Université Pierre & Marie Curie 2e semestre 2013-2014

Master de Mathématiques (M1)

THÉORIE DES NOMBRES

Daniel BERTRAND

Janvier 2014

.

Plan

Sommaire

Ce cours porte sur la théorie algébrique des nombres, certaines de ses applications en algorithmique et en cryptographie, et une introduction à la théorie analytique des nombres.

Bibliographie

H. Cohen: *A course in computational number theory*; Springer, 1993.
M. Demazure: *Cours d'Algèbre*; Cassini, 1997.
G. Hardy and E. Wright: *An introduction to the theory of numbers*; Oxford UP, 1979
P. Samuel: *Théorie algébrique des nombres*; Hermann, 1967.
J-P. Serre: *Cours d'arithmétique*; PUF, 1970.

CHAPITRE I

EXTENSIONS ALGÉBRIQUES

§1. Rappels sur les groupes. Fonction indicatrice d'Euler.

Un *groupe* est la donnée d'un ensemble G muni d'une loi de composition interne : $G \times G \to G : (x, y) \mapsto xy$ associative, admettant un élément neutre, noté e, et tel que tout élément x de G admette un inverse, noté x^{-1}. G est dit abélien (ou commutatif) si sa loi est commutative (on la note alors additivement). Le cardinal $|G|$ de G s'appelle l'*ordre* de G .

Exemples : le groupe symétrique S_n formé des permutations d'un ensemble à n éléments; $|S_n| = n!$; le sous-groupe $A_n \lhd S_n$ des permutations s paires (i.e. de signature $\varepsilon_s = 1$);

$GL_n(\mathbf{R})$ = groupe multiplicatif des matrices réelles $n \times n$ inversibles;

le groupe additif $\mathbf{Z}$ des entiers rationnels; tout sous-groupe non nul de $\mathbf{Z}$ est de la forme $a\mathbf{Z}$, pour un unique entier rationnel $a > 0$ (effectuer des divisions euclidiennes par son plus petit élément > 0);

$\mathbf{Z}/a\mathbf{Z}$ = groupe additif des classes de congruence d'entiers rationnels modulo un entier rationnel $a > 0$; il est d'ordre a;

$(\mathbf{Z}/a\mathbf{Z})^*$ = groupe multiplicatif des classes de congruence modulo a d'entiers b premiers à a (par Bézout, $(a, b) = 1$ ssi $\exists m, b' \in \mathbf{Z}, am + bb' = 1$, et $\overline{b'}$ est l'inverse de $\overline{b}$). Son ordre est noté $\phi(a)$ ($\phi :=$ *fonction indicatrice d'Euler*, avec $\phi(1) = 1$). Pour tout nombre premier p et tout entier $n > 0$, $\phi(p) = p - 1$, $\phi(p^n) = p^{n-1}(p - 1)$, et on déduit du lemme chinois (cf. §2) que pour tout couple d'entiers m, n:

$$(m, n) = 1 \ \Rightarrow \ \phi(m, n) = \phi(m)\phi(n)$$

(ϕ est une fonction arithmétique 'multiplicative'). Finalement, pour tout $a > 0$:

$$\phi(a) = a \prod_{p \text{ premier}, p|a} (1 - \frac{1}{p}).$$

Un *sous-groupe* H de G (notation : $H < G$) est une partie H contenant e , stable sous la loi de groupe de G et sous l'inversion. Comme une intersection de sous-groupes

est un sous-groupe, on peut parler du plus petit sous-groupe $< S >$ de G contenant une partie S de G , qu'on appelle aussi sous-groupe engendré par S .

On appelle *classe (à gauche) modulo* H toute partie de G de la forme $xH := \{xh; h \in H\}$ pour un élément x de G , appelé représentant dans G de la classe en question (chacun de ses éléments en est donc un représentant). Deux classes distinctes sont disjointes, et le cardinal de l'ensemble G/H des classes modulo H, appelé *indice* de H dans G et noté $[G : H]$, vérifie la relation :

$$|G| = |H| \times [G : H].$$

(En particulier, si G est fini, $|H|$ divise $|G|$.)

On dit que H est distingué dans G (ou normal; notation: $H \triangleleft G$), si $\forall x \in G, xH = Hx$, i.e. $\forall h \in H, xhx^{-1} \in H$. L'ensemble G/H est alors naturellement muni d'une structure de groupe, appelée *quotient* de G par H , telle que la surjection canonique $\pi : G \to G/H : x \mapsto (x) = xH$ est un homomorphisme (voir infra) de groupes. De plus, π établit une bijection entre l'ensemble des sous groupes A de G contenant H et l'ensemble des sous-groupes A/H de G/H; pour $A \triangleleft G$, A/H est normal dans G/H, et $(G/H)/(A/H) \simeq G/A$.

Un *homomorphisme* d'un groupe G vers un groupe G' est une application $f : G \to G'$ telle que $f(xy) = f(x)f(y)$ pour tout $(x, y) \in G \times G$. (Alors $f(e) = e'; f(x^{-1}) = f(x)^{-1}$.) On parle d'endomorphisme si $G = G'$, d'isomorphisme s'il existe un homomorphisme g de G' dans G tel que $f \circ g = id_{G'}, g \circ f = id_G$, d'automorphisme si f est à la fois un endo. et un iso..

Proposition: (décomposition canonique d'un homomorphisme) : *Soit f un homomorphisme de G vers G' . Alors :*

i) $Ker(f)$ = noyau de $f := \{x \in G : f(x) = e'\}$ est un sous groupe distingué de G . Notons π la surjection canonique de G sur $G/Ker(f)$;

ii) $Im(f)$ = image def $:= \{f(x), x \in G\}$ est un sous-groupe de G' . Notons i l'injection canonique de $Im(f)$ dans G' .

iii) il existe un unique isomorphisme $\overline{f}$ de $G/Ker(f)$ sur $Im(f)$ tel que $f = i \circ \overline{f} \circ \pi$.

Soit I un ensemble, et $\{G_i; i \in I\}$ une collection de groupes. On définit une structure de *groupe produit* sur le produit des ensembles G_i en posant $(x_i)_{i\in I}.(y_i)_{i\in I} = (x_i y_i)_{i\in I}$. Si tous les G_i sont égaux à un G , ce groupe, noté G^I (ou G^n si $card(I) = n$ est fini), s'identifie au groupe des applications de I dans G .

On dit qu'un sous-groupe de G est *cyclique* (ou monogène) s'il est engendré par un élément x de G . L'application $n \mapsto x^n$ définit alors un homomorphisme surjectif du groupe additif $\mathbf{Z}$ sur $< x >$, de noyau $a\mathbf{Z}$, avec $a \in \mathbf{Z}_{\geq 0}$, et $< x > \simeq \mathbf{Z}/a\mathbf{Z}$. Si a est non

nul, $< x >= \{x, x^2, ..., x^{a-1}, x^a = e\}$ est un groupe fini, et son ordre a s'appelle l'*ordre* de l'élément x de G. En particulier, si $|G|$ est fini, l'ordre de tout élément de G divise l'ordre de G. Appliqué au groupe $(\mathbf{Z}/a\mathbf{Z})^*$, ceci entraîne que pour tout entier x premier à a, $x^{\phi(a)} \equiv 1 \pmod{a}$, et en particulier (petit théorème de Fermat):

$$\forall\ p \text{ premier}, \forall\ x \text{ t.q. } p \not| x\ :\ x^{p-1} \equiv 1 \bmod p.$$

Les générateurs d'un groupe cyclique fini $< x >$ d'ordre a sont de la forme x^b, où b est premier à a; il y en a donc $\phi(a)$. Soit alors n un entier > 0. Pour tout diviseur $d = n/d'$ de n, il existe (exercice) un unique sous-groupe $C_d = d'\mathbf{Z}/n\mathbf{Z} \simeq \mathbf{Z}/d\mathbf{Z}$ d'ordre d du groupe cyclique $C_n = \mathbf{Z}/n\mathbf{Z}$. Par conséquent, les éléments d'ordre d de C_n sont les générateurs de C_d, et l'on a:

$$\forall n > 0,\ \Sigma_{d|n}\ \phi(d)\ =\ n.$$

La *fonction de Möbius* est l'application $\mu : \mathbf{Z}_{>0} \to \mathbf{Z}$ définie par $\mu(n) = 0$ si n est divisible par un carré (i.e. par le carré d'un entier > 1); $\mu(n) = (-1)^r$ si n est le produit de r nombres premiers distincts (donc $\mu(1) = 1$). C'est une fonction arithmétique multiplicative, qui vérifie $\Sigma_{d|n}\mu(d) = 0$ pour tout $n \geq 2$. On en déduit: $\phi(n) = \Sigma_{d'|n}(\Sigma_{d|\frac{n}{d'}}\mu(d))\phi(d') = \Sigma_{dd'|n}\mu(d)\phi(d')$, soit

$$\forall n > 0,\ \phi(n)\ =\ \Sigma_{d|n}\ \mu(d)\frac{n}{d}.$$

Application aux polynômes cyclotomiques

Lemme: *soit K un corps commutatif (cf. §2). Tout sous-groupe fini G du groupe multiplicatif K^* est cyclique.*

Démonstation: soit d un diviseur de l'ordre n de G tel qu'il existe un élément x d'ordre d de G. Puisque K est un corps comm., l'équation $X^d = 1$ a au plus d solutions dans K, et tous les éléments d'ordre d de G sont des générateurs de $< x >$. Ainsi, pour tout $a|n$, il y a 0 ou $\phi(a)$ éléments d'ordre a de G, et la formule $n = \Sigma_{a|n}\phi(a)$ impose qu'il y en ait $\phi(a)$ pour tout a. En particulier, il en existe d'ordre n, et G est cyclique.

Soient n un entier > 0, et K un corps algébriquement clos de caractéristique nulle ou première à n (cf. §2, ou prendre $K = \mathbf{C}$). Le polynôme $X^n - 1 \in K[X]$ est alors séparable (cf. fin du §2), et a n racines distinctes dans K, qui forment le sous-groupe $\mu_n(K)$ de K^* des racines n-ièmes de l'unité dans K. C'est un groupe d'ordre n, donc cyclique, dont les générateurs s'appellent les racines primitives n-ièmes de l'unité; ce sont les racines n-ièmes ζ vérifiant $\zeta^d \neq 1\ \forall d|n, d \neq n$. On appelle n-ième polynôme cyclotomique le polynôme, de degré $\phi(n)$:

$$\Phi_n(X)\ =\ \prod_{\zeta \in \mu_n(K),\ \zeta \text{ primitive}} (X - \zeta).$$

On a $\Pi_{d|n}\Phi_d(X) = X^n - 1$, et

$$\Pi_{d|n}(X^{\frac{n}{d}} - 1)^{\mu(d)} = \Pi_{d'|n}\Phi_{d'}^{\Sigma_{d|(n/d')}\mu(d)} \;=\; \Phi_n(X).$$

En particulier, $\Phi_n(X)$ est un polynôme unitaire à coefficients dans $\mathbf{Z}$ si $car(K) = 0$, et son image dans $\mathbf{F}_p[X]$ par réduction modulo p si $car(K) = p$ est un nombre premier.

Application à la cryptographie

La système RSA de chiffrage à clef publique consiste à assigner à chaque membre M_i d'un réseau $\{M_1, ..., M_n\}$ un couple p_i, q_i de grands nombres premiers secrets, de produit $n_i = p_i q_i$ public, et un entier public c_i premier à $\phi(n_i)$. Seul M_i peut alors calculer l'inverse d_i de c_i dans $\mathbf{Z}/\phi(n_i)\mathbf{Z}$, car la connaissance de $\phi(n_i) = (p_i - 1)(q_i - 1)$ exigerait de factoriser n_i (on peut alternativement supposer que M_i a reçu d_i en secret; en tous cas, une fois qu'il connaît d_i, M_i peut oublier p_i et q_i).

Un message est la donnée d'un entier N, inférieur à tous les n_i (supposés du même ordre de grandeur) et tel que $(N, n_i) = 1$; en particulier, $N^{\phi(n_i)} \equiv 1 \pmod{n_i}$ pour tout i. Les dictionnaires "Entiers - Français; Français - Entiers" sont publics.

Si M_1 veut envoyer secrètement N à M_2, il lui adresse de façon publique le reste N_2 de la division de N^{c_2} par n_2. Pour récupérer N, M_2 calcule alors $N_2^{d_2} \equiv N \pmod{n_2}$, ce qu'ignorant d_2, les autres membres $M_i, i > 2$, du réseau ne sauront faire. Et s'ils tentent $N_2^{d_i}$ avec leur propre clef, l'incohérence du résultat obtenu leur montrera que le message ne leur était pas destiné.

Pour s'assurer que le message envoyé par M_1 lui est dû, on peut aussi procéder comme suit (on suppose ici que $n_1 < n_2$, de sorte que deux restes de division par n_1 distincts ne peuvent représenter la même classe modulo n_2): M_1 calcule le reste N^1 de la division de N^{d_1} par n_1, et envoie $N_2^1 := (N^1)^{c_2} \pmod{n_2}$ à M_2. Alors, M_2 peut récupérer N en calculant $(N_2^1)^{d_2} \equiv N^1 \pmod{n_2}$, puis $(N^1)^{c_1} \equiv N \pmod{n_1}$. La cohérence du message obtenu confirme que M_1 en était bien l'expéditeur.

§2 Rappels sur les anneaux. Lemme chinois. PGCD.

Un anneau (unitaire) A est un groupe abélien (de loi notée + , d'élément neutre 0), muni d'une loi de composition interne $A \times A \to A : (x, y) \mapsto xy$ associative, distributive par rapport à la loi + , admettant un élément neutre noté 1 . Si x et y commutent, on a la formule de Newton: $(x+y)^n = \Sigma_{k=0,..n}\ C_n^k\ x^k y^{n-k}$. On dit que A est commutatif si tous ses éléments commutent entre eux, et que A est *intègre* s'il est non nul et s'il n'a pas de diviseur de 0 autre que 0 (i.e. $xy = 0 \Rightarrow x = 0$ ou $y = 0$).

On appelle *unité* de A tout élément x de A tel qu'il existe $y \in A$ vérifiant $xy = yx = 1$ (y est alors unique, et noté x^{-1}). L'ensemble des unités de A forme un groupe (pour la loi multiplicative), noté A^*. Un *corps* est un anneau non nul K tel que $K^* = K \setminus \{0\}$.

Exemples: $M_n(\mathbf{R})$ = anneau (non commutatif pour $n > 1$) des matrices carrées réelles d'ordre n ; alors, $(M_n(\mathbf{R}))^* = GL_n(\mathbf{R})$. Plus généralement, pour un anneau A commutatif, $M_n(A)$ = anneau des matrices carrées d'ordre n à coeff. dans A; alors $(M_n(A))^*$, noté $GL_n(A)$, est l'ensemble des matrices x de déterminant

$$det(x) = \sum_{s \in S_n}\ \varepsilon_s x_{1,s(1)} ... x_{n,s(n)} \in A^*.$$

L'application $det : GL_n(A) \to A^*$ est un homomorphisme de groupes. On note $SL_n(A) = \{x \in M_n(A), det(x) = 1\}$ son noyau. Par exemple, $\mathbf{Z}^* = \{1, -1\}$, et $SL_n(\mathbf{Z})$ est un sous-groupe d'indice 2 de $GL_n(\mathbf{Z})$.

le groupe $\mathbf{Z}/a\mathbf{Z}$ est naturellement muni d'une structure d'anneau, dont le groupe des unités est le groupe noté $(\mathbf{Z}/a\mathbf{Z})^*$ au §1. Pour $a = p$ premier, $\mathbf{Z}/p\mathbf{Z}$ est un corps, noté $\mathbf{F}_p$;

si A_1 et A_2 sont deux anneaux, on munit (en calculant coordonnée par coordonnée comme pour les groupes) le produit $A_1 \times A_2$ d'une structure d'anneau, non intègre si les A_i sont non nuls. On a: $(A_1 \times A_2)^* = (A_1)^* \times (A_2)^*$;

l'anneau $A[X_1, ..., X_n]$ des polynômes en n variables à coeff. dans un anneau commutatif intègre A admet A^* pour groupes des unités;

le corps des fractions $K = Fr(A)$ d'un anneau commutatif intègre A est défini par les règles de calcul usuelles sur les fractions $\frac{a}{b}, a \in A, b \in A \setminus \{0\}$. On note $K(X_1, ..., X_n) := Fr(A[X_1, ..., X_n])$ le corps des *fractions rationnelles* en n variables à coefficients dans K.

Un *idéal* (à gauche) J de l'anneau A est un sous-groupe du groupe $(A, +)$ tel que $\forall (a, x) \in A \times J, ax \in J$. Comme une intersection d'idéaux est un idéal, on peut parler du plus petit idéal (S) de A contenant une partie S de A, qu'on appelle aussi idéal engendré par S . Si $J = Ax := (x)$ est engendré par un élément x, on dit que J est un idéal principal. Un anneau A est dit *principal* s'il est intègre et si tous ses idéaux sont principaux.

Pour tout idéal bilatère J de A, le groupe quotient A/J est naturellement muni d'une structure d'anneau, faisant de la surjection canonique $\pi : A \to A/J$ un homomorphisme d'anneau (voir infra). De plus, π établit une bijection entre l'ensemble des idéaux de A contenant J et l'ensemble des idéaux de A/J.

Un *homomorphisme* d'un anneau A vers un anneau A' est une homo. de groupes additifs $f : A \to A'$ tel que $f(xy) = f(x)f(y)$ pour tout $(x, y) \in A \times A$, et $f(1) = 1'$. Endomorphisme, isomorphisme, automorphisme se définissent comme pour les groupes, et on a de même un théorème de décomposition canonique des homomorphismes d'anneaux (noter que $Ker(f)$ sera ici un idéal bilatère de A). Pour A, A' commutatifs, on étend f en un homomorphisme de $A[X]$ vers $A'[X]$ en associant à $P(X) = \Sigma_{i=0,..,n} a_i X^i \in A[X]$ le polynôme $f(P) = P^f := \Sigma_{i=0,...,n} f(a_i)X^i$ de $A'[X]$.

On suppose désormais nos anneaux *commutatifs*. Soient $J_1, ..., J_n$ des idéaux de A. On définit leur *produit* $J_1...J_n$ comme l'idéal *engendré par* les éléments de la forme $x_1...x_n$, où x_i parcourt J_i pour tout $i = 1, ..., n$. Leur intersection ensembliste $J_1 \cap ... \cap J_n$ est un idéal de A , qui contient, en général strictement, $J_1...J_n$. L'ensemble $\{\Sigma_{i=1,...,n} x_i;\ \forall i,\ x_i \in J_i\}$ est un idéal de A, noté $J_1 + ... + J_n$. Dans ces conditions:

Lemme chinois : *Supposons que pour tout couple* $(i, j), 1 \leq i < j \leq n$, *d'indices distincts,* $J_i + J_j = A$. *Alors,* $J_1...J_n = J_1 \cap ... \cap J_n$, *et l'application* ϕ *qui, à un élément* x *de* A, *associe l'élément* (x mod $J_1, ..., x$ mod J_n) *de l'anneau produit des* A/J_i *établit par passage au quotient un isomorphisme d'anneaux :*

$$A/J_1...J_n \simeq (A/J_1) \times ... \times (A/J_n).$$

Démonstration. Pour $r = 2$, soient $a_i \in J_i, i = 1, 2$ tels que $a_1 + a_2 = 1$, et $y \in J_1 \cap J_2$; alors, $y = ya_1 + ya_2 \in J_1J_2$, qui coïncide donc avec $J_1 \cap J_2$; pour tout couple x_1, x_2 d'élts de A, $\phi(a_1x_2 + a_2x_1) =$ (classe de x_1 mod J_1, classe de x_2 mod J_2), et ϕ est bien surjective. Pour $r > 2$, il existe par hypothèse $a_i \in J_i$ et $\alpha_i \in J_1$ tels que $\alpha_i + a_i = 1$ pour tout $i > 1$. Alors, $1 - a_2...a_r \in J_1$, donc $J_1 + J_2...J_r = A$, et on conclut par récurrence.

Applications: - $\mathbf{Z}/3\mathbf{Z} \times \mathbf{Z}/4\mathbf{Z}$ est isomorphe à $\mathbf{Z}/12\mathbf{Z}$ (et est en particulier un groupe cyclique). Mais $\mathbf{Z}/3\mathbf{Z} \times \mathbf{Z}/3\mathbf{Z}$ n'est pas cyclique (et n'est donc pas isomorphe à $\mathbf{Z}/9\mathbf{Z}$).

- soient $a_1, ..., a_n$ des entiers ≥ 1, et premiers entre eux deux à deux (i.e. tels que $pgcd(a_i, a_j) = 1$ pour tout $i \neq j$). Alors, $(\mathbf{Z}/a_1...a_n\mathbf{Z})^* \simeq (\mathbf{Z}/a_1\mathbf{Z})^* \times ... \times (\mathbf{Z}/a_n\mathbf{Z})^*$. En particulier, la fonction indicatrice d'Euler est multiplicative (cf. §1).

On dit qu'un idéal J d'un anneau A est *maximal* (resp. *premier*) si A/J est un corps (resp. un anneau intègre). Donc maximal $\Rightarrow$ premier (réciproque fausse en général: penser

à $\mathbf{C}[X,Y]$). J est maximal ssi $J \neq A$ et le seul idéal le contenant strictement est A tout entier; J est premier ssi $J \neq A$ et $xy \in J \Rightarrow x \in J$ou $y \in J$. On peut déduire du lemme de Zorn (énoncé équivalent à l'axiome du choix) que tout idéal de A distinct de A est contenu dans un idéal maximal.

Soient A un anneau *intègre*, et a, b deux éléments de A. On dit que a *divise* b dans A (notation : $a|b$) si (a) contient (b), qu'ils sont *associés* si $(a) = (b)$ (i.e. s'il existe une unité u de A telle que $a = ub$), que a est *irréductible* si $a \neq 0, a \notin A^*$ et les seuls éléments de A divisant a sont les unités de A et les éléments associés à a, et enfin que a est *premier* si (a) est un idéal premier non nul (i.e. $a \neq 0, a \notin A^*$ et $a|xy \Rightarrow a|x$ ou $a|y$). Tout élément premier est irréductible, mais l'inverse est faux en général (consider la relation $2.3 = (1+\sqrt{-5}).(1-\sqrt{-5})$ dans $A = \mathbf{Z}[\sqrt{-5}]$).

On dit qu'un anneau intègre A est *factoriel* si tout élément non nul et non unité admet une décomposition en produit d'éléments irréductibles, *unique* à permutation et multiplications par des unités près.

Proposition: *soit A un anneau intègre, et a un élément de A non nul et non unité.*

i) si A est factoriel, a irréductible $\Leftrightarrow$ a premier;

ii) A principal $\Rightarrow$ A factoriel;

iii) si A est principal, tout idéal premier non nul est maximal, donc a irréductible $\Leftrightarrow$ a premier $\Leftrightarrow$ (a) maximal.

iv) si A admet un algorithme de division euclidienne, il est principal.

Dans un anneau factoriel, on associe de la façon usuelle à toute famille $a_1, ..., a_n$ d'éléments (éventuellement nuls) de A leur *pgcd* et leur *ppcm*, bien définis à multiplication par une unité près. On a $pgcd(a,b) \times ppcm(a,b) = ab$ (= signifiant ici associé). On dit que a et b sont *premiers entre eux* (ou: étrangers) si $pgcd(a,b) = 1$. L'idéal principal engendré par $pgcd(a,b)$ contient en général strictement l'idéal (a,b). On a néanmoins bien égalité si l'anneau est principal, d'où , *dans un anneau principal*, la notation $(a,b) = 1$ pour dire que a et b sont premiers entre eux. Enfin, si A est euclidien, l'algorithme de division euclidienne fournit un procédé de calcul effectif du pgcd.

Comme ils sont munis d'un algorithme de division euclidienne (de 'stathme' la valeur absolue pour l'un, le degré pour l'autre), les anneaux $\mathbf{Z}$ et $\mathbf{Q}[X]$ sont principaux. En revanche, $\mathbf{Z}[X]$ est seulement factoriel. On établira cette propriété au chapitre 5, en liaison avec d'utiles *critères d'irréductibilité*. Retenons dès à présent:

1) *Lemme de Gauss.*- Soit A un anneau factoriel, et $P = a_o + a_1X + ... + a_nX^n$ un élément de $A[X]$. On appelle *contenu* de P le pgcd $cont(P) = pgcd(a_0, ..., a_n)$ de ses coefficients. Pour

P, Q dans $A[X]$, on a: $cont(PQ) = cont(P)cont(Q)$. En particulier, $A[X]$ est factoriel, et un élément P de $\mathbf{Z}[X]$ est irréductible dans $\mathbf{Q}[X]$ si (et quand $cont(P) = 1$, seulement si) il est irréductible dans $\mathbf{Z}[X]$.

