U.F.R. DES SCIENCES

MASTER DE MATHÉMATIQUES

SECONDE ANNÉE

2004-2005

COMPLÉMENTS D'ALGÈBRE

Jean COUGNARD

Master 2ème année
Programme des enseignements fondamentaux

Ces cours permettrons aux étudiants du master de renforcer leurs connaissances de base. Ils sont, en particulier, fortement recommandés aux étudiants souhaitant passer le concours de l'agrégation.

Compléments d'analyse (20 heures de cours+20 heures de TD).

Compléments de géométrie (20 heures de cours+20 heures de TD).

Compléments d'algèbre (20 heures de cours+20 heures de TD).

• **Compléments de théorie des groupes :** groupes finis, théorèmes de Sylow ; groupes opérants sur un ensemble.

• **Théorème des facteurs invariants** (cas d'un anneau de base euclidien). Application aux groupes abéliens et aux endomorphismes d'espaces vectoriels.

• **Réseaux de $\mathbb{R}^n$.**

• **Répartition des nombres premiers :** Théorème de Dirichlet au moins dans le cas $p \equiv 1 \pmod{n}$, formes faibles du théorème des nombres premiers (théorème de Tchebitcheff).

• **Quaternions** et applications.

Les TD pourront aussi être l'occasion d'étudier quelques théorèmes classiques :
- Transcendance de e et π.
- Théorème de Perron-Frobenius
- Comparaison des preuves du théorème d'Alembert-Gauss...

BIBLIOGRAPHIE :

[AB] J.-M. ARNAUDIÈS & J. BERTIN, *Groupes, Algèbres et Géométries* T 1. Ellipses 1993.

[C] J. CALAIS, *Éléments de la théorie des groupes*. P.U.F. 1996

[Cl] A.CLARK, *Elements of Abstract Algebra*. Dover Pub. 1971

[D] D. DUVERNEY, *Théorie des nombres*. Dunod 1998.

[EMF] W.J. ELLISON & M. MENDÈS FRANCE, *Les nombres premiers*. Hermann 1975.

[G1] R. GOBLOT, *Algèbre Commutative*. MASSON 1996.

[G2] R. GOBLOT, *Algèbre Linéaire*. MASSON 1995

[GS] G. GRAS & M.-N. GRAS, *Algèbre fondamentale, Arithmétique*. Ellipses 2004.

[HW] G. H. HARDY & E.M. WRIGHT, *Une introduction à la théorie des nombres*. Springer verlag 2001.

[IR] K.IRELAND &M. ROSEN *A clasical introduction to modern number theory*. Springer Verlag 1982.

[L]] T.Y. LAM, *The algebraic Theory of quadratic forms*. Benjamin 1973.

[Li] A. ALAOUI & H.QUEFFÉLEC & C. SACRÉ & V. VASSALLO, *Quelques aspects des mathématiques actuelles* Ellipses 1998.

[LTJ] L. LESIEUR & R. TEMAM & J.LEFEBVRE, *Compléments d'algèbre linéaire*. A. Colin 1978

[M] R. MNEIMNÉ, *Éléments de géométrie*. Cassini 1997.

[P] D. PERRIN, *Cours d'Algèbre*. Seconde édition Ellipses 1998.

[Sa] P. SAMUEL, *Théorie algébrique des nombres*. Méthodes, Hermann 1967.

[Se] D. SERRE, *Les Matrices, Théorie et pratique*. Dunod 2001.

[St] I. STEWART, *Galois theory* Second edition. Chapman and Hall 1989.

[T] G. TENENBAUM *Introduction à la théorie analytique et probabilistique des nombres*. 2ème édition. Cours spécialisés vol 1 Société Mathématique de France 1995.

[TMF] G. TENENBAUM & M. MENDÈS FRANCE, *Les nombres premiers*. Que sais-je vol. 571, P.U.F. 1997.

[V] R. VIDONNE, *Groupe circulaire, rotations et quaternions*. Ellipses 2001.

CHAPITRE I : ACTIONS DE GROUPE

C'est l'intérêt premier d'un groupe que d'agir sur un ensemble (muni ou non d'une structure). Tout les ouvrages de la bibliographie traitant des groupes y consacrent au moins un chapitre.

§1 ACTION D'UN GROUPE SUR UN ENSEMBLE.

Définition I-1 : *Une* **action** (à gauche) *d'un groupe G sur un ensemble X est une application φ :*

$$\varphi : \begin{array}{ccc} G \times X & \to & X \\ (g,x) & \mapsto & g \cdot x := \varphi(g,x) \end{array}$$

telle que quels que soient $g, h \in G$ et $x \in X$, alors $(gh) \cdot x = g \cdot (h \cdot x)$ et $e \cdot x = x$ où e est l'élément neutre de G. On l'appelle encore **opération** *de G sur X et on dit que G agit ou opère sur X.*

On définit de même une action à droite comme une application $(g,x) \mapsto x \cdot g$ telle que $x \cdot gh = (x \cdot g) \cdot h$ et $x \cdot e = x$.

Remarques :

1 On notera parfois gx l'élément $g \cdot x$ de X mais cette notation est à éviter lorsque X est égal à G (à cause des risques de confusions).

2 Dans tout ce qui suit on ne considère que des actions à gauche.

On peut voir une action comme un morphisme de G dans le groupe $\mathfrak{S}_X$ des permutations de X :

Proposition I-2 : *Si G agit sur X par $(g,x) \mapsto g \cdot x$, alors pour tout $g \in G$ l'application $\pi_g : x \mapsto g \cdot x$ est une permutation de X, et $g \mapsto \pi_g$ est un morphisme de G dans $\mathfrak{S}_X$* i.e. *$\pi_{gh} = \pi_g \circ \pi_h$. Réciproquement, si $g \mapsto p_g$ est un morphisme de G dans $\mathfrak{S}_X$ alors l'application $(g,x) \mapsto p_g(x)$ est une action de G sur X. Cela établit deux bijections réciproques entre l'ensemble des actions de G sur X et l'ensemble des morphismes de G dans $\mathfrak{S}_X$.*

Preuve : Il est clair que $\pi_e(x) = e \cdot x = x$ donc $\pi_e = id$. Par ailleurs, $\pi_g \circ \pi_h(x) = \pi_g(h \cdot x) = g \cdot (h \cdot x) = gh \cdot x = \pi_{gh}(x)$ et en particulier $\pi_g \circ \pi_{g^{-1}} = id$ donc π_g est inversible : c'est une permutation de X et $g \mapsto \pi_g$ est un morphisme.
Réciproquement : soit $g \mapsto p_g$ un morphisme $G \to \mathfrak{S}_X$. Si on pose $g \cdot x := p_g(x)$, on a $e \cdot x = p_e(x)$. Or, p_e est l'image de e par un morphisme, donc c'est l'identité et $e \cdot x = x$. De même, $gh \cdot x = p_{gh}(x) = p_g \circ p_h(x) = p_g(p_h(x)) = g \cdot (h \cdot x)$ ce qui prouve qu'on a une action de G sur X. □

Définition I-3 : *Si G agit sur X, la relation $x \sim y$ définie par : il existe un élément $g \in G$ tel que $y = g \cdot x$ est une relation d'équivalence sur X. La classe de $x \in X$ pour cette action s'appelle l'***orbite** *de x,on la note $G \cdot x$; l'ensemble des orbites forme une partition de X.*
On dit que l'action est **transitive** *ou que G agit transitivement s'il n'y a qu'une seule orbite .*
Le **noyau** *de l'action est le noyau du morphisme $G \to \mathfrak{S}_X$ associé, c'est-à-dire l'ensemble :*
$\{g \in G \ ; \ \forall x \in X, \ g \cdot x = x\}$ des g qui fixent X. *On dit que l'action est* **fidèle** *(ou que G agit fidèlement) si son noyau est réduit à $\{e\}$.*

Remarque : L'action est transitive si elle peut transformer tout élément en tout autre. **Attention !** Cela ne veut pas dire que le morphisme dans $\mathfrak{S}_X$ est surjectif !!! regarder C_3 et $\mathfrak{S}_3$ agissant sur {1,2,3}.

Exemples :
1 $GL_n(\mathbb{R})$ agit sur $\mathbb{R}^n$.
2 Si K est un corps, $(K^*, \times)$ agit par multiplication sur $K^n - \{0\}$.
3 Le groupe G des rotations de l'espace vectoriel euclidien $\mathbb{R}^3$ agit sur $\mathbb{R}^3$ (chaque rotation envoie un point de $\mathbb{R}^3$ sur un autre point). Les orbites sont les sphères centrées à l'origine. L'action n'est donc pas transitive. Elle est fidèle car la seule rotation fixant tout point de $\mathbb{R}^3$ est l'identité.
4 Si X est un ensemble, $\mathfrak{S}_X$ agit sur X par permutation. L'action est transitive et fidèle.
5 Tout groupe G agit sur lui-même par multiplication à gauche : $g \cdot h = gh$. Cette action est transitive et fidèle. Ce dernier exemple montre donc le :

Théorème I-4 : (**Cayley**) *Tout groupe G est isomorphe à un groupe de permutations, c'est-à-dire à un sous-groupe du groupe des permutations sur un ensemble X* (à savoir $X = G$).

Preuve : L'action fidèle de G sur lui-même donne un morphisme injectif de G dans $\mathfrak{S}_G$.

Remarque : Il suffit donc pour étudier tous les groupes de connaître les sous-groupes des groupes symétriques, d'où l'importance de ces derniers (cf. §3 et 4).

Exemples : (suite)
6 Tout groupe G agit sur lui-même par conjugaison : $(g,x) \mapsto g \cdot x := gxg^{-1}$ (ce dernier produit étant le produit dans le groupe G) ; on voit ici le danger de noter gx à la place de $g \cdot x$. Cette action n'est pas transitive si $G \neq \{e\}$ (l'orbite de e est réduite à e lui-même). Le noyau de l'action est le centre du groupe, l'action n'est donc fidèle que si ce centre est réduit à $\{e\}$.
7 $GL_n(k)$ opère sur $M_n(k)$ par conjugaison : $(P,M) \mapsto PMP^{-1}$. Les orbites sont les classes de similitude (si $k = \mathbb{C}$ il y en a autant que de formes de Jordan).
8 Le groupe des homographies $z \mapsto \dfrac{az+b}{cz+d}$ (a, b, c, $d \in \mathbb{C}$, $ad-bc \neq 0$) agit transitivement et fidèlement sur $\mathbb{C} \cup \{\infty\}$.
9 Soit $f : \mathbb{R}^n \to \mathbb{R}^n$ telle que pour tout point $P \in \mathbb{R}^n$ il existe une unique solution x_P de l'équation différentielle $x'(t) = f(x(t))$ avec la condition initiale $x(0) = P$. Alors $\mathbb{R}$ agit sur l'espace $\mathbb{R}^n$ par $t.P = x_P(t)$ et les orbites sont les trajectoires, c'est-à-dire les courbes intégrales de cette équation différentielle.
10 Soit G un groupe opérant sur un ensemble X et f un morphisme d'un groupe Γ dans G. Montrez que $\begin{array}{ccc} \Gamma \times X & \to & X \\ (\gamma, x) & \mapsto & f(\gamma)x \end{array}$ définit une action de Γ sur X (si f est injective on parle de restriction de l'action, si f est surjective on parle d'inflation).

§2 FORMULES DES CLASSES.

Ce chapitre est classique, le 2 du théorème 8 est moins souvent évoqué, on pourra consulter [AB], [M] à son sujet. Dans tout ce paragraphe, G est un groupe agissant (à gauche) sur un ensemble X.

Définition I-5 : *Le* **stabilisateur** *d'un élément $x \in X$ pour l'action de G est l'ensemble $G_x = \{g \in G \ g \ g \cdot x = x\}$, c'est un sous-groupe de G.*

Remarque : **Attention !** il n'y a aucune raison, à priori, pour qu'il soit distingué.

Proposition I-6 : *Pour x fixé dans G, l'application $\begin{array}{ccc} g & \mapsto & g \cdot x \\ G & \to & X \end{array}$ définit une bijection de l'ensemble G/G_x des classes à droite modulo G_x sur l'orbite de x. Ainsi le cardinal de l'orbite $G \cdot x$ est égal à l'indice $(G : G_x)$.*

Preuve : L'application $g \mapsto g \cdot x$ est une surjection de G sur $G \cdot x$. Par ailleurs :
$g' \cdot x = g \cdot x \iff g^{-1}g' \cdot x = x \iff g^{-1}g' \in G_x \iff g\,G_x = g'\,G_x$. □

Remarque : Cela montre en particulier que si G est un groupe fini et si x,y sont dans la même orbite, alors les stabilisateurs G_x et G_y ont même cardinal puisqu'ils ont même indice. Mais il y a mieux :

Proposition I-7 : *Si $y = g \cdot x$, la conjugaison $h \mapsto ghg^{-1}$ par g dans G définit un isomorphisme de G_x sur G_y.*

Preuve : Si $h \in G_x$, alors $ghg^{-1} \cdot y = gh \cdot (g^{-1} \cdot y) = gh \cdot x = g \cdot x = y$ donc $ghg^{-1} \in G_y$. Ainsi la conjugaison définit bien un morphisme (*car conserve loi de G*) de G_x dans G_y, d'inverse $k \mapsto g^{-1}kg$. □

Le résultat fondamental dans l'utilisation des actions de groupes finis est le suivant, connu sous le nom de **formules des classes** :

Théorème I-8 : (**formules des classes**) *Soit G un groupe fini agissant sur un ensemble fini X, alors :*
1 $\#X = \displaystyle\sum_{i=1}^{n} (G : G_{x_i})$ *où les x_i forment un système de représentants des orbites.*
2 *Le nombre n d'orbites est donné par la formule :*

$$n = \frac{1}{\#G} \sum_{g \in G} \#X_g$$

où $X_g = \{x \in X \ ; \ g \cdot x = x\}$ est l'ensemble des points de X fixés par g (au contraire du stabilisateur : c'est ici g qui est fixé).

Preuve :
1 Les orbites forment une partition de X donc le résultat découle de la proposition 2.
2 La somme $\sum_{g\in G} \#X_g$ est le cardinal de l'ensemble $A = \{(g,x) \in G \times X \ ; \ g \cdot x = x\}$. Si on choisit un système (x_i) de représentants des orbites, alors :

$$\#A = \sum_{x\in X} \#G_x = \sum_{i=1}^{n} (\#G_{x_i})(\#\text{orbite de } x_i) = \sum_{i=1}^{n} (\#G_{x_i})(G : G_{x_i}) = n\,\#G.$$

□

Ces formules ont de très nombreuses applications dont nous verrons quelques unes plus loin dans ce cours. Citons en particulier :

- Un groupe d'ordre $p^r m$ (p premier, m premier à p) admet un sous-groupe d'ordre p^r.
- Le théorème de Wedderburn : tout corps fini est commutatif.
- Le calcul du nombre de coloriages différents d'un cube (par action du groupe d'isométries laissant le cube stable sur l'ensemble des sextuplets de couleurs).

Démontrons simplement une application, liée à la théorie de Sylow (chapitre VI) :

Proposition I-9 : *Un groupe G de cardinal p^n où p est un nombre premier et $n > 0$ a un centre non trivial.*

Preuve : On fait agir G sur lui-même par conjugaison. Le stabilisateur d'un $x \in G$ est G (c'est-à-dire l'orbite n'a qu'un élément) si et seulement si x est dans le centre $Z(G)$. Tout autre stabilisateur a un indice qui est un diviseur de p^n différent de 1, donc multiple de p. Ainsi la première formule montre que $p^n = \#Z(G)$+multiple de p et donc $\#Z(G)$ est un multiple de p (non nul car $Z(G)$ contient au moins e). □

Complétons l'introduction au groupe symétrique faite dans le premier chapitre.

§3 CONJUGAISON DANS LE GROUPE SYMÉTRIQUE.

On étudie les classes d'équivalence pour la relation de conjugaison. Rappelons que $\mathfrak{S}_n$ (comme tout groupe) agit sur lui-même par conjugaison. On cherche à étudier les orbites de $\mathfrak{S}_n$ pour cette action *c'est-à-dire reconnaître quelles sont les permutations conjuguées*. La permutation conjuguée de σ par une permutation τ est $\tau\sigma\tau^{-1}$. Il suffit pour la calculer de savoir conjuguer un cycle puisque :
$\tau c_1 \dots c_r \tau^{-1} = \tau c_1 \tau^{-1} \dots \tau c_r \tau^{-1}$, or

$$\tau(i_1 \dots i_r)\tau^{-1} = (\tau(i_1) \dots \tau(i_r))$$

(comparer à gauche et à droite l'image des éléments $\tau(i_1), \dots, \tau(i_r)$ ainsi que des autres entiers)

Exemple : Dans $\mathfrak{S}_4$ on a $(134)(13)(24)(134)^{-1} = (34)(21)$.
En remarquant que la conjugaison ne change pas la longueur d'un cycle, on obtient ainsi :

Proposition I-10 : *Si on appelle type d'une permutation $\sigma = c_1 \dots c_r$ la suite $(l_1, \dots, l_r)$ des longueurs des cycles c_i classées par ordre croissant (i.e. $1 < l_1 \leq l_2 \leq \dots \leq l_r$), alors deux permutations sont conjuguées dans $\mathfrak{S}_n$ si et seulement si elle ont même type* (ici encore il est important de prendre de vrais cycles).

Preuve : Deux permutations conjuguées sont de même type par l'étude précédente.
Réciproquement, donnons nous deux permutations $\sigma = c_1 \dots c_r$ et $\sigma' = c'_1 \dots c'_s$ de même type décomposées en cycles de supports disjoints. L'égalité des types montre que $r = s$ et quitte à ordonner les cycles par longueur croissante (*on le peut car ils commutent*), deux cycles c_t et c'_t ont même longueur, soit $c_t = (i_{t,1} \dots i_{t,l_t})$ et $c'_t = (j_{t,1} \dots j_{t,l_t})$. Tous les entiers $i_{t,u}$ sont distincts et en même nombre que les $j_{t,u}$; il y a donc une permutation τ qui envoie chaque $i_{t,u}$ sur le $j_{t,u}$ correspondant et le complémentaire des $i_{t,u}$ sur le complémentaire des $j_{t,u}$. D'après le calcul du conjugué d'un cycle, on en déduit que $\tau c_t \tau^{-1} = c'_t$ et par suite que $\tau\sigma\tau^{-1} = \sigma'$. □

Autrement dit : si $\sigma = (i_1 \dots i_k)(\dots)(\dots i_l)$ et $\sigma' = (j_1 \dots j_k)(\dots)(\dots j_l)$ alors $\tau = \begin{pmatrix} i_1 & \dots & i_l \\ & & \\ j_1 & \dots & j_l \end{pmatrix}$ convient.

Application au calcul du cardinal des classes de conjugaison

Théorème I-11 : *Les r-cycles de $\mathfrak{S}_n$ forment une classe de conjugaison de cardinal $\frac{1}{r}\frac{n!}{(n-r)!}$. Plus généralement, les permutations de type $(\underbrace{l_1,\dots,l_1}_{n_1\ fois},\dots,\underbrace{l_k,\dots,l_k}_{n_k\ fois})$ forment une classe de conjugaison de cardinal $\frac{n!}{(n-s)!}\prod_1^k \frac{1}{n_i!\,l_i^{n_i}}$ où $s=\sum l_i n_i$ est le cardinal du support de σ.*

Preuve : Tout r-uplet $(i_1,\dots,i_r)$ définit un cycle $(i_1\dots i_r)$ et chaque cycle provient de r tels r-uplets puisque $(i_1\dots i_r)=(i_2\dots i_r i_1)=\dots$. Or il y a $\frac{n!}{(n-r)!}$ r-uplets, d'où le nombre de r-cycles.

Dans le cas général, si $s=\sum n_i l_i$, tout s-uplet $(i_1,\dots,i_s)$ définit une permutation de type $(\underbrace{l_1,\dots,l_1}_{n_1\ fois},\dots,\underbrace{l_k,\dots,l_k}_{n_k\ fois})$, à savoir

$$\underbrace{(i_1\dots i_{l_1})(i_{l_1+1}\dots)\dots(\dots i_{n_1 l_1})}_{n_1}\underbrace{(\dots)\dots(\dots)}_{n_2}\dots\underbrace{(\dots)\dots(\dots i_s)}_{n_k}.$$

Le nombre de s-uplets est $\frac{n!}{(n-s)!}$ et deux s-uplets définissent la même permutation si et seulement si on permute circulairement les éléments d'un même paquet (ce qui correspond aux différentes écritures d'un même cycle), ou qu'on permute entre eux deux paquets de même longueur (ce qui ne fait qu'échanger deux cycles, qui commutent). Il y a $\prod l_i^{n_i}$ permutations de la première sorte, et $\prod n_i!$ de la seconde, d'où le résultat. □

Exercice **1** : Combien il y a-t-il de classes de conjugaison dans $\mathfrak{S}_5$, vérifier que la somme de leurs cardinaux vaut 120 (on trouve $1+10+15+20+20+30+24=120$).

§4 GÉNÉRATEURS – SIGNATURE – GROUPE ALTERNÉ.

