5) icosaèdre $\triangle_{20}$

Si l'on pose le centre d'un polyèdre régulier à l'origine de $\mathbb{R}^3$, alors les rotations de ce polyèdre vont former un sous-groupe fini de SO(3).

De telle manière on obtient tout de même que trois groupes finis non-isomorphes, car les groupes correspondants du cube et du octaèdre, et aussi ceux du octaèdre et du dodecaèdre sont les mêmes.

EXERCICES

12.1 (CLASSES DE CONJUGAISON DU GROUPE S_n)

(a) Deux cycles $\tau = (j_1, \cdots, j_k)$ et $\tau' = (j'_1, \cdots, j'_{k'})$ ont dit *indépendants* si

$$\{j_1, \cdots, j_k\} \cap \{j'_1, \cdots, j'_{k'}\} = \emptyset.$$

Montrer que toute permutation $\tau \in S_n$ se décompose en produit $\tau_1 \cdot \tau_2 \cdot \ldots \cdot \tau_r$ des cycles indépendants de longueur $k_1 \geq k_2 \geq \cdots \geq k_r > 1$. On dit que τ est de type cyclique $(k_1, k_2, \cdots, k_r, 1, \cdots, 1)$ avec une *partition* $k_1 + k_2 + \cdots + k_r + 1 + \cdots + 1 = n$.

(b) Soit $\sigma = \begin{pmatrix} 1, \cdots, n \\ i_1, \cdots, i_n \end{pmatrix}$, $\tau = (j_1, \cdots, j_k)$ un cycle d'ordre k dans le groupe S_n des permutations de l'ensemble $X_n = \{1, \cdots, n\}$. Montrer que la permutation $\sigma\tau\sigma^{-1}$ coincide avec le cycle $(i_{j_1}, \cdots, i_{j_k})$ d'ordre k.

(c) Soit $\tau = (12)(345)$, $\tau' = (16)(234)$ deux éléments de S_6 de type cyclique $(3,2,1)$. Trouver toutes les $\sigma \in S_6$ telles que

$$\sigma\tau\sigma^{-1} = \tau'.$$

(d) Montrer que deux éléments τ et τ' de S_n sont conjugés si et seulement si ils ont le même type cyclique.

(e) En déduire que le nombre de classes de conjugaison de S_n est égal au nombre $p(n)$ des partitions de n. Trouver $p(3), p(4), p(5), p(6)$.

(f) Trouver toutes les classes de conjugaison du groupe S_n. Pour $n = 3, 4, 5, 6$ déterminer le nombre d'éléments dans chaque classe.

(g) Démontrer l'identité de Euler : on pose $p(0) = 1$, alors

$$\sum_{n\geq 0} p(n)x^n = \prod_{m\geq 1}(1-x^m)^{-1}.$$

12.2 On considère le groupe de dièdre D_n d'ordre $2n$ presenté par deux générateurs a, b et par les relations suivantes

$$D_n = \langle a, b \mid a^n = b^2 = baba = e\rangle.$$

(a) Montrer que

$$D_n = \left\{b^l a^k \mid k = 0, 1, \cdots, n-1, l = 0, 1\right\}.$$

(b) Trouver toutes les classes de conjugaison de D_n (il faut traiter séparement le cas de $n = 2m$ pair et de $n = 2m+1$ impair).

12.3 Soit A_n le sous-groupe des permutations paires de S_n.

(a) Trouver toutes les classes de conjugaison des groupes A_4 et A_5.

(b) En déduire tous les sous-groupes distingues de A_4 et de A_5.

12.4 (a) Trouver toutes les classes de conjugaison du groupe $\mathrm{GL}_n(\mathbb{C})$.

(b) Trouver toutes les classes de conjugaison du groupe $\mathrm{SL}_n(\mathbb{C})$.

(c) Trouver toutes les classes de conjugaison du groupe $\mathrm{GL}_2(\mathbb{F}_3)$.

12.5 (a) Pour tout nombre complexe $\lambda \in \mathbb{C}$ on pose

$$J_n(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda & 1 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 0 & \lambda & 1 \\ 0 & 0 & \cdots & 0 & 0 & \lambda \end{pmatrix} = \lambda \cdot I_n + N_n,\ N_n = J_n(0).$$

Montrer une condition nécessaire et suffisante pour qu'une matrice complexe $A \in \mathrm{Mat}_n(\mathbb{C})$ soit conjugée à la matrice $J_n(\lambda)$ (c'est-à-dire, il existe une matrice $C \in \mathrm{GL}_n(\mathbb{C})$ telle que $C^{-1}AC = J_n(\lambda)$) est que A n'a qu'un seul vecteur colonne propre (à proportionnalité près), de valeur propre égale à λ.

(b) On admet que toute matrice complexe $A \in \mathrm{Mat}_n(\mathbb{C})$ est conjugée à une matrice

$$B = \mathrm{diag}\{B_1, \cdots, B_r\}$$

diagonale bloc par bloc avec tout bloc donnée par une matrice $J_{n_j}(\lambda_j)$ (c'est-à-dire, on admet qu'il existe une matrice $C \in \mathrm{GL}_n(\mathbb{C})$ telle que $C^{-1}AC = B$). Montrer que

$$C^{-1}A^mC = B^m = \mathrm{diag}\{B_1^m, \cdots, B_r^m\}.$$

(c) Montrer que

$$\mathrm{rk}(J_n(\lambda))^k = \begin{cases} n, & \text{si } \lambda \neq 0 \\ n-k, & \text{si } \lambda = 0 \text{ et } k < n \\ 0, & \text{si } \lambda = 0 \text{ et } k \geq n \end{cases}$$

(d) On suppose qu'il existe une matrice $C \in \mathrm{GL}_n(\mathbb{C})$ telle que $C^{-1}AC = B$ est une matrice diagonale bloc par bloc

$$B = \mathrm{diag}\{B_1, \cdots, B_r\}$$

avec tout bloc donné par une matrice $J_{n_j}(\lambda_j)$. On note par $N_m(\lambda_j, B)$ le nombre de blocs de B de type $J_m(\lambda_j)$, puis on définit

$$r_m(\lambda_j, A) = r_m(\lambda_j, B) = \mathrm{rk}(B - \lambda_j I_n)^m = \mathrm{rk}(A - \lambda_j I_n)^m, \text{ et on pose } r_0(\lambda_j, A) = n.$$

Montrer que pour tous les $m = 1, 2, \cdots$ on a

$$\begin{aligned} r_m(\lambda_j, A) - r_{m+1}(\lambda_j, A) &= N_{m+1}(\lambda_j, B) + N_{m+2}(\lambda_j, B) + \cdots, \\ r_{m-1}(\lambda_j, A) - r_m(\lambda_j, A) &= N_m(\lambda_j, B) + N_{m+1}(\lambda_j, B) + \cdots, \end{aligned}$$

donc

$$N_m(\lambda_j, B) = r_{m-1}(\lambda_j, A) - 2r_m(\lambda_j, A) + r_{m+1}(\lambda_j, A).$$

(c) Soient A et A' deux matrices telles qu'il existe C, $C' \in \mathrm{GL}_n(\mathbb{C})$,

$$C^{-1}AC = B, C'^{-1}A'C' = B'$$

sont deux matrices diagonales bloc-par-bloc

$$B = \mathrm{diag}\{B_1, \cdots, B_r\}, \quad B' = \mathrm{diag}\{B'_1, \cdots, B'_{r'}\}.$$

Montrer que A et A' sont conjuguées si et seulement si pour tous les m et j,

$$N_m(\lambda_j, B) = N_m(\lambda_j, B') \Longleftrightarrow r_m(\lambda_j, A) = r_m(\lambda_j, A').$$

En utilisant la *forme normale* B d'une matrice complexe A trouver toutes les classes de conjugaison du groupe $\mathrm{GL}_n(\mathbb{C})$.

(d) En utilisant la forme canonique d'une matrice orthogonale A trouver toutes les classes de conjugaison du groupe $\mathrm{SO}_n(\mathbb{R})$.

(e)En utilisant la forme normale d'une matrice complexe A trouver toutes les classes de conjugaison du groupe $\mathrm{SL}_n(\mathbb{C})$.

(f) Trouver toutes les classes de conjugaison du groupe $\mathrm{GL}_2(\mathbb{F}_3)$.

12.6 Classes de conjugaison du groupe des *quaternions de Cayley*

$$Q_8 = \{\pm 1, \pm i, \pm j, \pm k \mid ij = k = -ji, jk = i = -kj, ki = j = -ik, i^2 = j^2 = k^2 = -1\}.$$

(a) Trouver toutes les classes de conjugaison du groupe Q_8.

(b) Trouver le centre du groupe Q_8.

(c) Montrer que les groupes Q_8 et D_4 d'ordre 8 ne sont pas isomorphes.

(d) Montrer que le groupe Q_8 est isomorphe au groupe formé par les matrices

$$\pm \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \pm \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, \pm \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \pm \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix},$$

(comparer avec les *matrices de Pauli*) :

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \ \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \ \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}).$$

12.7 Soit $K \subset E$ une extension de corps. Montrer que $\overline{K_E} = \{x \in E, x \text{ est algébrique sur } K\}$ est un sous-corps de E.

12.8 Montrer qu'il existe des polynômes irréductibles sur $\mathbb{Q}$ de tout degré $n \geq 1$, par exemple $X^n - p$, où p est un nombre premier.

12.9 Soit K un corps. En considérant l'ordre des éléments d'un groupe cyclique, montrer que $n = \sum_{d|n} \varphi(d)$. En déduire une autre preuve que tout sous-groupe d'ordre n de K^* est cyclique.

12.10 Soit K un corps à q éléments, $q \geq 4$. Montrer que $\sum_{x \in K} x^2 = 0$. Plus généralement, calculer, pour $s \geq 1$, la somme $\sum_{x \in K} x^s$.

12.11 Si H est un sous-groupe de $\mathbb{C}^*$ tel que $\mathbb{C}^*/H$ est fini, monter que $H = \mathbb{C}^*$.

12.12 Soit G un groupe abélien. Montrer que $\mathrm{Hom}(G, K^*)$ est un groupe abélien avec la multiplication $\chi_1\chi_2(g) = \chi_1(g)\chi_2(g)$

12.13 Si $|G| < \infty$, montrer que $|\mathrm{Hom}(G, K^*)| \leq |G|$.

12.14 Soient $K = \mathbb{C}$, G un groupe abélien. On pose $G^\vee = \mathrm{Hom}(G, \mathbb{C}^*) = X^*(G)$.

a) Montrer qu'il existe un isomorphisme (non-canonique) $G^\vee \xrightarrow{\sim} G$, donc $|G^\vee| = |G|$.

b) Démontrer les relations d'orthogonalité :

$$\sum_{g \in G} \chi(g) = \begin{cases} |G|, & \text{si } \chi \equiv 1 \\ 0, & \text{sinon} \end{cases}, \quad \sum_{\chi \in G^\vee} \chi(g) = \begin{cases} |G|, & \text{si } g = e \\ 0, & \text{sinon} \end{cases}$$

12.15 Soit

$$1 \to A \xrightarrow{f} B \xrightarrow{g} C \to 1$$

une suite exacte de trois groupes abéliens finis ; c'est-à-dire, que $\mathrm{Im}\, f = \mathrm{Ker}\, g$, $\mathrm{Ker}\, f = \{1\}$, $\mathrm{Im}\, g = C$. Montrer qu'il existe une suite exacte des groupes des caractères :

$$1 \to C^\vee \xrightarrow{g^\vee} B^\vee \xrightarrow{f^\vee} A^\vee \to 1,$$

où $g^\vee$, $f^\vee$ sont définis pour tout $\chi \in C^\vee$, $\psi \in B^\vee$ et pour tout $b \in B$, $a \in A$ par

$$g^\vee(\chi)(b) = \chi(g(b)), \quad f^\vee(\psi)(a) = \psi(f(a)).$$

Cours N 6. Lundi 2 mars 2015

12.8.1 Simplicité du groupe projectif spécial $PSL(2,F) = SL(2,F)/\{\pm I_2\}$

THÉORÈME 12.8.1 *Le groupe projectif spécial $PSL(2,F) = SL(2,F)/\{\pm I_2\}$ sur tout corps de cardinal $|F| > 3$ est simple.*

PREUVE. On considère les sous-groupes et les éléments suivants :

$$U = \left\{ u(\alpha) = \begin{pmatrix} 1 & \alpha \\ 0 & 1 \end{pmatrix} \;\middle|\; \alpha \in F \right\},$$
$$\overline{U} = \left\{ \overline{u}(\alpha) = \begin{pmatrix} 1 & 0 \\ \alpha & 1 \end{pmatrix} \;\middle|\; \alpha \in F \right\},$$
$$D = \left\{ d(\lambda) = \begin{pmatrix} \lambda & 0 \\ 0 & \lambda^{-1} \end{pmatrix} \;\middle|\; \lambda \in F^* \right\},$$
$$B = DU = UD = \left\{ \begin{pmatrix} \alpha & \beta \\ 0 & \alpha^{-1} \end{pmatrix} \right\}$$

(le sous-groupe de Borel). On remarque que

$$d(\lambda) = u(\lambda - 1)\overline{u}(1)u(\lambda^{-1} - 1)\overline{u}(-\lambda),$$

donc le sous-groupe de Borel est engendré par les sous-groupes unipotents U et $\overline{U}$. On considère aussi l'élément

$$w = u(1)\overline{u}(-1)u(1) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

2) Le groupe $G = SL(2,F)$ possède la décomposition (dite décomposition de Bruhat) suivante :

$$G = B \cup BwB, \quad B \cap BwB = \emptyset. \tag{12.1}$$

3) Le sous-groupe de Borel B est maximal dans G.

En effet, soit $H \supset B$ un sous-groupe, alors la décomposition (12.1) implique que tout élément $h \in H$, n'est pas contenu dans B, se trouve dans BwB, c'est-à-dire, $h = b_1 w b_2$, d'où $w \in H$, et donc $H = G$.

4) Si $|F| \geq 4$, alors $G = SL(2,F) = G'$. On prend $0 \neq \lambda \in F$, $\lambda^2 \neq 1$, ce qui est possible pour $|F| > 3$. Puis, on utilise la relation de commutation

$$d(\lambda)u(\alpha)d(\lambda)^{-1}u(\lambda)^{-1} = u(\alpha(\lambda^2 - 1)),$$

pour voir que $B' = U$ et $G' \supset U$. Puisque $G' \lhd G$, on a $G' \supset wUw^{-1} = \overline{U}$ On a vu par 1) et 2) que U et $\overline{U}$ engendrent G, donc $G' = G$.

5) Si $|F| \geq 4$, alors $PSL(2,F) = SL(2,F)/Z$ est simple (où $Z = \{\pm I_2\}$ est le centre).

On utilise l'égalité (facile à vérifier)

$$\bigcap_{x \in G} xBx^{-1} = Z.$$

Il faut montrer que si $H \lhd G = SL(2,F)$, alors soit $H \subset Z$, soit $H \supset G'$. La maximalité de B implique $HB = B$ ou $HB = G$. Si $HB = B$, alors $H \subset B$. Puisque $H \lhd G$, $H = xHx^{-1} \subset xBx^{-1}$ pour tout $x \in G$, on a $H \subset Z$.

D'autre part,

$$HB = G \Longrightarrow w = hb,\ h \in H, b \in B.$$

Dans ce cas

$$\overline{U} = wUw^{-1} = hbUb^{-1}h^{-1} = hUh^{-1} \subset HU,$$

puisque $H \lhd G$. L'inclusion $U \subset HU$ implique $HU = G$, puisque U et $\overline{U}$ engendrent G. Ceci implique que le groupe quotient

$$G/H = HU/H \cong U/(U \bigcup H)$$

est abélien, d'où $H \supset G'$. Maintenant la simplicité du groupe $PSL(2,F)$ est évidente. ∎

12.9 Groupes classiques (définition préliminaire)

Soit k un corps commutatif.

(a) *Le groupe général linéaire* $GL(n,k)$ est le groupe de toutes les matrices inversibles de taille $n \times n$ sur un corps commutatif k. Ce groupe $G = GL(n,k)$ agit sur l'ensemble $\Omega = k^n$ de la manière suivante.

Soit $e_1, \cdots, e_n$ la base standard pour k^n

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ \cdots \\ 0 \end{pmatrix}, e_2 = \begin{pmatrix} 0 \\ 1 \\ \cdots \\ 0 \end{pmatrix}, \cdots, e_n = \begin{pmatrix} 0 \\ 0 \\ \cdots \\ 1 \end{pmatrix}.$$

Soit $\varphi : k^n \to k^n$ une application linéaire inversible $\varphi = \varphi_A$ de k^n telle que

$$\varphi \mapsto A = A_\varphi,\ A_i = \varphi(e_i) \text{ pour } i = 1, \cdots, n,$$

où $A = A_\varphi$ est la matrice de l'application linéaire φ dans la base standard pour k^n.

Ici on on note $A_1, \cdots, A_n$ les colonnes d'une matrice

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix},\ A_1 = \begin{pmatrix} a_{11} \\ a_{21} \\ \cdots \\ a_{n1} \end{pmatrix}, A_2 = \begin{pmatrix} a_{12} \\ a_{22} \\ \cdots \\ a_{n2} \end{pmatrix}, \cdots, A_n = \begin{pmatrix} a_{1n} \\ a_{2n} \\ \cdots \\ a_{nn} \end{pmatrix}.$$

Pour tout vecteur colonne $X \in k^n$, on a

$$\varphi(X) = \varphi(\begin{pmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{pmatrix}) = x_1\varphi(e_1) + x_2\varphi(e_2) + \cdots + x_n\varphi(e_n) = A \begin{pmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{pmatrix} = AX.$$

Par ce choix de bases ordonnées on obtient donc un isomorphisme $GL_k(k^n) \to GL(n,k)$, où $GL_k(k^n)$ le groupe d'applications linéaires inversibles φ de k^n :

$$\varphi \mapsto A = A_\varphi,\ A_i = \varphi(e_i) \text{ pour } i = 1, \cdots, n,$$

où $A = A_\varphi$ est dite la matrice de l'application linéaire φ dans la base standard pour k^n.

Toute matrice inversible C est la matrice de passage de la base standard $e_1, \cdots, e_n$ pour k^n vers une autre base, notée $e'_1 = C_1, \cdots, e'_n = C_n$ (les colonnes de C).

Dans la base $e'_1, \cdots, e'_n$, la matrice de l'application φ est $C^{-1}AC$, c'est-à-dire, que

$$(\varphi(e'_1), \cdots, \varphi(e'_n)) = (e'_j, \cdots, e'_n)C^{-1}AC = (AC_1, \cdots, AC_n).$$

En effet, on voit $(e'_j, \cdots, e'_n)$ comme la matrice C, $e'_j = Ce_j = C_j$, $\varphi(Ce_j) = AC_j$.

Pour une marice A inversible on peut aussi voir A comme la matrice de passage de la base standard $e_1, \cdots, e_n$ pour k^n vers une autre base $A_1, \cdots, A_n$. On retrouve que dans la base $e'_1, \cdots, e'_n$, la matrice de l'application φ est $C^{-1}AC$:

$$e'_1, \cdots, e'_n \to e_1, \cdots, e_n \to A_1, \cdots, A_n \to AC_1, \cdots, AC_n$$

(en base $e'_1, \cdots, e'_n$).

(b) *Le groupe spécial linéaire* $SL(n,k)$ est le groupe de toutes les matrices (inversibles) de taille $n \times n$ sur un corps commutatif k, et de déterminant 1.

On suppose pour l'instant dans cette section que $k = \mathbb{R}$ ou $\mathbb{C}$, et on définit :

(c) *Le groupe orthogonal* $O(n,k)$ est le groupe de toutes les matrices (inversibles) de taille $n \times n$ sur k, et avec la condition

$$AA^t = I_n \tag{12.2}$$

De telles matrices forment un groupe, puisque

$$I_n I_n^t = I_n, A^{-1}(A^{-1})^t = A^{-1}(A^t)^{-1} = (A^t A)^{-1} = (I_n)^{-1} = I_n,$$

$$(AB)(AB)^t = ABB^t A^t = AA^t = I_n$$

(on appelle ce groupe orthogonal réel ou complexe dans les cas $k = \mathbb{R}$ ou $\mathbb{C}$ respectivement).

Exemples :

a) Trouver les valeurs propres et vecteurs propres de la matrice

$$A = \begin{pmatrix} \frac{5}{4} & \frac{3i}{4} \\ \frac{-3i}{4} & \frac{5}{4} \end{pmatrix} \in SO(2,\mathbb{C});$$

Conclure.

b) Montrer que

$$A := \begin{pmatrix} 1-\frac{a^2}{4} & -\frac{\sqrt{2}\,a}{2} & \frac{-1}{4}\,i\,a^2 \\ \frac{\sqrt{2}\,a}{2} & 1 & \frac{1}{2}\,i\,a\,\sqrt{2} \\ \frac{-1}{4}\,i\,a^2 & \frac{-1}{2}\,i\,\sqrt{2}\,a & 1+\frac{a^2}{4} \end{pmatrix} \in SO(3,\mathbb{C})$$

est une matrice non-diagonalisable pour tout $a \neq 0$. Trouver les valeurs propres et vecteurs propres. Conclure.

Pour une matrice

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \in O(n,k),$$

on considère les colonnes $A_1, \cdots, A_n$ qui forment une base orthonormée pour l'espace vectoriel k^n, muni du produit scalaire standard

$$((x_1, x_2, \cdots, x_n)^t, (y_1, y_2, \cdots, y_n)^t) = x_1y_1 + x_2y_2 + \cdots + x_ny_n.$$

On dit qu'une application $\varphi : k^n \to k^n$ est une isométrie, si

$$(\varphi(u), \varphi(v)) = (u, v) \text{ pour tout } u, v \in k^n.$$

Soit $e_1, \cdots, e_n$ la base standard pour k^n

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ \cdots \\ 0 \end{pmatrix}, e_2 = \begin{pmatrix} 0 \\ 1 \\ \cdots \\ 0 \end{pmatrix}, \cdots, e_n = \begin{pmatrix} 0 \\ 0 \\ \cdots \\ 1 \end{pmatrix}.$$

Par ce choix de bases orthonormales, on obtient un isomorphisme $O_k(k^n) \to O(n,k)$, où $O_k(k^n)$ est le groupe d'applications linéaires isométriques φ de k^n :

$$\varphi \mapsto A = A_\varphi,\ A_i = \varphi(e_i) \text{ pour } i = 1, \cdots, n,$$

et $A = A_\varphi$ est dite la matrice de l'application linéaire φ dans la base standard pour k^n.

On peut aussi voir A comme la matrice de passage de la base standard $e_1, \cdots, e_n$ pour k^n vers une autre base orthonormée $A_1, \cdots, A_n$.

(d) *Le groupe spécial orthogonal* $SO(n,k)$ est le groupe de toutes les matrices orthogonales de déterminant 1 :

$$SO(n,k) = O(n,k) \cap SL(n,k).$$

On écrit $SO(n)$ au lieu de $SO(n,\mathbb{R})$.

(e) *Le groupe unitaire* $U(n)$ est le groupe de toutes les matrices complexes de la taille $n \times n$ sur k, et avec la condition$A\overline{A^t} = I_n$, $\overline{A}$ est la matrice avec les éléments complexes-conjugués. On utilise souvent la notation $A^* = \overline{A^t}$, donc $(AB)^* = B^*A^*$.

On vérifie de la même façon que les matrices unitaires forment un groupe :

$$I_n I_n^* = I_n, A^{-1}(A^{-1})^* = A^{-1}(A^*)^{-1} = (A^*A)^{-1} = (I_n)^{-1} = I_n,$$

$$(AB)(AB)^* = ABB^*A^* = AA^* = I_n.$$

Pour une matrice $A \in U(n)$, on vérifie par définition que les colonnes $A_1, \cdots, A_n$, forment une base orthonormée pour l'espace vectoriel $\mathbb{C}^n$, muni du produit scalaire unitaire standard

$$((x_1, x_2, \cdots, x_n)^t, (y_1, y_2, \cdots, y_n)^t) = \overline{x_1}y_1 + \overline{x_2}y_2 + \cdots + \overline{x_n}y_n.$$

On dit qu'une application $\varphi : \mathbb{C}^n \to \mathbb{C}^n$ est une isométrie unitaire, si

$$(\varphi(u), \varphi(v)) = (u, v) \text{ pour tous } u, v \in k^n.$$

Soit $e_1, \cdots, e_n$ la base standard pour k^n

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ \cdots \\ 0 \end{pmatrix}, e_2 = \begin{pmatrix} 0 \\ 1 \\ \cdots \\ 0 \end{pmatrix}, \cdots, e_n = \begin{pmatrix} 0 \\ 0 \\ \cdots \\ 1 \end{pmatrix}.$$

Par ce choix de bases orthonormales on obtient un isomorphisme $U(\mathbb{C}^n) \to U(n)$, où $U(\mathbb{C}^n)$ le groupe d'applications $\mathbb{C}$-linéaires isométriques unitaires φ de $E = \mathbb{C}^n$:

$$\varphi \mapsto A = A_\varphi,\ A_i = \varphi(e_i) \text{ pour } i = 1, \cdots, n,$$

où $A = A_\varphi$ est dite la matrice de l'application linéaire φ dans la base standard pour $\mathbb{C}^n$.

(f) *Le groupe spécial unitaire* $SU(n)$ est le groupe de toutes les matrices unitaires de déterminant 1 :

$$SU(n) = U(n) \cap SL(n, \mathbb{C}).$$

La définition implique directement que

$$O(n) = U(n) \cap GL(n, \mathbb{R}),\ SO(n) = U(n) \cap SL(n, \mathbb{R}).$$

g) *Le groupe symplectique* $Sp(n,k)$ est un sous-groupe de $GL(2n,k)$ défini comme le groupe de toutes les matrices

$$\left\{ M \in GL(2n,k) \;\middle|\; M \in Mat(n,k), M^t J_n M = J_n \right\},$$

à l'aide de la matrice

$$J_n = \begin{pmatrix} 0 & 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \\ 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 1 \\ -1 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ 0 & -1 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \\ 0 & 0 & \cdots & -1 & 0 & 0 & \cdots & 0 \end{pmatrix} = \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix}.$$

Autrement dit, ce groupe coincide avec toutes les matrices de bloc $2n \times 2n$:

$$Sp(n,k) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in GL(2n,k) \;\middle|\; a,b,c,d \in Mat(n,k), {}^t ac = {}^t ca, {}^t ad - {}^t cb = I_n \right\}.$$

12.9.1 Rappels sur les classes d'éléments congugués

dans les groupes $G = GL(n,k), O(n), U(n)$ a) Toute matrice $A \in Mat(n,k)$ sur un corps commutatif k est annulée par son polynôme caractéristique

$$p_A(T) = \det(A - T \cdot I_n) = \sum_{j=0}^{n} \tau_j(A)(-T)^{n-j}$$

(le théorème de Hamilton-Cayley), où $\tau_j(A) \in k$ coincide avec la somme des $\binom{n}{j}$ mineurs diagonaux de la taille j. En particuler, $\tau_1(A) = \operatorname{Tr} A$, $p_A(0) = \tau_n(A) = \det A$. Soit

$$(-1)^n p_A(T) = p_1^{l_1}(T) \cdot \ldots \cdot p_s^{l_s}(T)$$

la décomposition en produit de puissances de polynômes irréductibles unitaires distincts sur k, et on note $P_i(T) = p_A(T)/p_i^{l_i}(T)$ pour $(i = 1, \cdots, s)$.