2) Soit p un nombre premier, $\pi : \mathbf{Z}[X] \to \mathbf{F}_p[X]$ l'homomorphisme de réduction modulo p (réduire modulo p chacun des coefficients), et $P \in \mathbf{Z}[X]$ de coefficient dominant a_n premier à p. Alors P est irréductible dans $\mathbf{Q}[X]$ dès que $\pi(P)$ l'est dans $\mathbf{F}_p[X]$.

3) *Critère d'Eisenstein.*- Soit $P = a_o + a_1X + ... + a_nX^n \in \mathbf{Z}[X]$, et p un nombre premier tel que $p \not| a_n, p|a_i$ pour $i < n, p^2 \not| a_0$; alors P est irréductible dans $\mathbf{Q}[X]$.

On en déduit (poser $Y = X - 1$) que pour p premier, le polynôme cyclotomique $\Phi_p(X) = 1 + X... + X^{p-1}$ est irréductible dans $\mathbf{Q}[X]$. L'irréductibilité, plus subtile, de tous les polynômes cyclotomiques $\Phi_n(X)$ dans $\mathbf{Q}[X]$ se déduit des résultats du chap. 3.

4) *Discriminant.*- Soient K un corps, $P = a_nX^n + a_{n-1}X^{n-1} + ... + a_0 \in K[X]$ un polynôme de degré $n \geq 1$, de *polynôme dérivé* $DP(X) = \Sigma_{i=1,..,n} i a_i X^{i-1}$. On dit que P est *séparable* si $(P, DP) = 1$. Alors (exercice), P n'est pas divisible par un carré dans $K[X]$ (i.e. par le carré d'un élément non unité), et la réciproque est vraie quand K est algébriquement clos. Ainsi, $P \in K[T]$ est séparable ssi il n'admet pas de racine multiple dans la clôture algébrique de K (ou dans son corps de décomposition sur K, cf. §3).

De plus, il existe une expression polynômiale 'universelle' $D_n(a_0, ..., a_{n-1}, a_n)$ à coefficients dans $\mathbf{Z}$ (dont la valeur dans K s'appelle le discriminant $Disc(P)$ de P), telle que, toujours sous l'hypothèse $a_n \neq 0$: $(P, DP) \neq 1 \Leftrightarrow Disc(P) = 0$. Par exemple,

$$Disc(X^n + a_1X + a_0) = (-1)^{n(n-1)/2}((1-n)^{n-1}a_1^n + n^n a_0^{n-1}).$$

Un homomorphisme d'anneaux dont la source est un corps, est injectif ou nul. Si K est un corps, le noyau de l'homomorphisme de $\mathbf{Z}$ dans K défini par $n \mapsto n.1$ est un idéal premier de $\mathbf{Z}$, donc égal à 0, ou à $p\mathbf{Z}$, pour un nombre p premier. On dit respectivement que K est de *caractéristique* nulle, ou égale à p. Dans ce deuxième cas, K contient le corps $\mathbf{F}_p$, et l'application

$$F : K \to K : x \mapsto x^p$$

est un endomorphisme d'anneau, induisant l'identité sur $\mathbf{F}_p$ (d'après Fermat), et appelé *Frobenius* de K (cf. chapitre 2). Si $car(K) = 0$, tout élément irréductible P de $K[X]$ vérifie $(P, DP) = 1$, et est donc séparable. Si $car(K) = p$, il en est de même quand le Frobenius de K est un automorphisme de K (par exemple, pour $K = \mathbf{F}_p$): en effet, un élément P irréductible non séparable vérifie $DP = 0$, donc $\exists S \in K[X]$ tel que $P(X) = S(X^p) = ((F^{-1}(S))(X))^p$, qui n'est pas irréductible dans $K[X]$. Mais si F n'est pas surjectif sur K, il existe des polynômes irréductibles non séparables (cf. chapitre 3).

§3. Extensions algébriques

Soit K un corps (commutatif). On appelle *extension* de K la donnée d'un corps L et d'un homomorphisme d'anneaux (unitaires) i de K dans L. Comme K est un corps, i est nécessairement injectif, et on identifie en général K à $i(K)$, en notant simplement L/K une telle extension.

La multiplication des éléments de L par ceux de $i(K)$ munit L d'une structure naturelle d'espace vectoriel sur K . La dimension (éventuellement infinie) de L en tant que K-esp. vect. est appelée *degré* de l'extension L/K, et est notée $[L : K]$ (ne pas confondre avec l'indice d'un sous-groupe). On dit qu'une extension est *finie* si son degré l'est.

On appelle extension intermédiaire M de L/K la donnée de deux extensions M/K et L/M telles que la composée des injections sous-entendues correspondantes soit l'injection sous-entendue de L/K. Pour toute partie S de L , on désigne par $K[S]$ le sous-anneau de L engendré par K et S, et par $K(S) = Fr(K[S])$ l'extension intermédiaire de L/K engendrée par K et S dans L. Si $S = \{s\}$ est réduite à un élément, on dit que $K(s)$ est une extension *monogène*. Si S est fini, on dit que $K(S)/K$ est une extension *de type fini* (mais $K(S)/K$ ne sera pas forcément une extension finie; exemple: $S = \{T\}$ dans le corps $K(T)$ des fractions rationnelles à coefficients dans K.)

Théorème 1: *Soient L/K une extension, et M une extension intermédiaire. Alors,* $[L : K] = [L : M][M : K]$.

Démonstration: dans le cas où $[M : K] = m$ et $[L : M] = n$ sont finis, soit $\{\alpha_1, ..., \alpha_m\}$ (resp. $\{\beta_1, ..., \beta_n\}$) une base de M de K (resp. de L sur M). Alors, la famille $\{\alpha_i\beta_j; 1 \leq i \leq m, 1 \leq j \leq n\}$ forme une base de l'espace vectoriel L sur K, qui a donc pour dimension mn. Même argument dans le cas où une des extensions est infinie, en remplaçant les entiers m, n par des cardinaux.

Application (exercices): le problème de trisection de l'angle ou de la duplication du cube ne peut être résolu par des constructions à la règle et au compas.

Soit L/K une extension, et α un élément de L. L'ensemble

$$J_\alpha = \{P \in K[T],\ P(\alpha) = 0 \text{ dans } L\}$$

est un idéal de l'anneau pricipal $K[T]$, distinct de $K[T]$. Si $J_\alpha = 0$, i.e. si α ne vérifie aucune relation polynômiale non triviale à coefficients dans K, on dit que α est *transcendant* sur K. Sinon, on dit que α est algébrique sur K; l'idéal J_α, distinct de 0 et de $K[T]$, admet alors un générateur unitaire de degré ≥ 1, appelé *polynôme minimal* de α sur K, et noté $Min_{\alpha,K}$ (ou M_α si la référence à K est claire); son degré s'appelle le *degré* de α sur K.

Théorème 2: *i) Soient L/K une extension, et α un élément de L algébrique sur K. Alors, $Min_{\alpha,K}$ est irréductible dans $K[T]$, de degré $n = [K(\alpha) : K]$, l'extension $K(\alpha)/K$ admet $\{1, \alpha, ..., \alpha^{n-1}\}$ pour base sur K, et $K(\alpha) = K[\alpha]$.*
ii) Soit $P \in K[T]$, irréductible et unitaire. Alors, il existe une extension M/K et un élément α de M tels que $P = Min_{\alpha,K}$ et $M = K(\alpha)$. Une telle extension est appelée corps de rupture de P.

Démonstration: i) soit $ev_\alpha : K[T] \to L : Q \mapsto Q(\alpha)$ l'homomorphisme d'anneau 'évaluation en α'. Par passage au quotient par son noyau J_α, il définit un isomorphisme de $K[T]/J_\alpha$ sur un sous-anneau (forcément intègre) de L, donc J_α est un idéal premier, non nul, donc maximal ((cf. §2, Proposition, iii). Ainsi, son générateur M_α est irréductible et $K[T]/J_\alpha$ est un corps, dont l'image $K[\alpha]$ par ev_α coïncide donc avec $K(\alpha)$, et est un espace vectoriel sur K de dimension n, engendré par les images de $1, T, ..., T^{n-1}$.

ii) l'anneau $M = K[T]/(P)$ est un corps si P est irréductible, et la classe α de T modulo (P) vérifie les conditions souhaitées.
[Dans la pratique, on calcule l'inverse d'un élément $\beta \neq 0$ de M, donné sous la forme $\beta = Q(\alpha)$, avec $P \not| Q$, en appliquant Bézout: $\exists\ A, Q' \in K[T], AP + QQ' = 1$; alors $\beta^{-1} = Q'(\alpha)$.]

Un *homomorphisme* d'une extension (i, L) de K vers une extension (i', L') de K est un homomorphisme d'anneau $\phi : L \to L'$ tel que $\phi o i = i'$. On dit alors que ϕ est un K-homomorphisme de L vers L' (ou: un homomorphisme de L/K vers L'/K; comme ϕ est injectif, on parle aussi de *plongement* au lieu d'homomorphisme). Plus généralement, si ϕ est un homomorphisme de K vers un corps K', et (i, L) (resp. (i', L') une extension de K (resp. K'), un homomorphisme d'anneau $\psi : L \to L'$ tel que $\psi o i = i' o \phi$ s'appelle un homomorphisme de L vers L' *au-dessus* de ϕ (ou: étendant ϕ). Idem pour les iso., les endo., les automorphismes.

Le théorème de prolongement suivant entraîne en particulier que pour $P \in K[T]$ irréductible donné, tous les corps de rupture de P sont K-isomorphes.

Théorème 3 : *soient K, K' deux corps, ϕ un homomorphisme de K dans K', $L/K, L'/K'$ deux extensions, et α un élément de L algébrique sur K.*

i) (1er théorème de prolongement) *Soit $\alpha' \in L'$ algébrique sur K', tel que $\phi(Min_{\alpha,K}) = Min_{\alpha',K'}$. Alors, il existe un unique homomorphisme ψ de $K(\alpha)$ dans L' au-dessus de ϕ tel que $\psi(\alpha) = \alpha'$.*

ii) Soit ψ un homomorphisme de $K(\alpha)$ dans L' au-dessus de ϕ. Alors, $\alpha' := \psi(\alpha)$ est algébrique sur $\phi(K)$ (donc sur K'), et vérifie $Min_{\alpha',\phi(K)} = \phi(Min_{\alpha,K})$.

[Le nombre d'homomorphismes distincts de $K(\alpha)$ dans L' au dessus de ϕ est donc égal au nombre de racines distintes de $\phi(Min_{\alpha,K})$ dans L'.]

Démonstration: i) l'homomorphisme d'évaluation $Q \in K[T] \mapsto (\phi(Q))(\alpha') \in L'$ passe au quotient $K[T]/J_\alpha$ par son noyau, et induit un plongement $\psi : K(\alpha) \to L'$ envoyant α sur α', et dont la restriction à K vaut ϕ. Cela détermine entièrement ψ sur $K(\alpha)$.

ii) $(\phi(Min_{\alpha,K}))(\alpha') = \psi(Min_{\alpha,K}(\alpha)) = 0$, et l'homomorphisme d'anneau $\phi : K[T] \to \phi(K)[T]$ préserve l'irréductibilité.

On dit qu'une *extension L/K est algébrique* si *tous* ses éléments sont algébriques sur K; elle est dite transcendante dans le cas contraire (i.e. s'il existe un élement de L transcendant sur K). Dans le cas d'une extension monogène, les propriétés suivantes sont équivalentes: $K(\alpha)/K$ est algébrique $\Leftrightarrow$ $K(\alpha)/K$ est finie $\Leftrightarrow$ $K[\alpha] = K(\alpha)$. D'où pour une extension quelconque:

$$L/K \text{ est finie} \Leftrightarrow L/K \text{ est algébrique } \textit{et} \text{ de type fini.}$$

Soient L/K une extension, et $x, y \in L$ algébriques sur K. Alors, $K(x,y)/K$ est une extension finie (de degré $\leq [K(x) : K].[K(y) : K]$), donc algébrique. En particulier, $x+y$ et xy sont algébriques sur K. Ainsi, *l'ensemble des éléments de L algébriques sur K forme un sous-corps de L*, appelé fermeture algébrique de K *dans* L. C'est évidemment un extension algébrique de K, mais elle n'est pas nécessairement finie (exemple: la fermeture algébrique $\overline{\mathbf{Q}}$ de $\mathbf{Q}$ dans $\mathbf{C}$). Par ailleurs, elle est sa propre fermeture algébrique dans L. En effet, pour toute extension intermédiaire $L/M/K$, avec M/K *algébrique, tout $x \in L$ algébrique sur M est algébrique sur K* (si S désigne l'ensemble des coefficients de $Min_{x,M}$, x est algébrique sur $K(S)$, et $K(S,x)/K$ est une extension finie, donc algébrique.)

On dit qu'un corps Ω est *algébriquement clos* si les éléments irréductibles de $\Omega[T]$ sont tous de degré 1, i.e. si tout polynôme de degré ≥ 1 admet une racine dans Ω (tout polynôme se décompose alors en facteurs linéaires dans Ω). On dit qu'une extension $\overline{K}/K$ est une *clôture algébrique de K* si les deux conditions suivantes sont simultanément réalisées: $\overline{K}/K$ est une extension algébrique et $\overline{K}$ est un corps algébriquement clos.

Théorème: *tout corps K admet une clôture algébrique $\overline{K}$. De plus, deux clôtures algébriques de K sont K-isomorphes.*

Démonstration: nous nous contentons ici de vérifier l'existence d'une clôture algébrique en supposant qu'on dispose d'une extension Ω de K algébriquement close, mais peut-être transcendante (grâce à l'Analyse ou la Topologie, c'est le cas de $K = \mathbf{Q}$, avec $\Omega = \mathbf{C}$). Soient donc $\overline{K}$ la fermeture algébrique de K dans Ω, F un élément de $\overline{K}[T]$ et ω une racine

de $P \in \Omega[T]$ dans le corps alg. clos Ω. Comme on vient de le voir, ω appartient à $\overline{K}$, qui est donc à la fois algébrique sur K et algébriquement clos.

Soit $F \in K[T]$ un polynôme de degré $n \geq 1$, non nécessairement irréductible, de terme dominant a_n. On dit qu'une extension L/K est un *corps de décomposition* de F sur K s'il existe n éléments (non nécessairement distincts) $\alpha_1, ..., \alpha_n$ de L tels que les deux conditions suivantes soient simultanément réalisées:

$$L = K(\alpha_1, ..., \alpha_n) \text{ et } F(T) = a_n \Pi_{i=1,...,n}(T - \alpha_i) \text{ dans } L[T].$$

La première condition entraîne alors que L/K est une extension finie. On exprime la seconde en disant que F se décompose en facteurs linéaires dans L, ou, plus simplement, qu'il *se décompose* (sous-entendu: complètement) dans L.

Théorème 4: *Soient K un corps, et F un élément de $K[T]$ de degré $n \geq 1$.*

i) F admet un corps de décomposition L sur K.

ii) (2ème théorème de prolongement) *Soient $\phi : K \to K'$ un homomorphisme, et L' une extension de K' dans laquelle $\phi(F)$ se décompose en facteurs linéaires. Alors, il existe un homomorphisme (non unique) ψ de L dans L' au-dessus de ϕ. De plus:*

iii) Si ϕ est un isomorphisme, et si L' est un corps de décomposition de $\phi(F)$ sur K', alors, ψ est un isomorphisme.

[Deux corps de décomposition de F sur K sont donc K-isomorphes, et on peut parler, à K-isomorphisme près, 'du' corps de décomposition de F sur K.]

Démonstration. i) C'est clair pour $n = 1$, et on procède par récurrence sur n (sans fixer K). Supposons que F ne se décompose pas en facteurs linéaires sur K. Il admet alors un facteur M_1 irréductible sur K de degré > 1. D'après le théorème 2, il existe K_1/K et $\alpha_1 \in K_1$ racine de M_1 tels que $K_1 = K(\alpha_1)$, de sorte que $F(T)$ s'écrit $(T - \alpha_1)F_1(T)$ dans $K_1[T]$. On conclut en appliquant l'hypothèse de récurrence au couple (F_1, K_1).

ii) Par récurrence sur n. Soit $\phi(F) = \phi(a_n)\Pi_{i=1,...,n}(T - \beta_i)$ la décomposition de $\phi(F)$ dans L', et M_1 le polynôme minimal de α_1 sur K, de degré $r \leq n$. Après réindexation, on peut écrire $\phi(M_1) = \Pi_{i=1,...,r}(T - \beta_i)$ et $\phi(M_1)$, irréductible sur $\phi(K)$, est le polynôme minimal de β_1 sur $\phi(K)$. D'après le théorème 3.i, il existe un homomorphisme ϕ_1 de $K_1 = K(\alpha_1)$ dans L' au-dessus de ϕ envoyant α_1 sur β_1. On conclut en appliquant l'hypothèse de récurrence à $F_1(T) := F(T)/(T - \alpha_1) \in K_1[T])$, $\phi_1 : K_1 \to K_1' = K'(\beta_1)$.

iii) Soient $\phi(F) = \phi(a_n)\Pi_{i=1,...,n}(T - \beta_i)$ la factorisation de $\phi(F)$ dans le corps de décomposition $L' = K(\beta_1, ..., \beta_n)$, et ψ un homomorphisme de L dans L' au-dessus de ϕ. D'après le théorème 3.ii, ψ envoie les α_i sur certains des β_j et par injectivité, préserve les multiplicités. Donc tous sont atteints, et ψ est un isomorphisme.

CHAPITRE II

CORPS FINIS

§1. Les corps $\mathbf{F}_q$.

Soient p un nombre premier, et K un corps de caractéristique p. Comme on l'a vu, K contient $\mathbf{F}_p$, et est donc une extension de $\mathbf{F}_p$. Dans ces conditions, K est un corps fini (c'est-à-dire: possède un nombre fini q d'éléments), si et seulement si l'extension $K/\mathbf{F}_p$ est finie (c'est-à-dire: est de degré d fini). Le $\mathbf{F}_p$-espace vectoriel K a alors $q = p^d$ éléments. Ainsi un corps fini a pour cardinal une puissance pure de sa caractéristique (noter qu'un corps fini est forcément de caractéristique $\neq 0$). Inversément:

Théorème 1: *soient $q = p^d$ une puissance de p. Le corps de décomposition du poly. $G_d(T) = T^{p^d} - T$ sur $\mathbf{F}_p$ est, à isomorphisme près, l'unique corps de cardinal q. On le note $\mathbf{F}_q$.*
[Remarque: pour $d > 1$, il n'y a aucun rapport entre $\mathbf{F}_{p^d}$ et $\mathbf{Z}/p^d\mathbf{Z}$: le premier est un corps, le deuxième un anneau non intègre; le groupe additif de $\mathbf{Z}/p^d\mathbf{Z}$ est cyclique, celui de $\mathbf{F}_{p^d}$ un produit de d groupes cycliques.]

La démonstration du théorème 1 repose sur le

Lemme 1: *Soit K un corps fini, et p^n son cardinal. Alors,*
i) le Frobenius F de K est un automorphisme de K, qui vérifie $F^n = id_K$; en particulier, tout élément irréductible de $K[T]$ est séparable.
ii) pour tout entier $m \geq 1$, le polynôme $G_m(X) = X^{p^m} - X$ est séparable (ses racines dans toute extension de K sont donc simples), et pour $m = n$, $G_n(T) = \Pi_{x \in K}(T - x)$.

Démonstration. i) Rappelons (I, §2) que le Frobenius de K est l'application $F : x \mapsto x^p$. C'est bien un endomorphisme de l'anneau K, puisque $F(xy) = (xy)^p = F(x)F(y)$, $F(1) = 1$, et $F(x + y) = \Sigma_{k=0,...,p} C_p^k x^k y^{p-k} = x^p + y^p = F(x) + F(y)$ (noter que $\forall k = 1, ..., p-1, p | C_p^k$). Comme sa source K est un corps, F est injectif, et donc surjectif puisque c'est un endomorphisme du groupe fini $(K, +)$. Ainsi, F est un automorphisme. Comme le groupe multiplicatif K^* est d'ordre $p^n - 1$, ses éléments x vérifient $x^{p^n - 1} = 1$, d'où

pour tout $x \in K, F^n(x) = x^{p^n} = x = id_K(x)$. (NB: le fait que F^n est l'identité entraîne d'ailleurs directement que F est un automorphisme.) Comme F est un automorphisme de K, la deuxième assertion résulte de I, fin du §2.

ii) La première assertion est claire ($DG_m(X) = -1$ est premier à G_m), et on vient de voir que les $p^n = deg(G_n)$ éléments de K sont racines de G_n.

Démonstration du théorème 1. Soient L le corps de décomposition du polynôme $G_d(T)$ sur $\mathbf{F}_p$, F son Frobenius, et M l'ensemble des racines de G_d dans L. Par définition, L est le sous-corps de L engendré sur $\mathbf{F}_p$ par l'ensemble M. Mais on vérifie à la main (ou en disant que c'est l'ensemble des éléments x de L tels que l'automorphisme $\sigma = F^d$ de l'extension $L/\mathbf{F}_p$ vérifie $\sigma(x) = x$) que M est lui-même un corps. Donc $M = L$. Par ailleurs, M a q éléments, puisque G_d est un polynôme séparable, et L est bien un corps fini à q éléments. Enfin, d'après le Lemme 1.ii), tout corps K de cardinal $q = p^d$ est un corps de décomposition du polynôme G_d sur $\mathbf{F}_p$, et est donc isomorphe à L.

Fixons une clôture algébrique $\overline{\mathbf{F}}_p$ de $\mathbf{F}_p$. Le théorème 1 entraîne que pour tout entier d, il y a une et une seule extension $\mathbf{F}_{p^d}$ de $\mathbf{F}_p$ de degré d contenue dans $\overline{\mathbf{F}}_p$: le corps $\mathbf{F}_p(\zeta)$ engendré par une racine primitive (p^d-1)-ième de l'unité dans $\overline{\mathbf{F}}_p$. (Rien de tel, bien sûr, pour $\overline{\mathbf{Q}}$, qui contient pour tout $d > 1$, une infinité d'extensions de $\mathbf{Q}$ de degré d, deux à deux non isomorphes.) Ces extensions s'imbriquent de la façon suivante, qui permet d'ailleurs de définir $\overline{\mathbf{F}}_p$ comme la réunion croissante des corps $\mathbf{F}_{p^{n!}},\ n \geq 1$.

Proposition 1: *soient n et m deux entiers ≥ 1. Alors, $\mathbf{F}_{p^n} \subset \mathbf{F}_{p^m}$ si et seulement si $n|m$.*

Démonstration: si $\mathbf{F}_{p^m}$ est une extension de $\mathbf{F}_{p^n}$, son degré d vérifie $p^m = (p^n)^d$, donc $m = dn$. Inversément, si $n|m$, $p^n - 1$ divise $p^m - 1$, donc $\mu_{p^n-1}(\overline{\mathbf{F}}_p) \subset \mu_{p^m-1}(\overline{\mathbf{F}}_p)$.

Remarque 1: considérons, avec les notations $m = dn$ précédentes, le n-ième itéré F^n du Frobenius de $\mathbf{F}_{p^m}$. D'après le lemme 1, c'est un élément du groupe $G = Aut(\mathbf{F}_{p^m}/\mathbf{F}_{p^n})$ des automorphismes de l'extension $\mathbf{F}_{p^m}/\mathbf{F}_{p^n}$, i.e. des automorphismes du corps $\mathbf{F}_{p^m}$ induisant l'identité sur $\mathbf{F}_{p^n}$. On déduit des résultats du Chap. 3 que G est un groupe cyclique, d'ordre d, engendré par l'automorphisme F^n. (Voir feuilles TD pour une preuve directe.)

En particulier, pour tout entier m, le groupe $Aut(\mathbf{F}_{p^m}/\mathbf{F}_p) \simeq \mathbf{Z}/m\mathbf{Z}$ est engendré par le Frobenius F de $\mathbf{F}_{p^m}$. On peut néanmoins montrer que $Aut(\overline{\mathbf{F}}_p/\mathbf{F}_p)$ est beaucoup plus gros que le groupe cyclique (isomorphe à $\mathbf{Z}$) qu'y engendre le Frobenius de $\overline{\mathbf{F}}_p$.