On sait que le groupe $\mathfrak{S}_n$ est engendré par les cycles. Il est souvent utile, pour montrer certaines propriétés des permutations, de pouvoir se restreindre à un ensemble plus petit de générateurs (*par exemple pour le cube dont le groupe des isométries est isomorphe à $\mathfrak{S}_4$*). La proposition suivante fournit de tels ensembles :

Théorème I-12 :
1 *Le groupe $\mathfrak{S}_n$ est engendré par les transpositions.*
2 *Le groupe $\mathfrak{S}_n$ est engendré par les transpositions de la forme $(1\,i)$.*
3 *Le groupe $\mathfrak{S}_n$ est engendré par les transpositions, dites élémentaires, de la forme $(i\ \ i+1)$.*
4 *Le groupe $\mathfrak{S}_n$ est engendré par les deux permutations $(1\,2)$ et $(1\,2\,\dots\,n)$.*

Preuve :
1 Il suffit d'écrire tout cycle comme produit de transpositions, ce qui est possible car :

$$(i_1\dots i_k)=(i_1 i_k)\dots(i_1 i_3)(i_1 i_2)$$

(*vérifier sur chaque élément. Attention ! on commence par la droite car c'est un composé d'applications*).
2 D'après 1, il suffit d'écrire toute transposition comme produit de transpositions de la forme $(1i)$ ce qui est aussi possible puisque $(ij)=(1i)(1j)(1i)$ (*on passe de $(1j)$ à (ij) par conjugaison, cf. le calcul d'un conjugué*).
3 Puisque par conjugaison $(1\,i)=(i-1\ i)(1\ i-1)(i-1\ i)$, on voit par récurrence sur i que toute transposition $(1i)$ est produit de transpositions élémentaires.
4 La transposition (12) et le cycle $c=(12\dots n)$ engendrent les transpositions élémentaires (et donc $\mathfrak{S}_n$ entier) grâce à la relation de conjugaison $(i+1\ i+2)=c^i(1\,2)c^{-i}$. □

Remarque : Toute permutation est donc produit de transpositions, mais elles ne sont pas uniques (voir la preuve du 2), pas plus que leur nombre. Cependant ce nombre a une parité bien définie qui découle de la notion suivante :

Définition I-13 : *Une* **inversion** *d'une permutation σ est un couple (i,j) tel que $i<j$ et $\sigma(i)>\sigma(j)$, ou encore tel que $\frac{\sigma(i)-\sigma(j)}{i-j}<0$. La* **signature** *de σ est le nombre*

$$\epsilon(\sigma)=(-1)^{\text{nombre d'inversions de }\sigma}=\prod_{\substack{(i,j)\\ i<j}}\frac{\sigma(i)-\sigma(j)}{i-j}$$

(car σ est une bijection donc la valeur absolue du produit vaut 1) . *On dit que σ est* **paire** *(resp.* **impaire***) si $\epsilon(\sigma) = 1$ (resp. -1).*

Proposition I-14 : *La signature $\epsilon : \mathfrak{S}_n \to \{\pm 1\}$ est un morphisme de groupes* (en particulier $\sigma\tau$ et $\tau\sigma$ ont même signature). *Deux permutations conjuguées ont même signature :* $\epsilon(\sigma\tau\sigma^{-1}) = \epsilon(\sigma)\epsilon(\tau)\epsilon(\sigma)^{-1} = \epsilon(\tau)$).

Preuve : Si $\tau, \sigma \in \mathfrak{S}_n$, l'application $(i,j) \mapsto \{\tau(i), \tau(j)\}$ est une permutation des couples d'entiers, que l'on peut réordonner :

$$\prod_{\substack{(i,j)\\ i<j}} \frac{\sigma\tau(i) - \sigma\tau(j)}{i-j} = \prod_{\substack{(i,j)\\ i<j}} \frac{\sigma\tau(i) - \sigma\tau(j)}{\tau(i) - \tau(j)} \prod_{\substack{(i,j)\\ i<j}} \frac{\tau(i) - \tau(j)}{i-j} = \epsilon(\sigma)\epsilon(\tau).$$

Car dans le premier produit du second terme on peut, s'il le faut changer l'ordre au numérateur et au dénominateur pour avoir $\displaystyle\prod_{\substack{(k,l)\\ \tau(k)<\tau(l)}} \frac{\sigma\tau(k) - \sigma\tau(l)}{\tau(k) - \tau(l)}$ □

Proposition I-15 : *Toute transposition est impaire. Ainsi, pour $n > 1$, ϵ est un morphisme surjectif, et une permutation est paire si et seulement si elle est produit d'un nombre pair de transpositions.*

Preuve : Toute transposition est conjuguée de $(1\,2)$ donc a même signature. Or pour $(1\,2)$, seule la paire $\{1,2\}$ fournit une inversion. □

Définition I-16 : *Pour $n \geq 2$, le noyau $\mathfrak{A}_n$ de la signature $\mathfrak{S}_n \to \{\pm 1\}$ est un sous-groupe normal d'indice 2 de $\mathfrak{S}_n$, appelé le n-ième* **groupe alterné**. *C'est donc l'ensemble des permutations paires.*

Preuve : Pour vérifier qu'il est d'indice 2, il suffit de remarquer que le morphisme surjectif ϵ définit un isomorphisme de $\mathfrak{S}_n/\mathfrak{A}_n$ sur $\{\pm 1\}$. □

De même que dans le cas du groupe symétrique tout entier, il est intéressant de trouver dans $\mathfrak{A}_n$ des générateurs privilégiés :

Proposition I-17 : *Si $n \geq 3$, le groupe $\mathfrak{A}_n$ est engendré par les 3-cycles.*

Preuve : Toute permutation paire étant un produit d'un nombre pair de transpositions de la forme $(1\,a)$, il suffit de remarquer que $(1\,b)(1\,a) = (1\,a\,b)$. □

Remarque : La preuve du théorème II-11 montre que la signature d'un cycle de longueur n est $(-1)^{n-1}$. Nous montrons maintenant un résultat essentiel dans l'étude des équations algébriques. Il est dû à Galois et permet de montrer qu'un polynôme de degré 5 ou plus n'a en général pas de racine exprimable par radicaux à partir de ses coefficients comme c'est le cas en degré 2, 3 ou 4.

Théorème I-18 : *$\mathfrak{A}_n$ est un groupe simple si et seulement si $n \neq 4$.*

Preuve : Les trois premiers sont simples pour raison de cardinal.
$\mathfrak{A}_4$ n'est pas simple car l'identité et les trois produits de 2 transpositions forment un sous-groupe (*en fait isomorphe à* $(\mathbb{Z}/2\mathbb{Z})^2$), évidemment stable par conjugaison donc normal.
On suppose donc maintenant que $n \geq 5$.
On commence par montrer que les 3-cycles sont tous conjugués dans $\mathfrak{A}_n$ si $n \geq 5$. En effet deux 3-cycles c, c' sont conjugués dans $\mathfrak{S}_n$: $c' = \sigma c \sigma^{-1}$. Si σ est impaire, on choisit des entiers $a, b \notin Supp(c)$ (*ici $n > 4$ est crucial*) de sorte que la transposition $\tau = (a\,b)$ commute à c, donc $c' = (\sigma\tau)c(\sigma\tau)^{-1}$ et $\sigma\tau$ est paire.

Le lemme suivant permet d'étudier les autres classes de conjugaison dans $\mathfrak{A}_n$: (*rappel* : le cardinal de l'orbite est l'indice du stabilisateur)

Lemme I-19 : *Si $\sigma \in \mathfrak{A}_n$, les conjugués de σ dans S_n forment une (resp. deux) classe(s) de conjugaison dans $\mathfrak{A}_n$ s'il existe une permutation impaire commutant à σ (resp. sinon).*

Preuve : Soit $A_{n,\sigma}$ (resp. $C_{n,\sigma}$) le stabilisateur de σ pour l'action de $\mathfrak{A}_n$ (resp. $\mathfrak{S}_n$) sur lui-même par conjugaison, c'est-à-dire les permutations de A_n (resp. S_n) commutant à σ. Puisque le nombre de conjugués (cardinal de l'orbite) est l'indice du stabilisateur, on voit que si $A_{n,\sigma} = C_{n,\sigma}$ alors $\#\mathfrak{A}_n \cdot \sigma = \frac{1}{2}\#\mathfrak{S}_n \cdot \sigma$, et que si $A_{n,\sigma}$ est d'indice ≥ 2 dans $C_{n,\sigma}$ alors $\#\mathfrak{A}_n \cdot \sigma \geq \#\mathfrak{S}_n \cdot \sigma$, d'où égalité car $\mathfrak{A}_n \cdot \sigma \subset S_n \cdot \sigma$. □

Ainsi les 15 produits de 2 transpositions sont tous conjugués dans $\mathfrak{A}_5$ puisque (12) commute à (12)(34). Par ailleurs, les permutations commutant au cycle $\sigma = (12345)$ forment un groupe d'indice 24 (nombre de

5-cycles) de $\mathfrak{S}_5$, donc de cardinal 5. Il n'y a donc que les 5 puissances de σ, qui sont toutes dans $\mathfrak{A}_5$. Par suite les vingt-quatre 5-cycles forment deux classes de conjugaison de $\mathfrak{A}_5$. Les trois cycles sont conjugués dans $\mathfrak{A}_5$ et au nombre de 20.
La partition de $\mathfrak{A}_5$ en classes de conjugaison est donc :

classe :	$\{id\}$	C_3	C_{22}	C_5	C_5'
cardinal :	1	20	15	12	12

Les seules réunions de classes contenant $\{id\}$ ayant un cardinal divisant 60 sont donc $\{id\}$ et $\mathfrak{A}_5$. Un sous-groupe de $\mathfrak{A}_5$ étant réunion de telles classes s'il est normal, on en déduit que $\mathfrak{A}_5$ est simple.

Lemme I-20 : *Si $n > 4$ et H est un sous-groupe normal de $\mathfrak{A}_n$ contenant un 3-cycle, alors $H = \mathfrak{A}_n$.*

Preuve : Les 3-cycles sont conjugués dans $\mathfrak{A}_n$ et engendrent $\mathfrak{A}_n$. □

De façon analogue, en utilisant ce lemme, montrons maintenant que $\mathfrak{A}_6$ est simple :

On fixe un sous-groupe H distingué dans $\mathfrak{A}_6$.

Si les σ dans $H \setminus \{e\}$ ont tous support $\{1, \dots, 6\}$, ils sont de la forme (12)(3456) ou (123)(456). Il y a 90 (resp. 40) conjugués de ces permutations dans $\mathfrak{S}_6$, et comme plus haut le nombre de conjugués dans $\mathfrak{A}_6$ est 90 car commutant avec (12) (resp. 40 car commutant avec (14)(25)(36)). Donc le cardinal de H est somme de certains des nombres 1,40, 90 y compris 1, et une telle somme ne divise $\#\mathfrak{A}_6 = 360$ que si elle vaut 1.

On suppose donc que H contient un $\sigma \neq id$ fixant par exemple 1 (*Idée* : H contient un sous-groupe normal de $\mathfrak{A}_5$ donc $\mathfrak{A}_5$ donc un 3-cycle. Idem avec $\mathfrak{A}_6$.) Soit $G_1 \subset \mathfrak{A}_6$ les permutations qui fixent 1 (c'est le stabilisateur pour quelle action ?). Alors G_1 est isomorphe à $\mathfrak{A}_5$, et $H \cap G_1$ est un sous-groupe normal de G_1 (car stable par conjugaison par G_1), donc égal à G_1 donc H contient le 3-cycle $(2\,3\,4)$ d'où $H = \mathfrak{A}_6$.

On peut maintenant supposer que $n > 6$ et montrer que $\mathfrak{A}_n$ est simple : soit H un sous-groupe normal, et $\sigma \neq id$ une permutation de H. Soit τ un 3-cycle qui ne commute pas à σ (si $\sigma(a) \neq a$, soit b, c deux autres entiers, alors $\tau = (\sigma(a), b, c)$ convient).
Alors $\sigma_1 = \sigma(\tau\sigma^{-1}\tau^{-1})$ est élément de H (*car H est normal*) et produit de deux 3-cycles $(\sigma\tau\sigma^{-1})\tau^{-1}$. Il existe donc un ensemble de 6 entiers $i_1, \dots, i_6$ contenant son support. Soit F l'ensemble des permutations de $\mathfrak{A}_n$ de support dans $\{i_1, \dots, i_6\}$. C'est un sous-groupe de $\mathfrak{A}_n$ isomorphe à $\mathfrak{A}_6$ et $H \cap F$ en est un sous-groupe normal (*stable par conjugaison*) contenant $\sigma_1 \neq id$ (*c'est pas id car ils ne commutent pas*) donc $H \cap F = F$ soit $H \supset F$. Donc H contient un 3-cycle par exemple $(i_1\, i_2\, i_3)$, et $H = \mathfrak{A}_n$. □

Corollaire I-21 : *Les polynômes de $\mathbb{Q}[X]$ de degré ≥ 5 ne sont pas tous résolubles par radicaux.*

Preuve : Voir au paragraphe 6.

Corollaire I-22 : *Pour $n > 4$, les groupes dérivés de $\mathfrak{A}_n$ et $\mathfrak{S}_n$ sont égaux à $\mathfrak{A}_n$.*

Preuve : La signature d'un commutateur est paire donc $D(\mathfrak{A}_n) \subset D(\mathfrak{S}_n) \subset \mathfrak{A}_n$. Par ailleurs, $D(\mathfrak{A}_n)$ est normal dans $\mathfrak{A}_n$ et non réduit à $\{id\}$ car $\mathfrak{A}_n$ n'est pas abélien. □

Remarque : Autre démonstration : tout 3-cycle est un commutateur. Or, si σ est un 3-cycle, alors σ^2 aussi, donc conjugué dans $\mathfrak{A}_n$ à σ, soit $\sigma = \tau\sigma\tau^{-1}\sigma^{-1}$.

Exercice **2** : $D(\mathfrak{A}_4)$ est le groupe de Klein V_4. En effet c'est un sous-groupe distingué, donc soit $\mathfrak{A}_4$ soit V_4 (pas 1 car $\mathfrak{A}_4$ n'est pas abélien) ; c'est le plus petit à quotient abélien or $|\mathfrak{A}_4/V_4| = 12/4 = 3$ donc est abélien.

Corollaire I-23 : *Si $n \geq 5$, les seuls sous-groupes distingués de $\mathfrak{S}_n$ sont $\{id\}, \mathfrak{S}_n$ et $\mathfrak{A}_n$.*

Preuve : $\mathfrak{A}_n$ est un sous-groupe normal de $\mathfrak{S}_n$. Soit H un sous-groupe distingué de $\mathfrak{S}_n$. Alors $H \cap \mathfrak{A}_n \triangleleft \mathfrak{A}_n$ donc vaut $\{id\}$ ou $\mathfrak{A}_n$. Dans le second cas, H vaut $\mathfrak{S}_n$ ou $\mathfrak{A}_n$. Dans le premier cas, la signature $H \to \{\pm 1\}$ a un noyau trivial donc $\#H \leq 2$. Il reste à exclure le cas $\#H = 2$. Or dans ce cas $H = \{id, \sigma\}$, donc pour $\tau \in \mathfrak{S}_n$ le conjugué $\tau\sigma\tau^{-1}$ est encore dans H et d'ordre 2 donc vaut σ, ce qui veut dire que σ est dans le centre de $\mathfrak{S}_n$. On conclut en utilisant le :

Lemme I-24 : *Si $n \geq 3$, le centre de $\mathfrak{S}_n$ est réduit à l'identité*

Preuve : Si $\sigma \neq id$, soit $a \in \{1, \dots, n\}$ tel que $\sigma(a) \neq a$ et b un élément différent de a et de $\sigma(a)$. Alors si $\tau = (b\ \sigma(a))$, il vérifie $\tau\sigma(a) = b$ et $\sigma\tau(a) = \sigma(a)$ donc σ n'est pas dans le centre.

Ceci achève la preuve du corollaire. □

§5 THÉORIE DE SYLOW.

La théorie de Sylow permet de compter dans certains cas le nombre de sous-groupes de cardinal donné d'un groupe G. Ce cardinal doit bien sûr diviser celui de G (théorème de Lagrange), la théorie de Sylow etudie cette question lorsque le cardinal est une puissance d'un nombre premier.

Dans tout ce paragraphe, p désigne un nombre premier. Commençons par quelques résultats utiles sur les éléments et groupes d'ordre une puissance de p.

Théorème I-25 : (**Cauchy**) *Tout groupe G d'ordre multiple de p contient un élément d'ordre p.*

Preuve : On procède par récurrence sur $\#G$. Commençons par le cas abélien plus facile. Si G, abélien contient un élément g d'ordre multiple de p, soit pk, alors g^k convient. Sinon, fixons un élément non neutre g qui est donc d'ordre t premier à p. Ainsi le quotient $G/ < g >$ (*groupe car G abélien*) est d'ordre $\#G/t$ multiple de p et par récurrence contient un élément $\overline{h} = h < g >$ (où $h \in G$) d'ordre p. Ceci prouve que h est d'ordre multiple de p (car $\overline{h}^n = \overline{h^n}$) et comme plus haut une puissance de h est d'ordre p.
Passons au cas général, G agit sur lui-même par conjugaison. Les éléments du centre $Z(G)$ sont ceux dont l'orbite n'a qu'un élément. Si ce centre (abélien) est d'ordre multiple de p, le cas abélien achève la preuve. Sinon la formule des classes, qui s'écrit ici

$$\#G = \#Z(G) + \sum (G : G_x)$$

montre qu'il existe un élément $x_i \notin Z(G)$ dont le stabilisateur G_{x_i} est d'indice non multiple de p, donc de cardinal multiple de p. Puisque $\#G_{x_i} < \#G$ (car non dans le centre) la récurrence s'applique à G_{x_i}. □

Définition I-26 : *Un p-**groupe** est un groupe dont tout élément est d'ordre puissance de p.*

Remarque : Il peut être infini par exemple $\{n/p^k \in \mathbb{Q}/\mathbb{Z}\ ;\ n \in \mathbb{Z}, k \in \mathbb{N}\}$.

Proposition I-27 : *Un groupe fini G est un p-groupe si et seulement s'il est d'ordre puissance de p.*

Preuve : Tout groupe d'ordre puissance de p est un p-groupe par le théorème de Lagrange. Réciproquement, si un premier $q \neq p$ divise $\#G$, le groupe G contient un élément d'ordre q (*par Cauchy*) donc n'est pas un p-groupe. □

Proposition I-28 : *Si G est un groupe d'ordre p^r, et si $0 \leq s \leq r$, alors G contient un sous-groupe d'ordre p^s.*

Preuve : Par récurrence sur $\#G$. Il n'y a rien à démontrer si $r = 0$. Faisons agir G par conjugaison sur lui-même. Le stabilisateur G_x de chaque élément $x \notin Z(G)$ est un sous-groupe strict de G donc d'indice multiple de p. La formule des classes prouve ainsi que $p \mid \#Z(G)$ et donc que le centre contient un élément c d'ordre p. Le sous-groupe $< c >$ est alors normal (*car central*) et le groupe $G/ < c >$ d'ordre p^{r-1} contient par récurrence un sous-groupe H d'ordre p^{s-1}. L'image inverse $\mathcal{H} = \pi^{-1}(H)$ de H par la projection $\pi : G \to G/ < c >$ est un sous-groupe de G contenant c, et la projection $\pi : \mathcal{H} \to H$ est surjective de noyau $< c >$ ce qui prouve que $H \simeq \mathcal{H}/ < c >$. Par suite $\#\mathcal{H} = p.p^{s-1} = p^s$. □

Définition I-29 :
1 *Si H est un sous-groupe d'un groupe G ses* **conjugués, dans** G, *sont les images gHg^{-1} de H par conjugaison par les éléments de G.*
2 *Si G est un groupe fini, p un nombre premier et p^r la plus grande puissance de p qui divise $\#G$, alors tout sous-groupe de G de cardinal p^r s'appelle un p-***(sous-groupe de) Sylow** *de G.*

Remarques :
1 En particulier H est normal si et seulement s'il est son seul conjugué.
2 Les conjugués de H sont isomorphes à H car la conjugaison est un automorphisme.
3 Si p ne divise pas $\#G$ le p-Sylow est $\{e\}$.

Exemple : Les 2-Sylow de S_5 sont d'ordre 8, les 3-Sylow de S_5 sont d'ordre 3, les 5-Sylow de S_5 sont d'ordre 5. Ils sont d'ordre 1 pour les autres valeurs de p. Faire la même chose avec S_6.

Le résultat central de ce chapitre est le suivant :

Théorème I-30 : (Sylow) *Soit G un groupe fini, p un nombre premier.*
1 *Pour tout entier s tel que p^s divise $\#G$, il existe un sous-groupe de G d'ordre p^s. En particulier il existe un p-Sylow de G.*
Le nombre n_p des p-Sylow de G vérifie $n_p \equiv 1(p)$ et $n_p \mid \#G$.
2 *Le conjugué d'un p-Sylow en est aussi un et réciproquement tous les p-Sylow sont conjugués dans G. Enfin tout p-sous-groupe de G est contenu dans un p-Sylow.*

Remarque : Si p^r est la plus grande puissance de p divisant le cardinal de G, le nombre n_p des p-Sylow de G divise donc $\#G/p^r$ et est congru à 1 modulo p ce qui est une restriction très forte.

