Sur le 17ème problème de Hilbert

Danielle Gondard-Cozette

ORSAY Série A
N° d'Ordre 1398

SUR LE 17eme PROBLEME DE HILBERT

THESE

présentée au centre d'Orsay de
l'Université de Paris-Sud pour
obtenir le titre de

DOCTEUR de 3eme CYCLE

spécialité : Mathématiques Pures

par

DANIELLE GONDARD

Assistante à l'Université de Bretagne Occidentale

soutenue le 24 janvier 1973 devant la Commission d'Examen :

MM.	P. SAMUEL	Président
	A. NERON	Examinateurs
	G. POITOU	
	P. RIBENBOIM	Invité

10

SUR LE 17EME PROBLEME DE HILBERT

THESE

PRESENTEE AU CENTRE D'ORSAY DE

L'UNIVERSITE DE PARIS-SUD POUR

OBTENIR LE TITRE DE

DOCTEUR DE 3ÈME CYCLE

par

DANIELLE GONDARD

Assistante à l'Université de Bretagne Occidentale.

Je remercie vivement Monsieur Samuel de m'avoir donné le sujet de cette thèse, d'avoir dirigé mes travaux et de m'avoir aidée de ses conseils.

Je remercie aussi Monsieur Poitou et Monsieur Néron de s'être intéressés à cette thèse et d'avoir accepté de faire partie du jury.

J'exprime également ici ma reconnaissance à Monsieur Ribenboim qui, après avoir dirigé mon D.E.A. et m'avoir orientée vers ces problèmes d'Arithmétique des Corps, continue de m'aider et de me conseiller dans mon travail.

Il serait injuste de ne pas remercier aussi les Mathématiciens de l'Université de Bretagne Occidentale pour la gentillesse avec laquelle ils ont toujours accepté de répondre à mes questions.

Enfin je remercie Madame Quéré pour le soin qu'elle a apporté à la réalisation matérielle de cette thèse.

En 1888 dans un article [1] Hilbert définissait une forme réelle définie :

Une forme f *de degré* n *(homogène) de* m *variables à coefficients réels sera dite définie si pour tout ensemble de valeurs réelles des variables elle prend une valeur positive ou nulle.*

Une somme de carrés de formes réelles est évidemment une forme définie ; Hilbert posait alors la question suivante :

Une forme définie peut-elle être représentée comme somme de carrés de formes réelles ?

Toujours dans le même article Hilbert répondait par la négative à cette question. On peut toujours trouver parmi les formes définies de degré n et de m variables des formes qui ne peuvent pas être décomposées en une somme finie de carrés de formes réelles sauf dans les trois cas particuliers donnés ci-dessous :

1. Les formes quadratiques définies de m variables, qui sont représentables par une somme de m carrés de formes.

2. Les formes définies de deux variables et de degré quelconque n, qui sont représentables par une somme de deux carrés de formes.

3. Les formes biquadratiques définies de trois variables, qui sont représentables par une somme de trois carrés de formes.

Remarquons premièrement qu'une forme définie ne peut être que de degré pair ; en effet, si f est une forme de degré impair 2n+1 et si on donne à toutes les variables la même valeur a, alors $\varphi(a, \ldots, a) = \lambda\, a^{2n+1} = \varepsilon|\lambda|\, a^{2n+1}$. Si on fait alors tendre a vers $\pm\infty$ on en déduit que φ tend vers $\pm\varepsilon\infty$. Dans l'un des deux cas la forme tend vers $-\infty$ et elle ne peut donc pas être définie. Donc si f est une forme définie alors le degré de f est pair, soit 2 n.

Il démontre d'abord le résultat dans le cas de formes Φ de degré 6 et de trois variables et dans le cas de formes ψ biquadratiques et de 4 variables. Puis il considère le cas général.

Soit une forme Φ de degré 6 et de trois variables non décomposables en somme de carrés de formes. Nous allons montrer qu'il est alors toujours possible de construire une forme définie de degré $2n \geq 6$ et de $m \geq 3$ variables telle que si on la suppose décomposable en somme de carrés de formes cela entraîne que Φ serait elle-même décomposable en somme de carrés de formes, ce qui est impossible par hypothèse.

Tout d'abord montrons qu'il est possible de trouver une forme de degré $2n$, $n > 3$, de trois variables non décomposables en somme de carrés de formes. Soit la forme $x^{2n-6}\,\Phi(x, y, z)$; supposons-la décomposée, nous aurions donc

$$x^{2n-6}\,\Phi(x, y, z) = \sum_{i=1}^{p} (\varphi_i(x, y, z))^2$$

Dans cette expression soit a_i le terme constant en x de $\varphi_i(x, y, z)$. Comme il n'y a pas de tel terme dans le membre de gauche nous devons donc avoir :

$$0 = \Sigma\, a_i^2 \quad \text{où} \quad a_i \in \mathbb{R}[y, z] \subset \mathbb{R}(y, z)$$

Comme $\mathbb{R}(y, z)$ est un corps ordonnable cette égalité entraîne que

$$\forall\, i \qquad a_i = 0 \qquad (*)$$

On peut donc alors mettre x en facteur dans $\varphi_i(x, y, z)$ et l'égalité devient :

$$x^{2n-6}\,\Phi(x, y, z) = x^2 \sum_{i=1}^{p} (\varphi_{i1}(x, y, z))^2$$

et donc

$$x^{2n-8}\,\Phi(x, y, z) = \sum_{i=1}^{p} (\varphi_{i1}(x, y, z))^2$$

(*) La théorie des corps ordonnables étant postérieure à cet article j'ignore quelle était la démonstration de Hilbert car il ne fait que citer le résultat dans [1]

Et recommençant le raisonnement et en l'itérant on épuise les puissances de x en facteur à gauche et on finit par trouver une décomposition de Φ en somme de carrés de formes, ce qui est impossible.

Montrons maintenant qu'on peut construire une forme de degré $2n \geqslant 6$ et de $m \geqslant 3$ variables qui ne soit pas décomposable en somme de carrés de formes.

Considérons la forme
$x^{2n-6}\ \Phi(x, y, z) + t^{2n} + ... + w^{2n}$, elle est évidemment définie puisque Φ l'est elle-même. Supposons cette forme décomposable en somme de carrés de formes réelles ; nous aurions :

$$x^{2n-6}\ \Phi(x, y, z) + t^{2n} + ... + w^{2n} = \sum_{i=1}^{p} (\varphi_i(x, y, z, ...,w))^2$$

Egalons alors à zéro les variables t, ..., w, nous obtenons :

$$x^{2n-6}\ \Phi(x, y, z) = \sum_{i=1}^{p} (\varphi_i(x, y, z, 0, 0 ...., 0))^2$$

Or nous venons de voir que une telle décomposition était impossible. On en déduit que la forme $x^{2n-6}\ \Phi(x, y, z) + t^{2n} + ... + w^{2n}$ est non décomposable en somme de carrés de formes.

Ceci résout donc le cas des formes de degré $2n \geqslant 6$ et de $m \geqslant 3$ variables. Pour terminer il convient d'examiner le cas des formes définies de degré 4 et de $m \geqslant 4$ variables.

Soit ψ une forme biquadratique définie de 4 variables non décomposable en somme de carrés. Par le même procédé que ci-dessus la forme $\psi(x, y, z, t) + u^4 + ... + w^4$ sera une forme définie de degré 4 et de m variables non décomposable en somme de carrés de formes.

En résumé, nous pouvons donner le tableau suivant :

degré	2	2 n	4	4	4	6	2n ⩾ 6
nombre de variables	m	2	3	4	m ⩾ 4	3	m ⩾ 3
décomposable	oui	oui	oui	non	non	non	non
nombre de carrés	m	2	3	/	/	/	/
Démonstration	théorie des formes quadratiques	élémentaire voir II - ch. 4	Hilbert [1]	Hilbert [1]	Hilbert [1]	Hilbert [1]	Hilbert [1]

Il est clair que ce tableau apporte une réponse à tous les cas possibles.

En 1900 à Paris, Hilbert posait son 17è problème [3] -

Il est évident qu'une somme de carrés de formes réelles est une forme définie ; cependant une forme définie ne peut pas toujours être écrite comme somme de carrés de formes.

Considérant l'ensemble des fonctions rationnelles de plusieurs variables à coefficients réels qui sont définies (i.e, qui pour tout ensemble de valeurs réelles des variables où la fonction a un sens prend une valeur positive ou nulle), cet ensemble est stable par addition et multiplication et pour le quotient.

On peut alors se poser la question suivante :

une forme réelle définie peut-elle être représentée comme quotient de deux sommes de carrés de formes réelles ?

Il serait aussi souhaitable, pour certaines questions de constructions géométriques (*), de savoir si les coefficients des formes employées dans la représentation peuvent s'exprimer rationnellement en fonction des coefficients de la forme représentée.

Quand Hilbert posait ce problème il l'avait résolu dans le cas des formes de trois variables dans son article de 1893 [2]

Il y montre qu'une forme définie de trois variables et de degré n peut se mettre sous la forme $F = \frac{\Phi^2 + \Psi^2 + X^2}{H}$ où Φ Ψ et X sont des formes de degré n-2 et où H est une forme définie de degré n-4.

En reprenant le raisonnement pour H on obtient $H = \frac{\varphi^2 + \psi^2 + x^2}{h}$ où φ ψ et x sont des formes de degré n-6 et où h est une forme définie de degré n-8.

En itérant le procédé, la dernière forme jouant le rôle de h sera une forme définie quadratique ou biquadratique de 3 variables dont on sait que c'est une somme de 3 carrés de formes réelles.

(*) Voir [4].

En tenant compte du fait qu'un produit de sommes de carrés de formes est une somme de carrés de formes, on a le théorème suivant :

Soit F une forme définie quelconque de 3 variables. On peut mettre F sous la forme :

$$F = \frac{\Phi_1^2 + \ldots\ldots + \Phi_p^2}{\varphi_1^2 + \ldots\ldots + \varphi_q^2}$$ où Φ_i et φ_j sont des formes à coefficients réels

Le 17è problème de Hilbert formulé en termes de n variables peut s'exprimer en termes de polynômes de n-1 variables.

D'autre part il suffit d'étudier la possibilité d'une telle décomposition pour des polynômes définis pour obtenir un résultat sur les fractions rationnelles définies :

Si on sait que tout polynôme P défini peut s'écrire $P = \frac{\sum_{i=1}^{p} \Phi_i^2}{\sum_{j=1}^{q} \varphi_j^2}$

où les Φ_i et φ_j sont des polynômes, une fraction rationnelle définie s'écrira $\frac{P}{Q} = \frac{1}{Q^2} PQ$, où PQ est un polynôme défini, d'où :

$$PQ = \frac{\sum_{i=1}^{p'} \Phi_i^2}{\sum_{j=1}^{q'} \varphi_j^2} \quad \text{et} \quad \frac{P}{Q} = \frac{\sum_{i=1}^{p'} \Phi_i^2}{\sum_{j=1}^{q'} (\varphi_j\, Q)^2}$$

Enfin au lieu de chercher si une fraction rationnelle définie peut s'écrire comme quotient de sommes de carrés de polynômes, on peut chercher si une fraction rationnelle définie peut s'écrire comme somme de carrés de fractions rationnelles.

En effet si

$$f = \frac{\sum_{i=1}^{p} \Phi_i^2}{\sum_{j=1}^{q} \varphi_j^2} = \frac{\sum_{i=1}^{p} \Phi_i^2 \sum_{j=1}^{q} \varphi_j^2}{(\sum_{j=1}^{q} \varphi_j^2)^2} = \frac{\sum_{k=1}^{p\times q} \psi_k^2}{(\sum_{j=1}^{q} \varphi_j^2)^2}$$

et donc

$$f = \sum_{k=1}^{pq} \left(\frac{\psi_k}{\sum_{j=1}^{q} \varphi_j^2} \right)^2$$

la réciproque est immédiate.

LE 17è PROBLEME DE HILBERT PEUT ALORS S'ENONCER SOUS LA FORME SUIVANTE :

SOIT $f(X_1, \ldots, X_n) \in \mathbb{R}(X_1, \ldots, X_n)$ *TELLE QUE*

$$\forall (x_1, \ldots, x_n) \in \mathbb{R}^n \qquad f(x_1, \ldots, x_n) \geqslant 0$$

PEUT-ON ALORS TOUJOURS ECRIRE f *SOUS LA FORME D'UNE SOMME FINIE DE CARRES DE FRACTIONS RATIONNELLES*

$$f = \sum_{i=1}^{p} g_i^2 \qquad g_i \in \mathbb{R}(X_1, \ldots\ldots, X_n) \quad ?$$

ET, SI LE PROBLEME ADMET UNE SOLUTION, COMBIEN POUR UN NOMBRE DE VARIABLES DONNÉ, n, *FAUT-IL DE CARRÉS POUR POUVOIR EXPRIMER N'IMPORTE QUELLE FRACTION RATIONNELLE DÉFINIE DE* $\mathbb{R}(X_1, \ldots, X_n)$?

On peut aussi se poser ce même problème pour des fractions rationnelles à coefficients dans Q ou encore dans d'autres corps et chercher s'il admet des solutions. C'est le point de vue qualitatif adopté dans la 1ère partie, puis s'il en admet, chercher combien de carrés sont nécessaires à la décomposition. C'est le point de vue quantitatif adopté dans la 2è partie.

Enfin, dans la troisième partie nous nous intéresserons à un problème un peu différent car le corps de fonctions considéré ne sera plus le corps des fractions rationnelles à coefficients dans un corps donné, mais $\mathbb{R}(V)$ corps des fractions de l'anneau $\mathbb{R}[V]$ quotient de $\mathbb{R}[X_1, \ldots, X_n]$ par l'idéal $\mathcal{J}(V)$ de la variété algébrique réelle irréductible V.

1ERE PARTIE

RESOLUTIONS QUALITATIVES

Nous allons tout d'abord examiner sur quels corps K il est intéressant de se poser le 17è problème de Hilbert sous cette nouvelle forme :

Soit K *un corps,* $f \in K(X_1, \ldots, X_n)$ *telle que*
$f(a_1, \ldots, a_n) \geqslant 0$ *quel que soit* $(a_1, \ldots, a_n) \in K^n$;
Alors a-t-on $f = \sum_{i=1}^{p} g_i^2$ *avec* $g_i \in K(X_1, \ldots, X_n)$?

En premier lieu remarquons que pour pouvoir parler de positivité de $f(a_1, \ldots, a_n)$ il faut bien que le corps K soit ordonné.

D'autre part il faut que les fonctions constantes définies admettent une décomposition en somme de carrés d'éléments du corps ce qui n'est possible que si les éléments positifs de K sont somme de carrés.

Si K était non commutatif (voir [9] Exercice 32), l'ensemble des éléments positifs de K contiendrait l'ensemble des sommes de produits de carrés qui contient strictement l'ensemble des sommes de carrés. Dans ce cas il est donc absurde de poser le 17è problème de Hilbert sous la forme ci-dessus puisqu'on a déjà des constantes positives non sommes de carrés d'éléments du corps.

Si K est commutatif et ordonné, on a bien sûr K ordonnable ; or dans un corps ordonnable les sommes de carrés d'éléments du corps représentent les éléments totalement positifs, c'est-à-dire les éléments de K positifs pour tout ordre de K. On en déduit que s'il existe deux ordres différents sur K, alors il existera des éléments positifs pour le premier ordre qui ne seront pas positifs pour le second et qui ne pourront donc pas être écrits comme somme de carrés d'éléments du corps K.

Le 17è problème de Hilbert tel qu'il est énoncé ci-dessus n'est donc possible que pour des corps K *commutatifs possédant un et un seul ordre*. Comme exemples de tels corps nous pouvons citer le corps des nombres réels, le corps des nombres rationnels, les corps ordonnés maximaux.

Dans un article paru en 1902, [5], Landau avait obtenu le résultat suivant :
Soit Ω un corps ordonné contenu dans $\mathbb{R}$, alors toute fonction définie $f(X) \in \Omega(X)$ peut-être représentée comme somme de carrés de fonctions de $\Omega(X)$ si et seulement si les nombres positifs de Ω sont représentables comme somme de carrés d'éléments de Ω.

Ceci donne donc une réponse positive au problème dans le cas des corps contenus dans $\mathbb{R}$ et n'ayant qu'un seul ordre et des fonctions définies d'une seule variable.

Ce résultat n'est en fait qu'un cas particulier du résultat qu'Artin a obtenu en 1926 et que nous allons étudier au chapitre 1.
Au chapitre 2, nous étudierons une résolution du 17è problème par une méthode de Logique qui nous permettra de retrouver en partie le résultat d'Artin mais aussi d'obtenir un résultat nouveau.
Au chapitre 3, nous étudierons un contre exemple d'un théorème de Lang qui donnait un résultat plus général englobant les précédents.
Au chapitre 4, nous obtiendrons un résultat plus général qui permettra de retrouver les résultats des chapitres 1 et 2.
Au chapitre 5, nous donnerons quelques remarques sur le cas particulier de la décomposition en somme de carrés des polynômes définis.

CHAPITRE 1 RESOLUTION QUALITATIVE D'ARTIN (1926)

THEOREME 1 : THEOREME D'ARTIN

Soit K *un sous corps de* $\mathbb{R}$ *n'ayant qu'un seul ordre.*
Soit $f \in K(X_1, \ldots, X_n)$ *telle que pour tout*
$(x_1, \ldots, x_n) \in \mathbb{Q}^n$ *où* f *a un sens on ait* $f(x_1, \ldots, x_n) \geqslant 0$
Alors $f = g_1^2 + g_2^2 + \ldots + g_s^2$ *où* $g_i \in K(X_1, \ldots, X_n)$

Remarquons que l'hypothèse peut sembler plus faible que celle énoncée dans le 17è problème car f est positive sur les éléments de $\mathbb{Q}^n$ seulement. Mais en fait cette hypothèse entraîne f positive sur les éléments de K^n où f est définie car $\mathbb{Q} \subset K \subset \mathbb{R}$, et si f est définie en un point de K^n elle est définie et continue dans un voisinage de ce point et $\mathbb{Q}$ est dense dans $\mathbb{R}$ donc dans K.

Ce théorème nous permet donc de répondre par l'affirmative au 17è problème de Hilbert dans le cas des corps $\mathbb{Q}$ et $\mathbb{R}$ et de tous les corps intermédiaires qui possèdent un et un seul ordre tels que par exemple le corps de tous les nombres algébriques réels, les corps de nombres algébriques réels dont tous les autres corps conjugués sont non réels.

Pour démontrer son théorème Artin utilise la proposition suivante :

Proposition

Soit K *un corps de nombres réels ordonné par l'ordre induit par celui de* $\mathbb{R}$*;*
soit $K(X_1, \ldots, X_n)$ *extension ordonnée de* K. *Soient* $f_1, \ldots, f_k$
des fonctions de $K(X_1, \ldots, X_n)$
Alors il existe $(a_1, \ldots, a_n) \in \mathbb{Q}^n$ *tel que pour tout* j , $f_j(X_1, \ldots, X_n)$
est défini en $(a_1, \ldots, a_n)$ *et le signe de* f_j *pour l'ordre de* $K(X_1, \ldots, X_n)$
et celui de $f_j(a_1, \ldots, a_n)$ *pour l'ordre de* K *sont les mêmes.*

Démontrons alors le théorème en admettant cette proposition :
Soit f définie de $K(X_1, \ldots, X_n)$ non somme de carrés dans $K(X_1, \ldots, X_n)$. Alors f est non totalement positif dans $K(X_1, \ldots, X_n)$ et il existe un ordre sur $K(X_1, \ldots, X_n)$ tel que $f < 0$.

Comme K n'a qu'un seul ordre, $K(X_1, \ldots, X_n)$ muni de l'ordre pour lequel $f < 0$ est extension ordonnée de K. Appliquons alors la proposition ci-dessus ; on en déduit qu'il existe $(a_1, \ldots, a_n) \in \mathbb{Q}^n$ tel que $f(a_1, \ldots, a_n)$ ait même signe que f et comme $f < 0$, $f(a_1, \ldots, a_n) < 0$ ce qui est contraire à l'hypothèse.