Remarque 2: On dit qu'une extension algébrique L/K est *séparable* si pour tout $\alpha \in L$, le polynôme $Min_{\alpha,K}$ est séparable. Soit alors K un corps *fini*. Le lemme 1 entraîne que

toute extension algébrique L/K est séparable; si de plus L/K est une extension finie, on a vu qu'il existe un élément ζ de L tel que $L = K(\zeta)$. Ainsi, toute extension finie d'un corps fini est monogène. Plus généralement, on montrera au chapitre 3 que pour tout corps K, toute extension finie et séparable de K est monogène. En revanche, le corps $K = \mathbf{F}_p(X)$, de caractéristique p mais infini, admet des extensions finies non séparables (example: L = le corps de rupture du polynôme $P(T) = T^p - X$, qui est irréductible sur K mais est une puissance p-ième dans $L[T]$), tandis que le corps $K = \mathbf{F}_p(X, Y)$ admet même des extensions finies non monogènes (exemple: L = le corps de décomposition du polynôme $F(T) = (T^p - X)(T^p - Y)$ sur K).

§2. La loi de réciprocité quadratique.

Soit p un nombre premier. On dit qu'un entier rationnel a est un *résidu quadratique modulo* p si sa classe $\overline{a}$ dans $\mathbf{Z}/p\mathbf{Z}$ y est un carré, i.e. si l'équation $X^2 - \overline{a} = 0$ a une racine dans le corps $\mathbf{F}_p$ (donc deux pour $p \not| a$). Si $p = 2$, tout entier est résidu quadratique modulo 2 (le seul polynôme irréductible de degré 2 de $\mathbf{F}_2[X]$ est $X^2 + X + 1$). C'est aux nombres premiers impairs qu'on s'intéresse donc ici.

Soit p un nombre premier *impair* et a un entier rationnel, de classe $\overline{a}$ modulo p non nulle. On appelle *symbole de Legendre* de a (ou de $\overline{a}$), et on note $(\frac{a}{p})$ (ou $(\frac{\overline{a}}{p})$) le nombre défini par

$$(\tfrac{a}{p}) = 1 \quad \text{si } a \text{ est résidu quadratique modulo } p \ , \ (\tfrac{a}{p}) = -1 \quad \text{sinon} \ .$$

On étend le symbole de Legendre à tous les entiers en posant $(\frac{a}{p}) = 0$ si p divise a. Par ailleurs, on identifie les 3 valeurs $\{\pm 1, 0\}$ du symbole de Legendre tantôt à des nombres complexes, tantôt à des éléments du corps $\mathbf{F}_p$ (ou de n'importe quel corps de caract. $\neq 2$).

Lemme 2: *i) Pour tout entier* a, $(\frac{a}{p}) = \overline{a}^{\ \frac{p-1}{2}}$; *en particulier,* $(\frac{1}{p}) = 1, (\frac{-1}{p}) = (-1)^{\frac{p-1}{2}}$;
ii) $\forall a, b \in \mathbf{Z}, (\frac{ab}{p}) = (\frac{a}{p})(\frac{b}{p})$, *et* $\Sigma_{x \in \mathbf{F_p}} (\frac{x}{p}) = 0$;
iii) $(\frac{2}{p}) = (-1)^{\frac{p^2-1}{8}}$.

Démonstration: i) considérons les endomorphismes χ et ψ du groupe $\mathbf{F}_p^*$ défini par $\chi(x) = x^{\ \frac{p-1}{2}}, \psi(x) = x^2$. D'après Fermat, $Im\chi \subset Ker\psi = \{\pm 1\}$, et $Im\psi \subset Ker\chi$. Mais $Ker\chi$, formé des éléments d'ordre divisant $(p-1)/2$, coïncide donc avec l'unique sous-groupe d'ordre $(p-1)/2$ du groupe cyclique $\mathbf{F}_p^*$. Idem pour $Im\psi$, d'indice $|Ker\psi| = 2$ dans $\mathbf{F}_p^*$. Finalement, $Ker\chi = Im\psi$, et $(\frac{a}{p}) = 1 \Leftrightarrow \overline{a} \in Im\psi \Leftrightarrow \chi(\overline{a}) = 1$, d'où $(\frac{a}{p}) = \chi(\overline{a})$ pour tout a premier à p, soit $(\frac{a}{p}) = \overline{a}^{\ \frac{p-1}{2}}$ pour tout a.
ii) La première relation est claire sur χ. Comme $Ker\chi$ est d'indice 2, il y a autant de résidus quadratiques que de non résidus quadr. dans $\mathbf{F}_p^*$, d'où la seconde relation.

iii) Soit ζ une racine primitive 8-ième de l'unité dans $\overline{\mathbf{F}}_p$. Alors, $\zeta^4 = -1$ et $\theta := \zeta + \zeta^{-1}$ a pour carré 2. On conclut en notant que $(\frac{2}{p}).\theta = \theta^{p-1}\theta = \zeta^p + \zeta^{-p}$ est égal à θ si $p \equiv \pm 1$ mod 8 , et à $-\theta$ si $p \equiv \pm 3$ mod 8.

Remarque 3: soient G un groupe fini. On appelle *caractère* de G tout homomorphisme de groupe de G dans $\mathbf{C}^*$, i.e. toute application $\chi : G \to \mathbf{C}^*$ telle que pour tout $x, y \in G : \chi(xy) = \chi(x)\chi(y)$. En définissant le produit χ de deux caractères χ_1, χ_2 par $\chi(x) = \chi_1(x)\chi_2(x)$, on munit l'ensemble $\hat{G}$ des caractères de G d'une structure de groupe, d'élément neutre le caractère trivial $\mathbf{1} : x \mapsto \mathbf{1}(x) := 1$. Avec $\{\pm 1\} \subset \mathbf{C}^*$, le symbole de Legendre devient un caractère du groupe $G = \mathbf{F}_p^*$, d'ordre 2 dans $\hat{G}$, et le lemme 2.ii est un cas particulier des relations générales (exercice):

$$\forall \chi \in \hat{G}, \chi \neq \mathbf{1} : \ \Sigma_{x \in G}\ \chi(x) = 0\ ;\ \Sigma_{x \in G}\ \mathbf{1}(x) = |G|.$$

Théorème 2 (Gauss): *soient p et ℓ deux nombres premiers impairs distincts. Alors*

$$(\frac{\ell}{p})(\frac{p}{\ell}) = (-1)^{\frac{p-1}{2}\frac{\ell-1}{2}}.$$

Démonstration: soit ζ une racine primitive p-ième de l'unité dans $\overline{\mathbf{F}}_\ell$, ou dans $\mathbf{C}$. On va calculer de deux façons la "somme de Gauss", bien définie dans chacun de ces corps par la formule

$$S = \sum_{x=0}^{p-1} (\frac{x}{p})\ \zeta^x = \sum\nolimits_{x \in \mathbf{F}_p^*} (\frac{x}{p})\ \zeta^x.$$

(1) *Qu'on lise dans $\overline{\mathbf{F}}_\ell$ ou dans $\mathbf{C}$, on a :* $S^2 = (-1)^{\frac{p-1}{2}} p$.

En effet, $S^2 = \Sigma_{(x,y)}(\frac{xy}{p})\ \zeta^{x+y} = \Sigma_{t \neq 0}(\frac{t}{p})(\Sigma_{x \neq 0}\zeta^{x(1+t)})$, puisque pour $y = xt, (\frac{xy}{p}) = (\frac{x^2}{p})(\frac{t}{p}) = (\frac{t}{p})$. Comme $x \mapsto \zeta^{x(1+t)}$ est un caractère de $(\mathbf{F}_p, +)$, $\Sigma_{x \neq 0}\zeta^{x(1+t)}$ vaut $p-1$ si $p|1+t$, et -1 sinon. Ainsi, $S^2 = (\frac{-1}{p})(p-1) + (-1)\Sigma_{t \neq -1}(\frac{t}{p})$, qui vaut $(\frac{-1}{p})p$ d'après le lemme 2, ii).

(2) *Dans $\overline{\mathbf{F}}_\ell$, on a :* $S^{\ell-1} = (\frac{\ell}{p})$.

En effet, comme on est maintenant en caractéristique ℓ, $S^\ell = \Sigma_x(\frac{x}{p})\ \zeta^{\ell x} = (\frac{\ell}{p})\Sigma_x(\frac{\ell x}{p})\ \zeta^{\ell x} = (\frac{\ell}{p})S$, et on peut diviser par S, qui d'après (1), est non nul dans $\overline{\mathbf{F}}_\ell$.

En regroupant (1) et (2), on obtient finalement:

$$(\frac{\ell}{p}) = S^{\ell-1} = (-1)^{\frac{p-1}{2}\frac{\ell-1}{2}} p^{\frac{\ell-1}{2}} = (-1)^{\frac{p-1}{2}\frac{\ell-1}{2}} (\frac{p}{\ell}).$$

Cette identité, établie modulo ℓ, ne relie que des 1 et des -1, et persiste donc en toute caractéristique.

Remarque 4: la formule (1), lue en caractéristique 0, montre que pour tout nombre premier p congru à 1 (resp. à 3) modulo 4, l'extension quadratique $\mathbf{Q}(\sqrt{p})$ (resp. $\mathbf{Q}(\sqrt{-p})$ de $\mathbf{Q}$ est contenue dans le corps $\mathbf{Q}(e^{2i\pi/p})$; dans tous les cas, on a donc: $\mathbf{Q}(\sqrt{p}) \subset \mathbf{Q}(e^{2i\pi/8p})$. Plus généralement, on déduit de la "théorie du corps de classes" que toute extension abélienne de $\mathbf{Q}$ est contenue dans une extension cyclotomique.

§3. Factorisation dans $\mathbf{F}_p[X]$. Algorithme de Berlekamp.

Soient p un nombre premier, et P un polynôme unitaire à coefficients dans $\mathbf{Z}$. Comme on l'a dit en I, §2, il suffit, pour que P soit irréductible dans $\mathbf{Q}[X]$, que son image $\pi(P)$ par l'application de réduction modulo p soit irréductible dans $\mathbf{F}_p[X]$. Bien sûr, c'est loin d'être nécessaire: prendre $P(T) = T^2 - \ell$, avec ℓ premier, $(\frac{\ell}{p}) = 1$; ou encore: $P(T) = \Phi_{p^d-1}(T)$, irréductible sur $\mathbf{Q}$ d'après la Proposition 3 ci-dessous, bien que tous ses facteurs irréductibles modulo p aient pour degré d (un tel facteur irréductible $\overline{A}$ est le polynôme minimal sur $\mathbf{F}_p$ d'une racine de Φ_{p^d-1}, i.e. d'une racine primitive $p^d - 1$-ième de l'unité dans $\overline{\mathbf{F}}_p$; comme le corps engendré sur $\mathbf{F}_p$ par une telle racine est $\mathbf{F}_{p^d}$, c'est que $deg(\overline{A}) = d$; en particulier, $d|\phi(p^d - 1)$, ce qu'on peut retrouver en notant que p est un élément d'ordre d de $(\mathbf{Z}/(p^d - 1)\mathbf{Z})^*$). Néanmoins, le critère suivant d'irréductibilité sur $\mathbf{F}_p$ s'avère parfois utile.

Proposition 2: *Soit $A \in \mathbf{F}_p[X]$ un polynôme de degré m. Alors, A est irréductible si et seulement si l'on a simultanément*

$$(X^{p^m} - X, A(X)) = A(X), \text{ et } \forall \ell \text{ premier }, \ell | m : (X^{p^{m/\ell}} - X, A(X)) = 1.$$

Démonstration: si A est irréductible, son corps de rupture $\mathbf{F}_p[X]/A(X)\mathbf{F}_p[X] = \mathbf{F}_p(\overline{X})$ sur $\mathbf{F}_p$ a p^m éléments, donc coïncide avec $\mathbf{F}_{p^m}$, et l'itéré de F d'ordre minimal stabilisant $\overline{X}$ est F^m. Inversément, tout facteur irréductible de $X^{p^m} - X$ qui ne divise aucun des $X^{p^n} - X$, $n|m, n \neq m$, admet $\mathbf{F}_{p^m}$ pour corps de rupture sur $\mathbf{F}_p$, et est donc de degré m.

Voici, à titre d'exercice, un autre type d'application de la réduction modulo p à l'irréductibilité dans $\mathbf{Z}[X]$.

Proposition 3: *pour tout entier $n \geq 1$, le polynôme cyclotomique $\Phi_n(X)$ est irréductible dans $\mathbf{Q}[X]$.*

Démonstration: soient ζ une racine primitive n-ième de l'unité dans $\overline{\mathbf{Q}}$, et P son polynôme minimal sur $\mathbf{Q}$. D'après le lemme de Gauss (1, §2), on a $\Phi_n = PQ$, avec P et Q unitaires à

coefficients dans $\mathbf{Z}$. On va montrer que toutes les racines primitives n-ième de l'unité sont racines de P, de sorte que $\Phi_n = P$ est bien irréductible sur $\mathbf{Q}$. Par induction, il suffit de montrer que pour tout nombre premier p ne divisant pas n, la racine ζ^p de Φ_n est racine de P. Sinon, $Q(\zeta^p) = 0$, ζ est racine du polynôme $R(X) := Q(X^p)$, qui est donc divisible par P, et d'après le lemme de Gauss, $\exists S \in \mathbf{Z}[X]$ tel que $R = PS$. D'où par réduction modulo p: $\pi(R) = \pi(P)\pi(S)$. Mais $\pi(R(X)) = \pi(Q(X^p)) = \pi(Q(X)^p)$ dans $\mathbf{F}_p[X]$, soit $\pi(P)\pi(S) = (\pi(Q))^p$. Ainsi, tout facteur irréductible $\overline{A}$ de $\pi(P)$ dans $\mathbf{F}_p[X]$ divise $\pi(Q)$, et $\pi(\Phi_n) = \pi(P)\pi(Q)$ est divisible par $\overline{A}^2$. Mais $p \not| n$, donc Φ_n est un polynôme séparable de $\mathbf{F}_p[X]$, donc sans facteur carré.

Enfin, l'énoncé suivant fournit à la fois un critère d'irréductibilité et un algorithme de factorisation dans $\mathbf{F}_p[X]$.

Proposition 4 (Berlekamp): *Soit $A \in \mathbf{F}_p[X]$ un polynôme sans facteur carrés, et $A_1...A_r$ sa décomposition en polynômes irréductibles dans $\mathbf{F}_p[X]$. Les polynômes $Q \in \mathbf{F}_p[X]$ nuls ou de degré $< deg(A)$, qui, pour tout $i = 1, ..., r$, sont congrus modulo A_i à un élément de $\mathbf{F}_p$, sont exactement les p^r polynômes Q nuls ou de degré $< deg(A)$, tels que $Q(X)^p \equiv Q(X)$ $mod(A(X))$.*

Démonstration: puisque les A_i sont irréductibles et distincts, ils sont premiers entre eux deux à deux, et le lemme chinois, appliqué à l'anneau principal $\mathbf{F}_p[X]$, permet d'attacher à tout r-uplet $(s_1, ..., s_r) \in (\mathbf{F}_p)^r$ de polynômes constants un unique polynôme Q nul ou degré $< deg(A)$ tel que A_i divise $Q - s_i$ pour tout i. Alors, $Q^p \equiv s_i^p = s_i \equiv Q \bmod(A_i)$ pour tout i, et A divise $Q^p - Q$. Inversément, $X^p - X = \Pi_{s \in \mathbf{F}_p}(X - s)$ dans $\mathbf{F}_p[X]$, donc tout polynôme Q vérifie $Q^p - Q = \Pi_{s \in \mathbf{F}_p}(Q - s)$. Par conséquent, si A divise $Q^p - Q$, chaque facteur irréductible A_i de A divise l'un des facteurs de ce produit, i.e. il existe pour tout $i = 1, ...r$ un polynôme constant s_i tel que $Q \equiv s_i \bmod A_i$. Restreinte aux polynômes Q nuls ou de degré $< deg(A)$, l'application $Q \mapsto (s_1, ..., s_r)$ est la réciproque de la précédente, d'où la bijection recherchée.

Application: si S désigne l'endomorphisme $Q \mapsto Q^p$ du $\mathbf{F}_p$-espace vectoriel $\mathbf{F}_p[X]/(A)$, le noyau de $S - Id$ a pour dimension $\rho = r$, d'où comme promis un critère d'irreductibilité ($\rho = 1$) sur $\mathbf{F}_p$. Et si $\rho \geq 2$, la proposition montre que pour tout élément Q du noyau, il existe $s \in \mathbf{F}_p$ tel que A et $Q - s$ ont un pgcd non trivial, d'où un algorithme de factorisation dans $\mathbf{F}_p[X]$.

CHAPITRE III

THÉORIE DE GALOIS

§1. Extensions galoisiennes.

Soient K un corps et L/K une extension algébrique. On dit que L/K est *normale* si tout élément irréductible de $K[T]$ qui admet une racine dans L se décompose dans $L[T]$ en facteurs linéaires, autrement dit si, pour tout élément α de L, il existe des éléments (non nécessairement distincts) $\alpha_1 = \alpha$, α_2, .., α_d de L tels que $Min_{\alpha,K}(T) = \Pi_{i=1,...,d}(T - \alpha_i)$; ces racines de $Min_{\alpha,K}$ sont appelées les *conjugués* de α sur K (dans L). On dit que L/K est une extension *galoisienne* si elle est à la fois normale et séparable, autrement dit si tout $\alpha \in L$ de degré d sur K admet d conjugués distincts dans L. (Rappelons que L/K est dite *séparable* si $\forall \alpha \in L$, le polynôme $Min_{\alpha,K}$ est séparable.)

Lemme 1: *i) soit L/K une extension algébrique. Alors, L/K est normale et finie si et slt s'il existe un élément F de $K[T]$ tel que L soit un corps de décomposition de F sur K.*

ii) Soit L/K une extension normale. Pour tout élément α de L, l'ensemble des conjugués de α dans L coïncide avec le sous-ensemble de L formé par les images de α sous les différents automorphismes de L/K.

Démonstration: i) si L/K est finie, il existe un ensemble fini S d'éléments de L algébriques sur K tels que $L = K(S)$, et si L/K est de plus normale, $F(T) := \Pi_{\alpha \in S} Min_{\alpha,K}(T)$ vérifie la condition demandée.

Inversement, soit L le corps de décomposition d'un élément F de $K[T]$. C'est évidemment une extension finie de K. Soient alors P un élément irréductible de $K[T]$ admettant une racine α dans L, N un corps de décomposition de P *sur* L, et α' une racine quelconque de P dans N. Le 1er théorème de prolongement fournit un K-isomorphisme ϕ de $K(\alpha)$ sur $K(\alpha')$. Puisque L (resp. $L(\alpha')$) est un corps de décomposition de F (resp. $\phi(F)$) sur $K(\alpha)$ (resp. $K(\alpha')$), on déduit du 2ème théorème de prolongement un isomorphisme ψ de L sur $L(\alpha')$ au-dessus de ϕ. En comparant les degrés, on conclut que $\alpha' \in L$.

ii) Pour tout automorphisme σ de L/K, $Min_{\alpha,K}(\sigma\alpha) = \sigma(Min_{\alpha,K}(\alpha)) = 0$, et $\sigma\alpha$ est un conjugué de α sur K. Inversement, soit α' un conjugué de α sur K. Traitons

d'abord le cas où L/K est une extension finie, donc par i), le corps de décomposition d'un $F \in K[T]$. D'après le 1er théorème de prolongement, il existe un K-plongement ϕ de $K(\alpha)$ dans L envoyant α sur α'. D'après le 2e théorème de prolongement, ϕ se prolonge en un automorphisme σ de L/K, et $\sigma(\alpha) = \phi(\alpha) = \alpha'$. D'où la bijection ensembliste annoncée (qui explique qu'on appelle souvent *conjugaisons de* L/K les automorphismes d'une extension normale L/K). Le cas général s'en déduit au moyen du lemme de Zorn (considérer l'ensemble des couples (L', ψ') formés d'une extension L'/K normale, contenue dans L et d'un automorphisme ψ' de L' au-dessus de ϕ, muni de la relation d'ordre $(L_1, \psi_1) < (L_2, \psi_2)$ si $L_1 \subset L_2$ et ψ_2 prolonge ψ_1.)

Pour toute extension intermédiaire M/K d'une extension normale L/K, l'extension L/M est normale (car $\forall \alpha \in L$, $Min_{\alpha,M} | Min_{\alpha,K}$; en revanche, M/K n'est en général pas normale). Par ailleurs, M/K et L/M peuvent être normales sans que L/K le soit. Si L/K est séparable, M/K l'est évidemment, ainsi que L/M (car tout diviseur d'un polynôme séparable est séparable). On verra plus bas que M/K et L/M séparables $\Rightarrow L/K$ séparable.

Lemme 2 : *Soient L/K une extension* normale, *et M/K une extension intermédiaire.*

i) Soient ϕ un K-homomorphisme de M dans L, et α un élément de L, de degré d sur M. Le nombre d' de racines distinctes de $\phi(Min_{\alpha,M})$ dans L vérifie $1 \leq d' \leq d$, et est égal au nombre d'homomorphismes de $M(\alpha)$ dans L au dessus de ϕ; de plus, $d' = d$ si et seulement si $Min_{\alpha,M}$ est séparable. Plus généralement:

ii) Supposons $m = [M : K]$ fini. Le nombre m' de K-plongements de M dans L vérifie $1 \leq m' \leq m$, et vaut m si et seulement si M/K est séparable.

iii) M/K est normale si et seulement si pour tout K-plongement ϕ de M dans L , on a: $\phi(M) \subset M$ (et ϕ est alors un automorphisme de M/K).

En regroupant les deux dernières conclusions de ce lemme, et en notant $Aut(L/K)$ le groupe des automorphismes d'une extension algébrique L/K (i.e. des automorphismes σ du corps L tels que $\forall x \in K, \sigma(x) = x$), on obtient finalement:

Théorème 1 : *une extension finie L/K est galoisienne si et seulement si le groupe $Aut(L/K)$ est d'ordre $[L : K]$.*

Démonstration du lemme 2: i) D'après le 1er théorème de prolongement, le nombre de plongements de $M(\alpha)$ dans L au-dessus de ϕ est égal d'. De plus, L/K étant normale, $P = Min_{\alpha,K}$ se décompose en facteurs linéaires dans L, et il en est de même de $Q = Min_{\alpha,M}$ et de $\phi(Q)$, qui divisent $P = \phi(P)$. En particulier $d' \geq 1$, et $d' = d$ ssi $\phi(Q)$ est séparable sur K. Puisque $\phi(Q, DQ) = (\phi(Q), \phi(DQ))$, cela équivaut à demander que Q soit séparable.

ii) récurrence sur m. Si $m > 1$, il existe une extension M_0 de K et un élément α de M de degré $d > 1$ sur M_0 tel que $M = M_0(\alpha)$. Par i), tout K-plongement ϕ de M_0 dans L admet au plus d prolongements ψ à M, et exactement d si M/M_0, est séparable. Comme tout K-plongement ψ de M dans L induit un ϕ sur M_0, on déduit de l'hypothèse de récurrence qu'il y a $1 \leq m' \leq d[M_0 : K] = m$ K-plongements de M dans L, et exactement m si M/K, donc M/M_0 et M_0/K, sont séparables.

Supposons maintenant que M/K ne soit pas séparable. Alors, il existe $\alpha \in M$, de degré $d > 1$ sur K, non séparable sur K, et par i), $d' < d$ K-plongements de $K(\alpha)$ dans L. On vient de voir que chacun d'eux a au plus m/d prolongements à M. Le nombre de K-plongements de M dans L est au donc au plus égal à $d'm/d < m$.
[NB: si M/K n'est pas séparable, de sorte que K est de caractéristique p non nulle, on peut en fait démontrer que $m'|m$, et que $\frac{m}{m'}$ est une puissance pure de p.]

iii) L'image, par un K-plongement ϕ de M dans L, d'un élément α de M est un conjugué dans L de α sur K. Si M/K est normale, ces conjugués sont déjà tous dans M, donc $\phi(M) \subset M$, et ϕ est un K-endomorphisme de M. Enfin, *tout endomorphisme ϕ d'une extension* algébrique M/K *est un automorphisme*: c'est clair si $[M : K] < \infty$; dans le cas général, remarquer que pour tout $\beta \in M$, l'ensemble B des racines de $Min_{\beta,K}$ dans M est fini, et que ϕ induit une injection, donc une bijection, de B sur lui-même; par conséquent, β admet dans $B \subset M$ un antécédent sous ϕ, et ϕ est surjectif.