Preuve :
a. Montrons l'existence d'un p-Sylow (et donc de sous-groupes d'ordre p^s par la prop. 4).
Notons $\#G = n = p^r.m$ où $(m, p) = 1$. On peut supposer $r > 0$ sans quoi $\{e\}$ est un p-Sylow. Soit X l'ensemble des parties de G (*on dit bien parties, pas sous-groupes*) à p^r éléments. Son cardinal

$$\#X = \binom{n}{p^r} = \prod_0^{p^r-1} \frac{n-i}{p^r-i}$$

est premier à p car $p^r m - i$ et $p^r - i$ sont divisibles par la même puissance de p tant que $i < p^r$. Le groupe G agit sur X par multiplication :

$$(g, \mathcal{A}) \mapsto g.\mathcal{A} = \{ga \ ; \ a \in \mathcal{A}\}$$

(*qui a bien p^r éléments*).
Puisque p ne divise pas $\#X$, une orbite $G \cdot \mathcal{A}_0$ au moins a un cardinal $(G : G_{\mathcal{A}_0})$ non multiple de p d'après la formule des classes. Par suite $\#G_{\mathcal{A}_0}$ est multiple de p^r. Par ailleurs, si $a \in \mathcal{A}_0$, les $g.a$ $(g \in G_{\mathcal{A}_0})$ sont distincts (*car les ga $(g \in G)$ sont distincts*) et dans $\mathcal{A}_0$ d'où $\#G_{\mathcal{A}_0} \leq \#\mathcal{A}_0 = p^r$ ce qui prouve qu'il y a égalité et que $G_{\mathcal{A}_0}$ est un p-Sylow.
b. La conjugaison ne change pas le cardinal donc tout conjugué d'un p-Sylow est un p-Sylow.
c. Montrons que tout p-sous-groupe est dans un p-Sylow et que les p-Sylow sont conjugués dans G.
Fixons pour cela un p-Sylow S, et un p-sous-groupe H de G (qui peut être un p-Sylow ou non). Il suffit de montrer que H est dans un conjugué de S.
Faisons agir H sur l'ensemble X_1 des classes modulo S par $(h, gS) \mapsto hgS$. Puisque X_1 est de cardinal $(G : S) = m$ premier à p et que les orbites ont des cardinaux puissances de p (les stabilisateurs sont des sous-groupes de H), l'une au moins a cardinal 1. Notons la g_0S, de sorte que $g_0^{-1}(hg_0) \in S$ pour tout $h \in H$ soit $h \in g_0 S g_0^{-1}$.
d. Il reste à montrer les conditions sur n_p.
Le p-Sylow $S = S_1$ agit par conjugaison sur l'ensemble de tous les p-Sylow $X_2 = \{S_1, \ldots, S_{n_p}\}$. L'unique orbite à 1 élément est celle de S_1 : en effet si $S_i = sS_is^{-1}$ pour tout $s \in S_1$, alors $S_iS_1 = S_1S_i$ donc ce produit ensembliste est un sous-groupe (*pourquoi ?*) et S_i est normal dans S_1S_i (*faire le calcul*) ; mais S_i et S_1 sont des Sylow de S_1S_i et par suite sont conjugués dans S_1S_i, donc égaux (*par normalité*). Il résulte de cela que les autres orbites sont de cardinal une puissance non triviale de p donc que $n_p = \#X_2 \equiv 1(\mathrm{mod} p)$.
Enfin si G agit sur X_2 par conjugaison, on a vu qu'il n'y avait qu'une orbite qui a donc n_p éléments et $n_p = (G : G_{S_1})$ divise $\#G$. □

Exemple : Il n'existe pas de groupe simple d'ordre 15 ; en effet, dans un groupe d'ordre 15 le nombre n_5 de 5-Sylow vérifie $n_5 \equiv 1(\mathrm{mod}\ 5)$ et $n_5 \mid 3$ donc $n_5 = 1$. Il y a ainsi un seul 5-Sylow ce qui veut dire qu'il est son seul conjugué et par suite est normal.

Exercice **3** : Faire la même chose pour 63 (voir n_7). Pour 56 c'est un peu plus subtil : il ne peut y avoir à la fois sept 2-Sylow et huit 7-Sylow (compter les éléments d'ordre 7 ou divisant 2).

Exercice **4** : Montrez que si $G/Z(G)$ est cyclique, G est abélien (et réciproquement !). En déduire que tout groupe d'ordre p^2 est abélien. Cette propriété est-elle encore vraie si on remplace $G/Z(G)$ est cyclique par $G/Z(G)$ est abélien (réutilisez le groupe des isométries du plan affine euclidien conservant un carré).

Exercice **5** : Soit G un groupe non abélien d'ordre 6. Montrez qu'il existe un élément de G d'ordre 3 et que cet élément engendre un sous-groupe distingué. En déduire que G est isomorphe au groupe S_3 des permutations de trois éléments.

§ 6 SUITES DE COMPOSITION – GROUPES RÉSOLUBLES.

Il s'agit, suivant l'expression consacrée, de dévisser un groupe pour faire apparaître des groupes simples. On pourra consulter [C], [Cl].

Définition I-31 : *Soit G un groupe fini, une suite croissante $\{e\} = G_0 \subset G_1 \subset \ldots \subset G_{n-1} \subset G_n = G$ de sous-groupes est une suite normale si quel que soit i $(0 \leq i \leq n-1)$ G_i est distingué dans G_{i+1}, le nombre n est la longueur de la suite et les G_{i+1}/G_i sont ses quotients.*

Définition I-32 : *Une seconde suite normale est dite plus fine que la première si tous les G_i de la première apparaissent dans la seconde ; ceci donne une relation d'ordre sur l'ensemble des suites normales.*

Définition I-33 : *Une suite normale, sans répétition, qui est un élément maximal pour la relation d'ordre s'appelle une suite de composition.*

La proposition qui suit est évidente :

Proposition I-34 : *Une suite est une suite de composition si et seulement si ses quotients sont des groupes simples.*

Corollaire I-35 : *Toute suite normale peut être incluse dans une suite de composition.*

Définition I-36 : *On dit que deux suites normales sont isomorphes si et seulement si elles ont même longueur et mêmes quotients, à l'ordre près.*

Théorème I-37 : *Deux suites normales d'un même groupe fini possèdent des suites plus fines isomorphes.*

Preuve : Soit les deux suites normales :

$$\{e\} = G_0 \subset G_1 \subset \ldots \subset G_{n-1} \subset G_n = G$$
$$\{e\} = H_0 \subset H_1 \subset \ldots \subset H_{m-1} \subset H_m = G$$

Pour tout entier k de la forme $k = qm + r$ avec $0 \leq q < n$, $0 \leq r \leq m$ on définit $\hat{G}_k = G_q(G_{q+1} \cap H_r)$; c'est un sous-groupe de G : $g\gamma\gamma_1^{-1}g_1^{-1} = g((\gamma\gamma_1^{-1})g_1^{-1}(\gamma_1\gamma^{-1}))\gamma\gamma_1^{-1}$. Si on écrit $qm + 0 = (q-1)m + m$ on a $\hat{G}_{qm} = G_q(G_{q+1} \cap H_0) = G_q$ et $\hat{G}_{qm} = G_{q-1}(G_q \cap H_m) = G_q$ ce qui rend la notation cohérente, montre que la suite est plus fine que la suite de départ et possède $mn + 1$ termes. De même pour tout entier ℓ de la forme $\ell = qn + s$ avec $0 \leq q < m$, $0 \leq s \leq n$ on définit le sous-groupe $\hat{H}_\ell = H_q(H_{q+1} \cap G_s)$ avec les mêmes propriétés. On a construit deux suites de sous-groupes de G. Montrons (sur l'une seulement) que ce sont deux suites normales. Soit $g\gamma \in G_u(G_{u+1} \cap H_v)$ et $g_1\gamma_1 \in G_u(G_{u+1} \cap H_{v+1})$ regardons $(g_1\gamma_1)g\gamma(\gamma_1^{-1}g_1^{-1})$ qui se récrit $g_1(\gamma_1 g\gamma_1^{-1})(\gamma_1\gamma\gamma_1^{-1})g_1^{-1} = g_1(\gamma_1 g\gamma_1^{-1})\left((\gamma_1\gamma\gamma_1^{-1})g_1^{-1}(\gamma_1\gamma^{-1}\gamma_1^{-1})\right)(\gamma_1\gamma\gamma_1^{-1})$ ce qui montre bien que $G_u(G_{u+1} \cap H_v) \triangleleft G_u(G_{u+1} \cap H_{v+1})$. On a donc deux suites normales. Utilisons maintenant le troisième théorème d'isomorphisme.

Lemme I-38 : *Soit H_1, H_2 deux sous-groupes d'un groupe G, N_1, N_2 deux sous-groupes distingués respectivement de H_1 et de H_2, on a les isomorphismes :*

$$\frac{N_1(H_1 \cap H_2)}{N_1(H_1 \cap N_2)} \simeq \frac{H_1 \cap H_2}{(H_1 \cap N_2)(N_1 \cap H_2)} \simeq \frac{N_2(H_1 \cap H_2)}{N_2(H_2 \cap N_1)}$$

Preuve : On a déjà vu que $N_1(H_1 \cap H_2)$, $N_1(H_1 \cap N_2)$ sont des sous-groupes de G et que le second est distingué dans le premier. On pose alors $H = H_1 \cap H_2$, $N = N_1(H_1 \cap N_2)$ et on utilise $\frac{NH}{N} \simeq \frac{H}{H \cap N}$, on obtient alors :

$$\begin{aligned}\frac{N_1(H_1 \cap N_2)(H_1 \cap H_2)}{N_1(H_1 \cap N_2)} &\simeq \frac{H_1 \cap H_2}{H_1 \cap H_2 \cap N_1(H_1 \cap N_2)} \simeq \frac{H_1 \cap H_2 \cap H_2}{H_2 \cap H_1 \cap (H_1 \cap N_2)N_1} \\ &\simeq \frac{H_1 \cap H_2}{H_2 \cap (H_1 \cap N_2)N_1} \simeq \frac{H_1 \cap H_2}{H_2 \cap N_1(H_1 \cap N_2)} \\ &\simeq \frac{H_1 \cap H_2}{(H_2 \cap N_1)(H_1 \cap N_2)} \simeq \frac{H_1 \cap H_2}{(H_1 \cap N_2)(H_2 \cap N_1)}\end{aligned}$$

On utilise l'égalité $X \cap YZ = Y(X \cap Z)$ lorsque $Y \subset X$, Z sont des sous-groupes, avec $X = H_2$, $Y = H_1 \cap N_2$, $Z = N_1$. □

Fin de la démonstration du théorème : Pour $k = um + v$ et $\ell = vn + u$ on a les isomorphismes :

$$\frac{\hat{G}_{k+1}}{\hat{G}_k} = \frac{G_u(G_{u+1} \cap H_{v+1})}{G_u(G_{u+1} \cap H_v)} \simeq \frac{H_v(H_{v+1} \cap G_{u+1})}{H_v(H_{v+1} \cap G_u)} \simeq \frac{\hat{H}_{\ell+1}}{\hat{H}_\ell}$$

Ces isomorphismes montrent que les doublons dans les deux nouvelles suites sont en nombres égaux, il suffit de les éliminer pour obtenir le résultat. □

Quand on part de deux suites de composition, comme on ne peut pas trouver de suite plus fine, on a :

Corollaire I-39 : (THÉORÈME DE JORDAN-HÖLDER) *Deux suite de composition d'un même groupe sont isomorphes.*

Les groupes résolubles furent introduits par Galois dans son mémoire sur les équations résolubles par radicaux (voir le paragraphe suivant).

Définition I-40 : *Un groupe G est dit résoluble s'il existe une suite finie de sous-groupes G_i de G*

$$\{e\} = G_0 \subsetneq G_1 \subsetneq \ldots \subsetneq G_{n-1} \subsetneq G_n = G$$

telle que :
1 . $G_i \triangleleft G_{i+1}$, $0 \leq i \leq n-1$,
2 . G_{i+1}/G_i *est abélien.*

Exemple : Les groupes abéliens sont résolubles, les groupes $\mathfrak{S}_3$, $\mathfrak{S}_4$, $\mathfrak{A}_4$ sont résolubles.

On déduit de ce qui précède

Corollaire I-41 : *Soit G un groupe, H un sous-groupe de G et N un sous-groupe distingué de G.*
1 . *Si G est résoluble, H est résoluble,*
2 . *Si G est résoluble, G/N est résoluble,*
3 . *Si N et G/N sont résolubles, G est résoluble.*

On peut construire une suite normale de sous-groupes d'un groupe fini G en prenant $G_1 = D(G)$, $G_2 = D(D(G)) := D^2(G)$, $D^n(G) := D(D^{n-1}(G))$. Cette suite est décroissante, dans un groupe fini, donc stationnaire, soit $D^n(G)$ le dernier groupe obtenu : $G \supsetneq D(G) \supsetneq D^2(G) \supsetneq \ldots \supsetneq \{e\}$ les quotients sont abéliens sauf éventuellement le dernier isomorphe à $D^n(G)$. Si ce dernier est abélien, le groupe est résoluble, on a $D^{n+1}(g) = \{e\}$.
Si le groupe est résoluble, il existe une suite normale plus fine que la précédente et telle que les quotients soient abéliens. Si $D^n(G) \neq \{e\}$, on en extrait $D^n(G) \supsetneq G_k \supsetneq \ldots \{e\}$; comme $D^n(G)/G_k$ est abélien, $G_k \subset D(D^n(G))$ ce qui est contradictoire puisque $D^n(G)$ est censé être le dernier. On en déduit :

Corollaire I-42 : *Le groupe fini G est résoluble si et seulement s'il existe n tel que $D^n(G) = \{e\}$.*

On introduit d'autres familles de groupes :
Les groupes hyper-résolubles : *les G_i sont supposés distingués dans G et G_i/G_{i+1} est cyclique* (le groupe $\mathfrak{A}_4$ est résoluble mais pas hyper-résoluble).
Les groupes nilpotents : *les G_i sont distingués dans G et G_i/G_{i+1} est inclus dans le centre de G/G_i.*

§ 7 PRODUITS SEMI-DIRECTS.[P]

Ayant dévissé un groupe pour faire apparaître des groupes simples, on peut essayer de construire des groupes à partir d'autres qui jouent le rôle de briques. La manière la plus simple consiste à faire le produit direct. On expose une méthode un peu plus élaborée. On est conduit à la définition suivante :

Définition I-43 : *Soit H, K deux sous-groupes de G. On dit que G est* **produit semi-direct** *de H par K et on note $G = K \rtimes H$ ou $G = H \ltimes K$ si les trois conditions suivantes sont vérifiées :*
1 *K est distingué dans G.*
2 *$KH = G$, autrement dit tout élément $g \in G$ s'écrit sous la forme $g = kh$.*
3 *$H \cap K = \{e\}$, i.e. l'écriture précédente est unique.*

Remarques :
1 Remarquer la dissymétrie, seul K est supposé normal.
2 Si $G = KH$ c'est automatiquement égal à HK car $hk = hkh^{-1}\, h$.

3 Attention l'ordre est important, signe ouvert côté K (comme ker).
4 Pour $h \in H$ l'application $k \mapsto hkh^{-1}$ est un automorphisme de K noté $\varphi(h)$; $h \mapsto \varphi(h)$ est un morphisme de H dans $\mathrm{Aut}(K)$.

En résumé, on a un sous-groupe distingué K, un sous-groupe H tels que $H \cap K = \{e\}$, $KH = G$ et un morphisme φ de H dans $\mathrm{Aut}(K)$; le produit $khk'h'$ est égal à $k\varphi(h)(k')hh'$ qui est bien le produit d'un élément $k\varphi(h)(k')$ de K par $hh' \in H$.

Proposition I-44 : *Si G est produit semi-direct (en particulier si c'est un produit direct) de K par H, alors le groupe quotient G/K est isomorphe à H. En particulier, si les groupes sont finis alors $\#G = \#K\ \#H$.*

Preuve : D'après **2)** et **3**, tout élément g s'écrit uniquement sous la forme $g = kh$. On considère alors l'application de G dans H qui à g associe le facteur h. C'est un morphisme car si $g_1 = k_1h_1$ et $g_2 = k_2h_2$ alors $g_1g_2 = \left[k_1(h_1k_2h_1^{-1})\right](h_1h_2)$ et le deuxième facteur est bien dans H. Ce morphisme est surjectif car sa restriction à H est l'identité. Le noyau est $\{g\ ;\ g = kh \text{ avec } h = e\} = K$. Ainsi par passage au quotient par le noyau, φ définit un isomorphisme de G/K sur H.

Exercice **6** : Le groupe symétrique S_3 est produit semi-direct du sous-groupe d'ordre 2 engendré par la transposition (12) par le sous-groupe d'indice 2 engendré par (123), mais n'est pas produit direct.

De même qu'on peut définir le produit direct de deux groupes abstraits (non a priori sous-groupes d'un même groupe), on peut définir des produits semi-directs de groupes abstraits. Il peut y en avoir plusieurs, qui ne dépendent pas que des groupes de départ. Cette construction est basée sur le point 4 de la remarque précédente :

Définition I-45 : *Soit H, K deux groupes, et φ un morphisme de H dans* $\mathrm{Aut}(K)$. *Le* **produit semi-direct** (externe) *de H par K associé à φ est l'ensemble* (pas le groupe !) *$K \times H$ muni de la loi de composition $(k,h)(k',h') := (k.\varphi(h)(k'), hh')$. Ce produit semi-direct est noté $K \rtimes_\varphi H$.*

Exercice **7** :
1 Le produit semi-direct externe forme bien un groupe.
• associativité :
$[(k,h)(k',h')](k'',h'') = (k\varphi(h)(k'), hh')(k'',h'') = (k\varphi(h)(k')\varphi(hh')(k''), hh'h'')$
$= (k\varphi(h)(k')\varphi(h) \circ \varphi(h')(k''), hh'h'')$ tandis que
$(k,h)[(k',h')(k'',h'')] = (k,h)(k'\varphi(h')(k''), h'h'') = (k\varphi(h)(k'\varphi(h')(k'')), hh'h'')$
$= (k\varphi(h)(k')\varphi(h) \circ \varphi(h')(k''), hh'h'')$
• (e,e) est neutre :
$(e,e)(k,h) = (e\varphi(e)(k), eh) = (\mathrm{id}(k), h) = (k,h)$, $(k,h)(e,e) = (k\varphi(h)(e), he) = (k,h)$
• l'inverse de (k,h) est $(\varphi(h^{-1})(k^{-1}), h^{-1})$.
2 Si G est le produit semi-direct (interne) d'un sous-groupe H par un sous-groupe normal K, alors $G \simeq K \rtimes_\varphi H$ par $(k,h) \mapsto kh$ pour l'automorphisme φ défini par $\varphi(h)(k) = hkh^{-1}$ (ainsi $\varphi(h)$ est la conjugaison par h).
3 Si G est le produit semi-direct externe de K par H pour φ ($G \simeq K \rtimes_\varphi H$), l'application de K dans G : $k \mapsto (k,e)$ est un morphisme injectif et l'image est un sous-groupe distingué de G ; de même l'application de H dans G : $h \mapsto (e,h)$ est un morphisme injectif. L'intersection des images est réduite à l'élément neutre, tout élément de G est produit d'un élément de l'image de K par un appartenant à celle de H. En résumé, G est produit semi-direct interne de l'image de H par celle de K.

Exemple : Soit $K = \mathbb{Z}$, $H = \{\pm 1, \times\}$ on a un morphisme de groupe de H dans $\mathrm{Aut}(\mathbb{Z})$: en envoyant 1 sur l'identité et -1 sur la multiplication par -1. Sur $K \times H$ on définit l'opération $(n,\epsilon)(n'\epsilon') = (n + \epsilon n', \epsilon\epsilon')$ ce qui donne un exemple de produit semi-direct.

Exemple : Les similitudes planes et autres exemples géométriques : le groupe affine est semi-direct de GL_n par les translations (qui sont distinguées).

Exemple : Recherchez dans les groupes déjà rencontrés ceux qui peuvent se mettre sous la forme d'un produit semi-direct.

CHAPITRE II : DIVISEURS ÉLÉMENTAIRES

Plusieurs théorèmes de structure, celui sur les groupes abéliens de type fini, celui sur la réduction de Jordan se ramènent en dernier ressort à des propriétés de modules de type fini sur un anneau principal. C'est ce thème qui est exposé ici. La démarche suivie est celle exposée dans [GS] centrée sur l'algorithme de Smith. Il prend toute son importance lorsque l'anneau est euclidien, cas le plus usuel. On donne en conclusion le théorème de structure des groupes abéliens et la réduction de Jordan. Celui qui doit faire un exposé sur ces théorèmes peut bien entendu éviter cette généralisation et proposer une approche directe (voir, par exemple, [GS] chapitre VI pour le premier, [G2] chapitre 7 pour le second).

Dans le premier paragraphe nous rappelons quelques points élémentaires sur les modules, dans le but de fixer les notations, ces notions ayant été acquises en première année de Master (comme on dit en franglais). Pour nous les anneaux sont commutatifs et unitaires.

§1 RAPPELS SUR LE RANG DES MODULES.

Définition II-1 : *On dit qu'un A module M est de type fini s'il existe des éléments $m_1, \dots, m_r$ en nombre fini tels qu'il soit égal au sous-module $< m_1, \dots, m_r >$.*

Définition II-2 : *On dit qu'un A module M est libre de type fini, s'il est de type fini engendrable par des éléments $e_1, \dots, e_n$ tels que tout élément m de M s'écrive de manière unique $m = \sum_{i=1}^{n} a_i e_i$ $(a_i \in A)$. La famille $\{e_1, \dots, e_n\}$ s'appelle une base de M.*

Remarque : L'exemple immédiat est le module A^n.