Pour démontrer la proposition, nous allons d'abord en transformer l'énoncé après avoir donné quelques définitions. (*).

Soit $K(X_1, \ldots, X_n)$ extension ordonnée de K ; soit L un sous-corps de K ; soit $f \in K(X_1, \ldots, X_n)$; on appelle L-domaine de f l'ensemble
$Dom_L(f) = \{(x_1, \ldots, x_n) \in L^n \ / \ f(x_1, \ldots, x_n) \text{ défini}\}$

On appelle Bande de f sur L l'ensemble
$Bd_L(f) = \{(a_1, \ldots, a_n) \in Dom_L(f) \ / \ f(a_1, \ldots, a_n) > 0\}$

(*) On peut trouver la démonstration originale du théorème dans Artin [6] en allemand, et sa transcription en anglais dans [7]. La démonstration que nous donnons ici est celle de M. Ribenboim et sera publiée dans [8]. Son schéma est le même que celui de la démonstration d'Artin.

On appelle ensemble L-distingué les intersections finies de bandes sur L correspondant à des fonctions positives pour l'ordre de $K(X_1, \ldots, X_n)$, c'est-à-dire les ensembles E de L^n suivants :

$$E = \bigcap_{i=1}^{k} Bd_L(f_i) \qquad \text{avec} \qquad f_i > 0$$

La proposition peut alors s'énoncer sous la forme suivante :

Proposition :
Soit K *un sous-corps de* $\mathbb{R}$ *muni de l'ordre induit par celui de* $\mathbb{R}$.
Soit $K(X_1, \ldots, X_n)$ *extension ordonnée de* K, *alors tout ensemble* Q*-distingué est non vide.*

La proposition est évidemment vraie pour n = 0 car alors les seules fonctions sont les fonctions constantes. Supposons la proposition vraie pour n et démontrons la pour n + 1.

La démonstration utilise deux lemmes que nous démontrerons d'abord.

Lemme 1

Si $f \in K(X_1, \ldots, X_n)[Y]$ a $t \geqslant 0$ racines distinctes dans un corps ordonné maximal extension algébrique de $K(X_1, \ldots, X_n)$ alors il existe un Q-ensemble distingué E tel que

1. $\emptyset \neq E \subset Dom_Q(f)$ (*)
2. Si $(a_1, \ldots, a_n) \in E$ alors $\alpha_{(a_1, \ldots, a_n)}(f) \in K[Y]$ a t racines distinctes dans un corps ordonné maximal extension algébrique de K, où $\alpha_{(a_1, \ldots, a_n)}$ désigne l'extension de l'homomorphisme d'anneaux défini ci-après :

(*) Lorsque $g \in K(X_1, \ldots, X_n)[Y]$ on appelle Q-domaine de g, par abus de langage, l'ensemble de Q^n où tous les coefficients de g sont définis. D'après la définition le Q-domaine de g devrait être un ensemble de Q^{n+1} ; mais il est évident que g est défini quelle que soit la valeur dans Q donnée à Y. La dernière composante des éléments du Q-domaine de g n'a donc aucun intérêt, c'est pourquoi nous avons appelé Q-domaine de g un ensemble de Q^n afin de ne pas surcharger les démonstrations.

Soit $(a_1, \ldots, a_n)$ donné dans $\mathbb{Q}^n$.

Soit $K[X_1, \ldots, X_n]_S = \{f \in K(X_1, \ldots, X_n) \ / \ (a_1, \ldots, a_n) \in Dom_L(f)\}$

C'est aussi le sous-anneau de $K(X_1, \ldots, X_n)$ noté usuellement $S^{-1} K[X_1, \ldots, X_n]$ où S est la partie multiplicative

$S = \{h \in K[X_1, \ldots, X_n] / \ h(a_1, \ldots, a_n) \neq 0 \}$

On désigne par $\alpha_{(a_1, \ldots, a_n)}$ l'homomorphisme d'anneau suivant ;

$$K[X_1, \ldots, X_n]_S \longrightarrow K$$

$$f \longrightarrow f(a_1, \ldots, a_n)$$

Cet homomorphisme s'étend en un homomorphisme de $K[X_1, \ldots, X_n]_S [Y] \longrightarrow K[Y]$ par application de $\alpha_{(a_1, \ldots, a_n)}$ aux coefficients du polynome en Y.

Démonstration du lemme 1

Soit $f = Y^m + Y^{m-1} \varphi_1 + \ldots + \varphi_m \qquad \varphi_i \in K(X_1, \ldots, X_n)$

Soit la suite canonique de f :

$f_o = f \qquad f_1 = f' \qquad f_o = q_o f_1 - f_2$

$\ldots \qquad f_{j-1} = q_{j-1} f_j - f_{j+1} \quad \ldots f_{s-1} = q_{s-1} f_s .$

avec $\quad d^o f_{k+1} < d^o f_k$

Posons (*) $\quad \eta = 1 + m + \sum_{i=1}^{m} \varphi_i^2 \qquad \eta \in K(X_1, \ldots, X_n)$

Soit $\{ \psi_1 , \ldots, \psi_r \}$ l'ensemble d'éléments de $K(X_1, \ldots, X_n)$ formé des valeurs absolues de tous les coefficients non nuls des f_j et q_j et des valeurs absolues des $f_j(X_1, \ldots, X_n, -\eta)$ et $f_j(X_1, \ldots, X_n, \eta)$ non nuls.

(*) Voir D.E.A. page 27 ; ou propriété p, page 160 [8].

Donc $\forall j,\ \psi_j > 0.\quad \psi_j \in K(X_1, \ldots, X_n)$

la proposition étant par hypothèse vraie pour n variables, l'ensemble Q-distingué

$$E = \bigcap_{j=1}^{r} Bd_{Q}\ \psi_j \quad \text{est non vide.}$$

Soit $(a_1, \ldots, a_n)$ appartenant à E, alors $(a_1, \ldots, a_n)$ appartient au Q-domaine de tous les coefficients non nuls de f_j et de q_j donc $(a_1, \ldots, a_n)$ appartient au Q-domaine de f_j et de q_j et donc au Q-domaine de f et on a bien :

$$E \subset Dom_{Q}(f)$$

Soit $(a_1, \ldots, a_n)$ donné dans E

Notons $\bar{h}$ l'image de $h \in K[X_1, \ldots, X_n]_S\ [Y]$

par l'extension de $\alpha_{(a_1, \ldots, a_n)}$ vue ci-dessus.

Alors :

$\bar{f} = \bar{f}_o = Y^m + \varphi_1(a_1, \ldots, a_n)\ Y^{m-1} + \ldots + \varphi_m(a_1, \ldots, a_n)$

$\bar{f}' = \bar{f}_1 = (\bar{f})'$

.....

$\bar{f}_{j-1} = \bar{q}_{j-1}\ \bar{f}_j - \bar{f}_{j+1}$

..... $\bar{f}_{s-2} = \bar{q}_{s-2}\ \bar{f}_{s-1} - \bar{f}_s$ et enfin

$\bar{f}_{s-1} = \bar{q}_{s-1}\ \bar{f}_s$ enfin $\deg \bar{f}_{k+1} < \deg \bar{f}_k$ pour tout k de 0 à s-1.

La suite $(\bar{f}_o, \bar{f}_1, \ldots, \bar{f}_s)$ est donc la suite canonique de $\bar{f}$.

On sait que toute racine de f dans une extension algébrique ordonnée maximale $K(X_1, \ldots, X_n)$ de $K(X_1, \ldots, X_n)$ est dans l'intervalle $(-\eta, +\eta)$ et que, d'après le théorème de Sturm, le nombre de racines dans cet intervalle, qui est donc t d'après l'hypothèse, peut s'exprimer par :

$t = V(f_o(-\eta), f_1(-\eta), \ldots, f_s(-\eta)) - V(f_o(\eta), f_1(\eta), \ldots, f_s(\eta))$

V représentant le nombre de changements de signe dans chacune des deux suites.

Remarquons alors que $f_j(-\eta)$ et son image $\alpha_{(a_1,\ldots,a_n)} f_j(-\eta) = \bar{f}_j(-\bar{\eta})$ ont le même signe :

En effet $(a_1, \ldots, a_n) \in Bd_Q \ |f_j(-\eta)|$

. Si $f_j(-\eta) > 0$ alors $|f_j(-\eta)| = f_j(-\eta)$

et donc $(a_1, \ldots, a_n) \in Bd_Q \ f_j(-\eta)$ ce qui veut dire que

$\alpha_{(a_1, \ldots, a_n)} f_j(-\eta) > 0$

. Si $f_j(-\eta) < 0$ alors $|f_j(-\eta)| = - f_j(-\eta)$

et donc $(a_1, \ldots, a_n) \in Bd_Q - f_j(-\eta)$

donc $\alpha_{(a_1, \ldots, a_n)} -f_j(-\eta) > 0$ donc $\alpha_{(a_1, \ldots,a_n)} f_j(-\eta) < 0$.

On en déduit que

$V(f_o(-\eta), \ldots, f_s(-\eta)) = V(\bar{f}_o(-\bar{\eta}), \ldots, \bar{f}_s(-\bar{\eta}))$

Et le même raisonnement pour $f_j(\eta)$ permettrait de montrer que

$V(f_o(\eta), \ldots, f_s(\eta)) = V(\bar{f}_o(\bar{\eta}), \ldots., \bar{f}_s(\bar{\eta}))$

Remarquons alors que toute racine de $\bar{f}$ dans une extension algébrique ordonnée maximale $\tilde{K}$ de K est dans l'intervalle $(-\bar{\eta}, \bar{\eta})$ et le nombre des racines dans cet intervalle est donné par le théorème de Sturm. Finalement, le nombre de racines de $\bar{f}$ dans $\tilde{K}$ est

$V(\bar{f}_o(-\bar{\eta}), \ldots, \bar{f}_s(-\bar{\eta})) - V(\bar{f}_o(\bar{\eta}), \ldots, \bar{f}_s(\bar{\eta}))$ qui est donc aussi

$V(f_o(-\eta), \ldots, f_s(-\eta)) - V(f_o(\eta), \ldots., f_s(\eta))$

C'est donc bien t.

Lemme 2

Soient $f_1, \ldots, f_t$ des polynômes unitaires, non nécessairement distincts, de $K(X_1, \ldots, X_n)[Y]$. Soient $\rho_1, \ldots, \rho_t$ des éléments d'une extension algébrique ordonnée maximale $\widetilde{K(X_1, \ldots, X_n)}$ de $K(X_1, \ldots, X_n)$ tels que ρ_j soit racine de f_j et que $\rho_1 < \rho_2 < \ldots < \rho_t$.

Alors il existe un $\mathbb{Q}$-ensemble distingué E tel que :

1. $\emptyset \neq E \subset \bigcap_{i=1}^{t} \mathrm{Dom}_{\mathbb{Q}}(f_i)$

2. Si $(a_1, \ldots, a_n) \in E$ et $\bar{f}_j = \alpha_{(a_1, \ldots, a_n)} f_j$, alors il existe $\beta_1, \beta_2, \ldots, \beta_t$ dans une extension algébrique ordonnée maximale $\tilde{K}$ de K tels que β_j soit racine de $\bar{f}_j$ avec $\beta_1 < \beta_2 < \ldots < \beta_t$.

Démonstration du lemme 2

$\rho_j < \rho_{j+1}$ donc $\rho_{j+1} - \rho_j > 0$ dans $\widetilde{K(X_1, \ldots, X_n)}$ on en déduit que :

$\sqrt{\rho_{j+1} - \rho_j} \in \widetilde{K(X_1, \ldots, X_n)}$

Soit alors le corps

$$L = K(X_1, \ldots, X_n)(\rho_1, \ldots, \rho_t, \sqrt{\rho_2 - \rho_1}, \ldots, \sqrt{\rho_t - \rho_{t-1}})$$

L est une extension finie de $K(X_1, \ldots, X_n)$ qui est de caractéristique nulle donc il existe un élément primitif θ tel que

$$L = K(X_1, \ldots, X_n)(\theta)$$

Soit $g \in K(X_1, \ldots, X_n)[Y]$ le polynôme minimal de θ.

D'après le lemme 1, il existe un $\mathbb{Q}$-ensemble distingué E' tel que :

1. $\emptyset \neq E' \subset \mathrm{Dom}_{\mathbb{Q}}(g)$
2. Si $(a_1, \ldots, a_n) \in E'$ et $\bar{g} = \alpha_{(a_1, \ldots, a_n)}(g)$ alors $\bar{g}$ a une racine γ dans $\tilde{K}$.

Puisque $\rho_i \in L$ et que $\sqrt{\rho_{j+1} - \rho_j} \in L$ on peut poser

$\rho_i = p_i(\theta)$ et $\sqrt{\rho_{j+1} - \rho_j} = q_j(\theta)$ avec p_i et q_j appartenant à

$K(X_1, \ldots, X_n)[Y]$. On sait que ρ_i est racine de f_i donc on déduit de

$f_i(X_1, \ldots, X_n, \rho_i) = 0$ que $f_i(X_1, \ldots, X_n, p_i(\theta)) = 0$

Posons $v_i \in K(X_1, \ldots, X_n)[Y]$ $v_i = f_i(X_1, \ldots, X_n, p_i(X_1, \ldots, X_n, Y))$

Alors $v_i(\theta) = 0$; comme g est le polynôme minimal de θ on a

$v_i = g \tilde{v}_i$ dans $K(X_1, \ldots, X_n)[Y]$

De même $\rho_{j+1} - \rho_j = (q_j(\theta))^2$ donc

$p_{j+1}(\theta) - p_j(\theta) = (q_j(\theta))^2$. Soit alors $\ell_j = p_{j+1} - p_j - (q_j)^2$

Nous avons $\ell_j(\theta) = 0$ et donc $\ell_j = g \tilde{\ell}_j$ dans $K(X_1, \ldots, X_n)[Y]$

Remarquons encore que $q_j(\theta) \neq 0$ car $\rho_{j+1} \neq \rho_j$ donc $\frac{1}{q_j(\theta)} \in L$ et il existe
$m_j \in K(X_1, \ldots, X_n)[Y]$ tel que $m_j(\theta) = \frac{1}{q_j(\theta)}$.
Soit alors $u_j = m_j q_j - 1$ dans $K(X_1, \ldots, X_n)[Y]$
$u_j(\theta) = 0$ et donc $u_j = g \tilde{u}_j$ dans $K(X_1, \ldots, X_n)[Y]$

Considérons alors l'ensemble d'éléments de $K(X_1, \ldots, X_n)$: $\{\Psi_1, \ldots, \Psi_r\}$
formé des valeurs absolues des coefficients non nuls des polynômes
f_i, g, v_i, $\tilde{v}_i$, p_j, q_j, ℓ_j, $\tilde{\ell}_j$, m_j, u_j, $\tilde{u}_j$.

Soit $E'' = \bigcap_{i=1}^{r} Bd_{\mathbb{Q}}(\Psi_i)$

Considérons $E = E' \cap E''$, E est un ensemble $\mathbb{Q}$.distingué et d'après
l'hypothèse d'induction E est non vide.

D'autre part d'après la construction $E \subset E''$ et $E'' \subset \mathrm{Dom}_Q(f_i)$ pour tout i, donc $\emptyset \neq E \subset \bigcap_{i=1}^{t} \mathrm{Dom}_Q(f_i)$

Soit alors $(a_1, \ldots, a_n) \in E$; en appliquant l'homomorphisme $\alpha_{(a_1, \ldots, a_n)}$ on a $\bar{v}_i = \bar{g}\,\tilde{v}_i$, $\bar{\ell}_j = \bar{g}\,\tilde{\ell}_j$, $\bar{u}_j = \bar{g}\,\tilde{u}_j$

On en déduit puisque γ est racine de $\bar{g}$ dans $\tilde{K}$ que $\bar{v}_i(\gamma) = 0$ et donc que $\bar{f}_i$ possède une racine $\beta_i = \bar{p}_i(\gamma)$

De même $\bar{\ell}_j(\gamma) = 0$ donc $\bar{p}_{j+1}(\gamma) - \bar{p}_j(\gamma) = (q_j(\gamma))^2$ et donc $\beta_{j+1} - \beta_j = (q_j(\gamma))^2 \geqslant 0$

Enfin $\bar{u}_j(\gamma) = 0$, c'est-à-dire $\bar{m}_j(\gamma)\bar{q}_j(\gamma) = 1$ ce qui entraîne $\bar{q}_j(\gamma) \neq 0$ et donc $\beta_{j+1} - \beta_j > 0$ soit $\beta_j < \beta_{j+1}$.

Démonstration de la proposition

Tout d'abord montrons que démontrer la proposition pour $f_1, \ldots, f_k$ appartenant à $K(X_1, \ldots, X_n)[Y]$ suffit pour la démontrer pour des fonctions de $K(X_1, \ldots, X_n, Y)$. En effet, soit $f_i = \frac{g_i}{h_i}$, g_i et h_i dans $K(X_1, \ldots, X_n)[Y]$. S'il existe $(a_1, \ldots, a_{n+1}) \in \bigcap_{i=1}^{k} \mathrm{Bd}_Q(g_i\ h_i)$ alors on a $g_i(a_1, \ldots, a_{n+1})\ h_i(a_1, \ldots, a_{n+1}) > 0$ pour tout $i = 1, \ldots, k$ donc $h_i(a_1, \ldots, a_{n+1}) \neq 0$ d'où $\frac{g_i(a_1, \ldots, a_{n+1})}{h_i(a_1, \ldots, a_{n+1})}$ est défini.

De plus nous pouvons écrire pour tout i :

$$\frac{g_i(a_1, \ldots, a_{n+1})}{h_i(a_1, \ldots, a_{n+1})} = \frac{1}{[h_i(a_1, \ldots, a_{n+1})]^2} \times g_i(a_1, \ldots, a_{n+1})\ h_i(a_1, \ldots, a_{n+1})$$

Comme $g_i(a_1, \ldots, a_{n+1})\ h_i(a_1, \ldots, a_{n+1}) > 0$ on a aussi

$$\frac{g_i(a_1, \ldots, a_{n+1})}{h_i(a_1, \ldots, a_{n+1})} > 0 \quad \text{et donc} \quad (a_1, \ldots, a_{n+1}) \in \bigcap_{i=1}^{k} \mathrm{Bd}_Q\left(\frac{g_i}{h_i}\right)$$

Soient donc $f_1, \ldots, f_k$ des éléments de $K(X_1, \ldots, X_n)[Y]$ avec $f_i > 0$ pour $i = 1, \ldots, k$

Soient R' une extension algébrique ordonnée maximale de $K(X_1, \ldots, X_n, Y)$ et $R \subset R'$ une extension ordonnée maximale algébrique de $K(X_1, \ldots, X_n)$.

Pour tout i on peut décomposer f_i sous la forme
$f_i = \varphi_i\, p_{i1}^{e_{i1}}\, p_{i2}^{e_{i2}} \ldots p_{iq_i}^{e_{iq_i}}$ où $e_{ij} \geq 1$, $\varphi_i \in K(X_1, \ldots, X_n)$

et où les $p_{ij} \in K(X_1, X_2, \ldots, X_n)[Y]$ sont des polynômes irréductibles unitaires distincts. Puisque par hypothèse f_i est positif, on peut avoir $\varphi_i > 0$ et $p_{ij} > 0$.

Si on a un élément $(a_1, \ldots, a_{n+1})$ de $\mathbb{Q}^{n+1}$ tel que

$$(a_1, \ldots, a_{n+1}) \in \left(\bigcap_i Bd_{\mathbb{Q}}(\varphi_i)\right) \cap \left(\bigcap_{i,j} Bd_{\mathbb{Q}}(p_{ij})\right)$$

alors $\varphi_i(a_1, \ldots, a_n) > 0$ pour tout i et $p_{ij}(a_1, \ldots, a_{n+1}) > 0$ pour tout i et tout j donc $f_i(a_1, \ldots, a_{n+1}) > 0$ pour $i = 1, \ldots, k$ et donc $(a_1, \ldots, a_{n+1}) \in \bigcap_{i=1}^{k} Bd_{\mathbb{Q}}(f_i)$

Il suffit donc de trouver un élément $(a_1, \ldots, a_{n+1})$ appartenant à
$\left(\bigcap_i Bd_{\mathbb{Q}}\, \varphi_i\right) \cap \left(\bigcap_{ij} Bd_{\mathbb{Q}}(p_{ij})\right)$

Soit Φ un sous-ensemble fini de $K(X_1, \ldots, X_n)$ et P un sous-ensemble fini de $K(X_1, \ldots, X_n)[Y]$ tels que si $\varphi \in \Phi$ alors $\varphi > 0$ et si $p \in P$ alors p est irréductible, unitaire et $p > 0$.