Pour trouver une forme canonique de A, on utilise l'action de l'anneau $R = k[T]$ sur le k-espace vectoriel $V = k^n$, donnée par $T \cdot X = AX$, et on note V_A ce $k[T]$-module. Lorsque $V = k^n$ est annulé par $p_A(A) = p_i^{l_i}(A)P_i(A)$, on utilise la décomposition en sous-espaces invariants

$$V = \bigoplus_i V_i = P_i(A)V, \quad X = P_1(A)Q_1(A)X + \cdots + P_s(A)Q_s(A)X$$

en utilisant la décomposition $P_1(T)Q_1(T) + \cdots + P_s(T)Q_s(T) = 1$, où $P_i(T) = p_\varphi(T)/p_i^{l_i}(T)$. Tout sous-espace est annulé par $p_i^{l_i}(A)$; c'est un sous-R-module sur l'anneau $R = k[T]$, qui se décompose en une somme directe des sous-R-modules monogènes isomorphes à $R = k[T]/(p_i^{l_{ij}}(T))$, avec $l_{ij} \leq l_i$, et

$$p_i^{l_i}(T) = \prod_{j=1}^{s_i} p_i^{l_{ij}}(T), V \cong \bigoplus_{i=1}^{s} \bigoplus_{j=1}^{s_i} k[T]/(p_i^{l_{ij}}(T)).$$

On en déduit que A_1 et $A_2 \in G = GL(n,k)$ sont conjuguées si et seulement si les $k[T]$-modules V_{A_1} et V_{A_2} sont isomorphes :

$$A_1 = C^{-1} A_2 C, \text{ pour un isomorphisme } C : X \mapsto CX, \; C : k^n \to k^n.$$

Rappels sur la forme normale des matrices complexes, $k = \mathbb{C}$ (voir [Godement], §35) : pour tout nombre complexe $\lambda \in \mathbb{C}$ on pose

$$J_n(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda & 1 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 0 & \lambda & 1 \\ 0 & 0 & \cdots & 0 & 0 & \lambda \end{pmatrix} = \lambda \cdot I_n + N_n,\ N_n = J_n(0).$$

Une condition nécessaire et suffisante pour qu'une matrice complexe $A \in \mathrm{Mat}_n(\mathbb{C})$ soit conjugée à la matrice $J_n(\lambda)$ (c'est-à-dire, il existe une matrice $C \in \mathrm{GL}_n(\mathbb{C})$ telle que $C^{-1}AC = J_n(\lambda)$) est que A n'a qu'un seul vecteur colonne propre (à proportionnalité près), de valeur propre égale à λ.

On montre que toute matrice complexe $A \in \mathrm{Mat}_n(\mathbb{C})$ est conjugée à une matrice

$$B = \mathrm{diag}\{B_1, \cdots, B_r\}$$

diagonale bloc par bloc avec tout bloc donné par une matrice $J_{n_j}(\lambda_j)$ (c'est-à-dire, il existe une matrice $C \in \mathrm{GL}_n(\mathbb{C})$ telle que $C^{-1}AC = B$).

b) On montre que toute matrice orthogonale réelle $A \in O(n, \mathbb{R})$ est conjugée à une matrice

$$B = \mathrm{diag}\{B_1, \cdots, B_r\}$$

diagonale bloc par bloc avec tout bloc donné soit par (± 1), soit par une matrice de rotation $B_j = \begin{pmatrix} \cos\theta_j & -\sin\theta_j \\ \sin\theta_j & \cos\theta_j \end{pmatrix}$ (c'est-à-dire, il existe une matrice $C \in O(n, \mathbb{R})$ telle que $C^{-1}AC = B$).

Corollaire 12.9.1 *Toute matrice* $A \in SO(3)$ *est conjugée à une matrice*

$$B = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

c) On montre que toute matrice unitaire complexe $A \in U(n)$ est conjugée à une matrice

$$B = \mathrm{diag}\{B_1, \cdots, B_r\}$$

diagonale avec tout bloc donné soit par $(\alpha_j), |\alpha_j| = 1$, (c'est-à-dire, il existe une matrice $C \in U(n)$ telle que $C^{-1}AC = B$).

Homomorphisme remarquable de $SU(2)$ dans $SO(3)$

Soit

$$g = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \in SU(2) \text{ alors } g^{-1} = \begin{pmatrix} \delta & -\beta \\ -\gamma & \alpha \end{pmatrix} \in SU(2)$$

On montre que

$$g = \begin{pmatrix} \alpha & \beta \\ -\overline{\beta} & \overline{\alpha} \end{pmatrix}, |\alpha|^2 + |\beta|^2 = 1.$$

On en déduit que $SU(2)$ est homéomorphe à $S^3 \subset \mathbb{R}^4$.

On pose

$$b_\varphi = \begin{pmatrix} e^{i\frac{\varphi}{2}} & 0 \\ 0 & e^{-i\frac{\varphi}{2}} \end{pmatrix} \in SU(2), \quad c_\theta = \begin{pmatrix} \cos\frac{\theta}{2} & i\sin\frac{\theta}{2} \\ i\sin\frac{\theta}{2} & \cos\frac{\theta}{2} \end{pmatrix} \in SU(2)$$

On vérifie que

$$a(\varphi,\theta,\psi) = b_\varphi c_\theta b_\psi = \begin{pmatrix} \cos\frac{\theta}{2}e^{i\frac{\varphi+\psi}{2}} & i\sin\frac{\theta}{2}e^{i\frac{\varphi-\psi}{2}} \\ i\sin\frac{\theta}{2}e^{-i\frac{\varphi-\psi}{2}} & \cos\frac{\theta}{2}e^{-i\frac{\varphi+\psi}{2}} \end{pmatrix}$$

On en déduit que toute matrice

$$g = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \in SU(2)$$

se décompose sous la forme

$$g = a(\varphi,\theta,\psi) = b_\varphi c_\theta b_\psi$$

(INDICATION :Il suffit de poser

$$|\alpha| = \cos\frac{\theta}{2}, \text{Arg}\alpha = \frac{\varphi+\psi}{2}, \quad |\beta| = \sin\frac{\theta}{2}, \text{Arg}\beta = \frac{\varphi-\psi+\pi}{2}).$$

Ceci implique l'existence d'un homomorphisme

$$\Phi : SU(2) \to SO(3)$$

tel que

$$\Phi(b_\varphi) = B_\varphi, \quad \Phi(c_\theta) = C_\theta,$$

(utiliser l'action du groupe $SU(2)$ par conjugaison sur les matrices hermitiennes de trace nulle

$$H_x = \begin{pmatrix} x_3 & x_1 + ix_2 \\ x_1 - ix_2 & -x_3 \end{pmatrix}$$

en base de matrices de Pauli

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \ \sigma_2 = \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}, \ \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix})$$

On trouve le noyau de Φ, et on en déduit que $SO(3)$ est homéomorphe à l'espace projectif réel $\mathbb{RP}^3$.

12.10 Simplicité du groupe SO(3)

Par le théorème connu d'Euler, tout élément du groupe SO(3) des rotations propres de l'espace euclidien de dimension 3 est une rotation par rapport à un axe. Disons que les matrices

$$B_\varphi = \begin{pmatrix} \cos\varphi & -\sin\varphi & 0 \\ \sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{pmatrix}, C_\theta = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix} \tag{1}$$

correspondent aux rotations relatives des axes z et x respectivement. Rapellons que par la paramétrisation des rotations par les angles d'Euler φ, θ, ψ ($0 \leq \varphi, \psi < 2\pi, 0 \leq \theta < \pi$), toute matrice $A \in \mathrm{SO}(3)$ peut-être représentée comme le produit

$$A = B_\varphi C_\theta B_\psi.$$

THÉORÈME 12.10.1 *Le groupe* SO(3) *est simple, i.e. pour tout sous-groupe distingué* $H \subset \mathrm{SO}(3)$ *soit* $H = \mathrm{SO}(3)$ *soit* $H = \{1\}$.

Démonstration. Si $H \neq \{1\}$ on peut supposer que

$$H \ni B = \begin{pmatrix} \cos\varphi & -\sin\varphi & 0 \\ \sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

avec $\cos\varphi \neq 1$, parce que avec tout élément du sous-groupe distingué H son élément conjugé doit être contenu dans H. D'autre part, pour tout $A \in \mathrm{SO}(3)$ on a $BAB^{-1}A^{-1} \in H$, et on voit directement qu'il existe $A \in \mathrm{SO}(3)$ tel que $BAB^{-1}A^{-1} \neq 1$ (le centre du groupe SO(3) est trivial). Alors on peut supposer que $BAB^{-1}A^{-1}$ et une rotation non-triviale par l'angle $\theta \neq 2\pi k$, c'est à dire, $tr(BAB^{-1}A^{-1}) = 1 + 2\cos\theta$.

Considérons la famille continue $A_t \in \mathrm{SO}(3)$ des matrices paramétrisées par $t \in [0, 1]$ telle que $A_1 = A$, $A_0 = 1_3$. Par exemple, si pour une matrice $C \in SO(3)$ on a $C^{-1}AC = B_\psi$ on peut prendre $A_t = CB_{t\psi}C^{-1}$.

La fonction $f(t) = tr(BA_tB^{-1}A_t^{-1})$ est alors une fonction continue telle que $f(0) = 3, f(1) = 1 + 2\cos\theta$.

Ceci implique que pout tout θ' tel que $0 \leq \theta' \leq \theta$ le sous–groupe H contient une rotation par l'angle θ'.

En considerant les puissances de ces éléments on voit bien que H contient des rotations par tout angle donné, c'est à dire $H = \mathrm{SO}(3)$ car il est distingué.

12.11 Formes quadratiques

12.11.1 Définitions

(voir [Se70], p.51, [Bourbaki], §3, n°4, [Godement], §36, [Lang], Chapter XIV).

Cours N 7 Lundi 2 mars 2015

DÉFINITION 12.11.1 *Soit V un module sur un anneau commutatif A. Une application $Q : V \to A$ est appelée une forme quadratique sur V si*
1) On a $Q(ax) = a^2Q(x)$ pour tout $a \in A$, et $x \in V$,
2) L'application $(x, y) \mapsto Q(x + y) - Q(x) - Q(y)$ est une application A-bilinéaire.

Un tel couple est appelé un module quadratique.

Nous nous limiterons au cas où l'anneau A est un corps k de caractéristique $\neq 2$ et que V est de dimension finie sur k.

On posera alors

$$\langle x, y\rangle = \frac{1}{2}[Q(x + y) - Q(x) - Q(y)]$$

(le produit scalaire associé à Q). On a $Q(x) = \langle x, x\rangle$, cela établit une correspondance bijective entre formes quadratiques et formes bilinéaires symétriques.

DÉFINITION 12.11.2 *Si (V, Q) et (V', Q') sont deux modules quadratiques, on appelle* morphisme métrique *de (V, Q) dans (V', Q') toute application A-linéaire $f : V \to V'$ telle que $Q' \circ f = Q$; on a alors $\langle f(x), f(y)\rangle = \langle x, y\rangle$.*

Soit $(e_i)_{1\leq i\leq n}$ une base de V, $A = (a_{ij})$ la matrice de Q par rapport à cette base : $a_{ij} = \langle e_i, e_j\rangle$; c'est une matrice symétrique. Si $x = \sum x_i e_i$ est un élément de V, on a

$$Q(x) = \sum_{i,j} a_{ij}x_i x_j$$

ce qui montre que $Q(x)$ est une forme quadratique.

Si l'on modifie la base (e_i) au moyen d'une matrice inversible C, la matrice A' de Q par rapport à la nouvelle base est $C \cdot A \cdot C^t$. On a en particulier

$$\det(A') = \det(A)\det(C)^2$$

ce qui montre que $\det(A)$ est déterminé à la multiplication près par un élément de k^2 ; on l'appelle le discriminant de Q et on le note $\mathrm{disc}(Q)$.

12.11.2 Orthogonalité

Soit (V, Q) un module quadratique sur k. Deux éléments x, y de V sont dits orthogonaux si $\langle x, y\rangle = 0$. L'ensemble des éléments orthogonaux à une partie H de V est noté $H^\perp$: c'est un sous-espace vectoriel de V, et on dit que V_1 et V_2 sont orthogonaux si $V_1 \subset V_2^\perp$. L'orthogonal $V^\perp$ de V tout entier est appelé le radical (ou le noyau) de V, et on le note $\mathrm{rad}(V)$. Sa codimension s'appelle le rang de Q. Si $V^\perp = \{0\}$; cela équivaut à dire que le discrimnant de Q est $\neq 0$ (auquel cas on peut le considérer comme un élément du groupe k^*/k^{*2}).

Soit U un sous-espace vectoriel de V, et soit U^* le dual de U. Soit $q_U : V \to U^*$ l'application qui associe à tout $x \in V$ la forme linéaire $y \mapsto \langle x, y\rangle$. Le noyau de q_U est $U^\perp$. En particulier on voit que Q est non-dégénérée si et seulement si $q_V : V \to V^*$ est un isomorphisme.

DÉFINITION 12.11.3 *Soient $U_1, \cdots, U_m$ des sous-espaces vectoriels de V. On dit que V est somme directe orthogonale des U_i, si ceux-ci sont orthogonaux deux à deux. On écrit alors*

$$V = U_1 \hat{\oplus} \cdots \hat{\oplus} U_m.$$

REMARQUE 12.11.4 *Si $x \in V$ a pour composantes x_i dans U_i, on a*

$$Q(x) = Q_1(x_1) + \cdots + Q_m(x_m),$$

où $Q_i = Q|_{U_i}$ désigne la restriction de Q à U_i.

Inversement, si (U_i, Q_i) est une famille des modules quadratiques, la formule ci-dessus munit $V = \oplus_i U_i$ d'une forme quadratique Q, dite somme directe des Q_i, et l'on a $V = U_1 \hat{\oplus} \cdots \hat{\oplus} U_m$.

PROPOSITION 12.11.5 *Si U est supplémentaire de $\mathrm{rad}(V)$ dans V, on a $V = U \hat{\oplus} \mathrm{rad}(V)$.*
∎

PROPOSITION 12.11.6 *Supposons (V, Q) non-dégénéré. Alors*
i) *Tous morphisme métrique de V dans un module quadratique (V', Q') est injectif.*
ii) *Pour tous sous-espace vectoriel U de V, on a*

$$U^{\perp\perp} = U, \dim U + \dim U^{\perp} = \dim V,$$
$$\mathrm{rad}(U) = \mathrm{rad}(U^{\perp}) = U \cap U^{\perp}.$$

Pour que U soit non-dégénéré, il faut et il suffit que $U^{\perp}$ le soit, auquel cas $V = U \hat{\oplus} U^{\perp}$.
iii) *Si V est somme directe orthogonale de deux sous-espaces, ceux-ci sont non-dégénérés, et chacun d'eux est l'orthogonal de l'autre.*

Si $f : V \to V'$ est un morphisme métrique, et si $f(x) = 0$, on a $\langle x, y \rangle = \langle f(x), f(y) \rangle = 0$ pour tout $y \in V$, d'où $x = 0$ puisque (V, Q) est non-dégénéré.

Si U est un sous-espace vectoriel de V, l'homomorphisme $q_U : V \to U^*$ défini plus haut est surjectif ; en effet il s'obtient en composant $q_V : V \to V^*$ avec la *surjection canonique* $V^* \to U^*$, et l'on suppose que q_V est bijectif. On a donc une suite exacte

$$0 \to U^{\perp} \to V \to U^* \to 0,$$

d'où $\dim V = \dim U^* + \dim U^{\perp} = \dim U + \dim U^{\perp}$.

Ceci montre que U et $U^{\perp\perp}$ ont même dimension ; comme U est contenu dans $U^{\perp\perp}$, on a $U = U^{\perp\perp}$, et on en déduit que $\mathrm{rad}(U) = \mathrm{rad}(U^{\perp})$, d'où en même temps la première assertion de ii). Enfin iii) est triviale. ∎

12.11.3 Vecteurs isotropes

DÉFINITION 12.11.7 *Un élément x d'un module quadratique (V, Q) est isotrope si l'on a $Q(x) = 0$. Un sous-espace U de V est dit isotrope si tous les éléments sont isotropes.*

On a évidemment :

$$U \text{ isotrope } \iff U \subset U^{\perp} \iff Q|_U = 0.$$

DÉFINITION 12.11.8 *On appelle plan hyperbolique ayant une base formée de deux éléments isotrope x, y tels que $\langle x, y\rangle \neq 0$.*

Quitte à multiplier y par $1/\langle x, y\rangle$, on peut supposer que $\langle x, y\rangle = 1$. La matrice de la forme quadratique par rapport à x, y est alors simplement $\left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$; son discriminant est -1 (en particulier, elle est non-dégénérée).

PROPOSITION 12.11.9 *Soit x un élément isotrope $\neq 0$ d'un module quadratique non-dégénéré (V, Q). Il existe un sous-espace U de V qui contient x et qui est un plan hyperbolique.*

Puisque V est non-dégénéré, il existe $z \in V$ tel que $\langle x, z\rangle = 1$. L'élément $y = 2z - \langle z, z\rangle x$ est isotrope et $\langle x, y\rangle = 2$. Le sous-espace $U = kx + ky$ repond à la question.

COROLLAIRE 12.11.10 *Si (V, Q) est non-dégénéré et contient un élément isotrope non-nul, on a $Q(V) = k$.*

12.11.4 Base orthogonales

DÉFINITION 12.11.11 *Une base $(e_1, e_2, \cdots, e_n)$ d'un module quadratique (V, Q) est dite orthogonale, si elle est formée d'éléments deux à deux orthogonaux, i.e. si $V = ke_1 \hat{\oplus} \cdots \hat{\oplus} e_n$.*

THÉORÈME 12.11.12 *Tout module quadratique (V, Q) possède une base orthogonale.*

Cela se démontre par récurrence sur $n = \dim V$, le cas $n = 0$ étant trivial. Si V est isotrope, toute base de V est orthogonale. Sinon, on choisit un element $e_1 \in V$ tel que $\langle e_1, e_1\rangle \neq 0$. L'orthogonal H de e_1 est un hyperplan, et comme e_1 n'appartient pas à H, on a $V = ke_1 \hat{\oplus} H$; vu l'hypothèse de récurrence, H possède une base orthogonale $(e_2, \cdots, e_n)$; il est clair que $(e_1, e_2, \cdots, e_n)$ répond à la question. ∎

Méthodes d'orthogonalisation* Gram-Schmidt, Jacobi, (voir [W] :
Jørgen Pedersen Gram (27 juin 1850- 29 avril 1916),
Erhard Schmidt (13 janvier 1876 - 6 décembre 1959)
Carl Gustav Jacob Jacobi (10 décembre 1804 -18 février 1851))

12.12 Espace d'Euclide et mécanique quantique*

Selon §11 de [KosMan], on va décrire réalisations et interptétations physiques de l'espace d'Euclide $\mathcal{E}$ de dimension trois sur $\mathbb{R}$, et de l'espace-temps $\mathcal{M}$ de Minkowski(-Poincaré), de dimension quatre sur $\mathbb{R}$, mini d'une métrique de signature $(r_+, r_-) = (1, 3)$. Du point de vue mathématique, leurs propriétés essentielles sont : le lien des rotations avec les quaternions, l'existence du produit vectoriel, et la géométrie des vecteurs de longueur nulle dans $\mathcal{M}$.

Pour formuler ces propriétés, on introduit l'epace unitaire auxiliaire $\mathcal{H}$ de dimension deux (sur $\mathbb{C}$), amettant une intérprétation dans le cadre de mécanique quantique.

PROPOSITION 12.12.1 *On considère l'espace réel $\mathcal{E}$ d'opérateurs autoadjoints dans $\mathcal{H}$ de trace nulle, et on définit $|f|$ comme $\sqrt{|\det f|}$ (=la valeur propre positive de f). Alors $\mathcal{E}$ est un espace d'Euclide $\mathcal{E}$ de dimension trois sur $\mathbb{R}$.*

En effet, les opérateurs f sont présentés dans une base orthonormée de $\mathcal{H}$ par matrices de type

$$\begin{pmatrix} a & \bar{b} \\ b & -a \end{pmatrix} = \mathrm{Re}b \cdot \sigma_1 + \mathrm{Im}\, b \cdot \sigma_2 + a \cdot \sigma_3, a \in \mathbb{R}, b \in \mathbb{C},$$

où

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \ \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \ \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

sont les *matrices de Pauli*, voir Exercice 12.6. On pose

$$\langle f, g \rangle = \frac{1}{2} \mathrm{Tr}(fg).$$

C'est un produit symétrique scalaire, et $|f|^2 = \frac{1}{2}\mathrm{Tr}(f^2) = \frac{1}{2}(\lambda^2 + \lambda^2) = |\det f|$.

On appelle *direction* dans $\mathcal{E}$ l'ensemble $\mathbb{R}_+ = \{af \mid a > 0\}$, pour un vecteur non-nul $f \in \mathcal{E}$.

PROPOSITION 12.12.2 *Il existe une bijection entre les directions dans $\mathcal{E}$ et les décompositions de $\mathcal{H}$ dans une somme directe orthogonale de sous-espaces $\mathcal{H}_+ \oplus \mathcal{H}_-$. Notamment, on fait corresponre à une direction $\mathbb{R}_+ f$ le sous-espace propre $\mathcal{H}_+$ de valeur propre positive et le sous-espace propre $\mathcal{H}_-$ de valeur propre négative.* ∎

Une interprétation physique : on identifie $\mathcal{H}$ avec l'espace d'états du système quantique "une particule de spin 1/2, localisée autour de l'origine" (par exemple, l'éléctron). En choisissant une direction $\mathbb{R}_+ f \subset \mathcal{E}$, on fait passer des particules de spin $1/2$ dans un champ magnétique non uniforme de direction verticale. Dans ce champ le système possède deux états stationnaires, notamment $\mathcal{H}_+$ et $\mathcal{H}_-$.

L'expérience de Stern et Gerlach (voir [W]) "est une expérience de mécanique quantique démontrant la quantification du spin. L'expérience, mise au point par Otto Stern et Walther Gerlach en 1920, consiste à faire passer des particules de spin $1/2$ (en l'occurrence des atomes d'argent) dans un champ magnétique non uniforme de direction verticale. Dans le modèle classique de l'atome de Niels Bohr, le faisceau de particules devrait être dispersé verticalement en raison de la composante verticale du spin de l'atome qui prend un continuum de valeurs entre $-1/2$ et $+1/2$ en fonction de l'orientation de l'atome. En revanche, l'expérience montre que le faisceau se sépare en deux, indiquant que la composante verticale du spin ne peut prendre que les valeurs $+1/2$ et $-1/2$."

PROPOSITION 12.12.3 *On a $(f, g) = 0$ si et seulement si $fg + gf = 0$.*

On a

$$(f, g) = \frac{1}{2}\mathrm{Tr}(fg) = \frac{1}{4}\mathrm{Tr}(fg + gf) = \frac{1}{4}\mathrm{Tr}[(f+g)^2 - f^2 - g^2],$$

mais tous les opérateurs de type f^2 sont scalaires, donc $fg+gf$ est scalaire, et il s'annule si et seulement si sa trace s'annulle. ∎

Cette preuve implique qu'une base $\{e_1, e_2, e_3\}$ est orthonormale si et seulement si

$$e_1^2 = e_2^2 = e_3^2 = \mathrm{Id}; \quad e_i e_j + e_j e_i = 0 (i \neq j).$$

On particulier, les matrices de Pauli $\sigma_1, \sigma_2, \sigma_3$ forment une base orthonormée de $\mathcal{E}$:

$$\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \sigma_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}; \quad e_i e_j + e_j e_i = 0 (i \neq j).$$

Une assertion réciproque montre la signification mathématique des matrices de Pauli

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \ \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \ \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

PROPOSITION 12.12.4 *Pour toute base orthonormée $\{e_1, e_2, e_3\}$ de $\mathcal{E}$ il existe une base orthonormée $\{h_1, h_2\}$ de l'espace $\mathcal{H}$ telle que*

$$A_{e_1} = \sigma_1, \ A_{e_2} = \sigma_2 \ ou \ -\sigma_2, \ A_{e_3} = \sigma_3,$$

où A_e est la matrice d'un opérateur e dans la base $\{h_1, h_2\}$; une telle base est déterminée à multiplication près par un nombre complexe de module 1.

PREUVE. Les valeurs propres de e_i sont ± 1, et soit $\mathcal{H} = \mathcal{H}_+ \oplus \mathcal{H}_-$, ou e_3 agit sur $\mathcal{H}_+$ triviallement, et sur $\mathcal{H}_-$ par le changement de signe. On choisit d'abord les vecteurs $h_1' \in \mathcal{H}_+$, et $h_2' \mathcal{H}_-$, avec $|h_1'| = |h_2'| = 1$. Ils sont déterminés à une multiplication près par les nombres complexes $e^{i\varphi_1}, e^{i\varphi_2}$, la matrice e_3 en $\{h_1', h_2'\}$ est $\sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$.

Puis, l'orthogonalité implique que

$$e_1(h_1') = e_1 e_3(h_1') = -e_3 e_1(h_1'),$$

donc $e_1(h_1')$ est un vecteur propre de e_3 de valeur propre -1. Donc, $e_1(h_1') = \alpha h_2'$. De même façon, $e_1(h_2') = \beta h_1'$. La matrice de e_1 en base $\{h_1', h_2'\}$ est hermitienne, donc $\alpha = \overline{\beta}$. La condition $e_1^2 = \mathrm{Id}$ donne $\alpha\beta = 1 = |\alpha|^2 = |\beta|^2$. Il reste à changer $\{h_1', h_2'\}$ par $\{h_1, h_2\} = \{x h_1', y h_2'\}$, où $|x| = |y| = 1$, de telle façon que la matrice de e_1 en $\{h_1, h_2\}$ devient $\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, ceci donne les conditions :

$$e_1(h_1) = x e_1(h_1') = x\alpha h_2' = \alpha x y^{-1} h_2, \tag{12.3}$$

$$e_1(h_2) = y e_1(h_2') = y\beta h_1' = \beta y x^{-1} h_1. \tag{12.4}$$

Ceci dit, $xy^{-1} = \alpha^{-1}$ implique $\alpha x y^{-1} = \beta y x^{-1} = 1$, et on peut poser $x = 1, y = \alpha$, et la base $\{h_1, h_2\}$ est déterminé à multiplication près par un nombre complexe de module 1.

Et en base $\{h_1, h_2\}$ on a $A_{e_3} = \sigma_3, A_{e_1} = \sigma_1$, Le même raisonnement que pour e_1 monte que dans cette base $A_{e_2} = \begin{pmatrix} 0 & \gamma \\ \overline{\gamma} & 0 \end{pmatrix}$, où $|\gamma|^2 = 1$. De plus, la condition d'orthogonalité $e_1 e_2 + e_2 e_1 = 0$ donne

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & \gamma \\ \overline{\gamma} & 0 \end{pmatrix} + \begin{pmatrix} 0 & \gamma \\ \overline{\gamma} & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = 0$$

ceci implique $\gamma + \overline{\gamma} = 0$, donc $\gamma = i$ ou $-i$, et $A_{e_2} = \sigma_2$ ou $-\sigma_2$. ∎

COROLLAIRE 12.12.5 *L'espace $\mathcal{E}$ est muni d'une orientation distinguée : une base $\{e_1, e_2, e_3\}$ de $\mathcal{E}$ appartienne à la classe de telle orientation distinguée, s'il existe une base orthonormée $\{h_1, h_2\}$ de l'espace $\mathcal{H}$ telle que $A_{e_1} = \sigma_1, \ A_{e_2} = \sigma_2, \ A_{e_3} = \sigma_3$.* ∎

Les opérateurs $\sigma_1, \sigma_2, \sigma_3$ sont appelés les observables des projections du spin sur les axes de $\mathcal{E}$: cette terminologie est expliquée par l'interprétation de mécanique quantique ci-dessus. Le facteur $1/2$ est introduit pour que les valeurs propres soient $\pm 1/2$.