Proposition II-3 : *Si un A-module M est libre toutes les bases ont le même nombre d'éléments, ce nombre s'appelle la dimension du module.*

Preuve : Soit $\mathcal{M}$ un idéal maximal de A, on construit le sous-module $\mathcal{M}M$ de M puis le module quotient $M/\mathcal{M}M$. De par sa construction, c'est aussi un $A/\mathcal{M}$-module. Comme ce dernier anneau est un corps, $M/\mathcal{M}M$ est un $A/\mathcal{M}$-espace vectoriel. Soit $\mu \in M/\mathcal{M}M$, c'est l'image d'un élément $m = \sum_{i=1}^{n} a_i e_i$ de M ; les images de e_i engendrent l'espace vectoriel : il est de dimension finie. Montrons que les images $\overline{e}_i$ des e_i sont $A/\mathcal{M}$-libres. Soit $\sum_{i=1}^{n} \alpha_i \overline{e_i} = 0$ avec $\alpha_i = \overline{a_i}$ $(a_i \in A, 1 \leq i \leq n)$ c'est dire que $\sum_{i=1}^{n} a_i e_i \in \mathcal{M}M$, donc que les $a_i \in \mathcal{M}$ et finalement que les α_i sont nuls. La famille des $\overline{e}_i$ est une base du $A/\mathcal{M}$-espace vectoriel. Le nombre d'éléments de la base de M est égal à la $A/\mathcal{M}$-dimension de $M/\mathcal{M}M$, donc invariante. □

Remarque : Pour les applications entre A-modules libres de dimension finie, on peut utiliser le calcul matriciel, en prenant bien garde de ne diviser que par des éléments inversibles.

Corollaire II-4 : *Tout A-module de type fini est quotient d'un A-module-libre de dimension finie.*

Définition II-5 : *Soit A un anneau intègre, M un A-module, Un élément m de M est de torsion s'il existe $a \in A \setminus \{0\}$ tel que $ax = 0$.*

Lemme II-6 : *Si A est un anneau intègre, l'ensemble $\mathrm{Tor}(M)$ des éléments de torsion de M est un sous A-module de M appelé le sous-module de torsion de M .*

Preuve : Soit $m_1, m_2 \in \mathrm{Tor}(M)$, il existe $a, b \neq 0 \in A$ tels que $am_1 = bm_2 = 0$, comme l'anneau est intègre $ab \neq 0$ et $ab(m_1 + m_2) = b(am_1) + a(bm_2) = 0$. Enfin si $c \neq 0 \in A$, on a $a(cm_1) = c(am_1) = 0$. □

Remarque : L'hypothèse A intègre est importante : prenons $A = \mathbb{Z}/6\mathbb{Z}$ les classes de 2 et 3 sont de torsion mais pas $3 - 2$ (les éléments de torsion ne forment pas un idéal de A).

On cite sans démonstration :

Corollaire II-7 : *Si A est un anneau intègre et M un A-module, le A-module quotient $M/\mathrm{Tor}(M)$ est sans torsion.*

Soit A un anneau intègre, son corps des fractions K est « naturellement » un A-module, de même K^n ; l'injection naturelle de A dans K permet alors de considérer A^n comme un sous-A-module libre de K^n et plus généralement tout K-espace vectoriel comme un A-module.

Proposition II-8 : *Soit A un anneau intègre de corps des fractions K et M un A-module de type fini, il existe un K-espace vectoriel de dimension finie V et une injection de A-module de $M/\mathrm{Tor}(M)$ dans V.*

Preuve : Puisque M est un A-module de type fini il existe un entier n et un morphisme surjectif π du A-module de A^n sur M, soit N son noyau, le module M est isomorphe à A^n/N. Dans cet isomorphisme, à

un élément m de M correspond une classe $x+N$. On considère A^n comme un sous-A-module du K-espace vectoriel K^n et on construit KN le sous-espace vectoriel de K^n engendré par N et l'espace vectoriel quotient $V = K^n/KN$. À la classe $x+N$, on associe celle de $x+KN$ dans V. Un vérification de routine montre que l'on a construit un morphisme f de A-modules de M dans V. Cherchons le noyau de ce morphisme. Pour que m appartienne au noyau de f, il faut et il suffit que x appartienne à KN. On peut donc écrire $x = \sum_{j=1}^{s} v_j n_j$ $(v_j \in K, n_j \in N)$ avec les $v_j = \frac{a_j}{d}$ $(a_j, d \in A,\ d$ un dénominateur commun) ce qui donne $dx = \sum_{j=1}^{s} a_j n_j \in N$ et donc $dm = 0$. L'élément x est dans le noyau de f si et seulement si $dx = 0$, donc si et seulement si $x \in \mathrm{Tor}(M)$. Le passage au quotient donne donc un morphisme injectif de A-module de $M/\mathrm{Tor}(M)$ dans V. □

La construction que l'on vient d'effectuer n'est pas canonique mais le nombre $n - \dim_K KN$ qui est la dimension du K-espace vectoriel engendré par l'image de $M/\mathrm{Tor}(M)$ est indépendant de cette construction.

Théorème II-9 : *Soit M un A module de type fini (A un anneau intègre de corps des fractions K), V un K-espace vectoriel et $u : M/\mathrm{Tor}(M) \to V$ un morphisme injectif de A-modules. La dimension du K-sous-espace vectoriel engendré par $u\,(M/\mathrm{Tor}(M))$ est indépendante du couple (V, u) ; cette dimension s'appelle le rang de M.*

Preuve : Soit $u : Y = M/\mathrm{Tor}(M) \to V$ et $u' : Y = M/\mathrm{Tor}(M) \to V'$ deux morphismes injectifs de A modules. Soit U, U' les sous-espaces vectoriels de V et V' engendrés par les images respectives de Y par u et u'. Puisque $u(Y)$ engendre le sous-espace vectoriel U on peut en extraire une base $(u(y_i)_{i\in I})$. On considère la famille $(u'(y_i)_{i\in I})$ et on suppose qu'il y a une relation de dépendance linéaire entre ces éléments : $\sum_j \lambda_j u'(y_j) = 0$ $(J \subset I)$. On écrit $\lambda_j = \frac{a_j}{b}$ $(a_j \in A, b \in A \setminus \{0\})$, dans le K-espace vectoriel U' on obtient :

$$0 = \sum_j \frac{a_j}{b} u'(y_j) = \frac{1}{b} \sum_j a_j u'(y_j) = \frac{1}{b} u'(\sum_j a_j y_j)$$

comme u' est injective $\sum_j a_j y_j = 0$ et par u, $\sum_j a_j u(y_j) = 0$. Comme, par hypothèse, les $u(y_j)$ sont linéairement indépendants tous les a_j sont nuls et donc tous les λ_j. On en déduit $\dim_K(U') \geq \dim_K(U)$. Comme U et U' peuvent jouer un rôle symétrique on en déduit l'égalité voulue. □

§2 MODULES SUR LES ANNEAUX PRINCIPAUX.

On suppose désormais que A est un anneau principal. On se propose de démontrer le résultat de structure suivant :

Théorème II-10 : *Soit A un anneau principal, $M = A^n$ un A-module libre de rang n et N un sous-module de M, alors :*

1 - N est un sous-module libre de rang $s \leq n$,

2 - il existe une base $e_1, \ldots, e_n$ de M et des éléments $q_1, \ldots, q_s$ de A tels que $q_1 | q_2 | \ldots | q_s$ et tels que $\{q_1 e_1, q_2 e_2, \ldots, q_s e_s\}$ est une base de N.

Commentaire : les q_i sont appelés les diviseurs élémentaires et la base $e_1, \ldots, e_n$ une base de M adaptée à N. La base adaptée n'est pas unique, les diviseurs élémentaires le sont (à multiplication par des inversibles près). Nous allons d'abord donner une démonstration du point 1, celle du point 2 sera achevée par la construction de l'algorithme de Smith au paragraphe suivant.

Corollaire II-11 : *Le quotient M/N est A-isomorphe à $(\prod_{i=1}^{s} A/q_i A) \times A^{n-s}$.*

Lemme II-12 : *Si A est un anneau principal alors tout sous-module N de A^n est libre de type fini et peut être engendré par n éléments au plus.*

Preuve : Soit $S = \{\beta_j, j \in J\}$ un système de générateurs de N (on ne le suppose pas fini). Écrivons $\beta_j = \sum_{i=1}^{n} \lambda_{i,j} \epsilon_i$, $\lambda_{i,j} \in A$ sur la base canonique de A^n. L'idéal de A engendré par les $\lambda_{1,j}$ est de la forme $a_1 A$ (A est principal). On a les faits suivants :

(i) il existe des $u_{1,j} \in A$ tels que $\lambda_{1,j} = a_1 u_{1,j}$,

(ii) il existe des v_j presque tous nuls tels que $a_1 = \sum_{j\in J} v_j \lambda_{1,j}$.

On construit $\sum_{j\in J} v_j \beta_j$ qui est alors de la forme $a_1\epsilon_1 + z_1$ où $z_1 \in A\epsilon_2 \oplus \ldots \oplus A\epsilon_n$. On vérifie immédiatement que N est aussi engendré par la partie $S_1 = \{a_1\epsilon_1 + z_1, \beta_j - u_{1,j}(a_1\epsilon_1 + z_1)\}$ de la forme $\{a_1\epsilon_1 + z_1, \gamma_j, j \in J\}$ où les $\gamma_j \in A\epsilon_2 \oplus \ldots \oplus A\epsilon_n$. On recommence avec le sous-module N' de $A\epsilon_2 \oplus \ldots \oplus A\epsilon_n$ engendré par les

$\gamma_j, j \in J$ (on passe directement à cette étape si $a_1 = 0$). On obtient alors un système S_n de générateurs de N de la forme $\{a_1\epsilon_1 + z_1, a_2\epsilon_2 + z_2, \ldots, a_n\epsilon_n\}$ où $z_k \in A\epsilon_{k+1} \oplus \ldots \oplus A\epsilon_n$. La remarque de la fin de la première étape montre que l'on peut faire disparaître les générateurs avec $a_k = 0$. Si on veut chercher une relation de dépendance entre ces générateurs on constate immédiatement que l'on a un système linéaire échelonné, la famille de générateurs est libre. □

Remarque : On trouve dans [Sa] une démonstration de théorème de structure basée sur ce schéma, non algorithmique et où les questions d'unicité sont laissées au lecteur.

Début de la démonstration du théorème : On peut supposer que N est engendré par n éléments : on ajoute des 0 s'il y en a moins et on se place dans $A^n \times A^{m-n}$ s'il y en a $m > n$, en remplaçant les générateurs α_i par $(\alpha_i, 0)$.

On pose $\alpha_j = \sum_{i=1}^{n} a_{i,j}\epsilon_i$, ($a_{i,j} \in A$, les ϵ_i étant la base canonique de A^n). On construit la matrice $R = (a_{i,j}) \in M_n(A)$, il lui correspond un endomorphisme f de A^n ayant en colonne j : $\alpha_j = f(\epsilon_j)$. L'image $f(M)$ est N. On va s'appuyer sur le lemme suivant :

Lemme II-13 : *Soit g, h deux isomorphismes de modules de M sur M alors : $M/N \simeq M/g \circ f \circ h(M)$.*

Preuve : On a $h(M) = M$, puis $f \circ h(M) = f(M) = N$ et enfin $M/g(N) \simeq g \circ g^{-1}(M)/g(N)$ et comme $g^{-1}(M) = M$ on obtient $g(M)/g(N) \simeq M/N$.

Le problème est de trouver g et h tels $g \circ f \circ h$ soit le plus simple possible, ce que l'on peut traduire par trouver des matrices U et $V \in GL_n(A)$ telles que URV soit « proche » de la matrice identité. On va montrer que l'on peut construire U et V telles que :

$$URV = \begin{pmatrix} q_1 & \ldots & 0 & 0 & \ldots & 0 \\ \vdots & \ddots & \vdots & \vdots & & \vdots \\ 0 & \ldots & q_s & 0 & \ldots & 0 \\ 0 & \ldots & 0 & 0 & \ldots & 0 \\ \vdots & \ddots & \vdots & \vdots & & \vdots \\ 0 & \ldots & 0 & 0 & \ldots & 0 \end{pmatrix}$$

comme on a :

$$N = f(M) = f \circ h(M) = g^{-1}(g \circ f \circ h(M)) = g^{-1}(\oplus_{i=1}^{s} Aq_i\epsilon_i) = \oplus_{i=1}^{s} q_i A g^{-1}(\epsilon_i)$$

la base adaptée est celle des $e_i = g^{-1}(\epsilon_i)$. □

§3 L'ALGORITHME DE SMITH.

On suppose donnée la matrice $R \in M_n(A)$ et on veut construire des matrices $U, V \in Gl_n(A)$ telles que URV ait la forme voulue. Ce travail se fait par des manipulations sur les lignes et les colonnes. Ces opérations peuvent se traduire par des produits de matrices que l'on va rappeler. On va modifier les matrices de permutations de telles sorte que U et V soient des matrices dans $Sl_n(A)$.

Permutations.

On se donne $1 \leq \alpha < \beta \leq n$ et on construit la matrice :

$$\Pi_\alpha^\beta = \begin{pmatrix} 1 & \ldots & 0 & \ldots & 0 & \ldots & 0 \\ \vdots & \ddots & \vdots & & \vdots & & \vdots \\ 0 & \ldots & 0 & \ldots & 1 & \ldots & 0 \\ \vdots & & \vdots & \ddots & \vdots & & \vdots \\ 0 & \ldots & -1 & \ldots & 0 & \ldots & 0 \\ \vdots & & \vdots & & \vdots & \ddots & \vdots \\ 0 & \ldots & 0 & \ldots & 0 & \ldots & 1 \end{pmatrix}$$

où les coefficients de la diagonale principale valent 1 sauf ceux en (α,α) et (β,β) qui valent 0, les autres coefficients sont nuls sauf celui en (α,β) qui vaut 1 et celui en (β,α) égal à -1. Cette matrice est de déterminant 1, si on multiplie la matrice R à gauche par Π_α^β on échange les lignes α et β et la nouvelle ligne β est multipliée par -1. On constate que l'opération faite sur R est exactement celle faite sur les lignes de la matrice identité pour obtenir Π_α^β. Si on multiplie la matrice R à droite par Π_α^β on échange les colonnes sα et β et la nouvelle colonne α est multipliée par -1. On constate que l'opération faite sur R est exactement celle faite sur les colonnes de la matrice identité pour obtenir Π_α^β. L'inverse de Π_α^β est sa transposée.

Opérations linéaires.

On se donne α,β, $1\leq\alpha\neq\beta\leq n$ (pour l'illustration on fixe $\alpha<\beta$), $\lambda\in A$ et on construit la matrice dont les coefficients de la diagonale valent 1, les autres sont nuls sauf celui sur la ligne α et la colonne β qui vaut λ. On peut obtenir cette matrice soit en ajoutant à la ligne α de la matrice identité λ fois sa ligne β soit en ajoutant à la colonne β de la matrice identité λ fois sa colonne α. On obtient la matrice $\Lambda_{\alpha,\beta}^\lambda$:

$$\Lambda_{\alpha,\beta}^\lambda=\begin{pmatrix}1&\dots&0&\dots&0&\dots&0\\ \vdots&\ddots&\vdots&&\vdots&&\vdots\\ 0&\dots&1&\dots&\lambda&\dots&0\\ \vdots&&\vdots&\ddots&\vdots&&\vdots\\ 0&\dots&0&\dots&1&\dots&0\\ \vdots&&\vdots&&\vdots&\ddots&\vdots\\ 0&\dots&0&\dots&0&\dots&1\end{pmatrix}$$

Cette matrice est de déterminant 1. Quand on multiplie R à gauche par $\Lambda_{\alpha,\beta}^\lambda$ on ajoute λ fois la ligne β à la ligne α. Quand on multiplie R à droite par $\Lambda_{\alpha,\beta}^\lambda$ on ajoute λ fois la colonne α à la colonne β. La remarque précédente sur le passage de la matrice identité à $\Lambda_{\alpha,\beta}^\lambda$ est aussi valable.

Nous allons pouvoir aborder les opérations qui conduisent à l'algorithme de Smith. Auparavant :

Définition II-14 : *Soit $a\in A\setminus\{0\}$, comme A est principal, on peut écrire $a=u\prod_p p^{n_p}$ avec $u\in A^*$, les p parcourant un système de représentants des éléments irréductibles, les $n_p\in\mathbb{N}$. On pose alors $\sigma(a)=\sum_p n_p$.*

Définition II-15 : *Soit R, R' deux matrices de $M_n(A)$, on dit que R et R' sont équivalentes s'il existe deux matrices P et $Q\in Gl_n(A)$ telles que $R'=PRQ$.*

L'algorithme de Smith repose sur l'arrêt de la décroissance de la fonction σ.

Lemme II-16 : *Soit $R\in M_n(A)$ de la forme $\begin{pmatrix}a&\dots&.\\ \vdots&&\vdots\\ \vdots&\dots&\vdots\end{pmatrix}$ avec $a\in A\setminus\{0\}$; soit b un coefficient de R situé sur la première ligne ou la première colonne, alors il existe dans $M_n(R)$, une matrice R' équivalente à R de la forme $\begin{pmatrix}d&\dots&.\\ \vdots&&\vdots\\ \vdots&\dots&\vdots\end{pmatrix}$ où d est un pgcd de a et b.*

Preuve : Traitons le cas où b est sur la première colonne (la transposition permet de traiter celui où il est sur la première ligne) $R={}^t\begin{pmatrix}a&\dots&b\dots\\ \vdots&&\vdots&\vdots\\ \vdots&&\vdots&\vdots\end{pmatrix}$. Posons $a=da'$, $b=db'$ avec a' et b' premiers entre eux. Il existe alors u et v tel que $ua'+vb'=1$. Considérons la matrice :

$$U=\begin{pmatrix}u&\dots&v&\dots&.\\ \vdots&\ddots&\vdots&&\vdots\\ -b'&\dots&a'&\dots&.\\ \vdots&&\vdots&\ddots&\vdots\\ .&\dots&\dots&\dots&.\end{pmatrix}$$

Les coefficients non écrits étant égaux à 1 sur la diagonale, à 0 ailleurs. La matrice est de déterminant 1, pour avoir son inverse, on échange a' et u et on remplace b' et v par leurs opposés. On vérifie que $R' = UR$ a la forme voulue. □

Première étape de l'algorithme.

On suppose $R \neq 0$ (sinon $N = 0$ et il n'y a rien à démontrer). Un des coefficients est non nul et on peut supposer que c'est $a_{1,1}$ (sinon, on utilise des matrice Π_α^β pour se ramener à ce cas).

Lemme II-17 : *Soit* $R = \begin{pmatrix} a & \dots \\ \vdots & \dots \end{pmatrix}$ *avec* $a \in A \setminus \{0\}$. *Si un coefficient de la première ligne ou de la première colonne n'est pas divisible par* a, *il existe une matrice* $R' = \begin{pmatrix} d & \dots \\ \vdots & \dots \end{pmatrix}$ *équivalente à* R *avec* $\sigma(d) < \sigma(a)$.

Preuve : Il suffit d'appliquer le lemme 15 pour remplacer a par le pgcd de a et du coefficient en question. □

Comme $\sigma(\ )$ ne peut décroitre indéfiniment le processus va se terminer.

Seconde étape de l'algorithme.

On utilise les matrices de combinaison linéaire pour annuler tous les coefficients de la première ligne et de la première colonne en dehors de celui égal à d. On a obtenu une matrice $R' = \begin{pmatrix} d & 0 & \dots & 0 \\ 0 & . & \dots & . \\ \vdots & \vdots & \vdots & . \\ 0 & \dots & \dots & \dots \end{pmatrix}$.

Lemme II-18 : *Soit* R *une matrice de* $M_n(A)$ *de la forme* $\begin{pmatrix} a & 0 & \dots & 0 \\ 0 & . & \dots & . \\ \vdots & \vdots & \vdots & . \\ 0 & \dots & \dots & \dots \end{pmatrix}$. *Si tous les coefficients de* R *ne sont pas multiples de* a, *il existe une matrice* $R' = \begin{pmatrix} d & \dots \\ \vdots & \dots \end{pmatrix}$ *équivalente à* R *avec* $\sigma(d) < \sigma(a)$.

Preuve : Si un des coefficients non nul n'est pas multiple de a, on l'amène sur la première ligne en multipliant R à gauche par une matrice $\Lambda^1_{1,\beta}$. On utilise alors le lemme 15 pour remplacer a par le pgcd d de a et du coefficient. □

Troisième étape de l'algorithme.

On reprend la seconde étape de l'algorithme qui remplace d par d' avec $\sigma(d') < \sigma(d)$ et les autres coefficients de la première ligne et de la première colonne nuls. On utilise alors à nouveau le lemme 17. La décroissance stricte de σ ne pouvant se poursuivre indéfiniment on va obtenir une matrice :

$$R' = \begin{pmatrix} a & 0 & \dots & 0 \\ 0 & . & \dots & . \\ \vdots & \vdots & \vdots & . \\ 0 & \dots & \dots & \dots \end{pmatrix}$$

avec a qui divise tous les coefficients.

Dernière étape de l'algorithme.

Il suffit de reprendre le procédé avec la matrice $n-1 \times n-1$ obtenue en supprimant la première ligne et la première colonne.

L'algorithme de Smith est construit, la preuve du théorème 10 est maintenant complète.