Remarquons que les racines de chaque polynôme p sont distinctes puisque p est irréductible, et K de caractéristique nulle ; de plus si $p \neq p'$ alors p et p' n'ont pas de racine commune.

Soient $\rho_1, \ldots, \rho_t$ les racines distinctes dans R des polynômes $p \in P$, numérotées de telle sorte que $\rho_1 < \rho_2 \ldots < \rho_t$. Soit alors pour chaque i p_i l'unique polynôme de p tel que $p_i(\rho_i) = 0$. (Les p_i ne sont alors pas nécessairement distincts).
L'ensemble $P - \{p_1 \ldots p_t\}$ est l'ensemble des polynômes $p \in P$ qui n'ont aucune racine dans R. Soient $p_{t+1} \ldots p_u$ ces polynômes.

Si $t = 0$ posons $E' = E'' = \mathbb{Q}^n$

Si $t > 0$ soit q le produit des polynômes distincts parmi les polynômes $p_1, \ldots, p_t$. Les racines de q sont donc exactement $\rho_1, \ldots, \rho_t$.

. Appliquons le lemme 1 à q : Il existe un ensemble $\mathbb{Q}$-distingué E' tel que

1. $\emptyset \neq E' \subset \mathrm{Dom}_{\mathbb{Q}}\, q$.

2. Si $(a_1, \ldots, a_n) \in E'$ alors $\bar{q} = \alpha_{(a_1, \ldots, a_n)}(q)$ a exactement t racines distinctes dans $\tilde{K}$.

. Appliquons le lemme 2 aux p_j, j variant de 1 à t : il existe un ensemble $\mathbb{Q}$-distingué E'' tel que

1. $\emptyset \neq E'' \subset \bigcap_{j=1}^{t} \mathrm{Dom}_{\mathbb{Q}}(p_j)$

2. Si $(a_1, \ldots, a_n) \in E''$ et si $\bar{p}_j = \alpha_{(a_1, \ldots, a_n)}(p_j)$ il existe $\beta_1, \ldots, \beta_t$ dans $\tilde{K}$ tels que β_j soit racine de $\bar{p}_j$ et tels que $\beta_1 < \beta_2 < \ldots < \beta_t$.

. Appliquons encore le lemme 1 aux p_j $j = 1, \ldots t, t+1, \ldots u$: il existe un ensemble $\mathbb{Q}$-distingué E_j pour tout j tel que :

1. $\emptyset \neq E_j \subset \mathrm{Dom}_{\mathbb{Q}}(p_j)$

2. Si $(a_1, \ldots, a_n) \in E_j$ et si $\bar{p}_j = \alpha_{(a_1, \ldots, a_n)}(p_j)$

alors le nombre de racines de p_j dans R est égal au nombre de racines de $\bar{p}_j$ dans $\tilde{K} \subset \mathbb{R}$.

. Soit Ψ l'ensemble des valeurs absolues des coefficients non nuls des polynômes $p \in P$ et soit

$$E''' = \left[\bigcap_{\varphi \in \Phi} Bd_{\mathbb{Q}}(\varphi)\right] \cap \left[\bigcap_{\psi \in \Psi} Bd_{\mathbb{Q}}(\psi)\right]$$

Soit alors $E = E' \cap E'' \cap E''' \cap E_1 \cap \ldots \cap E_u$

E est évidemment un ensemble $\mathbb{Q}$-distingué et donc d'après l'hypothèse de récurrence $E \neq \emptyset$.

Soit $(a_1, \ldots, a_n) \in E$. Nous allons déterminer $d \in \mathbb{Q}$ tel que

$$(a_1, \ldots, a_n, d) \in \left[\bigcap_{\varphi \in \Phi} Bd_{\mathbb{Q}}(\varphi)\right] \cap \left[\bigcap_{p \in P} Bd_{\mathbb{Q}}(p)\right]$$

Si $p \in P$, $p \in K(X_1, \ldots, X_n)[Y]$ et nous pouvons factoriser dans $R[Y]$ sous la forme $p = (Y - \rho_{j_1})(Y - \rho_{j_2}) \ldots (Y - \rho_{j_r}) v_1 \ldots v_s \begin{cases} r \geqslant 0 \\ s \geqslant 0 \end{cases}$

où les v_i sont des polynômes du second degré unitaires et irréductibles dans $R[Y]$ avec par choix $\rho_{j_1} < \rho_{j_2} < \ldots < \rho_{j_r}$.

R est un corps ordonné maximal donc chaque v_i peut se mettre sous la forme $v_i = (Y - \sigma_i)^2 + \tau_i^2$ où σ_i et τ_i appartiennent à R (avec $\tau_i \neq 0$ puisque v_i n'a pas de racine dans R). On a donc dans R' : $v_i > 0$.

Soit $(a_1, \ldots, a_n) \in E \subset \bigcap_{\psi \in \Psi} Bd_{\mathbb{Q}}(\Psi)$ les coefficients de p sont donc définis sur $(a_1, \ldots, a_n)$. Appliquons alors $\alpha_{(a_1, \ldots, a_n)}$ aux polynômes p de P.

a) $r > 0$ alors $p = p_{j_1} = \ldots = p_{j_r}$ avec $j_1 < \ldots < j_r \leqslant t$ alors $\bar{p} = \bar{p}_{j_1} = \ldots = \bar{p}_{j_r}$

se factorise en $\bar{p} = (Y - \beta_{j_1})(Y - \beta_{j_2}) \ldots (Y - \beta_{j_r}) \bar{v}_1 \ldots \bar{v}_s$

dans $\tilde{K}[Y] \subset \mathbb{R}[Y]$

Puisque on sait que le nombre de racines de $\bar{p}$ dans $\tilde{K}$ est égal à celui de p dans R alors on a chaque $\bar{v}_i$ qui est un polynôme du second degré irréductible dans $\tilde{K}[Y]$. $\bar{v}_i$ peut alors se mettre sous la forme :

$\bar{v}_i = (Y - b_i)^2 + c_i^2$ $\quad b_i$ et c_i appartenant à $\tilde{K} \subset \mathbb{R}$ et on a donc $\bar{v}_i(d) > 0$ quel que soit $d \in \tilde{K}$.

Chaque racine ρ_j, $\rho_j \in R$, est algébrique sur $K(X_1, \ldots, X_n)$; mais Y est transcendant sur $K(X_1, \ldots, X_n)$; On en déduit que $\rho_j \neq Y$ pour $j = 1, \ldots, t$ et que dans le corps R' nous avons

$Y < \rho_1$ ou $\rho_h < Y < \rho_{h+1}$ ou $\rho_t < Y$.

$\mathbb{R}$ est archimédien, $\mathbb{Q}$ est dense dans $\mathbb{R}$, alors il existe donc $d \in \mathbb{Q} \subset \tilde{K} \subset \mathbb{R}$ tel que

$d < \beta_1$ quand $Y < \rho_1$

ou $\beta_h < d < \beta_{h+1}$ quand $\rho_h < Y < \rho_{h+1}$

ou $\beta_t < d$ quand $\rho_t < Y$

Comme $p = (Y - \rho_{j_1}) \ldots (Y - \rho_{j_r}) v_1 \ldots v_s$ $\quad r \geqslant 1 \quad s \geqslant 0$

et $\bar{p} = (Y - \beta_{j_1}) \ldots (Y - \beta_{j_r}) \bar{v}_1 \ldots \bar{v}_s$

Il est clair alors que, v_i étant positif dans $K(X_1, \ldots, X_n)[Y]$ et $\bar{v}_i(d)$ dans $\tilde{K}$, p dans $K(X_1, \ldots, X_n)[Y]$ et $\bar{p}(d)$ dans $\tilde{K} \subset \mathbb{R}$ ont le même signe. Comme par hypothèse $p > 0$ on a $\bar{p}(d) > 0$.

b) $r = 0$. De même si $p = p_j$ avec $t + 1 \leqslant j \leqslant u$. Appliquons $\alpha_{(a_1, \ldots, a_n)}$ à p qui alors est factorisé sous la forme $p = v_1 \ldots v_s$ dans $R[Y]$ $\alpha_{(a_1, \ldots, a_n)}\, p = \bar{p}$ n'a aucune racine dans $\tilde{K}$ car p n'avait aucune racine dans R.

$\bar{p}$ peut donc se factoriser de la manière suivante :

$\bar{p} = \bar{v}_1 \bar{v}_2 \ldots \bar{v}_s$ où chaque $\bar{v}_i$ est un polynôme irréductible du second degré de $\tilde{K}[Y]$. On a donc aussi $\bar{v}_i(d) > 0$ quel que soit $d \in \tilde{K}$ et donc $\bar{p}(d)$ est positif dans $\tilde{K} \subset \mathbb{R}$.

Dans tous les cas $\bar{p}(d) > 0$ ce qui signifie $p(a_1, \ldots, a_n, d) > 0$ et donc $(a_1, \ldots, a_n, d) \in Bd_{\mathbb{Q}}(p)$ ceci quel que soit $p \in P$,

donc $(a_1, \ldots, a_n, d) \in \bigcap_{p \in P} Bd_{\mathbb{Q}}(p)$.

Comme $(a_1, \ldots, a_n, d)$ appartenait à $\bigcap_{\varphi \in \Phi} Bd_{\mathbb{Q}}(\Phi)$ quel que soit d.

On a bien trouvé un élément $(a_1, \ldots, a_n, d) \in \mathbb{Q}^{n+1}$ tel que

$(a_1, \ldots, a_n, d) \in (\bigcap_{\varphi \in \Phi} Bd_{\mathbb{Q}}(\varphi)) \cap (\bigcap_{p \in P} Bd_{\mathbb{Q}}(p))$

Chapitre 2 RESOLUTION PAR UNE METHODE DE LOGIQUE

Nous allons ici démontrer que le 17è problème de Hilbert admet une solution dans le cas d'un corps ordonné maximal. Le résultat était déjà connu, d'après le théorème d'Artin, dans le cas de corps ordonnés maximaux contenus dans $\mathbb{R}$.

L'origine de ce genre de démonstration se trouve dans les travaux de Tarski dont on peut énoncer ainsi l'un des résultats : "Toute propriété "élémentaire" vraie dans un corps ordonné maximal est vraie dans tout corps ordonné maximal". Robinson a repris cette théorie dans [10]. La démonstration que nous donnons ici est inspirée d'un exercice de [11].

THEOREME 2

Soit K *un corps ordonné maximal. Soit* p *appartenant à* $K[X_1, \ldots, X_n]$; *si* $p(X_1, \ldots, X_n)$ *est positif ou nul quels que soient* $(X_1, \ldots, X_n) \in K^n$, *alors il existe* $g_1, \ldots, g_k$ *dans* $K(X_1, \ldots, X_n)$ *tels que* $p = {g_1}^2 + \ldots + {g_k}^2$.

Donnons quelques définitions et lemmes préliminaires.

Définition :

On dit qu'un ensemble $\mathcal{A}$ de formules permet l'élimination des quantificateurs dans une formule F du langage $\mathcal{L}$ s'il existe une formule F' de $\mathcal{L}$, sans quantificateurs telle que $F \leftrightarrow F'$ soit une conséquence de $\mathcal{A}$.

On dira que $\mathcal{A}$ permet l'élimination des quantificateurs dans $\mathcal{L}$ si $\mathcal{A}$ permet l'élimination des quantificateurs pour toute formule F de $\mathcal{L}$.

Définition

Un ensemble de formules $\mathcal{K}$ est dit saturé dans $\mathcal{L}$ si pour toute formule close F de $\mathcal{L}$, F ou $\neg$ F est conséquence de $\mathcal{K}$.

Lemme 1

Si $\mathcal{K}$ permet l'élimination des quantificateurs dans $\mathcal{L}$, si $\mathcal{D}_{\mathcal{M}}$ est le diagramme d'un modèle $\mathcal{M}$ de $\mathcal{K}$, $(\mathcal{K}, \mathcal{D}_{\mathcal{M}})$ est saturé pour le langage $\mathcal{L}'$ obtenu en ajoutant aux symboles de constantes du langage $\mathcal{L}$ les éléments de l'ensemble de base de $\mathcal{M}$.

Pour la démonstration de ce lemme voir [11]

Soitalors $\mathcal{K}$ l'ensemble de formules formé par la réunion des ensembles de formules (a), (b) et (c) suivants écrits dans le langage égalitaire $\mathcal{L}$ formé de :

0 et 1 symboles de constantes,

- symbole fonctionnel à 1 variable,

+ et × symboles fonctionnels à 2 variables,

> 0 symbole relationnel à 1 variable.

(a)
$$\begin{cases} \Lambda x \ \Lambda y \ \Lambda z \ ((x + y) + z = x + (y + z)) \\ \Lambda x \ \Lambda y \ (x + y = y + z) \\ \Lambda x \ (x + 0 = x) \\ \Lambda x \ (x + (- x) = 0) \end{cases}$$

$$\begin{cases} \Lambda x \ \Lambda y \ \Lambda z \ ((xy)z = x(yz)) \\ \Lambda x \ \Lambda y \ (xy = yx) \\ \Lambda x \ (x \ 1 = x) \\ \Lambda x \ \forall y \ ((x = 0) \vee (xy = 1)) \\ \Lambda x \ \Lambda y \ \Lambda z (x(y+z) = xy + xz)) \\ \neg (0 = 1) \end{cases}$$

(b) $\Lambda\, x\ \Lambda\, y((x > 0) \wedge (y > 0) \rightarrow x + y > 0)$

$\Lambda\, x\ \Lambda\, y((x > 0) \wedge (y > 0) \rightarrow x\, y > 0)$

$\Lambda\, x\ (x = 0 \vee x > 0 \vee -x > 0)$

$\Lambda\, x\ \neg\ (\ (x > 0) \wedge (-x > 0)\ )$

(c) $\Lambda\, x\ \bigvee y\ \ ((x = y^2) \vee (-x = y^2))$

L'ensemble des formules

$\Lambda\, x_o \ldots \Lambda\, x_{2n}\ \bigvee x\ \ (x_o + x_1\, x + \ldots + x_{2n}\, x^{2n} + x^{2n+1} = 0)$

écrites pour chaque n positif.

Nous utilisons ici les abréviations qui consistent à appeler n le terme $1 + 1 + \ldots + 1$ et à appeler t^p le terme $t \times \ldots \times t$. Il est clair que les formules de (c) peuvent s'écrire dans le langage $\mathcal{L}$.

Il est clair que les modèles de (a) sont les corps commutatifs, que les modèles de (a) $\cup$ (b) sont les corps commutatifs ordonnés et que les modèles de $\mathcal{A}$ = (a) $\cup$ (b) $\cup$ (c) sont les corps commutatifs ordonnés maximaux. (on peut pour cela se reporter à la caractérisation d'Euler Lagrange dans [9]).

Lemme 2

$\mathcal{A}$ permet l'élimination des quantificateurs dans $\mathcal{L}$.

Pour la démonstration de ce lemme, voir [11].

NOTE

Nous utiliserons la notation $x \geqslant 0$ qui remplacera la formule $(x > 0) \vee (x = 0)$

Démonstration du théorème

Soit D_K le diagramme de K, K étant un modèle de $\mathcal{A}$ (car ordonné maximal) et $\mathcal{A}$ permettant l'élimination des quantificateurs, on en déduit que $(\mathcal{A}, D_K)$ est saturé.

Comme la formule $\Lambda x_1 \ldots \Lambda x_n \ (p(x_1, \ldots, x_n) \geqslant 0)$ est satisfaite dans K qui est un modèle de $(\mathcal{A}, D_K)$ alors elle est satisfaite dans tout modèle de $(\mathcal{A}, D_K)$.
Un modèle de $(\mathcal{A}, D_K)$ est un corps ordonné maximal extension ordonnée de K.

Mais toute extension ordonnée de K se plonge dans une extension ordonnée maximale où la formule $\Lambda x_1 \ \ldots \ \Lambda x_n (p(x_1, .., x_n) \geqslant 0)$ est vraie.
Cette formule est donc aussi vraie dans toute extension ordonnée de K.

K n'a qu'un seul ordre donc tout corps ordonné contenant K est extension ordonnée de K et on a donc finalement :

$\Lambda x_1 \ \ldots \ \Lambda x_n \ (p(x_1, \ldots, x_n) \geqslant 0)$ qui est vraie dans tout corps ordonné contenant K.

$K(X_1, \ldots, X_n)$ est un corps ordonnable contenant K. Soit un ordre quelconque sur $K(X_1, \ldots, X_n)$ d'après ce qui précède

$\Lambda x_1 \ \ldots \Lambda x_n (p(x_1, \ldots, x_n) \geqslant 0)$ est vraie dans $K(X_1, \ldots, X_n)$.

Choisissons comme éléments de $K(X_1, \ldots, X_n)$ le n-uple
$(X_1, \ldots, X_n) \in [K(X_1, \ldots, X_n)]^n$ alors $p(X_1, \ldots, X_n) \geqslant 0$
Ceci étant vrai pour tout ordre sur $K(X_1, \ldots, X_n)$, $p(X_1, \ldots, X_n)$ est un élément totalement positif de $K(X_1, \ldots, X_n)$ et c'est donc une somme de carrés d'éléments de $K(X_1, \ldots, X_n)$.

Nous avons donc bien $p = \sum_{i=1}^{k} g_i^2$ où $g_i \in K(X_1, \ldots X_n)$

Il est clair que le fait d'avoir une décomposition en somme de carrés dans $K(X_1, \ldots, X_n)$ pour tout polynôme p positif ou nul

sur les éléments de K^n suffit pour obtenir une décomposition en somme de carrés de toute fonction définie de $K(X_1, \ldots, X_n)$ (raisonnement déjà vu p. 7)
Nous avons donc le résultat suivant :

THEOREME 2'

Soit K *un corps ordonné maximal, soit* $f \in K(X_1, \ldots, X_n)$ *telle que pour tout* $(a_1, \ldots, a_n) \in K^n$ *où* f *a un sens, alors* $f(a_1, \ldots, a_n) \geqslant 0$, *alors nous avons*

$$f = \sum_{i=1}^{p} g_i^2 \qquad g_i \in K(X_1, \ldots, X_n)$$

Chapître 3

SUR UN THEOREME DE LANG

A ce point nous avons donc obtenu que le 17è problème de Hilbert sous la forme adoptée ici, ne peut être résolu que dans un corps commutatif ordonné et n'ayant qu'un seul ordre.

Au chapître 1, nous avons vu que ce problème admettait une solution pour tout corps n'ayant qu'un seul ordre et contenu dans $\mathbb{R}$. C'est-à-dire finalement pour tout corps archimédien n'ayant qu'un seul ordre.

Au chapître 2, nous avons vu que dans le cas des corps ordonnés maximaux on pouvait se passer de l'hypothèse que le corps était contenu dans $\mathbb{R}$.

On peut alors se poser la question : l'hypothèse K n'a qu'un seul ordre, ne suffirait-elle pas pour répondre par l'affirmative au 17è problème de Hilbert ?

S. Lang dans [12] énonce le théorème suivant :

Théorème 3

Soit K un corps ordonnable n'admettant qu'un seul ordre. Soit $f(X_1, \ldots, X_n) \in K(X_1, \ldots, X_n)$ telle que pour tout $(a_1, \ldots, a_n) \in K^n$ où f a un sens, on ait $f(a_1, \ldots, a_n) \geqslant 0$. Alors f est une somme de carrés d'éléments de $K(X_1, X_2, \ldots, X_n)$

Mais ce théorème est faux et Dubois dans [13] donne un contre exemple, c'est-à-dire *un corps* F *n'ayant qu'un seul ordre et un polynôme* f *d'une variable* X *à coefficients dans ce corps qui est positif sur les éléments du corps* F *et qui n'est pas une somme de carrés d'éléments de* F(X).