12.13 Espace-temps de Minkowski*

On appelle l'espace-temps $\mathcal{M}$ de Minkowski, un espace de dimension quatre sur $\mathbb{R}$, mini d'une métrique de signature $(r_+, r_-) = (1, 3)$ (parfois on travaille avec la signature $(r_+, r_-) = (3, 1)$

Selon [W], Relativité restreinte, on nomme Relativité restreinte une première version de la théorie de la Relativité, émise en1905 par Albert Einstein, qui ne considérait pas la question des accélérations d'un référentiel, ni les interactions d'origine gravitationnelles. Cette théorie a introduit . . .la notion d'espace-temps et expliqué quelques phénomènes étonnants, mais vérifiés expérimentalement, de variation des mesures de longueur et de durée entre un observateur et un autre, chacun d'eux étant situé dans un référentiel différent.

Les postulats d'Einstein (1905)

P1 : Les lois de la physique sont les mêmes dans tous les référentiels inertiels. Rappelons que depuis Newton, un référentiel est dit inertiel si tout corps isolé y possède un mouvement rectiligne uniforme, c'est à dire un vecteur vitesse constant

P2 : La vitesse de la lumière dans le vide a la même valeur dans tous les référentiels inertiels."

Les principes de bases :

a) *Points de* $\mathcal{M}$ (évenements). Dans un référentiel un événement est caractérisé par ses coordonnées, spatio-temporelles, « tel endroit, tel instant ».

b) *Unités de mesure.* On utilise c = vitesse de la lumière (299 792 458 m/s) qui donne un moyen pour passer des unités de temps t_0 aux unités de longueur $l_0 = ct_0$ (par exemple, 1 seconde-lumiére). Dans ces unités, la vitesse de la lumière est égale 1.

On utilisera la notation $\ell = (x_0, x_1, x_2, x_3) \in \mathcal{M}$ (avec $c = 1$) au lieu de la notation traditionnelle (t, x, y, z), avec les coordonnées convenables expliquées ci-dessous, voir e).

c) *Intervalle d'espace-temps.* Deux événements situés respectivement en ℓ_1 et ℓ_2 seront séparés par un intervalle d'espace-temps dont le carré est $(\ell_1 - \ell_2, \ell_1 - \ell_2)$. Les intervalles tels que : $(\ell_1 - \ell_2, \ell_1 - \ell_2) < 0$ (avec une signature $(1, 3) = (+, -, -, -)$) sont appelés intervalles de genre espace.

Les intervalles tels que : $(\ell_1 - \ell_2, \ell_1 - \ell_2) = 0$ sont appelés intervalles de genre lumière.

Les intervalles tels que : $(\ell_1 - \ell_2, \ell_1 - \ell_2) > 0$ (avec une signature $(1, 3) = (+, -, -, -)$) sont appelés intervalles de genre temps.

d) *Lignes de l'univers.* Selon [W], Ligne d'univers, en physique, la ligne d'univers d'un objet est l'unique trajectoire de cet objet lorsqu'il voyage à travers l'espace-temps en 4 dimensions.

S'il existe un vecteur de genre temps sur une ligne droite $L \subset \mathcal{M}$, tous tels vecteurs sont de genre temps.

On appelle telles lignes lignes d'univers des référentiel inertiels (tout corps isolé y possède un mouvement rectiligne uniforme, c'est à dire un vecteur vitesse constant).

Un référentiel inertiel, qui n'est pas de « tel endroit, tel instant », se déplace sur un décalage $\ell + L$ d'une ligne d'univers de genre temps. Soient ℓ_1, ℓ_2 deux points sur la lignes d'univers d'un référentiel inertiel. Alors $(\ell_1 - \ell_2, \ell_1 - \ell_2) > 0$, et l'intervalle $|\ell_1 - \ell_2| = (\ell_1 - \ell_2, \ell_1 - \ell_2)^{1/2}$ est le temps propre de ce référentiel mesuré à l'aide de son horloge.

e) *Espace physique d'un référentiel inertiel.* Le sous-ensemble

$$\mathcal{E}_\ell = \ell + L^\perp \subset \mathcal{M}$$

s'interprète comme l'ensemble des points de "l'espace instantané physique" du référentiel inertiel, qui se trouve au point ℓ de sa ligne d'univers, et le complémentaire est pris par rapport à la métrique de Minkowski. On voit que $\mathcal{M} = L \oplus L^\perp$, et que $L^\perp$ est muni d'une structure d'un espace euclidien de dimension trois (avec une métrique euclidienne négative définie)

f) *Systèmes de coordonnées intértiels.* Soit L un vecteur de genre temps muni d'une orientation, e_0 un vecteur positive de longueur 1, $\{e_1, e_2, e_3\}$ une base orthonormée dans $L^\perp$: $(e_i, e_i) = -1$ pour $i = 1, 2, 3$. Un système de coordonnées d'une telle base et dite inertiel.

On utilisera la notation $\ell = (x_0, x_1, x_2, x_3) = x_0e_0 + x_1e_1 + x_2e_2 + x_3e_3 \in \mathcal{M}$ (avec $c = 1$) au lieu de notation traditionnelle (t, x, y, z). Si $\ell = (x_0, x_1, x_2, x_3), \ell' = (y_0, y_1, y_2, y_3)$, le produit scalaire alors s'écrit

$$(\ell, \ell') = ((x_0, x_1, x_2, x_3), (y_0, y_1, y_2, y_3)) = x_0y_0 - x_1y_1 - x_2y_2 - x_3y_3.$$

Puisque $x_0 = ct_0$, où t_0 le temps propre, la longueur de l'intervalle espace-temps de l'origine au point $\ell = \sum_{i=0}^3 x_ie_ii \in \mathcal{M}$ est égale à $(c^2t_0^2 - \sum_{i=1}^3 x_i^2)^{1/2}$.

Tout système de coordonnées inertiel dans $\mathcal{M}$ identifie $\mathcal{M}$ avec l'espace de Minkowski $(\mathbb{R}^4, x_0^2 - \sum_{i=1}^3 x_i^2)$. Les isométries de cet espace forme le *groupe de Lorentz*; les isométries, respectant l'orientation du temps, forment son *sous-groupe orthochrone*

g) *Cône de lumière.* L'ensemble de $\ell \in \mathcal{M}$ avec $(\ell, \ell) = 0$ est dit cône de lumière C (en origine). Dans tous système inertiel de coordonnées C est donné par l'equation $x_0^2 = \sum_{i=1}^3 x_i^2$.

Pour $x_0 > 0$ le point (x_0, x_1, x_2, x_3) de C est séparé du référentiel de $(x_0, 0, 0, 0)$ par l'intervalle de genre temps de carré $(\sum_{i=1}^3 x_i^2)^{1/2} = -x_0^2$, c'est-à-dire, ce point se trouve à distance, parcouru par le temps x_0 par un photon, emis de l'origine au moment de temps initial.

Les vecteurs $\ell \in \mathcal{M}$ avec $(\ell, \ell) = 0$ n'ont pas d'interprétation physique. Les lignes de tels vecteurs doivent correspondre aux lignes d'univers de "tachyons" (voir [W]) : "En physique des particules on nomme tachyon une particule qui, si elle existait, se déplacerait à une vitesse supérieure à celle de la lumière."

Étude mathématique de $\mathcal{M}$.

Réalisation de $\mathcal{M}$ comme l'espace des métriques. On considère de nouveau l'epace unitaire auxiliaire $\mathcal{H}$ de dimension deux (sur $\mathbb{C}$), admettant une intérprétation dans le cadre de mécanique quantique, et on considère l'espace réel $\mathcal{M}$ de produits scalaires hermitiens dans $\mathcal{H}$. En choisissant une base $\{h_1, h_2\}$ de $\mathcal{H}$, la matrice de Gram G de telle métrique $\ell \in \mathcal{M}$ donne une bijection de $\mathcal{M}$ avec toutes les matrices hermitiennes de taille 2×2. On fait correspondre à $\ell \in \mathcal{M}$ le déterminant de G, noté ℓ.

Le passage de $\{h_1, h_2\}$ vers $\{h_1', h_2'\} = \{h_1, h_2\}V$ mêne au changement de G par $G' = V^tGV$, et $\det G' = |\det G|^2 \det G$, donc le calcul de $\det \ell$ dans toute base de $\mathcal{H}$

dans la même classe par l'action de $\mathrm{SL}_2(2,\mathbb{C})$, donne le même résultat. Desormais on fixe une telle classe de bases de $\mathcal{H}$, et un changement de la classe donne la miltiplication de $\det \ell$ par un scalaire posivive.

PROPOSITION 12.13.1

a) $\mathcal{M}$ est un espace de dimension 4 sur $\mathbb{R}$.

b) Il existe une seule métrique symétrique (ℓ, m) pour laquelle $(\ell, \ell) = \det \ell$. Sa signature est égale $(1,3)$, donc $\mathcal{M}$ est un espace de Minkowski.

PREUVE. a) L'espace de matrices hermitiennes de taille 2×2 admèt la base $\sigma_0 = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$, $\sigma_1 = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$, $\sigma_2 = \left(\begin{smallmatrix} 0 & -i \\ i & 0 \end{smallmatrix}\right)$, $\sigma_3 = \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$ (matrices de Pauli). Donc $\dim \mathcal{M} = 4$.

b) Montrons, que dans la réalisation matricielle de $\mathcal{M}$ la fonction $\det \ell$ est une forme quadratique dont la fonction polarisée correspondant a la forme

$$(\ell, m) = \frac{1}{2}(\mathrm{Tr}\, \ell \,\mathrm{Tr}\, m - \mathrm{Tr}\, \ell m),$$

apparemment symétrique et bilinéaire. En effet, si λ et μ les valeurs propres de ℓ, on a $\det \ell = \lambda\mu$, $\mathrm{Tr}\, \ell = \lambda + \mu$, $\det \ell^2 = \lambda^2 + \mu^2$, donc

$$\lambda\mu = \det \ell = \frac{1}{2}(\mathrm{Tr}(\lambda+\mu)^2 - \lambda^2 - \mu^2) = \frac{1}{2}(\mathrm{Tr}\, \ell \,\mathrm{Tr}\, m - \mathrm{Tr}\, \ell m). \tag{12.5}$$

Ceci implique que $\{\sigma_0, \sigma_1, \sigma_2, \sigma_3\}$ est une base orthonormée de $\mathcal{M}$ de signature $(1,3)$. ∎

COROLLAIRE 12.13.2 *Soit $L \subset \mathcal{M}$ une droite de genre temps. Alors $L^\perp$ muni de la métrique $-(\ell, m)$ est un espace euclidien, et on a $\mathcal{M} = L \oplus L^\perp$.*

PREUVE. L'affirmation $\mathcal{M} = L \oplus L^\perp$ vient du fait que les droites de genre temps sont non-dégénérées. Puisque la signature de $\mathcal{M}$ est $(0,3)$, et celle de L est $(1,0$, il est $(0,3)$ sur $L^\perp$ ∎.

Passons à la signification géométrique des produits scalaires. Il y a des différences remarquables entre la métrique de Minkowski et la métrique euclidienne, avec une signification physique. La plus impréssionante est liée au fait, que l'inégalité de Cauchy-Schwarz pour les vecteurs de genre temps s'avère d'être inversée.

PROPOSITION 12.13.3 *Soit $\langle \ell_1, \ell_1 \rangle > 0$, $\langle \ell_2, \ell_2 \rangle > 0$, $\ell_i \in \mathcal{M}$. Alors*

$$\langle \ell_1, \ell_2 \rangle^2 \geq \langle \ell_1, \ell_1 \rangle \cdot \langle \ell_2, \ell_2 \rangle.$$

L'égalité est achevée si et seulement si ℓ_1 est ℓ_2 sont linéairement dépendants.

PREUVE. On vérifie tout d'abord que la fonction $\langle t\ell_1 + \ell_2, t\ell_1 + \ell_2 \rangle$ toujours possède une racine réelle t_0. Dans la réalisation matricielle $\mathcal{M}$ la condition $\langle \ell_2, \ell_2 \rangle > 0$ signifie que $\det \ell_2 > 0$, i.e. les valeurs caractéristiques de ℓ_2 sont de même signe, disons ε_2 (± 1). Alors pour $t \to -(\varepsilon_1\varepsilon_2)\infty$ les valeurs propres de la matrice $t\ell_1 + \ell_2$ sont approximativement proportionnelles de celles de ℓ_1. Donc le discriminant du polynôme $\langle t\ell_1 + \ell_2, t\ell_1 + \ell_2 \rangle$ est non-négatif, ce qui implique

$$\langle \ell_1, \ell_2 \rangle^2 \geq \langle \ell_1, \ell_1 \rangle \cdot \langle \ell_2, \ell_2 \rangle.$$

Corollaire 12.13.4 ("l'inégalité du triangle inversée") *Soit ℓ_1, ℓ_2 de genre temps et $\langle \ell_1, \ell_2 \rangle \geq 0$, alors $\ell_1 + \ell_2$ de genre temps et*

$$|\ell_1 + \ell_2| \geq |\ell_1| + |\ell_2|,$$

où $|\ell|^2 = \langle \ell, \ell \rangle$, et l'égalité est achevé si et seulement si ℓ_1 est ℓ_2 sont linéairement dépendants.

Paradoxe des jumeaux, voir [W]

On appelle deux vecteurs ℓ_1, ℓ_2 ayant la même orientation de genre temps, si $\langle \ell_1, \ell_2 \rangle > 0$. Imaginons deux référentiel jumeaux : l'un est inértiel et suit sa ligne d'univers à partir de 0 jusqu'à $\ell_1 + \ell_2$, tandis que l'autre arrive à ce point en deux étapes : d'abord par un mouvement inértiel de 0 jusqu'à ℓ_1, et puis de ℓ_1 jusqu'à $\ell_1 + \ell_2$. Il subit l'accélération lorsqu'il se trouve autour de 0 et autour de ℓ_1 pour pouvoir d'abord quitter son frère, et puis pour revenir à lui. On voit, par le corollaire, que le temps écoulé pour le voyageur est plus petit que pour le sédentaire.

> Parmi ces expériences de pensée, le paradoxe des jumeaux, énoncé en 1911 par Paul Langevin, a focalisé et focalise encore les pensées de ceux qui réfléchissent sur la relativité restreinte. On se trouve en fait devant deux aspects paraissant antinomiques (paradoxaux tous deux pour un physicien classique). D'un côté, hors des phases d'accélération (départ, retournement et freinage de celui qui s'en va), chacun des jumeaux est en mouvement relatif uniforme par rapport à l'autre : son horloge, vue du référentiel de l'autre est ralentie. D'un autre côté, lors de la rencontre finale, le *voyageur*, parti et revenu, est plus jeune que le *sédentaire*. Il existe bien sûr une asymétrie réelle entre les expériences vécues par chacun des deux jumeaux. Le jumeau voyageur subit une phase d'accélération, contrairement à son jumeau sédentaire : à ce titre, la découverte de la relativité générale, en 1915, amène une explication du vieillissement différent des jumeaux. Il faut cependant avoir conscience que ce n'est pas la phase d'accélération qui joue un rôle essentiel dans le ralentissement de l'horloge du jumeau voyageur, mais bien la dilatation des temps durant l'aller et le retour qu'il effectue dans deux référentiels inertiels différents.

Lé coefficient de Lorentz

Vu l'inégalité $\frac{\langle l_1, l_2 \rangle}{|l_1||l_2|} \geq 1$, on ne peut pas intérpréter cette quantité comme le cosinus d'un angle. L'explication est donnée par une intérprétation physique.

Soient $|\ell_1| = 1, |\ell_2| = 1$; en particuler, le référentiel inértiél ℓ_1 a vécu une unité de temps de son propre horloge à partir de debut de mouvement à l'origine. Au point ℓ_1, l'espace physique instantané pour lui est $\ell_1 + (\mathbb{R}\ell_1)^{\perp}$. La ligne univers pour $\mathbb{R}\ell_2$ coupe cet espace au point $x\ell_2$, où x se calcul par la condition

$$\langle x\ell_2 - \ell_1, \ell_1 \rangle = 0,$$

i.e. $x = \langle \ell_1, \ell_2 \rangle^{-1}$. La distance entre ℓ_1 à $x\ell_2$ est de genre ~~temps~~ : pour le référentiel $\mathbb{R}\ell_1$ est la distance, par laquelle $\mathbb{R}\ell_2$ s'éloigne de lui pour l'unité de temps, i.e. la vitesse relative de $\mathbb{R}\ell_2$. Elle est égale

$$\begin{aligned} v =& [-\langle x\ell_2 - \ell_1, x\ell_2 - \ell_1 \rangle]^{1/2} = [-\langle x\ell_2 - \ell_1, x\ell_2 \rangle]^{1/2} = \\ =& [-x^2\langle \ell_2, \ell_2 \rangle + x\langle \ell_1, \ell_2 \rangle]^{1/2} = [-\langle \ell_1, \ell_2 \rangle^{-2} + 1]^{1/2}, \end{aligned}$$

d'où

$$\langle \ell_1, \ell_2 \rangle = \frac{1}{\sqrt{1 - v^2}},$$

le célèbre coefficient de Lorentz écrit souvent sous la forme $\frac{1}{\sqrt{1 - v^2/c^2}}$, indiquant explicitement la mesurement de vitesses par rapport à celle de lumière. En particulier,

$$x = \sqrt{1 - v^2},$$

i.e. au moment de temps l'unité pour le premier référentiel, l'horloge du deuxième référentiel qui se trouve dans son espace physique instantané montre $\sqrt{1 - v^2}$ (la version quantitave du phénomène de dilatation des temps).

Quatre orientations de l'espace-temps

Soit $\{e_i\}$, $\{e'_i\}$, $(i = 0, \cdots, 1)$ deux bases orthonormées de $\mathcal{M}$:

$$\langle e_0, e_0 \rangle = \langle e'_0, e'_0 \rangle = 1, \ \langle e_i, e_i \rangle = \langle e'_i, e'_i \rangle = -1 \ (i = 1, 2, 3).$$

On les appelle ayant la même orientation, s'il existe un système continu d'isométries $f_t : \mathcal{M} \to \mathcal{M}$, $0 \leq t \leq 1$, tel que $f_0 = \mathrm{id}$, $f_1(e_i) = e'_i$.

Deux conditions sont nécessaires :

a) $\langle e_0, e'_0 \rangle > 0$. En effet, $\langle e_0, f_t(e_0) \rangle^2 \geq 1$ par la proposition 12.13.3, donc le signe de $\langle e_0, f_t(e_0) \rangle$ ne peut pas changer puisque $\langle e_0, f_0(e_0) \rangle = 1$; i.e. e_0, e'_0 ont la même orientation de genre temps.

b) Le déterminant de l'application de la projection orthogonale

$$\sum_{i=1}^{3} \mathbb{R}e_i \to \sum_{i=1}^{3} \mathbb{R}e'_i$$

est positif. On appelle deux bases avec cette propriété ayant la même orientation de genre espace.

Réciproquement, si deux bases possèdent la même orientation de de genre espace et de genre temps, elles ont la même orientation dans $\mathcal{M}$. Pour construire un système continu d'isométries $f_t : \mathcal{M} \to \mathcal{M}$, $0 \leq t \leq 1$, tel que $f_0 = \mathrm{id}$, $f_1(e_i) = e'_i$, on pose tout d'abord $f_t(e_0) = \frac{te'_0 + (1-t)e_0}{|te'_0 + (1-t)e_0|}$. Puis, on choisit comme $f_t(e_1)$, $f_t(e_2)$, $f_t(e_3)$ la base orthonormée de $f_t(e_0)^\perp$ par l'orthogonalisation de Gram-Schmidt.

On va noter $\Lambda = O(1, 3)$ le groupe de Lorentz, i.e. le groupe des isométries de l'espace $\mathcal{M}$. Soit

$\Lambda_+^\uparrow$ le sous-groupe de Λ respectant l'orientation d'une base orthonormée ;

$\Lambda_-^\uparrow$ la partie d'éléments du groupe Λ respectant l'orientation de genre temps, et inversant l'orientation de genre espace ;

$\Lambda_+^\downarrow$ la partie d'éléments du groupe Λ respectant l'orientation de genre espace, et inversant l'orientation de genre temps ;

$\Lambda_-^\downarrow$ la partie d'éléments du groupe Λ inversnt l'orientation de genre temps, et inversant l'orientation de genre espace.

THÉORÈME 12.13.5 *Le groupe de Lorentz est constitué de quatre composantes connexes*

$$\Lambda = \Lambda_+^\uparrow \cup \Lambda_-^\uparrow \cup \Lambda_+^\downarrow \cup \Lambda_-^\downarrow.$$

THÉORÈME 12.13.6 *Dans la réalisation de $\mathcal{M}$ comme l'espace de matrices de Gram des métriques hermitiens de $\mathcal{H}$, dans une base $\{h_1, h_2\}$, on fait correspondre à une matrice $V \in SL(2,\mathbb{C})$ la transformation*

$$s(V)\ell = V^t \ell \overline{V}.$$

L'application s donne un homomorphisme surjectif de $SL(2,\mathbb{C})$ sur $\Lambda_+^\uparrow$ de noyau $\pm I_2$.

12.14 Rotations euclidiennes et boosts*

Soient e_0 et e_0' deux vecteurs de genre temps ayant la même orientation de genre temps. On considère les complémentaires orthogonaux L_0 et L_0' correspondants. Il existe une transformation standard de Lorentz de $\Lambda_+^\uparrow$, qui transforme e_0 dans e_0', et qui s'appelle "boost" (en anglais). Pour $e_0 = e_0'$, cela est id. Pour $e_0 \neq e_0'$ on la définit de la manière suivante : on considère le plan $(L_0 \cap L_0')^\perp$, contenant e_0 et e_0'. La signature de la métrique de Minkowski est égale $(1,1)$ sur ce plan. Donc il existe un couple de vecteurs de genre espace $e_1, e_1' \in (L_0 \cap L_0')^\perp$, orthogonaux à e_0 et e_0' respectivement. Pour calculer la matrice de passage

$$\{e_0', e_1'\} = \{e_0, e_1\} \begin{pmatrix} a & c \\ b & d \end{pmatrix},$$

on note que $a = \langle e_0, e_0' \rangle = \frac{1}{\sqrt{1-v^2}}$, où v est la vitesse relative entre les référentiels e_0, e_0'. Puis, les matrices de Gram de $\{e_0', e_1'\}$ et de $\{e_0, e_1\}$ sont $\left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$, donc

$$a^2 - b^2 = 1,\ ac - bd = 0,\ c^2 - d^2 = -1.$$

Sachant a on trouve $b = \frac{v}{\sqrt{1-v^2}}$. De plus, $\det \left(\begin{smallmatrix} a & c \\ b & d \end{smallmatrix}\right) = 1$ donne $d = a$, $c = b$. La matrice de boost est donc

$$\begin{pmatrix} \frac{1}{\sqrt{1-v^2}} & \frac{v}{\sqrt{1-v^2}} & 0 & 0 \\ \frac{v}{\sqrt{1-v^2}} & \frac{1}{\sqrt{1-v^2}} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

avec une "rotation hyperbolique" $\left(\begin{smallmatrix} \cosh\theta & \sinh\theta \\ \sinh\theta & \cosh\theta \end{smallmatrix}\right)$, où

$$\cosh\theta = \frac{1}{\sqrt{1-v^2}}, \sinh\theta = \frac{v}{\sqrt{1-v^2}}.$$

13 Algèbre géométrique

(voir [Dieudonné], [Garrett])

- $GL(n)$ (étude géométrique).
- Formes bilinéaires et hermitiennes, groupes classiques.
- Théorème de Witt et l'extension d'isométries

On va étudier en détail les notions de l'algèbre géométrique.

Concernant la notation matricielle pour une matrice rectangulaire $A = (a_{ij})$ soit ${}^tA$ la transposée de A. Si les entrées de A se trouvent dans un anneau D muni d'involution σ, soit A^σ donnée par $(A^\sigma)_{ij} = a_{ij}^\sigma$.

13.1 Étude géométrique du groupe $GL(n)$ et de ses sous-groupes

Le groupe $GL(n)$ est le groupe classique le plus facilement étudié, mais il indique déjà les phénomènes les plus intéressants pour l'utilisation dans beaucoup d'autres situations. Le groupe général linéaire $GL(n,k)$ est le groupe de toutes les matrices inversibles de la taille $n \times n$ avec les entrés dans un corps commutatif k. Le groupe spécial linéaire $SL(n,k)$ est le groupe de toutes les matrices (inversibles) de la taille $n \times n$ avec les entrées dans un corps commutatif k, et de déterminant 1.

Pour une approche moins dépendante de coordonnées, on fixe un k-espace vectoriel V de dimension n et soit $GL_k(V)$ le groupe de tous les automorphismes k-lineaires de V. Tout choix d'une base de V sur k-basis donne un isomorphisme $GL_k(V) \to GL(n,k)$ en utilisant la matrice de l'application linéaire par rapport aux bases choisies. Soit $e_1, \cdots, e_n$ la base standard pour k^n

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ \cdots \\ 0 \end{pmatrix}, e_2 = \begin{pmatrix} 0 \\ 1 \\ \cdots \\ 0 \end{pmatrix}, \cdots, e_n = \begin{pmatrix} 0 \\ 0 \\ \cdots \\ 1 \end{pmatrix}.$$

Par ce choix de bases ordonnées on obtient un isomorphisme $GL_k(k^n) \to GL(n,k)$.

Un *drapeau* $\mathcal{F}$ dans V est une chaîne

$$\mathcal{F} = (V_{d_1} \subset V_{d_2} \subset \cdots \subset V_{d_m})$$

de sous-espaces, où V_{d_i} est de dimension d_i, et

$$d_1 < \cdots < d_m.$$

On dit que le type de ce drapeau est $(d_1, \cdots, d_m)$. Dans k^n le drapeau standard de type $(d_1, \cdots, d_m)$ est le drapeau de type $(d_1, \cdots, d_m)$ avec

$$V_{d_i} = ke_1 + \cdots + ke_{d_{i-1}} + ke_{d_i}.$$

On définit le *sous-groupe parabolique* $P_{\mathcal{F}}$ associé au drapeau $\mathcal{F}$ par :

$$P = \{g \in GL_k(V) : \forall i, gV_{d_i} = V_{d_i}\}$$

Si $V = k^n$ et $\mathcal{F}$ le *drapeau standard* de type $(d_1, \cdots, d_m)$, le sous-groupe parabolique $P_{\mathcal{F}}$ est formé par tous les éléments admettant un développement en blocs :

$$\begin{pmatrix} d_1 \times d_1 & & & * \\ & (d_2 - d_1) \times (d_2 - d_1) & & \\ & & \ddots & \\ 0 & & & (n - d_m) \times (n - d_m) \end{pmatrix}$$

où le $i^{\text{ème}}$ entrée diagonale est (comme indiqué) de type $(d_i - d_{i-1}) \times (d_i - d_{i-1})$, les entrées inférieures sont 0, les entrées supérieures sont arbitraires. Tout $g \in P = P_{\mathcal{F}}$ induit une application naturelle sur les quotients $V_{d_i}/V_{d_{i-1}}$, où on définit $V_{d_0} = 0$ et $V_{d_{m+1}} = V$). Alors le *radical unipotent* $R_u P$ est

$$R_u P = \{p \in P_{\mathcal{F}} : p = id \text{ sur tous } V_{d_i}/V_{d_{i-1}} \text{ et sur } V/V_{d_m}\}.$$

Le radical unipotent $R_u P$ est un sous groupe distingué de P. Dans le cas du sous-groupe parabolique standard P de type $(d_1, \cdots, d_m)$ sur k^n le radical unipotent est formé par les éléments de la forme

$$\begin{pmatrix} 1_{d_1} & * & \cdots & & \\ & 1_{d_2 - d_1} & * & \cdots & \\ & & \ddots & * & \cdots \\ & & & \ddots & * \\ 0 & & & & 1_{n-d_m} \end{pmatrix}$$

où 1_d désigne la matrice identité de type $d \times d$. Choisissons les sous-espaces V'_{n-d_i} de V de telle façon que V'_{n-d_i} est un sous-espace complémentaire de V_{d_i} dans V et tel que

$$V'_{n-d_m} \subset \cdots \subset V'_{n-d_1}$$

est un drapeau de type opposé au drapeau de V_{d_i}. On pose

$$P' = \{g \in GL_k(V) : \forall i, g V'_{n-d_i} = V'_{n-d_i}\}$$

$$M = P \cap P'$$

Alors M est dit une composante de Levi complementaire de P, et $P = P_{\mathcal{F}}$ est le produit semi-direct standard

$$P = M \ltimes R_u P$$

de M et $R_u P$, où M normalise $R_u P$. Pour le sous-groupe standard parabolique P dans $GL(n, k)$ de type $(d_1, \cdots, d_m)$ le choix standard du sous-espace complementaire est

$$V'_{n-d_i} = k e_{d_i+1} + \cdots + k e_n$$

Alors la composante de Levi standard est le groupe des matrices de la forme

$$\begin{pmatrix} d_1 \times d_1 & & & & \\ & d_2 - d_1 \times d_2 - d_1 & & & \\ & & \ddots & & \\ & & & \ddots & \\ 0 & & & & n - d_m \times n - d_m \end{pmatrix}$$

où le $i^{\text{ème}}$ élément diagonale est de type $(d_i - d_{i-1}) \times (d_i - d_{i-1})$, $i = 1, \ldots, m+1$, et tous les autres blocs sont nuls. Dans le cas du groupe $GL_k(V)$ la composante de Levi d'un sous-groupe minimal parabolique est dite un torus maximal (k-)déployé.