Remarques :
1 L'algorithme ne fonctionne que si on sait calculer le pgcd de deux éléments de A ; en dehors du théorème de structure, intéressant par lui-même, il est pratique dans le cas des anneaux euclidiens. Dans ce cas la matrice U du lemme II-15 est produit de matrices élémentaires.
2 Si $\{\epsilon_j\}$, $(1 \leq j \leq n)$ désigne la base canonique de A^n, la base adaptée construite est formée des $\{e_j = g^{-1}(\epsilon_j)\}$, $(1 \leq j \leq n)$ comme on l'a vu à la fin du second paragraphe. La construction matricielle montre que ces e_j sont les vecteurs colonnes de la matrice U^{-1}. Cette dernière est le produit des matrices élémentaires dont on connaît les inverses. Cette remarque permet de construire aisément la matrice U^{-1}.

§4 APPLICATIONS.

On établit le théorème de structure des modules de type fini sur un anneau principal :

Corollaire II-19 : *Étant donné L un A-module de type fini sur un anneau principal, il existe des éléments $q_1|q_2|\ldots|q_s$ et un entier r tels que :*

$$L \simeq \prod_{i=1}^{s} A/q_iA \times A^r.$$

Preuve : Le module L est le quotient de A^n par un sous-module N. Dans une base adaptée $\{e_1, \ldots, e_n\}$ de A^n de pour $N \simeq q_1Ae_1 \oplus \ldots \oplus q_sAe_s$, avec $q_1|q_2|\ldots|q_s$; on pose $q_{s+1}, \ldots, q_n = 0$. En appliquant le lemme ci-dessus on obtient :

$$A^n/N \simeq \oplus_{i=1}^{n} A\epsilon_i / \oplus_{i=1}^{n} q_iA\epsilon_i \simeq \prod_{i=1}^{n} A/q_iA \simeq \prod_{i=1}^{s} A/q_iA \times A^r.$$

Bien entendu la question se pose de l'unicité des entiers r, s et des éléments q_i.

Proposition II-20 : *Dans le théorème de structure des modules de type fini sur un anneau principal, les entiers r et s sont unique de même pour les idéaux $(q_1) \supset (q_2) \supset \ldots \supset (q_s)$.*

Preuve : On suppose que l'on a deux décompositions pour un module X de type fini sur A (anneau principal) : $X \simeq \prod_{i=1}^{s} A/q_iA \times A^r \simeq \prod_{i=1}^{s'} A/q'_iA \times A^{r'}$. Le A-module sans torsion $X/\mathrm{Tor}(X)$ est isomorphe à la fois à $(X/\prod_{i=1}^{s} A/q_iA) \simeq A^r$ et à $(X/\prod_{i=1}^{s'} A/q'_iA) \simeq A^{r'}$. Comme les bases d'un module libre sur un anneau ont le même nombre d'éléments on a l'égalité de $r = r'$ (c'est le rang de X). Il suffit maintenant de traiter le cas d'un module de torsion. On suppose désormais $X \simeq \prod_{i=1}^{s} A/q_iA \simeq \prod_{i=1}^{s'} A/q'_iA$. On se propose, dans un premier temps de démontrer que $s = s'$.

On suppose qu'il existe un même élément irréductible p diviseur commun de q_1 et de q'_1. La propriété de divisibilité des q_i et des q'_i permet d'écrire $q_i = pr_i$, $q'_i = pr'_i$ avec la condition de divisibilité $r_i|r_{i+1}$ $(1 \leq i \leq s-1)$, $r'_i|r'_{i+1}$ $(1 \leq i \leq s'-1)$. Construisons le module quotient X/pX, on a les isomorphismes $X/pX \simeq \prod_{i=1}^{s} \frac{A/pr_iA}{pA/pr_iA} \simeq \prod_{i=1}^{s} A/pA$ et de même pour l'autre décomposition. Le quotient X/pX est un A/pA-espace vectoriel de de dimension s et s', donc $s = s'$.

Si q_1 et q'_1 sont premiers entre eux, soit p un diviseur irréductible de q_1, p' un diviseur irréductible de q'_1. On suppose que $p \not| q'_i$ $(1 \leq i \leq u)$. En remarquant que pour $1 \leq i \leq u$, p est inversible dans A/q'_iA on a d'une part $X/pX \simeq \prod_{i=1}^{s} A/pA$ et d'autre part $X/pX \simeq \prod_{i=u+1}^{s'} A/pA$ ce qui nous donne en regardant la dimension de ces A/pA-espaces vectoriels $s = s' - u$ soit $s < s'$. Le même travail avec p' en échangeant les rôles des deux décomposition montre $s' < s$ ce qui est contradictoire. Les diviseurs q_1 et q'_1 ont un facteur irréductible commun et donc $s = s'$.

Il reste à établir l'unicité des diviseurs élémentaires. On va raisonner par récurrence en utilisant la fonction σ pour $\sigma(q_1 \ldots q_s q'_1 \ldots q'_s)$. Le résutat est immédiat si elle vaut 0 (car alors $X = 0$). On garde les notations précédentes et on calcule $pX \simeq \prod_{i=1}^{s} pA/pr_iA \simeq \prod_{i=1}^{s} A/r_iA \simeq \prod_{i=1}^{s} A/r'_iA$. Pour ce module la fonction σ a diminué de $2s$, on peut lui appliquer l'hypothèse de récurrence les suites des r_i et des r'_i sont identiques, il en est donc de même pour celles des q_i et des q'_i. □

Le cas $A = \mathbb{Z}$ nous donne immédiatement la structure des groupes abéliens de type fini :

Corollaire II-21 : *Soit G un groupe abélien de type fini, il existe un unique entier r et une suite unique d'entiers $1 < q_1|q_2|\ldots|q_s$ tels que $G \simeq \mathbb{Z}^r \times \prod_{i=1}^{s} \mathbb{Z}/q_i\mathbb{Z}$.*

Preuve : C'est une application immédiate de la classification précédente. On en trouvera une démonstration directe dans [C]. □

Un autre cas particulièrement intéressant est celui où $A = K[X]$, anneau des polynômes sur un corps K. Soit E un K-espace vectoriel de dimension finie, f un K endomorphisme de E. On munit E d'une structure de $K[X]$-module via f de la façon suivante : soit $m \in E$ et $P = \sum_{i=1}^{r} a_i X^i$, on pose $P.m = \sum_{i=1}^{r} a_i f^i(m)$. La vérification de la structure de $K[X]$-module ne présente aucune difficulté. Ce module est de type fini : une base $\epsilon_1, \ldots, \epsilon_n$ engendre E comme K-espace vectoriel et donc comme $K[X]$-module (penser aux polynômes constants). Le théorème de structure nous dit que $E \simeq_{K[X]} K[X]^r \times \prod_{i=1}^{s} K[X]/Q_i$, comme $K[X]$, au contraire de E, est de dimension infinie l'entier r est obligatoirement nul soit :

$$E \simeq_{K[X]} \prod_{i=1}^{s} K[X]/Q_i K[X] \quad Q_1|Q_2|\ldots|Q_s.$$

avec une base adaptée où, pour faciliter, on choisit les Q_i unitaires : $Q_i(X) = \sum_{j=0}^{d_i} c_{i,j} X^j$ avec $c_{i,d_i} = 1$.

Corollaire II-22 : *Le polynôme minimal de f est Q_s.*

Preuve : Par définition, le polynôme minimal de f est l'annulateur de $E \simeq_{K[X]} \prod_{i=1}^{s} K[X]/Q_i$ c'est le ppcm des Q_i. Étant donné leurs divisibilité, c'est bien Q_s. □

Les sous-$K[X]$-modules $K[X]/Q_i e_i$ sont stables par multiplication par X, ce sont donc des sous-espaces de E stables par f. Les bases naturelles pour ces K-espaces vectoriels sont les $\{e_i, Xe_i, \ldots, X^{d_i-1}e_i\}$ l'action de f étant la traduction de la multiplication par X, on obtient une matrice $d_i \times d_i$ de la forme :

$$\begin{pmatrix} 0 & 0 & 0 & \ldots & 0 & -c_{i,0} \\ 1 & 0 & 0 & \ldots & 0 & -c_{i,1} \\ 0 & 1 & 0 & \ldots & 0 & -c_{i,2} \\ \vdots & \ddots & \ddots & \ldots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & 1 & -c_{i,d_i-1} \end{pmatrix}$$

En recollant les divers blocs de cette forme, on obtient la décomposition de Frobenius de la matrice de f. Toute décomposition de ce type s'interprète comme une structure de $K[X]$-module de type fini pour E, comme on en a vu l'unicité, on obtient l'unicité de la décomposition de Frobenius.

Lorsque K est algébriquement clos, on retrouve la décomposition de Jordan. On note Ω l'ensemble des racines de Q_s, les racines des Q_i appartiennent toutes à cet ensemble ce qui permet d'écrire $Q_i = \prod_{\lambda \in \Lambda} (X - \lambda)^{s_{i,\lambda}}$. En utilisant le théorème des restes chinois dans l'anneau principal $K[X]$, puis en réordonnant les termes :

$$E \simeq \prod_{i=1}^{s} \prod_{\lambda \in \Lambda} \frac{K[X]}{(X-\lambda)^{n_{i,\lambda}} K[X]} \simeq \prod_{\lambda \in \Lambda} \prod_{i=1}^{s} \frac{K[X]}{(X-\lambda)^{n_{i,\lambda}} K[X]}.$$

Comme K-base de $\frac{K[X]}{(X-\lambda)^{n_{i,\lambda}} K[X]}$ on choisit $\{(X-\lambda)^{n_{i,\lambda}-1}, (X-\lambda)^{n_{i,\lambda}-2}, \ldots, (X-\lambda), 1\}$. On sait que l'action de f sur ce K-espace vectoriel équivaut à la multiplication par X, or $X(X-\lambda)^r = (X-\lambda)^{r+1} + \lambda(X-\lambda)^r$. On obtient la restriction la matrice de la restriction de f à ce sous-espace :

$$\begin{pmatrix} \lambda & 1 & 0 & \ldots & 0 & 0 \\ 0 & \lambda & 1 & \ldots & 0 & 0 \\ 0 & 0 & \lambda & \ddots & 0 & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & \lambda & 1 \\ 0 & 0 & 0 & \ldots & 0 & \lambda \end{pmatrix}$$

qui est la forme annoncée.

En pratique, on a remarqué que les ϵ_i $(1 \leq i \leq n)$ forment un système de générateurs de E comme $K[X]$-module. On construit $M = \oplus_{i=1}^{n} K[X]\epsilon_i \simeq K[X]^n$ et π un morphisme surjectif de $K[X]$-modules de M sur E. Notons $(a_{i,j})$ la matrice de f relativement à la base $\epsilon_1, \ldots, \epsilon_n$, comme l'action de X sur ϵ_j est la même que celle de f. On constate que les vecteurs $\sum_{i=1}^{n}(a_{i,j} - \delta_{i,j}X)\epsilon_i$ sont dans le noyau de σ.

On va montrer :

Lemme II-23 : *le noyau N de π est engendré par les n vecteurs $\sum_{i=1}^{n}(a_{i,j} - \delta_{i,j}X)\epsilon_i$.*

On en tire une première conséquence : La matrice R de l'algorithme de Smith est donc la matrice caractéristique de f relativement à la base $\epsilon_1, \ldots, \epsilon_n$! On applique l'algorithme de Smith et on obtient :

$$URV = \begin{pmatrix} Q_1 & \ldots & 0 & 0 & \ldots & 0 \\ \vdots & \ddots & \vdots & \vdots & & \vdots \\ 0 & \ldots & 1 & 0 & \ldots & 0 \\ 0 & \ldots & 0 & Q_i & \ddots & 0 \\ \vdots & \ddots & \vdots & \ddots & \ddots & \vdots \\ 0 & \ldots & 0 & 0 & \ldots & Q_n \end{pmatrix}$$

On a vu qu'il n'y a pas de 0 sur la diagonale (ils donneraient une dimension infinie à E). Les matrices U et V étant de déterminant 1 on obtient $\det R = Q_1 \ldots Q_n$ le produit des diviseurs élémentaires est égal au polynôme caractéristique !

Preuve : (d'après [G2]) Le $K[X]$-module $\oplus K[X]\epsilon_i$ est un K-espace vectoriel de dimension infinie dont une base est formée des vecteurs $X^j\epsilon_i$ $(j \in \mathbb{N}, 1 \leq i \leq n)$. Tout $W \in M$ peut donc s'écrire de manière unique $W = \sum_j X^j w_j$ où les w_j sont des vecteurs de E.

Soit donc $W = \sum X^k w_k \in \ker \sigma$, on veut montrer qu'il est de la forme $W = \sum_i X^i f(v_i) - X^{i+1}v_i$, les v_i étant des vecteurs de E.

Si on identifie $\sum X^k w_k = \sum_i X^i f(v_i) - X^{i+1}v_i$ on obtient les relations

$$\begin{aligned} w_0 &= f(v_0) \\ \vdots &= \vdots \\ w_k &= f(v_k) - v_{k-1} \\ \vdots &= \vdots \end{aligned}$$

qui se transforment en :

$$\begin{aligned} w_0 &= f(v_0) \\ f(w_1) &= f^2(v_1) - f(v_0) \\ f^2(w_2) &= f^3(v_2) - f^2(v_1) \\ \vdots &= \vdots \\ f^k(w_k) &= f^{k+1}(v_k) - f^k(v_{k-1}) \\ \vdots &= \vdots \end{aligned}$$

Mais, comme W est dans le noyau de σ on a $\sum f^i(w_i) = 0$ on obtient :

$$\sum_{0 \leq i \leq k} f^i(w_i) = -\sum_{k+1 \leq i} f^i(w_i) = f^{k+1}(-\sum_{k+1 \leq i} f^{i-k-1}(w_i))$$

ce qui conduit à poser $v_k = -\sum_{k+1\leq i} f^{i-k-1}(w_i)$. On a alors :

$$f(v_k) - v_{k-1} = -\sum k+1 \leq i f^{i-k}(w_i) + \sum_{k\leq i} f^{i-k}(w_i) = w_k$$

qui est bien ce que l'on recherchait. Par linéarité, on peut se restreindre aux générateurs de l'énoncé.□Tirons les conséquences de l'application de l'algorithme de Smith à la matrice $R = A - XI_n$. On obtient une matrice diagonale URV et une base adaptée formée des vecteurs $e_i = U^{-1}(\epsilon_i)$. Lorsque le degré de Q_i est égal à 0, e_i est dans le noyau de σ. L'image de σ est donc engendrée par les vecteurs $\sigma(X^j e_i)$ avec $0 \leq j \leq n_i = \text{degré} Q_i$, ils sont au nombre de n et ce sont les $f^j(\sigma(e_i))$. Comme $\sigma(Q_i(X)e_i) = 0 = Q_i(f)(\sigma(e_i))$ les vecteurs $\sigma(e_i)$ sont les générateurs de sous-espaces cycliques de f.

CHAPITRE III : QUATERNIONS ET ROTATIONS.

Le but de ce chapitre de compléments est de donner une autre approche du groupe des rotations de $\mathbb{R}^3$. On peut consulter [G2], [GS], [L], [LTJ], [P], [V].

§1 ALGÈBRE DES QUATERNIONS.

On considère l'ensemble $\mathbb{H}$ des matrices 2×2 à coefficients complexes de la forme $\begin{pmatrix} \alpha & \beta \\ -\overline{\beta} & \overline{\alpha} \end{pmatrix}$. C'est visiblement un sous $\mathbb{R}$-espace vectoriel de $M_2(\mathbb{C})$ de dimension 4 (vérifiez qu'il ne s'agit pas d'un sous-$\mathbb{C}$-espace vectoriel de $M_2(\mathbb{C})$). Il possède une base « évidente » : $I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, $E_1 = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}$, $E_2 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$, $E_3 = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}$.

Lemme III-1 : *$\mathbb{H}$ est un anneau.*

Preuve : Il suffit de vérifier que c'est un sous-anneau de $M_2(\mathbb{C})$. Comme c'est un sous-espace vectoriel réel et que I_2 en est un élément. Il suffit de vérifier sa stabilité par multiplication. Or :

$$\begin{pmatrix} \alpha & \beta \\ -\overline{\beta} & \overline{\alpha} \end{pmatrix} \begin{pmatrix} \alpha' & \beta' \\ -\overline{\beta'} & \overline{\alpha'} \end{pmatrix} = \begin{pmatrix} \alpha\alpha' - \beta\overline{\beta'} & \alpha\beta' + \overline{\alpha'}\beta \\ -\overline{\alpha}\overline{\beta'} - \alpha'\overline{\beta} & \overline{\alpha}\overline{\alpha'} - \overline{\beta}\beta' \end{pmatrix} = \begin{pmatrix} \alpha\alpha' - \beta\overline{\beta'} & \alpha\beta' + \overline{\alpha'}\beta \\ -\left(\overline{\alpha\beta' + \overline{\alpha'}\beta}\right) & \left(\overline{\alpha\alpha' - \beta\overline{\beta'}}\right) \end{pmatrix}$$

qui est bien de la même forme. □

Si on considère la forme hermitienne f sur $\mathbb{C}^2$ dont la matrice relativement à la base canonique est $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$: $f\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}\right) = {}^t\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \overline{\begin{pmatrix} y_1 \\ y_2 \end{pmatrix}}$ la matrice adjointe de l'endomorphisme représenté par $q = \begin{pmatrix} \alpha & \beta \\ -\overline{\beta} & \overline{\alpha} \end{pmatrix}$ – la matrice q^* telle que $f(q(X), Y) = f(X, q^*(Y))$ – est $q^* = \begin{pmatrix} \overline{\alpha} & -\beta \\ \overline{\beta} & \alpha \end{pmatrix} = {}^t\overline{q}$.

Définition III-2 : *Le quaternion q^* s'appelle le conjugué du quaternion q.*

Il est immédiat que $(q^*)^* = q$, $(q_1q_2)^* = q_2^*q_1^*$. En effectuant le produit qq^* on constate immédatement que $qq^* = q^*q = \det(q)I_2 = (\alpha\overline{\alpha} + \beta\overline{\beta})I_2$. On en déduit :

Corollaire III-3 : *$\mathbb{H}$ est un corps.*

Remarques :
1- Si $0 \neq q$, $q^{-1} = \frac{1}{\det(q)}q^*$.
2 La multiplication dans $\mathbb{H}$ se déduit par distributivité de la multiplication entre les éléments E_i, E_j de la base :

$$E_1^2 = E_2^2 = E_3^2 = -I_2; \quad E_1E_2 = -E_2E_1 = E_3; \quad E_2E_3 = -E_3E_2 = E_1; \quad E_3E_1 = -E_1E_3 = E_2$$

Ce corps est non commutatif.

En particulier si $q = x_0I + x_1E_1 + x_2E_2 + x_3E_3$ on a $q^* = x_0I - x_1E_1 - x_2E_2 - x_3E_3$ et $\det(q) = qq^* = x_0^2 + x_1^2 + x_2^2 + x_3^2$. Le déterminant munit $\mathbb{H}$ d'une forme quadratique définie positive. On comprend alors :

Définition III-4 : *Pour un quaternion q, le nombre $\det(q)$ s'appelle la norme de q. Les quaternions de norme 1 sont appelés les quaternions unitaires.*

On a la proposition évidente :

Proposition III-5 : *Les quaternions de norme 1 forment un groupe, le groupe $SU(2, \mathbb{C})$.*

§2 QUATERNIONS PURS.

Un quaternion étant une matrice carrée, on peut parler de sa trace. Il est évident que $Tr(q)I = q + q^*$ ($Tr(q) = 2x_0$ en utilisant l'écriture suivant la base).

Définition III-6 : *On appelle quaternion pur les quaternions de trace nulle. On appelle quaternions réels ceux égaux à leur conjugué.*

Les quaternions purs forment un sous-espace de $\mathbb{H}$ de dimension 3 de base E_1, E_2, E_3.

Proposition III-7 : *Pour un quaternion pur u, les propriétés suivantes sont équivalentes :*

i) u est unitaire

ii) $u^2 = -I$

iii) $x_1^2 + x_2^2 + x_3^2 = 1$.

Preuve : L'équivalence de i) et de iii) se ramène aux définitions. Pour i) et ii) on remarque $u^* = -u$ puisque u est pur, il est unitaire si et seulement si $I = uu^* = -u^2$. □

L'espace vectoriel des quaternions pur est donc un espace euclidien A_0 de dimension 3 pour la forme quadratique det dont une base orthonormée est formée des vecteurs E_1, E_2, E_3. On note $<\quad,\quad>$ le produit scalaire correspondant. On oriente l'espace en choisissant $\{E_1, E_2, E_3\}$ comme base directe.

Théorème III-8 : *Soit u et v deux quaternions purs, alors :*

a) $< u, v >= 0$ *(u et v sont orthogonaux) si et seulement* $uv + vu = 0$.

b) *Si u et v sont unitaires et orthogonaux les vecteurs u, v, $w = uv$ forment une base orthonormée directe.*

Preuve :

a) posons $u = x_1E_1 + x_2E_2 + x_3E_3$, $v = y_1E_1 + y_2E_2 + y_3E_3$. On calcule uv et vu :

$uv = -(x_1y_1 + x_2y_2 + x_3y_3) + (x_2y_3 - x_3y_2)E_1 + (x_3y_1 - x_1y_3)E_2 + (x_1y_2 - x_2y_1)E_3$

$vu = -(x_1y_1 + x_2y_2 + x_3y_3) + (x_3y_2 - x_2y_3)E_1 + (x_1y_3 - x_3y_1)E_2 + (x_2y_1 - x_1y_2)E_3$

En additionnant on obtient $uv + vu = -2 < u, v >$ ce qui démontre le a).

b) Si u et v sont orthogonaux la formule calculée précédemment montre que $uv = u \wedge v$, ce qui termine la démonstration. □

§3 QUATERNIONS UNITAIRES.

