COURS DE L'INSTITUT FOURIER

JEAN-RENÉ JOLY

Chapitre 7 Le théorème d'Ax

Cours de l'institut Fourier, tome 4 (1971), p. 1-11

<http://www.numdam.org/item?id=CIF_1971__4__A7_0>

Chapitre 7

Le théorème d'Ax

(7.1). Introduction.

Le but de ce chapitre est de démontrer le théorème suivant:

Théorème 1 (Ax). - Soit F_1 , F_2 , ... , F_r une famille de polynômes appartenant à $K[X] = K[X_1 , \dots , X_n]$, K étant un corps fini à $q = p^f$ éléments. Soit d_i le degré de F_i , et posons

$d = d_1 + d_2 + \dots + d_r$;

b = le plus grand strictement inférieur à n/d ;

N = le nombre de solutions dans K^n du système d'équations

$$(\Sigma) \quad \begin{cases} F_1(X_1 , \dots , X_n) = 0 \\ \dots\dots\dots \\ F_r(X_1 , \dots , X_n) = 0 \ ; \end{cases}$$

on a alors la congruence

$$(1) \qquad N \equiv 0 \pmod{q^b}.$$

Naturellement, la congruence (1) ne nous apprend quelque chose que si $b > 0$, donc si $n > d$; elle implique alors en particulier, sans autre hypothèse, que N est divisible par q , ce qui précise le théorème de Chevalley-Warning. A ce propos, notons deux choses:

Remarque 1. - Une assertion du type N est divisible par q a un caractère plus naturel qu'une assertion du type N est divisible par p : car N et

q ont des significations géométriques analogues: respectivement, nombre de points dans K^n de la variété algébrique affine définie par (Σ), et nombre de points de la droite affine sur K ; mais p n'admet pas d'interprétation de ce genre.

Remarque 2. - La combinaison du théorème de Warning (tout seul: théorème 2 du chapitre 3) et du théorème d'Ax donne le résultat suivant:

Corollaire. - Si le système (Σ) admet au moins une solution (par exemple, si tous les F_i sont sans terme constant), alors on peut écrire

$$N = q^b N' ,$$

avec

$$N' \geqslant q^{n-d-b} .$$

A titre d'exercice, le lecteur pourra essayer de "dire le maximum de choses" sur le nombre de solutions d'une équation telle que

$$a_1 X_1^3 + a_2 X_2^3 + \ldots + a_7 X_7^3 = 0$$

en utilisant le corollaire ci-dessus, ainsi que le théorème 1 du chapitre 6.

*

Le théorème 1 est démontré au paragraphe 2; la démonstration proposée est essentiellement celle d'Ax, à deux détails près:

on évite l'introduction de corps p-adiques;

on utilise les relations de Stickelberger "$\mathrm{ord}(\tau(j)) = \sigma(j)$" telles qu'elles ont été démontrées directement (et assez élémentairement) au chapitre 5, paragraphe 6, et sans les déduire des congruences (67) de ce même chapitre 5, congruences dont la démonstration est nettement barbante!
A part cela, les calculs sont exactement les mêmes, c'est-à-dire astucieux mais rébarbatifs. Le paragraphe 3 montre comment le cas général (r quel-

conque) peut se déduire par un argument combinatoire du cas particulier où $r = 1$.

(7.2). Démonstration du théorème 1 pour $r = 1$.

On s'intéresse donc à une seule équation

$$F(X_1, \dots, X_n) = 0, \tag{2}$$

à n variables, de degré d, et N désigne le nombre de solutions de (2) dans K^n, b le plus grand entier strictement inférieur à n/d. On utilisera systématiquement les notations et résultats du chapitre 5, paragraphe 6: on conseille donc au lecteur de s'y reporter pour se rafraîchir la mémoire! Rappelons quand même que F est le corps cyclotomique des racines $p(q-1)^{\text{ièmes}}$ de l'unité sur Q, que $\zeta \in F$ est une racine primitive $p^{\text{ième}}$ de l'unité, que $T^* \subset F$ est le groupe des racines $(q-1)^{\text{ièmes}}$ de l'unité, que $\mathfrak{P}$ est un diviseur premier quelconque de (p) dans O_F, qu'on identifie K à $O_F/\mathfrak{P}$, et que $\chi : K^* \longrightarrow T^*$ est le caractère multiplicatif "canonique" de K correspondant à cette identification. On posera d'autre part $T = T^* \cup \{0\}$, et on prolongera χ à K tout entier en décrétant comme d'habitude que $\chi(0) = 0$: χ devient ainsi une bijection de K sur T, et si, pour tout $t \in T$, $\bar{t}$ désigne la classe de t modulo $\mathfrak{P}$, alors $\chi : K \longrightarrow T$ est la bijection inverse de $t \longmapsto \bar{t}$ ($T \longrightarrow K$). Enfin, on notera θ le caractère additif de K à valeurs dans F défini comme précédemment par $\theta(x) = \zeta^{Tr(x)}$ ($x \in K$, $Tr = Tr_{K/F_p}$).