Supposons maintenant que tout K-plongement de M dans L laisse stable M. Soient α un élément de M et α' un conjugué de α sur K. D'après le lemme 1.ii, il existe un automomorphisme ψ de L/K tel que $\psi(\alpha) = \alpha'$. Par hypothèse, la restriction ϕ de ψ à M laisse stable M, donc $\alpha' = \psi(\alpha) = \phi(\alpha) \in M$. Ainsi, $Min_{\alpha,K}$ se décompose en facteurs linéaires dans M, et M/K est normale.
Preuve du Théorème 1 : on vient de voir que pour toute extension algébrique, $Aut(L/K)$ coïncide avec l'ensemble des K-plongements de L dans L. Si L/K est galoisienne finie, il a, d'après le lemme 2.ii, $[L : K]$ éléments. Inversement, soient L/K une extension de degré fini $m = |Aut(L/K)|$, et $\tilde{L}/K$ une extension normale contenant L (par exemple, une clôture algébrique de L). Alors, l'ensemble de K-plongements de L dans $\tilde{L}$ a au plus m éléments. Comme il contient $Aut(L/K)$, il en a donc m (d'où la séparabilité de L/K grâce au lemme 2.ii), et il coïncide avec ce groupe (d'où la normalité de L/K grâce au lemme 2.iii).

Remarque 1: i) soient L une extension normale de K (par exemple, une clôture algébrique de K), et M/K une extension intermédiaire. D'après le lemme 2.iii, l'ensemble des extensions normales de K contenant M et contenues dans L admet pour la relation d'inclusion

un plus petit élément $\tilde{M}$, appelé *clôture normale de M dans L*. Si M/K est finie, $\tilde{M}/K$ l'est aussi: en effet, $M = K(S)$, où $S \subset M$ est fini, le polynôme $\Pi_{\alpha \in S} Min_{\alpha,K}$ se décompose en facteurs linéaires dans L, et son corps de décomposition $\tilde{M} \subset L$ sur M répond à la question.

ii) On déduit du lemme 2.ii que si M/K et L/M sont deux extensions séparables, alors L/K est séparable. (En considérant les polynômes minimaux, on se ramène au cas où M/K est finie, et $L = M(\alpha)$. Alors, M admet $[M : K]$ K-plongements distincts dans $\overline{K}$, et chacun d'eux se prolonge en $[L : M]$ plongements distincts de L dans $\overline{K}$.)

Lemme 3: *Toute extension L/K finie et séparable est monogène (i.e. admet un élément γ, dit primitif pour L/K, tel que $L = K(\gamma)$.)*

Démonstration: on l'a vérifié au chapitre II quand K est un corps fini. Supposons maintenant K infini. L/K étant de type fini, il suffit par induction de montrer l'assertion quand $L = K(\alpha, \beta)$. Soit $\tilde{L}$ un corps de décomposition de $F = Min_{\alpha,K} Min_{\beta,K}$ sur L. D'après le lemme 2.ii, il existe $n = [L : K]$ K-plongements distincts $\phi_1, ..., \phi_n$ de L dans $\tilde{L}$. Le polynôme

$$P(T) = \Pi_{1 \leq i \neq j \leq n}[(\phi_i(\alpha) - \phi_j(\alpha)) + (\phi_i(\beta) - \phi_j(\beta))T] \in \tilde{L}[T]$$

est alors non nul; K étant infini, il existe donc $c \in K$ tel que $P(c) \neq 0$, autrement dit tel que les éléments $\phi_i(\alpha + c\beta), i = 1, ..., n$ de $\tilde{L}$ sont distincts. Mais ce sont des racines du polynôme minimal de $\gamma := \alpha + c\beta$ sur K, qui est donc de degré $\geq n$. Comme $\gamma \in L$, ce degré est majoré par n, et $L = K(\gamma)$.

§2. La correspondance de Galois.

Soient L/K une extension algébrique, et $G = Aut(L/K)$ le groupe des automorphismes de l'extension L/K. Pour toute extension intermédiaire M, le groupe $Aut(L/M) = \{\sigma \in G; \forall x \in M, \sigma(x) = x\}$ des automorphismes de l'extension L/M est un sous-groupe de G. Inversement, pour tout sous-groupe H de G, le corps $L^H := \{x \in L; \forall \sigma \in H, \sigma(x) = x\}$ est une extension intermédiaire de l'extension L/K, appelé *sous-corps de L fixé par H*. On a:

$$H < H' < G \Rightarrow K \subset L^{H'} \subset L^H \subset L \quad ;$$
$$K \subset M \subset M' \subset L \Rightarrow Aut(L/M') < Aut(L/M) < G.$$

Toute extension intermédiaire M est contenue dans $L^{Aut(L/M)}$, mais il se peut qu'elle le soit strictement (exemple: $K = M = \mathbf{Q} \subset L = \mathbf{Q}(\sqrt[3]{2})$, pour lequel $Aut(L/K) = \{id_L\}$, $L^{Aut(L/K)} = L$). De même, tout sous-groupe H de G est contenu dans $Aut(L/L^H)$, mais

peut l'être strictement (exemple hors programme: $K = \mathbf{F}_p \subset L = \overline{K}, H = \{F^n, n \in \mathbf{Z}\}$, où F désigne le Frobenius de L; alors $L^H = K$, mais $Aut(L/K)$ est bien plus gros que H). Pour simplifier l'exposé, nous nous limiterons pour l'essentiel aux groupes finis, où ce deuxième type de contre-exemples n'apparaît pas. Quant au premier type de contre-exemples, il disparaît lorsqu'on se restreint à des extensions L/K *galoisiennes*.

Pour L/K galoisienne, le groupe $Aut(L/K)$ s'appelle le *groupe de Galois de l'extension galoisienne* L/K, et est souvent noté $Gal(L/K)$.

Théorème 2: *i) Soient L/k une extension quelconque, G un sous-groupe* fini *de $Aut(L/k)$, et K le corps L^G. Alors, L/K est une extension galoisienne finie, et $Aut(L/K) = G$.*

ii) Soient L/k une extension algébrique, *G un sous-groupe quelconque de $Aut(L/k)$, et K le corps L^G. Alors, L/K est une extension galoisienne*

iii) Soit L/K une extension galoisienne. Alors, $L^{Aut(L/K)} = K$.

On obtient ainsi une nouveau critère pour qu'une extension algébrique L/K, finie ou non, soit galoisienne: il faut et il suffit que $L^{Aut(L/K)} = K$.

Démonstration: i) Soient α un élément de L, et $\sigma_1 = \mathrm{id}_L, \sigma_2, ...\sigma_r$ un sous-ensemble *maximal* de G tel que $\sigma_1\alpha, ..., \sigma_r\alpha$ soient distincts. Posons $P(T) = \Pi_{i=1,...r}(T - \sigma_i\alpha) \in L[T]$. Alors, $\sigma(P) = P$ pour tout $\sigma \in G$, de sorte que $P \in K[T]$, et α est algébrique sur K. De plus, $Min_{\alpha,K}$ divise P, mais est aussi divisible par P puisque $Min_{\alpha,K}(\sigma_i\alpha) = \sigma_i(Min_{\alpha,K}(\alpha)) = 0$; bref, $Min_{\alpha,K} = P$. Ainsi, le polynôme minimal de α sur K se décompose en facteurs linéaires distincts dans L, et L/K est galoisienne.

On vient de montrer que le polynôme minimal sur K de tout élément de L est de degré $\leq |G|$, i.e. que tous les éléments de L sont de degé $\leq |G|$ sur K. Comme L/K est séparable, il résulte alors du lemme 3 que $[L : K] \leq |G|$. Enfin, $[L : K] = |Aut(L/K)|$ d'après le théorème 1, et G est un sous-groupe de $Aut(L/L^G)$. Par conséquent, L/K est degré $\geq |G|$, donc finalement égal à $|G|$, et G remplit tout $Aut(L/K)$.

ii) Si L/k est algébrique, l'orbite sous G, éventuellement infini, de tout élément α de L est formée de racines de $Min_{\alpha,k}$, et est donc finie. La première partie du raisonnement de i) reste donc valable, et L/K est galoisienne.

iii) Soit α un élément de L, de degré d sur K. D'après les lemmes 2.i et 1.ii, il existe d automorphismes $\sigma_1 = \mathrm{id}, ..., \sigma_d$ de L/K tels que $\{\sigma_i\alpha; i = 1, ...d\}$ forme l'ensemble des racines de $Min_{\alpha,K}$ dans L. Si maintenant $\alpha \in K^G$, on a $\sigma_i(\alpha) = \alpha$ pour tout $i = 1, ..., d$, et $d = 1$. Ainsi, α appartient à K, et $K^G = K$.

Remarque 3: pour démontrer l'inégalité $[L : K] \geq |G|$ au théorème 2.i, on peut également procéder de la manière suivante. Voyons $G \subset Aut(L/K)$ comme une famille d'éléments du

L-espace vectoriel $F(\Gamma, L)$ des fonctions sur $\Gamma := L^*$, à valeurs dans L. Si $[L : K] < |G|$, elle est linéairement dépendante sur L. Mais les éléments de G sont aussi des caractères du groupe Γ (à valeurs dans le groupe multiplicatif du corps L), et un théorème classique, valable pour tout groupe Γ et tout corps L, affirme que des caractères distincts de Γ forment une famille L-libre dans $F(\Gamma, L)$.

Théorème de Galois: *soient L/K une extension galoisienne finie, et $G = Gal(L/K)$ son groupe de Galois.*

i) Soit M un extension intermédiaire, et $H = Gal(L/M) < G$ le groupe de Galois de l'extension (galoisienne) L/M; alors, $[L : M] = |H|$, et $M = L^H$;

ii) Soit H un sous-groupe de G, et $M = L^H$ le sous-corps de L fixé par H; alors L/M est une extension galoisienne de degré $[L : M] = |H|$, et $H = Gal(L/M)$;

iii) Une extension intermédiaire M/K est galoisienne si et seulement si $H := Gal(L/M)$ est un sous-groupe distingué de G; dans ce cas, son groupe de Galois $Gal(M/K)$ s'identifie au groupe quotient G/H.

Démonstration: i) et ii) sont des réécritures du théorème 2. (Mais en ii), le fait que L/L^H est galoisienne n'est pas profond, puisqu'on part d'une extension L/K galoisienne.) Passons à iii).

Supposons M/K galoisienne, et soient $\tau \in G, \sigma \in H$. Il s'agit de voir que $\tau^{-1}\sigma\tau$ appartient à H, i.e. fixe chaque élément x de M. Comme $\tau(x)$ est une racine de $Min_{x,M}$ et que M/K est normale, $\tau(x) \in M$, et est donc fixé par σ. Alors, $\tau^{-1}\sigma\tau(x) = \tau^{-1}\tau(x) = x$.

Supposons H distingué dans G. Il s'agit de voir que M/K est normale (comme L/K est séparable, on sait déjà que M/K l'est aussi). Les conjugués dans L d'un élément $x \in M$ sont de la forme $\tau(x), \tau \in G$. Mais pour tout $\sigma \in H, \tau^{-1}\sigma\tau := \sigma'$ appartient à $H \lhd G$, donc fixe x, et $\sigma\tau(x) = \tau\sigma'(x) = \tau(x)$, soit $\tau(x) \in L^H = M$.

Dans ces conditions, tout automorphisme σ de L/K induit sur M un automorphisme $\sigma_{|M}$ de M (cf. lemme 2.iii), de sorte que l'application $\pi : \sigma \mapsto \sigma_{|M}$ définit un homomorphisme de groupe de G dans $Gal(M/K)$. Mais π est surjective d'après le deuxième théorème de prolongement, et son noyau est formé des automorphismes de L/K induisant l'identité sur M. Autrement dit, $Ker(\pi) = Gal(L/M)$, et $Gal(M/K) \simeq G/H$.

Le théorème de Galois montre que l'étude des extensions intermédiaires d'une extension galoisienne finie L/K équivaut à celle des sous-groupes de $Gal(L/K)$. Par exemple, on en déduit immédiatement que L/K n'a qu'un nombre fini d'extensions intermédiaires (cela reste vrai si L/K est seulement séparable, sinon, c'est en général faux). De ce fait, on désigne souvent du même adjectif les propriétés d'une extension galoisienne et celles de

leurs groupes de Galois. Ainsi, une extension galoisienne est dite abélienne (resp. cyclique) si son groupe de Galois l'est. Dans l'autre sens, on voit pourquoi 'normal' et 'distingué' sont synonymes en théorie des groupes; quant au mot 'résoluble', il fait l'objet du §3.

§3. Applications

Pour alléger l'exposé, nous présentons ces applications sous forme de séries d'exercices, et nous nous plaçons en caractéristique nulle (toutes les extensions sont donc séparables).

a) Resolubilite par radicaux

Soit K un corps (de caractéristique nulle). Une extension M/K est dite *résoluble* s'il existe une tour d'extensions $K = K_0 \subset K_1 \subset ... \subset K_m = M$ telle que pour tout $i = 1, ..., m$, il existe un élément β_i de K_i et un entier $n_i > 0$ tel que $\beta_i^{n_i} \in K_{i-1}$ et $K_i = K_{i-1}(\beta_i)$. Pour $F \in K[T]$, de corps de décomposition K_F sur K, on dit que l'équation $F(x) = 0$ est *résoluble par radicaux* si l'extension K_F/K est contenue dans une extension résoluble de K. Par ailleurs, on appelle souvent *groupe de Galois de* F le groupe $Gal(K_F/K)$.

Dans le problème étudié ici, on peut se ramener au cas où K contient pour tout entier $n > 0$ une racine primitive n-ième de l'unité. C'est ce que nous supposerons désormais. Si M/K est résoluble, chacune des extensions intermédiaires K_i/K_{i-1} est alors galoisienne, de groupe de Galois isomorphe à un sous-groupe de $\mathbf{Z}/n_i\mathbf{Z}$, donc abélien (et même cyclique). Inversement, on peut démontrer, au moyen du "théorème 90 de Hilbert", que toute extension cyclique K''/K', de degré n, d'un corps K' contenant une racine primitive n-ième de l'unité, est de la forme $K'' = K'(\beta^{\frac{1}{n}})$, où $\beta \in K'$.
Ex. 1: soit M/K une extension résoluble, et N/K une clôture normale de M/K. Alors, N/K est une extension résoluble.

Soit G un groupe. On dit que G est *résoluble* s'il admet une suite de sous-groupes $G_m = \{e\} < G_{m-1} < ... < G_1 < G_0 = G$ telle que pour tout $i = 1, ..., m$, G_i soit distingué dans G_{i-1} et G_{i-1}/G_i soit abélien.

On démontre en théorie des groupes que si G est un groupe résoluble, et si $H \triangleleft G$, alors G/H est résoluble.

Ex. 2: Soit M/K une extension résoluble , et L/K une extension intermédiaire. Alors, $Aut(L/K)$ est un groupe résoluble.
[Noter que si $K' = L^{Aut(L/K)}$, M/K' est résoluble, et L/K' est galoisienne, de groupe de Galois $Aut(L/K') = Aut(L/K)$. Par l'ex. 1 et la remarque précédente sur G/H, il suffit alors de montrer que si N/K est galoisienne et résoluble, le groupe $Gal(N/K)$ est résoluble.]

On obtient en définitive:

Théorème A: *une équation $F(x) = 0$ à coefficients dans K est résoluble par radicaux si et seulement si son groupe de Galois est résoluble.*

b) L'équation générale de degré n.

Soient k un corps de caractéristique 0, et $\mathbf{K} = k(u_1, ..., u_n)$ le corps des fractions rationnelles en n variables. *L'equation générale de degré n* est l'équation $\mathbf{P}_n(x) = 0$, où $\mathbf{P}_n(T) = T^n - u_1 T^{n-1} + ... + (-1)^n u_n \in \mathbf{K}[T]$. Soit $\mathbf{L} = \mathbf{K}(v_1, ..., v_n)$ le corps de décomposition de $\mathbf{P}_n$ sur $\mathbf{K}$, avec $\mathbf{P}_n(T) = \Pi_{i=1,...,n}(T - v_i)$. On a: $u_1 = v_1 + ... + v_n$, $u_2 = \Sigma_{1 \leq i < j \leq n} v_i v_j$, ... , $u_n = \Pi_{i=1,...,n} v_i$.

Soit par ailleurs $L = k(x_1, ..., x_n)$ le corps des fractions rationnelles en n nouvelles variables $x_1, ..., x_n$. L'action du groupe symétrique S_n sur L, définie, pour tout $f \in L$, par $(\sigma f)(x_1, ..., x_n) := f(x_{\sigma(1)}, ..., x_{\sigma(n)})$, identifie S_n à un sous-groupe de $Aut(L/k)$. Les élements du corps $S = L^{S_n}$ sont appelés les *fonctions symétriques* en les x_i. Soit $K := k(s_1, ..., s_n)$ le sous-corps de S engendré par les fonctions symétriques *élémentaires* $s_1 = x_1 + ... + x_n$, $s_2 = \Sigma_{1 \leq i < j \leq n} x_i x_j$, ... , $s_n = \Pi_{i=1,...,n} x_i$.

Théorème B': *toute fonction symétrique s'exprime comme une fraction rationelle en les fonctions symétriques élémentaires. Autrement dit, on a $K = S$ (et $Aut(L/K) = S_n$).*
[D'après le thm. 3.i, $[L : S] = |S_n| = n!$, donc $[L : K] \geq n!$. Pour l'inégalité inverse, noter que pour tout corps K', le corps de décomposition sur K' de tout élément de $K'[T]$ de degré n est de degré au plus $n!$. Considérer alors $P_n(T) = T^n - s_1 T^{n-1} + ... + (-1)^n s_n \in K[T]$.]

On peut montrer que l'homomorphisme d'anneaux $ev : k[u_1, ..., u_n] \to K : P \mapsto P(s_1, ..., s_n)$ est injectif. Il s'etend donc en un isomorphisme ϕ de $\mathbf{K}$ sur K.
Ex. 3: les extensions $\mathbf{L}/\mathbf{K}$ et L/K sont isomorphes. En particulier, *l'équation générale de degré n admet S_n pour groupe de Galois.*

Or on démontre en théorie des groupes que S_n est un groupe résoluble si $n \leq 4$, et non résoluble dès que $n \geq 5$. Du théorème A, on déduit donc:

Théorème B: *l'équation générale de degré n est résoluble par radicaux si et slt si $n \leq 4$.*

Reste à trouver les formules par radicaux donnant les solutions de ces équations. Expliquons la méthode pour $n = 3$ (formule de Cardan). On suppose comme toujours que $car(k) = 0$, et pour simplifier, que k contient les racines cubiques ρ, ρ^2 de l'unité. On met sans difficulté $\mathbf{P}_3$ sous la forme $\mathcal{P}(T) = T^3 + pT + q = (T - v_1)(T - v_2)(T - v_3)$, de discriminant $D = -4p^3 - 27q^2$, qui n'est pas un carré dans le corps de base $\mathcal{K} = k(p, q)$. Le sous-corps de $\mathcal{L} = \mathcal{K}(v_1, v_2, v_3)$ fixé par le groupe alterné A_3, d'indice 2 dans $S_3 =$

$Gal(\mathcal{L}/\mathcal{K})$, est l'extension quadratique $\mathcal{K}(\sqrt{D})$, avec $\sqrt{D} = (v_1 - v_2)(v_2 - v_3)(v_3 - v_1)$. Considérons alors les éléments $\eta = v_1 + \rho v_2 + \rho^2 v_3$, $\zeta = v_1 + \rho^2 v_2 + \rho v_3$ de $\mathcal{L}$. Leurs cubes sont invariants sous A_3, donc appartiennent à $\mathcal{K}(\sqrt{D})$. On vérifie en explicitant le théorème B (formules de Newton, avec $v_1 + v_2 + v_3 = 0$) que

$$\eta^3 = \frac{1}{2}(-27q + 3\sqrt{-3}\sqrt{D}) \ , \ \zeta^3 = \frac{1}{2}(-27q - 3\sqrt{-3}\sqrt{D}),$$

où $\sqrt{-3} = \rho - \rho^2$. De plus, $\eta\zeta$ est invariant sous S_3, donc appartient à $\mathcal{K}$ (le calcul donne $\eta\zeta = -3p$). Ainsi,

$$v_1 = \frac{1}{3}(\eta + \zeta), \ v_2 = \frac{1}{3}(\rho^2\eta + \rho\zeta) \ , v_3 = \frac{1}{3}(\rho\eta + \rho^2\zeta).$$

La condition sur $\eta\zeta$ assure qu'on obtient ainsi, à permutation paire près, un unique triplet de solutions de $\mathcal{P}(x) = 0$.

c) Constructions à la règle et au compas

Soient S (resp. P) une collection finie de points (resp. un point) du plan, et K (resp. $K(P)$) le corps engendré par les coordonnées des points de S (resp. l'extension de K engendrée par celles de P) dans un repère orthonormé. 0n suppose que l'origine et le point $(0, 1)$ appartiennent à S. Comme l'équation aux abscisses des points d'intersection d'une droite ou d'un cercle avec une droite ou un cercle est au plus quadratique, dire que P est constructible à la règle et au compas à partir de S équivaut à dire qu'il existe une tour d'extensions $K_0 = K \subset K_1 \subset ... \subset K_m$ telle que pour tout $i = 1, ..., m$, K_i/K_{i-1} soit une extension quadratique et que K_m contienne $K(P)$.

Ex. 4: avec les notations ci-dessus, soit N/K une clôture normale de K_m/K. Montrer que $[N : K]$ est une puissance de 2.

Soient p un nombre premier, et G un groupe fini. On dit que G est un p-groupe si son ordre est une puissance de p. On montre en théorie des groupes que tout p-groupe est résoluble. En considérant les extensions intermédaires attachées à la suite de sous-groupes correspondants, on en déduit:

Théorème C : *P est constructible à la règle et au compas à partir de S si et seulement s'il existe une extension N de $K(P)$, galoisienne sur K, telle que $Gal(N/K)$ soit un 2-groupe.*

On appelle nombre premier de Fermat tout nombre premier de la forme $2^{2^n} + 1$, où $n \geq 0$. On n'en connait pour l'instant que 5 ($n = 0, 1, 2, 3, 4$).

Ex. 5: Soit N un entier ≥ 1. On considère dans le plan complexe l'ensemble $S = \{0, 1\}$, et le point $P_N = exp(2i\pi/N)$. Montrer que P_N est constructible à la règle et au compas

à partir de S si et seulement si N est le produit d'une puissance de 2 par un produit de nombres premiers de Fermat distincts.
[Indication: utiliser II, Proposition 3.] On laisse au lecteur le soin de découper une pizza (à la règle et au compas) en 65 537 parties égales.

CHAPITRE IV

QUELQUES ALGORITHMES SUR Z

§1. Modules sur les anneaux principaux.

La notion de module sur un anneau A (utile pour étudier les extensions d'anneaux) est modelée sur celle des espaces vectoriels sur un corps (dont on a vu le rôle pour les extensions de corps). Nous donnons ici un survol de cette théorie, dont nous utiliserons la terminologie générale; les énoncés eux-mêmes ne nous seront utiles que pour $A = \mathbf{Z}$, où elle se réduit à l'étude des groupes abéliens.

Soient A un anneau commutatif et unitaire. Un A-module est la donnée d'un groupe commutatif $(M, +)$, muni d'une loi de composition externe $A \times M \to M : (a, m) \mapsto am$ distributive pour les lois d'addition de A et de M et vérifiant pour tout $a, b \in A, m \in M$: $a(bm) = (ab)m, 1.m = m$. Un homomorphisme $f : M \to M'$ est un homomorphisme de groupe vérifiant $f(am) = af(m)$ pour tout $a \in A, m \in M$. Les notions de famille libre, de famille génératrice, de base, de somme directe, se définissent comme pour les espaces vectoriels. On dit qu'un A-module est de type fini (resp. libre) s'il admet une famille génératrice finie (resp. une base). Le rang d'un A-module est le maximum (fini ou non) des cardinaux de ses familles libres finies. Si un A-module M est libre de type fini, il est de rang n fini, et M est alors isomorphe au A-module A^n.

Pour tout élément x de M, l'*annulateur* $Ann(x) = \{a \in A, ax = 0\}$ est un idéal de A. L'idéal $Ann(M) = \cap_{x \in M} Ann(x)$ s'appelle l'annulateur de M. Les éléments x de M tels que $Ann(x) \neq (0)$ s'appellent les éléments de torsion de M. Si A est intègre, l'ensemble des éléments de torsion de M forme un sous-module M_{tor} de M. Si $M_{tor} = 0$, on dit que M est un module sans torsion; c'est le cas des espaces vectoriels, même sur les corps finis.

On dit qu'un sous-module N de M est facteur direct s'il existe un sous-module N' (qu'on appelle alors un supplémentaire de N) tel que $M = N \oplus N'$.