Soit $x = x_1E_1 + x_2E_2 + x_3E_3$ un quaternion pur. On note $\det(x) = x_1^2 + x_2^2 + x_3^2 = \delta^2$; si $x \neq 0$, le quaternion $\frac{x}{\delta}$ est pur et unitaire et tout quaternion pur peut s'écrire $x = \delta u$ avec $\delta \in \mathbb{R}$ et u unitaire et pur.

Soit $q = x_0I + x_1E_1 + x_2E_2 + x_3E_3$ un quaternion unitaire, on peut l'écrire $q = x_0I + \delta u$ avec u quaternion pur et unitaire et $x_0^2 + \delta^2 = 1$. Il existe donc un unique θ défini modulo 2π tel que $q = \cos\theta I + \sin\theta u$ (cette formule reste valable dans le cas d'un quaternion unitaire et réel avec $\theta = \pm\pi$).

Si $q = \cos\theta I + \sin\theta u$ est un quaternion unitaire écrit avec les conventions précédentes, on a $q^{-1} = \cos\theta I - \sin\theta u$

§4 LE MORPHISME de $SU(2,\mathbb{C})$ sur $SO(3,\mathbb{R})$.

Définition III-9 : *On fixe a un quaternion unitaire et à tout quaternion pur x on associe $\gamma_a(x) = axa^{-1}$.*

Théorème III-10 : *L'application γ_a est une rotation de l'espace euclidien A_0.*

Preuve : Il est évident que γ_a est linéaire bijective ; comme tous les éléments sont des matrices 2×2 à coefficients complexes $Tr(\gamma_a(x)) = Tr(x) = 0$, $\det(\gamma_a(x)) = \det(x)$ et γ_a est une isométrie de l'espace euclidien A_0 dans lui-même. Enfin si on choisit une base orthonormée directe x, y, $w = xy$ de A_0. Comme on a une isométrie, $\gamma_a(x)$ et $\gamma_a(y)$ sont des quaternions purs, unitaires et orthogonaux ; on a une nouvelle base orthonormée directe : $\gamma_a(x)$, $\gamma_a(y)$, $\gamma_a(x)\gamma_a(y) = axa^{-1}aya^{-1} = awa^{-1} = \gamma_a(xy)$. L'isométrie conserve l'orientation de l'espace euclidien A_0 : c'est une rotation. □

La proposition suivante est évidente.

Proposition III-11 : *L'application $a \mapsto \gamma_a$ est un morphisme de groupe de $SU(2,\mathbb{C})$ dans $SO(3,\mathbb{R})$.*

Précisons maintenant l'axe et l'angle de la rotation γ_a. Si a est réel et unitaire $a = \pm I$ et γ_a est l'identité sur A_0.

Théorème III-12 : *La rotation γ_a de A_0 associée au quaternion unitaire $a = \cos\theta I + \sin\theta u$ a pour axe la droite portée par u et pour angle 2θ.*

Preuve : Déjà vu lorsque a est réel. Si $a = \cos\theta I + \sin\theta u$ n'est pas réel son inverse est $a^{-1} = \cos\theta I - \sin\theta u$ et on a $\gamma_a(u) = (\cos\theta I + \sin\theta u)u(\cos\theta I - \sin\theta u) = u$. L'axe de la rotation est bien porté par u.

Pour déterminer l'axe de la rotation, il suffit d'étudier l'image d'un vecteur orthogonal à u. Soit v un tel vecteur : $\gamma_a(v) = (\cos\theta I + \sin\theta u)v(\cos\theta I - \sin\theta u) = \cos^2\theta v + \sin\theta\cos\theta uv - \sin\theta\cos\theta uv - \sin^2\theta uvu$. Comme u et v sont orthogonaux $uv = -vu$ et on trouve finalement $\gamma_a(v) = \cos(2\theta)v + \sin 2\theta uv$. □

Corollaire III-13 : *Le morphisme de groupe $a \in SU(2,\mathbb{C}) \to \gamma_a \in SO(3,\mathbb{R})$ est surjectif, de noyau $\pm I$.*

§5 **SIMPLICITÉ de** $SO(3,\mathbb{R})$ (démonstration géométrique).

Voir [P] où on trouvera des compléments et des références bibliographiques.

On suppose connu que le groupe des rotations $SO(3,\mathbb{R})$ est engendré par les retournements (rotations d'angle π).

Théorème III-14 : *Le groupe $SO3(\mathbb{R})$ est un groupe simple.*

Preuve : On suppose donné H sous-groupe distingué de $SO3(\mathbb{R})$, $H \neq \{I_3\}$ et on se propose de démontrer que $H = SO3(\mathbb{R})$. Pour cela, il suffit de prouver que les retournements de l'espace euclidien E sont tous dans H.

Lemme III-15 : *Deux retournements de E sont toujours conjugués dans $SO3(\mathbb{R})$.*

Preuve : Soit r (resp. s) un retournement et $\vec{r}$ (resp. $\vec{s}$) un vecteur unitaire de l'axe du retournement. On remarque que $\vec{r} = \frac{\vec{r}+\vec{s}}{2} + \frac{\vec{r}-\vec{s}}{2}$, que $\vec{s} = \frac{\vec{r}+\vec{s}}{2} - \frac{\vec{r}-\vec{s}}{2}$ et que les vecteurs $\frac{\vec{r}+\vec{s}}{2}$, $\frac{\vec{r}-\vec{s}}{2}$ sont orthogonaux ; le retournement σ d'axe $\vec{r}+\vec{s}$ est donc tel que $\sigma(\vec{r}) = \vec{s}$ et $\sigma(\vec{s}) = \vec{r}$.

Calculons la rotation $\sigma r \sigma^{-1}$ (remarquons que σ étant un retournement son inverse lui est égal). C'est une rotation comme produit de rotations. Si α est l'angle de cette rotation $1 + 2\cos(\alpha) = \mathrm{Tr}(\sigma r \sigma^{-1})$, les propriétés de la trace montrent que c'est égal à la trace de r soit $1 + 2\cos(\pi) = -1$. Il s'ensuit que $\cos(\alpha) = -1$ et que $\alpha = \pi$. L'image de $\vec{s}$ par $\sigma r \sigma^{-1}$ est $\sigma r \sigma^{-1}(\vec{s}) = \sigma r(\vec{r}) = \sigma(\vec{r}) = \vec{s}$. La rotation $\sigma r \sigma^{-1}$ est le retournement d'axe $\vec{s}$. □

Il suffit donc, puisque H est distingué, de montrer qu'un retournement est inclus dans H. Ceci résulte de l'enchaînement de plusieurs lemmes.

Lemme III-16 : *Supposons qu'il existe $h \in H$ et $\vec{x} \in E$ tels que $\vec{x}$ et $h(\vec{x})$ soient orthogonaux, il existe un retournement dans H.*

Preuve : Comme $\vec{x}$ et $h(\vec{x})$ sont orthogonaux, il en est de même pour $h^{-1}(\vec{x})$ et $h^{-1}h(\vec{x}) = \vec{x}$. Soit le retournement ρ d'axe porté par $\vec{x}$ et construisons $r = \rho h \rho^{-1} h^{-1} = \rho h \rho h^{-1}$. Si on l'écrit $(\rho h \rho^{-1})h^{-1}$ on voit que c'est un élément de H.

Calculons l'image de $h(\vec{x})$ par r : $\rho h \rho^{-1} h^{-1}(h(\vec{x})) = \rho h \rho^{-1}(\vec{x}) = \rho(h(\vec{x})) = -h(\vec{x})$.

Calculons l'image de $\vec{x}$ par r : $\rho h \rho(h^{-1}(\vec{x})) = \rho h(-h^{-1}(\vec{x}) = -\rho(\vec{x}) = -\vec{x}$.

Deux vecteurs orthogonaux sont transformés par r en leurs opposés, le vecteur orthogonal à ces deux vecteurs est fixe (le déterminant est 1, la troisième valeur propre est égale à 1) ; r est un retournement d'axe orthogonal à $\vec{x}$ et $h(\vec{x})$. □

Lemme III-17 : *Si une rotation f est telle que $\cos(\theta) < 0$ alors il existe un vecteur $\vec{v}$ tel que $\vec{v}$ et $f(\vec{v})$ soient orthogonaux.*

Preuve : Soit $\vec{p}$ l'axe de la rotation et $\vec{u}$, $\vec{w}$, $\vec{p}$ une base orthonormé de E. Écrivons un vecteur $\vec{v}$
$\vec{v} = x\vec{u} + y\vec{w} + z\vec{p}$, $f(\vec{v}) = (x\cos(\theta) - y\sin(\theta))\vec{u} + (x\sin(\theta) + y\cos(\theta))\vec{w} + z\vec{p}$. Le produit scalaire $< \vec{v}, f(\vec{v}) >$ est égal à $z^2 + \cos(\theta)(x^2 + y^2)$. Comme $\cos(\theta) < 0$ les valeurs x, y, z peuvent être choisies de sorte que le produit scalaire soit nul. □

Lemme III-18 : *Soit g une rotation d'angle θ non nul modulo 2π. Il existe n tel que $\cos(2^n\theta) < 0$.*

Preuve : Quitte à changer l'axe de la rotation, on peut supposer l'angle compris entre 0 et π. La somme $\sum_{n=1}^{\infty} \frac{1}{2^n}$ vaut 1, on écrit $\theta = \sum_{n=1}^{\infty} \frac{\epsilon_n}{2^n}\pi$ avec $\epsilon_n = 0$ ou 1.

Supposons $\cos(2^n\theta) > 0$ pour tout n (s'il est nul pour un n, au coup suivant, il vaut -1). Pour $n = 0$ $\cos(\theta) > 0$ et $0 < \theta < \pi/2$, comme la somme $\sum_{n=2}^{\infty} \frac{1}{2^n}$ vaut $1/2$ c'est que $\epsilon_1 = 0$.

Soit n le premier indice tel que $\epsilon_n = 1$, $2^{n-1}\theta = \frac{\pi}{2} + \ldots$ où la somme restante est inférieure à $\pi/2$ on a donc $\cos(2^{n-1}\theta) \leq 0$ ce qui est contraire à l'hypothèse que nous venons de faire. Tous les ϵ_k sont nuls et $\theta = 0$. □

Cette étape complète la démonstration. □

CHAPITRE IV : RÉSEAUX

Ce chapitre est une introduction aux sous-groupes de $(\mathbb{R}^n, +)$, nous ne intéressons qu'à des sous-groupes de type finis, on utilisera naturellement le théorème de structure des groupes abéliens de type fini (ici, sans torsion !). Nous verrons comment les appliquer à des questions d'approximation et à de problèmes de représentation d'entiers comme somme de carrés, il faudra alors faire appel à des propriétés des corps finis et des quaternions. La référence principale est [GS], on pourra aussi consulter [G2], [Sa]. Le lecteur qui voudrait approfondir ses connaissances sur la question pourra consulter avec profit

CONWAY & SLOANE *Sphere Packings, Lattices and Groups* Springer Verlag (1993),

MILNOR & HUSEMOLLER *Symmetric bilinear forms* Springer (1973) ou

MARTINET *Les réseaux parfaits des espaces euclidiens* Masson (1996).

Dans tout ce qui suit $\mathbb{R}^n$ est muni de la structure euclidienne pour laquelle la base canonique $\{\epsilon_1, \ldots, \epsilon_n\}$ est une base orthonormée.

§1 SOUS GROUPES DE $\mathbb{R}^n$, SOUS GROUPES DISCRETS.

Définition IV-1 : *Soit G un sous-groupe de $\mathbb{R}^n$, on dit que G est discret si tout compact de $\mathbb{R}^n$ ne contient qu'un nombre fini de points de G.*

Par exemple $\mathbb{Z}^n$ est un sous-groupe discret de $\mathbb{R}^n$. On peut donner des propriétés équivalentes à la définition :

Lemme IV-2 : *Soit G un sous-groupe de $\mathbb{R}^n$, les conditions suivantes sont équivalentes :*

(i) *pour tout compact C de $\mathbb{R}^n$, $C \cap G$ est fini ;*

(ii) *pour tout $\epsilon > 0$ la boule fermée $B(0, \epsilon)$ ne contient qu'un nombre fini de points de G ;*

(iii) *il existe $\eta > O$ tel que $B(0, \eta) \cap G = \{0\}$.*

Preuve : (i)⇒(ii) car $B(0, \epsilon)$ est un compact.

(ii)⇒(iii) soit $\epsilon > 0$, l'ensemble $B(0, \epsilon) \cap G$ est fini, on note $B(0, \epsilon) \cap G = \{0, g_1, g_2, \ldots, g_k\}$ et on pose $\eta = \frac{1}{2} \inf_i \{\|g_1\|, \|g_2\|, \ldots, \|g_k\|\}$, la boule $B(0, \eta)$ convient.

(iii)⇒(i) Soit C un compact de $\mathbb{R}^n$ tel que $C \cap G$ soit infini. Cet ensemble possède un point d'accumulation x, il existe une suit $\{g_i\}_{i \in \mathbb{N}}$ de points distincts qui convergent vers x. Si on se donne $\eta > 0$, il existe une suite g_{i_k} tel que $\|g_{i_k} - g_{i_k+1}\| < \frac{\|g_{i_k-1} - g_{i_k}\|}{2}$. La suite des $(g_{i_k} - g_{i_k+1})_k$ contient une suite infinie d'éléments non nuls et distincts de G qui converge vers 0, ce qui contredit (iii). □

Faire une démonstration utilisant le théorème de Borel-Lebesgue.

On peut associer, à un sous-groupe G de $\mathbb{R}^n$, deux notions de dimension :

Définition IV-3 : *On note $\mathbb{R}G$ le sous-espace vectoriel de $\mathbb{R}^n$ engendré par G, sa dimension (inférieure ou égale à n) s'appelle la dimension réelle de G.*

Définition IV-4 : *Puisque l'on a supposé G de type fini (et sans torsion, puisqu'inclus dans $\mathbb{R}^n$) il est libre de type fini et possède une base $\{g_1, \ldots, g_r\}$, r est le rang de G, c'est aussi la dimension du $\mathbb{Q}$-espace vectoriel engendré par G dans $\mathbb{R}^n$.*

Il y a bien sur dans $\mathbb{R}^n$ des sous-groupes qui ne sont pas de type fini, par exemple $(\mathbb{Q}, +)$ n'est pas un sous-groupe de type fini de $(\mathbb{R}, +)$.

Exercice **1** : Soit G le sous-groupe de $\mathbb{R}$ engendré par 1 et $\sqrt{2}$, montrez que son rang est 2 et que sa dimension réelle est 1.

§2 RÉSEAUX.

Définition IV-5 : *Soit $Gl_n(\mathbb{R})$ le groupe des $\mathbb{R}$-automorphismes du $\mathbb{R}$-espace vectoriel $\mathbb{R}^n$, on appelle réseau de $\mathbb{R}^n$ toute image de $\mathbb{Z}^n$ par un élément $f \in Gl_n(\mathbb{R})$.*

Il est alors $\mathbb{Z}$-libre de rang n, de rang réel n et c'est un sous-group discret ($\mathbb{Z}^n$ est discret et le réseau en est l'image par un homéomorphisme).

Définition IV-6 : *On appelle réseau relatif de $\mathbb{R}^n$ tout sous-groupe d'un réseau L de $\mathbb{R}^n$, si le rang de ce sous-groupe est n c'est aussi un réseau, on dit que c'est un sous-réseau de L.*

Lemme IV-7 : *Soit L un réseau de $\mathbb{R}^n$, $f, g \in Gl_n(\mathbb{R})$ tels que $f(\mathbb{Z}^n) = g(\mathbb{Z}^n) = L$, alors* $\det(f) = \pm \det(g)$.

Preuve : L'endomorphisme $f \circ g^{-1}$ de $\mathbb{R}^n$ est un automorphisme de L, $\mathbb{Z}$-module libre de rang n, c'est donc une matrice de $Gl_n(\mathbb{Z})$ donc de déterminant ± 1. □

Définition IV-8 : *L'invariant* $\det(f)$ *défini à partir de n'importe quel* $f \in Gl_n(\mathbb{R})$ *tel que* $f(\mathbb{Z}^n) = L$ *s'appelle le déterminant du réseau.*

Définition IV-9 : *Soit* $L = f(\mathbb{Z}^n)$ $(f \in Gl_n(\mathbb{R}))$ *un réseau,* $\{e_1 = f(\epsilon_1), \ldots, e_n = f(\epsilon_n)\}$ *une base de ce réseau, on appelle maille du réseau (ou parallélotope fondamental) l'ensemble* $\mathcal{P}(e_1, \ldots, e_n) = \{x \in \mathbb{R}^n | x = \sum_{i=1}^{n} \lambda_i e_i, 0 \leq \lambda_i < 1\}$.

Proposition IV-10 : *L'ensemble* $\mathcal{P}(e_1, \ldots, e_n)$ *est mesurable et sa mesure (pour la mesure de Lebesgue) ne dépend pas du choix de la base du réseau, elle est égale à* $|\det(M)|$ *où* M *est une quelconque des matrices de* $Gl_n(\mathbb{R})$ *telle que* $M(\mathbb{Z}^n) = L$*, on la note* μ_L.

Preuve : La maille $\mathcal{P}(\epsilon_1, \ldots, \epsilon_n)$ de $\mathbb{Z}^n$ est mesurable, par construction de la mesure de Lebesgue sur $\mathbb{R}^n$. La maille de L est son image par un difféomorphisme, elle est donc mesurable, la formule du changement de variable donne le résultat. □

Corollaire IV-11 : Soit L' un sous-réseau de L, l'indice $[L : L']$ est fini et $\mu_L = [L : L']\mu_{L'}$.

Preuve : Puisque L' est un sous-$\mathbb{Z}$-module de L de même rang, il existe une base $\{e_1, \ldots, e_n\}$ de L adaptée à L', donc des entiers $d_1, \ldots, d_n$ non nuls tels que $\{d_1 e_1, \ldots, d_n e_n\}$ est une base de L'. On en déduit $f \in Gl_n(\mathbb{R})$ tel que $f(\mathbb{Z}^n) = L$ en envoyant ϵ_j sur e_j, si on multiplie la $j$ème colonne de f par d_j, on obtient $f' \in Gl_n(\mathbb{Z})$ tel que $f'(\mathbb{Z}^n) = L'$. Les définitions précédentes donnent $\mu_{L'} = d_1 \ldots d_n \mu_L$. On termine en remarquant que $[L : L'] = |L/L'| = \prod_{i=1}^{n} d_i$. □

§3 THÉORÈMES DE JACOBI-BRAVAIS ET DE MINKOWSKI.

Énonçons le théorème de Jacobi-Bravais :

Théorème IV-12 : *Soit* G *un sous-groupe discret de* $\mathbb{R}^n$ *de dimension réelle* r*, alors* G *est un réseau relatif de rang* r.

Preuve : Puisque l'espace vectoriel engendré par les éléments de G est de dimension r, on peut extraire de G un système maximal de r éléments de G linéairement indépendants. Notons $e_1, \ldots, e_r$ ces éléments. On construit $\mathcal{P}_1(e_1, \ldots, e_r)$ comme on a construit la maille d'un réseau. L'adhérence $\overline{\mathcal{P}}_1$ est un compact et donc $\mathcal{P}_1 \cap G$ est un ensemble fini $(e'_1, \ldots, e'_k)$; on réunit ces deux ensembles $\{e_1, \ldots, e_r, e'_1, \ldots, e'_k\}$. On va montrer que l'on a construit un système de générateurs de G ; on aura établi que G est de type fini.

Soit $x \in G$, on peut l'écrire $x = \sum_{i=1}^{r} \lambda_i e_i$, $\lambda_i \in \mathbb{R}$ que l'on transforme en $\sum_{i=1}^{r} [\lambda_i] e_i + \sum_{i=1}^{r} (\lambda_i - [\lambda_i]) e_i$ avec $[\lambda_i]$ la partie entière de λ_i. L'élément $\sum_{i=1}^{r} (\lambda_i - [\lambda_i]) e_i$ est, par soustraction, dans G et par construction dans $\mathcal{P}_1$ c'est donc un e'_j $(1 \leq j \leq k)$.

Il reste à prouver que le rang de G comme $\mathbb{Z}$-module est bien r. Reprenons $x \in G$, $x = \sum_{i=1}^{r} \lambda_i e_i$ et construisons la suite $x^{(m)} = mx - \sum_{i=1}^{r} [m\lambda_i] e_i$ d'éléments de $\mathcal{P}_1 \cap G$. Comme cet ensemble est fini il existe $j \neq \ell$ tels que $x^{(j)} = x^{(\ell)}$ ce qui donne $(j - \ell)x = \sum_{i=1}^{r} ([j\lambda_i] - [\ell\lambda_i]) e_i$. Si on applique ce procédé aux e'_s, on en déduit qu'il existe d tel que $dG \subset \mathbb{Z}e_1 \oplus \ldots \oplus \mathbb{Z}e_r$, le groupe G est donc de rang au plus r, or il contient déjà r éléments indépendants sur $\mathbb{Z}$ (puisqu'indépendants sur $\mathbb{R}$). □

Un autre théorème important en théorie des réseaux est celui de Minkowski (pour des applications, voir [Sa] et Milnor-Husemoller cité plus haut).