Le résultat suivant est un prototype du résultat analogue pour les classes plus larges de groupes.

PROPOSITION 13.1.1
a) Tous les sous-groupes paraboliques de type donné sont conjugués dans $GL_k(V)$
b) Toutes les composantes de Levi dans un sous-groupe parabolique P sont conjugués par éléments de P
c) Tous les tores maximaux k-déployés sont conjugués dans $GL_k(V)$

Pour montrer que tous les sous-groupes paraboliques de type donné sont conjugués, il suffit de voir que pour tout choix de deux drapeaux

$$V_{d_1} \subset V_{d_2} \subset \cdots \subset V_{d_m}$$
$$V'_{d_1} \subset V'_{d_2} \subset \cdots \subset V'_{d_m}$$

de même type il existe un $g \in GL_k(V)$ tel que $gV_{d_i} = V'_{d_i}$ pour tous i. On choisit deux bases $\{v_i\}$, $\{v'_i\}$ de V, de telle façon que

$$V_{d_i} = kv_1 + \cdots + kv_{d_i},$$
$$V'_{d_i} = kv'_1 + \cdots + kv'_{d_i}.$$

Alors on définit g par $gv_i = v'_i$. Ceci prouve que tous les sous-groupes paraboliques de type donné sont conjugués.

Pour démontrer que toutes les composantes de Levi dans un sous-groupe parabolique P sont conjugués par des éléments de P soit

$$V_{d_1} \subset \cdots \subset V_{d_m}$$

le drapeau pour lequel P est le stabilisateur, et soient $V^1_{n-d_i}, V^2_{n-d_i}$, (avec $1 \leq i \leq m$) deux choix de familles de sous-espaces complémentaires qui définissent les composantes de Levi correspondantes. Il suffit de trouver $p \in P$ de telle façon que $pV^1_{n-d_i} = V^2_{n-d_i}$ (pour tous les indices i).

Pour $\ell = 1, 2$, on définit $W^\ell_1, \ldots, W^\ell_{m+1}$ comme, respectivement,

$$V_{d_1}, V_{d_2} \cap V^\ell_{n-d_1}, V_{d_3} \cap V^\ell_{n-d_2}, \cdots, V_{d_m} \cap V^\ell_{n-d_{m-1}}, V^\ell_{n-d_m}.$$

Pour $\ell = 1, 2$ nous avons $V = \oplus W^\ell_i$. Par l'hypothèse, $\dim_k W^\ell_1 = \dim_k W^\ell_2$ pour tous j. Alors, il existe une infinité d'éléments $g \in GL_k(V)$ tels que $gW^1_j = W^2_j$ pour tous j. Pour chaque tel g, certainment $g \in P$, et puisque V^ℓ_j est une somme de W^ℓ_j, on a assurément $pV^1_{n-d_i} = V^2_{n-d_i}$ pour un $p \in P$ et pour ~~pour~~ tout i.

Soient T_1, T_2 deux tores maximaux k-déployés, on choisit des sous-groupes paraboliques minimaux P_i contenant T_i. Par la première partie de la proposition, il existe

$h \in GL_k(V)$ tel que $hP_1h^{-1} = P_2$. Alors hT_1h^{-1} est une autre composante de Levi (un torus maximal déployé) à l'intérieur de P_2, donc par la seconde assertion de la proposition il existe $p \in P_2$ tel que $p(hP_1h^{-1})p^{-1} = T_2$. Ceci donne la troisième assertion de la proposition. ∎

Maintenant on va généraliser le précédant directement en remplaçant le corps commutatif k par un corps gauche (un anneau de division) D. On reprend la version sans coordonnées de la discussion précédente ; les illustrations matricielles resterons les mêmes.

On définit un espace vectoriel V de dimension finie sur un corps gauche (anneau de division) D (c'est-à-dire, V est un module de génération finie sur D). La notion de dimension a un sens, étant définie comme le rang d'un module libre. Les résultats élémentaires sur l'indépendance linéaire et sur les bases sont les mêmes que sur les corps commutatifs.

La perte de la commutativité de D devient importante lorsque l'on considère les endomorphismes D-linéaires. Si D n'est pas commutatif, alors l'anneau $\mathrm{End}_D(V)$ de tous les endomorphismes D-linéaires de V ne contient pas D d'une façon naturelle. Alors, un choix de D-bases pour un espace vectoriel D de dimension $n$donne un isomorphisme

$$\mathrm{End}_D(V) \to \{n \times n \text{ matrices à coefficients dans } D^{opp}\}$$

où D^{opp} est l'anneau opposé à D. Ceci dit, D^{opp} est le même groupe additif D, mais avec la multiplication $*$, donnée par

$$x * y = yx$$

où yx est la multiplication dans D. (Parfois on peut éviter cette complication (innocente) en décrivant V comme étant un D-module "droit", mais de toute façon la définition d'un module "droit" est réellement celle d'un module sur l'anneau opposé D^{opp}.) Le groupe général linéaire $GL(n, D)$ sur D est le groupe de toutes les matrices inversibles de $n \times n$ à coefficients dans D. Une version sans coordonnées du groupe général linéaire est $GL_D(V)$, le groupe de tous les automorphismes D-linéaires de V. Un choix de D-bases pour V donne un isomorphisme

$$GL_D(V) \to GL(n, D^{opp})$$

Les définitions concernant les drapeaux et les sous-groupes paraboliques sont identiques à celles du cas où D est commutatif. Un drapeau $\mathcal{F}$ dans V est une chaîne

$$\mathcal{F} = (V_{d_1} \subset V_{d_2} \subset \cdots \subset V_{d_m})$$

de sous-espace , où V_i est de dimension i et

$$d_1 < \cdots < d_m.$$

Le type d'un drapeau est la suite $(d_1, \cdots, d_m)$.

Un sous-groupe parabolique $P = P_{\mathcal{F}}$ in $GL_D(V)$ est le stabilisateur d'un drapeau

$$\mathcal{F} = (V_{d_1} \subset V_{d_2} \subset \cdots \subset V_{d_m})$$

C'est-à-dire,

$$P_{\mathcal{F}} = \{g \in GL_D(V) : \forall i, gV_{d_i} = V_{d_i}\}$$

Tout $g \in P = P_{\mathcal{F}}$ induit une application naturelle sur les quotients $V_{d_i}/V_{d_{i-1}}$ (où on définit $V_{d_0} = 0$ et $V_{d_{m+1}} = V$). Le radical unipotent $R_u P$ est

$$R_u P = \{p \in P_{\mathcal{F}} : p = id \text{ sur } V_{d_i}/V_{d_{i-1}}\}$$

Le *radical unipotent* $R_u P$ est un sous-groupe distingué de P. En choisissant les sous-espaces V'_{n-d_i} de V de telle façon que V'_{n-d_i} est un sous-espace complémentaire de V_{d_i} dans V. Alors

$$V'_{n-d_i} \subset \cdots \subset V'_{n-d_1}$$

est un drapeau de type opposé de celui de V_{d_i}.

On pose

$$P' = \{g \in GL_D(V) : \forall i, gV'_{n-d_i} = V'_{n-d_i}\}$$

$$M = P \cap P'$$

Alors M est appelée une *composante de Levi* ou *un complémentaire de Levi* dans P, et $P = P_{\mathcal{F}}$ est le produit semi-direct

$$P = M \ltimes R_u P$$

de M et de $R_u P$, où M normalise $R_u P$.

PROPOSITION 13.1.2 *a) Tous les sous-groupes paraboliques de type donné sont conjugués dans* $GL_D(V)$

b) Toutes les composantes de Levi dans un sous-groupe parabolique P *sont conjugués par éléments de* P

Cours N 10 Lundi 23 mars 2015

13.2 Formes bilinéaires et formes hermitiennes, groupes classiques.

Dans cette section, on considère les groupes classiques définis comme des isométries ou des similitudes de « formes » sur les espaces vectoriels. On définit premièrement les groupes orthogonaux et les groupes symplectiques. Cette famille de descriptions peut être simplifiée, au prix d'une obscuration des membres les plus simples.

13.2.1 Formes bilinéaires, formes symétriques

Soit k un corps de caractéristique différente de 2 et soit V un k-espace vectoriel de dimension finie. Une forme (k-)*bilinéaire* sur V est une fonction à valeurs dans k sur $V \times V$ donc pour tous $x, y \in k$ et $v, v_1, v_2 \in V$

$$\begin{aligned}\langle v_1 + v_2, v\rangle &= \langle v_1, v\rangle + \langle v_2, v\rangle\\ \langle v, v_1 + v_2\rangle &= \langle v, v_1\rangle + \langle v, v_2\rangle\\ \langle xv, yv_1\rangle &= xy\langle v, v_1\rangle.\end{aligned}$$

Si on a toujours

$$\langle v_1, v_2\rangle = \langle v_2, v_1\rangle,$$

alors la forme bilinéaire est dite symétrique. La fonction

$$Q[v] = \langle v, v\rangle$$

est dite la forme quadratique associée, à partir de laquelle $\langle\cdot,\cdot\rangle$ peut être récupérée par

$$4\langle v_1, v_2\rangle = Q[v_1 + v_2] - Q[v_1 - v_2]$$

Le groupe orthogonal associé est le groupe d'isométries de Q (ou de $\langle\cdot,\cdot\rangle$), étant défini comme

$$O(Q) = O(\langle\cdot,\cdot\rangle) = \{g \in GL_k(V) : \forall v_1, v_2 \in V,\ \langle gv_1, gv_2\rangle = \langle v_1, v_2\rangle\}$$

Le groupe de similitudes associé est défini comme

$$\begin{aligned}GO(Q) = GO(\langle\cdot,\cdot\rangle) =&\{g \in GL_k(V) :\ \exists \nu(g) \in k^\times \text{ tel que}\\ &\forall v_1, v_2 \in V, \langle gv_1, gv_2\rangle = \nu(g)\langle v_1, v_2\rangle\}.\end{aligned}$$

Si on a pour tout
$\langle v_1, v_2\rangle = -\langle v_2, v_1\rangle$ pour tout $v_1, v_2 \in V$, alors la forme bilinéaire
$f : V \times V \to k$, $f(v_1, v_2) = \langle v_1, v_2\rangle$ est dite alternée ou symplectique.

Le groupe symplectique associé à f est le groupe d'isométrie de la forme $f = \langle,\rangle$ défini comme

$$Sp(f) = \{g \in GL_k(V) : \forall v_1, v_2 \in V,\ \langle gv_1, gv_2\rangle = \langle v_1, v_2\rangle\}$$

Le groupe de similitudes associé est défini comme

$$\begin{aligned}GSp(f) = \{&g \in GL_k(V) : \exists \nu(g) \in k^\times\\ &\text{tel que } \forall v_1, v_2 \in V,\ \langle gv_1, gv_2\rangle = \nu(g)\langle v_1, v_2\rangle\}\end{aligned}$$

13.2.2 Formes sesquilinéaires, formes hermitiens

Soit K un corps, une extension quadratique de k, son sous-corps trivial par l'action d'un automorphisme k-linéaire σ.

DÉFINITION 13.2.1
a) Une forme k-bilinéaire $f : V \times V \to K$, $f(v_1, v_2) = \langle v_1, v_2 \rangle$ sur un K-espace vectoriel de dimension finie V est dite sesquilinéaire *(avec une référence implicite à σ) si*

$$\langle xv_1, yv_2 \rangle = xy^\sigma \langle v_1, v_2 \rangle$$

(pour tous $x, y \in K$ et pour tous $v_1, v_2 \in V$).
b) Une forme sesquilinéaire $f = \langle , \rangle$ sur un K-espace vectoriel de dimension finie V est dite hermitienne, *si pour tous $v_1, v_2 \in V$ on a*

$$\langle v_2, v_1 \rangle = \langle v_1, v_2 \rangle^\sigma$$

Le groupe unitaire associé est le groupe d'isométries de $\langle , \rangle$, étant défini comme

$$U(f) = \{g \in GL_K(V) : \langle gv_1, gv_2 \rangle = \langle v_1, v_2 \rangle, \ \forall v_1, v_2 \in V\}$$

Le groupe de similitudes associé est défini comme

$$GU(f) = \{g \in GL_K(V) : \exists \nu(g) \in k^\times \text{ tel que } \forall v_1, v_2 \in V, \ \langle gv_1, gv_2 \rangle = \nu(g)\langle v_1, v_2 \rangle\}.$$

On peut traîter les groupes précédents simultanement, en incluant aussi des groupes plus généraux de la façon suivante : soit D est une algèbre de division avec une (anti-)involution σ. Notons que, $D \to D$ est telle que les propriétés

$$(\alpha)^{\sigma\sigma} = \alpha \text{ et } (\alpha + \beta)^\sigma = \alpha^\sigma + \beta^\sigma \text{ et } (\alpha\beta)^\sigma = \beta^\sigma \alpha^\sigma$$

pour tous $\alpha, \beta \in D$. Soit Z le centre de D. On suppose que D soit de dimension finie sur Z, et que

$$k = \{x \in Z \mid x^\sigma = x\}.$$

Soit V un D-espace vectoriel de dimension finie, et fixons $\varepsilon = \pm 1$. Soit

$$f = \langle \cdot, \cdot \rangle, f : V \times V \to D$$

une forme k-bilinéaire à valeurs dans D sur V de telle façon que

$$\langle v_2, v_1 \rangle = \varepsilon \langle v_1, v_2 \rangle^\sigma$$
$$\langle \alpha v_1, \beta v_2 \rangle = \alpha^\sigma \langle v_1, v_2 \rangle \beta$$

pour tous $\alpha, \beta \in D$ et $v_1, v_2 \in V$.

Une telle forme est dite ε-hermitienne sur V. On appelle un tel espace V (muni de $\langle , \rangle$) un (D, σ, ε)-espace.

13.2.3 Groupes d'isométries généraux

Soit V_i deux (D,σ,ε)-espaces muni de formes $f_i = \langle , \rangle_i$ pour $(i = 1, 2)$. Une application D-linéaire $\varphi : V_1 \to V_2$ est une isométrie si, pour tous $u, v \in V_1$,

$$\langle \varphi u, \varphi v\rangle_2 = \langle u, v\rangle_1$$

L'applicaton φ est une similitude s'il existe $\nu \in k^\times$ tel que, pour tous $u, v \in V_1$,

$$\langle \varphi u, \varphi v\rangle_2 = \nu\langle u, v\rangle_1$$

Écrivons $\varphi : V_1 \cong V_2$ si φ est une isométrie.

DÉFINITION 13.2.2
a) Le groupe d'isométries *associé à* $f = \langle , \rangle$ *sur* $V = V_1 = V_2$ *est défini comme*

$$U(f) = \{g \in GL_D(V) : \forall v_1, v_2 \in V, \langle gv_1, gv_2\rangle = \langle v_1, v_2\rangle, \ \}.$$

b) Le groupe de similitudes *associé est défini comme*

$$GU(f) = \{g \in GL_D(V) : \exists \nu(g) \in k^\times \text{ tel que } \forall v_1, v_2 \in V, \ \langle gv_1, gv_2\rangle = \nu(g)\langle v_1, v_2\rangle\}.$$

13.2.4 Orthogonalisation, vecteurs isotropes*

Un D-sous-espace U dans un (D,σ,ε)-espace V possède un complémentaire orthogonal

$$U^\perp = \{u' \in V \ \ \langle u', u\rangle = 0, \forall u \in U\}$$

Noter que $U \cap U^\perp = 0$ n'est pas assuré en général. Le noyau de tout l'espace V est noté par $V^\perp$. La forme est dite non dégénérée si $V^\perp = 0$. Souvent on va supposer (sans référence à la forme) que simplement l'espace V est non dégénéré.

Si V_1, V_2 sont deux (D,σ,ε)-espaces avec les formes, respectivement, $\langle , \rangle_1, \langle , \rangle_2$, alors la somme directe $V_1 \oplus V_2$ de D-espaces vectoriels est un (D,σ,ε)-espace muni de la forme

$$\langle v_1 + v_2, v_1' + v_2'\rangle = \langle v_1, v_1'\rangle_1 \langle v_2, v_2'\rangle_2$$

On l'apppelle la somme orthogonale. En général, deux sous-espaces V_1, V_2 d'un (D,σ,ε)-espace sont orthogonaux si

$$V_1 \subset V_2^\perp,$$

ou, de façon équivalente, si $V_2 \subset V_1^\perp$.

Si $\langle v, v\rangle = 0$ pour $v \in V$, alors v est un vecteur isotrope. Si$\langle v, v'\rangle = 0$ pour tout $v, v' \in U$ pour un sous-espace U de V alors U est un sous-espace (totalement) isotrope. S'il n'existe pas de vecteur isotrope non nul dans U, alors on dit que U est anisotrope.

PROPOSITION 13.2.3 *Soit V un (D,σ,ε)-espace non-dégénéré avec un sous-espace U. Alors U est non-dégénéré si et seulement si $V = U \oplus U^\perp$, et si et seulement si $U^\perp$ est non-dégénéré*

Preuve. On utilise l'application $\Lambda : V \to \mathrm{Hom}_D(U, D)$ donnée par $v \to \lambda_v$ où

$$\lambda_v(u) = \langle u, v \rangle.$$

Il découle de la non-dégénéréscance de V que Λ est surjective. Le noyau est donc $U^\perp$ Alors, par l'algèbre linéaire on a

$$\dim_D U^\perp + \dim_{D^{opp}} \Lambda(U) = \dim_D V$$

Donc, comme la dimension de $\Lambda(V)$ est la même que la dimension de U, par décompte de dimensions, on a $U \cap U^\perp = 0$ si et seulement si $U + U^\perp$ est une somme directe (et donc une somme orthogonale). Puisque $U \subset U^{\perp\perp}$, le fait que U est dégénéré implique que $U \cap U^\perp$ est non nul. Alors $U^\perp \cap U^{\perp\perp}$ est non nul puisqu'il contient $U \cap U^\perp$ donc $U^\perp$ est dégénéré. De l'autre coté, U est non-dégénéré implique que $U + U^\perp$ est une somme directe, donc $\dim U = \dim V - \dim U^\perp$.

Puisque $\dim U^{\perp\perp} = \dim V - \dim U^\perp$ on déduit de la non-dégénéréscance de V, que $U^{\perp\perp} = U$, donc $U^{\perp\perp} + U^\perp$ est une somme directe, et $U^\perp$ est non-dégénéré. ∎

Une D-base $e_1, \ldots, e_n$ pour un (D, σ, ε)-espace V est dite orthogonale si $\langle e_i, e_j \rangle$ pour $i \neq j$.

PROPOSITION 13.2.4 *Soit V un (D, σ, ε)-espace. Supposons que le cas où $\varepsilon = -1$, $D = k$, et σ et trivial est exclu. Si le produit $\langle , \rangle$ n'est pas identiquement nul, alors il existe $v \in V$ avec $\langle v, v \rangle \neq 0$. Si V est non-dégénéré, alors il possède une base orthogonale.*

Preuve. Supposons que $\langle v, v \rangle = 0$ pour tous $v \in V$. Alors

$$0 = \langle x + y, x + y \rangle = \langle x, x \rangle + \langle x, y \rangle + \varepsilon \langle x, y \rangle^\sigma + \langle y, y \rangle = \langle x, y \rangle + \varepsilon \langle x, y \rangle^\sigma$$

Si $\varepsilon = 1$ et le produit $\langle , \rangle$ n'est pas identiquement nul, il existe x, y tels que $\langle x, y \rangle = 1$.

Alors on a

$$0 = \langle x, y \rangle + \varepsilon \langle x, y \rangle^\sigma = 1 + 1,$$

Contradiction. Supposons que $\varepsilon = -1$ et σ n'est pas triviale sur D. Alors il existe $\alpha \in D$ tel que $\alpha^\sigma \neq \alpha$, et avec $\omega = \alpha - \alpha^\sigma$, $\omega^\sigma = -\omega$. Si $\langle , \rangle$ n'est pas identiquement nul alors il existe x, y tels que $\langle x, y \rangle = 1$. Alors on a

$$0 = \langle \omega x, y \rangle + \varepsilon \langle \omega x, y \rangle^\sigma = \omega^\sigma \langle x, y \rangle - \langle x, y \rangle^\sigma \omega =$$
$$-\omega + \varepsilon\omega = -2\omega,$$

Contradiction.

Pour construire une base orthogonale, on utilise la récurrence par dimension. Si la dimension d'un espace non-dégénéré V est 1, alors tout vecteur non nul forme une base orthogonale. En général, par la discussion précédente, on peut trouver $v \in V$ de telle façon que $\langle v, v \rangle \neq 0$. Alors $Dv^\perp$ est non–dégénéré et V est une somme directe orthogonale de Dv et de $Dv^\perp$, par la proposition précédente. ∎

Supposons que V est de dimension deux, avec une base ordonnée x, y de telle façon que

$$\langle x, x \rangle = \langle y, y \rangle = 0 \text{ et } \langle x, y \rangle = 1$$

Alors V est un plan hyperbolique et x, y est un paire hyperbolique dans V. Un (D, σ, ε)-espace est hyperbolique, s'il est une somme orthogonale de plans hyperboliques.

PROPOSITION 13.2.5 *Soit V et W deux espaces hyperboliques de même dimension (avec les même données (D, σ, ε). Alors il existe une isométrie $f : V \to W$. Ceci dit, la dimension est un seul invariant d'un espace hyperbolique.*

Preuve : Comparer les paires hyperboliques. ∎

PROPOSITION 13.2.6 *On considère un espace V non-dégénéré avec $\varepsilon = -1, D = k$, et soit σ trivial. Alors V est* hyperbolique, *ceci dit, V est une somme orthogonale de plans hyperboliques.*

Preuve. Puisque σ est triviale, $\alpha\beta = \beta\alpha$ pour tous $\alpha, \beta \in D$, alors D est un corps. Puisque

$$\langle x, x\rangle = -\langle x, x\rangle$$

et la caractéristique est différente de 2, tout vecteur est isotrope. Fixons $x \in V$ non-nul, et considérons $y \in V$ tels que $\langle x, y\rangle \neq 0$. Alors, quitte à remplacer y par un élément de D on peut supposer $\langle x, y\rangle = 1$, c'est-à-dire, un paire hyperbolique. Alors $Dx + Dy$ et $(Dx + Dy)^{\perp}$ sont non-dégénérés, et on fait la récurrence sur la dimension. ∎

PROPOSITION 13.2.7 *Soit V est un espace non-dégénéré, et soit $-V$ le même espace muni de la forme opposée sur V. Alors la somme orthogonale*

$$W = V \oplus -V$$

est hyperbolique.

Preuve. Dans le cas d'espaces alternés non-dégénérés (avec $D = k, \varepsilon = 1, \sigma$ triviale), l'espace V lui-même est déjà hyperbolique, donc $-V$ l'est. De l'autre coté, pour un espace non-alterné non-dégénéré V, on peut choisir une base orthogonale $\{e_i\}$ (pour les deux V et $-V$. Alors on affirme que dans $V \oplus -V$ les sous-espaces

$$H_i = De_i \oplus De_i$$

sont des plans hyperboliques, pour tous les indices i. (Ceci donnerait la preuve de la proposition). Puisque la caractéristique est différente de 2, on peut considérer les vecteurs

$$x_i = \frac{1}{2}e_i \oplus e_i, y_i = \langle e_i, e_i\rangle^{-1}e_i \oplus -e_i$$

qui sont linéairement independents (puisque $1 \neq -1$). Ils constitutent deux isotropes par construction. Puis, les constantes sont telles que pour la forme $\langle , \rangle$ sur $V \oplus -V$ on a $\langle x_i, y_i\rangle = 1$. ∎

PROPOSITION 13.2.8 *Soit V un espace non-dégénéré, et soit W un sous-espace. Soit W_0 le noyau de W. Alors il existe un sous-espace non-dégénéré W_1 of W tel que $W_0 + W_1 = W$ est une somme directe. De plus,pour toute base $x_1, \cdots, x_n$ de W_0, et pour tout tel W_1 il existe une partie $\{y_i\} \subset W_1^{\perp}$, telle que les sous-espaces $Dx_i + Dy_i$ sont des plans hyperboliques mutuellement orthogonaux. En particulier,*

$$W + \sum_i Dy_i = W \bigoplus \oplus_i Dy_i$$

est non-dégénéré et $W_0 + \sum_i Dy_i$ est un espace hyperbolique.

COROLLAIRE 13.2.9 *Soit V un espace non-dégénéré. Alors il existe un sous-espace hyperbolique H de V et un sous-espace anisotrope A de V tel que V est la somme directe orthogonale $V = H \oplus A$.*

13.2.5 Classification d'espaces orthogonaux et hermitiens.

On rapelle d'abord une classification d'espaces orthogonaux et hermitiens à isométrie près ((voir [Lang], Ch.XIV, §7 , §11).

On donnera une classification plus fine d'espaces quadratiques et une généralisation de la loi d'inértie et de la notion de la signature (voir §1 du Chapitre IV de [Se70] et §3 de Partie II de [KosMan]).