C'est ce contre-exemple que nous allons étudier maintenant.

Soit $\mathbb{Q}$ le corps des nombres rationnels, soit t une indéterminée et $\mathbb{Q}(t)$ le corps des fractions rationnelles à coefficients dans $\mathbb{Q}$. Ordonnons $\mathbb{Q}(t)$ de telle façon que t soit positif et infinitésimal.

Soit alors K un corps ordonné maximal extension algébrique ordonnée de $\mathbb{Q}(t)$.

Soit F le corps formé des éléments de K qui peuvent être obtenus par une suite finie d'opérations rationnelles et d'extractions de racines carrées à partir des éléments de $\mathbb{Q}(t)$ (exactement comme dans les constructions avec la règle et le compas) qui sont positifs. Puisque tout $a > 0$ de F a une racine carrée dans F, F a un seul ordre.

Considérons alors le polynôme en X :

$$f(X) = (X^3 - t)^2 - t^3$$

f ne peut pas être une somme de carrés dans F(X) qui est contenu dans K(X), puisque f(1) et $f(t^{1/3})$ ont des signes opposés dans K.

$$f(1) = (1 - t)^2 - t^3 > 0$$

$$f(t^{1/3}) = - t^3 < 0$$

Montrons alors que f(X) est positive sur les éléments de F.

Considérons l'ensemble B de tous les éléments finis de K (u est fini si $|u| < n$ pour un certain entier n). B est un anneau de valuation de K. La valuation induite v est une mesure de l'ordre de grandeur ; $v(a) < v(b)$ signifie que $a^{-1}b$ est infinitésimal.

Soit G le groupe de valeur (en notation additive). Le caractère algébrique de K sur $\mathbb{Q}(t)$ entraîne que le rang de G est 1.

En notant que G est sans torsion, on en déduit (voir [14]) que G est isomorphe à un sous-groupe de $\mathbb{Q}$ et qu'on peut avoir $v(t) = 1$; de plus comme K est ordonné maximal les seuls polynômes irréductibles sont de degré 1 ou 2 et puisque t est positif, t admet une racine carrée dans K. Donc K contient une racine n-ième de t quel que soit n et l'application de G dans $\mathbb{Q}$ est surjective ; on peut désormais identifier G à $\mathbb{Q}$.

On peut alors vérifier aisément que si f(y) est négatif, alors $v(y) = \frac{1}{3}$. En effet :

. Si $v(y) > \frac{1}{3}$ alors $v(y^3) > 1$ donc $v(y^3) > v(t)$, ce qui signifie que $\frac{y^3}{t}$ est infinitésimal. Or $f(y) = (y^3 - t)^2 - t^3 = t^2(\frac{y^3}{t} - 1)^2 - t^3$

$$f(y) = t^2 \left[(\frac{y^3}{t})^2 - 2\frac{y^3}{t} + 1 - t\right] > 0$$

. Si $v(y) < \frac{1}{3}$ alors $v(y^3) < v(t)$ et $\frac{t}{y^3}$ est infinitésimal.

Or $f(y) = t^2\left[(\frac{y^3}{t})^2 - 2\frac{y^3}{t} + 1 - t\right]$ peut s'écrire

$f(y) = t^2 \frac{y^6}{t^2}\left[1 - 2\frac{t}{y^3} + \frac{t^2}{y^6} - \frac{t^2}{y^6} \times t\right] > 0$

Donc si $v(y) \neq \frac{1}{3}$ f(y) est positif.

Soit z un élément quelconque de F, alors z, d'après la construction de F, appartient à un corps H_r qui est le dernier terme d'une suite finie de sous-corps de K où chaque extension est quadratique :

$$\mathbb{Q}(t) = H_o \subset H_1 \subset \ldots \subset H_r$$

En appliquant la relation de ramification ($ef \leqslant n$, [15]) pour $n = 2$ on déduit que le groupe de valeur de H_i a pour indice 1 ou 2 dans le groupe de H_{i+1}. Le groupe de valeur de la valuation sur $\mathbb{Q}(t)$ étant $\mathbb{Z}$, on en déduit que nécessairement $v(z)$ est de la forme $\frac{m}{2^k}$ où $m \in \mathbb{Z}$ et $k \in \mathbb{N}$.

En remarquant que $\frac{m}{2^k}$ ne peut pas être égal à $\frac{1}{3}$, on en déduit que $f(z)$ est positif, ce qui termine la démonstration z ayant été pris quelconque dans F.

Il est donc inutile maintenant de chercher à montrer que le 17è problème de Hilbert peut être résolu pour n'importe quel corps K commutatif et n'ayant qu'un seul ordre.
On ne peut espérer qu'un résultat plus faible concernant certains de ces corps seulement.

CHAPITRE 4 RESOLUTION DANS UN CAS PLUS GENERAL

Nous démontrerons ici le théorème suivant :

Théorème 4

Soit k *un corps ordonné n'ayant qu'un seul ordre, dense (au sens de la topologie définie par les intervalles) dans sa clôture réelle* $\bar{k}$. *(C'est une extension algébrique ordonnée de* k *qui est un corps ordonné maximal).*

Soit $f \in k(X_1, \ldots, X_n)$ *telle que pour tout* $(x_1, \ldots, x_n) \in k^n$ $f(x_1, \ldots, x_n) \geqslant 0$.

Alors il existe des éléments $h_i \in k(X_1, \ldots, X_n)$ *tels que* $f = \sum_1^r h_i^2$.

Ce théorème permet de retrouver le résultat obtenu au chapitre 2 pour tout corps ordonné maximal k. En effet si k est ordonné maximal, $k = \bar{k}$ est bien dense dans $\bar{k}$.

Ce théorème permet aussi de retrouver le résultat obtenu au chapitre 1. Si k est un corps tel que $\mathbb{Q} \subset k \subset \mathbb{R}$, k n'ayant qu'un seul ordre, alors une clôture réelle de k $\bar{k}$ est telle que $k \subset \bar{k} \subset \mathbb{R}$ et k est bien dense dans $\bar{k}$. (voir [12], page 279).

Le théorème 4 est en fait un corollaire de la proposition.

Proposition 1

Soit k *un corps ordonné dense dans sa clôture réelle* $\bar{k}$; *soit* $f \in k(X_1, \ldots, X_n)$ *telle que pour tout* $(x_1, \ldots, x_n) \in k^n$, $f(x_1, \ldots, x_n) \geqslant 0$. *Alors il existe des éléments* $p_i \in k$, $p_i > 0$ *et des éléments* $g_i \in k(X_1, \ldots, X_n)$ *tels que* $f = \sum_{i=1}^{q} p_i g_i^2$.

Cette proposition est démontrée par P. Ribenboim dans un exposé fait en juin 1971 à Paris [17]

Le théorème 4 en est un corollaire immédiat : si k n'a qu'un seul ordre alors l'ensemble des éléments positifs pour cet ordre est $\Sigma\, k^2$. Donc $p_i \in k$ et $p_i > 0$ entraîne l'existence d'éléments $a_i^j \in k$ tels que :

$$p_i = \sum_{j=1}^{s} (a_i^j)^2$$

En remplaçant p_i par cette expression dans le résultat ci-dessus, nous obtenons :

$$f = \sum_{i=1}^{r} h_i^2 \quad \text{où} \quad h_i \in k(X_1, \ldots, X_n)$$

I - Définitions et résultats utilisés pour la démonstration de la proposition 1

Définition 1

Soient K/k et L/k des extensions.
Soit A un sous-anneau de K contenant k.
Tout k-homomorphisme $\sigma : A \to L$ s'appelle une L-spécialisation de K/k ; et A est le domaine de σ .

Définition 2

Soit $E \subset K$. Soit σ une L-spécialisation de K/k dont le domaine contient E.
Si K/k et L/k sont des extensions ordonnées on dira que σ préserve E lorsque :

a) $x \in E$ et $x < 0$ ==> $\sigma(x) < 0$

b) $x \in E$ et $x > 0$ ==> $\sigma(x) > 0$

Définition 3

Soit L/k une extension ordonnée. Soit K/k une extension ordonnée ; on dira que l'extension K/k est L-artinienne lorsque la propriété suivante sera vérifiée : si E est un sous-ensemble fini de K, il existe une L-spécialisation de K/k qui préserve E.

Lemme

Soit L/k une extension ordonnée telle que L soit dense dans $\bar{L}$. Alors pour tout ordre de $k(X_1, \ldots, X_n)$ $(n \geq 0)$ prolongeant l'ordre de k, l'extension $k(X_1, \ldots, X_n)/k$ est L-artinienne.

Pour $n = 0$ $\quad k(X_1, \ldots, X_n) = k$, et k/k est évidemment L-artinienne.
Le lemme est donc (par récurrence) un corollaire de la proposition suivante :

Proposition 2

Soit L/k *une extension ordonnée telle que* L *soit dense dans* $\bar{L}$. *Si* K/k *est une extension ordonnée* L-*artinienne, et* Y *une indéterminée, alors* K(Y)/k *est une extension* L-*artinienne pour tout ordre de* K(Y) *qui prolonge l'ordre de* K.

II . Démonstration de la proposition 1

Si f ne s'écrit pas $\sum_{i=1}^{q} p_i g_i^2$ alors puisque $k(X_1, \ldots, X_n)$ est une extension ordonnable de k, c'est que f n'est pas positif pour tous les ordres de $k(X_1, \ldots, X_n)$ qui prolongent celui de k (*); donc il existe un ordre sur $k(X_1, \ldots, X_n)$ prolongeant celui de k tel que $f = \frac{f_1}{f_2} < 0$.

(*) Ceci d'après les résultats d'Artin.
Si K/k est une extension ordonnable de k (c'est-à-dire qu'il existe sur K un ordre qui prolonge celui de k).
Alors $x \in K$ est positif pour tout ordre sur K prolongeant l'ordre sur k si et seulement si $x = \sum_{i=1}^{m} p_i x_i^2$ où $p_i \in k$ et $p_i > 0$ et $x_i \in K$.

Puisque k est dense dans $\bar{k}$, d'après le lemme, $k(X_1, \ldots, X_n)/k$ est k-Artinienne pour tout ordre de $k(X_1, \ldots, X_n)$ prolongeant celui de k.

Donc il existe une k-spécialisation σ de $k(X_1, \ldots, X_n)/k$ qui préserve l'ensemble $E = \{X_1, \ldots, X_n, f, f_2\}$.

Soit $x_i = \sigma(X_i)$. $x_i \in k$. Alors :

$f_2(x_1, \ldots, x_n) = f_2(\sigma(X_1), \ldots, \sigma(X_n)) = \sigma(f_2) \neq 0$ puisque $f_2 \neq 0$ dans $K(X_1, \ldots, X_n)$.

Donc f est définie sur l'élément $(x_1, \ldots, x_n) \in k^n$.

De plus $f(x_1, \ldots, x_n) = f(\sigma(X_1), \ldots, \sigma(X_n)) = \sigma(f) < 0$ pour l'ordre indiqué au début qui rend $f < 0$ dans $K(X_1, \ldots, X_n)$.

Or $f(x_1, \ldots, x_n) < 0$ en un point $(x_1, \ldots, x_n)$ où f est définie est contraire à l'hypothèse faite sur f.

Donc f peut s'écrire sous la forme :

$$f = \sum_{i=1}^{q} p_i \, g_i^2 \qquad p_i \in k,\ p_i > 0 \quad \text{et} \quad g_i \in k(X_1, \ldots, X_n)$$

III - Démonstration de la proposition 2

La démonstration utilise les deux lemmes suivants dus à Artin.

Lemme 1 :

Soit K/k et L/k des extensions ordonnées.

Soit $f = Y^m + a_1 Y^{m-1} + \ldots + a_m \in K[Y]$

un polynôme ayant exactement $t \geqslant 0$ racines distinctes dans $\bar{K}$ clôture réelle de K.

Alors il existe un ensemble fini $E \subset K$, contenant les coefficients de f et tel que, si σ est une L-spécialisation de K/k qui préserve E, alors $f^\sigma = Y^m + \sigma(a_1)Y^{m-1} + \ldots + \sigma(a_m)$ a exactement t racines distinctes dans $\bar{L}$ clôture réelle de L.

Démonstration du lemme 1 :

. Soit $(f_o, f_1, \ldots, f_s)$ la suite canonique de f, c'est-à-dire la suite $f_o = f$, $f_1 = f'$, f_2 défini par $f_o = q_o f_1 - f_2$, ... , f_{j+1} défini par $f_{j-1} = q_{j-1} f_j - f_{j+1}$ avec deg f_{j+1} < deg f_j pour j de 1 à s - 1 et enfin $f_{s-1} = q_{s-1} f_s$ (qui définirait $f_{s+1} = 0$).

Posons $\eta = 1 + m + \sum_{i=1}^{m} a_i^2$, $\eta \in K$.

. Soit E le sous-ensemble de K formé de tous les coefficients des polynômes f_j et q_j ainsi que des éléments $f_j(-\eta)$ et $f_j(\eta)$.

. Soit σ une L-spécialisation de K/k qui préserve E. Donc le domaine de σ contient E . Il est clair que la suite canonique de f^σ est $(f_o^\sigma, \ldots, f_s^\sigma)$ puisque E contient les coefficients des f_j et q_j.
On sait aussi que toute racine de f dans la clôture réelle $\bar{K}$ de K est dans l'intervalle $]-\eta, +\eta[$ (voir propriété p page 160 [8])

. Par le théorème de Sturm, le nombre de racines distinctes de f dans $]-\eta, +\eta[$ (intervalle de $\bar{K}$) est égal à la différence $V_{f,-\eta} - V_{f,\eta}$ où $V_{f,-\eta}$ est le nombre de changements de signes de la suite

$$f_o(-\eta), \quad f_1(-\eta), \ldots, f_s(-\eta)$$

et $V_{f,\eta}$ celui de la suite

$f_o(\eta), f_1(\eta), \ldots, f_s(\eta)$

Soit $t = V_{f,-\eta} - V_{f,+\eta}$

. Puisque $\sigma(\eta) = 1 + m + \sum_{i=1}^{m} \sigma(a_i)^2$, alors toute racine du polynôme f^σ dans $\bar{L}$ est dans l'intervalle $]-\sigma(\eta), \sigma(\eta)[$.

Le nombre de ces racines est égal à $V_{f^\sigma,-\sigma(\eta)} - V_{f^\sigma,\sigma(\eta)}$.

. Or la suite canonique de f^σ est on l'a vu f_o^σ, f_1^σ , ..., f_s^σ.

De plus $(f_j(-\eta))^\sigma = f_j^\sigma(-\sigma(\eta))$ et puisque $f_j(-\eta) \in E$ le signe de $f_j(-\eta)$ est le même que celui $f_j^\sigma(-\sigma(\eta))$.

Donc $V_{f^\sigma, \sigma(-\eta)} = V_{f, -\eta}$

De même on a $V_{f^\sigma, \sigma(\eta)} = V_{f,\eta}$

Donc $t = V_{f, -\eta} - V_{f,\eta} = V_{f^\sigma,\sigma(-\eta)} - V_{f^\sigma, \sigma(\eta)}$ et f^σ a bien exactement t racines distinctes dans la clôture réelle $\bar{L}$ de L.

Lemme 2 :

Soient K/k et L/k des extensions ordonnées. Soient f_1, ..., f_t des polynômes de $K[Y]$. (non nécessairement distincts). Soient $r_1 < r_2 < .... < r_t$ des éléments de $\bar{K}$ tels que r_i soit une racine de f_i. Alors il existe un ensemble fini $E \subset K$, contenant les coefficients de chaque f_i, tel que, si σ est une L-spécialisation de K/k qui préserve E, il existe des éléments b_1, ..., b_t dans $\bar{L}$ tels que $b_1 < b_2 < .... < b_t$ et b_i est racine de f_i^σ.

Démonstration du lemme 2 :

On a $r_j < r_{j+1}$ donc $\sqrt{r_{j+1} - r_j} \in \bar{K}$.

Soit $K' = K(r_1, ..., r_t, \sqrt{r_2 - r_1}, ..., \sqrt{r_t - r_{t-1}})$

K' est une extension algébrique finie de K, et il existe un élément

primitif tel que $K' = K(u)$.

Soit $g \in K[X]$ le polynôme minimal de u sur K.
D'après le lemme 1 il existe un ensemble fini $E' \subset K$, E' contenant les coefficients de g, et E' tel que si σ est une L-spécialisation de K/k préservant E' alors g^σ a une racine dans $\bar{L}$.

. Soient p_i, q_j dans $K[Y]$ tels que $r_i = p_i(u)$ et $\sqrt{r_{j+1} - r_j} = q_j(u)$.

. Soit $v_i = f_i(p_i)$ dans $K[Y]$; donc $v_i(u) = f_i(p_i(u)) = f_i(r_i) = 0$
et donc $v_i = g\, v'_i$ où $v_i \in K[Y]$ puisque g est le polynôme minimal de u sur K.

. De même soit $\ell_j = p_{j+1} - p_j - q_j^2$, alors $\ell_j(u) = 0$. Donc $\ell_j = g\, \ell'_j$ avec $\ell'_j \in K[Y]$ puisque g est le polynôme minimal de u sur K.

. Remarquons que $q_j(u) \neq 0$, donc $\frac{1}{q_j(u)} \in K' = K(u)$ et donc il existe $m_j \in K[Y]$ tel que $m_j(u) = \frac{1}{q_j(u)}$.

. Soit $z_j = m_j\, q_j - 1$; $z_j \in K[Y]$ et $z_j(u) = 0$
Donc $z_j = g\, z'_j$ où $z'_j \in K[Y]$.

Soit alors E l'ensemble union de E' et de l'ensemble des coefficients des polynômes f_i, g, p_i, q_j, v_i, v'_i, ℓ_j, ℓ'_j, m_j, z_j et z'_j.

Soit σ une L-spécialisation de K/k qui préserve E et donc aussi E'.
Alors g^σ a une racine γ dans $\bar{L}$.
De $v_i^\sigma = g^\sigma\, v'^\sigma_i$, résulte que γ est une racine de v_i^σ. Donc que $f_i^\sigma(p_i^\sigma(\gamma)) = v_i^\sigma(\gamma) = 0$.

Alors $b_i = p_i^\sigma(\gamma)$ est un élément de $\bar{L}$ et est une racine de f_i^σ.

On a aussi $\ell_j^\sigma = g^\sigma\, \ell'^\sigma_j$ et $z_j^\sigma = g^\sigma\, z'^\sigma_j$ donc γ est racine de ℓ_j^σ et de z_j^σ.
On a $\ell_j^\sigma(\gamma) = 0$ donc $b_{j+1} - b_j = p_{j+1}^\sigma(\gamma) - p_j^\sigma(\gamma) = (q_j^\sigma(\gamma))^2$

donc $b_{j+1} - b_j \geqslant 0$.

On a aussi $z_j^{\sigma}(\gamma) = 0$ donc $m_j^{\sigma}(\gamma)\, q_j^{\sigma}(\gamma) = 1$ et donc $q_j^{\sigma}(\gamma) \neq 0$

d'où $b_{j+1} \neq b_j$.

On a donc bien finalement $b_j < b_{j+1}$ dans $\bar{L}$.

Démonstration de la proposition 2

Soit E un sous-ensemble fini de K(Y), il faut montrer qu'il existe une L-spécialisation de K(Y) qui préserve E.
Il est clair qu'on peut se ramener au cas où les éléments de E sont des polynômes unitaires irréductibles de K[Y] ou des éléments de K.

. Soit $\overline{K(Y)}$ la clôture réelle de K(Y) muni d'un ordre prolongeant celui de K ; alors la clôture algébrique $\bar{K}$ de K dans $\overline{K(Y)}$ est une clôture réelle de K.