Théorème IV-13 : *Soit* G *un réseau de* $\mathbb{R}^n$ *et* S *un sous-ensemble mesurable de* $\mathbb{R}^n$*. On suppose que* S *est symétrique par rapport à* 0*, qu'il est convexe et enfin que sa mesure de Lebesgue est telle que* $\mu(S) > 2^n \mu_G$ *(égalité large si on suppose* S *compact), alors* S *contient un élément de* G *autre que* 0.

Preuve : Soit $\{e_1, \ldots, e_n\}$ une base du réseau. C'est aussi une base de l'espace vectoriel $\mathbb{R}^n$. Montrons que les translatées de la maille associée à cette base forment une partition de $\mathbb{R}^n$. Tout d'abord, tout $x \in \mathbb{R}^n$ s'écrit de manière unique $\sum_{=1}^{n} \lambda_i e_i = \sum_{i=1}^{n} [\lambda_i] e_i + \sum_{i=1}^{n} (\lambda_i - [\lambda_i]) e_1$ donc $x = g + y$ avec $y \in \mathcal{P}$, la maille associée au réseau. Les translatées de la maille recouvrent $\mathbb{R}^n$. Ensuite, soit g et $g' \in G$, s'il existe $z \in (g + \mathcal{P}) \cap (g' + \mathcal{P})$ on a $z = g + u = g' + u'$ soit $g - g' = u' - u$. Or si $u = \sum_{=1}^{n} \lambda_i e_i$, $u' = \sum_{=1}^{n} \lambda'_i e_i$ $(0 \leq \lambda_i, \lambda'_i < 1)$ on aura à la fois $\lambda_i - \lambda'_i \in \mathbb{Z}$ et $|\lambda_i - \lambda'_i| < 1$ ce qui implique que pour tout i, $\lambda_i = \lambda'_i$ soit $u = u'$ et $g = g'$. Deux translatées de la maille ont une intersection non vide si et seulement si elles sont égales.

Donnons d'abord la démonstration dans le cas de l'inégalité stricte. Soit $S' = \frac{1}{2}S$. Cette partie est recouverte par les translatées de la maille : $S' = \cup_{g\in G}(g+\mathcal{P})\cap S'$ la réunion étant disjointe. Comme S est mesurable, il en est de même de S' et des $(g+\mathcal{P})\cap S'$. La réunion étant disjointe $\mu(S') = \sum_g \mu((g+\mathcal{P})\cap S')$, en utilisant l'invariance de la mesure par translation, $\mu(S') = \sum_g \mu(\mathcal{P}\cap(S'-g))$. L'additivité de la mesure implique qu'il existe g et g' distincts tel que $(S'-g)\cap(S'-g') \neq \emptyset$: il existe $s, s' \in S'$ tels que $0 \neq s-s' \in G$. Établissons que $s-s' \in S$. Il suffit d'écrire $s-s' = \frac{1}{2}(2s-2s')$: les éléments $2s, 2s' \in S$, comme S est symétrique $-2s' \in S$, comme S est convexe $\frac{1}{2}(2s+(-2s')) \in S$. Ceci termine la démonstration dans le cas de l'inégalité stricte.

Supposons S compact et l'inégalité large. On construit l'homothétique $(1+\epsilon)S$, il vérifie les hypothèse avec l'inégalité stricte. On déduit l'existence de points $f_\epsilon \in G\cap(1+\epsilon)S$. Si on choisit les $\epsilon \leq 1$, tous ces f_ϵ appartiennent à l'ensemble fini $2S\cap(G\setminus\{0\})$. Il existe donc une infinité de $k>0$ tel que $f_{1/k} = f \in (G\setminus\{0\})$ soit fixe. ceci implique $f \in \cap_k(1+\frac{1}{k})S = S$. □

§4 APPLICATIONS À DES PROBLÈMES DIOPHANTIENS.

Premier problème :

Soit $x = (x_1, \ldots, x_n) \in \mathbb{R}^n \setminus \mathbb{Q}^n$. On construit le sous-groupe $\mathbb{Z}^n + \mathbb{Z}x$.

Lemme IV-14 : *Le sous-groupe G est de type fini et non discret.*

Preuve : Qu'il soit de type fini est évident. S'il était discret, il serait libre de rang n (d'après le théorème de Jacobi-Bravais) et admettrait $\mathbb{Z}^n$ comme sous-réseau. Il possède une base $\{e_1, \ldots, e_n\}$ adaptée à ce sous réseau. Il existe des entiers $d_1|\ldots|d_n$ tels que $\{d_1e_1, \ldots, d_ne_n\}$ soit une base de $\mathbb{Z}^n$ et des entiers $a_1, \ldots, a_n$ tels que $x = \sum_{i=1}^n a_ie_i = \sum_{i=1}^n \frac{a_i}{d_i}d_ie_i$, l'élément $x \in \frac{1}{d_n}\mathbb{Z}^n$ ce qui contredit l'hypothèse. □

Sous les hypothèses que l'on vient de faire, $\forall\epsilon > 0$, le fermé $\{z = (z_1, \ldots, z_n) \in \mathbb{R}^n |\ \forall i\ |z_i| \leq \epsilon\}$ contient un élément y non nul de G, puisque ce dernier n'est pas discret. On peut écrire $y = \sum_{i=1}^n a_ie_i + ax$ avec $a \neq 0$ les a_i non tous nuls et la propriété $|a_i + ax_i| \leq \epsilon$. On en déduit :

Théorème IV-15 : *Soit $x_1, \ldots, x_n$, n réels, alors pour tout ϵ tel que $0 < \epsilon < 1$, il existe des rationnels $\frac{p_i}{q}$, q indépendant de i tels que :*

$$\text{pour } i = 1, \ldots, n \quad \left| x_i - \frac{p_i}{q} \right| < \frac{\epsilon}{q}.$$

Preuve : Le cas où les x_i sont tous rationnels est évident et sans intérêt. Si on suppose que $x = (x_i, \ldots, x_n) \notin \mathbb{Q}^n$ on applique le résultat de la discussion précédente avec $\epsilon' < \epsilon < 1$ que l'on récrit $|\frac{a_i}{a} - x_i| \leq \frac{\epsilon'}{a} < \frac{\epsilon}{a}$. □

Montrez que si $\alpha/\pi \notin \mathbb{Q}$ alors les $e^{in\alpha}$ forment une partie dense du cercle trigonométrique lorsque n parcourt $\mathbb{Z}$.

Second problème :

On se donne des entiers a_i, $i = 1, \ldots, n$, non tous nuls et un autre entier b ; on recherche les n-uples $(x_1, \ldots, x_n) \in \mathbb{Z}^n$ tels que :

$$(*) \qquad \sum_{i=1}^n a_ix_i = b.$$

Il s'agit donc de trouver les points à coordonnées entières de l'hyperplan affine $\sum_{i=1}^n a_ix_i = b$. On peut commencer par la démarche habituelle dans ce genre de situation :

Lemme IV-16 : *Si $x^0 = (x_1^0, \ldots, x_n^0)$ est une solution particulière de $(*)$ alors toute solution est de la forme $x = x^0 + y$ où y est solution de l'équation homogène : $\sum_{i=1}^n a_ix_i = 0$.*

L'existence d'une solution particulière est assurée par :

Lemme IV-17 : *Une condition nécessaire et suffisante pour que $(*)$ admette au moins une solution est que le pgcd d des a_i divise b.*

Preuve : La condition est nécessaire, puisque d divise le membre de gauche de $(*)$. Réciproquement si cette condition est vérifiée,on se ramène au cas où $d = 1$. Le membre de gauche représente les éléments de l'idéal de

$\mathbb{Z}$ engendrés par les a_i, comme ils sont premiers entre eux c'est $\mathbb{Z}$, et il existe $\xi_1, \ldots, \xi_n$ tels que $\sum_{i=1}^n a_i\xi_i = 1$, les $b\xi_i$ donne la solution cherchée. En utilisant l'associativité du pgcd on peut construire explicitement une telle solution. □

Lemme IV-18 : *L'ensemble des solutions de l'équation homogène est un réseau relatif de rang $n-1$ inclus dans $\mathbb{Z}^n$.*

Preuve : L'ensemble des solutions est stable par addition et par multiplication par un entier, c'est un sous-groupe de $\mathbb{Z}^n$. Il engendre un sous-espace vectoriel de dimension $n-1$. Si on suppose $a_1 \neq 0$ on a $n-1$ solutions linéairement indépendantes $y^i = (y_1^i, \ldots, y_n^i)$ $(2 \leq i \leq n)$ avec $y_1^i = -a_i$, $y_i^i = a_1$. Ces solutions engendrent un réseau relatif Γ de rang $n-1$. □

Le problème est maintenant de trouver une base de l'ensemble du réseau relatif G des solutions de l'équation homogène.

Théorème IV-19 : *Soit $\{e_1, \ldots, e_n\}$ une base de $\mathbb{Z}^n$ adaptée à Γ, il existe $q_1|q_2|\ldots|q_{n-1}$ tels qu'une base de Γ est $\{q_1e_1, \ldots, q_ne_n\}$. Alors $\{e_1, \ldots, e_{n-1}\}$ est une $\mathbb{Z}$-base de G.*

Preuve : Puisque $q_je_j \in \Gamma$, pour $j = 1, \ldots, n-1$, on a $\sum_{i=1}^n a_i(q_je_j)_i = 0$ où $(\ )_i$ désigne la composante du vecteur sur ϵ_i d'où $q_j \sum_{i=1}^n a_i(e_j)_i = 0$ et les e_j sont des solutions du système homogène.

Réciproquement, si $y = (y_1, \ldots, y_n)$ est une solution du système homogène, on a $\sum_{i=1}^n a_iy_i = 0$. Comme $\{e_1, \ldots, e_n\}$ est une base de $\mathbb{Z}^n$, on peut écrire $y = \sum_{i=1}^n \lambda_ie_i$, $\lambda_i \in \mathbb{Z}$. Comme y et $e_1, \ldots, e_{n-1}$ sont des solutions, par différence λ_ne_n en est une aussi. Mais par construction $e_n \notin \mathbb{R}G$ donc $\lambda_n = 0$ et $y = \sum_{i=1}^n \lambda_ie_i$. □

Troisième problème : on va maintenant établir le théorème des quatre carrés.

Théorème IV-20 : *Tout entier positif est somme de quatre carrés d'entiers.*

La démonstration fait appel, outre la théorie des réseaux, aux corps finis et aux quaternions.

Lemme IV-21 : *Si tout nombre premier p est somme de quatre carrés d'entiers, il en est de même pour tout entier positif.*

Preuve : Introduisons le sous-anneau $A = \mathbb{Z} + \mathbb{Z}i + \mathbb{Z}_j + \mathbb{Z}k$ du corps des quaternions $\mathbb{H} = \mathbb{R} + \mathbb{R}i + \mathbb{R}j + \mathbb{R}k$. Dire que $p = a_p^2 + b_p^2 + c_p^2 + d_p^2$ se traduit par p est la norme réduite de $\pi_p = a_p + b_pi + c_pj + d_pk$ $(p = N_r(\pi_p))$. On utilise la factorisation de l'entier n en produit de nombres premiers, $n = \prod_p p^{v_p(n)}$ et la multiplicativité de la norme réduite : $n = \prod_p N_r(\pi_p)^{v_p(n)} = N_r(\prod_p \pi_p^{v_p(n)})$. On écrit alors $\prod_p \pi_p^{v_p(n)} = a + bi + cj + dk$. □

Comme $2 = 1 + 1 + 0 + 0$, on se restreint aux nombres premiers impairs.

Lemme IV-22 : *Dans le corps fini à p éléments $\mathbb{F}_p$ (p premier, impair), tout élément est somme de deux carrés.*

Preuve : Soit $a \in \mathbb{F}_p$, on construit les ensembles $U = \{u^2 - a \ | u \in \mathbb{F}_p\}$, $V = \{-v^2 \ | v \in \mathbb{F}_p\}$. Le groupe $\mathbb{F}_p^*$ est cyclique d'ordre $p-1$, comme p est impair le groupe $\mathbb{F}_p^{*2}$ est d'ordre $\frac{p-1}{2}$. L'ensemble des carrés de $\mathbb{F}_p$ est d'ordre $\frac{p-1}{2} + 1 = \frac{p+1}{2}$. Les ensembles U et V sont de cardinal $\frac{p+1}{2}$, ils ne peuvent donc être disjoints. Il existe $w, u, v \in \mathbb{F}_p$ tels que $w = u^2 - a = -v^2$, soit $a = u^2 + v^2$. □

On peut maintenant terminer la démonstration du théorème.

Preuve : Soit p premier impair, il existe $\overline{u}$ et $\overline{v} \in \mathbb{F}_p$ tels que $\overline{u}^2 + \overline{v}^2 + 1 = 0$. Choisissons, dans $\mathbb{Z}$, des représentants u et v de $\overline{u}$, $\overline{v}$ et construisons le morphisme de $\mathbb{Z}$-modules de $\mathbb{Z}^4$ dans $\mathbb{F}_p^2$ défini par :

$$\begin{array}{ccc} \mathbb{Z} \times \mathbb{Z} \times \mathbb{Z} \times \mathbb{Z} & \longrightarrow & \mathbb{F}_p \times \mathbb{F}_p \\ (a, b, c, d) & \mapsto & (ua + vb - c, -va + ub - d) \end{array}$$

Puis que $u^2 + v^2 = -1$, le morphisme est surjectif. Son noyau G est un sous-module d'indice p^2 donc un sous-réseau de $\mathbb{Z}^4$ (que l'on a implicitement considéré comme inclus dans $\mathbb{R}^4$).

Soit, dans $\mathbb{R}^4$, la sphère S de rayon $r = \sqrt{2p}$, c'est une partie convexe, symétrique de $\mathbb{R}^4$ de mesure $\mu(S) = \frac{1}{2}\pi^2\sqrt{2p}^4 = 2\pi^2p^2 > 16p^2$. Le théorème de Minkowski montre qu'il existe un point g non nul de

$G \cap S$, donc (a, b, c, d) avec $a^2 + b^2 + c^2 + d^2 < 2p$. Utilisons maintenant la définition de G, on $c \equiv ua + vb \mod (p)$, $d \equiv -va + ub \mod (p)$. D'où :

$$\begin{aligned} a^2 + b^2 + c^2 + d^2 &\equiv a^2 + b^2 + (ua + vb)^2 + (ub - va)^2 \\ &\equiv a^2 + b^2 + u^2a^2 + v^2b^2 + u^2b^2 + v^2a^2 \\ &\equiv a^2(1 + u^2 + v^2) + b^2(1 + v^2 + u^2) \equiv 0 \mod p \end{aligned}$$

On a $0 < a^2 + b^2 + c^2 + d^2 < 2p$ et $a^2 + b^2 + c^2 + d^2 \equiv 0 \mod (p)$, d'où $p = a^2 + b^2 + c^2 + d^2$. □

CHAPITRE V : RÉPARTITION DES NOMBRES PREMIERS

On connaît la preuve d'Euclide de l'existence d'une infinité de nombres premiers : Soit $p_1, \ldots, p_n$ n nombres premiers, on construit $1 + \prod_{i=1}^{n} p_i$ si ce nombre n'est pas premier il admet un diviseur premier qui n'est pas dans la liste et s'il est premier, là aussi, on a un nombre premier qui n'est pas dans la liste. La question de la répartition des nombres premiers se pose depuis longtemps, faute de formule permettant d'en construire la liste.

Une première question est celle de l'estimation de la fonction $\pi(x)$ qui donne le nombre de nombres premiers $\leq x$: la réponse , devinée par Gauss, est $\pi(x) \sim \frac{x}{log(x)}$ elle fut établie par Hadamard et De La Vallée Poussin (1896). Un premiere approche est due à Tcebychev, nous donnerons une version édulcorée :

$$\log(2)\frac{x}{\log(x)} + O(x^{1/2}\log(x)) \leq \pi(x) \leq 4\log(2)\frac{x}{\log(x)} + O(x^{1/2}\log(x)).$$

Une autre question naturelle : existe-t-il un nombre infini de nombre premier dans une famille infinie, sans contrainte évidente (les multiple d'un entier !). Dirichlet a montré que dans une progression arithmétique $\{na+b | n \in /Zdb$ (pgcd$(a,b) = 1\}$ de raison $a \neq 0$, il existe une infinité de nombres premiers. Nous donnerons quelques exemples et une preuve d'une forme faible du théorème de Dirichlet ($b = 1$). Par contre, on ne sait pas s'il y a une infinité de nombres premier de la forme $n^2 + 1$.

§1 UNE SÉRIE DIVERGENTE (Réference [IR])

On pourra aussi consulter [Li]. Les carrés des entiers sont éloignés les uns des autres : $\sum_{\substack{n\in\mathbb{N}\\ n>0}} \frac{1}{n^2}$ converge alors que $\sum_{\substack{n\in\mathbb{N}\\ n>0}} \frac{1}{n}$ diverge. Qu'en est-il des nombres premiers ?

Théorème V-1 : *Soit p_k (k entier ≥ 1) la suite croissante des nombres premiers, la série $\sum_k \frac{1}{p_k}$ diverge.*
Preuve : Soit $p_1, \ldots, p_{\pi(n)}$ la suite croissante des nombres premiers $\leq n$. On construit :

$$\lambda(n) = \prod_{i=1}^{\pi(n)} \frac{1}{(1 - \frac{1}{p_i})}$$

Comme $0 < \frac{1}{p_i} \leq \frac{1}{2} < 1$, on a $\frac{1}{(1-\frac{1}{p_i})} = \sum_{r=0}^{\infty} \frac{1}{p_i^r}$. En effectuant le produit de $i = 1$ à $\pi(n)$, on obtient la somme des inverses des entiers n'ayant que ces p_i dans leur décomposition en facteurs premiers. Parmi ceux là, il y a tous les entiers compris entre 1 et n. On a donc $\lambda(n) \geq \sum_{j=1}^{n} \frac{1}{j}$ et par conséquent $\lambda(n)$ tend vers l'infini avec n.

Intéressons-nous maintenant à $\log(\lambda(n)) = -\sum_{i=1}^{\pi(n)} \log(1 - p_i^{-1})$. On a :

$$\begin{aligned}\log(\lambda(n)) &= -\sum_{i=1}^{\pi(n)} \log(1 - p_i^{-1}) = \sum_{i=1}^{\pi(n)} \sum_{m=1}^{\infty} (mp_i^m)^{-1} \\ &= \frac{1}{p_1} + \ldots + \frac{1}{p_{\pi(n)}} + \sum_{i=1}^{\pi(n)} \sum_{m=2}^{\infty} (mp_i^m)^{-1}\end{aligned}.$$

Mais $\sum_{m=2}^{\infty}(mp_i^m)^{-1} < \sum_{m=2}^{\infty}(p_i^m)^{-1} = p_i^{-2}(1 - \frac{1}{p_i})^{-1} \leq 2p_i^{-2}$. En remplaçant on a :

$$\log(\lambda(n)) \leq \sum_{i=1}^{\pi(n)} \frac{1}{p_i} + 2\sum_{i=1}^{\pi(n)} \frac{1}{p_i^2}.$$

Lorsque l'on fait tendre n vers l'infini $\log(\lambda(n))$ tend vers $+\infty$, la seconde somme étant inférieure à $\frac{\pi^2}{6}$ c'est que $\lim_{n\to\infty} \sum_{i=1}^{\pi(n)} \frac{1}{p_i}$ tend vers l'infini. □

Remarques :
1 Une façon nettement plus compliquée que celle d'Euclide de montrer qu'il y a une infinité de nombres premiers !

2 Si, dans la définition de $\lambda(n)$, on remplace p_i par p_i^s avec $\Re(s) > 1$, on arrive facilement à la formule du produit pour la fonction ζ de Riemann.
3 Et pourtant, il y a des intervalles arbitrairement grand sans nombre premier, par exemple tous les entiers de $[n! + 2, \ldots, n! + n]$ de longueur $n - 1$ ont un diviseur compris entre 2 et n.

§**2 LE THÉORÈME DE TCHEBYTCHEV.** (Réferences [HW], [TMF])

On peut aussi consulter [EMF], [T]. Rappelons que l'on note, pour $x \geq 2$, $\pi(x)$ le nombre des nombres premier p, $2 \leq p \leq x$. On introduit une fonction auxiliaire θ définie par $\theta(x) = \sum_{\substack{p \leq x \\ p\text{premier}}} \log p$ et on compare ces deux fonctions : $\theta(x) = \sum_{\substack{p \leq x \\ p\text{premier}}} \log p \leq \theta(x) \sum_{\substack{p \leq x \\ p\text{premier}}} 1 = \pi(x)\log(x)$. Par ailleurs,

$$\begin{aligned}\theta(x) &\geq \sum_{\substack{x^{\frac{1}{2}} \leq p \leq x \\ p\text{premier}}} \log(p) \geq \frac{1}{2}\log(x) \sum_{\substack{x^{\frac{1}{2}} \leq p \leq x \\ p\text{premier}}} 1 \\ &= \frac{1}{2}\log(x)(\pi(x) - \pi(x^{\frac{1}{2}})) \geq \frac{1}{2}\log(x)(\pi(x) - x^{\frac{1}{2}})\end{aligned}.$$

ce qui se transforme, en tenant compte de la première inégalité, en :

$$\frac{\theta(x)}{\log(x)} \leq \pi(x) \leq x^{\frac{1}{2}} + \frac{2\theta(x)}{\log(x)}.$$

Pour obtenir le comportement asymptotique de θ, on va la comparer avec la fonction ψ définie par $\psi(x) = \sum_{\substack{(p,m), p^m \leq x \\ p\text{premier}}} \log(p)$. L'équivalence de $p^m \leq x$ avec $p \leq x^{\frac{1}{m}}$ donne $\psi(x) = \sum_{m=1}^{\infty} \theta(x^{\frac{1}{m}})$ dont les termes sont nuls dès que $m > \frac{\log(x)}{\log(2)}$. Par la définition de θ, on a évidemment $\theta(x) < x\log(x)$ et donc à fortiori $\theta(x^{\frac{1}{m}}) < x^{\frac{1}{m}}\log(x) \leq x^{\frac{1}{2}}/log(x)$. Compte tenu de la majoration des m intervenant dans la somme $\sum_{m \geq 2} \theta(x^{\frac{1}{m}}) = O(x^{\frac{1}{2}}\log(x)^2)$, ce qui donne la comparaison des comportements asymptotiques de ψ et θ :

Proposition V-2 : *On a l'égalité :* $\theta(x) = \psi(x) + O(x^{\frac{1}{2}}\log(x)^2)$.