Pour simplifier la démonstration, qui est du type calculatoire, nous la découperons en plusieurs étapes.

1ère étape. - Introduction du polynôme $C(Y)$.

Soit Y une indéterminée; comme $\mathrm{card}(T) = q$, il existe évidemment un (et un seul) polynôme $C(Y)$, de degré $q-1$ et à coefficients dans F , et tel que $C(t) = \theta(\bar{t})$ pour tout $t \in T$. Posons

$$C(Y) = \sum_{j=0}^{q-1} c_j Y^j . \tag{3}$$

Lemme 1. - Les notations $\sigma(j)$, $\tau(j)$ et ord ayant la même signification qu'au chapitre 5, paragraphe 6, on a

$$\begin{cases} c_0 = 1 , \quad c_{q-1} = -q/(q-1) , \\ \text{et} \quad c_j = \tau(j)/(q-1) \quad \text{pour} \quad 1 \leq j < q-1 . \end{cases} \tag{4}$$

En particulier, pour tout j tel que $0 \leq j \leq q-1$, on a

$$\mathrm{ord}(c_j) = \sigma(j) . \tag{5}$$

Démonstration. (5) résulte immédiatement de (4), de la définition de $\sigma(j)$, du théorème 5 du chapitre 5, pour $1 \leq j < q-1$ et

pour $j = 0$, du fait que $\mathrm{ord}(1) = 0$;

pour $j = q-1$, du fait que $\mathrm{ord}(-q/(q-1)) = \mathrm{ord}(q) = \mathrm{ord}(p^f) = f\,\mathrm{ord}(p) \Rightarrow f(p-1) = \sigma(q-1)$.

Prouvons donc (4). Remarquons d'abord que puisque T^* est le groupe des racines $(q-1)^{\text{ièmes}}$ de l'unité, on a

$$\sum_{t \in T^*} t^u = \begin{cases} q , & \text{si } u = 0 ; \\ 0 , & \text{si } q-1 \text{ ne divise pas } u ; \\ q-1 , & \text{dans les autres cas.} \end{cases} \tag{6}$$

Si alors on suppose $1 \leq j < q-1$, et si on développe $\tau(j) = \tau(\chi^{-j}) = \sum_{x \in K^*} \chi^{-j}(x)\,\theta(x) = \sum_{t \in T^*} t^{-j}\,\theta(\bar{t}) = \sum_{t \in T^*} t^{-j}\,C(t)$ grâce à (3),

on trouve, compte tenu de (6),

(7) $$(q-1)c_j = \tau(j) ;$$

de même, si on développe $\sum_{t \in T} C(t) = \sum_{t \in T} \theta(\bar{t}) = 0$ grâce à (3) et à (6), on trouve

(8) $$qc_0 + (q-1)c_{q-1} = 0 ;$$

enfin, il est évident que

(9) $$c_0 = \theta(0) = 1 ;$$

les égalités (7), (8) et (9) donnent immédiatement les relations (4).

2ème étape. - Evaluation de N à l'aide du caractère θ.

Lemme 2. - On a l'égalité

(10) $$qN = \sum_{x_0} \sum_{x_1} \dots \sum_{x_n} \theta(x_0 F(x_1, \dots, x_n)) .$$

Démonstration. Déjà faite, puisque ce lemme coïncide avec le lemme 1 du chapitre 6, paragraphe 2. (Dans la somme de droite, chaque x_i parcourt K tout entier).