Proposition 1 : *i) soit N un sous-module d'un A-module M, tel que le quotient M/N soit un A-module* libre. *Alors, N est facteur direct dans M.*

ii) Soit M un module sur un anneau intègre *A. Alors, M/M_{tor} est sans torsion. Si M est libre, alors M est sans torsion.*

Démonstration: i) soit $\mathcal{E}$ un système représentatif dans M d'une base de M/N. On vérifie aisément que le A-module N' engendré par $\mathcal{E}$ dans M est un supplémentaire de N.

ii) est élémentaire. Noter qu'en général, un A-module sans torsion n'est pas forcément libre (exemples: les idéaux non principaux de A; le $\mathbf{Z}$-module sans torsion $\mathbf{Q}$).

Nous nous restreignons désormais au cas des anneaux *principaux*, et même, en ce qui concerne les algorithmes, à $\mathbf{Z}$, ou tout au moins à un anneau euclidien. Le théorème 1.1 (ii) (sous la version équivalente 1.3) ne sera démontré que dans ce cas. (Voir par exemple le livre de P. Samuel *Théorie algébrique des nombres* pour le cas général.)

Théorème 1.1 : *Soient A un anneau principal, $M \simeq A^n$ un A-module libre de type fini, de rang n, et N un sous-A-module de M. Alors*

i) N est libre de type fini, de rang $s \leq n$

ii) il existe une base $\{e_1, \ldots, e_n\}$ de M et des éléments $a_1, \ldots, a_s$ de A tels que a_i divise a_{i+1} pour tout $i = 1, \ldots, s-1$, et que $\{a_1 e_1, \ldots, a_s e_s\}$ soit une base de N.

Les idéaux $\{(a_1), \ldots, (a_s)\}$ ne dépendent que M et N, et s'appellent les facteurs invariants de N dans M. Tout module de type fini étant quotient d'un module libre de type fini, on déduit du théorème 1.1 le corollaire fondamental suivant.

Théorème 1.2 : *Soit M un module de type fini sur un anneau principal A. Alors,*

i) M_{tor} est facteur direct dans M, et M est libre si et seulement s'il est sans torsion;

ii) plus précisément, il existe un entier $r \geq 0$ et des idéaux $(a_s) \subset (a_{s-1}) \subset \ldots \subset (a_1) \neq A$ tels que $M \simeq A^r \oplus A/(a_1) \oplus \ldots \oplus A/(a_s)$. L'entier r est le rang de M, et les idéaux (a_i), qu'on appelle les facteurs invariants de M, ne dépendent que de M; par exemple, $(a_s) = Ann(M_{tor})$.

En voici, pour $A = \mathbf{Z}$, trois applications immédiates:

Éléments primitifs de $\mathbf{Z}^n$: ce sont les vecteurs $x = (x_1, \ldots, x_n)$ de $\mathbf{Z}^n$ vérifiant les propriétés équivalentes suivantes:

i) x est la première colonne d'un élément de $GL_n(\mathbf{Z})$;

ii) les entiers $x_1, \ldots, x_n$ sont premiers entre eux dans leur ensemble;

iii) $\mathbf{Z}^n/\mathbf{Z}x$ est un $\mathbf{Z}$-module sans torsion;

iv) $\mathbf{Z}^n \cap \mathbf{R}x = \mathbf{Z}x$

Groupes abéliens finis : tout groupe abélien $G \neq 0$ fini est isomorphe à un produit de groupes cycliques $\mathbf{Z}/a_1\mathbf{Z} \times \ldots \times \mathbf{Z}/a_s\mathbf{Z}$, où $a_1, \ldots, a_s$ sont des entiers > 1 tels que a_i divise a_{i+1} pour tout $i = 1, \ldots, s-1$. Ces entiers, ainsi que s, ne dépendent que de G (ainsi, leur produit est l'ordre de G; a_s est l'exposant de G.)

Forme normale de Smith: soit B une matrice $m \times n$ (m lignes, n colonnes) à coefficients dans $\mathbf{Z}$, de rang s. On dit qu'elle est sous forme de Smith si tous ses coefficients sont nuls en dehors de ses s premiers élements diagonaux, qui sont positifs et vérifient pour tout $i < s$: $b_{i,i}$ divise $b_{i+1,i+1}$.

Théorème 1.3 : *Soit A une matrice $m \times n$ à coefficients dans $\mathbf{Z}$, de rang s.*

i) Il existe $V \in GL_m(\mathbf{Z})$ et $U \in GL_n(\mathbf{Z})$ telle que $B = VAU$ soit sous forme de Smith.

ii) Pour tout entier $r = 0, 1, ..., s$, soit $\Delta_r(A)$ le pgcd des mineurs d'ordre r de A, avec $\Delta_0 = 1$. Alors, $b_{r,r} = \frac{\Delta_r(A)}{\Delta_{r-1}(A)}$. La matrice B de i) est donc indépendante de U et de V.

Démonstration: i) par récurrence, il suffit de montrer que l'ensemble $\mathcal{A}$ des matrices VAU semblables à A contient une matrice B dont les coefficients $b_{1,j}, b_{i,1}$ sont nuls pour $i, j > 1$. C'est clair si $A = 0$. Sinon, des inversions de lignes et de colonnes montrent que $\mathcal{A}$ contient une matrice B, dont le coefficient $b_{1,1}$ réalise le minimum des valeurs absolues non nulles de tous les coefficients $b'_{i,j}$ de tous les éléments de $\mathcal{A}$. Comme la matrice B' obtenue en ajouant un multiple de la première colonne à la colonne d'indice $j > 1$ de B appartient à $\mathcal{A}$, on voit en effectuant la division euclidienne de $b_{1,j}$ par $b_{1,1}$, puis une nouvelle inversion de colonnes sur B', que tous les coefficients $b_{1,j}$ ($j > 1$) de B sont nuls. De même, $b_{i,1} = 0$ pour tout $i > 1$.

ii) résulte de la formule de Lagrange: soient Y une matrice $q \times t$, Z une matrice $t \times p$, r un entier $\leq inf(p, q, t)$, K une injection croissante: $[1, r] \hookrightarrow [1, q]$, et L une injection croissante: $[1, r] \hookrightarrow [1, p]$; alors,

$$Det(YZ)_{KL} = \Sigma_H Det(Y_{KH}) Det(Z_{HL}),$$

où la sommation porte sur l'ensemble des injections croissantes $H : [1, r] \hookrightarrow [1, t]$.

Nous décrivons maintenant une forme plus simple de réduction des matrices, la mise sous *forme normale d'Hermite*, qui suffit pour de nombreux problèmes: soit B une matrice $m \times n$, à coefficients dans $\mathbf{Z}$, qu'on suppose, pour alléger les énoncés, de rang m (de sorte que $m \leq n$). On dit qu'elle est sous forme d'Hermite si ses $n - m$ dernières colonnes sont nulles, tandis que ses m premières colonnes forment une matrice triangulaire supérieure vérifiant: $\forall 1 \leq i < j$, $0 \leq b_{i,j} < b_{i,i}$.

Théorème 2 : *Soit A une matrice $m \times n$ de rang m, à coefficients dans $\mathbf{Z}$. Alors, il existe une unique matrice B sous forme d'Hermite telle que $B = AU$, où $U \in GL_n(Z)$.*

Démonstration: voir cours d'algèbre du Master. [NB: on notera que les coefficients diagonaux de B ne sont en général pas ceux de la matrice de Smith associée à A.]

Application: soit $A, B = AU$ comme supra. Une base sur $\mathbf{Z}$ du noyau de l'application $\mathbf{Z}$-linéaire attachée à $A : \mathbf{Z}^n \to \mathbf{Z}^m$ est donnée par les $n - m$ dernières colonnes de U.

§2. Géométrie des nombres.

Soit V un espace vectoriel sur $\mathbf{R}$, de dimension n finie. On appelle *réseau* de V tout sous-groupe de V engendré comme $\mathbf{Z}$-module par une base de V sur $\mathbf{R}$. Un réseau L de V est donc discret, et le quotient V/L est compact. Inversement:

Proposition 2 : *Tout sous-groupe discret M de V est un réseau du $\mathbf{R}$-sous-espace vectoriel qu'il engendre dans V. En particulier, le groupe topologique V/M est isomorphe à $(\mathbf{R}/\mathbf{Z})^r \times \mathbf{R}^{n-r}$, où r désigne le rang de M sur $\mathbf{Z}$.*

Démonstration: soient $\{e_1, ..., e_r\}$ un système maximal d'éléments de M linéairement indépendants sur $\mathbf{R}$, W le sous-$\mathbf{R}$-espace vectoriel de V qu'ils engendrent, et $\overline{P}$ la partie compacte $\{\Sigma_{i=1,...,r} x_i e_i, 0 \leq x_i \leq 1\}$ de V. Alors, $\overline{P} \cap M$ est fini (car M est discret), engendre M sur $\mathbf{Z}$ et est contenu dans l'espace vectoriel engendré sur $\mathbf{Q}$ par les e_i. Il existe donc un entier $d > 0$ tel que M soit contenu dans le $\mathbf{Z}$-module libre de rang r engendré par $\frac{1}{d}e_1, ..., \frac{1}{d}e_r$. Le théorème 1.1 entraîne alors que M est engendré sur $\mathbf{Z}$ par r éléments, qui forment encore une base de W sur $\mathbf{R}$.

Fixons un élément de volume sur V, d'où une mesure invariante par translation μ sur V, et soit L un réseau de V. Pour toute base $\{b_1, ..., b_n\}$ de L sur $\mathbf{Z}$, considérons le parallèlépipède $P = \{\Sigma_{i=1,...,n}\ x_i b_i\ ,\ 0 \leq x_i < 1\}$; sa mesure est indépendante de la base choisie, et s'appelle le covolume $\mu(V/L)$ de L. En effet, pour $V = \mathbf{R}^n$, muni de la mesure de Lebesgue μ, $\mu(V/L)$ est donné par la valeur absolue du déterminant de la matrice B représentant $\{b_1, ..., b_n\}$ dans la base canonique de $\mathbf{R}^n$. On l'appelle dans ce cas *déterminant* de L, et on le note $d(L)$. En termes du produit scalaire usuel $(.)$ sur $\mathbf{R}^n$, on peut aussi exprimer $d(L)$ comme la racine carrée du déterminant de la matrice de Gram associée à la base choisie de L:

$$\mu(V/L) = d(L) = |Det(B)| = (det(B^t B))^{1/2} = (det((b_i.b_j)_{1 \leq i,j \leq n}))^{1/2}.$$

(Un changement de base de L sur $\mathbf{Z}$ et un changement de base orthogonale de $\mathbf{R}^n$ transforment B en UBV, où U est une matrice orthogonale, et $V \in GL_n\mathbf{Z}$; cela ne modifie pas $d(L)$.) Enfin, si L est contenu dans le réseau canonique $\mathbf{Z}^n$ de $\mathbf{R}^n$, on peut également voir $d(L)$ comme l'indice $[\mathbf{Z}^n : L]$ de L dans $\mathbf{Z}^n$.

Fixons d'autre part une norme N sur l'espace vectoriel V. Le but de la géométrie des nombres est d'évaluer les éléments de petite norme des réseaux de V. Dans le cas le plus général, N est la jauge d'un corps convexe symétrique par rapport à 0, et on a:

Théorème 2 (Minkowski): *Soient μ une mesure de Haar sur V, $\mathbf{K}$ une partie convexe bornée symétrique par rapport à l'origine de V, et L un réseau de V. On suppose que $\mu(\mathbf{K}) > 2^n \mu(V/L)$. Alors, $\mathbf{K}$ contient un élément non nul de L.*

Démonstration: montrons tout d'abord que si L est un réseau de $\mathbf{R}^n$, muni de la mesure de Lebesgue μ, et si S est une partie μ-intégrable de $\mathbf{R}^n$ telle que $\mu(S) > d(L)$, il existe deux éléments $x \neq y$ de S tels que $x - y \in L$. En effet, soit P un parallèlépipède représentant $\mathbf{R}^n/L$ comme ci-dessus, de sorte que $\mu(P) = d(L)$ et que S est la réunion disjointe des parties $S \cap (h + P)$, où h parcourt L. Supposons que les parties $(-h + S) \cap P$ soient disjointes. Alors,

$$\mu(S) = \Sigma_{h \in L}\ \mu(S \cap (h + P)) = \Sigma_{h \in L}\ \mu((-h + S) \cap P) < \mu(P),$$

ce qui contredit l'hypothèse . Il existe donc $h \neq h' \in L$ et $x, y \in S$ tels que $x - y = h - h' \in L \setminus \{0\}$.

Le théorème 2 s'en déduit en considérant la partie intégrable $S = \frac{1}{2}\mathbf{K}$, qui vérifie par hypothèse $\mu(S) = (\frac{1}{2})^n \mu(\mathbf{K}) > d(L)$. L'élément non nul $z = \frac{1}{2}(2x - 2y)$ de $\mathbf{K} \cap L$ répond alors à la question.

Remarque: supposons $\mathbf{K}$ fermé. Pour tout $i = 1, \ldots, n$, soit λ_i le plus petit nombre réel tel que $L \cap \lambda_i \mathbf{K}$ contient i éléments linéairement indépendants sur $\mathbf{Z}$ (ou sur $\mathbf{R}$ – cela revient au même, puisque L est un réseau). Le théorème 2 équivaut à la majoration: $\lambda_1^n \mu(\mathbf{K}) \leq 2^n \mu(V/L)$. Plus généralement, le *deuxième théorème de Minkowski sur les minima successifs* énonce:

$$\lambda_1 \lambda_2 \ldots \lambda_n \mu(\mathbf{K}) \leq 2^n \mu(V/L).$$

On montre également que $\lambda_1 \lambda_2 \ldots \lambda_n \mu(\mathbf{K}) \geq \frac{2^n}{n!} \mu(V/L)$.

Applications:

i) tout nombre premier $\equiv 1 \bmod 4$ est somme de deux carrés; tout entier positif est somme de 4 carrés.

ii) Supposons que $\mathbf{K}$ soit la boule unité de $\mathbf{R}^n$, muni de la mesure de Lebesgue, de sorte que $\mu(\mathbf{K}) = \frac{\pi^{\frac{n}{2}}}{\Gamma(1+\frac{n}{2})}$. Le théorème 2 entraîne que le carré de la norme euclidienne du plus petit élément non nul de L est majoré par $\gamma_n d(L)^{\frac{2}{n}}$, où γ_n vaut au plus $\frac{4}{\pi}\Gamma(1 + \frac{n}{2})^{\frac{2}{n}}$.

La meilleure valeur possible de γ_n (valable pour tout réseau L de $\mathbf{R}^n$) s'appelle la n-ième constante d'Hermite. La détermination de γ_n est un problème ouvert, mais on en connait les premières valeurs: $\gamma_1 = 1, \gamma_2^2 = \frac{4}{3}, \gamma_3^3 = 2, ..., \gamma_8^8 = 256$. (Et γ_n vaut au plus $(\frac{4}{3})^{\frac{n-1}{2}}$.)

iii) Pour les applications à l'arithmétique des corps de nombres, voir le chapitre 5.

Les algorithmes concernent (comme dans l'application (ii) ci-dessus), le cas où la norme $N = \|.\|$ de V est la racine carrée d'une forme quadratique définie positive, c'est-à-dire où V est une espace euclidien, de produit scalaire (.). Rappelons tout d'abord le procédé d'orthogonalisation de Gram-Schmidt.

Proposition 3 : *Soit $\{b_1, ..., b_n\}$ une base d'un espace euclidien V. Définissons par récurrence sur $i = 1, ..., n$, les vecteurs $b'_i = b_i - \Sigma_{j=1,...,i-1}\mu_{i,j}b'_j$, où $\mu_{i,j} = \frac{(b_i \cdot b'_j)}{(b'_j \cdot b'_j)}$. Pour tout $i = 1, ..., n$, b'_i est la proj. orthogonale de b_i sur l'orthogonal de $\oplus_{j=1,...,i-1}\mathbf{R}b_j = \oplus_{j=1,...,i-1}\mathbf{R}b'_j$ dans $\oplus_{j=1,...,i}\mathbf{R}b_j$, et $\{b'_1, ..., b'_n\}$ est une base orthogonale de V.*

Le corollaire suivant fournit, à l'instar de l'inégalité inverse du deuxième théorème de Minkowski, une minoration du produit des minima successifs de la norme $\|.\|$ sur un réseau de V.

Corollaire (Hadamard) : *Pour tout réseau L de l'espace euclidien V, et toute famille $B = \{b_1, ..., b_n\}$ d'éléments de L linéairement indépendants sur $\mathbf{Z}$: $d(L) \leq \Pi_{i=1,...,n}\|b_i\|$, et ces expressions sont égales si et seulement si B est à la fois une base de L et une base orthogonale de V.*

On s'attend donc à ce que les bases de L formées de 'petits' vecteurs soient presque orthogonales. C'est ce que formalise la notion suivante. [**NB**: la fin de ce §2 ne fait pas partie du programme du cours.]

Définition: soit L un réseau de l'espace euclidien V. On dit qu'une base $\{b_1, ..., b_n\}$ de L est *reduite sous forme LLL* si l'on a, avec les notations du procédé de Gram-Schmidt: $|\mu_{i,j}| \leq \frac{1}{2}$ pour tout $1 \leq j < i \leq n$, et, pour tout $i = 2, ..., n$,

$$\|b'_i + \mu_{i,i-1}b'_{i-1}\|^2 \geq \frac{3}{4}\|b'_{i-1}\|^2,$$

ou de façon équivalente:

$$\|b'_i\|^2 \geq (\frac{3}{4} - \mu_{i,i-1}^2)\|b'_{i-1}\|^2.$$

[Noter que les vecteurs $b'_i + \mu_{i,i-1}b'_{i-1}$ et b'_{i-1} sont les projections orthoganales de b_i et de b'_{i-1} sur l'orthogonal de $\oplus_{j=1,....,i-2}\mathbf{R}b_j$.]

Théorème 3 : *Soit* $\{b_1, ..., b_n\}$ *une base réduite sous forme LLL d'un réseau* L *de* V. *Alors*

i) $d(L) \leq \Pi_{i=1,...,n}\|b_i\| \leq 2^{\frac{n(n-1)}{4}} d(L)$;

ii) $\forall\ 1 \leq j \leq i \leq n\ ,\ \|b_j\| \leq 2^{\frac{i-1}{2}}\|b'_i\|$, *de sorte que* $\|b_1\| \leq 2^{\frac{n-1}{4}} d(L)^{1/n}$;

iii) $\forall\ x \in L \setminus \{0\}\ ,\ \|b_1\| \leq 2^{\frac{n-1}{2}}\|x\|$; *plus généralement, pour tout t-uplet* $\{x_1, ...x_t\}$ *d'éléments de* L *lin. indép. sur* $\mathbf{Z}$, *et tout* $j \leq t$, *on a* $\|b_j\| \leq 2^{\frac{n-1}{2}} max(\|x_1\|, ..., \|x_t\|)$.

On trouvera dans les livres de H. Cohen (*A course in computational algebraic number theory*) et de M. Mignotte (*Mathématiques pour le calcul formel*) un algorithme de construction d'une base *LLL* d'un réseau arbitraire L de $\mathbf{R}^n$. Le fait qu'il converge (et par conséquent, que *tout réseau admet des bases réduites sous forme LLL*) repose sur le théorème de Minkowski (sous la version donnée dans sa troisième application). On notera qu'en vertu du Théorème 3.v, l'algorithme *LLL* permet de trouver, si ce n'est les minima successifs du réseau L, du moins des familles d'éléments de L qui ne sont pas loin de les réaliser.

Application: étant donnés n nombres réels non nuls $z_1, ..., z_n$, l'algorithme *LLL* permet souvent de déterminer s'ils sont linéairement dépendants sur $\mathbf{Z}$, et de trouver alors une relation $a_1z_1 + ... + a_nz_n = 0$ à coefficients $(a_1, ..., a_n) = a \in \mathbf{Z}^n$ les liant. Pour cela, on considère, pour tout entier $M > 0$, la forme quadratique $Q_M(a) = a_2^2 + a_3^2 + ... + a_n^2 + M(a_1z_1 + ... + a_nz_n)^2$. Elle est définie positive, et munit donc $\mathbf{R}^n$ d'une structure euclidienne. Si M est très grand, les éléments a du réseau $\mathbf{Z}^n$ tels que $Q_M(a)$ est assez petit sont candidats à fournir de telles relations.

§3 Algorithmes pour les polynômes.

L'algorithme *LLL* a été introduit en vue de factoriser les polynômes $A(X)$ à coefficients dans $\mathbf{Z}$ (cf. A. Lenstra, H. Lenstra, L. Lovász , *Factoring polynomials with rational coefficients*, Math. Ann., 261, 1982, 515-534). Voici une autre méthode, fondée sur l'algorithme de factorisation des polynômes sur les corps finis $\mathbf{F}_p$ donné par la Proposition 4 du Chap. II (Berlekamp).

De $\mathbf{F}_p[X]$ *(pour un ou plusieurs p) à* $\mathbf{Z}[X]$.

Un exemple: un polynôme $A(X)$ de degré 4 qui se décompose modulo p_1 (resp. p_2) en produits de deux facteurs irréductibles de degrés 2 (resp. de degrés 1 et 3) est irréductible. Mais on ne trouve pas forcément de tels couples: par exemple, on déduit de la loi de réciprocité quadratique que le polynôme $\Phi_8(X) = X^4 + 1$ se décompose modulo p en

produit de 4 facteurs de degré 1 si $p \equiv 1$ mod. 8 (ou si $p = 2$), et en produit de 2 facteurs irréductibles de degré 2 sinon. On ne peut en déduire l'irréductibilité de Φ_8 dans $\mathbf{Z}[X]$.

Dans la pratique, on combine les informations données par réduction modulo p avec la majoration suivante des valeurs absolues des facteurs eventuels de A.

Soient $A(X) = \Sigma_{i=0,...,n} a_i X^i$, $B(X) = \Sigma_{j=0,...,m} b_j X^j$ *deux polynômes non nuls à coefficients dans* $\mathbf{Z}$, *tels que* B *divise* A. *Pour tout* $j = 0, ..., m$,

$$|b_j| \leq \binom{n-1}{j}(\Sigma_{i=0,...,n}|a_i|^2)^{\frac{1}{2}} + \binom{n-1}{j-1}|a_n|.$$

Pour d'autres algorithmes, et une étude comparative de leurs performances, voir les livres de H. Cohen et de M. Mignotte cités plus haut. Plutôt que la considération de plusieurs nombres premiers, il est ainsi souvent plus efficace de tâcher de relever aux anneaux (non intègres) $(\mathbf{Z}/p^s\mathbf{Z})[X]$, avec s suffisamment grand, la décomposition donnée par l'algorithme de Berlekamp en un seul nombre premier p.

De $\mathbf{Z}[X]$ *à* $\mathbf{Q}[X]$ (cf. chap. I, §2).

Pour tout polynôme non nul $A \in \mathbf{Z}[X]$, on rappelle que le contenu de A, noté $cont(A)$, est le pgcd des coefficients de A.

Théorème 4 (lemme de Gauss) : *i) soient* P *et* Q *deux polynômes à coefficients dans* $\mathbf{Z}$; *alors,* $cont(PQ) = cont(P)cont(Q)$;

ii) soient $F \in \mathbf{Z}[X]$ *de contenu* 1, *et* $P \in \mathbf{Q}[X], Q \in \mathbf{Q}[X]$ *tels que* $F = PQ$. *Alors, il existe* $\lambda \in \mathbf{Q}, \lambda \neq 0$ *tel que* λP *et* $\lambda^{-1}Q$ *soient des éléments de* $\mathbf{Z}[X]$ *de contenu* 1;

iii) en particulier, les éléments irréductibles de l'anneau $\mathbf{Z}[X]$ *sont les nombres premiers (au signe près) et les polynômes de contenu* 1 *irréductibles dans* $\mathbf{Q}[X]$.

iv) $\mathbf{Z}[X]$ *est factoriel; plus généralement, pour tout anneau factoriel* A, *l'anneau* $A[X]$ *est factoriel.*

(Par conséquent, un polynôme unitaire de $\mathbf{Z}[X]$ est irréductible dans $\mathbf{Z}[X]$ si et seulement s'il est irréductible dans $\mathbf{Q}[X]$; et tout algorithme de factorisation dans $\mathbf{Z}[X]$ fournit un algorithme de factorisation dans $\mathbf{Q}[X]$.)