13.2.6 Formes symétriques sur les corps ordonnés

THÉORÈME 13.2.10 (LOI D'INERTIE DE SYLVESTER*) *Soit k un corps ordonné, et E un k-espace vectoriel muni d'une forme symetrique non-gégénérée. Alors il existe un entier $r \geq 0$ tel que pour toute base orthogonale $\{v_1, \cdots, v_n\}$ de E, il y a exactement r éléments $\langle v_i, v_i\rangle$ strictement positifs, et exactement $n - r$ éléments $\langle v_i, v_i\rangle$ strictement négatifs.*

Preuve. On pose $a_i = \langle v_i, v_i\rangle$ et $b_i = \langle w_i, w_i\rangle$ pour toute autre base orthogonale. On raisonne par l'absurde. Si $b_i = \langle w_i, w_i\rangle$ sont strictement positifs pour exactement s éléments avec $i = 1, 2, \cdots, s$, il suffit de montrer que

$$v_1, \cdots, v_r, w_{s+1}, \cdots, w_n$$

sont linéairement indépendents, alors $r + n - s \leq n$, donc $r \leq s$ et $r = s$ par symétrie. Si

$$u = x_1 v_1 + \cdots + x_r v_r = -y_{s+1} w_{s+1} - \cdots - y_n w_n,$$

$$\langle u, u\rangle = x_1^2 \langle v_1, v_1\rangle + \cdots + x_r^2 \langle v_r, v_r\rangle = y_{s+1}^2 \langle w_{s+1}, w_{s+1}\rangle + \cdots + y_n^2 \langle w_n, w_n\rangle,$$

d'où la contradiction : on a $\langle u, u\rangle > 0$ et $\langle u, u\rangle < 0$ au même temps. ∎

DÉFINITION 13.2.11

a) La signature *de g est le couple $(r_+, r_-) = (r, rk(g) - r)$ qui défini le nombre $\sigma(g) = r_+ - r_-$. Pour toute forme hyperbolique h, $\sigma(h) = 0$ et $\sigma(g \oplus h) = \sigma(g)$.*

b) Une forme positive définie *est de signature $(r_+, r_-) = (n, 0)$.*

COROLLAIRE 13.2.12 (CRITÈRE DE SYLVESTER) *Soit k un corps ordonné, et E un k-espace vectoriel muni d'une forme symetrique non-gégénérée g. Alors g est positive définie si et seulement si tous les mineurs principaux Δ_i sont strictement positifs.* ∎

James Joseph Sylvester (3 septembre, 1814, Londre - 15 mars, 1897 Oxford)

13.2.7 Cas des formes hermitiennes

Soit k_0 un corps ordonné, et soit $K = k_0(\alpha)$, où $\alpha^2 < 0$, et on pose $\alpha^\sigma = -\alpha$. On considère un K-espace vectoriel hermitien E muni d'une forme hermitienne $(x, y) \mapsto \langle x, y\rangle$, donc $\langle y, x\rangle = \langle x, y\rangle^\sigma$. On observe que $\langle x, x\rangle \in k_0$ pour tout $x \in E$.

THÉORÈME 13.2.13 (LOI D'INERTIE HERMITIENNE) *Soit E un K-espace vectoriel muni d'une forme hermitienne g. Il existe une base orthogonale de E.*

Si g est non-gégénérée, il existe un entier $r \geq 0$ tel que pour toute base orthogonale $\{v_1, \cdots, v_n\}$ de E, il y a exactement r éléments $\langle v_i, v_i\rangle$ strictement positifs, et exactement $n - r$ éléments $\langle v_i, v_i\rangle$ strictement négatifs.

Preuve. On pose $a_i = \langle v_i, v_i\rangle$ et $b_i = \langle w_i, w_i\rangle$ pour toute autre base orthogonale. On raisonne par l'absurde. ∎

La signature de g est le couple $(r_+, r_-) = (r, rk(g) - r)$ et on pose $\sigma(g) = r_+ - r_-$. Pour toute forme hyperbolique h, $\sigma(h) = 0$ et $\sigma(g \oplus h) = \sigma(g)$.

Identité de polarisation :

$$\langle u + \alpha v, u + \alpha v\rangle = \langle u, u\rangle - \alpha\langle u, v\rangle^\sigma + \alpha\langle u, v\rangle + \alpha\alpha^\sigma\langle v, v\rangle$$
$$\langle u - \alpha v, u - \alpha v\rangle = \langle u, u\rangle + \alpha\langle u, v\rangle^\sigma - \alpha\langle u, v\rangle + \alpha\alpha^\sigma\langle v, v\rangle$$
$$\alpha\langle u + v, u + v\rangle = \alpha\langle u, u\rangle + \alpha\langle u, v\rangle^\sigma + \alpha\langle u, v\rangle + \alpha\langle v, v\rangle$$
$$\alpha\langle u - v, u - v\rangle = \alpha\langle u, u\rangle - \alpha\langle u, v\rangle^\sigma - \alpha\langle u, v\rangle + \alpha\langle v, v\rangle, \text{ d'où}$$
$$\langle u, v\rangle = \frac{1}{4\alpha}[\langle u + \alpha v, u + \alpha v\rangle - \langle u - \alpha v, u - \alpha v\rangle + \alpha\langle u + v, u + v\rangle - \alpha\langle u - v, u - v\rangle]$$

13.2.8 Formes matricielle d'isométries*

Groupes $O(r_+, r_-)$ et $U(r_+, r_-)$.

13.2.9 Groupes symplectiques en coordonnées*

RAPPEL : Soit V un K-espace vectoriel de dimension finie sur un corps K de caractéristique différente de deux (par exemple, de caractéristique 0). Une forme K-bilinéaire

$$f : V \times V \to K$$

est dite *alternée* si pour tout $u \in V$, on a $f(u, u) = 0$.

(a) Montrer que pour tout corps K de caractéristique différernte de 2 cette condition est équivalente à l'identité : pour tout $u, v \in V$

$$f(u, v) = -f(v, u).$$

(b) On appelle une forme bilinéaire f *symplectique* si elle est antisymétrique et non-dégénérée, c'est-à-dire

$$\forall u \in V\backslash\{0\}, \exists u' \in V, f(u, u') \neq 0.$$

Montrer que l'application $\varphi : V \to V^*$, $\varphi(u)(v) = f(u, v)$ est un isomorphisme de K-espaces vectoriels.

(c) On considère le sous-espace $W \subset V$ engendré par u et u', et soit

$$V' = \{v \in V \mid \forall w \in W, f(v,w) = 0\}.$$

Montrer qu'il y a une décomposition $V = W \oplus V'$.

(d) En déduire que la dimension de V sur K est paire.

(f) En déduire qu'il existe une base $\{e_i\}$ de V telle que la matrice $A_f = (a_{i,j})$ de la forme f dans cette base est

$$J_n = \begin{pmatrix} 0 & 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \\ 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 1 \\ -1 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ 0 & -1 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \\ 0 & 0 & \cdots & -1 & 0 & 0 & \cdots & 0 \end{pmatrix}$$

(f) Pour toute forme hermitienne positive H sur un espace vectoriel complexe V on pose

$$f(z,w) = \text{Im } H(z,w).$$

Montrer que la forme f est symplectique (sur $\mathbb{R}$).

Déscription matricielle du groupe symplectique $G_n = \text{Sp}(V)$, à l'aide des matrices blocs :

(a) On considère l'ensemble

$$\text{Sp}(V) = \{g \in GL(V) \mid \forall u, v \in V, f(g(u), g(v)) = f(u,v)\}.$$

Montrer que $\text{Sp}(V)$ est un groupe dit symplectique (un sous-groupe de $\text{GL}(V)$).

(b) En utilisant une base de I d), montrer qu'il existe un isomorphisme

$$\text{Sp}(V) \xrightarrow{\sim} \text{Sp}_n(K) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \text{GL}_{2n}(K) \;\middle|\; a,b,c,d \in \text{Mat}_n(K), {}^t ac = {}^t ca, {}^t ad - {}^t cb = I_n \right\}$$

Soit $A = (a_{ij}) \in GL_{2n}(K)$ une matrice *antisymétrique inversible*, c'est-à-dire, $a_{ii} = 0$, ${}^t A = -A$, et $\det(A) \neq 0$. Pour toute base $\{e_i \mid i = 1, \cdots, 2n\}$, on a alors une forme symplectique f dont la matrice coincide avec A. Montrer qu'il existe une matrice inversible $C \in GL_{2n}(K)$, telle que ${}^t CAC = J_n \in GL_{2n}(K)$. En déduire que $\det A = (\det C)^{-2}$ est un carré dans le groupe miltiplicatif K^*.

(c) *Le pffafian.* Tout d'abord, soient t_{ij} $(1 \leq i < j \leq 2n)$ $n(2n-1)$ variables indépendantes, et soit $K = \mathbb{Q}(t) = \mathbb{Q}(t_{ij})$ le corps des fractions des variables t_{ij} $(1 \leq i < j \leq n)$. On considère la matrice $T = (t_{ij})$, où on pose $t_{ii} = 0$ pour $i = 1, 2, \cdots, 2n$, et $t_{ji} = -t_{ij}$ pour $1 \leq i < j \leq n$. Alors la matrice T est *antisymétrique inversible* sur K, donc d'après I(b), il existe une matrice inversible $C \in GL_{2n}(K)$, telle que ${}^t CTC = J_n \in GL_{2n}(K)$, et le polynôme $\det T = (\det C)^{-2}$ coincide donc avec le carré d'une *fraction rationnelle* dans K.

Montrer qu'il existe un polynôme $\mathrm{Pf}(T) \in \mathbb{Z}[t] = \mathbb{Z}[t_{ij}]$ dit le *pffafian générique* de T, tel que $\det(T) = \mathrm{Pf}(T)^2$.

(d) Montrer que si $n = 1$, $\mathrm{Pf}(T) = t_{12}$, et si $n = 2$,

$$\mathrm{Pf}(T) = t_{12}t_{34} - t_{13}t_{24} + t_{14}t_{23}.$$

(voir [Dieudonné], [Lang], [KosMan]).

(e) Pour tout anneau commutatif R et pour toute matrice antisymétrique $A \in GL_{2n}(R)$ on considère l'homomorphisme

$$\varphi \ : \mathbb{Z}[t] \to R,\ t_{ij} \mapsto a_{ij},$$

et on pose $\mathrm{Pf}(A) = \varphi(\mathrm{Pf}(T))$. Montrer qu'on a $\det(A) = \mathrm{Pf}(A)^2$.

(f) Montrer que pour toute matrice $C \in \mathrm{Mat}_{2n}(R)$ on a

$$\mathrm{Pf}({}^tCAC) = \det(C)\mathrm{Pf}(A).$$

Solution. On déduit tout d'abord une formule générique : soient u_{ij} $(1 \leq i, j \leq 2n)$ $4n^2$ variables algébriquement indépendantes sur $\mathbb{Q}$, et telles que t_{ij} et u_{ij} sont $4n^2+n(2n-1)$ variables algébriquement indépendantes sur $\mathbb{Q}$. On pose $U = (u_{ij}) \in \mathrm{Mat}_{2n}(K')$, où $K' = \mathbb{Q}(t_{ij}, u_{ij})$ est le corps des fractions des variables t_{ij} $(1 \leq i < j \leq 2n)$ et u_{ij} $(1 \leq i, j \leq 2n)$.

On déduit de I(e) que

$$\mathrm{Pf}({}^tUTU)^2 = \det(U)^2\mathrm{Pf}(T)^2 \Rightarrow \mathrm{Pf}({}^tUTU) = \pm\det(U)\mathrm{Pf}(T)$$

Ensuite, on substitue $U = I_{2n}$ et $T = J_n$, ceci implique immédiatement que le signe ci-dessus est $+$.

Enfin, pour tout anneau commutatif R, pour toute matrice antisymétrique $A \in GL_{2n}(R)$ et pour toute matrice carrée $C \in \mathrm{Mat}_{2n}(R)$ on considère l'homomorphisme

$$\psi \ : \mathbb{Z}[t, u] \to R,\ t_{ij} \mapsto a_{ij}, u_{ij} \mapsto c_{ij},$$

et on déduit

$$\mathrm{Pf}({}^tCAC) = \det(C)\mathrm{Pf}(A).$$

(g) En déduire que pour toute matrice symplectique $G = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{Sp}_n(K)$ on a $\det(G) = 1$.

13.3 Théorème de Witt et l'extension d'isométries

Ici, on donne un résultat traditionnel sur les extensions d' isométries dans les espaces non-dégénérés munis de la forme ci-dessus.

Ce résultat implique dans un cas spécial, que tous les sous-groupes paraboliques « de même type » dans les groupes d'isométries (et dans les groupes de similitudes) sont conjugués. On exclut le cas de la caractéristique 2, et on utilise les même notations que ci-dessus. Pour un (D, σ, ε)-espace V muni de forme $\langle , \rangle$, soit $-V$ note le (D, σ, ε)-espace qui est le même D-espace vectoriel mais muni de forme $-\langle , \rangle$. Soit V_0 désigne le noyau d'un (D, σ, ε)-espace V.

THÉORÈME 13.3.1 (WITT)

Soient U, W deux sous-espaces d'un espace non-dégénéré V. Toute isométrie $\phi : U \to W$ admet une extension à une isométrie $\Phi : V \to V$. (Ceci dit, la restriction de Φ sur U est ϕ).

Si U, V, W sont espaces tels que $U \oplus V \cong U \oplus W$, alors $V \cong W$.

Pour simplicité, on va traîter uniquement le cas symetrique (des formes quadratiques en caractéristique $\neq 2$).

13.3.1 Théorème de Witt, [Se70]

Soient (V, Q) et (V', Q') deux modules quadratiques *non-dégénérés* ; soit U un sous-espace vectoriel de V, et soit $s : U \to V'$ un morphisme injectif de U dans V'. On cherche à prolonger s à un sous-espace plus grand que U, et si possible à V tout entier. On commence par le cas où U est dégénéré :

LEMME 13.3.2 *Si U est dégénéré, on peut prolonger s en un morphisme métrique injectif $s_1 : U_1 \to V'$, où U_1 contient U comme hyperplan.*

Soit x un élément non nul de $\mathrm{rad}(U)$. Comme x est isotrope, la proposition 12.11.9 montre qu'il existe un plan hyperbolique de V qui le contient ; on peut donc trouver $y \in V$ tel que $\langle x, y\rangle = 1$ et $\langle y, y\rangle = 0$. Puisque y n'est pas orthogonal à x, on a $y \notin U$, et le sous-espace $U_1 = U \oplus ky$ contient U comme hyperplan. On construit de même un élément $y' \in V'$ tel que $\langle s(x), y'\rangle = 1$ et $\langle y', y'\rangle = 0$. On pose $s_1(y) = y'$. ∎

THÉORÈME 13.3.3 *Si (V, Q) et (V', Q') sont deux modules quadratiques isomorphes et non-dégénérés, tout morphisme injectif*

$$s : U \to V'$$

d'un sous-espace vectoriel U de V peut être prolongé en un isomorphisme de V sur V'.

PREUVE. Puisque V et V' sont isomorphes,on peut supposer $V = V'$. D'autre part, en appliquant le lemme ci-dessus, on voit que l'on peut se borner au cas où U est non-dégénéré. On raisonne par récurrence sur $\dim U$.

Si $\dim U = 1$, U est engendré par un élément non isotrope. Si $y = s(x)$, on a $\langle y, y\rangle = \langle x, x\rangle$. On peut choisir $\varepsilon = \pm 1$ tel que $x + \varepsilon y$ ne soit pas isotrope ; sinon, en effet, on aurait

$$2\langle x, x\rangle + 2\langle x, y\rangle = 2\langle x, x\rangle - 2\langle x, y\rangle = 0,$$

ce qui entraînerait $\langle x, x\rangle = 0$. Choisissons un tel ε et soit H l'hyperplan orthogonal à $z = x + \varepsilon y$; on a $V = kz \hat{\oplus} H$. Soit σ la symétrie par rapport à H, i.e. un automorphisme de V qui est identique sur H et qui change z en $-z$. Comme $x - \varepsilon y$ appartient à H, on a

$$\sigma(x - \varepsilon y) = x - \varepsilon y, \quad \text{et } \sigma(x + \varepsilon y) = -x - \varepsilon y,$$

d'où $\sigma(x) = -\varepsilon y$. L'automorphisme $-\varepsilon\sigma$ prolonge donc s.

Si $\dim U > 1$, on décompose U sous la forme $U_1 \hat{\oplus} U_2$ avec $U_1, U_2 \neq 0$.

D'après l'hypothèse de récurrence la restriction s_1 de s à U_1 se prolonge en un automorphisme σ_1 de V ; quitte à remplaçer s par $\sigma_1^{-1} \circ s$, on peut donc supposer que s est identité sur U_1. Le morphisme s applique alors U_2 dans l'orthogonal V_1 de U_1 ; d'après l'hypothèse de récurrence, la restriction de s sur U_2 se propolge donc en un automorphisme σ_2 de V_1 ; l'automorphisme σ de V qui est identité sur U_1 , et σ_2 sur V_1 repond alors à la question. ∎

COROLLAIRE 13.3.4 ("LOI DE SIMPLIFICATION") *Deux sous-espaces isomorphes d'un module quadratique non-dégénéré ont des orthogonaux isomorphes*

On prolonge un isomorphisme entre les deux sous-espaces en un automorphisme du module, et on restreint ce dernier aux orthogonaux.

En particulier, si U, V, W sont espaces tels que $U \oplus V \cong U \oplus W$, alors $V \cong W$.

COROLLAIRE 13.3.5 (CLASSIFICATION D'ESPACES ORTHOGONAUX) *Tout module quadratique L admet une décomposition orthogonale*

$$L = L_0 \oplus L_h \oplus L_a$$

où L_0 est isotrope, L_h hyperbolique et L_a anisotrope. Pour toutes deux telles décompositions il existe une isométrie $f : L \to L$, respectant cette décompositions.

En effet, $L_0 = \mathrm{rad} L$, et $L = L_0 \hat{\oplus} L_1$. Si L_a n'est pas anisotrope, on considère dans L_1 un espace maximal isotrope U qui existe par le lemme de Zorn. Complétons ce sous-espace à un espace hyperbolique $L_h \subset L_1$ de dimension double, et soit $L_a = L_h^{\perp}$ dans L_1. Alors L_a n'a pas des vecteurs isotropes (par la maximalité de U).

Unicité : pour deux décompositions

$$L = L_0 \oplus L_h \oplus L_a = L_0 \oplus L'_h \oplus L'_a$$

il existe une isométrie $f : L_0 \oplus L_h \to L_0 \oplus L'_h$, et on la complète par le théorème de Witt. ∎

Ce corollaire est une généralisation de la loi d'inertie, reduisant la classification d'espaces orthogonaux à celle d'espaces anisotropes.

13.3.2 Groupe de Witt*

(voir [Lang], [KosMan], [Mi-Hu]). On considère l'ensemble $W(k)$ des classes d'espaces orthogonaux anisotropes sur un corps k (à isométrie près).

On considère l'opération d'addition sur $W(k)$: pour deux espaces orthogonaux anisotropes L_1 et L_2, de classes $[L_1], [L_2] \in W(k)$ on définit $[L_1] + [L_2]$ comme la partie anisotrope de $L_1 \hat{\oplus} L_2$ (bien définie par le corollaire 13.3.1).

THÉORÈME 13.3.6

a) L'ensemble $W(k)$ muni d'addition est un groupe abélien

b) Soit L_α l'espace de dimension 1 muni du produit scalaire $\langle x, y \rangle = \alpha xy$. Alors la classe $[L_\alpha]$ ne dépend que de la classe de $\alpha (k^*)^2$, est les $[L_\alpha]$ engendrent $W(k)$. ∎

13.3.3 Exemples : $W(\mathbb{F}_q)$, $W(\mathbb{R})$, $W(\mathbb{Q})$

(voir Ch.4 de [Mi-Hu]).

THÉORÈME 13.3.7 (SANS DÉMONSTRATION)

a) Le groupe $W(\mathbb{F}_q) \cong \begin{cases} \mathbb{Z}/4\mathbb{Z}, & \text{si } q \equiv 1 \bmod 4 \\ \text{non-cyclique d'ordre 4}, & \text{si } q \equiv 3 \bmod 4 \end{cases}$

b) Le groupe $W(\mathbb{R}) \xrightarrow{\sim} \mathbb{Z}$ *par l'application de la signature* $\sigma(Q) = r_+ - r_-$ *;*

c) Il existe une suite exacte

$$0 \to \mathbb{Z} \to W(\mathbb{Q}) \ \to \bigoplus_p W(\mathbb{F}_p) \to 0.$$

Cours N 11 Lundi 20 avril 2015 (suite)

Géométrie projective. Coniques, quadriques

14 Géométrie projective

14.1 Espace projectif $\mathbb{P}^n$, variétés algébriques

Soit K un corps et $n \geq 1$ un entier. On considère l'espace projectif de dimension n sur K et on note $\mathbb{P}^n_K$ ou simplement $\mathbb{P}^n$ l'ensemble des classes d'équivalence de $(n+1)$-uplets

$$(X_0, \cdots, X_n) \in K^{n+1} \setminus (0, \cdots, 0),$$

sous la relation $(X_0, \cdots, X_n) \sim (\lambda X_0, \cdots, \lambda X_n)$ pour tout $\lambda \in K^*$. Une classe d'équivalence x est un point de $\mathbb{P}^n$.

Soit $K[X] = K[X_0, \cdots, X_n]$. On interprète les éléments de $K[X]$ comme des fonctions régulières de l'espace affine $\mathbb{A}^{n+1}$, et on interprète les éléments de

$$K\left(\frac{X_1}{X_0}, \frac{X_2}{X_0}, \cdots, \frac{X_n}{X_0}\right),$$

comme des fonctions rationnelles de $\mathbb{P}^n$

DÉFINITION 14.1.1 *Une partie $X \subset \mathbb{P}^n$ est dite une* variété algébrique projective *si*

$$X = X_P = \{x \in \mathbb{P}^n \mid \forall F \in P, F(x) = 0\}$$

où P est un idéal premier homogène de $K[X_0, \cdots, X_n]$, c'est à dire, un idéal premier engendré par des polynômes homogènes, par exemple, $P = (F)$, F un polynôme homogène irréductible.

14.2 Courbes planes projectives.

Rappelons qu'un point P du plan projectif $\mathbb{P}^2 = \mathbb{P}^2_K$ est donné comme la classe d'équivalence, notée $(X : Y : Z)$, d'un triplet non-nul $(X, Y, Z) \in K^3 \backslash \{(0, 0, 0)\}$, de telle façon que $(X, Y, Z) \sim (X', Y', Z') = (\lambda X, \lambda Y, \lambda Z)$ $(\lambda \in K^*)$.

On a l'inclusion

$$\mathbb{A}^2 \hookrightarrow \mathbb{P}^2, \quad (x, y) \mapsto (x, y, 1),$$

qui donne tous les points de $\mathbb{P}^2$ avec $Z \neq 0$.

On a les relations suivantes entre les coordonnées affines (x, y) et les coordonnées projectives $(X : Y : Z)$:

$$X = xZ, Y = yZ, \quad x = \frac{X}{Z}, y = \frac{Y}{Z}$$

Cartes affines

On a les trois parties suivantes de $\mathbb{P}^2$:

$$\mathbb{A}^2_1, \mathbb{A}^2_2, \mathbb{A}^2_3 \subset \mathbb{P}^2, \text{ telles que } \mathbb{A}^2_1 : X \neq 0,\ \mathbb{A}^2_2 : Y \neq 0,\ \mathbb{A}^2_3 : Z \neq 0.$$

isomorphes a K^2, et on a les coordonnées

$$\begin{aligned}
&\text{sur } \mathbb{A}^2_1 : X = xZ, Y = yZ, \quad x = \frac{X}{Z}, y = \frac{Y}{Z} \\
&\text{sur } \mathbb{A}^2_2 : X = x'Y, Z = y'Y, \quad x' = \frac{X}{Y}, y' = \frac{Z}{Y} \\
&\text{sur } \mathbb{A}^2_3 : Y = x''X, Z = y''X, \quad x'' = \frac{Y}{X}, y'' = \frac{Z}{X}
\end{aligned}$$

EXERCICE . Réécrire l'équation de la courbe de Fermat en coordonnées (x', y') et en (x'', y'').

Une courbe algébrique dans $\mathbb{P}^2$ est donnée par une équation homogène dans les coordonnées projectives : $F(X, Y, Z) = 0$, alors l'égalité $F(X, Y, Z) = 0$ ne dépend pas du choix des coordonnées projectives : si $(X', Y', Z') = (\lambda X, \lambda Y, \lambda Z)$, alors $F(X', Y', Z') = \lambda^d F(X, Y, Z) = 0$, où d est le degré homogène du polynôme F.

14.3 Fonctions affines et quadratiques, et quadriques affines

(voir [KosMan], §5, Partie 3, p.221). Soit A un espace affine associé à un k-espace vectoriel sur un corps commutatif k, (c'est-à-dire, muni d'une action additive $A \times V \to A$ de V, avec $A = a_0 + V$). Une fonction $Q : A \to k$ de la forme

$$Q(a_0 + x) = q(x) + l(x) + c$$

est dite quadratique, où $a_0 \in A$, $l \in V^*$, $q : V \to k$ une forme quadratique. L'ensemble

$$X = X^{\text{aff}}_Q = \{a \in A \mid Q(a) = 0\},$$

est dite *quadrique affine* dans A. Pour un choix de base (e_i) de V, on a

$$Q(a_0 + x) = \sum_{i,j=1}^{n} a_{ij} x_i x_j + 2 \sum_{i=1}^{n} b_i x_i + c,$$

où $a_{ij} = a_{ji}$. Soit

$$A_Q = \left(\begin{array}{c|c} & b_1 \\ A_q & \vdots \\ & b_n \\ \hline b_1 \cdots b_n & c \end{array}\right)$$

et on pose $\Delta = \det A_Q$, $\delta = \det A_q$. Si $\Delta \neq 0$, la quadrique Q est dite non-dégénérée.

On point a_0 est dit central si $l(x) = 0$. Q est dit central s'il existe un centre de Q.

Classification des quadriques

Pour amener une quadrique affine à une forme canoniqie dans $\mathbb{A} \cong K^n$ on utilise les transformations (bijectives) affines

$$\varphi : \mathbb{A} \to \mathbb{A}, \varphi(a_0 + x) = \varphi(a_0) + \psi(x), \text{où } \psi \in GL_K(V).$$

Toute transformation affine donne un nouveau système de coordonnées $(y_1, \cdots, y_n)$ associé à l'origine $\varphi(a_0)$ et la nouvelle base $\psi(e_1), \cdots, \psi(e_n)$:

$$a_0 + x_1e_1 + \cdots + x_ne_n = \varphi(a_0) + y_1\psi(e_1) + \cdots + y_n\psi(e_n) \Longleftrightarrow \begin{pmatrix} x_1 \\ \cdots \\ x_n \end{pmatrix} = C \begin{pmatrix} y_1 \\ \cdots \\ y_n \end{pmatrix} + v',$$

où $v' = (v_1, \cdots, v_n)^t, v_1e_1 + \cdots + v_ne_n = \varphi(a_0) - a_0$.
La matrice de Q en nouvelles coordonnées devient

$$A'_Q = \left(\begin{array}{c|c} & b'_1 \\ A'_q & \vdots \\ & b'_n \\ \hline b'_1 \cdots b'_n & c' \end{array}\right) = C^t A_Q C, \text{ où } C = \left(\begin{array}{c|c} & v_1 \\ C_\psi & \vdots \\ & v_n \\ \hline 0 \cdots 0 & 1 \end{array}\right).$$

On en déduit la classification des quadriques suivante :
a) Si q est non-dégénérée,

$$Q(a_0 + x) = \sum_{i=1}^{n} \lambda_i x_i^2 + c,$$

b) Si q est dégénérée de rang r, et Q est central,

$$Q(a_0 + x) = \sum_{i=1}^{r} \lambda_i x_i^2 + c,$$

c) Si q est dégénérée de rang r, et Q n'est pas central,

$$Q(a_0 + x) = \sum_{i=1}^{r} \lambda_i x_i^2 + x_{r+1},$$

Cas $k = \mathbb{R}$:

$$\begin{array}{ll} (\mathrm{I}_{r,s}) & x_1^2 + \cdots + x_s^2 - x_{s+1}^2 - x_r^2 = 1, \quad 0 \le s \le r \le n; \\ (\mathrm{I}'_{r,s}) & x_1^2 + \cdots + x_s^2 - x_{s+1}^2 - x_r^2 = 0, \quad 0 \le s \le r \le n, s \ge r/2; \\ (\mathrm{II}_{r,s}) & x_1^2 + \cdots + x_s^2 - x_{s+1}^2 - x_r^2 = 2x_{r+1}, \quad 0 \le s \le r \le n-1, s \ge r/2. \end{array}$$

Exercice

a) Montrer qu'il y a exactement 17 cas pour $n = 3$, 9 cas pour $n = 2$
b)* Donner une formule générale sur $\mathbb{R}$, sur $\mathbb{C}$ et sur $\mathbb{F}_q$

14.4 Coniques

De façon similaire, on considère l'équation homogène

$$Q(X_0, X_1, \cdots, X_n) = \sum_{i,j=0}^{n} a_{ij} X_i X_j = \sum_{i,j=1}^{n} a_{ij} X_i X_j + 2\sum_{i=1}^{n} a_{i0} X_i X_0 + a_{00} X_0^2 \qquad (14.1)$$

où $a_{0i} = a_{i0} = b_i$ pour $i = 1, 2, \ldots, n$, $a_{00} = c$. Les *coordonnées non–homogènes (affines)* $x_1, \cdots, x_n$ sont reliées aux *coordonnées homogènes (projectives)*$X_0, \ldots, X_n$ par

$$X_i = x_i X_0 \quad (i = 1, 2, \ldots, n).$$

La forme quadratique $Q(X)$ peut être écrite de façon commode sous la forme

$$Q(X) = X^t A_Q X, \quad X^t = (X_0, X_1, \ldots, X_n),$$

où $A_Q = (a_{ij})$ est la matrice des coefficients.