. Soient $r_1 < r_2 < \ldots < r_t$ toutes les racines dans $\bar{K}$ des polynômes appartenant à E. Indexons les polynômes de telle façon que r_i soit racine de $f_i \in E$; il n'est donc pas exclu que $f_i = f_j$ pour $i \neq j$.
Soit $g \in K[Y]$ le produit des polynômes non constants et distincts de E. Les facteurs de g sont donc unitaires irréductibles et distincts, donc toutes les racines de g sont simples et les racines de g dans $\bar{K}$ sont exactement $r_1, r_2, \ldots, r_t$.

. Soit E' un sous-ensemble fini de K tel que :
1. E' contient les coefficients de g, les éléments de $E \cap K$, et les coefficients des polynômes de E.
2. E' contient le discriminant d de g ($d = \prod_{i<j} (r_i - r_j)^2$) ; remarquons

que d est non nul puisque les racines de g sont distinctes.

3. E' est choisi (grâce aux lemmes 1 et 2) de telle façon que, si σ est une L-spécialisation de K/k qui préserve E', alors g^σ a précisément t racines distinctes dans $\bar{L}$ (qui sont toutes simples car alors le discriminant de g^σ est $d^\sigma \neq 0$), et il existe des éléments $b_1, \ldots, b_t$ dans $\bar{L}$ tels que b_i est racine de f_i^σ et $b_1 < b_2 < \ldots < b_t$.

. D'après l'hypothèse il existe une L-spécialisation σ de K/k qui préserve E' et donc la condition 3) est satisfaite par σ. Il en résulte donc que $b_1, \ldots, b_t$ sont toutes les racines de g^σ dans $\bar{L}$.

. Ecrivons $f_i = (Y - r_{i_1}) \ldots (Y - r_{i_{t(i)}}) \cdot q_i$ où $i_1, \ldots, i_{t(i)}$ appartiennent à $\{1, 2, \ldots, t\}$.

Alors $q_i \in \bar{K}[Y]$ et q_i n'a aucune racine dans $\bar{K}$.

Or q_i est unitaire, donc q_i est un produit de polynômes du second degré irréductibles de $\bar{K}[Y]$ (car $\bar{K}$ ordonné maximal).
Mais un tel polynôme du second degré irréductible sur un corps ordonné maximal est somme de deux carrés. Donc q_i est somme de carrés, donc q_i est positif dans $\bar{K}[Y]$ et q_i est positif dans $\overline{K(Y)}$.

. $b_{i_1}, \ldots, b_{i_{t(i)}}$ sont toutes les racines dans $\bar{L}$ de f_i^σ car sinon g^σ aurait plus de t racines dans $\bar{L}$.

Alors nous avons
$f_i^\sigma = (Y - b_{i_1})(Y - b_{i_2}) \ldots (Y - b_{i_{t(i)}})p_i$ où $p_i \in \bar{L}[Y]$ et p_i est unitaire et sans racine dans $\bar{L}$. On en déduit de même que p_i est somme de carrés dans $\bar{L}[Y]$.
Alors quel que soit l'élément $Z \in \bar{L}$, $p_i(Z)$ est une somme de carrés dans $\bar{L}$, donc $p_i(Z)$ est un élément positif de $\bar{L}$.

. Dans le corps $K(Y)$ on a $Y < r_1$, ou $Y > r_t$, ou il existe un indice j tel que $r_{j-1} < Y < r_j$ puisque Y est transcendant sur K alors que les r_j sont algébriques sur K.
D'après l'hypothèse L est dense dans $\bar{L}$, donc il existe $Z \in L$ tel que $b_{j-1} < Z < b_j$, ou $Z < b_1$, ou $Z > b_t$.

. Puisque q_i est positif dans $\overline{K(Y)}$ et que $p_i(Z)$ est positif dans $\bar{L}$ on en déduit d'après les expressions de f_i et de $f_i^{\sigma}(Z)$ que f_i et $f_i^{\sigma}(Z)$ ont le même signe puisque le nombre de racines de f_i supérieures à Y est le même que le nombre de racines de f_i^{σ} supérieures à Z.

. Soit $A \subset K[Y]$ l'anneau des polynômes à coefficients dans le domaine de σ. Alors $E \subset A$.
Soit $\zeta : A \to L$ l'application définie par $\zeta(h) = h^{\sigma}(Z)$. Alors ζ est une L-spécialisation de K(Y)/k qui préserve E d'après ce qui précède.
Donc K(Y)/k est L-artinienne.

Chapitre 5 REMARQUES SUR LE CAS PARTICULIER DES POLYNOMES

Soit K un corps commutatif ordonné tel que toute fonction rationnelle définie de $K(X_1, \ldots, X_n)$ se décompose en somme de carrés d'éléments de $K(X_1, \ldots, X_n)$, il est naturel de se demander si dans le cas d'un polynôme défini de $K[X_1, \ldots, X_n]$ on peut obtenir une décomposition en somme de carrés d'éléments de $K[X_1, \ldots, X_n]$.

En fait cette question admet une réponse positive dans le cas d'une seule variable et négative dans le cas de $\mathbb{R}[X_1, \ldots, X_n]$ où $n > 1$

1 CAS DE R[X] ET DE K[X]

L'étude à la fois qualitative et quantitative sera faite dans la partie II par une méthode très élémentaire qui permettra d'affirmer que dans le cas d'un corps R *ordonné maximal si* p *est un polynôme défini de* R[X] *alors* p *est somme de deux carrés dans* R[X].

Nous verrons également dans la partie II un résultat très général qui est : *soit* p *un élément de* K[X], K *étant un corps quelconque, si* p *est somme de* n *carrés dans* K(X), *alors* p *est somme de* n *carrés dans* K[X].

2 CAS DE $\mathbb{R}[X, Y]$

Nous avions vu qu'Hilbert a montré qu'il existait une forme définie de 3 variables et de degré 6 qui ne pouvait pas se décomposer en somme de carrés de formes mais qu'il n'en donnait pas d'exemple explicite. Nous allons ici donner un exemple de polynôme de deux variables à coefficients réels qui ne peut pas se décomposer en somme de carrés d'éléments de $\mathbb{R}[X, Y]$ quoique ce polynôme soit défini.

Nous utiliserons le théorème de Motzkin suivant [18] 1967 :

THEOREME :

Pour $n \geq 3$ *le polynôme de* n *variables à coefficients réels :*

$$h(t_1, t_2, \ldots, t_{n-1}, u) = (t_1^2 + \ldots + t_{n-1}^2 - n\,u^2)t_1^2 \ldots t_{n-1}^2 + u^{2n}$$

ne peut pas s'écrire sous forme d'une somme de carrés de polynômes de $\mathbb{R}[t_1, t_2, \ldots, t_{n-1}, u]$

Démonstration

Supposons que $h = \sum_k s_k^2$ dans $\mathbb{R}[t_1, \ldots, t_{n-1}, u]$ Soit 2 d le degré du terme de plus haut degré apparaissant dans cette somme.

En écrivant seulement les termes de degré 2 d nous aurons

$h^* = \sum_k s_k^{*2}$ où s_k^* est la partie de s_k de degré d.

Si $2d \neq 2n$ comme h est un polynôme homogène de d° 2 n, on a $\sum_k s_k^{*2} = 0$; cette égalité dans le corps ordonnable $\mathbb{R}(t_1, \ldots, t_{n-1}, u)$ entraîne $s_k^* = 0$ quel que soit k.

Donc $2d = 2n$ et comme h est homogène de degré 2 n : $h = h^*$

Nous avons alors la relation $h = \sum_k s_k^{*2}$ où s_k^* est un polynôme homogène de degré n.

Posons $s_k^* = C_{0k} + C_{1k}\,u + \ldots + C_{nk}\,u^n$ où les C_{ik} sont des polynômes de n-1 variables et de degré n - i.

Nous obtenons alors en explicitant l'égalité précédente la relation $(\ast)$

$$(t_1^2 + \ldots + t_{n-1}^2 - n\,u^2)t_1^2 \ldots t_{n-1}^2 + u^{2n} = \sum_k (C_{0k} + C_{1k}\,u + \ldots + C_{nk}\,u^n)^2$$

. Identifions les termes constants en u dans (✱)

$$\sum_k C_{0k}^2 = (t_1^2 + \ldots + t_{n-1}^2) t_1^2 t_2^2 \ldots t_{n-1}^2$$

Dans C_{0k} considérons le terme constant en t_1, soit a_{1k}, d'après l'identité ci-dessus, nous avons l'égalité $\sum_k a_{1k}^2 = 0$ qui a lieu dans le corps ordonnable $R(t_2, \ldots, t_{n-1})$ et qui entraîne donc $a_{1k} = 0$ pour tout k.

On en déduit immédiatement que t_1 divise C_{0k}.

On fait le même raisonnement pour $t_2, \ldots, t_{n-1}$ toutes ces variables jouant un rôle identique.

On obtient donc que t_1 puis t_2 puis $\ldots t_{n-1}$ divisent C_{0k} donc que $t_1 \times \ldots \times t_{n-1}$ divise C_{0k}.

. Considérons alors les termes de degré 2 en u dans (✱)

Nous avons :

$$\sum_k (2\, C_{0k}\, C_{2k} + C_{1k}^2) = - n\, t_1^2 \ldots t_{n-1}^2$$

D'où $$\sum_k C_{1k}^2 = - n\, t_1^2 \ldots t_{n-1}^2 - 2 \sum_k C_{0k}\, C_{2k}$$

C_{0k} étant divisible par $t_1 \ldots t_{n-1}$ $\sum_k C_{0k}\, C_{2k}$ l'est aussi. Donc $\sum_k C_{1k}^2$ est divisible par $t_1 \times \ldots \times t_{n-1}$ par un raisonnement analogue à celui fait ci-dessus, soit b_{1K} le terme constant en t_1 de C_{1K}. Nous avons donc $\sum_k b_{1k}^2 = 0$ et donc $b_{1k} = 0$ ce qui signifie que C_{1k} est divisible par t_1. Le même raisonnement pour t_2 puis $t_3, \ldots, t_{n-1}$ nous permet d'obtenir que finalement $t_1\, t_2 \ldots t_{n-1}$ divise C_{1K}.

. Considérons enfin les termes de degré 4 en u.

Nous aurons

$$\sum_k (2\, C_{0k}\, C_{4k} + 2\, C_{1k}\, C_{3k} + C_{2k}^2) = 0 \qquad \underline{\text{pour } n > 2}$$

Donc $$\sum_k C_{2k}^2 = - 2 \sum_k C_{0k}\, C_{4k} - 2 \sum_k C_{1k}\, C_{3k}$$

C_{0k} et C_{1k} étant divisibles par $t_1 \ldots t_{n-1}$, ΣC_{2k}^2 est aussi divisible par $t_1 \ldots t_{n-1}$ et par le même raisonnement déjà fait deux fois on en déduit que C_{2k} est divisible par $t_1 \times \ldots \times t_{n-1}$.

Mais C_{2k} est un polynôme de degré n-2 il ne peut donc pas être divisible par $t_1 \ldots t_{n-1}$ qui est un polynôme de degré n-1 d'où l'absurdité de l'hypothèse $h = \Sigma s_k^2$ dès que $n > 2$.

Appliquons maintenant le théorème au contre exemple cherché.

Si n = 3, nous avons donc d'après le théorème
$(t_1^2 + t_2^2 - 3u^2)t_1^2 t_2^2 + u^6$ qui n'est pas somme de carrés de polynômes en t_1, t_2, u.

Considérons le polynôme de deux variables obtenu en divisant le polynôme précédant par u^6 et en posant $x = \frac{t_1}{u}$ $y = \frac{t_2}{u}$

$f(x, y) = (x^2 + y^2 - 3)x^2 y^2 + 1$

Il ne peut être d'après le théorème précédent somme de carrés dans $\mathbb{R}[x, y]$

Montrons que f(x, y) est défini sur $\mathbb{R}$.

. Si $x^2 + y^2 \geqslant 3$, on a bien sûr $f(x, y) > 1$

. Si $x^2 + y^2 < 3$ soit $2k = x^2 + y^2$ $k \geqslant 0$

Pour k donné le maximum de $x^2 y^2$ est atteint pour $x^2 = y^2 = k$

Vérifions qu'alors
$(3 - x^2 - y^2)x^2 y^2 \leqslant 1$

Nous avons $0 \leqslant (3 - x^2 - y^2)x^2y^2 \leqslant (3 - 2k)k^2$, il suffit donc de montrer que
$(3 - 2k)k^2 \leqslant 1$
d'où $3k^2 - 2k^3 - 1 \leqslant 0$
et $-(k - 1)^2 (2k + 1) \leqslant 0$ ce qui est toujours vérifié quel que soit k réel positif ou nul.

Nous pouvons énoncer :

Théorème 5

$f(x, y) = (x^2 + y^2 - 3)x^2 y^2 + 1$ ***est défini sur*** $\mathbb{R}$ *et donc* $f(x, y)$ *est une somme de carrés dans* $\mathbb{R}(x, y)$ *mais* $f(x, y)$ *bien que polynôme n'est pas une somme de carrés dans* $\mathbb{R}[x, y]$

3 GENERALISATIONS

a) *Soit alors* K *un corps,* $\mathbb{Q} \subset K \subset \mathbb{R}$, *tel que le 17è problème de Hilbert admette une solution sur ce corps* K. *Puisque* $f(x, y) = (x^2 + y^2 - 3)x^2y^2 + 1$ *est défini sur* $\mathbb{R}$, $f(x, y)$ *l'est sur* K(K *n'a qu'un ordre, celui induit par* $\mathbb{R}$). *Donc* $f(x, y)$ *est une somme de carrés dans* $K(x, y)$ *mais* $f(x, y)$ *n'est pas une somme de carrés dans* $K[x, y]$ *car si* $f(x, y)$ *est une somme de carrés dans* $K[x, y]$ *ceci entraîne que* $f(x, y)$ *est une somme de carrés dans* $\mathbb{R}[x, y]$ *ce qui est faux comme nous venons de le voir.*

b) Enfin, considérons $p = f(x, y) + t^2 + u^2 + \ldots + w^2$ où $f(x, y) = (x^2 + y^2 - 3)x^2 y^2 + 1$ alors le polynôme $p \in K[x, y, t, \ldots, w]$ est défini sur K et est donc somme d'éléments de $K(x, y, t, \ldots, w)$ mais p n'est pas somme déléments de $K[x, y, t, .., w]$. En effet, il suffit de donner aux variables t, u, .. w la valeur nulle et on trouverait si p était somme de carrés dans $K[x, y, t, \ldots, w]$ que $f(x, y)$ est somme de carrés dans $K[x, y]$ ce qui est faux comme on vient de le voir.

4 Décomposition explicite en somme de carrés dans $\mathbb{R}(X, Y)$ du polynôme $f(X, Y) = 1 + X^2(X^2 - 3)Y^2 + X^2 Y^4$.

Multiplions $f(X, Y)$ par la somme de deux carrés $1 + X^2$:

$$\begin{aligned} f(X, Y)(1 + X^2) &= 1 - 3 X^2 Y^2 + X^2 Y^4 + X^2 + X^6 Y^2 - 2 X^4 Y^2 + X^4 Y^4 \\ &= (1 - X^2 Y^2)^2 + X^2[1 - Y^2 - 2 X^2 Y^2 + Y^4 + X^4 Y^2] \\ &= (1 - X^2 Y^2)^2 + X^2[1 - Y^2]^2 + X^2 Y^2[1 - 2 X^2 + X^4] \end{aligned}$$

Donc

$$f(X, Y) = \frac{(1 - X^2 Y^2)^2 + X^2[1 - Y^2]^2 + X^2 Y^2[1 - X^2]^2}{1 + X^2}$$

et en multipliant numérateur et dénominateur par $(1 + X^2)$

$$f(X, Y) = \frac{((1 - X^2 Y^2)^2 + (X[1 - Y^2])^2)(1 + X^2) + X^2Y^2[1 - X^2]^2 (1 + X^2)}{(1 + X^2)^2}$$

Nous savons que dans un anneau commutatif :
$(a^2 + b^2)(c^2 + d^2) = (ac - bd)^2 + (ad + bc)^2$

Nous en déduisons donc :

$$f(X, Y) = \frac{[(1-X^2Y^2)1-X^2(1-Y^2)]^2+[X(1-X^2Y^2)+X(1-Y^2)]^2+X^2Y^2(1-X^2)^2(1+X^2)}{(1 + X^2)^2}$$

$$f(X, Y) = \frac{(1-X^2)^2+X^2(2-X^2Y^2-Y^2)^2+X^2Y^2(1 - X^2)^2+X^4Y^2(1-X^2)^2}{(1 + X^2)^2}$$

et donc f(X, Y) est décomposée en somme de quatre carrés dans R(X, Y) :

$$f(X, Y) = \left(\frac{1-X^2}{1+X^2}\right)^2+\left(\frac{X(2-Y^2(1+X^2))}{1+X^2}\right)^2 + \left(\frac{X Y(1-X^2)}{1+X^2}\right)^2 + \left(\frac{X^2Y(1-X^2)}{1+X^2}\right)^2$$

5 REMARQUE SUR UN TRAVAIL DE R.M. ROBINSON 1969 [19]

A partir du travail de Hilbert qui ne donnait pas de contre exemple explicite Raphael M. Robinson a montré que les polynômes
$P(x, y) = x^2(x^2 - 1)^2 + y^2(y^2 - 1)^2 - (x^2 - 1)(y^2 - 1)(x^2 + y^2 - 1)$

et

$q(x, y, z) = x^2(x - 1)^2 + y^2(y - 1)^2 + z^2(z - 1)^2 + 2 xyz(x + y + z - 2)$

sont définis positifs mais ne peuvent pas s'exprimer comme somme de carrés de polynômes à coefficients réels.

Mais de plus, appliquant le théorème d'Artin, Robinson a ensuite cherché une représentation de ces polynômes en somme de carrés de fractions rationnelles à coefficients dans $\mathbb{Q}$ et a trouvé de telles représentations. Malheureusement, je n'ai connaissance que du résumé de l'article, adressé à "The American Mathematical Society" en 1969, sur lequel R.M. Robinson m'a indiqué qu'il ne savait quand cela allait paraître.

En conclusion à cette partie nous pouvons énoncer le résultat suivant :

Soit k *un corps ordonné n'ayant qu'un seul ordre, dense dans sa clôture réelle* $\bar{k}$ *(par exemple* $\mathbb{Q} \subset k \subset \mathbb{R}$, *ou* k *ordonné maximal).*

Soit $f \in k(X_1, \ldots, X_n)$ *telle que pour tout* $(x_1, \ldots, x_n) \in k^n$ $f(x_1, \ldots, x_n) \geqslant 0$.

Alors il existe des éléments $h_i \in k(X_1, \ldots, X_n)$ *tels que* $f = \sum_{i=1}^{r} h_i^2$.

De plus si $n = 1$, *et que* $f \in k[X_1]$ *alors on peut affirmer que* $f = \sum_{i=1}^{r} h_i^2$ *avec* $h_i \in k[X_1]$

Mais si $n > 1$, *et* $f \in k[X_1, \ldots, X_n]$
On ne peut pas affirmer que $f = \sum_{i=1}^{r} h_i^2$ *avec* $h_i \in k[X_1, \ldots, X_n]$

On peut se demander s'il existe des corps ordonnés n'ayant qu'un seul ordre, non denses dans leur clôture réelle et sur lequel le 17è problème de Hilbert admettrait une réponse positive.
Cette question n'est pas résolue.

2EME PARTIE

RESOLUTIONS QUANTITATIVES

Soit K un corps tel que toute fonction définie de $K(X_1, \ldots, X_n)$ se décompose en somme de carrés dans $K(X_1, \ldots, X_n)$.

Dans cette partie nous allons donc chercher à connaître le nombre de carrés nécessaires et suffisants pour pouvoir décomposer en somme de carrés toute fraction rationnelle définie à coefficients dans le corps donné K, le nombre de variables des fractions rationnelles étant fixé.

Au chapitre 1, nous donnerons un résultat de Cassels sur le cas particulier des polynômes d'une seule variable.