Démonstration: i) Il suffit de prouver la relation lorsque $cont(P) = cont(Q) = 1$. Si $cont(PQ)$ était > 1, il existerait un nombre premier p divisant tous les coefficients de PQ. Si π désigne l'homomorphisme de réduction modulo p, on aurait donc $\pi(P)\pi(Q) = \pi(PQ) = 0$ dans $\mathbf{F}_p[X]$. Mais l'anneau des polynômes à coefficients dans un corps (ou, plus généralement, dans un anneau intègre) est intègre. Contradiction.

ii) soient u (resp. v) un dénominateur commun des coefficients de P (resp. de Q). Alors, uP (resp. vQ) est un élément de $\mathbf{Z}[X]$, dont on note u' (resp. v') le contenu. D'après i), $uvPQ$ a pour contenu $u'v'$. Mais $cont(uvF) = uvcont(F) = uv$, donc $uv = u'v'$, et $\lambda = u/u'$ convient.

iii) tout $F \in \mathbf{Z}[X]$ s'écrit dans $\mathbf{Z}[X]$ sous la forme $cont(F)F'$, où $cont(F') = 1$; et on vient de voir qu'un élément de $\mathbf{Z}[X]$ de contenu 1 est irréductible dans $\mathbf{Z}[X]$ si et seulement s'il l'est dans $\mathbf{Q}[X]$.

iv) pour tout $F \in \mathbf{Z}[X]$, l'existence d'une décomposition en irréductibles de F découle de iii); son unicité du fait que $\mathbf{Z}$ et $\mathbf{Q}[X]$ sont tous deux factoriels. Même démonstration pour un anneau factoriel A quelconque, en remplaçant $\mathbf{Q}$ par le corps des fractions de l'anneau (intègre) A.

Pour conclure, donnons la preuve du critère d'irréductibilité d'Eisentein:

Proposition 5 (critère d'Eisenstein): *soient $A(X) = \Sigma_{i=0,...,n} a_i X^i$ un élément de $\mathbf{Z}[X]$, et p un nombre premier. On suppose que p divise $a_0, a_1, ..., a_{n-1}$ mais ne divise pas a_n, et que p^2 ne divise pas a_0. Alors, P est irréductible dans $\mathbf{Q}[X]$.*

Démonstration: soit $B(X) = \Sigma_{i=0,...,r} b_i X^i$ un facteur non constant de P dans $\mathbf{Z}[X]$. Alors, $p \not| b_r$ et on peut supposer que $p|b_0$. En considérant le plus grand entier t tel que p divise $b_0, ..., b_t$, on voit que $n = t + 1$, d'où $r = n$.

.

CHAPITRE V

ARITHMÉTIQUE DES CORPS DE NOMBRES

§1. Anneaux d'entiers.

Théorème 1 : *soient B un anneau commutatif et intègre, A un sous-anneau de B, et x un élément de B. Les propriétés suivantes sont équivalentes.*

i) x vérifie une relation de dépendance intégrale sur A (c'est-à-dire : il existe un polynôme unitaire $P \in A[T]$ tel que $P(x) = 0$);

ii) $A[x]$ est un A-module de type fini;

iii) il existe un A-module de type fini $M \neq 0$ inclus dans B tel que $xM \subset M$.

Démonstration: *i)* $\Rightarrow$ *ii)* $\Rightarrow$ *iii)* est clair. Écrivons iii) au moyen d'un système générateur de M sur A: $xm_i = \Sigma_{j=1,...,t}\, a_{ij}m_j, i = 1, ..., t$. Par conséquent, $\Delta(x) = \det(xI_t - (a_{i,j}))$ annule la partie $M \neq 0$ de l'anneau intègre B, et $\Delta(x) = 0$ fournit la relation de dépendance intégrale i) recherchée.

Les éléments x de B vérifiant les propriétés équivalentes du thm 1 sont dits *entiers* sur A. Leur ensemble s'appelle la fermeture intégrale A' de A *dans* B; le thm. 1 montre que A' est un sous-anneau de B (pour $x, y \in A'$, considérer $M = A[x, y]$), et que $(A')' = A'$.

Soient A un anneau commutatif intègre, et K son corps de fractions. On dit que A est *intégralement clos* s'il coïncide avec sa fermeture intégrale dans K. Par exemple, tout anneau factoriel, donc tout anneau principal, est intégralement clos.

Soient A et K comme ci-dessus, L une extension algébrique de K de degré n fini, B la fermeture intégrale de A dans L, Ω une clôture algébrique de K, et S l'ensemble des K-homomorphismes de L dans Ω. Pour tout élément x de L, on note η_x l'endomorphisme du K-espace vectoriel L défini par $\{y \in L\} \mapsto \{\eta_x(y) := xy\}$. Le polynôme caractéristique de l'endomorphisme η_x, multiplié par $(-1)^n$, s'appelle le *polynôme caractéristique* de x relativement à l'extension L/K:

$$\mathrm{Car}_{x,L/K}(T) = \det(Tid_L - \eta_x) = T^n - a_{n-1}T^{n-1} + ... + (-1)^n a_0 \in K[T].$$

La trace et la norme de x relativement à L/K sont alors respectivement définies par

$$Tr_{L/K}(x) := a_{n-1} = Tr(\eta_x)\ ,\ N_{L/K}(x) := a_0 = \mathrm{Det}(\eta_x).$$

Proposition 1 : *soit x un élement de L, de degré d sur K (de sorte que d divise n). Alors,*

i) $\mathrm{Car}_{x,L/K}(T) = (\mathrm{Min}_{x,K}(T))^{\frac{n}{d}}$, *de sorte que* $Tr_{L/K}(x) = \frac{n}{d} Tr_{K(x)/K}(x)$, $N_{L/K}(x) = (N_{K(x)/K}(x))^{\frac{n}{d}}$;

ii) si A est intégralement clos, et si x est entier sur A, $Min_{x,K}(T) \in A[T]$, *de sorte que* $Tr_{L/K}(x)$ *et* $N_{L/K}(x)$ *appartiennent à A.*

iii) si L/K est séparable, $Tr_{L/K}(x) = \Sigma_{\sigma \in S}\sigma(x), N_{L/K}(x) = \Pi_{\sigma \in S}\sigma(x)$.

Démonstration: i) si $n = d$, la matrice représentative de η_x dans la base cyclique $1, x = \eta_x(1), ..., x^{d-1} = \eta_x^{d-1}(1)$ est une matrice companion H, dont la dernière colonne est au signe près formée des coefficients de $Min_{x,K}$, et le résultat est bien connu. Dans le cas général, fixons une base $\omega_1, ..., \omega_{n/d}$ de L sur $K(x)$. Dans la base $\omega_j x^i; i = 0, ..., d-1, j = 1, ..., n/d$ de L sur K, la matrice représentative de η_x est composée de blocs diagonaux égaux à H, et son polynôme caractéristique est la puissance $\frac{n}{d}$-ième de celui de H.

ii, iii) Les coefficients de $Min_{x,K}$ sont les fonctions symétriques élémentaires de ses racines, et sont donc à la fois dans K et entiers sur A.

L'application $N_{L/K}$ est un homomorphisme du groupe multiplicatif L^* dans K^*. L'application $Tr_{L/K}$ est une forme K-linéaire sur L, non identiquement nulle si L/K est séparable. Pour simplifier, on suppose désormais que K est de *caractéristique nulle*. Alors, $Tr_{L/K}$ fournit un isomorphisme de K-espaces vectoriels de L vers $Hom_{K-\mathrm{lin}}(L, K)$, et pour toute base $\{\omega_1, \cdots, \omega_n\}$ de L sur K, il existe une unique base $\{\omega'_1, \cdots, \omega'_n\}$ de L sur K telle que $\forall\ i, j \in [1, n]$, $Tr_{L/K}(\omega_i \omega'_j) = \delta_{i,j}$ (symbole de Kronecker).

Définition : soit $\mathbf{B} = \{\omega_1, \cdots, \omega_n\}$ une base de L sur K. On appelle discriminant de $\mathbf{B}$ l'élément

$$disc(\mathbf{B}) = \mathrm{Det}(Tr_{L/K}(\omega_i \omega_j))_{1 \leq i,j \leq n}$$

de K . Comme cette matrice représente une forme bilinéaire non dégénérée, il est non nul.

Proposition 2 : *soient $\mathbf{B}, \mathbf{B}'$ deux bases de L sur K, M, M' les A-modules libres engendrés respectivement par $\mathbf{B}, \mathbf{B}'$ dans L. Alors:*

i) $disc(\mathbf{B}') = disc(\mathbf{B})(\mathrm{Det}P)^2$, *où P désigne la matrice de passage de $\mathbf{B}$ à $\mathbf{B}'$; en particulier,* $disc(\mathbf{B})$ *ne dépend, aux carrés des unités de A près, que de M;*

ii) $disc(\mathbf{B}) = [\mathrm{Det}(\sigma(\omega_i))_{\sigma \in S, 1 \leq i \leq n}]^2$; *en particulier, si* $\mathbf{B} = \{1, x, \cdots, x^{n-1}\}$ *pour un $x \in L$ de polynôme minimal P sur K, alors* $disc(\mathbf{B}) = (-1)^{\frac{n(n-1)}{2}} N_{L/K}(P'(x)) = Disc(P)$;

(iii) on suppose que les éléments de $\mathbf{B}$ appartiennent à l'anneau B, et que A est intégralement clos. Alors, $disc(\mathbf{B})$ *appartient à A. De plus, si* $\mathbf{B}' \subset M$, *et si* $disc(\mathbf{B}')$ *est sans facteur carré dans A, alors $M' = M$, et $\mathbf{B}'$ est une base de M sur A.*

Démonstration: i) et iii) sont clairs. Pour ii), considérer le produit à gauche de la matrice $(\sigma(\omega_i))_{\sigma,i}$ par sa transposée; si $\mathbf{B} = (x^i), i = 0, ..., n-1$, on obtient un déterminant de van der Monde, égal au signe près au produit des différences des racines de P, c'est-à-dire à son discriminant.

Le théorème suivant montre que quand $A = \mathbf{Z}$, l'anneau B lui-même est un A-module libre de type fini, de rang n.

Théorème 2 : *on suppose que* $A = \mathbf{Z}$, $K = \mathbf{Q}$. *Il existe une base* $\mathbf{B}$ *de* $L/\mathbf{Q}$ *formée d'éléments de* B, *et engendrant* B *sur* $\mathbf{Z}$.

Démonstration: notons tout d'abord que pour tout $x \in L$, il existe un entier $d \neq 0$ tel que dx soit entier sur $\mathbf{Z}$: par exemple, le *ppcm* des dénominateurs des coefficients de $Min_{x,\mathbf{Q}}$ convient. L'ensemble $\mathcal{B}$ des bases de L sur $\mathbf{Q}$ formées d'éléments de B est donc non vide. Leurs discriminants sont des entiers rationnels non nuls, et on peut en choisir une, soit $\mathbf{B}$, de discriminant D minimal en valeur absolue. Alors, les coordonnées de tout élément de B dans $\mathbf{B}$ sont entières, sans quoi on pourrait former, en considérant la partie fractionnaire de l'une d'elles, une base dans $\mathcal{B}$ de discriminant $< D$ en valeur absolue.

Terminologie des corps de nombres.

Dans toute la suite du chapitre, on considère un corps de nombres K, c'est-à-dire une extension de $\mathbf{Q}$, de degré fini $[K : \mathbf{Q}] = n$. On note $S_K = \{\sigma_1, \cdots, \sigma_n\}$ l'ensemble des plongements de K dans le corps $\mathbf{C}$ des nombres complexes, et $\mathbf{O}_K$ la fermeture intégrale de $\mathbf{Z}$ dans K, appelée *anneau des entiers* du corps de nombres K. D'après le théorème 2, il existe une base $\{\omega_1, \cdots, \omega_n\}$ de $\mathbf{O}_K$ sur $\mathbf{Z}$, et puisque les deux unités de l'anneau $\mathbf{Z}$ sont de carré égal à 1, toutes les bases de $\mathbf{O}_K$ sur $\mathbf{Z}$ ont le même discriminant D_K, appelé *discriminant de* K. On note en général U_K le groupe $\mathbf{O}_K^*$ des unités de l'anneau $\mathbf{O}_K$, et on supprime les indices K lorsqu'un seul corps de nombres est en jeu. Par abus de langage, les ideaux et les unités de $\mathbf{O}_K$ sont parfois appelées ideaux et unités de K.

Exemple 1 : les corps *quadratiques* sont les extensions de $\mathbf{Q}$ de degré 2. Ils sont de la forme $K = \mathbf{Q}(\sqrt{d})$, où $d \neq 1$ est un entier rationnel sans facteur carré. Si d est positif (resp. négatif), on dit que K est quadratique réel (resp. imaginaire).

Si $d \equiv 2$ ou 3 mod.4, $\mathbf{O}$ admet $\{1, \sqrt{d}\}$ pour base sur $\mathbf{Z}$, et $D = 4d$;

Si $d \equiv 1$ mod.4, $\mathbf{O}$ admet $\{1, \frac{1+\sqrt{d}}{2}\}$ pour base sur $\mathbf{Z}$, et $D = d$.

On reviendra en détail sur ce cas au chapitre VI.

Exemple 2 : les corps *cyclotomiques*. Ils sont de la forme $K_n = \mathbf{Q}(\zeta_n)$, où n est un entier positif, et $\zeta_n = e^{2i\pi/n}$. Le polynôme cyclotomique Φ_n (cf. chap. I et II) est irréductible sur

$\mathbf{Q}$, de sorte que $K_n/\mathbf{Q}$ est une extension galoisienne, de degré $\phi(n)$, et de groupe de Galois isomorphe à $(\mathbf{Z}/n\mathbf{Z})^*$. L'anneau des entiers $\mathbf{O}_n$ de K_n admet $\{1, \zeta_n, \zeta_n^2, ..., \zeta_n^{\phi(n)-1}\}$ pour base; autrement dit, $\mathbf{O}_n = \mathbf{Z}[\zeta_n] \simeq \mathbf{Z}[T]/\Phi_n(T)$. Pour $n = p \neq 2$ premier, son discriminant vaut $D_p = (-1)^{\frac{p-1}{2}} p^{p-2}$. Le corps K_p contient le corps quadratique $\mathbf{Q}(\sqrt{(-1)^{\frac{p-1}{2}} p})$ (cf. chap II), et aucun autre corps quadratique.

§2 Idéaux des corps de nombres.

Soit K un corps de nombres de degré n, d'anneau d'entiers $\mathbf{O}$. Un idéal $\mathbf{a}$ de $\mathbf{O}$ (on dira aussi: un idéal *entier*) est un $\mathbf{O}$-module contenu dans $\mathbf{O}$. Si $\mathbf{a}$ est non nul, il est, *en tant que* $\mathbf{Z}$*-module*, libre de rang n (d'après le thm. 2 ci-dessus, et le thm. 1.1 du chap. IV). En particulier, l'indice $N(\mathbf{a}) = [\mathbf{O} : \mathbf{a}]$, qu'on appelle la *norme* de $\mathbf{a}$, est *fini*. On en déduit que tout suite croissante d'idéaux de $\mathbf{O}$ est stationnaire (autrement dit, l'anneau $\mathbf{O}$ est *noethérien*). Comme un anneau intègre fini est un corps, on en déduit également:

Proposition 3 : *tout idéal premier non nul* $\mathbf{p}$ *de* $\mathbf{O}$ *est* maximal. *(Par ailleurs, il contient un unique nombre premier: le générateur de* $\mathbf{p} \cap \mathbf{Z} = (p)$*.)*

Les anneaux d'entiers de corps de nombres sont donc noethériens, intégralement clos, et leurs idéaux premiers non nuls sont maximaux. De tels anneaux sont appelés des *anneaux de Dedekind*. Des propriétés générales des anneaux de Dedekind (qu'on trouvera par exemple développées dans le livre de Samuel 'Théorie algébrique des nombres'), on tire les deux énoncés suivants, que nous admettrons.

Définition : soit $\mathbf{O}$ un anneau de Dedekind, de corps de fractions K. Un *idéal fractionnaire* de K est un $\mathbf{O}$-module $\mathbf{a}$ de type fini contenu dans K, ou, de façon équivalente, tel qu'il existe un élément $\gamma \neq 0$ de K pour lequel $\gamma\mathbf{a}$ soit un idéal de $\mathbf{O}$.

Soit $J_K = J$ l'ensemble des idéaux fractionnaires non nuls de K. Si $\mathbf{a}$, $\mathbf{b} \in J$, on pose $\mathbf{ab} := \{\Sigma_{i \in I \text{fini}} a_i b_i, a_i \in \mathbf{a}, b_i \in \mathbf{b}\}$; c'est encore un élément de J.

Théorème 3 : *pour* $\mathbf{O}$ *de Dedekind, ce produit munit* J *d'une structure de groupe abélien, d'élément neutre* $\mathbf{O}$*, où l'inverse d'un élément* $\mathbf{a}$ *de* J *est l'idéal fractionnaire*

$$\mathbf{a}^{-1} := \{x \in K, x\mathbf{a} \subset \mathbf{O}\}.$$

En particulier, pour $\mathbf{a}$ et $\mathbf{b}$ dans J, $\mathbf{b} \subset \mathbf{a}$ si et seulement s'il existe un idéal *entier* $\mathbf{c}$ tel que $\mathbf{b} = \mathbf{ac}$. On dit alors que $\mathbf{a}$ *divise* $\mathbf{b}$, et on note: $\mathbf{a} \mid \mathbf{b}$. Plus généralement, la propriété suivante remplace efficacement la non-factorialité éventuelle de $\mathbf{O}$.

Proposition 4: *soit* $\mathbf{O}$ *un anneau de Dedekind. Tout idéal entier (resp. fractionnaire) non nul et distinct de* $\mathbf{O}$ *admet une décomposition unique en produit de puissances* > 0 *(resp.* > 0 *ou* < 0*) d'idéaux premiers de* $\mathbf{O}$*.*

(Pour l'unicité, noter que si $\mathbf{p}$ est un idéal premier d'un anneau quelconque $\mathbf{O}$, et si $\mathbf{a}$, $\mathbf{b}$ sont des idéaux de $\mathbf{O}$ tels que $\mathbf{p} \supset \mathbf{ab}$, alors, $\mathbf{p}$ contient l'un au moins des idéaux $\mathbf{a}$, $\mathbf{b}$).

Pour $\mathbf{O}$ de Dedekind, on a donc:

$$\mathbf{a} \cap \mathbf{b} = \mathbf{ab} \Leftrightarrow \mathbf{a} + \mathbf{b} = \mathbf{O}.$$

(Rappelons que l'implication $\Rightarrow$ n'est pas correcte dans un anneau quelconque; l'implication $\Leftarrow$ est toujours vraie, cf. le lemme chinois.)

Revenons à l'anneau des entiers $\mathbf{O}$ d'un corps de nombres K, et à la norme $N(\mathbf{a}) = \mathrm{card}(\mathbf{O}/\mathbf{a})$ de ses idéaux $\mathbf{a}$ non nuls.

Proposition 5: *soit* $\mathbf{a}$ *et* $\mathbf{b}$ *deux idéaux non nuls de* $\mathbf{O}$*. Alors:*

i) $N(\mathbf{ab}) = N(\mathbf{a})N(\mathbf{b})$*;*

ii) si $\mathbf{b} \subset \mathbf{a}$*,* $N(\mathbf{a})$ *divise* $N(\mathbf{b})$ *et* $N(\mathbf{a}) = N(\mathbf{b})$ *si et seulement si* $\mathbf{a} = \mathbf{b}$*;*

iii) $N(\mathbf{a})$ *appartient à* $\mathbf{a}$*; en particulier, pour tout entier naturel* A*, il n'y a qu'un nombre fini d'idéaux de* $\mathbf{O}$ *de norme* A*;*

iv) soit $\{\omega_1, \cdots, \omega_n\}$ *une base de* $\mathbf{a}$ sur $\mathbf{Z}$*; alors* $N(\mathbf{a}) = \mid \frac{Disc(\{\omega_1,\cdots,\omega_n\}}{D_K} \mid^{\frac{1}{2}}$*; en particulier, si* $\mathbf{a} = \alpha\mathbf{O}$ *est un idéal principal,* $N(\mathbf{a}) = \mid N_{K/\mathbf{Q}}(\alpha) \mid$.

Démonstration: i) Grâce à la proposition 4 et au lemme chinois, il suffit de montrer que pour tout idéal premier $\mathbf{p}$, de norme q, et tout entier $m \geq 1$, $N(\mathbf{p}^m) = q^m$. Mais $\mathbf{p}^{m-1}/\mathbf{p}^m$ est naturellement muni d'une structure de module sur $\mathbf{O}/\mathbf{p}$, c'est-à-dire d'espace vectoriel sur le corps $\mathbf{F}_q$, et sa dimension est majorée son rang sur $\mathbf{O}$, donc par 1. Comme elle n'est pas nulle ($\mathbf{p}^{m-1} \neq \mathbf{p}^m$, cf. prop. 4), elle vaut 1. Ainsi, $[\mathbf{p}^{m-1} : \mathbf{p}^m] = q$, et on conclut par récurrence sur m.

ii) et iii) sont clairs ($N(\mathbf{a})$ annule le $\mathbf{Z}$-module $\mathbf{O}/\mathbf{a}$), et iv) résulte de la Prop. 2.

La multiplicativité de l'application 'norme' sur les idéaux entiers permet de l'étendre au groupe J_K tout entier, et montre qu'un idéal de $\mathbf{O}$ est premier dès que sa norme est un nombre premier. La réciproque est fausse, mais si $\mathbf{p}$ est un idéal de $\mathbf{O}$ premier, sa norme, cardinal du corps fini $\mathbf{O}/\mathbf{p}$, est une puissance pure p^f de l'unique nombre premier p que contient $\mathbf{p}$. On en déduit:

Proposition 6: *soient* p *un nombre premier,* $(p) = \mathbf{p}_1^{e_1} \cdots \mathbf{p}_r^{e_r}$ *la décomposition de l'idéal entier* $p\mathbf{O}$ *en produit de puissances d'idéaux premiers, et* p^{f_i} *la norme de* $\mathbf{p}_i$*. Alors, l'ensemble* $\{\mathbf{p}_1, \ldots, \mathbf{p}_r\}$ *coïncide avec l'ensemble des idéaux premiers de* $\mathbf{O}$ *contenant* p*, et* $n = \Sigma_{i=1,\ldots,r} e_i f_i$.

Pour tout i, le nombre e_i (resp. f_i) s'appelle l'*indice de ramification* (resp. le *degré résiduel*) de l'idéal premier $\mathbf{p}_i$ au-dessus de p. La proposition suivante permet de les calculer sous une hypothèse (assez restrictive) sur $\mathbf{O}$.

Proposition 7: *On* suppose *qu'il existe un élément θ de $\mathbf{O}$, de polynôme minimal $P \in \mathbf{Z}[T]$ sur $\mathbf{Q}$ (cf. Prop. 1.ii), tel que $\mathbf{O} = \mathbf{Z}[\theta]$, et on note $P_1, ..., P_s$ des élements de $\mathbf{Z}[T]$ tels que si π désigne la réduction modulo le nombre premier p, la décomposition de $\pi(P)$ en facteurs irréductibles dans $\mathbf{F}_p[T]$ soit donnée par $\pi(P) = \pi(P_1)^{e_1}...\pi(P_r)^{e_r}$. Alors, pour tout $i = 1, ..., r$, l'idéal $\mathbf{p}_i := (p, P_i(\theta))$ de $\mathbf{O}$ est premier, et la décomposition de $p\mathbf{O}$ en idéaux premiers de $\mathbf{O}$ est donnée par $(p) = \mathbf{p}_1^{e_1} \cdots \mathbf{p}_r^{e_r}$.*

Démonstration: si θ_i est une racine de $\pi(P_i)$ dans $\overline{\mathbf{F}}_p$ (de sorte que $\mathbf{F}_p(\theta_i) = \mathbf{F}_p[T]/(\pi(P_i))$), l'application $\pi_i : \mathbf{Z}[\theta] \to \mathbf{F}_p[\theta_i]$ définie par $Q(\theta) \mapsto \pi(Q)(\theta_i)$ a pour image un corps, donc son noyau $Ker(\pi_i)$ est un idéal premier de $\mathbf{O}$. Cet idéal contient clairement $\mathbf{p}_i$, et on vérifie aisément qu'il lui est égal. De plus, $\mathbf{p}_1^{e_1}...\mathbf{p}_r^{e_r}$ est contenu dans $(p, P_1(\theta)^{e_1}....P_r(\theta)^{e_r})$, donc dans $(p, P(\theta) + p\mathbf{O}) = (p)$. Ainsi, $\{\mathbf{p}_1, ..., \mathbf{p}_r\}$ est l'ensemble des idéaux premiers de $\mathbf{O}$ contenant p, et $(p) = \mathbf{p}_1^{e_1'}...\mathbf{p}_r^{e_r'}$ pour des entiers $e_i' \in [1, e_i]$. Or $p^{f_i} = [\mathbf{O} : \mathbf{p}_i] = |\mathbf{F}_p(\theta_i)| = p^{deg(\pi(P_i))}$. On conclut en combinant les relations $\Sigma_i e_i' f_i = n$ et $\Sigma_i e_i deg(\pi(P_i)) = deg(\pi(P)) = deg(P)$.