Nous allons transformer cette formule (10); introduisons pour cela quelques notations:

$x = (x_0, x_1, \dots, x_n)$ désignera le "point générique" de K^{n+1} ;

U désignera l'ensemble des suites $u = (u_1, \dots, u_n)$ ayant les deux propriétés suivantes: chaque u_i est un entier $\geqslant 0$; pour tout $u \in U$, la hauteur de u (égale par définition à $u_1 + \dots + u_n$, et notée $\| u \|$) est inférieure ou égale à $d = \deg(F)$;

si $u \in U$, u' désignera la suite $(1, u_1, \ldots, u_n)$; X^u signifiera $X_1^{u_1} \ldots X_n^{u_n}$, $X^{u'}$ signifiera $X_0 X_1^{u_1} \ldots X_n^{u_n}$; significations analogues pour x^u et $x^{u'}$ si $x \in K^{n+1}$ et $u \in U$.

Cela étant, on peut écrire successivement

$$F(X) = \sum_{u \in U} a_u X^u , \tag{11}$$

$$x_0 F(x_1, \ldots, x_n) = \sum_{u \in U} a_u x^{u'} , \tag{12}$$

$$\theta(x_0 F(x_1, \ldots, x_n)) = \prod_{u \in U} \theta(a_u x^{u'}) . \tag{13}$$

Dans le passage de (12) à (13), on utilise évidemment le fait que θ est un caractère additif. D'autre part, si $b_u = \chi(a_u)$, si $t_i = \chi(x_i)$ et si on pose $t = (t_0, t_1, \ldots, t_n)$, $b_u t^{u'}$ est évidemment un élément de T, d'image canonique dans $K = O_F/\mathfrak{P}$ égale à $a_u x^{u'}$; d'où, par définition de $C(Y)$ et en utilisant (3),

$$\theta(a_u x^{u'}) = C(b_u t^{u'}) = \sum_{j=0}^{q-1} c_j b_u^j t^{ju'} . \tag{14}$$

Les égalités (13) et (14) mènent ainsi à

$$\theta(x_0 F(x_1, \ldots, x_n)) = \prod_{u \in U} \sum_{j=0}^{q-1} c_j b_u^j t^{ju'} \tag{15}$$

Désignons alors par J l'ensemble de toutes les applications de U dans $\{0, 1, \ldots, q-1\}$, c'est-à-dire en somme l'ensemble de toutes les "façons d'associer un j à chaque u"; la distributivité de l'addition par rapport à la multiplication (dans F) permet de réécrire (15):

$$\theta(x_0 F(x_1, \ldots, x_n)) = \sum_{j \in J} \prod_{u \in U} c_{j(u)} b_u^{j(u)} t^{j(u)u'} . \tag{16}$$

(10) et (16) donnent ainsi

$$(17) \qquad qN = \sum_{j \in J} \sum_{t \in T^{n+1}} \prod_{u \in U} c_{j(u)} b_u^{j(u)} t^{j(u)u'} .$$

Pour chaque $j \in J$, désignons maintenant pour simplifier $\prod_{u \in U} b_u^{j(u)}$ par $b^{(j)}$ (c'est un élément de T) et posons

$$(18) \qquad \sum_{u \in U} j(u)u' = \left(\sum_u j(u) , \sum_u j(u)u_1 , \ldots , \sum_u j(u)u_n \right) = e'_j .$$

L'égalité (17) se simplifie alors en

$$(19) \qquad qN = \sum_{j \in J} b^{(j)} \prod_{u \in U} c_{j(u)} \sum_{t \in T^{n+1}} t^{e'_j} .$$

3ème étape. - Réduction du problème.

Comme tous les termes du membre de droite, dans (19), sont dans l'anneau O_F des entiers de F, il suffit pour prouver le théorème de montrer ceci:

(Div1) Quel que soit $j \in J$, l'entier algébrique $\prod_{u \in U} c_{j(u)} \sum_{t \in T^{n+1}} t^{e'_j}$ est divisible (dans O_F) par q^{b+1}.