Au chapitre 2, nous trouverons une minoration du nombre de carrés nécessaires pour décomposer en somme de carrés dans $K(X_1, \ldots, X_n)$ toute fonction définie de $K(X_1, \ldots, X_n)$.

Au chapitre 3, nous donnerons une majoration du nombre de carrés nécessaires à la décomposition de toute fonction définie de $K(X_1, \ldots, X_n)$ dans le cas où K est un corps ordonné maximal.

Au chapitre 4, nous obtiendrons par une méthode élémentaire une résolution complète dans le cas d'un corps $K(X)$ où K est ordonné maximal.

Au chapitre 5, nous étudierons un résultat récent qui permet d'obtenir une résolution complète du cas $\mathbb{R}(X, Y)$.

Au chapitre 6, nous citerons quelques résultats quantitatifs sur le 17è problème de Hilbert posé sur d'autres corps que $\mathbb{R}$ ou des corps ordonnés maximaux.

CHAPITRE 1 ETUDE D'UN RESULTAT DE CASSELS

Nous allons étudier ici un résultat obtenu par M. Cassels dans [21]

Théorème 6

Si un polynôme de $K[X]$ *(*K *étant un corps quelconque) est somme de* n *carrés dans* $K(X)$, *alors il est somme de* n *carrés dans* $K[X]$

Ce résultat n'est nouveau que par le fait que n carrés suffisent. Artin dans [6] avait déjà démontré l'aspect qualitatif de ce résultat en s'inspirant d'une démonstration de Landau dans [20] sur le cas de $\mathbb{Q}[X]$.

Démonstration

On traite d'abord trois cas particuliers simples qui seront écartés par la suite.

1) Si n = 1 C'est que $f \in K[X]$ s'écrit $f = (\frac{g}{h})^2$ g et $h \in K[X]$

En utilisant la décomposition en facteurs irréductibles :

$$f = \prod_{i=1}^{r} p_i^{e_i} \qquad p_i \text{ irréductible}, \quad e_i \in \mathbb{N}$$

$$\frac{g}{h} = \prod_{i=1}^{r} p_i^{\ell_i} \qquad p_i \text{ irréductible}, \quad \ell_i \in \mathbb{Z}$$

De l'identité on déduit $e_i = 2\,\ell_i$, ce qui entraîne donc $\ell_i \in \mathbb{N}$ et donc $\frac{g}{h}$ est un polynôme.

2) Si K est de caractéristique 2 et n > 1

$$f = \sum_{i=1}^{n} (\frac{g_i}{h_i})^2 = (\sum_{i=1}^{n} \frac{g_i}{h_i})^2 = (\frac{g}{h})^2 \text{ avec } g \text{ et } h \in K[X]$$

f est donc alors somme de 1 carré dans K(X) et l'on est ramené au cas 1).

3) K est de caractéristique différente de 2, n > 1 et - 1 est somme de n-1 carrés dans K.

Nous avons donc dans K :

$$-1 = a_2^2 + a_3^2 + \ldots + a_n^2$$

On peut alors écrire f sous la forme :

$$f = \left(\frac{f+1}{2}\right)^2 - \left(\frac{f-1}{2}\right)^2$$

$$= \left(\frac{f+1}{2}\right)^2 + \left(\frac{f-1}{2}\right)^2 \left[\sum_{j=2}^{n} a_j^2\right]$$

$$= \left(\frac{f+1}{2}\right)^2 + \sum_{j=2}^{n} \left(a_j \frac{f-1}{2}\right)^2$$ et donc f est une somme de n carrés

de polynômes de $K[X]$.

Ecartons maintenant ces trois cas particuliers et étudions donc le cas de :

K de caractéristique différente de 2, n > 1, et - 1 non somme de n-1 carrés dans K.

Par hypothèse $f = \sum_{i=1}^{n} \left(\frac{g_i}{h_i}\right)^2$

Nous avons donc aussi en réduisant au même dénominateur :

$$f z^2 = \sum_{i=1}^{n} y_i^2 \qquad z, y_i \in K[X]$$

Il faut alors montrer que l'on peut trouver un n+1-uple de $K[X]$ (y_i, z) tel que $f z^2 = \sum_{i=1}^{n} y_i^2$ avec $z \in K$ et $z \neq 0$

Considérons alors l'équation :

$$f Z^2 = \sum_{i=1}^{n} Y_i^2 \quad (*)$$

La démonstration se fait par l'absurde en supposant que (y_i, z) est une solution de l'équation telle que $z \notin K$ et z de degré le plus petit possible

dans l'ensemble des solution de (*) qui n'est pas vide puisque l hypothèse sur f donne une solution, et en montrant qu'on peut alors trouver une solution (y'_i, z') avec $d^o z' < d^o z$ et $z' \neq 0$, d'où la contradiction.

Cette solution (y'_i, z') va être obtenue comme le second point d'intersection avec $f Z^2 = \sum_{i=1}^{n} Y_i^2$ de la droite joignant (y_i, z) a un point $(\lambda_i, 1)$ dans l'espace projectif à n dimensions sur K(X).

Définissons ce point $(\lambda_i, 1)$:

Soient λ_i les éléments de $K[X]$ définis de façon unique par division euclidienne dans $K[X]$:

$y_i = z \lambda_i + r_i$ avec $r_i = 0$ ou $d^o r_i < d^o z$

Remarquons que si $\forall i$, $r_i = 0$ alors $y_i = z \lambda_i$ et de la relation $f z^2 = \sum_{i=1}^{n} y_i^2$ on déduit $f z^2 = \sum_{i=1}^{n} \lambda_i^2 z^2$ et donc puisque $z \neq 0$

$f = \Sigma \lambda_i^2$ ce qui donne une solution $(\lambda_i, 1)$ ce qui est impossible d'après l'hypothèse. Donc $f \neq \Sigma \lambda_i^2$ et les r_i sont non tous nuls.

Posons alors :

$$y'_i = y_i \{\sum_u \lambda_u^2 - f\} - 2 \lambda_i \{\sum_u \lambda_u y_u - f z\}$$

$$z' = z \{\sum_u \lambda_u^2 - f\} - 2 \{\sum_u \lambda_u y_u - f z\}$$

y'_i et z' appartiennent évidemmment à $K[X]$.

Montrons que (y'_i, z') est solution de l'équation (*) en montrant que c'est le second point d'intersection de la droite joignant (y_i, z) à $(\lambda_i, 1)$ avec l'hypersurface d'équation (*) dans l'espace projectif à n dimensions sur le corps K(X).

Soit donc Γ : $f Z^2 = \sum_{i=1}^{n} Y_i^2$

et D : $\begin{cases} Y_i = \lambda y_i + t \lambda_i \\ Z = \lambda z + t.1 \end{cases}$

Les intersections de Γ et de D sont données par les valeurs de t qui satisfont :

$$f (\lambda z + t)^2 = \sum_{i=1}^{n} (\lambda y_i + t \lambda_i)^2$$

en tenant compte du fait que $f z^2 = \sum_{i=1}^{n} y_i^2$ ceci devient :

$$f(2 \lambda z t + t^2) = \sum_{i=1}^{n} (2 \lambda t \lambda_i y_i + t^2 \lambda_i^2)$$

Simplifions par t ($t \neq 0$ car sinon on obtient le point que l'on connaît (y_i, z)).

$$2 \lambda(f z - \sum_{i=1}^{n} y_i \lambda_i) + t(f - \sum_{i=1}^{n} \lambda_i^2) = 0$$

On peut choisir $\lambda = \Sigma \lambda_i^2 - f$ qui est non nul d'après ce qui a été vu.

En simplifiant alors par $\Sigma \lambda_i^2 - f$ on obtient $t = 2(f z - \sum_{i=1}^{n} y_i \lambda_i)$

en reportant cette valeur de t et la valeur de λ choisie dans les coordonnées d'un point de D on obtient bien :

$$\begin{cases} y'_i = (\sum_{u=1}^{n} \lambda_u^2 - f) y_i + 2(f z - \sum_{u=1}^{n} \lambda_u y_u) \lambda_i \\ z' = (\sum_{u=1}^{n} \lambda_u^2 - f) z + 2(f z - \sum_{u=1}^{n} \lambda_u y_u) \end{cases}$$

vérifions maintenant que cette solution (y'_i, z') est bien telle que le degré de z' est inférieur au degré de z et que $z' \neq 0$. Nous avions défini les λ_i par la division euclidienne $y_i = \lambda_i z + r_i$.

Comme $z \neq 0$ on peut écrire ceci

$\frac{y_i}{z} = \lambda_i + \frac{r_i}{z}$. Soit $\Lambda_i = \lambda_i - \frac{y_i}{z} = \frac{-r_i}{z}$ comme le degré de r_i est plus petit que le degré de z, le degré de Λ_i est négatif. (**)

Calculons z' en fonction de Λ_i : $\lambda_i = \Lambda_i + \frac{y_i}{z}$ portons ceci dans l'expression de z'

$$z' = z\left(\sum_{u=1}^{n} (\Lambda_u + \frac{y_u}{z})^2 - f\right) + 2\left(fz - \sum_{u=1}^{n} y_u(\Lambda_u + \frac{y_u}{z})\right)$$

$$z' = \frac{1}{z} \sum_{u=1}^{n} (\Lambda_u^2 z^2 + 2\Lambda_u y_u z + y_u^2) - fz + 2fz - 2\sum_{u=1}^{n} (\Lambda_u y_u + \frac{y_u^2}{z})$$

$$z' = z \sum_{u=1}^{n} \Lambda_u^2 + 2\sum_{u=1}^{n} \Lambda_u y_u + \frac{1}{z}\sum_{u=1}^{n} y_u^2 + fz - 2\sum_{u=1}^{n} \Lambda_u y_u - \frac{2}{z}\sum y_u^2$$

$$z' = z \sum_{u=1}^{n} \Lambda_u^2 - \frac{f z^2}{z} + f z = z \sum_{u=1}^{n} \Lambda_u^2$$

en tenant compte de $\sum_{u=1}^{n} y_u^2 = f z^2$.

Donc puisque $z' = z \sum_{u=1}^{n} \Lambda_u^2$ et que le degré de $\Lambda_u \neq 0$, ceci signifie que le degré de z' est bien strictement inférieur au degré de z.

Vérifions enfin que $z' \neq 0$. Il suffit pour cela de vérifier que $\sum_{i=1}^{n} \Lambda_u^2$ est non nul.

(**) On définit évidemment le degré d'un élément de K(X) de la manière suivante :

Soit $f \in K(X)$ $\quad f = \frac{p}{q}$ p et q appartenant à $K[X]$, par définition

$d^o f = (d^o p) - (d^o q.)$

En effet, $\sum_{u=1}^{n} \Lambda_u^2 = \sum_{u=1}^{n} \left(\frac{r_u}{z}\right)^2 = \frac{1}{z^2} \sum_{u=1}^{n} r_u^2$

Donc z' est nul si et seulement si $\sum_{u=1}^{n} r_u^2 = 0$

On sait que les r_u sont non tous nuls.

Soit d le maximum des degrés des polynômes r_u non nuls. Alors

$$r_u = b_{uo} X^d + b_{u1} X^{d-1} + \ldots + b_{ud} \quad \text{les} \quad b_{uj} \in K$$

et les b_{uo} non tous nuls.

On déduit alors de $\sum_{u=1}^{n} r_u^2 = 0$ $\qquad \sum_{u=1}^{n} b_{uo}^2 = 0$

L'existence d'un $b_{uo} \neq 0$ entraîne alors que - 1 est somme de n - 1 carrés ce qui est contraire à l'hypothèse faite sur le corps K.

On a donc bien trouvé une solution de $f Z^2 = \sum_{i=1}^{n} Y_i^2$, (y'_i, z') qui est telle que deg z' < deg z et $z' \neq 0$, ce qui est contraire à l'hypothèse sur z.

Donc il existe une solution de degré le plus petit possible (y_i, z) qui est donc telle que $z \in K$ et $z \neq 0$. Cette solution permet de décomposer f en somme de n carrés de polynômes de $K[X]$.

CHAPITRE 2

MINORATION DU NOMBRE DE CARRES NECESSAIRES POUR POUVOIR DECOMPOSER TOUTE FONCTION DEFINIE DE $K(X_1, \ldots, X_n)$

Dans [22] Davenport avait démontré en 1963 que $X_1^2 + X_2^2 + X_3^2 + X_4^2$ ne pouvait pas être une somme de 3 carrés dans $\mathbb{R}(X_1, X_2, X_3, X_4)$. Nous obtiendrons ici un résultat plus général comme conséquence d'un résultat dû à Cassels et démontré dans [21]

Lemme (Cassels) 1964

Soit K un corps de caractéristique différente de 2, d un élément de K. $X^2 + d$ est une somme de n carrés dans K(X) si et seulement si -1 est une somme de $n-1$ carrés dans K où d est une somme de $n-1$ carrés dans K.

Démonstration :

1°) Si d est une somme de $n-1$ carrés dans K, il est évident que $X^2 + d$ est une somme de n carrés dans K(X)

2°) Si -1 est une somme de $n-1$ carrés dans K, nous pouvons écrire :

$$-1 = \sum_{j=2}^{n} a_j^2 \quad \text{et donc :}$$

$$X^2 + d = \left(\frac{(X^2+d) + 1}{2}\right)^2 - \left(\frac{(X^2+d) - 1}{2}\right)^2$$

$$= \left(\frac{(X^2+d) + 1}{2}\right)^2 + \sum_{j=2}^{n} a_j^2 \left(\frac{(X^2+d) - 1}{2}\right)^2$$

qui est bien une somme de n carrés de K(X).

Réciproquement :

Soit $X^2 + d$ une somme de n carrés dans K(X) et -1 non somme de $n-1$ carrés dans K. Montrons qu'alors nécessairement d est somme de n-1 carrés dans K.

D'après le résultat du chapitre 1 de cette partie si $X^2 + d$ est somme de n carrés dans K(X) c'est aussi une somme de n carrés dans $K[X]$. Donc

$$X^2 + d = \sum_{i=1}^{n} f_i(X)^2 \qquad f_i \in K[X]$$

Chaque polynôme f_i est de degré au plus égal à 1. En effet, sinon en considérant seulement les termes de plus haut degré dans l'égalité nous aurions par identification de leurs coefficients $0 = \sum_{i=1}^{n} \lambda_i^2$ avec les λ_i non tous nuls.

De ceci on déduit donc -1 est somme de $n-1$ carrés dans K, ce qui est impossible d'après l'hypothèse.

Soit donc

$$f_i = a_i X + b_i \qquad a_i \text{ et } b_i \in K$$

Considérons $f_n = a_n X + b_n$.

. Si $a_n \neq 1$ soit l'équation

$a_n x + b_n = x \qquad (a_n - 1)x + b_n = 0$

On en déduit la solution dans K

$$x = \frac{-b_n}{a_n - 1} \quad (\text{Car } a_n \neq 1)$$

. Si $a_n = 1$, considérons l'équation $a_n x + b_n = -x$, on en déduit

$$x = \frac{-b_n}{a_n + 1} \text{ solution dans K}$$

Soit α la valeur x trouvée selon celle de a_n en portant celle-ci dans la relation

$$X^2 + d = \sum_{i=1}^{n} f_i^2 \text{ on obtient}$$

$$\alpha^2 + d = \sum_{i=1}^{n} (a_i \alpha + b_i)^2$$

$$= \sum_{i=1}^{n-1} (a_i \alpha + b_i)^2 + (a_n \alpha + b_n)^2$$

$$= \sum_{i=1}^{n-1} (a_i \alpha + b_i)^2 + (\pm \alpha)^2$$

D'où $d = \sum_{i=1}^{n-1} (a_i \alpha + b_i)^2$ et d est donc somme de n-1 carrés dans K.

Théorème 7

Soit K *un corps ordonnable, alors le polynôme* $X_1^2 + X_2^2 + \ldots + X_n^2 + 1$ *n'est pas somme de* n *carrés dans* $K(X_1, \ldots, X_n)$

Démonstration

Elle se fait par récurrence :
pour n = 1, $X_1^2 + 1$ n'est pas somme d'un seul carré dans $K(X_1)$ d'après le lemme. En effet, K étant ordonnable est bien de caractéristique différente de 2, et 1 ou - 1 ne sont pas sommes de 0 carrés dans K.

Supposons le résultat acquis pour n - 1. Donc que $X_1^2 + \ldots + X_{n-1}^2 + 1$ n'est pas une somme de n - 1 carrés dans $K(X_1, \ldots, X_{n-1})$

Considérons
$X_n^2 + (X_1^2 + \ldots + X_{n-1}^2 + 1) = X_n^2 + d$

D'après l'hypothèse de récurrence d n'est pas somme de n - 1 carrés dans $K(X_1, \ldots, X_{n-1})$; d'autre part $K(X_1, \ldots, X_{n-1})$ étant ordonnable puisque K l'est, - 1 ne peut pas être somme de n - 1 carrés dans $K(X_1, \ldots, X_{n-1})$

En appliquant le lemme à $X_n^2 + d$ dans le corps $K(X_1, \ldots, X_{n-1})(X_n)$ nous en déduisons que $X_n^2 + d$ n'est pas somme de n carrés dans $K(X_1, \ldots, X_{n-1})(X_n) = K(X_1, \ldots, X_n)$. D'où le théorème.

En conclusion si K est un corps tel que le 17è problème de Hilbert admette sur $K(X_1, \ldots, X_n)$ une réponse positive, alors pour pouvoir décomposer toute fonction définie de $K(X_1, \ldots, X_n)$ en somme de p(n) carrés dans $K(X_1, \ldots, X_n)$ il faudra nécessairement que $p(n) \geq n+1$ car le polynôme $X_1^2 + \ldots + X_n^2 + 1$ qui est évidemment défini n'est pas somme de n carrés dans $K(X_1, \ldots, X_n)$.

CHAPITRE 3

MAJORATION DU NOMBRE DE CARRES NECESSAIRES A LA DECOMPOSITION DE TOUTE FONCTION DEFINIE DE $K(X_1, \ldots, X_n)$ K ETANT UN CORPS ORDONNE MAXIMAL.

Dans ce chapitre nous étudierons le résultat suivant énoncé par A. Pfister dans [23].

Théorème 8

Soit R *un corps ordonné maximal, et* $f \in R(X_1, \ldots, X_n)$ *telle que* $f(a_1, \ldots, a_n) \geqslant 0$ *pour tout* $(a_1, \ldots, a_n) \in R^n$ *où* f *a un sens ; alors* f *est somme de* 2^n *carrés dans* $R(X_1, \ldots, X_n)$

Ce résultat nous permet donc d'affirmer que si R est un corps ordonné maximal, et p(n) le nombre de carrés nécessaires à la décomposition en somme de carrés dans $R(X_1, \ldots, X_n)$ de toute fonction définie de $R(X_1, \ldots, X_n)$ alors nous avons

$$p(n) \leqslant 2^n.$$

En rapprochant ce résultat de celui obtenu au chapitre précédent, nous obtenons donc

$$\boxed{n + 1 \leqslant p(n) \leqslant 2^n}$$

1 . Démonstration du théorème 8

Nous utiliserons trois lemmes.

Lemme 1. Ce lemme est en fait le théorème 2' énoncé au chapitre 2 de la première partie.
Si R est un corps ordonné maximal, toute fonction définie de $R(X_1, \ldots, X_n)$ est une somme de carrés dans $R(X_1, \ldots, X_n)$

Lemme 2. (Pfister [23]).

Si R est un corps ordonné maximal et L une extension algébrique non ordonnable du corps $R(X_1, \ldots, X_n)$ alors - 1 est une somme de 2^n carrés dans L.

Lemme 3. (J. Ax [26])

Soit K un corps ordonnable ; soit $f \in K(X)$ telle que f soit une somme de carrés dans K(X).
Si - 1 est somme de 2^{n-1} carrés dans tout corps L extension algébrique non ordonnable de K, alors f est somme de 2^n carrés dans K(X).

Démonstration du théorème :

. Pour n = 0 $f \in R$, et f est positif donc f est un carré dans R et on a bien $1 = 2^0$.