§3 Les théorèmes de finitude

On reprend les notations de la fin du §1 et du §2. On désigne de plus par s le nombre de plongements réels de K, c'est-à-dire d'éléments σ de S tels que $\sigma(K) \subset \mathbf{R}$. Le nombre de plongements non réels est pair, et noté $2t$, de sorte que $n = s + 2t$. On appelle *constante de Minkowski* de K le nombre

$$M_K := (\frac{4}{\pi})^t \frac{n!}{n^n} \sqrt{D_K}.$$

On estimera à titre d'exercice les constantes de Minkowski des corps quadratiques et des corps cyclotomiques $\mathbf{Q}(\zeta_p)$.

N'étant en général pas factoriels, les anneaux d'entiers des corps de nombres ne sont en général pas principaux . L'énoncé fondamental suivant permet de remédier en grande partie à cette difficulté. Soit Pr_K le sous-groupe de J_K formé par les idéaux fractionnaires *principaux*, c'est-à-dire de la forme $\mathbf{a} = \alpha\mathbf{O}$, pour un élément α de K^*. Le quotient $C\ell_K := J_K/Pr_K$ s'appelle le *groupe des classes* (d'idéaux) de K.

Théorème 4 : *le groupe $C\ell_K$ est un groupe (abélien)* fini. *Son ordre h_K s'appelle le* nombre de classes *du corps de nombres K. Ainsi, pout tout idéal $\mathbf{a}$ de $\mathbf{O}$, $\mathbf{a}^{h_K}$ est un idéal principal, et l'anneau $\mathbf{O}_K$ est principal si et seulement si $h_K = 1$.*

Ce résultat découle de l'énoncé suivant (plus précis, et crucial pour le calcul effectif de h_K), joint à la Prop. 5.ii.

Théorème 5: *tout élément de $C\ell_K$ admet parmi ses représentants dans J_K un idéal entier de norme inférieure ou égal à la constante de Minkowski M_K de K.*

Démonstration: considérons l'application

$$\mathcal{S} : K \to \mathbf{R}^s \times \mathbf{C}^t : x \mapsto (\sigma_i(x), i = 1, ..., s; \sigma_j(x), j = s+1, ..., s+t),$$

où les t derniers plongements représentent les différentes paires $(\sigma_j, \overline{\sigma_j})$ de plongements complexes conjugués. C'est un homomorphisme de groupes (additifs) injectif, appelé plongement canonique de K dans $\mathbf{R}^s \times \mathbf{C}^t \simeq \mathbf{R}^n$ (on fixe ce dernier isomorphisme au moyen de la décomposition $\mathbf{C} = \mathbf{R} \oplus i\mathbf{R}$). La Prop. 2 entraîne que l'image de $\mathbf{O}$ sous $\mathcal{S}$ est un réseau de $\mathbf{R}^n$ de covolume $2^{-t}|D_K|^{1/2}$. Pour tout idéal $\mathbf{a}$ de $\mathbf{O}$, $\mathcal{S}(\mathbf{a})$ est donc un réseau de covolume $2^{-t}|D_K|^{1/2}N(\mathbf{a})$. Considérons alors, pour tout nombre réel $c > 0$, la partie de $\mathbf{R}^s \times \mathbf{C}^t$, convexe et symétrique par rapport à l'origine, $B_c = \{(y_1, ..., y_s, z_{s+1}, ..., z_{s+t}), \Sigma_i|y_i| + 2\Sigma_j|z_j| \leq c\}$. Son volume valant $2^s(\frac{\pi}{2})^t \frac{c^n}{n!}$, le théorème de Minskowski entraîne que dès que $c^n > (\frac{4}{\pi})^t n!|D_K|^{1/2}N(\mathbf{a})$, il existe un élément non nul α de $\mathbf{a}$ tel que $\mathcal{S}(\alpha) \in B_c$. D'après l'inégalité des moyennes arithmétique et géométrique, un tel α vérifie:

$$|N_{K/\mathbf{Q}}(\alpha)| = \Pi_i|\sigma_i(\alpha)|\Pi_j|\sigma_j(\alpha)|^2 \leq [\frac{1}{n}(\Sigma_i|\sigma_i(\alpha)| + 2\Sigma_j|\sigma_j(\alpha)|)]^n \leq \frac{c^n}{n^n}.$$

Ainsi, *tout idéal entier non nul* $\mathbf{a}$ *contient un élément non nul de norme majorée en valeur absolue par* $(\frac{4}{\pi})^t \frac{n!}{n^n}|D_K|^{1/2}N(\mathbf{a}) = M_K N(\mathbf{a})$.

Dans ces conditions, soient C un élément de $C\ell_K$, $\mathbf{a}$ un idéal entier représentant la classe inverse C^{-1} dans J_K, et α un élément non nul de $\mathbf{a}$ de norme $\leq M_K N(\mathbf{a})$. Alors, l'idéal $\alpha\mathbf{a}^{-1}$ est entier, et représente C dans J_K. D'après la Prop. 5, sa norme vaut $|N_{K/\mathbf{Q}}(\alpha)|(N(\mathbf{a}))^{-1}$, qui est bien majoré par M_K.

Corollaire (théorème d'Hermite): $\mathbf{Q}$ *est le seul corps de nombres de discriminant 1 .*

Démonstration: les idéaux entiers non nuls étant de norme ≥ 1, le théorème 5 entraîne l'inégalité $|D_K| \geq \frac{n^{2n}}{(n!)^2}(\frac{\pi}{4})^{2t} \geq \frac{n^{2n}}{(n!)^2}(\frac{\pi}{4})^n$, qui est > 1 pour $n \geq 2$.

Le théorème de Minskowski fournit un autre théorème de finitude fondamental.

Théorème 6 (théorème de Dirichlet): *le groupe U_K des unités de $\mathbf{O}_K$ est le produit direct du groupe fini μ_K (formé par les racines de l'unité dans K) et d'un groupe libre de rang $s + t - 1$.*

En particulier, pour K quadratique réel, il existe une unité u de $\mathbf{O}$ telle que $U_K = \{\pm 1\} \times u^{\mathbf{Z}}$. La détermination d'une telle unité, dite fondamentale, se ramène à la résolution des classiques *équations de Pell-Fermat*. De façon générale, il est clair qu'un élément α de $\mathbf{O}$ est une unité si et seulement si $N_{K/\mathbf{Q}}(\alpha) = \pm 1$.

Le groupe μ_K est le sous-groupe de torsion du $\mathbf{Z}$-module U_K. On déduit facilement de l'irréductibilité des polynômes cyclotomiques qu'il est fini. La partie profonde du théorème est donc la détermination du rang du $\mathbf{Z}$-module libre U_K/μ_K.

Démonstration: considérons l'application

$$\mathcal{L} : U_K \to \mathbf{R}^s \times \mathbf{R}^t : x \mapsto (Log|\sigma_i(x)|, i = 1, ..., s; Log|\sigma_j(x)|, j = s+1, ..., s+t).$$

C'est un homomorphisme de groupes, appelé par abus de langage 'plongement' logarithmique de U_K, dont le noyau coïncide avec μ_K: il n'y a en effet qu'un nombre fini d'éléments de $\mathbf{O}$ dont les images par tous les plongements de K soient bornés, les coefficients de leurs poly. minimaux étant alors bornés. Comme la norme d'une unité vaut ± 1, l'image de U_K sous $\mathcal{L}$ est incluse dans l'hyperplan W de $\mathbf{R}^{s+t}$ d'équation $\Sigma_i x_i + 2\Sigma_j y_j = 0$. La remarque précédente entraîne de plus qu'elle est discrète. D'après IV, Prop. 2, $U_K/\mu_K \simeq \mathcal{L}(U_K)$ est donc un $\mathbf{Z}$-module libre de rang $r \leq s+t-1$. Reste à voir que $\mathcal{L}(U_K)$ engendre W sur $\mathbf{R}$.

Pour alléger l'exposé, nous supposerons que $t = 0$, i.e. $n = s$. Montrons que dès que $s > 1$, il existe une unité η dont tous les conjugués, sauf $\sigma_1(\eta)$, sont de val. absolue < 1 (on dit que η est un nombre de Pisot). Soient $0 < \delta_0 < 1$ et M_0 deux réels tels que $M_0 \delta_0^{n-1} = |D_K|^{1/2}$. D'après Minskowski, il existe $\alpha_0 \neq 0$ dans $\mathbf{O}$ tel que $\sigma_1(\alpha_0) \leq M_0$, $\sigma_i(\alpha_0) \leq \delta_0$ pour $i \geq 2$, et donc $|N(\alpha_0)| \leq |D_K|^{1/2}$. De proche en proche, on construit de même une suite d'entiers $\alpha_n \neq 0$ dans $\mathbf{O}$ tels que $|N(\alpha_n)| \leq |D_K|^{1/2}$ et $sup_{i \geq 2} |\sigma_i(\alpha_n)| < inf_{i \geq 2} |\sigma_i(\alpha_{n-1})| < 1$. Comme il n'y a (d'après la prop. 5 iii) qu'un nombre fini d'éléments de $\mathbf{O}$ de norme bornée non associés, l'un des quotients $\alpha_{n'}/\alpha_n, n' > n \geq 0$, est une unité η répondant à la question. En particulier, $Log|\sigma_i(\eta)| < 0$ pour tout $i \geq 2$.

Supposons que $r < s-1$, i.e. que $\mathcal{L}(U_K)$ soit inclus dans un hyperplan W' distinct de W. Après réindexation, on peut supposer son équation de la forme $\Sigma_i c_i x_i = 0$, avec $c_1 \geq c_i$ pour $i \geq 2$. En retranchant c_1 fois l'équation de W, on voit que $\mathcal{L}(\eta)$ n'appartient pas à W'; contradiction.

CHAPITRE VI

ALGORITHMES QUADRATIQUES

§1. Corps quadratiques.

Soit $d \neq 1$ un entier rationnel sans facteurs carrés. Rappelons (cf. chap. V) que le corps $K = \mathbf{Q}(\sqrt{d})$, admet $\{1, \omega\}$ pour base sur $\mathbf{Z}$ de son anneau d'entiers $\mathbf{O}_K$, avec

$\omega = \sqrt{d}$ si $d \equiv 2$ ou 3 mod.4 ; le discriminant de K vaut alors $D = 4d$;

$\omega = \frac{1+\sqrt{d}}{2}$ si $d \equiv 1$ mod.4; le discriminant de K vaut alors $D = d$.

Les entiers D apparaisant de cette façon s'appellent les *discriminants fondamentaux*. Ils sont congrus à 0 ou 1 mod. 4 . Inversément, tout entier $D \neq 1$ qui est soit $\equiv 1$ mod. 4 et sans facteur carré, soit divisible par 4 et tel que $D/4$ soit $\equiv 2$ ou 3 mod. 4 et sans facteur carré, est un discriminant fondamental. On notera $C_D = C\ell_K$ et $h(D) = h_K$ le groupe de classes et le nombre de classes du corps quadratique K correspondant.

Suivant le type de décomposition de l'idéal (p) qu'il engendre dans $\mathbf{O}_K$, un nombre premier p est dit:

inerte si $(p) = \mathbf{p}$ (autrement dit, si (p) est encore premier);

décomposé si $(p) = \mathbf{pp}'$, où $\mathbf{p}' \neq \mathbf{p}$; dans ce cas, $\mathbf{p}'$ est l'image de $\mathbf{p}$ par l'automorphisme non identique de K, et les classes dans C_D de ces idéaux sont inverses l'une de l'autre;

ramifié si $(p) = \mathbf{p}^2$; dans ce cas, la classe de $\mathbf{p}$ dans C_D est d'ordre 1 ou 2.

On déduit de la proposition 7 du chap. IV, en notant $(\frac{\cdot}{p})$ le symbole de Legendre:

Proposition 1 : *i) Un nombre premier p impair est inerte si $(\frac{d}{p}) = -1$, décomposé si $(\frac{d}{p}) = 1$, et ramifié si p divise d.*

ii) Le nombre premier $p = 2$ est inerte si $d \equiv 5$ mod. 8, décomposé si $d \equiv 1$ mod. 8, et ramifié dans les autres cas.

En particulier, p est ramifié dans K si et seulement si p divise le discriminant D de K. C'est là d'ailleurs une propriété générale des corps de nombres.

L'algorithme de factorisation des entiers dont nous décrivons le principe au §3 repose sur l'énoncé suivant.

Théorème 1 : *Soient $D < 0$ un discriminant fondamental négatif, et t le nombre de facteurs premiers de D distincts. Alors,*

i) le groupe C_D/C_D^2 est isomorphe à $(\mathbf{Z}/2\mathbf{Z})^{t-1}$. En particulier, 2^{t-1} divise le nombre de classes $h(D)$, et $h(D)$ est impair si et seulement si $D = -4, -8$, ou l'opposé d'un nombre premier $\equiv 3$ mod. 4;

ii) plus concrètement, le sous-groupe $C_D[2]$ de C_D formé par les éléments d'ordre 1 ou 2 est engendré par les classes des idéaux premiers divisant D.

[Un énoncé similaire vaut quand $D > 0$, sous réserve de considérer les classes au sens 'restreint'. Sans cette modification, on obtient seulement: $2^{t-2}|h(D)$.]

Nous montrerons seulement que les classes considérées dans ii) engendrent un sous-groupe isomorphe à $(\mathbf{Z}/2\mathbf{Z})^{t-1}$. On se ramène facilement au cas où $D \leq -8$, pour lequel $\mathbf{O}_K^*$ $(= \mu_K$ puisque $D < 0)$ est réduit à $\{\pm 1\}$. Soient $p_1, ..., p_t$ les nombres premiers divisant D, $\mathbf{p_i}$ les idéaux premiers correspondant, et $\mathbf{a}$ le produit de $s \in [1, t]$ d'entre eux. Alors $\mathbf{a}^2 = (p_{i_1}...p_{i_s})$ est principal, mais $\mathbf{a}$ ne peut être principal que s'il existe $\zeta = \pm 1 \in \mathbf{O}_K^*$ tel que $\zeta p_{i_1}...p_{i_s}$ soit un carré dans K^*. Cela se produit si et seulement si $-dp_{i_1}...p_{i_s}$ est un carré dans $\mathbf{Q}^*$, c'est-à-dire ssi $s = t$ lorsque $d \equiv 1$ ou 2 mod. 4 (resp. $s = t-1$ et $2 \neq p_{i_1}, ..., p_{i_s}$ lorsque $d \equiv 3$ mod. 4). Dans chaque cas, les classes de $\mathbf{p}_1, ..., \mathbf{p}_t$ engendrent donc un sous-espace vectoriel de dimension $t-1$ du $\mathbf{F}_2$-espace vectoriel $C_D[2]$. [Contre-exemple dans le cas réel: la classe modulo $\mathbf{O}_K^*$ de $2 = (2-\sqrt{3})(1+\sqrt{3})^2$ est un carré dans $(\mathbf{Q}(\sqrt{3}))^*$, et l'idéal premier $\mathbf{p} = (1+\sqrt{3})$ qui divise 2 est principal.]

Le théorème 1 établit une correspondance entre des diviseurs de D et les éléments d'ordre 2 (appelées traditionellement classes *ambiges*) du groupe de classes C_D. Dans le paragraphe suivant, on montre comment représenter les éléments de C_D par des formes quadratiques de discriminant D, plutôt que par les idéaux de $\mathbf{O}_K$. Les classes ambiges y sont plus facilement reconnaissables, et le passage des formes qui les représentent à une factorisation de D est immédiat (cf. §3). Ce dictionnaire fournit en prime (cf. fin du §2) une majoration très précise du nombre de classes $h(D)$.

§2. Formes quadratiques.

Définition: soient $f(x,y) = ax^2 + bxy + cy^2, f'(x,y) = a'x^2 + b'xy + c'y^2$ deux formes quadratiques binaires à coefficients dans $\mathbf{Z}$. On dit qu'elles sont équivalentes s'il existe un élement $U = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \in SL_2(\mathbf{Z})$ (= sous-groupe de $GL_2(\mathbf{Z})$ formé des matrices de déterminant +1) tel que

$$f(U(x,y)) := f(\alpha x + \beta y, \gamma x + \delta y) = f'(x,y).$$

On vérifie que c'est bien une relation d'équivalence sur l'ensemble des formes quadratiques.

Pour tout entier n, l'étude de l'équation diophantienne $f(x,y) = n$ équivaut alors à celle de $f'(x,y) = n$.

Soit $D = b^2 - 4ac$ le *discriminant* de f. Alors, f' a même discriminant; de même, les pgcd de (a,b,c) et de (a',b',c') coïncident (et on dit que f est *primitive* si ce pgcd vaut 1); enfin, le signe de a' est celui de a si $D < 0$ (et est alors > 0 si et seulement si f est définie positive). Pour tout entier D, on peut donc considérer l'ensemble C'_D des classes d'équivalence de formes quadratiques primitives de discriminant D, auxquelles on demande de plus, si $D < 0$, d'être définies positives. L'ensemble C'_D, dont on note $h'(D)$ le cardinal, est non vide si et seulement si $D \equiv 0$ ou 1 modulo 4 (pour l'existence, considérer respectivement les formes primitives 'fondamentales' $x^2 - \frac{D}{4}y^2$ et $x^2 + xy + \frac{1-D}{4}y^2$). Par ailleurs, si D est un discriminant fondamental au sens du §1, toute forme f représentant un élément de C'_D est primitive.

Proposition 2: *soient $D < 0$ un discriminant fondamental et K le corps quadratique imaginaire $\mathbf{Q}(\sqrt{D})$. Il existe une bijection Ξ, explicitement décrite ci-dessous, entre l'ensemble C'_D des classes d'équivalence de formes quadratiques primitives définies positives de discriminant D, et l'ensemble C_D des classes d'idéaux fractionnaires de K. En particulier, les cardinaux $h(D)$ et $h'(D)$ de ces ensembles coïncident, et l'on peut par transport de structure munir C'_D d'une structure de groupe.*

Remarques: i) Un énoncé similaire vaut quand $D > 0$, sous réserve de considérer les classes d'idéaux au sens 'restreint' du corps quadratique réel K

ii) Soit Φ_D l'ensemble des formes quadratiques primitives de discriminant D (et définies positives si $D < 0$). On peut, suivant Gauss, munir l'ensemble Φ_D lui-même d'une loi de composition interne o . Celle-ci fournit par passage au quotient sous l'action de $SL_2(\mathbf{Z})$ la loi de groupe sur C'_D décrite dans l'énoncé.

Description de Ξ : soit $f(x,y) = ax^2 + bxy + cy^2$ une forme primitive de discriminant $D < 0$, définie positive, de sorte que $(a,b,c) = 1$ et $a > 0$. Soit $\tau = \frac{-b+\sqrt{D}}{2a}$ la racine du polynôme $aT^2 + bT + c$ de partie imaginaire positive. Alors $\mathbf{a} = \mathbf{Z} \oplus \mathbf{Z}\tau$ est un idéal fractionnaire de K, de norme $N(\mathbf{a}) = |\ \frac{Disc(\{1,\tau\})}{D}\ |^{\frac{1}{2}} = \frac{1}{a}$. Si $f' = foU$ désigne une forme équivalente à f, avec $U = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}^{-1} \in SL_2(\mathbf{Z})$, on aura $\tau' = \frac{\alpha\tau+\beta}{\gamma\tau+\delta}$, et l'idéal fractionnaire $\mathbf{a}' = \mathbf{Z} \oplus \mathbf{Z}\tau' = (\gamma\tau + \delta)^{-1}(\mathbf{Z} \oplus \mathbf{Z}\tau)$ définira dans C_D le même élément que $\mathbf{a}$. D'où une application Ξ : classe de $f \in C'_D \mapsto$ classe de $\mathbf{a} \in C_D$.

Soit N la norme relative à $K/\mathbf{Q}$. Pour $x, y \in \mathbf{Z}$, on a avec les notations précédentes:

$$f(x,y) = a(x^2 + \frac{b}{a}xy + \frac{c}{a}y^2) = \frac{N(x+y\tau)}{N(\mathbf{a})}.$$

On en déduit que l'application qui, à tout idéal fractionnaire non nul $\mathbf{a}$ de K, associe la forme quadratique $\check{f} : \mathbf{a} \to \mathbf{Z} : \check{f}(\xi) = \frac{N(\xi)}{N(\mathbf{a})}$, définit, après choix d'une base orientée de $\mathbf{a}$ sur $\mathbf{Z}$ et passage au quotient par l'action de $SL_2\mathbf{Z}$), une application de C_D dans C'_D, inverse de Ξ. Donc Ξ est bien une bijection.

Définition: soit $f(x,y) = ax^2 + bxy + cy^2$ une forme quadratique, de discriminant D (et définie positive si $D < 0$). On dira *ici* que f est *réduite* si $|b| \leq |a| \leq |c|$; si $D < 0$, et si $a = |b|$ ou c, on demande de plus que $b \geq 0$.

Théorème 2: *Soient D un entier rationnel qui n'est pas un carré, et $f(x,y) = ax^2 + bxy + cy^2$ une forme quadratique de discriminant D (définie positive si $D < 0$).*

i) Si f est réduite, $|a| \leq \sqrt{\frac{|D|}{3}}$.

ii) f admet dans sa classe d'équivalence une forme réduite (et une seule si $D < 0$).

Démonstration: i) si f est réduite, $|D| \geq 4|ac| - |b|^2 \geq 4|a|^2 - |a|^2 = 3|a|^2$.

ii) L'action de $U = \begin{pmatrix} 1 & -n \\ 0 & 1 \end{pmatrix}$) (resp. $U = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$) $\in SL_2(\mathbf{Z})$ sur f remplace ses coefficients a, b, c par $a, b - 2na, c - nb + n^2a$ (resp. par $c, -b, a$). Par ailleurs, le coefficient a des formes équivalentes à f n'est jamais nul, car D n'est pas un carré. L'algorithme de division euclidienne de b par $2a$ permet donc de supposer que $-|a| \leq b \leq |a|$. Si le nouveau coefficient c obtenu est de valeur absolue $\geq |a|$, on a gagné (aux cas d'égalité près si $D < 0$). Sinon, on remplace a, b, c par $c, -b, a$ avec $|c| < |a|$, et on recommence. On obtient ainsi en un temps fini une forme de la classe de f avec $|a| \geq 1$ minimal, qu'il reste éventuellement à mettre sous forme réduite au moyen d'une dernière division euclidienne.

Pour vérifier l'unicité dans le cas défini positif, on note que le coefficient $a' = a\alpha^2 + b\alpha\gamma + c\gamma^2$ de $f' = f \circ \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$ est $\geq a$ si f est réduite, et égal à a ssi $\gamma = 0$. (Autrement dit, le coefficient a d'une forme réduite f est la plus petite valeur prise par f (et donc par toute f' équivalente à f) sur $\mathbf{Z}^2 \setminus \{0\}$.)

Le théorème 2 entraîne immédiatement que pour tout D non carré, le cardinal $h'(D)$ de C'_D est fini. Plus précisément, désignons par $d(n)$ le nombre de diviseurs d'un entier $n > 0$; par exemple, $d(n) = 2$ ssi n est un nombre premier. Alors, $d(n)$ est une fonction arithmétique multiplicative, qui vérifie:

$$\forall \epsilon > 0, \exists c_\epsilon > 0 \text{ tel que } \forall n \in \mathbf{N}, d(n) < c_\epsilon n^\epsilon$$

(voir Hardy-Wright, *An introduction to the theory of number*, Theorem 315). Puisque les coefficients a, b, c d'une forme réduite de discriminant D vérifient $|b| \leq \sqrt{\frac{|D|}{3}}$ et $4ac = b^2 - D$, on en déduit:

$$h'(D) \leq 2\ \Sigma_{|b|\leq\sqrt{\frac{|D|}{3}}}\ d(|D-b^2|) \leq 4\sqrt{|D|}.c_\epsilon(2|D|)^\epsilon = \ \mathrm{O}(|D|^{\frac{1}{2}+\epsilon}).$$

Par conséquent, d'après la proposition 2:

Corollaire: *Soit K un corps quadratique imaginaire, de discriminant $D<0$. Le nombre de classes $h(D)$ de K est fini, et vérifie* $\overline{\lim}_{D\to-\infty}\frac{log(h(D))}{log|D|}\leq\frac{1}{2}$.

[On démontre que cette limite supérieure est en fait une limite, et qu'elle vaut $\frac{1}{2}$. En particulier, il n'y a qu'un nombre fini de corps quadratiques imaginaires d'anneau d'entiers principal. L'analogue de ce dernier résultat pour les corps quadratiques réels n'est pas connu - et est probablement incorrect.]

§3. Un algorithme de factorisation sur Z.