D'où le corollaire :

Corollaire V-3 : *On a l'encadrement :*

$$\frac{\psi(x)}{\log(x)} + O(x^{\frac{1}{2}}\log(x)) \leq \pi(x) \leq x^{\frac{1}{2}} + 2\frac{\psi(x)}{\log(x)} + O(x^{\frac{1}{2}}\log(x)) = 2\frac{\psi(x)}{\log(x)} + O(x^{\frac{1}{2}}\log(x)).$$

Il nous faut maintenant évaluer la fonction ψ. Elle va s'introduire, assez naturellement, par l'étude de $\log(n!)$. On a une première évaluation de cette quantité :

Lemme V-4 : *On a l'égalité :*

$$\log(n!) = \sum_{1 \leq m \leq n} \log(m) = n\log n - n + O(\log(n)).$$

Preuve : Cela provient d'une forme faible de la formule de Stirling :

$$\begin{aligned}0 &\leq \int_m^{m+1} (\log t)dt - \log(m) \\ &= \int_m^{m+1} \log\left(\frac{t}{m}\right) dt \leq \int_m^{m+1} \left(\frac{t}{m} - 1\right) dt = \frac{1}{2m}.\end{aligned}$$

Il reste à sommer pour $1 \leq m \leq n - 1$ et à utiliser $\int \log(t)dt = t\log(t) - t$. □

Une autre façon de calculer $\log(m)$ est de le factoriser en produit de puissances de nombres premiers.

$$\log(m) = \sum_{\substack{(p,\nu)|p^\nu|n \\ p\text{premier}}} \log(p).$$

On reporte dans $\log(n!)$et on intervertit les sommations :

$$\log(n!) = \sum_{1\leq m\leq n} \log(m) = \sum_{p^\nu \leq m} \log(p) \sum_{\substack{1\leq m\leq n \\ p^\nu | m}} 1$$

$$= \sum_{p^\nu \leq n} \log(p) \left[\frac{n}{p^\nu}\right].$$

On introduit la fonction :

$$\Lambda(d) = \begin{cases} \log(p) & \text{si } \exists \nu \geq 1 \ : d = p^\nu \\ 0 & \text{sinon} \end{cases}$$

On en déduit pour $n \geq 2$:

$$\sum_{d\leq n} \Lambda(d) \left[\frac{n}{d}\right] = n\log(n) - n + O(\log(n)).$$

On note $B(n)$ le membre de gauche et on pose $B(x) := B([x])$. On va utiliser l'estimation de $B(x)$ pour établir un encadrement de $\sum_{d\leq x} \Lambda(d)$ qui est justement $\psi(x)$.

Posons pour $u > 0$: $\chi(u) = [u] - 2[u/2]$. C'est une fonction périodique caractérsée par :

$$\chi(u) = \begin{cases} 0 & \text{si } 0 \leq u < 1 \\ 1 & \text{si } 1 \leq u < 2. \end{cases}$$

Nous allons maintenant calculer de deux façons la quantité $B_2(x) = B(x) - 2B(x/2)$. La première est de revenir à l'évaluation de $B(x)$, les termes en $x\log(x)$ disparaissent et on a :

$$B_2(x) = x\log 2 + O(\log(x).$$

La seconde consiste à prendre la définition de B et de χ :

$$B_2(x) = \sum_{d\leq x} \Lambda(d)\chi(x/d).$$

On obtient une majoration brutale et une minoration utilisant la propriété de χ :

$$\psi(x) - \psi(x/2) \leq B_2(x) \leq \psi(x).$$

La majoration et l'évaluation précédente de B_2 nous donne :

$$x\log 2 + O(\log(x)) \leq \psi(x).$$

On utilise la minoration sous la forme $\psi(x) \leq B_2(x) + \psi(x/2)$ et on itère ce qui fournit :
$\psi(x) \leq \sum_{0\leq j\leq k} B_2(x/2^j) + \psi(x/2^{k+1})$. Si on choisit $k = \left[\frac{\log x}{\log 2}\right]$ la valeur de ψ s'annule et on obtient :

$$\psi(x) \leq \sum_{0\leq j\leq \left[\frac{\log x}{\log 2}\right]} \left(\frac{x\log 2}{2^j} + O(\log x)\right)$$

$$\leq 2x\log 2 + O(\log(x)^2.$$

En résumé :

Lemme V-5 : *Pour $x \geq 2$, on a l'encadrement :*

$$x\log 2 + O(\log x) \leq \psi(x) \leq x\log 4 + O(\log(x).$$

Finalement, on reportant dans le corollaire 3 :

Théorème V-6 : *On a l'encadrement :*

$$\log 2\frac{x}{\log x} + O(x^{\frac{1}{2}}\log x) \leq \pi(x) \leq 2\log 4\frac{x}{\log x} + O(x^{\frac{1}{2}}\log x).$$

La démonstration de Tchebychev utilise une fonction χ moins grossière et donne un encadrement plus fin (voir [TMF]).

§3 DU CÔTÉ DES PROGRESSIONS ARITHMÈTIQUES. (Réferences [Li])

Les premiers exemples que l'on donne sont des retombées rapide du procédé d'Euclide évoqué au début de ce chapitre.

Proposition V-7 : *Il y a une infinité de nombres premiers congrus à* 3 *modulo* 4.

Preuve : Soit E l'ensemble des premiers de la forme $4n+3$. On prend $q_1, \ldots, q_m \in E$, on construit $p = 4\prod_{i=1}^{m} q_i - 1$, ce nombre p est congru à 3 modulo 4. Il admet obligatoirement un facteur premier congru à 3 modulo 4, on a un élément de E hors de la sous-famille. L'ensemble E ne peut donc pas être fini. □

Proposition V-8 : *Il y a une infinité de nombres premiers congru à* 5 *modulo* 8.

. . . et donc à 1 modulo 4 !

Lemme V-9 : *Si p est un nombre premier impair et si p divise a^2+b^2 (a et b premiers entre eux), alors $p \equiv 1$ modulo* 4.

Preuve : La condition implique que a et b sont premiers à p, donc inversibles dans $\mathbb{F}_p$ (le corps à p éléments). dans corps on a donc $a^2+b^2=0$ soit $(ab^{-1})^2 = -1$. Il y a un élément d'ordre 4 dans $\mathbb{F}_p^*$ et donc $4|p-1$. □

Preuve (de la proposition) : Soit E l'ensemble des nombres premiers congrus à 5 modulo 8 et $q_1, \ldots, q_r$ une sous-famille finie. On construit $x = \prod_{i=1}^{r} q_i$ et $n = x^2+4$. Comme x est impair x^2 est congru à 1 modulo 8 et $n \equiv 5$ modulo 8. Les facteurs premiers de n n'appartiennent pas à la sous-famille. De plus, le lemme montre qu'ils ne peuvent être congrus à 3 ou 7 modulo 8. Ils ne peuvent être tous congrus à 1 modulo 8, donc au moins l'un d'entre eux est congru à 5 modulo 8. □

Proposition V-10 : *Il y a une infinité de nombres premiers congrus à* 5 *modulo* 6 *(et donc à* 2 *modulo* 3*).*

Preuve : Soit E l'ensemble des nombres premiers congrus à 5 modulo 6 et $q_1, \ldots, q_r$ une sous-famille finie. On construit $x = 6\prod_{i=1}^{r} q_i$ et $n = x-1$ qui est congru à 5 modulo 6. Les facteurs premiers de n n'appartiennent pas à la sous-famille. Un nombre impair ne peut être congru qu'à 1 ou 5 modulo 6. Si tous les facteur premiers de n était de la forme $6m+1$, il en serait de même pour n, c'est contradictoire. Le nombre n admet un facteur premier congru à 5 modulo 6, en dehors de la sous-famille choisie. L'ensemble E est donc infini. □

Proposition V-11 : *Il y a une infinité de nombres premiers congrus à* 1 *modulo* 3.

Preuve : Soit E l'ensemble des nombres premiers congrus à 1 modulo 3 et $q_1, \ldots, q_r$ une sous-famille finie. On construit $x = 3\prod_{i=1}^{r} q_i$ et $n = x^2+x+1$. Soit p un facteur premier de n, il n'appartient pas à la sous-famille finie, il est différent de 3 qui divise x et donc x^2+x. Montrons que $p \equiv 1$ modulo 3. En effet $p|n$ implique $p|(x-1)n = x^3-1$: donc la classe de x est d'ordre 3 dans $\mathbb{F}_p^*$. Le théorème de Lagrange nous dit que $3|p-1$. □

Dans ces démonstration on a vu apparaître les polynômes $X-1$, X^2+1, X^2+X+1 qui sont des polynômes cyclotomiques (voir annexe, en fin de chapitre). Nous allons nous baser sur leurs propriétés pour démontrer que si a est un entier positif, il existe une infinité de nombres premiers de la forme $an+1$. Le théorème de Dirichlet remplace 1 par un entier b premier à a dans cet énoncé. La preuve, plus compliquée est exposée dans la première partie de [Li] (sections 3,4,5).

Théorème V-12 : *Soit $a > 1$ un entier, il existe une infinité de nombres premiers de la forme $an+1$.*

Preuve : On considère l'ensemble E des nombres premiers congus à 1 modulo a et $q_1, \ldots, q_r$ une sous-famille finie de E. On construit $x = a\prod_{i=1}^{r} q_i$ et on évalue $\Phi_a(x)$. On établit d'abord le lemme :

Lemme V-13 : *Si p est premier avec a et si $p|\Phi_a(x)$ alors $p \equiv 1$ modulo a.*

Preuve : On part de $\Phi_a(x)\prod_{\substack{d|a\\ d<a}} \Phi_d(x) = x^a - 1$ qui est divisible par p. L'ordre de x dans $\mathbb{F}_p^*$ est un diviseur de a, si c'est un diviseur strict δ de a alors dans $\mathbb{F}_p$, $x^\delta - 1 = 0$ et donc x annule un Φ_d pour $d < a$. La classe de x est une racine double de X^a-1, elle annule ax^{a-1}, donc $ax^a = a$, ce qui est impossible (a et p sont premiers entre eux). L'ordre de x est égal à a et c'est un diviseur de $p-1$. □

Démonstration du théorème : Soit ζ une racine primitive a-ème de l'unité

$$|\Phi_a(x)| = \left|\prod_{(k,a)=1} (x-\zeta^k)\right| > (x-1)^{\varphi(a)} > 1.$$

L'entier $\Phi_a(x)$ a donc des diviseurs premiers. Soit p l'un d'entre eux. On vérifie facilement, par construction de Φ_a, que pour $a \geq 2$, $\Phi_a(0) = 1$ et on sait (cf annexe) que Φ_a est à coefficients entiers. On écrit donc

$$\Phi_a(x) = 1 + \sum_{i=1}^{\varphi(a)} a_i x^i$$

Par choix de x, $\Phi_a(x)$ est congru à 1 modulo a, p est donc premier à a, le lemme nous dit alors que p appartient à E ; l'expression précédente de $\Phi_a(x)$ montre que p n'appartient à la sous-famille choisie. L'ensemble E est donc infini. □

ANNEXE : LES POLYNÔMES CYCLOTOMIQUES. (Référence [P]).

Soit K un corps, $n \in \mathbb{N} \setminus \{0\}$, on considère le polynôme $P_n = X^n - 1$. Le polynôme dérivé de $X^n - 1$ est nX^{n-1}, si la caractéristique de K ne divise pas n, la seule racine de P'_n est 0 qui n'est pas racine de P_n, le polynôme P_n ne possède pas de racine multiple. Si $p|n$, $n = pm$ on a $X^n - 1 = (X^m - 1)^p$ donc P_n n'a que des racines multiples dans un corps de décomposition de P_n. Dans la suite de ce paragraphe, on suppose la caractéristique de K première à n.

Lemme V-14 : *L'ensemble $\mu_n(K)$ des racines n-ièmes de l'unité appartenant à K^* est un sous-groupe cyclique de K^* d'ordre $d|n$.*

Preuve : Cela résulte de la structure des groupes abéliens finis et de ce qu'un polynômes de degré n possède au plus n racines. Ensuite, soit K_n un corps de décomposition de P_n, $\mu_n(K_n)$ est un sous-groupe cyclique d'ordre n de K_n^*, enfin $\mu_n(K) \subset \mu_n(K_n)$. □

Définition V-15 : *On appelle racine primitive n-ième de l'unité tout générateur de $\mu_n(K_n)$, on note $\mu_n^*(K_n)$ l'ensemble des racines primitives n-ièmes de 1 appartenant à K_n.*

Il y a donc $\varphi(n)$ (indicateur d'Euler) racines primitives n-ième de l'unité dans $\mu_n^*(K_n)$.

Définition V-16 : *Le polynôme $\Phi_{n,K}(X) = \prod_{\zeta \in \mu_n^*(K_n)} (X - \zeta)$ s'appelle le n-ième polynôme cyclotomique sur K.*

Le polynôme Φ_n est unitaire de degré $\varphi(n)$. Si k est le sous-corps premier de K, $P_n \in k[X]$, k_n le corps de décomposition sur k de P_n est inclus dans K_n. On en déduit que $\Phi_n \in k_n[X]$, on va montrer qu'en fait $\Phi_n \in k[X]$.

Proposition V-17 : *On a l'identité $X^n - 1 = \prod_{d|n} \Phi_d(X)$.*

Preuve : Cela résulte de l'égalité des ensemble $\mu_n(K_n) = \cup_{d|n} \mu_d^*(K_n)$ qui exprime que tout élément de μ_n et d'ordre d, diviseur de n.

Remarques :
1 . En comparant les degrés (i.e. en réinterprétant la formule de réunion de la démonstration) on retrouve la formule $n = \sum_{d|n} \varphi(d)$.
2 . La proposition donne une méthode de calcul des polynômes cyclotomiques, en procédant par récurrence :

$$\Phi_n(X) = \frac{X^n - 1}{\prod_{d|n, d \neq n} \Phi_d(X)}.$$

Exercice **1** : Calculez $\Phi_{24,\mathbb{Q}}$.

Proposition V-18 :
1 . *Le polynôme $\Phi_{n,\mathbb{Q}} \in \mathbb{Z}[X]$.*
2 . *Soit k un corps et $\sigma : \mathbb{Z} \to k$ l'homomorphisme canonique de $\mathbb{Z}$ dans k alors $\Phi_{n,k} = \sigma(\Phi_{n,\mathbb{Q}})$.*

En particulier, pour $p \not| n$ $\Phi_{n,\mathbb{F}_p}$ se déduit de $\Phi_{n,\mathbb{Q}}$ par réduction modulo p.

Preuve :
1 – On raisonne par récurrence sur n, on a $\Phi_1(X) = X - 1 \in \mathbb{Z}[X]$, supposons la propriété vraie pour tout $d < n$. Le produit $\prod_{d|n, d<n} \Phi_d(X)$ est un polynôme unitaire à coefficients dans $\mathbb{Z}$ qui divise Φ_n dans $\mathbb{Q}[X]$. Si on effectue cette division euclidienne, on reste dans $\mathbb{Z}$.

2 – On raisonne encore par récurrence, le cas $n = 1$ est trivial. Le point 1, permet d'écrire $X^n - 1 = \Phi_{n,\mathbb{Q}} F(X)$ dans $\mathbb{Z}[X]$. On a

$$\begin{aligned}\sigma(X^n - 1) = X^n - 1 &= \sigma(\Phi_{n,\mathbb{Q}})\sigma(F(X)) \\ &= \Phi_{n,k} \prod_{d<n, d|n} \Phi_{d,k}(X)\end{aligned}$$

Mais $\sigma(F) = \prod_{d<n,d|n} \sigma(\Phi_{d,\mathbb{Q}}(X)) = \prod_{d<n,d|n} \Phi_{d,k}(X)$. On obtient le résultat par comparaison des égalités obtenues. □

Intéressons-nous à un cas particulier :

Théorème V-19 : *Le polynôme $\Phi_{n,\mathbb{Q}}(X)$ est irréductible sur $\mathbb{Z}$ (donc sur $\mathbb{Q}$).*

Preuve : Soit K le corps de décomposition de Φ_n sur $\mathbb{Q}$, $\zeta \in K$ une racine primitive n-ième de l'unité. On se donne un nombre premier p ne divisant pas n.
1) Le choix de p implique que ζ^p est une autre racine primitive de l'unité.
2) Soit F (resp. G) le polynôme minimal (unitaire) de ζ (resp. ζ^p) sur $\mathbb{Q}$, ce polynôme appartient à $\mathbb{Z}[X]$. Puisque $\mathbb{Z}[X]$ est factoriel on décompose Φ_n en produit de facteurs irréductibles $\Phi_n = \prod_i P_i$, comme Φ_n est unitaire, il n'y a pas d'élément de $\mathbb{Z}$ dans la décomposition et les P_i sont unitaires ; mais F et G figurent dans cette décomposition. Enfin, on a vu qu'un polynôme irréductible sur $\mathbb{Z}$ l'est sur $\mathbb{Q}$.
3) Montrons que $F = G$. S'il n'en est pas ainsi le produit FG divise Φ_n dans $\mathbb{Z}[X]$. Comme $G(\zeta^p) = 0$ on en déduit que $G(X^p)$ est divisible par F dans $\mathbb{Q}[X]$, donc dans $\mathbb{Z}[X]$ puisqu'il s'agit de polynômes unitaires à coefficients entiers : $G(X^p) = FH$. On effectue une réduction modulo p. Dans le corps $\mathbb{F}_p$ l'élévation à la puissance p est égale à l'identité : $\overline{G}(X^p) = \overline{G(X)}^p$ la réduction modulo p de l'identité $G(X^p) = FH$ nous donne donc $\overline{G(X)}^p = \overline{FH}$. Soit P un facteur irréductible de $\overline{F}$, puisque $\mathbb{F}_p[X]$ est principal, P divise $\overline{G}$, on en déduit que dans $\mathbb{F}_p[X]$, $X^n - 1 = P^2 Q$ possède au moins une racine multiple, ceci est contradictoire avec le choix de p relativement à n.
4) Si ζ' est une autre racine primitive n-ième de l'unité, $\zeta' = \zeta^m$ avec $m = p_1^{a_1} \dots p_r^{a_r}$ et les p_i sont premiers à n, un rapide raisonnement par récurrence montre que ζ' et ζ ont le même polynôme irréductible F. La décomposition de Φ_n en facteurs irréductibles est donc $\Phi_n = F^t$, comme Φ_n n'a pas de racine multiple, $t = 1$, le polynôme est irréductible sur $\mathbb{Q}$, donc sur $\mathbb{Z}$ puisqu'il est unitaire. □

Exercice **2** : Montrez qu'il existe au plus une fonction μ définie sur l'ensemble des nombres entiers ≥ 1, à valeurs entières et vérifiant $\mu(1) = 1$ et si $n > 1$, $\sum_{d|n} \mu(d) = 0$.
Montrez qu'une telle fonction vérifie $\mu(1) = 1$, $\mu(p) = -1$ si p est premier, $\mu(p^r) = 0$ si p est premier et $r \geq 2$.
Montrez que si μ existe et si m et n sont premiers entre eux, on a $\mu(mn) = \mu(m)\mu(n)$ (décomposer tout diviseur d de mn en le produit d'un diviseur de m et d'un diviseur de n, montrez que la propriété est vraie si mn n'a que 4 diviseurs, puis procédez par récurrence).
Montrez que si μ existe $\mu(1) = 1$, $\mu(n) = (-1)^r$ si n est le produit de r nombres premiers distincts, $\mu(n) = 0$ si n est divisible par le carré d'un entier.
Montrez que μ existe (c'est la fonction de Möbius).
Soit f une fonction définie sur les entiers ≥ 1 à valeurs dans un groupe abélien, noté additivement, on définit une nouvelle fonction g par $g(n) = \sum_{d|n} f(d)$ où la somme est étendue à l'ensemble des diviseurs d de n, $1 \leq d \leq n$. Montrez que $f(n) = \sum_{d|n} g(d)\mu(\frac{n}{d})$.
Montrez que le n-ième polynôme cyclotomique vérifie $\Phi_n(X) = \prod_{d|n} (X^d - 1)^{\mu(\frac{n}{d})}$. Calculez Φ_{60}.