. Pour $n \geqslant 1$: Soit $K = R(X_1, \ldots, X_{n-1})$; K est un corps ordonnable. Soit L une extension algébrique non ordonnable de $R(X_1, \ldots, X_{n-1})$ on en déduit d'après le lemme 2 que - 1 est somme de 2^{n-1} carrés dans L. D'après le lemme 1, f est somme de carrés dans $R(X_1, \ldots, X_n) = K(X_n)$. En appliquant alors le lemme 3 on obtient que f est somme de 2^n carrés dans $K(X_n)$ et donc dans $R(X_1, \ldots, X_n)$.

2 . Résultats utilisés pour la démonstration des lemmes 2 et 3.

Proposition 1

Soit K un corps quelconque, k un entier alors le produit de deux sommes de 2^k carrés d'éléments de K est une somme de 2^k carrés d'éléments de k.

En utilisant la notation de Pfister, ceci s'écrit :

$$\boxed{2^k}_K \cdot \boxed{2^k}_K = \boxed{2^k}_K$$

Ce résultat est un corollaire d'un théorème de Pfister paru dans [24].

Avant d'énoncer ce théorème, précisons quelques notations. Soit K un corps quelconque. Considérons les polynômes $\varphi = \sum_{i=1}^{n} a_i X_i^2$; $\varphi \in K[X_1, \ldots, X_n]$ et on suppose que $\forall\, i \quad a_i \neq 0$.

On notera ce polynôme $\varphi = \langle a_1, \ldots, a_n \rangle$

On notera G_φ l'ensemble suivant :

$$G_\varphi = \{\varphi(x) \neq 0 \mid x \in K^n\}.$$

On définit aussi $\varphi \otimes \psi$ où $\varphi = \langle a_1, \ldots, a_m \rangle$

et $\psi = \langle b_1, \ldots, b_n \rangle$ par :

$$\varphi \otimes \psi = \langle a_1 b_1, \ldots, a_m b_1, a_1 b_2, \ldots, a_m b_2, \ldots, a_1 b_n, \ldots, a_m b_n \rangle$$

Le théorème de Pfister s'énonce alors sous la forme :

Soient $a_1, \ldots, a_k \in K^*$; soit $\varphi = \langle 1, a_1 \rangle \otimes \ldots \otimes \langle 1, a_k \rangle$

Alors

$G_\varphi = \{ \varphi(x) \neq 0 \mid x \in K^{2^k} \}$ est un sous-groupe de K^*

Pour démontrer la proposition 1, prenons

$\varphi = \langle 1, 1 \rangle \otimes \langle 1, 1 \rangle \otimes \ldots \otimes \langle 1, 1 \rangle$ (k fois)

alors $G_\varphi = \boxed{2^k}_K$; puisque d'après le théorème précédent G_φ est un groupe multiplicatif nous avons donc bien

$$\boxed{2^k}_K \cdot \boxed{2^k}_K = \boxed{2^k}_K$$

Proposition 2

Soit K un corps non ordonnable. On sait alors que - 1 est somme de carrés dans K. On appellera niveau du corps K le plus petit entier $\nu(K)$ tel que - 1 soit une somme de $\nu(K)$ carrés. Le niveau de K est alors toujours une puissance de 2.

Démonstration :

Soit k l'entier tel que $2^k \leqslant \nu(K) < 2^{k+1}$

Nous avons donc

$$-1 = x_1^2 + x_2^2 + \ldots + x^2_{\nu(K)} \quad \text{où } x_i \in K^*$$

Transformons la relation écrite ci-dessus en

$$0 = (1 + x_1^2 + \ldots + x^2_{2^k - 1}) + (x^2_{2^k} + \ldots + x^2_{\nu(k)})$$

$1 + x_1^2 + \ldots + x^2_{2^{k-1}}$ est non nul car sinon $\nu(K)$ vérifierait $\nu(K) \leqslant 2^{k-1}$ ce qui est contraire à l'hypothèse faite sur k.

Nous pouvons donc diviser par $1 + x_1^2 + \ldots + x^2_{2^{k-1}}$ et nous obtenons :

$$-1 = \frac{x^2_{2^k} + \ldots + x^2_{\nu(K)}}{1 + x_1^2 + \ldots + x^2_{2^{k-1}}}$$

D'après la proposition 1 on en déduit que - 1 est une somme de 2^k carrés dans K donc $\nu(K) = 2^k$.

Proposition 3 Pfister [23]

Soit L un corps non ordonnable tel que le corps L(i) satisfasse la propriété suivante :
Toute forme quadratique à coefficients dans L(i) de $\nu > 2^n$ variables a un zéro non trivial dans $(L(i))^\nu$ (propriété notée $C_n(2)$).
Alors - 1 est somme de 2^n carrés dans L.

Démonstration :

Si $i \in L$ ou si L est de caractéristique 2 on a évidemment - 1 est un carré dans L.

Si $i \notin L$ et si car $L \neq 2$ puisque L n'est pas ordonnable - 1 est une somme de 2^m carrés dans L pour un certain entier m (voir proposition 2)

. Si $m \leqslant n$ on a bien le résultat voulu.

. Si $m > n$. Soient $d_1, \ldots, d_{2^m}$ dans L tels que $d_1^2 + \ldots + d_{2^m}^2$ ne soit pas une somme de $2^m - 1$ carrés dans L.

Posons $b = d_1$

$c = d_2$

$a_1 = d_3^2 + d_4^2$

$a_2 = d_5^2 + d_6^2 + d_7^2 + d_8^2$

....

$a_{m-1} = d_{2^{m-1}+1}^2 + \ldots + d_{2^m}^2$

Aucun de ces éléments de L n'est nul car sinon on aurait $d_1^2 + \ldots + d_{2^m}^2$ qui serait une somme de $2^m - 1$ carrés, ce qui est contraire à l'hypothèse prise.

Posons alors $\varphi = \langle 1, a_1 \rangle \otimes \langle 1, a_2 \rangle \otimes \ldots \otimes \langle 1, a_{m-1} \rangle$ (notation définie page 67)

Il existe(*) $s \in L^{2^{m-1}}$ tel que $\varphi(s) = b^2 + c^2$. Donc on peut écrire en supposant $s = (x_1, \ldots, x_{2^{m-1}})$ l'égalité suivante :

(*) Voir [23] page 234 ou [8] propriété u : Soit L un corps tel que L(i) satisfasse la propriété $C_n(2)$. Soient $a_1, a_2, \ldots, a_n$, b et c dans L^* et soit $\varphi = \langle 1, a_1 \rangle \otimes \langle 1, a_2 \rangle \otimes \ldots \otimes \langle 1, a_n \rangle$, alors il existe $s \in L^{2^n}$ tel que $\varphi(s) = b^2 + c^2$

Comme ici $m-1 \geqslant n$ L a la propriété $C_{m-1}(2)$ d'où l'affirmation sur l'existence de s.

$$b^2 + c^2 = x_1{}^2 + a_1\, x_2{}^2 + a_2(x_3{}^2 + a_1\, x_4{}^2) + \ldots. +$$

$$+ a_{m-1}(x^2_{2^{m-2}} + 1 + \ldots + (a_1 \ldots a_{m-2})\, x^2_{2^{m-1}})$$

D'où puisque

$$d_1{}^2 + d_2{}^2 + \ldots + d^2_{2^m} = d_1{}^2 + d_2{}^2 + a_1 + \ldots + a_{m-1}$$

$$d_1{}^2 + \ldots + d^2_{2^m} = x_1{}^2 +$$

$$a_1(1 + x_2{}^2) + a_2(1 + x_3{}^2 + a_1\, x_4{}^2) + \ldots.$$

$$\ldots + a_{m-1}(1 + x^2_{2^{m-2}+1} + \ldots + (a_1 \ldots a_{m-2})x^2_{2^{m-1}})$$

Donc

$d_1{}^2 + \ldots + d^2_{2^m}$ est la somme des termes suivants :

$x_1{}^2$ qui est 1 carré dans L

$a_1(1 + x_2{}^2)$ qui est le produit de 2 sommes de deux carrés dans L

$a_2(1 + x_3{}^2 + a_1\, x^2{}_4)$ qui est le produit de 2 sommes de quatre carrés dans L puisque $a_1\, x_4{}^2$ est une somme de deux carrés.

. . . . . .

$a_{m-1}(1 + x^2_{2^{m-2}+1} + \ldots + (a_1 \ldots a_{m-2})\, x^2_{2^{m-1}})$ qui est le produit de

a_{m-1} qui est somme de 2^{m-1} carrés par une somme de 2^{m-1} carrés, en effet il suffit d'écrire les termes sous la forme suivante :

$$a_{m-1}\{\, 1 + x^2_{2^{m-2}+1} + a_1\, x^2_{2^{m-2}+2} + a_2(x^2_{2^{m-2}+3} + a_1\, x^2_{2^{m-2}+4})$$

$$+ a_3(x^2_{2^{m-2}+5} + a_1\, x^2_{2^{m-2}+6} + a_2(x^2_{2^{m-2}+7} + a_1\, x^2_{2^{m-2}+8}))+ \ldots\ldots. +$$

$$a_{m-2}\{\, x^2_k + a_1\, x^2_{k+1} + a_2(x^2_{k+2} + a_1\, x^2_{k+3}) + \ldots. + a_{m-3}[x_j{}^2 + a_1 x^2_{j+1} + a_2(\ldots) +$$

$$\ldots. + a_{m-4}(x^2 \ldots + a_1\, x^2 \ldots + a_2(\ldots) + \ldots + a_{m-5}(\ldots a_{m-6}(\ldots)))]\,\}$$

qui permet d'affirmer que la parenthèse { } est somme de
$1 + 1 + 2 + 4 + 8 + \ldots\ldots + 2^{m-2}$ carrés, donc de $1+\frac{2^{m-1}-1}{2-1} = 2^{m-1}$ carrés.

Par exemple, écrivons complètement le terme en a_4 :

$$a_4 \quad [1 + x_9^2 + a_1 x_{10}^2 + a_2 (x_{11}^2 + a_1 x_{12}^2) + a_3(x_{13}^2 + a_1 x_{14}^2 + a_2(x_{15}^2 + a_1 x_{16}^2))]$$

a_4 est somme de 16 carrés et le crochet somme des termes suivants :
1 ; x_9^2 ; $a_1 x_{10}^2$ qui est somme de 2 carrés comme a_1 ;
$a_2(x_{11}^2 + a_1 x_{12}^2)$ qui est le produit d'une somme de 4 carrés par une somme de 3 carrés donc est une somme de quatre carrés ;
$a_3(x_{13}^2 + a_1 x_{14}^2 + a_2(x_{15}^2 + a_1 x_{16}^2))$ où a_3 est somme de 8 carrés $a_1 x_{14}^2$ somme de 2 carrés, a_2 somme de 4 carrés et $x_{15}^2 + a_1 x_{16}^2$ somme de 3 carrés. La parenthèse est donc une somme de 7 carrés et l'ensemble est somme de 8 carrés ;
Le crochet coefficient de a_4 est donc finalement somme de
$(1 + 1 + 2 + 4 + 8) = 16$ carrés dans L.

On en déduit donc en utilisant le fait que le produit de deux sommes de 2^k carrés est une somme de 2^k carrés que

$d_1^2 + \ldots + d_{2^m}^2$ est somme de

$1 + 2 + 4 + \ldots + 2^{m-1} = 2^m - 1$ carrés dans L ce qui est contraire à l'hypothèse faite sur les d_i. Nous avons donc obtenu que pour $m > n$ les sommes de 2^m carrés sont aussi des sommes de 2^m-1 carrés.

Si -1 est somme de 2^m carrés dans L avec $m > n$ alors -1 est somme de $2^m - 1$ carrés dans L. Mais d'après la proposition 2 le niveau de L ne peut être qu'une puissance de 2 ; donc nous avons -1 qui est somme de 2^{m-1} carrés dans L. Si $m - 1 = n$ la démonstration est terminée,

sinon on en déduit que - 1 est somme de $2^{m-1} - 1$ carrés dans L donc de 2^{m-2} carrés dans L.
Si m - 2 = n c'est terminé sinon on recommence.
En itérant le procédé nous obtenons que - 1 est somme de 2^n carrés dans L.

Proposition 4

Soit K_o un corps algébriquement clos ; K une extension de K_o de degré de transcendance n. Alors si f est un polynôme homogène, de degré d à coefficients dans K, de $\nu > d^n$ variables, alors f a un zéro non trivial dans K^ν.

Cette proposition peut être considérée comme un corollaire d'un théorème que l'on peut trouver dans [8].

L'origine de ce résultat se trouve en fait dans un article de Lang [25]

Donnons quelques définitions préalables :

Soit K un corps, $i \geqslant 0$ un nombre réel et $d \geqslant 1$ un entier.
On dit que K a la propriété $C_i(d)$ (noté $K \in C_i(d)$) si tout polynôme f, homogène, à coefficients dans K, de degré d et de $n > d^i$ variables a un zéro non trivial dans K : $(x_1, \ldots, x_n) \neq (0, \ldots, 0)$ où $x_j \in K$.

On dit que K a la propriété C_i ($K \in C_i$) si pour tout $d \geqslant 1$, K a la propriété $C_i(d)$.

On appelle d-dimension diophantienne de K et on note $dd_d(K)$ le nombre réel $\mathrm{Inf}\{i / K \in C_i(d)\}$

On appelle dimension diophantienne de K et on note $dd(K)$ le nombre réel $\mathrm{Inf}\{i / K \in C_i\}$

Remarquons tout d'abord que

. $i < j$ alors $C_i(d) \Longrightarrow C_j(d)$

et $C_i \Longrightarrow C_j$

. $\forall\, i$ $K \in C_i(1)$ nous supposerons donc alors $d \geqslant 2$

. Si K est algébriquement clos alors $K \in C_o$.

Enoncé du théorème [8]

Soit K/K_o une extension de degré de transcendance $j \leqslant \infty$.
Si $K_o \in C_i$ alors $K \in C_{i+j}$

ou encore

$dd(K) \leqslant dd(K_o) + \deg tr\ (K/K_o)$; pour obtenir la proposition 4 il suffit de remarquer que K_o étant algébriquement clos $K_o \in C_o$, et K étant de degré de transcendance n sur K_o nous avons $K \in C_n$ et donc bien que tout polynôme homogène de $K[X_1, \ldots, X_\nu]$, de degré d avec $\nu > d^n$ a un zéro non trivial dans K^ν ceci quel que soit d entier.

3 Démonstration du Lemme 2

Lemme 2 : Soit R ordonné maximal, L une extension algébrique non ordonnable de $R(X_1, \ldots, X_n)$ alors -1 est somme de 2^n carrés dans L.

Démonstration :

Puisque R est ordonné maximal, R(i) est algébriquement clos.
L(i) / R(i) est une extension de degré de transcendance n, d'après la proposition 4 L(i) a la propriété C_n et donc évidemment la propriété $C_n(2)$.
L et L(i) vérifient donc les hypothèses de la proposition 3, ce qui entraîne donc que -1 est une somme de 2^n carrés dans L.

4 Démonstration du Lemme 3

Lemme 3 : K un corps ordonnable, $f \in K(X)$ une somme de carrés dans $K(X)$. Si -1 est somme de 2^{n-1} carrés dans tout corps L extension algébrique non ordonnable de K alors f est somme de 2^n carrés dans $K(X)$.

Démonstration :

. Il suffit de faire la démonstration dans le cas de $f \in K[X]$.
En effet, si $f \in K(X)$ alors $f = \frac{g}{h}$ g et $h \in K[X]$ et $f\,h^2 = gh$, $gh \in K[X]$, si on a déjà gh somme de 2^n carrés dans $K(X)$ alors fh^2 est somme de 2^n carrés dans $K(X)$ et donc f aussi en divisant chaque carré par h^2.

1. Si $f = 0$ c'est évident.

2. Si $f \neq 0$ on pose $f = a_o + a_1 X + \ldots + a_\ell X^\ell$ avec $a_i \in K$ et $a_\ell \neq 0$.

Montrons que nécessairement a_ℓ est somme de carrés dans K et que ℓ est pair.

Par hypothèse f est une somme de carrés dans $K(X)$ donc nous avons

$$f\,h^2 = \sum_{i=1}^{t} g_i^2 \qquad f,\ h \text{ et } g_i \in K[X]$$

Soit $r = \max_{1 \leqslant i \leqslant t} (\deg g_i)$

$$g_i = b_{io} + b_{i1} X + \ldots + b_{ir} X^r$$

$$h = C_o + C_1 X + \ldots + C_m X^m$$

En égalant les degrés des termes de plus haut degré nous avons $\ell + 2m = 2r$ car le terme de degré $2r$ ne peut disparaître dans $\Sigma\, g_i^2$, K étant un corps ordonnable.

Nous avons donc

$$a_\ell\, C_m^2 = \sum_{i=1}^{t} b_{ir}^2 \qquad \text{et} \qquad \ell + 2m = 2r$$

D'où ℓ est pair et a_ℓ est somme de carrés dans K.

. Posons alors $2k = \ell$.

a) Si $k = 0$ $\quad f = a_o$ et a_o est somme de carrés dans K donc dans K(X).

Vérifions que a_o est bien somme de 2^n carrés dans K(X).

Considérons le polynôme irréductible sur K $\quad X^2 + a_o$.
Soit α une racine de ce polynôme et $L = K(\alpha)$. L n'est évidemment pas ordonnable, on peut donc écrire d'après l'hypothèse du lemme 3 :

$$-1 = \sum_{i=1}^{2^{n-1}} (b_i + \alpha c_i)^2 \qquad b_i, c_i \in K$$

ou encore en identifiant les termes de K :

$$1 + \sum_{i=1}^{2^{n-1}} b_i^2 = -\alpha^2 \left(\sum_{i=1}^{2^{n-1}} c_i^2 \right) = a_o \left(\sum_{i=1}^{2^{n-1}} c_i^2 \right)$$

Nous avons $\sum_{i=1}^{2^{n-1}} c_i^2 \neq 0$ puisque K est ordonnable d'où

$$a_o = \frac{1 + \sum_{i=1}^{2^{n-1}} b_i^2}{\sum_{i=1}^{2^{n-1}} c_i^2}$$

qui est bien une somme de 2^n carrés en utilisant la proposition 1 et en écrivant :

$$d_o = \frac{(1 + \sum_{i=1}^{2^{n-1}} b_i^2) \left(\sum_{i=1}^{2^{n-1}} c_i^2 \right)}{\left(\sum_{i=1}^{2^{n-1}} c_i^2 \right)^2}$$

b) Si $k > 0$. $\quad f = a_o + a_1 X + \ldots + a_{2k} X^{2k}$

On sait que a_{2k} est somme de carrés dans K et en répétant le raisonnement fait pour a_o on obtient que a_{2k} est somme de 2^n carrés dans K.

. On peut supposer f unitaire : en effet
$\frac{f}{a_{2k}}$ est unitaire et si $\frac{f}{a_{2k}}$ est somme de 2^n carrés dans K(X),
par la proposition 1, $a_{2k} \frac{f}{a_{2k}}$ l'est aussi donc f est somme de 2^n carrés dans K(X).

. On peut supposer f sans facteur carré car si f est somme de 2^n carrés dans K(X), $g^2 f$ l'est aussi.

. D'après l'hypothèse nous pouvons donc écrire $f = \sum_{i=1}^{t} \frac{g_i^2}{h^2}$

Soit α une racine de f, p le polynome minimal de α sur K. Il nous suffit alors de montrer que p est somme de 2^n carrés dans K(X).
Car alors f étant le produit de polynômes tels que p, d'après la proposition 1, f sera somme de 2^n carrés dans K(X).

Il existe i $(1 \leqslant i \leqslant t)$ tel que $g_i(\alpha) \neq 0$. Car sinon p diviserait tous les g_i donc p^2 tous les g_i^2 et p^2 diviserait f, ce qui est contraire à l'hypothèse. On en déduit que :
$0 = f(\alpha)\ h^2(\alpha) = \sum_{i=1}^{t} (g_i(\alpha))^2$ les $g_i(\alpha)$ non tous nuls et donc le corps $K(\alpha)$ n'est pas ordonnable.