Voici le principe de cet algorithme, dû à D. Shanks. Soit N un grand entier impair, que, pour simplifier, on supposera ici sans facteurs carrés, et soit D le discriminant du corps $K=\mathbf{Q}(\sqrt{-N})$. Considérons une forme quadratique définie positive $f(x,y)=ax^2+bxy+cy^2$, de discriminant D. Alors l'opposée de $\Xi(f)$ dans C_D est donnée par la forme $ax^2-bxy+cy^2$. Lorsque f est réduite, $\Xi(f)$ est donc une classe ambige (ou triviale) si et seulement si

ou bien $b=0$, auquel cas $D=-4ac$, et $N=ac$;

ou alors $a=b$, auquel cas $D=b(b-4c)$, et $N=b(4c-b)$ pour b impair, $N=\frac{b}{2}(2c-\frac{b}{2})$ pour b pair;

ou enfin $a=c$, auquel cas $D=(b-2a)(b+2a)$, et $N=(2a-b)(b+2a)$ pour b impair, $N=(\frac{b}{2}+a)(a-\frac{b}{2})$ pour b pair.

On retrouve ainsi le fait, établi au §1, que toute classe ambige donne lieu à une factorisation de D; inversément, on déduit des formules précédentes (en distinguant les cas $N\equiv 1,3 \bmod 4$), que toute factorisation de D provient d'une forme réduite dans une classe ambige. Autrement dit, la mise sous forme réduite des formes quadratiques fournit une bijection, maintenant *explicite*, entre les factorisations de D et les classes ambiges.
[NB: les formes normales d'Hermite auraient en fait aussi permis de travailler avec les idéaux de $\mathbf{O}_K$.]

Reste à construire quelques classes ambiges. Pour cela, on calcule d'abord le nombre de classes $h(D)=2^sq$, avec q impair. Pour tout $x\in C_D$, l'ordre de x^q est alors une puissance de 2, et si x n'est pas la classe nulle, il existe $r<s$ tel que x^{2^r} soit d'ordre exactement 2. Pour plus de détails sur les algorithmes de Shanks sous-tendant ces constructions, voir le livre d'H. Cohen, §§5.4 et 8.6.

.

CHAPITRE VII

THÉORIE ANALYTIQUE DES NOMBRES

§1. Séries de Dirichlet.

Une série de Dirichlet est une série de fonctions de la variable complexe s, de la forme $\Sigma_{n\geq 1}\frac{a_n}{n^s}$, où les a_n sont des nombres complexes, et $n^s = exp(sLogn)$. On pose souvent $s = \sigma + it$, où $\sigma = Re(s) \in \mathbf{R}$.

Proposition 1 : *Soit $(a_n)_{n\geq 1}$ une suite de nombres complexes. On suppose que la série de Dirichlet $f = \Sigma\frac{a_n}{n^s}$ converge au point $s_0 = \sigma_0 + it_0$. Alors:*

i) elle converge uniformément sur tout secteur angulaire fermé de sommet s_0, de bissectrice $s_0 + \mathbf{R}_{\geq 0}$ et d'angle $< \pi$.

ii) en particulier, elle converge sur le demi-plan ouvert $Re(s) > \sigma_0$, et sa somme $f(s)$ y est holomorphe.

iii) il existe un plus petit nombre réel σ_c (eventuellement égal à $-\infty$), appelé abscisse de convergence de la série de Dirichlet, tel que $f(s)$ converge si $\sigma > \sigma_c$, et diverge si $\sigma < \sigma_c$.

On pose $\sigma_c = +\infty$ si la série ne converge nulle part. L'abscisse de convergence σ_a de la série $\Sigma_{n\geq 1}\frac{|a_n|}{n^s}$ s'appelle l'abscisse de convergence absolue de f. Bien sûr, $\sigma_a \geq \sigma_c$.

La preuve de la proposition 1 repose sur l'analogue discret suivant de l'intégration par partie.

Lemme d'Abel : *Soien (a_n) et (b_n) deux suites de nombres complexes. Posons $A_{n,n'} = \Sigma_{k=n,\ldots,n'}a_n$. Alors,*

$$\Sigma_{k=n,\ldots,n'}a_n b_n = A_{n,n'}b_{n'} - \Sigma_{k=n,\ldots,n'-1}A_{n,k}(b_{k+1} - b_k)).$$

On utilisera également la majoration élémentaire suivante: pour tout $\beta > \alpha > 0$, et s de partie réelle $\sigma > 0$,

$$|e^{-\alpha s} - e^{-\beta s}| \leq |\frac{s}{\sigma}|(e^{-\alpha\sigma} - e^{-\beta\sigma}).$$

Démonstration de la proposition 1 : i) On se ramène par translation au cas $s_0 = 0$, où l'hypothèse revient à la convergence de la série Σa_n. On doit montrer que pour tout T arbitrairement grand, et tout $\epsilon > 0$, il existe N tel que pour tout $n' > n > N$, le reste de Cauchy $S_{n,n'}(s) = \Sigma_{k=n,\ldots,n'} \frac{a_k}{k^s}$ est de module $< \epsilon$ pour tout s dans le secteur $\{\sigma \geq 0, \frac{|s|}{\sigma} \leq T\}$. Par hypothèse, il existe un tel N pour $s = 0$, i.e. pour la somme $A_{n,n'}$ (notation du lemme d'Abel). En appliquant ce lemme à $b_n = n^{-s}$, et la majoration élémentaire, on obtient:

$$|S_{n,n'}(s)| \leq \epsilon(1 + \frac{|s|}{\sigma}(\Sigma_{k=n,\ldots,n'-1} \ (\frac{1}{k^\sigma} - \frac{1}{(k+1)^\sigma})))$$

d'où $|S_{n,n'}(s)| \leq \epsilon(1+T)$ pour tout s dans le secteur considéré.

ii) et iii) D'après le théorème de Weiestrass, la somme $f(s)$ de la série est holomorphe sur tout domaine où elle converge uniformément, donc est holomorphe sur $Re(s) > \sigma_0$. Les autres assertions découlent de i).

Exemples: i) la série $\zeta(s) = \Sigma_{n \geq 1} \frac{1}{n^s}$ admet $\sigma_c = 1 = \sigma_a$ pour abscisse de convergence. Sa somme s'appelle la fonction zeta de Riemann. Elle admet un prolongement méromorphe sur $\mathbf{C}$ (voir bibliographie), avec pour unique pôle $s = 1$, de résidu 1 (voir Proposition 3 ci-dessous). Elle s'annule en tous les nombres entiers < 0 pairs. L'hypothèse de Riemann, ouverte au concours pour un million de dollars, affirme que ses seuls autres zéros sont situés sur la droite $Re(s) = 1/2$.

ii) (Exercices) si les a_n sont bornés, la série de Dirichlet admet une abscisse de convergence absolue $\sigma_a \leq 1$. Si toutes les sommes partielles $A_{n,n'}$ sont bornées, elle admet une abscisse de convergence $\sigma_c \leq 0$.

iii) Si sa somme $f(s)$ est identiquement nulle sur son demi-plan de convergence, alors tous les coefficients a_n de la série de Dirichlet sont nuls.

On dit qu'un point s_0 de la droite $Re(s) = \sigma_c$ est une singularité de f si pour tout voisinage ouvert Ω de s_0 dans $\mathbf{C}$, la restriction de f à $\Omega \cap \{Re(s) > \sigma_c\}$ n'admet pas de prolongement holomorphe sur Ω.

Proposition 2 (Landau) : *Soit $f(s) = \Sigma_{n \geq 1} \frac{a_n}{n^s}$ une série de Dirichlet à coefficients $a_n \geq 0$, et d'abscisse de convergence $\sigma_c \in]-\infty, +\infty[$. Alors le point $s = \sigma_c$ est une singularité de f.*

Démonstration : par translation, on peut supposer que $\sigma_c = 0$. On raisonne par l'absurde. Si 0 n'est pas une singularité, f est holomorphe sur un disque de centre 1 et de rayon $1+2\eta$, pour un $\eta > 0$, donc $f(-\eta)$ est somme de la série de Taylor $\Sigma_{k \geq 0} \frac{f^{(k)}(1)}{k!}(-1)^k(1+\eta)^k$ de

f en 1, qui est absolument convergente. Par ailleurs, le théorème de Weiestrass permet de dériver terme à terme la série de Dirichlet de f, et $(-1)^k f^{(k)}(1)$ est somme de la série $\Sigma_{n\geq 1} a_n (Logn)^k n^{-1}$, qui est une série convergente à termes positifs. La série double à termes positifs $f(-\eta) = \Sigma_{n,k} a_n (Logn)^k (1+\eta)^k/(n.k!)$ est donc convergente, et vaut par Fubini

$$f(-\eta) = \Sigma_n \frac{a_n}{n} (\Sigma_k \frac{[(1+\eta)Logn]^k}{k!}) = \Sigma_n \ \frac{a_n}{n} e^{(1+\eta)Logn} = \Sigma_{n\geq 1} \ a_n n^{\eta}.$$

Ainsi, la série de Dirichlet converge en $s = -\eta$. Son abscisse de convergence serait donc $\leq -\eta$, en contradiction avec l'hypothèse.

Des propriétés énoncées plus haut sur la fonction ζ de Riemann, nous n'aurons besoin que de la

Proposition 3 : *La fonction $\zeta(s) - \frac{1}{s-1}$ admet un prolongement holomorphe sur le demi-plan $Re(s) > 0$.*

Démonstration : Notons $\{x\} = x - [x]$ la partie fractionnaire du nombre réel x. On a

$$\zeta(s) = \Sigma_{n\geq 1} n^{-s} = \Sigma_n \ s \int_n^{+\infty} \frac{dx}{x^{s+1}} = s \int_1^{+\infty} (\Sigma_{n\leq x} 1) \frac{dx}{x^{s+1}} = s \int_1^{+\infty} \frac{[x]dx}{x^{s+1}} = \frac{s}{s-1} - g(s),$$

où $g(s) = s \int_1^{+\infty} \frac{\{x\}dx}{x^{s+1}}$. Comme $\{x\}$ est bornée, cette dernière intégrale converge pour $Re(s) > 0$, et définit sur ce demi-plan une fonction holomorphe. Il en est donc de même de $\zeta(s) - \frac{1}{s-1} = 1 - g(s)$.

Fonctions multiplicatives et produits eulériens.

Définition 1 : une fonction $a : \mathbf{N} \to \mathbf{C}$ est dite *multiplicative* si $a(nm) = a(n)a(m)$ pour tout couple d'entiers n, m *premiers entre eux*; exemples : les fonctions ϕ et μ du chapitre I, §1. Elle est dite *strictement multiplicative* si cette propriété est vérifiée pour tout n, m; exemples: la fonction 1, ou les caractères de Dirichlet (voir §2). On note $\mathcal{P}$ l'ensemble des nombres premiers.

Lemme 2 : *Soit a une fonction multiplicative bornée.*

i) La série de Dirichlet $f(s) = \Sigma_{n\geq 1} \frac{a(n)}{n^s}$ converge absolument sur le demi-plan $Re(s) > 1$, et y est égale au produit infini convergent $\Pi_{p\in\mathcal{P}} f_p(s)$, où $f_p(s)$ désigne la somme de la série absolument convergente $\Sigma_{k\geq 0} \frac{a(p^k)}{p^{ks}}$.

ii) En particulier, si a est strictement multiplicative, $f(s) = \Pi_{p\in\mathcal{P}} \frac{1}{1-a(p)/p^s}$ (produit infini convergent sur $Re(s) > 1$). Par exemple, sur ce demi-plan, $\zeta(s) = \Pi_{p\in\mathcal{P}} \frac{1}{1-1/p^s}$.

iii) Pour σ tendant vers 1^+, on a $\Sigma_{p\in\mathcal{P}}\ p^{-\sigma} \sim Log(\frac{1}{\sigma-1})$ et $\Sigma_{p\in\mathcal{P},k\geq 2} p^{-k\sigma} = O(1)$.

Démonstration : i) la convergence absolue résulte de ce que les $a(n)$ sont bornés. Soit alors F un ensemble fini de nombres premiers, et $< F >$ l'ensemble des entiers dont tous les facteurs premiers appartiennent à F. De la multiplicativité de a, on tire $\Sigma_{n\in<F>} \frac{a(n)}{n^s} = \Pi_{p\in F}\ f_p(s)$. Quand F croît, le terme de gauche tend vers $f(s)$ si $Re(s) > 1$. Le produit infini est donc convergent, et tend vers $f(s)$.

ii) Dans ce cas, $f_p(s) = \Sigma_{k\geq 0}(\frac{a(p)}{p^s})^k = \frac{1}{1-a(p)/p^s}$.

iii) $Log(\zeta(\sigma)) = \Sigma_{p\in\mathcal{P},k\geq 1}\frac{1}{kp^{k\sigma}} = \Sigma_{p\in\mathcal{P}}\frac{1}{p^\sigma} + g_2(\sigma)$, où $g_2(\sigma) < g_3(\sigma) := \Sigma_{p\in\mathcal{P},k\geq 2}\frac{1}{p^{k\sigma}}$. Pour $\sigma > 1$, cette série est majorée par la série $\Sigma_{n\geq 2,k\geq 2}\frac{1}{n^k}$, qui converge. Donc g_3 et g_2 sont bornées, et on conclut par la proposition 3.

§2. Nombres premiers dans les progressions arithmétiques.

Caractères de Dirichlet et fonctions L.

La définition 2 ci-dessous fournit un exemple important de fonctions strictement multiplicatives. Rappelons qu'on appelle *caractère* d'un groupe fini G tout homomorphisme de groupe $\chi : G \to \mathbf{C}^*$, i.e. vérifiant $\chi(xy) = \chi(x)\chi(y)$ pour tout (x, y) dans G. Les caractères de G forment eux-mêmes un groupe pour la loi $(\chi_1\chi_2)(x) := \chi_1(x)\chi_2(x)$, appelé groupe dual de G, et noté $\hat{G}$. L'élément neutre de $\hat{G}$ est le caractère unité χ_0, défini par $\chi_0(x) = 1$ pour tout $x \in G$. On a $\Sigma_{x\in G}\chi_0(x) = |G|$, tandis que pour tout

$$\chi \neq \chi_0\ ,\ \Sigma_{x\in G}\ \chi(x) = 0.$$

(Preuve: choisir $y \in G$ tel que $\chi(y) \neq 1$; le premier membre, multiplié par $\chi(y) - 1$, vaut 0.)

Soit G un groupe cyclique, d'ordre $|G| = n$, dont on fixe un générateur γ. Un caractère χ de G est entièrement déterminé par $\chi(\gamma) = w \in \mu_n(\mathbf{C})$. Inversément, toute racine n-ième de l'unité w détermine par cette formule un caractère de G. Ainsi, $\hat{G} \simeq \mu_n(\mathbf{C})$ est un groupe cyclique d'ordre n, donc isomorphe (non canoniquement) à G. Plus généralement, soit G un groupe *abélien* fini G. On a vu au chap. 4, §1, qu'il est isomorphe à un produit de groupes cycliques. On en déduit que $\hat{G}$ est encore isomorphe à G, et en particulier, qu'ils ont le même ordre. L'application $x \mapsto \{\chi \mapsto \chi(x)\}$ fournit alors un isomorphisme canonique de G avec le dual de $\hat{G}$, et la relation précédent entraîne, pour tout x différent de l'élément neutre e de G:

$$x \neq e,\ \Sigma_{\chi\in\hat{G}}\ \chi(x) = 0.$$

(et bien sûr, $\Sigma_{\chi\in\hat{G}}\ \chi(e) = |G|$.) On en déduit, en notant que $\chi(x^{-1}) = \chi(x)^{-1} = \overline{\chi(x)}$ (puisque χ prend ses valeurs dans le cercle unité):

Lemme 3 (relations d'orthogonalité) : *Soient G un groupe abélien fini, et x, y deux éléments de G. Alors, $\Sigma_{\chi\in\hat{G}}\ \overline{\chi(y)}\chi(x)$ vaut 0 si $x \neq y$, et 1 si $x = y$.*

Définition 2 : soit m un entier ≥ 1, et $U = U_m$ le groupe multiplicatif $(\mathbf{Z}/m\mathbf{Z})^*$, d'ordre $\phi(m)$. On appelle *caractère de Dirichlet (modulo m)* tout caractère du groupe U_m. On étend un tel caractère χ en une fonction $\chi : \mathbf{N} \to \mathbf{C}$ en posant $\chi(n) = \chi(\text{classe de } n \text{ mod. } m)$ si n est premier à m, et $\chi(n) = 0$ sinon. Une telle fonction est strictement multiplicative. Comme elle est m-périodique, elle est bornée; mieux, la relation donnée au début du paragraphe entraîne que les sommes $A_{n,n'} = \Sigma_{k=n,\ldots,n'}\chi(k)$ sont bornées si $\chi \neq \chi_0$.

On déduit donc de l'exemple ii) du §1 que *si* $\chi \neq \chi_0$, la série de Dirichlet associée

$$L(s,\chi) = \Sigma_{n\geq 1}\frac{\chi(n)}{n^s}$$

admet a une abscisse de convergence $\sigma_c(L(s,\chi)) \leq 0$; on a $\sigma_a(L(s,\chi)) \leq 1$ et $L(s,\chi) = \Pi_{p\in\mathcal{P}}\frac{1}{1-\chi(p)/p^s}$ si $Re(s) > 1$. En revanche, $\sigma_c(L(s,\chi_0)) \leq 1$; en fait,

$$\zeta(s) = L(s,\chi_0).\Pi_{p|m}\frac{1}{1-1/p^s},$$

de sorte que $\sigma_c(L(s,\chi_0)) = 1$.

Posons

$$\zeta_m(s) = \Pi_{\chi\in\hat{U}}L(s,\chi),$$

où χ parcourt l'ensemble des caractères de Dirichlet modulo m, et désignons, pour tout nombre premier p ne divisant pas m, par f_p l'ordre de la classe $\overline{p}$ de p dans U_m; alors, $g_p := \phi(m)/f_p$ est l'ordre du groupe quotient $U_m/ < \overline{p} >$, et l'on a:

Proposition 4 : *i) $\zeta_m(s) = \Pi_{(p,m)=1}(\frac{1}{1-p^{-f_p s}})^{g_p}$ pour $Re(s) > 1$;*
ii) pour tout $\chi \neq \chi_0$, $L(1,\chi) \neq 0$, et $\zeta_m(s)$ admet un pôle simple en $s = 1$.

Démonstration : i) On a $\Pi_{w\in\mu_{f_p}(\mathbf{C})}(1 - wT) = 1 - T^{f_p}$. De plus, g_p caractères de U_m prennent la valeur w en $\overline{p}$. Donc $\Pi_{\chi\in\hat{U}}(1 - \chi(p)T) = (1 - T^{f_p})^{g_p}$. On conclut en posant $T = p^{-s}$.

ii) Comme $L(s,\chi_0)$ admet un pôle simple en $s = 1$, il suffit de montrer la première assertion. Si elle n'était pas satisfaite, la fonction ζ_m serait holomorphe en $s = 1$, donc sur

$Re(s) > 0$. Mais vu i), c'est la somme d'une série de Dirichlet à coefficients réels positifs. D'après la Proposition 2, cette série admettrait donc une abscisse de convergence $\sigma_c \leq 0$. Or elle est minorée par la série $\Pi_{(p,m)=1}(1 + p^{-f_p s} + ...)^{g_p}$, donc aussi par $\Pi_{(p,m)=1}(1 + p^{-\phi(m)s} + ...)$, donc aussi par $\Sigma_{(n,m)=1} n^{-\phi(m)s}$, qui diverge pour $\sigma = 1/\phi(m)$.

Le théorème de Dirichlet

Nous sommes maintenant en mesure de démontrer le

Théorème (Dirichlet) *Soient m et a deux entiers naturels premiers entre eux. Il existe une infinité de nombres premiers p tels que $p \equiv a$ mod m.*

On va en fait obtenir un énoncé plus précis sur la répartition de ces nombres premiers, au moyen de la notion suivante. Rappelons que $\Sigma_{p\in\mathcal{P}}\ p^{-\sigma} \sim Log(1/(\sigma - 1))$ quand $\sigma \to 1^+$ (Lemme 2). On dit qu'une partie X de $\mathcal{P}$ admet une densité de Dirichlet δ si $\Sigma_{p\in X}\ p^{-\sigma} \sim \delta Log(1/(\sigma - 1))$ quand $\sigma \to 1^+$. Nous allons montrer que pour tout a premier à m, l'ensemble $\mathcal{P}_{a,m} := \mathcal{P}_a$ des nombres premiers congrus à a modulo m admet une densité de Dirichlet $\delta = 1/\phi(m)$ (et est donc bien infini).

Dans le cas du groupe U_m, les relations d'orthogonalité montrent que la fonction

$$n \mapsto \frac{1}{\phi(m)} \Sigma_{\chi\in\hat{U}}\ \overline{\chi(a)}\chi(n)$$

vaut 1 si $n \equiv a$ mod m, et 0 sinon. Par conséquent, sa restriction à l'ensemble $\mathcal{P}$ est la fonction caractéristique de $\mathcal{P}_a$.

Démonstration du Théorème : soit χ un caractère de Dirichlet modulo m. En prenant la dérivée logarithmique de $L(s, \chi)$ sur le domaine de convergence $Re(s) > 1$ de son développement en produit eulérien, on obtient

$$-\frac{L'}{L}(s, \chi) = \Sigma_{p\in\mathcal{P}}\ \frac{\chi(p)p^{-s}Logp}{1 - \chi(p)p^{-s}} = \Sigma_p \Sigma_{k\geq 1} \frac{\chi(p^k)Logp}{p^{ks}} = \Sigma_{n\geq 1} \frac{\chi(n)\Lambda(n)}{n^s},$$

où Λ désigne la fonction de von Mangoldt, définie par $\Lambda(n) = Log\ p$ si n est une puissance pure p^k d'un nombre premier p, $\Lambda(n) = 0$ sinon.

Multiplions la relation précédente par $\overline{\chi(a)}$, et sommons sur tous les $\chi \in \hat{U}$. De la relation d'orthogonalité, on déduit:

$$\Sigma_{n\equiv a \text{ mod } m}\ \frac{\Lambda(n)}{n^s} = -\frac{1}{\phi(m)} \Sigma_{\chi\in\hat{U}}\ \overline{\chi(a)} \frac{L'}{L}(s, \chi).$$

Pour $s = \sigma \to 1^+$, le membre de gauche de cette égalité est de la forme

$$\Sigma_{p \equiv a \bmod m} \frac{Logp}{p^\sigma} + O(1).$$

Les termes du membre de droite correspondant à des caractères $\chi \neq \chi_0$ sont tous bornés, puisque $L(1, \chi)$ ne s'annule pas. Quant à celui de χ_0, il se comporte comme la dérivée logarithmique de la fonction ζ, donc est de la forme

$$\frac{1}{\phi(m)} \frac{1}{\sigma - 1} + O(1).$$

Par conséquent, pour $\sigma \to 1^+$:

$$\Sigma_{p \in \mathcal{P}_a} \frac{Logp}{p^\sigma} = \frac{1}{\phi(m)} \frac{1}{\sigma - 1} + O(1),$$

ce qui entraîne déjà que $\mathcal{P}_a$ est infini. Mais mieux: en intégrant cette relation entre $\sigma > 1$ et 2, on obtient

$$\Sigma_{p \in \mathcal{P}_a} \frac{1}{p^\sigma} = \frac{1}{\phi(m)} Log(\frac{1}{\sigma - 1}) + O(1),$$

et $\mathcal{P}_a$ admet bien une densité de Dirichlet, égale à $\frac{1}{\phi(m)}$.

Remarque: d'après le théorème des nombres premiers (Hadamard - La Vallée Poussin), pour tout nombre réel x suffisamment grand, le nombre

$$\pi(x) = card\{p \in \mathcal{P}, p \leq x\}$$

de nombres premiers $\leq x$ est de l'ordre de $\frac{x}{Logx}$. Pour a et m premiers entre eux, soient alors

$$\pi(x; a, m) = card\{p \in \mathcal{P}, p \equiv a \bmod m, p \leq x\}$$

le nombre d'éléments $\leq x$ de $\mathcal{P}_a = \mathcal{P}_{a,m}$. On peut préciser le théorème de Dirichlet sous la forme

$$lim_{x \to +\infty} \frac{\pi(x; a, m)}{\pi(x)} = \frac{1}{\phi(m)},$$

ce qu'on exprime en disant que $\mathcal{P}_a$ admet une densité naturelle égale à $\frac{1}{\phi(m)}$.

Bibliographie complémentaire

G. Tennenbaum et M. Mendès-France : *Les nombres premiers*; PUF, coll. "Que sais-je ?", No 571, 1997.