Convenons d'écrire $q - 1 \mid e'_j$ si $q - 1$ divise chacune des $n + 1$ composantes de e'_j, et $q - 1 \nmid e'_j$ dans le cas contraire; alors, d'après les relations (6), on a

$$(20) \qquad \sum_t t^{e'_j} = q^{n+1} \qquad \text{si } e'_j = (0, 0, \ldots , 0) ;$$

$$(20') \qquad \sum_t t^{e'_j} = 0 \qquad \text{si } q - 1 \nmid e'_j$$

$$(21) \qquad \sum_t t^{e'_j} = (q - 1)^{s+1} q^{n-s} \qquad \text{si } e'_j \neq (0, 0, \ldots , 0) , \text{ si } q - 1 \mid e'_j$$

et si e_j (= e'_j privé de sa première composante) possède exactement s composantes non nulles (noter que la première composante de

e'_j est alors automatiquement différente de 0).

Pour prouver (Div1), donc le théorème, il suffit ainsi en fait, puisque évidemment $b \leq n$, de démontrer ceci:

(Div2) Si $j \in J$ est tel que e'_j soit différent de $(0, 0, \ldots, 0)$, soit divisible par $q - 1$, et que e_j possède exactement s composantes non nulles, alors l'entier algébrique $q^{n-s} \prod_{u \in U} c_{j(u)}$ est (dans l'anneau O_F) divisible par q^{b+1} .

4ème étape. Démonstration de (Div2).

Pour tout $u \in U$ et tout $j \in J$, écrivons l'entier $j(u)$ en base p :

$$j(u) = j_0(u) + j_1(u)p + \ldots + j_{f-1}(u)p^{f-1} . \tag{22}$$

(avec $0 \leq j_i(u) \leq p - 1$ pour chaque i). Ceci définit $j_i(u)$ pour $0 \leq i < f$; étendons cette définition en convenant que, quel que soit $z \in Z$, $j_z(u) = j_{i(z)}(u)$, où $i(z)$ est le reste de division de z par f . Enfin, pour $h = 0, 1, \ldots, f - 1$, posons

$$j^{(h)}(u) = \sum_{0 \leq i \leq f-1} j_{i-h}(u)p^i \tag{23}$$

(les $j^{(h)}(u)$ sont les f entiers déduits de $j(u)$ en permutant circulairement les chiffres de $j(u)$ en base p). Il est clair qu'on ne change rien aux égalités (20), (20') et surtout (21) en y remplaçant j par $j^{(h)}$, car ceci équivaut formellement à effectuer sur T la permutation $t \longmapsto t^{p^h}$; en particulier, cette substitution ne modifie pas la valeur de s : donc

$$s(q - 1) \leq \| e_{j(h)} \| = \Big\| \sum_u j^{(h)}(u)u \Big\| \leq d \sum_u j^{(h)}(u) . \tag{24}$$

Mais $\sum_u j^{(h)}(u)$ est la première composante de $e'_{j(h)}$, c'est donc un

entier positif divisible par $q - 1$ (cf. hypothèses de (Div2)); si $(s/d)^*$ désigne le plus petit entier supérieur ou égal à s/d, (24) implique donc

$$(25) \qquad (q-1)(s/d)^* \leq \sum_u j^{(h)}(u) .$$

Dans cette égalité, faisons $h = 0, 1, \ldots, f-1$ et additionnons:

$$(26) \qquad f(q-1)(s/d)^* \leq \sum_h \sum_u \sum_i j_{i-h}(u) p^i .$$

Une interversion de l'ordre des sommations, et l'utilisation de la notation $\sigma(j)$, transforment ceci en

$$(27) \qquad f(q-1)(s/d)^* \leq (1 + p + \ldots + p^{f-1}) \sum_u \sigma(j(u)) .$$

Mais d'une part $q = p^f$, ce qui permet de simplifier par le coefficient de la somme $\sum_u$ du second membre; et d'autre part $\sigma(j(u)) = \mathrm{ord}(c_{j(u)})$, d'après le lemme 1, (5), de ce paragraphe ($1^{\text{ère}}$ étape); (27) devient donc

$$(28) \qquad f(p-1)(s/d)^* \leq \mathrm{ord}\left(\prod_{u \in U} c_{j(u)} \right) .$$

Ainsi

$$(29) \qquad \mathrm{ord}\left(q^{n-s} \prod_{u \in U} c_{j(u)} \right) \geq f(p-1)\left[n - s + (s/d)^* \right]$$