Classification des coniques

Concernant l'équation

$$Q(X_0, X_1, \ldots, X_n) = 0 \qquad (14.2)$$

on peut commencer par la diagonalisation de A_Q avec une substitution linéaire non–dégénérée $X = CY$. La matrice C peut être trouvée effectivement par la méthode classique (de Lagrange) de l'extraction successive des carrées.

Classification des coniques

est plus simple dans le cas projectif que dans le cas affine

$$Q(X_0, X_1, \cdots, X_n) = \sum_{i=0}^{r} \lambda_i X_i^2, \quad r + 1 = \mathrm{rk} A_Q$$

Cas $k = \mathbb{R}$:

$$X_0^2 + \cdots + X_s^2 - X_{s+1}^2 - X_r^2 = 0, \quad 0 \le s \le r \le n, s \ge (r-1)/2;$$

EXERCICE

a) Montrer qu'il y a exactement 8 cas pour $n = 3$, et 5 cas pour $n = 2$. C'est à dire, que certaines quadriques affines non-equivalentes deviennent equivalentes après la clôture projective, c'est-à-dire, après le passage aux coordonnées projectives.
b)* Donner une formule générale sur $\mathbb{R}$, sur $\mathbb{C}$ et sur $\mathbb{F}_q$

15 Applications projectives et leurs utilisations

15.1 Groupes projectifs et projections

(voir [KosMan], Partie III, §8).

Soint L, M deux k-espaces vectoriels, $f : L \to M$ une application linéaire. Si $\operatorname{Ker} f = 0$, f envoie toute ligne droite de L dans une ligne droite dans M, donc f induit une application $P(f) : P(L) \to P(M)$, dite la projectivisation de$f$. En particulier, si f est un isomorphisme, $P(f)$ est dit un isomorphisme projectif.

Le groupe projectif, noté $PGL_k(L) \cong PGL(n+1, k) = GL(n+1, k)/Z$ est formé de tous les automorphismes projectifs de $P(L)$.

15.1.1 Action du groupe projectif sur les configurations projectives

On appelle configuration projective un système des sous-espaces profectifs de $P(L)$. Deux configurations sont dites congruentes si l'une transforme sur l'autre par une transformation projective.

a) Le groupe projectif $PGL_k(L)$ agit transitivement sur l'ensemble des sous-espaces projectifs de dimension donnée.

b) Le groupe projectif $PGL_k(L)$ agit transitivement sur l'ensemble des drapeaux de sous-espaces projectifs de type donné.

c) Le groupe projectif $PGL_k(L)$ agit transitivement sur l'ensemble des multiplets $(P_1, P_2, \cdots, P_m)$ des sous-espaces projectifs de dimension donnée $\dim P_i$ tels que pour tout i l'intersection de P_i avec l'enveloppe projective

$$(P_1, \cdots, P_{i-1}, \cdots, P_{i+1}, \cdots, P_m)$$

est vide. En effet, si $P_i = P(L_i)$,

$$L_i \cap \sum_{j \neq i} L_j = \{0\} \Longrightarrow \sum_j L_j = \bigoplus_j L_j \Longrightarrow L = (\bigoplus_j L_j) \oplus L'.$$

Définition 15.1.1 *Un multiplet $\{p_1, \cdots, p_N\}$ de points de $P(L)$ est dit en position générique si pour tout $m \leq \min\{N, n+1\}$ et pour tour toute partie $S \subset \{1, \cdots, N\}$ de cardinal m, l'enveloppe projective de $\{p_i \mid i \in S\} = P(L_S)$ est de dimension $m-1$ (c'est-à-dire, $\dim_k V_S = m$).*

d) Tous les systèmes de $n+2$ points en position générique sont congruents par c). Pour un tel système $\{p_1, \cdots, p_{n+2}\}$ on peut supposer que

$$p_1 = (1 : 0 : \cdots : 0), p_2 = (0 : 1 : \cdots : 0), \cdots, p_{n+1} = (0 : 0 : \cdots : 1), p_{n+2} = (1 : 1 : \cdots : 1).$$

Définition 15.1.2 *Le birapport d'un qudruplet $\{p_1, p_2, p_3, p_4\} \subset \mathbb{P}^1$ des points distincts est defini par*

$$[p_1, p_2, p_3, p_4] = f(p_4)$$

où f est une transformation projective unique de $\mathbb{P}^1$ avec
$f(p_1)=0$, $f(p_2)=1$, $f(p_3)=\infty$, *donc pour* $\{p_1,p_2,p_3,p_4\}=\{x_1,x_2,x_3,x_4\}$,

$$f: x \mapsto \frac{x_1-x}{x_3-x} : \frac{x_1-x_2}{x_3-x_2}, \quad [p_1,p_2,p_3,p_4] = \frac{x_1-x_4}{x_3-x_4} : \frac{x_1-x_2}{x_3-x_2}.$$

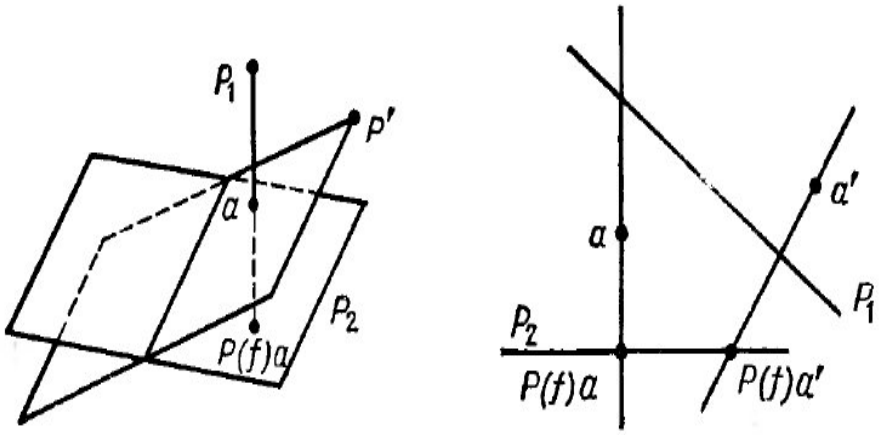
15.1.2 Projections

Soint L, M deux k-espaces vectoriels, $f: L \to M$ une application linéaire. Dans le cas où $\operatorname{Ker} f \neq 0$, $P(f)$ n'est défini que sur le complémentaire $U_f = P(L)\backslash P(\operatorname{Ker} f)$.

Soit $L = L_1 \oplus L_2$ la somme directe de deux sous-espaces ≥ 1. On pose $P = P(L)$, $P_i = P(L_i)$. La projection $f: L \to L_2$ induit l'application $P(f): P\backslash P_1 \to P_2$, dite *la projection du centre P_1 sur P_2.*

(a) On remarque que $\dim P_1 + \dim P_2 = \dim P - 1$ et $P_1 \cap P_2 = \emptyset$.

(b) Si $a \in P_2$, alors $P(f)a = a$; si $a \in P\backslash(P_1 \cup P_2)$, alors $P(f)a$ coïncide avec l'intersection de l'unique ligne droite dans P, qui coupe P_1 et P_2 et passe par a. En effet, par toute ligne $L_0 \subset L\backslash(L_1 \cup L_2)$, il existe un seul plan contenant L_0 qui coupe L_1 et L_2 par des lignes droites.



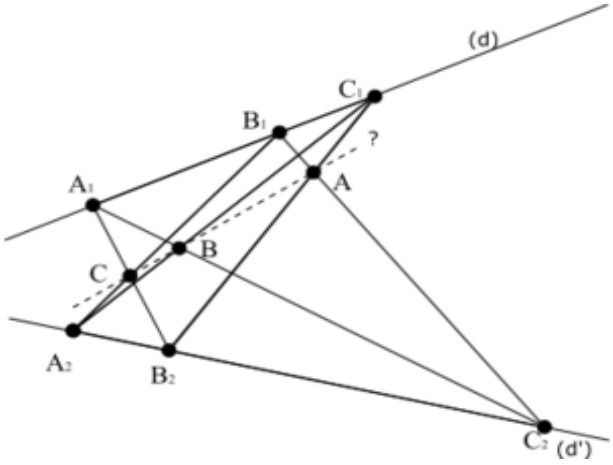
Ceci implique que la restriction de l'application sur sous-espace $P' \subset P\backslash P_1$ est une application projective $P(g)$.

15.2 Configurations de Pappus et de Desargues

15.2.1 Théorème de Pappus, voir [W]

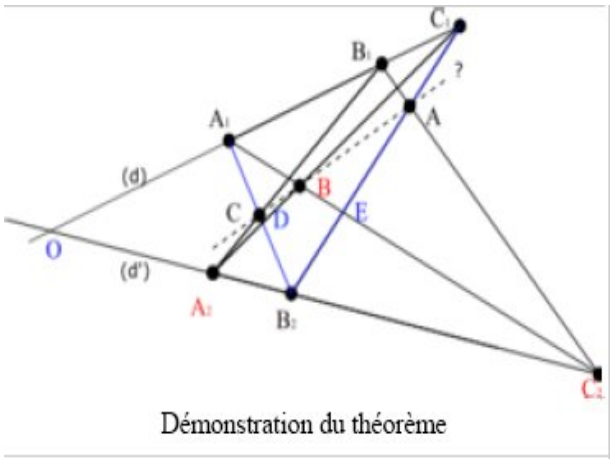
(voir aussi `http ://hexamys.free.fr/pappus.htm`

THÉORÈME 15.2.1 *Dans un plan, soient A_1, B_1, C_1 trois points distincts quelconques alignés sur une droite quelconque (d) , et soient A_2, B_2, C_2 trois autres points distincts quelconques alignés sur une autre droite quelconque (d') alors les points*
A intersection de (B_2C_1) avec (C_2B_1)
B intersection de (A_2C_1) avec (C_2A_1)
C intersection de (A_2B_1) avec (B_2A_1) sont alignés.



Il s'agit d'un théorème de géométrie projective donc les points considérés peuvent être propres ou impropres. Dans le cas où tous les points sont propres, on obtient une configuration du type ci-contre.

Démonstration à l'aide des applications projectives



On construit les points O intersection de (d) et (d'), D intersection de (A_1B_2) et (A_2C_1) et E intersection de (A_1C_2) et (C_1B_2).

On considère la projection centrale p de la droite (A_1B_2) sur la droite (d) de centre A_2

A_1 a pour image A_1
C a pour image B_1
D a pour image C_1
B_2 a pour image O

On considère la projection centrale q de la droite (d) sur la droite (B_2C_1) de centre C_2

A_1 a pour image E
B_1 a pour image A

C_1 pour image C_1
O a pour image B_2

Par l'application projective $q \circ p$ de la droite (A_1B_2) sur la droite (B_2C_1)
A_1 a pour image E
C a pour image A
D a pour image C_1
B_2 a pour image B_2

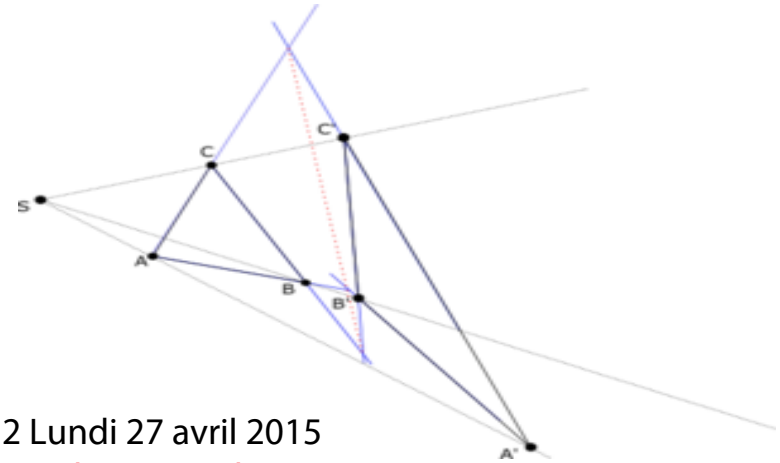
Si on regarde maintenant la projection centrale r de la droite (A_1B_2) sur la droite (B_2C_1) de centre B
A_1 a pour image E
D a pour image C_1
B_2 a pour image B_2
Or, une application projective d'une droite sur une autre est entièrement déterminée par l'image de trois points distincts. Les transformations $q \circ p$ et r coïncident sur A_1, D et B_2. Elles sont donc égales et $r(C) = A$. Les points A, B et C sont donc alignés. ∎

15.2.2 Théorème de Desargues, voir [W]

En géométrie projective

THÉORÈME 15.2.2 *Soient ABC et $A'B'C'$ deux triangles sans point commun tels que les droites (AA'), (BB') (CC') sont concourantes alors les points d'intersection des droites (AB) et $(A'B')$ (AC) et $(A'C')$ (BC) et $(B'C')$ sont situés sur une même droite. Dans le cas où aucun point d'intersection n'est un point impropre, on obtient la configuration ci-contre. Le dual de ce théorème en donne aussi la réciproque : ABC et $A'B'C'$ deux triangles sans point commun tels que les points d'intersection des droites (AB) et $(A'B')$ (AC) et $(A'C')$ (BC) et $(B'C')$ sont situés sur une même droite alors les droites (AA'), $(BB')(CC')$ sont concourantes (en un point propre ou impropre).*

En géométrie projective, une démonstration simple consiste à utiliser une application projective qui transforme les triangles (ABC) et $(A'B'C')$ en $(A_1B_1C_1)$ et $(A'_1B'_1C'_1)$ de telle sorte que les points d'intersection de (A_1B_1) et $(A'_1B'_1)$ et de (A_1C_1) et $(A'_1C'_1)$ soient impropres. Les droites (A_1B_1) et $(A'_1B'_1)$ sont alors parallèles ainsi que les droites (A_1C_1) et $(A'_1C'_1)$. On est alors revenu à la configuration précedente qui prouve que (B_1C_1) et $(B'_1C'_1)$ sont parallèles et que leur point d'intersection est donc impropre (situé sur la même droite que les deux précédents). On en déduit alors la propriété sur les droites de départ.



Cours N 12 Lundi 27 avril 2015

Révision pour l'examen du 4/5/2015

15.3 Théorème fondamental de la géométrie projective*

(voir [Dieudonné], p.77 ; [Sha87], §10, [KosMan], Partie III, §9). On considère les espaces projectifs $P(E)$, $P(E')$ sur les corps gauches K, K'.

THÉORÈME 15.3.1 *Soit $\varphi : P(E) \to P(E')$ une application bijective telle que trois points en ligne droite dans $P(E)$ soient transformés par φ en trois points en ligne droite dans $P(E')$. Alors, si $n \geq 3$, K et K sont isomorphes et $\varphi = P(g)$ pour une application semi-linéaire bijective $g : E \to E'$.*

15.3.1 Un espace projectif abstrait $\mathfrak{P}$

On considère un ensemble partiellement ordonné, tel que

1) Tout ensemble d'éléments x_α possède le suprémum $y = \cup_\alpha x_\alpha$ (l'union, ou "l'enveloppe projective"), noté $Sup(\{x_\alpha\})$;

2) Tout ensemble d'éléments x_α possède le infimum $y' = \cap_\alpha x_\alpha$ (l'intersection) noté $Inf(\{x_\alpha\})$; on note $O = Inf(\mathfrak{P})$ ("le sous-ensemble vide" vu comme un élément de $\mathfrak{P}$;

3) Pour tous $x, y \in \mathfrak{P}$, $y \leq x$ soit x/y l'ensemble d'éléments $z \in \mathfrak{P}$ tels que $y \leq z \leq x$. Pour tous $x, y \in \mathfrak{P}$, et $a \in x/y$, il existe $b \in x/y$, tel que

$$a \cup b = Sup(x/y) \text{ et } a \cap b = Inf(x/y).$$

4) *Finitude des longueurs des chaînes* :

$$a_1 \leq a_2 \leq \cdots \leq a_r, a_1 \neq a_2, a_2 \neq a_3, \cdots a_{r-1} \neq a_r,$$

alors on a $r \leq n + 2$ pour un n (la dimension de $\mathfrak{P}$).

5) Un $a \in \mathfrak{P}$ est dit *un point*, si $b \leq a, b \neq a$ implique $b = O = Inf(\mathfrak{P})$. Pour tout points distincts a et b il existe $c \neq a$, $c \neq b$, tel que $c \leq a \cup b$,

On peut déduire du théorème fondamental de la géométrie projective que $\mathfrak{P} \cong P(E)$ pour un espace projectif $P(E)$ sur un corps gauche K, si $n \leq 3$, et pour $n = 2$ on ajoute

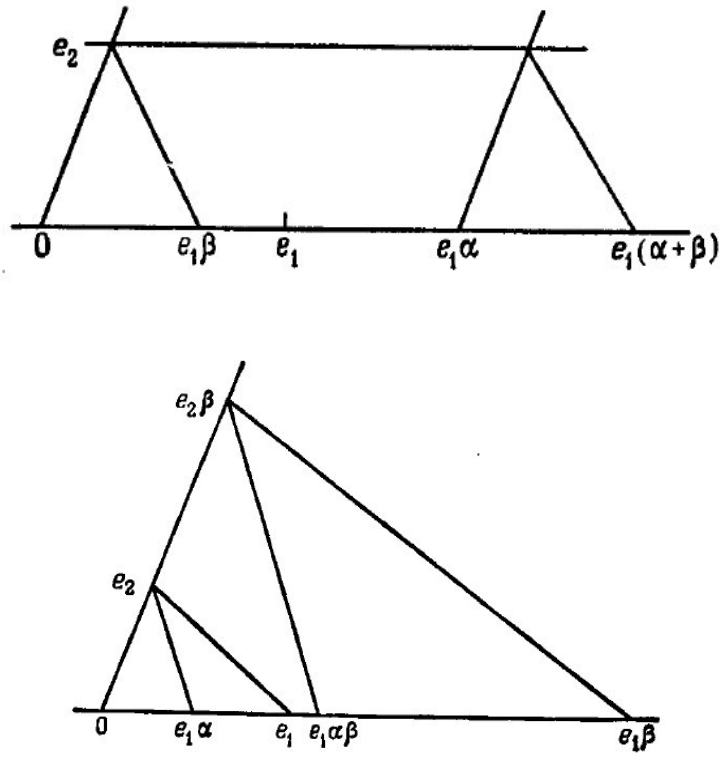
L'axiome de Desargues : Soient ABC et $A'B'C'$ trois triangles sans point commun tels que les droites $(AA'), (BB'), (CC')$ sont incidentes à un même point alors les points

d'intersection des droites (AB) et $(A'B')$, (AC) et $(A'C')$, (BC) et $(B'C')$ sont incidents à une même droite.

Pour construire le corps gauche K à partir d'un plan arguésien, on fixe une ligne droite muni d'un triplet de points distincts p_0, p_1, p_2 et on définit une structure de corps sur $D \backslash p_2$, avec 0 en p_0, et 1 en p_1 comme dans la figure suivante.

On peut montrer aussi qu'un plan projectif satisfait *l'axiome de Pappus* si et seulement si le corps K est commutatif.

REMARQUE . Tout corps fini est commutatif (voir [Wei74], §1).



Axiomes préalables d'un plan projectif (rappel, voir [W])

Un *plan projectif d'incidence* (PPI) est un PP qui vérifie les axiomes :

Il existe au moins 2 points.

Chaque droite possède au moins 3 points.

Pour deux points distincts il existe une et une seule droite qui leur est incidente.

Deux droites distinctes ont un et un seul point commun.

Pour toute droite il existe au moins un point non incident à cette droite.

(Auxquels on ajoute l'axiome « invisible » d'abondance : toute droite et tout point du PPI possèdent respectivement autant de points et de droites qu'il est nécessaire pour que la configuration étudiée ne soit pas nécessairement dégénérée).

Un *plan projectif arguésien* (PPA) est un PPI qui vérifie l'axiome de Desargues :
Soient ABC et $A'B'C'$ trois triangles sans point commun tels que les droites (AA'), (BB'), (CC') sont incidentes à un même point alors les points d'intersection des droites (AB) et $(A'B')$, (AC) et $(A'C')$, (BC) et $(B'C')$ sont incidents à une même droite.

Un *plan projectif de Pappus* (PPP) est un PPI qui vérifie l'axiome de Pappus : Dans un plan , soient A_1, B_1, C_1 trois points distincts quelconques alignés sur une droite quelconque (d), et soient A_2, B_2, C_2 trois autres points distincts quelconques alignés sur une autre droite quelconque (d'), alors les points A intersection de (B_2C_1) avec (C_2B_1), B intersection de (A_2C_1) avec (C_2A_1), C intersection de (B_2A_1) avec (A_2B_1) sont alignés.

16 Courbes planes*

16.1 Points singuliers des courbes projectives

Soit K un corps. Rappelons qu'une courbe projective plane $\mathcal{C}$ sur K est définie par une équation de type $F(X : Y : Z) = 0$, où $F(X : Y : Z) \in K[X, Y, Z]$ est une forme homogène des variables projectives X, Y, Z.

L'équation de la tangente, dans les coordonnées affines, a la forme

$$f'_x(P)(x - \alpha) + f'_y(P)(y - \beta) = 0.$$

Par la construction,

$$f(x, y) = F(x, y, 1), \text{ où } F(X, Y, Z) = 0 \text{ l'équation homogène de la courbe.}$$

Ceci implique : $f'_x = F'_x, f'_y = F'_y$, et selon le théorème connu de Euler (sur les fonctions homogènes) on a

$$XF'_X + YF'_Y + ZF'_Z = nF \text{ où } n \text{ est le degré de } F$$

Lorsque $P = (\alpha : \beta : 1)$ se trouve sur la courbe alors

$$\alpha F'_X(P) + \beta F'_Y(P) + F'_Z(P) = nF,$$

donc l'équation de la tangente se transforme en

$$xF'_X(P) + yF'_Y(P) + F'_Z(P) = 0 \Longleftrightarrow XF'_X + YF'_Y + ZF'_Z = 0.$$

Définition 16.1.1

(a) Un point singulier *sur une courbe projective plane $\mathcal{C}$ sur K est toute solution du système*

$$F = F'_X = F'_Y = F'_Z = 0$$

dans une extension de K.

(b) On dit qu'une courbe projective plane $\mathcal{C}$ sur K est lisse *si le système*

$$F = F'_X = F'_Y = F'_Z = 0$$

n'a pas de solutions non–triviales dans toute extension de K.

Exemple

(a) Soit $Q(X,Y,Z)$ une forme quadratique de matrice A_Q, sur un corps K de caractéristique $\mathrm{Car}(K) \neq 0$. Montrer que la conique $Q(X,Y,Z) = 0$ est non-singulière si et seulement si A_Q est inversible.

(b) Soit $Car(K) \neq 2, 3$, et soit $\mathcal{C}$ donnée par l'équation homogène correspondante

$$Y^2Z = X^3 + aXZ^2 + bZ^3,\ O = (0:1:0).$$

On vérifie que cette courbe est lisse si et seulement si le polynôme cubique $x^3 + ax + b$ n'a pas de racines multiplies (directement par la définiton des points singuliers comme des solutions de l'équation $F = F'_X = F'_Y = F'_Z = 0$ dans le cas général $F(X,Y,Z) = Y^2 + a_1XYZ + a_3YZ^2 - X^3 - a_2X^2Z - a_4XZ^2 - a_6Z^3$).

(c) Soit $K = \mathbb{F}_2$, et soit $\mathcal{C}$ une courbe sur K donnée par l'équation affine

$$y^2 + y = x^3 + ax + b,\ a, b \in \mathbb{F}_2$$

Montrer que cette courbe est toujours lisse (on n'a plus besoin d'exclure ici le cas des racines multiples).

(d) Soit $K = \mathbb{F}_3$, et soit $\mathcal{C}$ une courbe sur K donnée par l'équation affine :

$$y^2 = x^3 + ax^2 + bx + c,\ a, b, c \in \mathbb{F}_3$$

Montrer que cette courbe est lisse si et seulement si le polynôme cubique à droite n'a pas de racines multiples.

16.2 Equations cubiques

Le problème de l'existence d'une solution.

Pour les *formes cubiques* $F(X,Y,Z)$ en trois variables de coefficients entiers, on ne connait pas d'algorithme qui décide si l'équation $F = 0$ possède une solution non-triviale en nombres entiers en général. Les grandes classes de telles equations ont étés étudiées de point de vue théorique et numerique ; par exemple Selmer E.S. (1951–1954) a étudié en detail les équations de type

$$aX^3 + bY^3 + cZ^3 = 0.$$

Il arrive que même pour des équations simple comme $3X^3 + 4Y^3 + 5Z^3 = 0$ le principe de Minkowski–Hasse n'est pas valable : on peut montrer que cette équation n'a pas de solutions en nombres entiers, quoiqu'il existe des solutions réelles et des solution primitives modulo tout $N > 1$.

Addition de points sur une cubique plane

Toute forme cubique non-nulle $F(X,Y,Z)$ à coefficients entiers definit une courbe cubique $\mathcal{C}$ dans l'espace projectif $\mathbb{P}^2$ (sur $\mathbb{Q}$ et sur $\mathbb{C}$) :

$$\mathcal{C} = \{(X:Y:Z) \mid F(X,Y,Z) = 0\}. \tag{16.1}$$

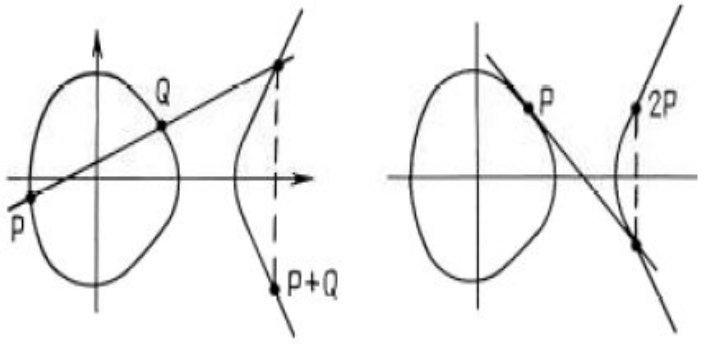
Si $\mathcal{C}$ est non-singulière et si $F = 0$ possède au moins une solution rationnelle, alors on peut trouver un changement de variables inversible à coefficients rationnels qui réduit F à la forme normale de Weierstrass

$$Y^2Z - X^3 - aXZ^2 - bZ^3 \qquad (a, b \in \mathbb{Q}). \tag{16.2}$$

On peut aussi supposer que la solution rationnelle de depart devient la solution évidente $(0 : 1 : 0)$ de l'équation ainsi obtenue (16.2). La condition de non-singularité de (16.2) est équivalente à la non-annulation du discriminant $4a^3 + 27b^2$. Une courbe cubique plane non-singulière est dite *elliptique*, si elle possède un point rationnel. En utilisant les coordonnées affines $x = X/Z, y = Y/Z$ on réduit $F = 0$ à la forme suivante :

$$y^2 = x^3 + ax + b, \tag{16.3}$$

où le polynôme cubique dans la partie droite n'a pas de racines multiples. Sous cette forme affine, la solution rationnelle ci-dessus devient le point à l'infini O. Il existe une jolie description gométrique de la loi de composition sur l'ensemble des points rationnels de $\mathcal{C}$ qui devient un groupe abélien avec le point à l'infini O comme élément neutre. Cette loi est donnée par la "méthode de sécantes et tangentes" de Poincaré. Notamment, pour une paire de points $P, Q \in \mathcal{C}(\mathbb{Q})$, on construit d'abord une droite passant par P, Q. Une telle droite coupe aussi $\mathcal{C}$ en un troisième point bien défini P'. Ensuite, on construit de nouveau une droite passant par P' et O. Enfin, son troisème point d' intersection avec $\mathcal{C}$ est dit la somme $P + Q$. Si $P = Q$, la première droite à construire doit bien-sûr être tangente à $\mathcal{C}$ en P.



Un calcul simple en coordonnées affines $P = (x_1, y_1)$, $Q = (x_2, y_2)$ montre que $P + Q = (x_3, y_3)$ où

$$\begin{aligned} x_3 &= -x_1 - x_2 + \left(\frac{y_1 - y_2}{x_1 - x_2}\right)^2, \\ y_3 &= \frac{y_1 - y_2}{x_1 - x_2}(x_1 - x_3) - y_1. \end{aligned} \tag{16.4}$$

Dans le cas limite $P = Q$ on remarque que $2y'_x y = 3x^2 + a$, et on obtient

$$x_3 = -2x_1 + \left(\frac{3x_1^2 + a}{2y_1}\right)^2, \; y_3 = \frac{3x_1^2 + a}{2y_1}(x_1 - x_3) - y_1. \tag{16.5}$$

Si $x_1 = x_2$ et $y_1 = -y_2$ alors $P + Q = O$, est le point à l'infini, qui est élément neutre pour la loi du groupe.

Cette méthode nous permet de construir de nouveaux points rationnels à partir des points connus. De tels points forment un sous-groupe engendré par des points de départ, par exemple, $mP,\ m \in \mathbb{Z}$, juste à partir du seul point P.

Pour les courbes cubiques singulières cette construction ne marche pas. Par exemple, on considère la courbe

$$\mathcal{C} :\ y^2 = x^2 + x^3, \tag{16.6}$$

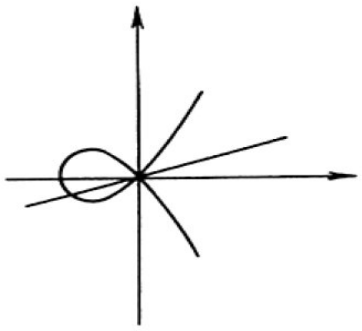


Fig. 8

qui est représentée par Fig. 8. Toute droite passant par $(0, 0)$ n'a qu'un seul autre point d'intersection avec $\mathcal{C}$: sur $y = tx$ il est donné par l'équation $x^2(t^2 - x - 1) = 0$. A part de la solution triviale $x = 0$, on obtient $x = t^2 - 1$ et $y = t(t^2 - 1)$ donc nous avons trouvé tous les points sur $\mathcal{C}$ à l'aide d'une parametrisation rationnelle. Dans le cas non-singulier il n'existe pas de telle parametrisation.

Une courbe admettant une parametrisation rationnelle est dite une courbe rationnelle. Rapellons qu'un autre exemple d'une courbe rationnelle est donné par une conique plane et sa projection stéréographique.

La structure de groupe des points rationnels sur une cubique plane

Une propriété très remarquable de la "méthode de sécantes et tangentes" de Poincaré est que cette méthode nous permet de construire tous les points rationnels à partir d'un nombre fini de points rationnels. Du point de vue de la théorie des groupes, cela signifie, que le résultat suivant a lieu :

THÉORÈME 16.2.1 (MORDELL, 1922) *Le groupe abélien $\mathcal{C}(\mathbb{Q})$ est de génération finie.*

On obtient alors à partir du théorème de structure des groupes abéliens que

$$\mathcal{C}(\mathbb{Q}) \cong \Delta \times \mathbb{Z}^r$$

où Δ est le sous-groupe fini formé par les points de torsion, et $\mathbb{Z}^r$ est le produit de r copies du groupe cyclique infini. Le nombre r est dit le rang de $\mathcal{C}$ sur $\mathbb{Q}$. Il est connu que le groupe de torsion Δ peut être déterminée explicitement. Par exemple, Nagell et Lutz

(Lutz E. (1937)) ont démontré que les points de torsion d'une courbe $y^2 = x^3 + ax + b$ avec a et b des nombres entiers, ont les coordonnées entiers x, y. De plus, l'ordonnée y d'un point de torsion est soit nulle, soit divise le nombre entier $D = -4a^3 - 27b^2$.

B.Mazur a démontré en 1976 que le sous-groupe de torsion Δ sur $\mathbb{Q}$ est isomorphe a l'un des quinze groupes suivants :

$$\mathbb{Z}/m\mathbb{Z}\ (m \leq 10, m = 12),\ \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2n\mathbb{Z}\ (n \leq 4), \tag{16.7}$$

et toutes ces possibilités se réalisent.

Une question ouverte importante est si r peut être infiniment grand. En 1982 J.-L. Mestre a construit des exemples de courbes de rang au moins 14. Il a donné aussi un exemple relativement simple d'une courbe de rang ≥ 9 : $y^2 + 9767y = x^3 + 3576x^2 + 425x - 2412$.

En 2000 Martin – Mcmillen ont trouvé une courbe elliptique de rang ≥ 24 :

$$y^2 \quad +xy + y = x^3 - 120039822036992245303534619191166796374x$$
$$+5042249924849106700108017991680827267594437562229114151 16$$

(voir `http ://www.math.hr/ duje/tors/rankhist.html` pour d'autres exemples).

Exemples.

1. Soit $\mathcal{C}$ donnée par l'équation

$$y^2 + y = x^3 - x.$$

dont les solutions en nombre entiers donnent la liste des cas où le produit de deux entiers consecutifs est égal au produit de trois entiers consecutifs. Alors le groupe Δ est trivial et $\mathcal{C}(\mathbb{Q})$ est cyclique de générateur $P = (0, 0)$.

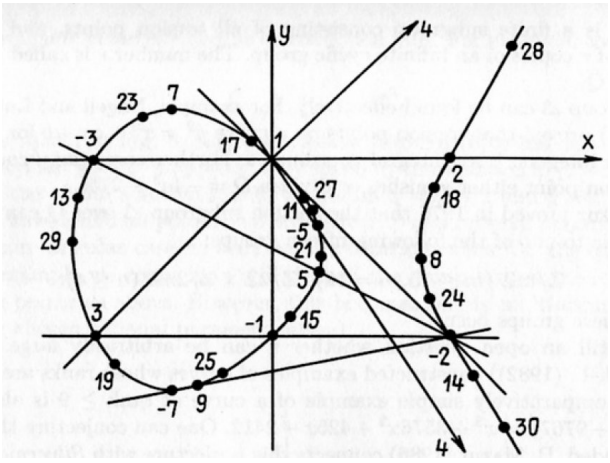


Fig. 9

2. Les points mP (numérotés par m) sont montrés dans le Fig. 9.
3. Soit $\mathcal{C}$ donnée par l'équation

$$y^2 + y = x^3 - 7x + 6.$$

Alors $\mathcal{C}(\mathbb{Q}) \cong \mathbb{Z}^3$, et les points (1,0), (0,2) forment une base de ce groupe.

4. Considérons la courbe $y^2 = x^3 + px,\ p = 877$. Un générateur modulo torsion du groupe des points rationnels de cette courbe a l'abscisse x

$$x = \frac{3754945281271621931055040699420927923462 01}{6215987776871505425463220780697238044100}.$$

Cet exemple montre que les méthodes naives de recherche des points rationnels deviennent rapidement inefficaces.

Congruences cubiques modulo un nombre premier p

Soit p un nombre premier et soit $F(X_0, X_1, X_2)$ une forme cubique à coefficients entiers. La réduction de F modulo p, donne une forme cubique sur le corps fini $\mathbb{F}_p$. Cette reduction est dit non-singulère si elle n'a pas de zéros communs avec ces dérivées partielles dans toute extension de $\mathbb{F}_p$. On peut montrer que la plupart des résultats de la géométrie algébrique complexe restent valables sur les corps de caractéristique positive. Cependant, les formes normales d'une courbe elliptique sont légérement plus compliquées. En utilisant un changement de variables inversibles des coordonneés projectives, on peut toujours réduire l'équation $F = 0$ en coordonnées affines pour $p \neq 2, 3$ à la forme de l'un des types suivantes (Koblitz N. (1987)) :

$$y^2 = x^3 + ax + b\ (4a^3 + 27b^2 \neq 0),\ a, b \in \mathbb{F}_p. \tag{16.8}$$

(on exclut le cas des racines multiples).

La courbe projective ainsi défini possède toujours le point $O = (0 : 1 : 0)$, rationnel sur $\mathbb{F}_p$ (rappelons que l'ensemble $\mathcal{C}(\mathbb{F}_p)$ des points rationnels sur $\mathbb{F}_p$ d'une courbe projective $\mathcal{C}\ : F(X, Y, Z) = 0$ est le sous-ensemble de $\mathbb{P}^2_{\mathbb{F}_p}$ des points

$$\Big\{(X : Y : Z) \in \mathbb{P}^2_{\mathbb{F}_p} \mid F(X, Y, Z) = 0\Big\}$$

Combien de points sur $\mathbb{F}_p$, c'est-à-dire, combien de solutions projectives de la congruence $F \equiv 0 \bmod\ p$, peut-on avoir ? Bien evidemment, le nombre total est au plus $2p + 1$ (on compt O), puisque sous la forme affine tout point fini x donne au plus deux valeurs de y. D'un autre coté, seulement une moitié des classes résiduelles sont des carrés (pour p impair). Donc on peut espérer que $x^3 + ax + b$ est un carré pour environ la moitié des x'.

Plus précisemment, soit $\chi(x) = \left(\frac{x}{p}\right)$ le symbole de Legendre. Alors, on a par définition, que le nombre de solutions de $y^2 = u$ dans $\mathbb{F}_p$ est $1 + \chi(u)$. Ceci implique,

$$\begin{aligned} \text{Card}\ \ \mathcal{C}(\mathbb{F}_p) &= 1 + \sum_{x \in \mathbb{F}_p} (1 + \chi(x^3 + ax + b)) \\ &= p + 1 + \sum_{x \in \mathbb{F}_p} \left(\frac{x^3 + ax + b}{p}\right). \end{aligned}$$

N. Koblitz (1987) compare la dernière somme avec un résultat sur les marches aléatoires sur une droite. Après p marches on peut espérer être à une distance d'environ $\sqrt{p}$ de zéro. En fait, on peut démontrer le théorème remarquable suivant :

THÉORÈME 16.2.2 (H.HASSE (1937)) *Soit* $N_p = \text{Card } \mathcal{C}(\mathbb{F}_p)$. *Alors*

$$|N_p - (p+1)| \leq 2\sqrt{p}.$$

Une preuve élémentaire a été donné par Yu.I.Manin (1956).

REMARQUE Les courbes elliptiques sur les corps finis trouvent beaucoup d'applications. En particulier, les cas où un tel groupe est cyclique de grande taille amènent aux cryptosystèmes ECDLP ("Elliptic curve discrete logarithm problem").

EXERCICES

16.1 On considère la courbe affine sur $\mathbb{Q}$ donnée par l'équation

$$\mathcal{C} : 2x^2 + 3y^2 = 5.$$

(a) Montrer qu'il existe un point rationnel $(x_0, y_0) \in \mathcal{C}(\mathbb{Q})$.
(b) Soit $t \in \mathbb{Q}$. Trouver tous les points d'intersection de la droite

$$y - y_0 = t(x - x_0)$$

avec la courbe $\mathcal{C}(\mathbb{Q})$.
(c) En déduire une parametrisation rationnelle de la courbe $\mathcal{C}$.
(d) Trouver tous les points rationnels de la courbe projective, donnée par l'équation homogène $2X^2 + 3Y^2 - 5Z^2 = 0$.

16.2 On considère la courbe affine sur $\mathbb{F}_{49}$, donnée par l'équation

$$\mathcal{C} : 2x^2 + 3y^2 = 5.$$

(a) Montrer qu'il existe un point rationnel $(x_0, y_0) \in \mathcal{C}(\mathbb{F}_{49})$.
(b) Soit $t \in \mathbb{F}_{49}$. Trouver tous les points d'intersection de la droite

$$y - y_0 = t(x - x_0)$$

avec la courbe $\mathcal{C}(\mathbb{F}_{49})$.
(c) En déduire une parametrisation rationnelle sur $\mathbb{F}_{49}$ de la courbe $\mathcal{C}$.
(d) Trouver tous les points rationnels sur $\mathbb{F}_{49}$ de la courbe projective donnée par l'équation homogène $2X^2 + 3Y^2 - 5Z^2 = 0$.

16.3 On considère la courbe cubique plane $\mathcal{E} \subset \mathbb{P}^2$ sur $\mathbb{Q}$, donnée sous la forme affine suivante $y^2 = x(x+9)(x-16)$.
(a) Trouver l'equation de $\mathcal{E}$ en coordonnées projectives.
(b) Montrer que la courbe $\mathcal{E}$ est lisse.
(c) Montrer que l'ensemble

$$G = \{(\infty, \infty), (0,0), (0,-9), (0,16), (-4,20), (-4,-20), (36,180), (36,-180)\} \subset \mathbb{P}^2$$

est un sous-groupe de $\mathcal{E}(\mathbb{Q})$.
(d) Le groupe G est-il cyclique ?

16.4 On considère une courbe cubique plane $\mathcal{F} \subset \mathbb{P}^2$ sur $\mathbb{F}_4$, donnée sous la forme affine suivante $y^2 + y = x^3 + x + 1$.
(a) Montrer que la courbe $\mathcal{F}$ est lisse.
(b) Trouver l'ordre du groupe $H = \mathcal{F}(\mathbb{F}_4)$.
(c) Le groupe H est-il cyclique ?

Sixième partie

Annexes

A Exercices

A.1 Examen du mardi 15 mai 2007, 9h–12h, AMPHI A

Le sujet est constitué de trois exercices et d'un problème tous indépendants.

Exercice I.

1. Trouver le nombre des polynômes unitaires irréductibles de degré 5 sur $\mathbb{F}_3$.
2. Trouver le produit de tous les polynômes unitaires irréductibles de degré 5 sur $\mathbb{F}_3$.

Exercice II.

Montrer que les extensions de corps $\mathbb{Q}(\sqrt{2},\sqrt{3})$ et $\mathbb{Q}(\sqrt{2}+\sqrt{3})$ sont égales.

Exercice III.

Dans $\mathbb{P}^2_{\mathbb{C}}$ de coordonnées homogènes $[X:Y:Z]$ on identifie le plan $\mathbb{C}^2$ de coordonnées x, y à la carte $Z \neq 0$.

Soient $C_1 = \{y^2 - x = 0\}$, $C_2 = \{xy - 1 = 0\}$, et $\bar{C_1}$, $\bar{C_2}$ les courbes projectives correspondantes.

1. Ecrire les équations homogènes qui définissent $\bar{C_1}$ et $\bar{C_2}$ et déterminer $\bar{C_1} \cap \bar{C_2}$.
2. Soit $p = [0:0:1]$ et $\bar{D} = \{Z = 0\}$ la droite à l'infini. On considère l'application

$$f : [a:b:0] \in \bar{D} \to [f_1 : f_2 : f_3] \in \bar{C_1}$$

   qui à tout point $q \in \bar{D}$ associe le point d'intersection de $\bar{C_1}$ avec la droite passant par p et q.

   Expliciter f_1, f_2 et f_3 et montrer que f est bijective.
3. Existe-t-il un automorphisme g de $\mathbb{P}^2_{\mathbb{C}}$ tel que $g(\bar{C_1}) = \bar{D}$? (Si oui expliciter g, si non donner un court argument).
4. Existe-t-il un automorphisme g de $\mathbb{P}^2_{\mathbb{C}}$ tel que $g(\bar{C_1}) = \bar{C_2}$? (Si oui expliciter g, si non donner un court argument).

Problème.

Soit k un corps commutatif. Le but de ce problème est de montrer la simplicité de $\mathrm{PSL}_n(k)$ pour $n \geqslant 3$ par une méthode due à Iwasawa.

(A) Résultats préliminaires sur les commutateurs

Pour tout groupe G, on note ici $D(G)$ le groupe des commutateurs de G, c'est-à-dire le sous-groupe de G engendré par les éléments de la forme $aba^{-1}b^{-1}$ pour $a, b \in G$. Un élément $u \in \mathrm{SL}_n(k)$, est dit une transvection, si $(u-I)^2 = 0$ et si l'image $\mathrm{Im}(u-I)$ de l'endomorphisme de k^n associé à $u-I$ est une droite $\langle x \rangle$, dite *droite de transvection*. Le noyau $\mathrm{Ker}\ (u-I)$ de l'endomorphisme $u-I$ est un hyperplan V_u de k^n (*base de la transvection*). L'objectif de cette partie est de montrer que $D(\mathrm{SL}_n(k)) = \mathrm{SL}_n(k)$ si $n \geqslant 3$.

1. On admettra que les transvections engendrent le groupe $\mathrm{SL}_n(k)$. Trouver la forme normale d'une transvection et montrer que les transvections sont toutes conjuguées entre elles pour $n \geq 3$. En déduire qu'il suffit de montrer qu'il existe un commutateur $aba^{-1}b^{-1}$ qui est une transvection.
2. Si k n'est pas de caractéristique 2, soit u une transvection. Montrer que u et u^2 sont deux transvections conjuguées dans $\mathrm{SL}_n(k)$. En déduire que u est un commutateur.
3. Si k est de caractéristique 2, considérer les matrices
$$u = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}, t = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, s = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$
et montrer que u est une transvection et un commutateur.

(B) La méthode d'Iwasawa

Soit G un groupe opérant sur un ensemble X. On suppose que l'action de G sur X est *doublement transitive* : si (x_1, x_2) et (y_1, y_2) sont deux couples de points distincts ($x_1 \neq x_2$ et $y_1 \neq y_2$) de X, alors il existe un $G \in G$ tel que $y_1 = gx_1$ et $y_2 = gx_2$.

Fixons un $x_0 \in X$ et notons H le stabilisateur de x_0 dans G.

1. Montrer que H agit transitivement sur $X \setminus \{x_0\}$.
2. Montrer que $g \mapsto gx_0$ induit une bijection de G/H sur X.
3. Montrer que pour tout $g \in G \setminus H$, on a $G = H \cup HgH$.
4. Montrer que H est maximal parmis les sous-groupes propres de G.
5. Soit N un sous-groupe distingué de G. Posons
$$NH = \{nh \in G; n \in N, h \in H\}.$$
Montrer que NH est un sous-groupe de G. Déduire de la question précédente que N opère transitivement ou trivialement sur X.
6. On suppose que l'on a une famille $(T_x)_{x \in X}$ de sous-groupes abéliens de G, tels que $T_{gx} = gT_xg^{-1}$ et tels que les T_x engendrent G. Soit N un sous-groupe distingué de G qui n'opère pas trivialement sur X. Montrer que l'on a $G = NT_x$.
7. Sous les hypothèses de la question précédente, montrer que N contient le groupe des commutateurs de G.

(C) Conclusion

Montrer que le groupe projectif spécial $\mathrm{PSL}_n(k)$ est simple si $n \geqslant 3$, où $\mathrm{PSL}_n(k) = \mathrm{SL}_n(k)/H_n$, $H_n = k^* I_n \cap \mathrm{SL}_n(k)$ est le sous-groupe des homothéties.

On pourra appliquer les résultats de la partie précédente (en particulier les questions 6 et 7) avec $G = \mathrm{PSL}_n(k)$, $X = \mathbb{P}^{n-1}(k)$ (identifié à l'ensemble des droites vectorielles de k^n).

Pour définir T_x pour tout $x \in X$, on note $e_0, \ldots, e_{n-1}$ la base canonique de k^n et on définit d'abord T_{ke_0} comme le groupe formé des transformations, représentées en coordonnées projectives par des matrices

$$\begin{pmatrix} 1 & \lambda_1 & \lambda_2 & \cdots & \lambda_{n-1} \\ 0 & 1 & 0 & \cdots & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & \cdots & 0 & 1 & 0 \\ 0 & \cdots & 0 & 0 & 1 \end{pmatrix}$$

dans la base $e_0, \ldots, e_{n-1}$ (avec $\lambda_1, \ldots, \lambda_{n-1} \in k$), puis on en déduit les autres T_x par conjugaison : si l'action de $g \in G$ envoie ke_0 (vu comme un élément de X) sur x, on pose $T_x = gT_{ke_0}g^{-1}$. (On montrera que le T_x ainsi défini ne dépend pas du choix de g et que T_x est un sous-groupe commutatif de G).

A.2 Corrigé de l'examen du mardi 15 mai 2007

Exercice I.

1. On a vu dans le cours la formule générale donnant le nombre de polynômes irréductibles unitaires de degré n sur $\mathbb{F}_q$:

$$\frac{1}{n}\sum_{d|n} \mu(n/d) q^d.$$

   Ici cette formule donne $\frac{1}{5}(\mu(5).3 + \mu(1).3^5) = \frac{3(3^4-1)}{5} = 48$.

2. On sait que

$$X^{3^5} - X = \prod_{\text{irréd deg 1}} P(X) \prod_{\text{irréd deg 5}} P(X).$$

   Ainsi le produit des polynômes irréductibles sur $\mathbb{F}_3$ de degré 5 est égal à

$$\frac{X^{3^5} - X}{X(X-1)(X+1)}.$$

Exercice II.

Rédigé par Stéphane Lamy et Lionel Fourquaux

D'une part $\sqrt{2}+\sqrt{3} \in \mathbb{Q}(\sqrt{2},\sqrt{3})$ ainsi clairement $\mathbb{Q}(\sqrt{2}+\sqrt{3}) \subset \mathbb{Q}(\sqrt{2},\sqrt{3})$. Pour avoir égalité il suffit de montrer l'égalité des dimensions comme espace vectoriel sur $\mathbb{Q}$. $\mathbb{Q}(\sqrt{2},\sqrt{3})$ étant clairement de dimension 4, il s'agit de démontrer que le polynôme minimal de $\sqrt{2}+\sqrt{3}$ est de degré 4. On calcule $(\sqrt{2}+\sqrt{3})^2 = 5+2\sqrt{6}$ et $(\sqrt{2}+\sqrt{3})^4 = 49+20\sqrt{6}$. On voit donc que le polynôme X^4-10X^2+1 annule $\sqrt{2}+\sqrt{3}$, reste à voir que ce polynôme est irréductible (sur $\mathbb{Z}$, ce qui implique l'irréductibilité sur $\mathbb{Q}$). Il n'admet pas de racines dans $\mathbb{Z}$ (les seuls candidats sont ± 1 à cause du coefficient constant), donc ne peut admettre qu'une décomposition sur $\mathbb{Z}$ de la forme

$$X^4 - 10X^2 + 1 = (X^2 + aX + b)(X^2 + cX + d)$$

avec $bd = 1$ et $a + c = 0$. En identifiant les termes en X^2 on trouve $10 = a^2 - 2b$, or $10 + 2b$ n'est pas un carré pour $b = \pm 1$, ce qui exclut une telle décomposition.

Exercice III.

1. $\bar{C_1} = \{Y^2 - XZ = 0\}$ et $\bar{C_2} = \{XY - Z^2 = 0\}$. On cherche les points $[X : Y : Z]$ dans l'intersection $\bar{C_1} \cap \bar{C_2}$. Si $Z \neq 0$ on peut supposer $Z = 1$ et on trouve les trois points $[1 : 1 : 1], [j : j^2 : 1]$ et $[j^2 : j : 1]$ où $j = e^{2i\pi/3}$. Si $Z = 0$ on trouve le quatrième point d'intersection $[1 : 0 : 0]$.
2. Soit $q = [a : b : 0]$ un point de $\bar{D}$, l'équation $aY - bX = 0$ est celle de la droite passant par $p = [0 : 0 : 1]$ et q, on doit donc résoudre le système $\begin{cases} aY - bX &= 0 \\ Y^2 - XZ &= 0 \end{cases}$ Cherchons d'abord les solutions avec $Z = 0$: il n'existe une telle solution que si $b = 0$, et dans ce cas $[1 : 0 : 0]$ est solution. On peut maintenant supposer $b \neq 0$ et chercher des solutions avec $Z = 1$. On obtient

   $$\begin{cases} aY - bX &= 0 \\ Y^2 - X &= 0 \end{cases} \Longrightarrow aY - bY^2 = 0 \Longrightarrow Y = 0 \text{ ou } Y = a/b.$$

   $Y = 0$ correspond au point $p = [0 : 0 : 1]$ évidemment solution, et $Y = a/b$ correspond au point cherché, ainsi

   $$f : [a : b : 0] \mapsto [\frac{a^2}{b^2} : \frac{a}{b} : 1] = [a^2 : ab : b^2]$$

   et cette dernière expression inclut aussi le cas $b = 0$. f est bijective : on vérifie que f^{-1} est définie par

   $$f^{-1} : [X : Y : Z] \in \bar{C_1} \mapsto \begin{cases} [X : Y : 0] \text{ si } [X : Y : Z] \neq [0 : 0 : 1] \\ [0 : 1 : 0] \text{ sinon.} \end{cases}$$

3. Les automorphismes de $\mathbb{P}^2_{\mathbb{C}}$ sont donnés par trois polynômes homogènes de degré 1, en particulier l'image réciproque d'une droite est une droite et donc on ne peut pas avoir $g(\bar{C_1}) = \bar{D}$
4. Considérons $g : [X : Y : Z] \to [X : -Z : iY]$, on vérifie immédiatement que la préimage de $\bar{C_2}$ par l'automorphisme g est bien $\bar{C_1}$.

Problème.

(A) Résultats préliminaires sur les commutateurs

1. La forme normale d'une transvection est

$$\begin{pmatrix} 1 & 1 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & \cdots & 0 & 1 & 0 \\ 0 & \cdots & 0 & 0 & 1 \end{pmatrix}$$

Les transvections sont toutes conjuguées entre elles pour $n \geq 3$, ainsi si une transvection s'écrit comme un commutateur alors en conjuguant on obtient que toute transvection appartient au groupe dérivé, et comme les transvections engendrent $\mathrm{SL}_n(k)$ on obtient $D(\mathrm{SL}_n(k)) = \mathrm{SL}_n(k)$.

2. On se place dans une base où u et u^2 s'écrivent

$$u = \begin{pmatrix} 1 & a & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & 0 & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{pmatrix} \qquad u^2 = \begin{pmatrix} 1 & 2a & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & 0 & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{pmatrix}$$

On vérifie que $u^2 = dud^{-1}$ où $d = \begin{pmatrix} 2 & 0 & \cdots & 0 \\ 0 & 1/2 & \cdots & 0 \\ \vdots & 0 & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{pmatrix}$ et donc $u = dud^{-1}u^{-1}$ est un commutateur.

3. On vérifie que u est une transvection en calculant $\begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}^2 = 0$ D'autre part on remarque que t et s sont des involutions et on calcule

$$tst^{-1}s^{-1} = (ts)^2 = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}^2 = u$$

(B) La méthode d'Iwasawa

1. Soient $x, y \in X \setminus \{x_0\}$. On choisit deux points auxiliaires z_1 et z_2 distincts de x, y et x_0. Il existe g_1 tel que $g_1.x = z_1$, $g_1.x_0 = z_2$, et il existe g_2 tel que $g_2.z_1 = y$, $g_2.z_2 = x_0$. La composée g_2g_1 est un élément de H (car fixant x_0) qui envoie x sur y.
2. L'application $g \in G \mapsto gx_0 \in X$ est surjective par hypothèse de transitivité, et donc induit une bijection $G/\sim \mapsto X$ ou $\sim$ est la relation d'équivalence : $g_1 \sim g_2$ si $g_1x_0 = g_2x_0$. Mais $g_1x_0 = g_2x_0 \Rightarrow g_1^{-1}g_2 \in H \Rightarrow g_1H = g_2H$. Ainsi $G/\sim = G/H$, et on a une bijection de G/H sur X.