Or, le symbole ord désigne la valuation $\mathfrak{P}$-adique normalisée associée à n'importe quel idéal premier $\mathfrak{P}$ de O_F divisant p (voir chapitre 5, paragraphe 6, propriété (iii) et la suite) et on a

$$(30) \qquad pO_F = \prod_{\mathfrak{P} \mid p} \mathfrak{P}^{p-1} ,$$

donc, puisque $q = p^f$,

$$(31) \qquad qO_F = \prod_{\mathfrak{P} \mid p} \mathfrak{P}^{f(p-1)} .$$

Ainsi, la propriété (Div2) (et par suite le théorème) sera prouvée si nous démontrons le lemme ci-dessous:

Lemme 3. - Pour tout entier s tel que $0 \leq s \leq n$, on a l'inégalité

$$n - s + (s/d)^* \geqslant b + 1 . \tag{32}$$

Démonstration. Il est clair que pour tout entier positif k, on a

$$k \geqslant ((s+k)/d)^* - (s/d)^* ;$$

car, pour $k = 0$, les deux membres sont égaux, et d'autre part le membre de gauche croît (en fonction de k) au moins aussi vite que le membre de droite. Faisons en particulier $k = n - s$: il vient, après transfert de $(s/d)^*$ dans le membre de gauche,

$$n - s + (s/d)^* \geqslant (n/d)^* = b + 1 , \quad \text{c.q.f.d.}$$

Le théorème d'Ax est ainsi démontré dans le cas où r = le nombre d'équations = 1 .

(7.3). Démonstration du théorème 1 dans le cas général (r quelconque).

Les hypothèses et notations sont celles du paragraphe 1; on se ramène en fait au cas d'une seule équation (cas réglé au paragraphe précédent) grâce à un lemme combinatoire bien connu:

Lemme 4. - L'entier r étant supposé $\geqslant 2$, posons $R = \{1, \dots, r\}$. Soient $V_1, \dots, V_r$ des ensembles finis, et soit

$$V = \bigcap_{j=1}^{r} V_j$$

l'intersection de tous les V_j. Soit d'autre part $\mathfrak{S}$ l'ensemble de toutes les parties finies non vides de R, et pour tout élément S de $\mathfrak{S}$, posons

$$U_S = \bigcup_{j \in S} V_j .$$

On a alors

$$\text{(33)} \qquad \operatorname{card}(V) = \sum_{S \in \mathcal{T}} (-1)^{\operatorname{card}(S)-1} \operatorname{card}(U_S) .$$

Démonstration. Par récurrence sur r. Pour $r = 2$, il s'agit de prouver une formule du type

$$\text{(34)} \qquad \operatorname{card}(V' \cap V'') = \operatorname{card}(V') + \operatorname{card}(V'') - \operatorname{card}(V' \cup V'') ,$$

ce qui est immédiat. Pour $r \geqslant 3$, on pose

$$V' = V_1 , \quad V'' = V_2 \cap \ldots \cap V_r ,$$

et on utilise successivement la formule (34), l'hypothèse de récurrence et la distributivité de la réunion par rapport à l'intersection.

Prouvons alors le théorème 1 dans le cas général, disons, pour $r \geqslant 2$. Appliquons le lemme 4 en prenant pour V_j l'ensemble des solutions dans K^n de l'équation (unique) $F_j(X) = 0$; V est alors l'ensemble des solutions dans K^n du système (Σ), et on a donc $\operatorname{card}(V) = N$. Par ailleurs, U_S ($S \in \mathcal{T}$) est égal à l'ensemble des solutions dans K^n de l'équation (unique) $F_S(X) = 0$, avec par définition $F_S = \prod_{j \in S} F_j$. Si on pose $N_S = \operatorname{card}(U_S)$, le lemme 4 donne

$$\text{(35)} \qquad N = \sum_{S \in \mathcal{T}} (-1)^{\operatorname{card}(S)-1} N_S .$$

D'autre part, le théorème 1 pour une équation montre que chaque entier N_S est divisible par q^b, puisque chaque polynôme F_S est un polynôme à n variables, mais de degré évidemment $\leq d = \deg(F_R) = \deg(F_1 F_2 \ldots F_r)$. Dans (35), chaque terme de la somme de droite est divisible par q^b, N est donc lui-même divisible par q^b, et c'est fini...

Signalons pour terminer ce chapitre que, du point de vue de la divisibilité de N par p, le théorème d'Ax est "le meilleur possible": voir [1], §4.