3. Soit $f \in G$, on veut montrer $f \in H \cup HgH$. Supposons $f \notin H$, par la question (1) il existe $h \in H$ tel que $f.x_0 = hg.x_0$. Ainsi $g^{-1}h^{-1}f$ fixe x_0 et est donc égal à $h' \in H$, finalement $f = hgh' \in HgH$ ce qu'on voulait.
4. Si K est un sous-groupe de G contenant strictement H, soit $g \in K \setminus H$. Alors K contient H et HgH, et par la question précédente $K = G$.
5. Soit nh et $n'h'$ deux éléments de NH. On a

$$nhn'h' = n(hn'h^{-1})hh' \in NH$$

car $hn'h^{-1} \in N$, de même

$$(nh)^{-1} = h^{-1}n^{-1}hh^{-1} \in NH.$$

Ainsi NH est un sous-groupe de G.

Supposons que N n'agisse pas trivialement sur X. Il existe $n \in N$ et $x_0 \in X$ tel que $n.x_0 \neq x_0$. En définissant H comme précédemment, on en déduit que NH est un sous-groupe de G qui contient strictement H, par la question (4) ceci implique $N = G$ et donc N agit transitivement sur X.
6. Soit $y \in X$, par la question précédente N agit transitivement sur X, il existe donc $n \in N$ tel que $n.y = x$. On a donc $T_y = nT_xn^{-1}$. Si $t' \in T_y$, on a $t' = ntn^{-1} = n(tn^{-1}t^{-1})t$ avec $t \in T_x$, ainsi $T_y \subset NT_x$. Finalement comme les T_y engendrent G, on a bien $G = NT_x$.
7. Soient $g_1, g_2 \in G$, par la question précédente on peut les écrire $g_1 = n_1t_1, g_2 = n_2t_2$ avec $n_i \in N, t_i \in T_x$. On a

$$\begin{aligned} g_1g_2g_1^{-1}g_2^{-1} &= n_1t_1n_2t_2t_1^{-1}n_1^{-1}t_2^{-1}n_2^{-1} \\ &= n_1(t_1n_2t_1^{-1})(t_1t_2t_1^{-1})n_1^{-1}t_2^{-1}n_2^{-1} \\ &= n_1(t_1n_2t_1^{-1})(t_2n_1^{-1}t_2^{-1})n_2^{-1} \\ &\in N \end{aligned}$$

(C) Conclusion

Vérifions que $\mathrm{PSL}_n(k)$ agit doublement transitivement sur $\mathbb{P}^{n-1}(k)$. Si $(x_1, x_2), (y_1, y_2)$ sont deux couples de points distincts dans $\mathbb{P}^{n-1}(k)$, ils correspondent à deux couples de points non-colinéaires sur $k^n \setminus \{0\}$, et peuvent donc chacun être complété en une base de k^n. Reste à choisir un élément de $GL(n, k)$ qui envoie la première base sur la deuxième (ce qui revient à se donner les deux premières colonnes de la matrice).

On déduit de la partie (B) que si N est un sous-groupe distingué non trivial de $\mathrm{PSL}_n(k)$, alors N contient le groupe dérivé $D(\mathrm{PSL}_n(k)$ (en effet N agit non trivialement sur $\mathbb{P}^{n-1}(k)$, car tout élément non trivial de $\mathrm{PSL}_n(k)$ agit non trivialement).

Enfin, la partie (A) permet d'affirmer que $N = \mathrm{PSL}_n(k)$.

A.3 Contrôle continu du mardi 13 mars 2007

Exercice 1 – Soit M un module libre de rang 4 sur l'anneau $A = \mathbb{Z}[X]$, et soit e_1, e_2, e_3, e_4 une base de M.

1. Calculer le produit extérieur
$$(e_1 + 2e_2 + 3e_3) \wedge (e_2 + 2e_3 + 3e_4) + (e_3 + 2e_4) \wedge (e_1 + 3e_4).$$
2. Montrer que le module $\wedge^2(M)$ est libre et en déterminer une base.
3. Calculer l'algébre symétrique de M. Montrer que $S(M)$ est isomorphe à
$$\mathbb{Z}[X_1, X_2, X_3, X_4, X_5].$$
4. Calculer l'algébre extérieure $\wedge(M)$ de M.

Exercice 2 – Soient n un entier naturel et P le polynôme
$$P(X) = X(X+1)\cdots(X+n-1) - 1$$
Supposons qu'il existe A, B non constants dans $\mathbb{Z}[X]$ tels que $P = AB$.

1. Montrer en évaluant P en des valeurs bien choisies que'on a $A + B = 0$.
2. En déduire que P est irréductible sur $\mathbb{Z}[X]$.
3. Montrer que P est irréductible dans $\mathbb{Q}[X]$ (on attend une justification en deux lignes maximum).
4. En s'inspirant de ce qui précède, montrer que $P + 2$ est irréductible dans $\mathbb{Q}[X]$ si $n \neq 4$ [on montrera d'abord que, dans le cas contraire, $P + 2$ serait un carré].
5. Factoriser $P + 2$ dans $\mathbb{Q}[X]$ en facteurs irréductibles si $n = 4$.

Exercice 3 – On suppose $x + y + z = 1$, $x^2 + y^2 + z^2 = 2$ et $x^3 + y^3 + z^3 = 3$. Calculer $x^4 + y^4 + z^4$.

Exercice 4 – (a) Exprimer le polynôme

$$P = (x_1+x_2-x_3-x_4)(x_1-x_2+x_3-x_4)(x_1-x_2-x_3+x_4)(x_1+x_2+x_3+x_4)+x_1x_2x_3x_4$$

à l'aide des polynômes symétriques élémentaires s_i.

(b) Montrer que le polynôme $f(X) = X^4 + 4X^3 + 6X^2 + 8X + 2$ possède 4 racines complexes distinctes $\alpha_1, \alpha_2, \alpha_3, \alpha_4$.

(c) Trouver la valeur $P(\alpha_1, \alpha_2, \alpha_3, \alpha_4)$.

Exercice 5 – Décomposer en élements simples la fraction rationnelle sur $\mathbb{Q}$:
$$\frac{X-7}{X^4 + X^3 - 3X^2 - X + 2}.$$

Exercice 6 – Soient k un corps et $n \in \mathbb{N}^*$.

Si $\sigma \in \mathfrak{S}_n$ est une permutation de $\{1, \ldots, n\}$, on définit une action de $\mathfrak{S}_n$ sur $k[X_1, \ldots, X_n]$ par
$$P^\sigma(X_1, \ldots, X_n) = P(X_{\sigma(1)}, \ldots, X_{\sigma(n)}).$$

1. Montrer que l'application $P \mapsto P^\sigma$ se prolonge en un automorphisme de la k-algèbre $k(X_1, \ldots, X_n)$, et que l'on définit ainsi une action de $\mathfrak{S}_n$ sur $k(X_1, \ldots, X_n)$.

Notons $k(X_1, \ldots, X_n)^{\mathrm{sym}}$ l'ensemble des fractions rationnelles symétriques, c'est-à-dire des élements de $k(X_1, \ldots, X_n)$ fixes sous l'action de $\mathfrak{S}_n$.

Le but de cet exercice est de prouver que toute fraction rationnelle symétrique s'écrit comme quotient de deux polynômes symétriques.

Soit $P/Q \in k(X_1, \ldots, X_n)^{\mathrm{sym}}$ (avec $P, Q \in k[X_1, \ldots, X_n]$).

2. Pour $\sigma \in \mathfrak{S}_n$, montrer que si $P^\sigma = P$ alors $Q^\sigma = Q$.
3. Montrer que l'on peut supposer que P et Q sont premiers entre eux.
4. Supposons que Q n'est pas symétrique. Montrer qu'il existe une transposition $\tau = (i, j)$ qui ne fixe pas Q.
5. Montrer que $X_i - X_j$ divise $P^\tau - P$ et $Q^\tau - Q$.
6. Montrer $(Q^\tau - Q)P = (P^\tau - P)Q$. En déduire qu'il existe une constante $\lambda \in k^*$ telle que $P = \lambda(P^\tau - P)$ et $\lambda(Q^\tau - Q)$.
7. Conclure.

A.4 Corrigé du partiel du mardi 13 mars 2007*

Exercice 1

On a

$$\begin{aligned}(e_1 + 2e_2 + 3e_3) \wedge (e_2 + 2e_3 + 3e_4) + (e_3 + 2e_4) \wedge (e_1 + 3e_4)& \\ = e_1 \wedge e_2 + 2e_1 \wedge e_3 + 3e_1 \wedge e_4 + 4e_2 \wedge e_3 + 6e_2 \wedge e_4 - 3e_2 \wedge e_3 + 9e_3 \wedge e_4& \\ - e_1 \wedge e_3 + 3e_3 \wedge e_4 - 2e_1 \wedge e_4& \\ = e_1 \wedge e_2 + e_1 \wedge e_3 + e_1 \wedge e_4 + e_2 \wedge e_3 + 6e_2 \wedge e_4 + 12e_3 \wedge e_4.&\end{aligned}$$

Comme M est isomorphe à A^4, il suffit d'étudier $\bigwedge^2(A^4)$. Considérons les morphismes de A-modules définis par

$$\begin{cases} \bigwedge^2(A^4) & \overset{(\text{« développement »})}{\longrightarrow} & A^6 \\ (x_1, x_2, x_3, x_4) \wedge (y_1, y_2, y_3, y_4) & \longmapsto & (x_1y_2 - x_2y_1, x_1y_3 - x_3y_1, x_1y_4 - x_4y_1, \\ & & x_2y_3 - x_3y_2, x_2y_4 - x_4y_2, x_3y_4 - x_4y_3) \end{cases}$$

et

$$\begin{cases} A^6 & \overset{(\text{« plongements »})}{\longrightarrow} & \bigwedge^2(A^4) \\ (z_1, z_2, z_3, z_4, z_5, z_6) & \longmapsto & (1,0,0,0) \wedge (0, z_1, z_2, z_3) + (0,1,0,0) \wedge (0,0,z_4,z_5) \\ & & +(0,0,1,0) \wedge (0,0,0,z_6). \end{cases}$$

Ces deux morphismes sont clairement inverses l'un de l'autre, donc $\bigwedge^2(A^4)$ est isomorphe à A^6.

L. Fourquaux

Dans le cas général, on en déduit que $\bigwedge^2(M)$ est libre de rang 6, et qu'une base de $\bigwedge^2(M)$ est :

$$e_1 \wedge e_2,\ e_1 \wedge e_3,\ e_1 \wedge e_4,\ e_2 \wedge e_3,\ e_2 \wedge e_4,\ e_3 \wedge e_4.$$

Comme M est un module libre de rang 4, de base e_1, e_2, e_3, e_4, le produit symétrique $S^n(M)$ est isomorphe au module des polynômes homogènes de degré n en e_1, e_2, e_3, e_4. La $\mathbb{Z}[X]$-algèbre $S(M)$ est donc isomorphe à l'algèbre des polynômes en e_1, e_2, e_3, e_4 sur $A = \mathbb{Z}[X]$, c'est-à-dire $\mathbb{Z}[X][e_1, e_2, e_3, e_4]$, qui est isomorphe à $\mathbb{Z}[X_1, X_2, X_3, X_4, X_5]$ (par exemple en envoyant X_i sur e_i pour $1 \leqslant i \leqslant 4$, et X_5 sur X).

Exercice 2

Notons que pour tout $x \in \{-(n-1), \ldots, -1, 0\}$, on a $P(x) = -1$, donc $A(x)B(x) = -1$. D'autre part, A et B sont à coefficients entiers, donc $A(x) \in \mathbb{Z}$ et $B(x) \in \mathbb{Z}$. Comme les seuls diviseurs entiers de -1 sont -1 et 1, on trouve :

$$(A(x), B(x)) \in \{(-1, 1), (1, -1)\},$$

donc $A(x) + B(x) = 0$.

Le polynôme $A + B$ a donc au moins n racines. Or $P = AB$ est de degré n, et A et B sont non constants, donc $A + B$ est de degré strictement inférieur à n. Donc $A + B = 0$.

On a donc $P = -A^2$, donc en particulier le coefficient du monôme de plus haut degré de P est strictement négatif, ce qui contredit la définition de P.

Donc de tels polynômes A et B n'existent pas, autrement dit P est irréductible dans $\mathbb{Z}[X]$.

Si P était réductible dans $\mathbb{Q}[X]$, alors il admettrait une factorisation non triviale $P = \widetilde{A}\widetilde{B}$, avec $\widetilde{A}$ et $\widetilde{B}$ dans $\mathbb{Q}[X]$. En posant $A = \frac{\widetilde{A}}{\mathrm{cont}(\widetilde{A})}$ et $B = \mathrm{cont}(\widetilde{A})\widetilde{B}$ (en notant cont le contenu, c'est-à-dire le pgcd de coefficients du polynôme), on aurait alors une factorisation $P = AB$ non triviale, dans $\mathbb{Z}[X]$, ce qui est impossible. Donc P est irréductible dans $\mathbb{Q}[X]$.

Si $P + 2$ était réductible dans $\mathbb{Q}[X]$, alors considérons une factorisation non triviale $P + 2 = CD$ dans $\mathbb{Q}[X]$. Quitte à diviser C par son contenu et à multiplier D par le même entier, on peut supposer que C et D sont dans $\mathbb{Z}[X]$.

Comme $P+2$ est unitaire, on peut supposer que C et D le sont (quitte à les multiplier tous les deux par -1).

Notons maintenant que pour tout $x \in \{-(n-1), \ldots, -1, 0\}$, on a $C(x)D(x) = P(x) + 2 = 1$. Comme C et D sont à coefficients entiers, on en déduit $C(x) = D(x) \in \{-1, 1\}$.

Comme $C - D$ est de degré strictement plus petit que n, on trouve donc $C = D$, et donc $P + 2 = C^2$, c'est-à-dire

$$(C - 1)(C + 1) = X(X + 1)\ldots(X + n - 1).$$

En particulier, n doit être pair, C est de degré $\frac{n}{2}$, et les racines de $C - 1$ et $C + 1$ sont deux parties complémentaires de $\{-(n-1), \ldots, -1, 0\}$.

Rappelons que l'on a $C(0) \in \{-1, 1\}$, et notons $\{\mathfrak{r} \subset \{-(n-1), \ldots, -1, 0\}$ l'ensemble des racines de $C - C(0)$.

Considérons le coefficient de degré 1 de $(C - C(0))(C + C(0))$. Il est égal à

$$2C(0) \prod_{x \in \mathfrak{r} \setminus \{0\}} (-x).$$

D'autre part, c'est aussi le coefficient de degré 1 de $X(X+1)\dots(X+n-1)$, c'est-à-dire $(n-1)!$. On trouve donc :

$$\begin{aligned} C(0) &= 1 \\ \prod_{x\in\mathfrak{r}\setminus\{0\}} (-x) &= \frac{(n-1)!}{2} \end{aligned}$$

La seconde égalité entraîne

$$\prod_{x\in\{-(n-1),\dots,-1,0\}\setminus\mathfrak{r}} (-x) = 2,$$

donc

$$\{-(n-1),\dots,-1,0\}\setminus\mathfrak{r} \subset \{-2,-1\},$$

donc $\frac{n}{2} \leqslant 2$ en considérant les cardinaux.

Dans le cas $n=2$, on trouve $C = X+1$, donc

$$(C-1)(C+1) = X(X+2) \neq X(X+1),$$

d'où une contradiction.

Donc pour $n \neq 4$, $P+2$ est irréductible dans $\mathbb{Q}[X]$.

Pour $n=4$, on doit avoir

$$\{-3,-2,-1,0\}\setminus\mathfrak{r} = \{-2,-1\},$$

donc $C-1 = C-C(0) = X(X+3)$, or, pour $C = X^2+3X+1$, on a bien

$$(C-1)(C+1) = X(X+1)(X+2)(X+3),$$

donc $P+2 = C^2 = (X^2+3X+1)^2$.

Exercice 3

Notons

$$\begin{aligned} \sigma_1 &= x+y+z \\ \sigma_2 &= xy+xz+yz \\ \sigma_3 &= xyz \end{aligned}$$

les polynômes symétriques élémentaires en x, y, z.

On a

$$\begin{aligned} x^2+y^2+z^2 &= \sigma_1^2 - 2\sigma_2 \\ x^3+y^3+z^3 &= \sigma_1(x^2+y^2+z^2) - \sigma_2(x+y+z) + 3\sigma_3 \\ x^4+y^4+z^4 &= \sigma_1(x^3+y^3+z^3) - \sigma_2(x^2+y^2+z^2) + \sigma_3(x+y+z), \end{aligned}$$

donc

$$\begin{aligned}\sigma_1 &= 1\\ \sigma_2 &= -\frac{1}{2}\\ \sigma_3 &= \frac{1}{6}\\ x^4+y^4+z^4 &= \frac{25}{6}.\end{aligned}$$

Exercice 4

On a

$$X^4+X^3-3X^2-X+2=(X-1)^2(X+1)(X+2).$$

Comme

$$\begin{aligned}\left.\frac{X-7}{(X+1)(X+2)}\right|_{X=1} &= -1\\ \left.\frac{X-7}{(X-1)^2(X+2)}\right|_{X=-1} &= -2\\ \left.\frac{X-7}{(X-1)^2(X+1)}\right|_{X=-2} &= 1,\end{aligned}$$

on a

$$\frac{X-7}{X^4+X^3-3X^2-X+2}=\frac{-1}{(X-1)^2}+\frac{c}{X-1}+\frac{-2}{X+1}+\frac{1}{X+2}.$$

Comme

$$\left(\frac{X-7}{X^4+X^3-3X^2-X+2}-\frac{-1}{(X-1)^2}-\frac{-2}{X+1}-\frac{1}{X+2}\right)\Big|_{X=0}=-1,$$

on trouve $c=1$, donc

$$\frac{X-7}{X^4+X^3-3X^2-X+2}=\frac{-1}{(X-1)^2}+\frac{1}{X-1}+\frac{-2}{X+1}+\frac{1}{X+2}.$$

Exercice 5

Soit $f=\frac{P}{Q}\in k(X_1,\dots,X_n)$ et soit $\sigma\in\mathfrak{S}_n$. On définit l'image de f sous σ comme $f^\sigma=\frac{P^\sigma}{Q^\sigma}$. Pour cela, il faut montrer que f^σ ne dépend pas du choix de P et Q.

Si $R\in k[X_1,\dots,X_n]$, on a $(PR)^\sigma=P^\sigma R^\sigma$ et $(QR)^\sigma=Q^\sigma R^\sigma$, donc $\frac{P^\sigma}{Q^\sigma}=\frac{(PR)^\sigma}{(QR)^\sigma}$, donc la définition de f^σ est bien défini.

Comme $f^{\sigma\sigma^{-1}}=f=f^{\sigma^{-1}\sigma}$, l'application $f\mapsto f^\sigma$ est inversible, d'inverse $f\mapsto f^{\sigma^{-1}}$.

Comme $f\mapsto f^\sigma$ est k-linéaire et compatible à la multiplication, c'est donc un automorphisme de $k(X_1,\dots,X_n)$.

Comme $(f^\sigma)^\tau=f^{\tau\sigma}$, on a bien une action de $\mathfrak{S}_n$ sur $k(X_1,\dots,X_n)$.

On considère maintenant une fraction rationnelle symétrique $\frac{P}{Q}$.

Si $P^\sigma = P$, on a $\left(\frac{P}{Q}\right)^\sigma Q^\sigma = \frac{P}{Q}Q$, donc $P = 0$ ou $Q^\sigma = Q$, puisque $k(X_1, \ldots, X_n)$ est intègre.

Si P et Q ont un facteur commun, quitte à diviser par se facteur, on se ramène au cas où ils n'ont pas de facteur commun non trivial.

Si Q n'est pas symétrique, comme les transpositions engendrent $\mathfrak{S}_n$, il existe une transposition τ telle que $Q^\tau \neq Q$.

Si $\tau = (i, j)$, avec $1 \leqslant i < j \leqslant n$, considérons $P^\tau - P$ (ou $Q^\tau - Q$) comme un polynôme en X_j, à coefficients dans l'anneau $k[X_1, \ldots, X_{j-1}, X_{j+1}, \ldots, X_n]$. on a $(P^\tau - P)(X_j) \equiv (P^\tau - P)(X_i) \pmod{X_i - X_j}$, or $P^\tau(X_i) = P(X_i)$, donc $X_i - X_j$ divise $P^\tau - P$ (et $Q^\tau - Q$, pour la même raison).

Comme $\left(\frac{P}{Q}\right)^\tau = \frac{P}{Q}$, on a $P^\tau Q = PQ^\tau$, donc $(P^\tau - P)Q = (Q^\tau - Q)P$. En particulier, P divise $(P^\tau - P)Q$ dans l'anneau factoriel $k[X_1, \ldots, X_n]$. Comme P et Q n'ont aucun facteur commun, on trouve que P divise $(P^\tau - P)$.

Le degré en X_m de $\frac{P^\tau - P}{P}$ est inférieur ou égal à 0, pour tout m, donc $\lambda = \frac{P^\tau - P}{P} \in k$.

On a alors $P^\tau - P = \lambda P$ et $Q^\tau - Q = \lambda Q$. Comme $Q \neq Q^\tau$, on a aussi $\lambda \in k^*$.

Comme $X_i - X_j$ divise $(P^\tau - P)$ et $(Q^\tau - Q)$, c'est un facteur commun à P et Q, d'où une contradiction.

Donc Q est symétrique, et donc P l'est aussi d'après la deuxième question.

A.5 Contrôle continu du mardi 14 mars 2006

1. Soient p un nombre premier, et $\mathbb{F}_{p^{10}}$ un corps de p^{10} éléments.

(a) Trouver tous les sous corps $F \subset \mathbb{F}_{p^{10}}$.

(b) Montrer que le groupe $\mathrm{Aut}(\mathbb{F}_{p^{10}})$ de tous les automorphismes de $\mathbb{F}_{p^{10}}$ est cyclique, et trouver tous ses générateurs.

(c) Trouver le nombre de tous les polynômes unitaires irréductibles de degré 10 sur $\mathbb{F}_{p^{10}}$.

2. a) Trouver le terme constant du polynôme minimal P sur $\mathbb{Q}$ du nombre algébrique $\alpha = \sqrt{2} + \sqrt[3]{3}$.

b) Trouver toutes les racines complexes de P.

c) Déterminer le groupe de Galois du corps engendré par toutes les racines complexes de P.

3.* On considère une extension galoisienne L/K de groupe de Galois $\mathrm{Gal}(L/K)$, isomorphe au groupe symétrique S_5.

a) Déterminer toutes les sous-extensions galoisiennes E/K, où $K \subset E \subset L$.

b) Donner un exemple de deux sous-extensions E_1/K et E_2/K, ($E_1 \subset L$, $E_2 \subset L$) telles que E_1 et E_2 ne sont pas incluses l'une dans l'autre, le composé $E_1 \cdot E_2$ n'est pas égal à L, et l'intersection $E_1 \cap E_2$ n'est pas égale à K.

2006

4.* On considère le polynôme à coefficients rationnels

$$P(T) = T^3 - 3T - 1.$$

(a) Trouver le discriminant de P.

(b) Montrer que P est irréductible sur $\mathbb{Q}$ et qu'il possède trois racines réelles α_1, α_2, α_3, telles que $\alpha_3 < \alpha_2 < 0 < \alpha_1$.

(c) Montrer que si α est racine P, il en est de même de $2 - \alpha^2$. On pose $K = \mathbb{Q}(\alpha_1)$.

(d) Montrer que l'extension $K/\mathbb{Q}$ est galoisienne, et que tout élément de son groupe de Galois induit une permutation paire sur l'ensemble $\{\alpha_1, \alpha_2, \alpha_3\}$.

Références

Ouvrages de base :

[Bosch] Siegfried Bosch, *Algebra*, 3rd Ed., 1999

[Bourbaki] Bourbaki N. *Algèbre, Chap.8 . "Modules et anneaux semi-simples"*, Masson, Paris 1981.

[Dieudonné] Jean Alexandre Dieudonné, *La géométrie des groupes classiques*, Springer, 1963.

[Godement] Roger Godement, *Cours d'algèbre.*, Hermann, Paris, 1969

[Lang] Serge Lang , *Algebra.* Reading, Mass. : Addison–Wesley, 3rd Ed., 1993.

[Se70] Jean– Pierre Serre, *Cours d'arithmétique.* Paris, Press Univ. France, 1970.

[VdW71] B.L. van der Waerden, *Algebra I,II*, New York : Springer Verlag, 1971, 1967.

Lectures complémentaires :

[AMcD] I. G. MacDonald et M. F.Atiyah, *Introduction to Commutative Algebra.* Reading, MA : Addison-Wesley, 1969.

[Bigard] Alain Bigard, Géométrie, Cours et exercices corrigés pour le Capes et l'agrégation, Masson, 1998

[BS85] Z.I. Borevich, I.R.Shafarevich, *Number Theory.* Traduction anglaise. : New York/London : Academic Press, 1966.

[ChL] Antoine Chambert-Loir *Algèbre commutative*, `http ://www.polytechnique.fr/ chambert/ teach/algcom.pdf`

[Coq02] Robert Coquereaux, *Espaces fibrés et connexions. Une introduction aux géométries classiques et quantiques de la physique théorique.* Centre de Physique Théorique, Luminy - Marseille, `http ://www.cpt.univ-mrs.fr/~coque/book/sourceforhtml.html`

[Garrett] Paul Garrett's Page `http ://www.math.umn.edu/~garrett/m/buildings/`

2006

[Jac] N. Jacobson, *Basic Algebra I and II*, New York, NY : W.H. Freeman, 1974, 1989. Second Edition.

[Kob87] Neal Koblitz, *A course of number theory and cryptography*, Graduate texts in mathematics 114, New York : Springer Verlag, 1987.

[Kos82] A.I. Kostrikin, *Introduction to Algebra New York*, NY : Springer-Verlag, 1982

[KosMan] A.I. Kostrikin, Yu. I. Manin, *Linear algebra and geometry*, Nauka, Moscow 1986 ; English translation, Gordon and Breach, New York-London 1989

[Li-Ni] Rudolf Lidl et Harald Niederreiter, *Introduction to finite fields and their applications.* Addison–Wesley : Reading, 1983

[Ma-Pa] Yu.I. Manin et A.A.Panchishkin, *Introduction to Modern Number Theory*, Encyclopaedia of Mathematical Sciences, vol. 49 (2nd ed.), Springer-Verlag, 2005, 514 p.

[Mi-Hu] J. Milnor et D.Hüsemoller, *Symmetric bilinear forms*, Springer-Verlag, 1973

[Sha87] I.R. Shafarevich, (1987) : Fundamental notions of algebra. Itogi Nauki, 11, 1987. English transl. : Encycl. Math. Sci 11. Berlin–Heidelberg–New York : Springer–Verlag, 1990.

[Weyl] Hermann Weyl, The Classical Groups : Their Invariants and Representations

[Tits] Jacques Tits, *Le Monstre (d'après R. Griess, B. Fischer et al.)* dans Séminaire Bourbaki, Vol. 1983/84. Astérisque no 121-122, (1985), 105–122.

[Wei74] A.Weil (1974) : Basic Number Theory. 3rd ed. Berlin–Heidelberg–New York : Springer–Verlag, 1974.

[W] Wikipédia, *Wikipédia, l'encyclopédie libre*
`http ://fr.wikipedia.org/wiki)`

[Stein] William Stein, *An Explicit Approach to Number Theory* (à paraître, voir page internet : `http ://modular.fas.harvard.edu/edu/Fall2001/124/lectures`).