Soit $\ell = \deg(p) = [K(\alpha) : K] \geqslant 2$, puisque $K(\alpha)$ est non ordonnable et que K est ordonnable. D'après l'hypothèse du lemme 3, -1 est somme de 2^{n-1} carrés dans $K(\alpha)$ donc

$$-1 = \sum_{i=1}^{2^{n-1}} (h_i(\alpha))^2 \qquad h_i \in K[X]$$

et on peut supposer $\deg h_i \leqslant \ell - 1$ (car sinon on divise h par p et on obtient comme reste un polynome qui a la même valeur que h sur α)
On en déduit que
$1 + \sum_{i=1}^{2^{n-1}} h_i^2$ admet la racine α et donc qu'il existe $q \in K[X]$ tel que

$$1 + \sum_{i=1}^{2^{n-1}} h_i^2 = p\,q$$ on a donc :

$\deg q \leqslant 2(\ell - 1) - \ell$

$\deg q \leqslant \ell - 2$ et en comparant les termes de plus haut degré on obtient que le coefficient du terme de plus haut degré de q est une somme de 2^n carrés dans K.

. Si $\ell = 2$ $\deg q \leqslant 0$ donc $q = q_o \in K$ et q_o d'après ce qui précède est une somme de 2^n carrés dans K

Donc $$p = \frac{1 + \sum_{i=1}^{2^{n-1}} h_i^2}{q_o}$$ est une somme de 2^n carrés dans K(X) en écrivant

$$p = \frac{(1 + \sum_{i=1}^{2^{n-1}} h_i^2)\, q_o}{(q_o)^2}$$ et en appliquant la propriété 1 .

. On effectue un raisonnement par récurrence sur le degré de p.

Supposons la propriété

"K corps ordonnable, $L = K(\alpha)$ corps non ordonnable, $p \in K[X]$ polynôme minimal de α et -1 somme de 2^{n-1} carrés dans L entraîne p somme de 2^n carrés dans L alors p est somme de 2^n carrés dans K(X)" vraie jusqu'à p de degré $\ell - 1$ et démontrons la pour ℓ.

Reprenons l'identité

$$1 + \sum_{i=1}^{2^{n-1}} h_i^2 = p\,q.$$

Soit β une racine de q et $q' \in K[X]$ le polynôme minimal de β nous avons

$$\deg(q') \leqslant \deg(q) \leqslant \ell - 2 < \ell$$

puisque $q(\beta) = 0$ nous avons

$-1 = \sum_{i=1}^{2^{n-1}} h_i^2(\beta)$ et donc $L' = K(\beta)$ est non ordonnable et -1 somme de 2^{n-1} carrés dans L'. D'après l'hypothèse de récurrence q' est donc une somme de 2^n carrés dans $K(X)$.

Comme ceci est vrai pour tout polynôme unitaire irréductible q' qui divise q et pour le coefficient du terme de plus haut degré de q en utilisant la propriété 1 nous obtenons que q est somme de 2^n carrés dans $K(X)$.

Or p peut s'écrire

$$p = \frac{1 + \sum_{i=1}^{2^{n-1}} h_i^2}{q} = \frac{(1 + \sum_{i=1}^{2^{n-1}} h_i^2)\, q}{q^2} \quad \text{et en}$$

appliquant à nouveau le fait qu'un produit de deux sommes de 2^n carrés est une somme de 2^n carrés nous avons bien le résultat voulu : p somme de 2^n carrés dans $K(X)$ et donc le polynôme f initial est aussi de 2^n carrés dans $K(X)$.

En conclusion :

Dans le cas d' un corps ordonné maximal K *pour pouvoir décomposer toute fonction définie de* $K(X_1, \ldots, X_n)$ *en* $p(n)$ *carrés dans* $K(X_1, \ldots, X_n)$ *il suffit de* $p(n) \leqslant 2^n$ *carrés.*

En reprenant le résultat obtenu au chapitre 2, nous obtenons que dans le cas de K *ordonné maximal les valeurs de* $p(n)$ *vérifient :*

$$n + 1 \leqslant p(n) \leqslant 2^n$$

Dans le cas de $n = 0$, on retrouve que dans un corps ordonné maximal tout élément positif est un carré.

Dans le cas de $n = 1$ $\quad n + 1 = 2^n = 2$ et donc toute fonction définie de $K(X)$, K étant ordonné maximal, est somme de deux carrés dans $K(X)$. Ce résultat peut être trouvé de façon élémentaire et fera l'objet du chapitre 4.

Dans le cas n = 2 nous obtenons pour K ordonné maximal :
$3 \leqslant p(2) \leqslant 4$.

CHAPITRE 4 RESOLUTION ELEMENTAIRE DU CAS DE K(X) OU K EST ORDONNE MAXIMAL.

Il suffit d'étudier le cas de $p \in K[X]$. p est évidemment de degré pair et pour que p soit défini il faut que le coefficient de son terme de plus haut degré soit positif. On peut alors supposer p unitaire car le coefficient du terme de plus haut degré étant positif dans K ordonné maximal est un carré dans K. On sait (*) alors qu'un tel polynome se décompose dans $K[X]$ sous la forme :

$$p(X) = \prod_{i \in I} (X - a_i)^{\alpha_i} \prod_{j \in J} (X^2 + b_j X + c_j)^{\beta_j}$$

où les trinomes $X^2 + b_j X + c_j$ sont sans racine dans K. Pour que $p(x)$ soit positif ou nul quel que soit $x \in K$ il faut que α_i soit pair donc $\alpha_i = 2\alpha'_i$. D'autre part le trinôme peut se mettre dans $K[X]$ sous la forme :

$X^2 + b_j X + c_j = (X - p_j)^2 + q_j^2$ puisqu'il n'a pas de racine dans K (*)

p s'écrit alors :

$$p(X) = \prod_{i \in I} (X - a_i)^{2\alpha'_i} \prod_{j \in J} ((X - p_j)^2 + q_j^2)^{\beta_j}$$

$$p(X) = \left[\prod_{i \in I} (X - a_i)^{\alpha'_i}\right]^2 \prod_{j \in J} ((X - p_j)^2 + a_j^2)^{\beta_j}$$

Or dans tout anneau commutatif nous avons l'égalité :

$(a^2 + b^2)(c^2 + d^2) = (ac - bd)^2 + (ad + bc)^2$

Donc le produit de deux sommes de deux carrés est une somme de deux carrés.

En itérant ce résultat nous pouvons donc affirmer que

$\prod_{j \in J} ((X - p_j)^2 + q_j^2)^{\beta_j}$ est une somme de deux carrés dans $K[X]$.

(*) Voir Bourbaki II-6-§ 2, n° 6.

Nous avons donc le résultat suivant :

Théorème 9

Soit K *ordonné maximal* $p \in K[X]$ *un polynôme défini alors* p *est somme de deux carrés dans* $K[X]$.

et nous en déduisons immédiatement :

Soit K *ordonné maximal* $f \in K(X)$ *une fonction définie alors* f *est somme de deux carrés dans* $K(X)$.

CHAPITRE 5 RESOLUTION DU CAS DE $\mathbb{R}(X, Y)$

Nous allons étudier ici un résultat dû à un travail récent de Messieurs Cassels, Ellison et Pfister.

Théorème 10 [27]

Le polynôme $f(X, Y) = 1 + X^2 (X^2 - 3)Y^2 + X^2 Y^4$ *de* $\mathbb{R}[X, Y]$ *n'est pas une somme de trois carrés dans* $\mathbb{R}(X, Y)$.

Nous avons vu que ce polynôme f(X, Y) était tel que pour tout $(x, y) \in \mathbb{R} \times \mathbb{R}$, $f(x, y) \geqslant 0$ (voir page 50). Donc d'après la conclusion du chapitre 3, f(X, Y) est somme de 3 ou 4 carrés dans $\mathbb{R}(X, Y)$.

Le théorème 10 affirmant que f(X, Y) n'est pas somme de trois carrés dans $\mathbb{R}(X, Y)$ nous pouvons alors énoncer le résultat suivant :

Théorème 11

Pour pouvoir décomposer en somme de carrés dans $\mathbb{R}(X, Y)$ *toute fonction* $f(X, Y) \in \mathbb{R}(X, Y)$ *telle que* $\forall (x, y) \in \mathbb{R} \times \mathbb{R}$, $f(x, y) \geqslant 0$, *4 carrés sont nécessaires (et suffisants d'après le chapitre 3).*

1 . Résultats préliminaires

Lemme 1 : Soit k un corps ordonnable. Soit $f(y) = 1 + ay^2 + by^4$ un élément de $k[y]$ avec $b \neq 0$ et $a^2 \neq 4b$.

f(y) est somme de trois carrés dans k(y) si et seulement si la courbe elliptique $\mathcal{C}^{-1}$: $-\eta^2 = \xi(\xi^2 - 2a\xi + a^2 - 4b)$ a un point k-rationnel (ξ, η) avec ξ, η dans k tels que ξ et $-\xi^2 + 2a\xi - a^2 + 4b$ soient sommes de deux carrés dans k.

Supposons que f(y) est somme de trois carrés dans k(y). Alors puisque f(y) est un polynôme d'après le résultat du théorème 6, f(y) est somme de trois carrés dans $k[y]$, donc :

$$f = f_1^2 + f_2^2 + f_3^2 \qquad \text{où} \quad f_i \in k[y]$$

Puisque k est ordonnable et que le degré de f est 4 les f_i sont de degré 2 au plus (sinon on aurait une identité du type $0 = d_1^2 + d_2^2 + d_3^2$, en identifiant les termes de plus haut degré des deux membres, ce qui est impossible dans un corps ordonnable).

Posons donc $f_i(y) = a_i + b_i\, y + c_i\, y^2 \quad (i = 1, 2, 3)$. En portant ceci dans la relation $f = f_1^2 + f_2^2 + f_3^2$ et en identifiant les termes de même degré nous obtenons :

$$\sum_1^3 a_i^2 = 1$$

$$\sum_1^3 a_i\, b_i = 0$$

$$\sum_1^3 b_i^2 + 2 \sum_1^3 a_i\, c_i = a$$

$$\sum_1^3 b_i\, c_i = 0$$

$$\sum_1^3 c_i^2 = b$$

Effectuons sur l'espace k^3 une transformation orthogonale telle que le vecteur (a_1, a_2, a_3) devienne le vecteur $(1, 0, 0)$

Le système précédent devient alors

$$\begin{cases} a_1\, b_1 = 0 \\ b_1^2 + b_2^2 + b_3^2 + 2\, a_1\, c_1 = a \\ b_1\, c_1 + b_2\, c_2 + b_3\, c_3 = 0 \\ c_1^2 + c_2^2 + c_3^2 = b \end{cases}$$

D'où (*)
$$\begin{cases} b_1 = 0 \\ b_2^2 + b_3^2 = a - 2c_1 \\ b_2 c_2 + b_3 c_3 = 0 \\ c_2^2 + c_3^2 = b - c_1^2 \end{cases}$$

Ceci entraîne :

$$\begin{aligned}(a - 2c_1)(b - c_1^2) &= (b_2^2 + b_3^2)(c_2^2 + c_3^2) \\ &= (b_2c_2 + b_3c_3)^2 + (b_2c_3 - b_3c_2)^2 \\ &= (b_2c_3 - b_3c_2)^2\end{aligned}$$

Posons $\xi = a - 2c_1$

$\eta = 2(b_2c_3 - b_3c_2)$

Alors $4(b - c_1^2) = 4b - 4c_1^2 = 4b - (\xi - a)^2$ en reportant dans la relation que nous venons de trouver on obtient :

$$\frac{\xi(4b - (\xi - a)^2)}{4} = \frac{\eta^2}{4} \quad \text{donc}$$

$$\xi((\xi - a)^2 - 4b) = -\eta^2$$

Vérifions les conditions sur ξ et $(\xi - a)^2 - 4b$:

$\xi = a - 2c_1 = b_2^2 + b_3^2$ est bien somme de deux carrés.

$-[(\xi - a)^2 - 4b] = 4(b - c_1^2) = 4(c_2^2 + c_3^2)$ est bien somme de deux carrés également.

Réciproquement :

Soient (ξ, η) vérifiant les hypothèses du lemme.

Alors

a) Si $\xi = 0$ $\quad 4b - a^2 = d^2 + e^2$ d'après l'hypothèse

on obtient une solution du système (*) en prenant $b_1 = b_2 = b_3 = 0$, $2c_1 = a$, $2c_2 = d$, $2c_3 = e$.

b) Si $\xi = b_2^2 + b_3^2 \neq 0$. On obtient

$$4b - (\xi - a)^2 = \left(\frac{\eta}{\xi}\right)^2 (b_2^2 + b_3^2)$$

Et on a une solution du système (*) en prenant :

$b_1 = 0$, $2c_1 = a - \xi$, $2c_2 = \frac{\eta}{\xi} b_3$, $2c_3 = -\frac{\eta}{\xi} b_2$.

Dans tous les cas on a une solution du système (*) et donc la possibilité de représenter f(y) comme somme de trois carrés.

Corollaire du lemme 1

$f(x, y) = 1 + x^2(x^2 - 3)y^2 + x^2 y^4$ est somme de trois carrés dans $\mathbb{R}(x, y)$ si et seulement si la courbe elliptique
$\mathcal{C}^{-1} : -\eta^2 = \xi(\xi - x^2(x^2 - 3) - 2x)(\xi - x^2(x^2 - 3) + 2x)$ a un point (ξ, η) rationnel sur $\mathbb{R}(x)$ tel que $\eta \neq 0$ et ξ est positive ou nulle sur $\mathbb{R}$.

Appliquons le lemme 1 avec $k = \mathbb{R}(x)$, $a = x^2(x^2 - 3)$ et $b = x^2$.

On en déduit que f(x, y) sera une somme de trois carrés dans $\mathbb{R}(x, y)$ si et seulement si la courbe elliptique
$\mathcal{C}^{-1} : \eta^2 = \xi(\xi^2 - 2x^2(x^2 - 3)\xi + x^4(x^2 - 3)^2 - 4x^2)$

$= \xi(\xi - x^2(x^2 - 3) - 2x)(\xi - x^2(x^2 - 3) + 2x)$

a un point (ξ, η) rationnel sur $\mathbb{R}(x)$ tel que ξ et

$-(\xi - x^2(x^2 - 3) - 2x)(\xi - x^2(x^2 - 3) + 2x)$ soient sommes de deux carrés dans $\mathbb{R}(x)$.

Si $\eta = 0$ alors $\xi = 0$ ou $\xi = x^2(x^2 - 3) \pm 2x$.

- Le premier point (0, 0) ne peut répondre aux conditions ci-dessus car alors il faudrait que
$4x^2 - (x^2(x^2 - 3))^2 = -x^2(x^2 - 1)^2(x^2 - 4)$ soit somme de deux carrés dans $\mathbb{R}(x)$.

Ceci est impossible puisque les éléments sommes de deux carrés dans $\mathbb{R}(x)$ sont toujours positifs sur $\mathbb{R}$, ce qui n'est pas le cas de $-x^2(x^2 - 1)^2(x^2 - 4)$.

- De même les deux autres points ne peuvent convenir non plus car
$\xi = x^2(x^2 - 3) \pm 2x$ doit être somme de deux carrés dans $\mathbb{R}(x)$.
Or $x^2(x^2 - 3) \pm 2x = x(x \mp 1)^2(x \pm 2)$ qui ne peut rester positif quelle que soit la valeur réelle donnée à x et qui ne peut donc pas être somme de deux carrés dans $\mathbb{R}(x)$.

Donc s'il existe un point (ξ, η) vérifiant les hypothèses alors $\eta \neq 0$.

Si $\xi = 0$ sans que $\eta = 0$ alors nous aurions $-\eta^2 = x^2(x^2 - 1)^2(x^2 - 4)$ ce qui est impossible car $-\eta^2$ est toujours négatif sur $\mathbb{R}$, ce qui n'est pas le cas du second membre.

Nous pouvons donc supposer $\xi \neq 0$

Alors de

$$-\eta^2 = \xi(\xi - x^2(x^2 - 3) - 2x)(\xi - x^2(x^2 - 3) + 2x)$$

on déduit

$$\frac{-\eta^2}{\xi} = (\xi - x^2(x^2 - 3) - 2x)(\xi - x^2(x^2 - 3) + 2x)$$

Mais si $\xi = A^2 + B^2$ dans $\mathbb{R}(x)$ alors

$$\frac{\eta^2}{A^2 + B^2} = \frac{A^2\eta^2 + B^2\eta^2}{(A^2 + B^2)^2} = \left(\frac{A\eta}{A^2 + B^2}\right)^2 + \left(\frac{B\eta}{A^2 + B^2}\right)^2$$

et nous avons donc

$-(\xi - x^2(x^2 - 3) - 2x)(\xi - x^2(x - 3) + 2x)$ qui est bien somme de deux carrés dans $\mathbb{R}(x)$.

Pour que ξ soit somme de deux carrés dans $\mathbb{R}(x)$ il faut et il suffit que ξ ne prenne que des valeurs positives ou nulles sur $\mathbb{R}$. D'où le résultat annoncé.

<u>Lemme 2</u> : S'il existe ξ et η dans $\mathbb{R}(x)$ tels que $-\eta^2 = \xi(\xi - x^2(x^2 - 3) - 2x)(\xi - x^2(x^2 - 3) + 2x)$, avec ξ somme de deux carrés dans $\mathbb{R}(x)$ alors ξ est égal à un carré dans $\mathbb{R}(x)$.

Tout d'abord on peut supposer que $\xi = \frac{P}{Q^2}$ avec PGCD $(P, Q) = 1$.

En effet si ξ est somme de deux carrés dans $\mathbb{R}(x)$ alors

$$\xi = \left(\frac{a}{b}\right)^2 + \left(\frac{c}{d}\right)^2 \quad \text{avec } a, b, c \text{ et } d \in \mathbb{R}[x]$$

PGCD $(a, b) = 1$, PGCD $(c, d) = 1$

d'où $\xi = \frac{a^2 d^2 + b^2 c^2}{(b\,d)^2} = \frac{\omega}{\psi^2}$.

Si p irréductible divise ψ^2, il divise $\psi = b\ d$ donc il divise b ou il divise d et de plus p^2 divise ψ^2.

Si p divise b et que p divise ω alors p divise a d mais p divisant b ne peut diviser a donc il divise d. Donc p divise b et d.

Si p divise d et ω on a de même p divise b et d.

Mais si p divise d, p^2 divise d^2 ; si p divise b, p^2 divise b^2 ; donc p^2 divise ω et puisque p^2 divise ω et ψ^2 on peut simplifier.

En itérant on peut donc supposer que

$$\xi = \frac{\omega}{\psi^2} \quad \text{avec PGCD } (\omega, \psi) = 1. \quad \omega \text{ et } \psi \in \mathbb{R}[x]$$

Remplaçons ξ par cette valeur dans l'équation de $\mathcal{C}^{-1}$. Nous obtenons :

$$-\eta^2(\psi^3)^2 = \omega(\omega^2 - 2x^2(x^2-3)\omega\psi^2 + x^2(x-1)^2(x+1)^2(x-2)(x+2)\psi^4)$$

Puisque le membre de droite est un polynôme, le membre de gauche est aussi un polynôme de $R[x]$.

Soit $\eta^2(\psi^3)^2 = H^2$. Soit $\omega = \omega_1\ \theta^2$ avec ω_1 sans facteur carré, alors dans $\mathbb{R}[X]$:

$$-H^2 = \omega_1\ \theta^2(\omega^2 - 2x^2(x^2-3)\omega\psi^2 + x^2(x-1)^2(x+1)^2(x-2)(x+2)\psi^4)$$

θ^2 divisant le membre de droite divise H^2 et en simplifiant nous obtenons dans $\mathbb{R}[X]$

$$-H_1^2 = \omega_1(\omega^2 - 2x^2(x^2-3)\omega\psi^2 + x^2(x-1)^2(x+1)^2(x-2)(x+2)\psi^4)$$

Soit alors p premier diviseur de ω_1. Alors p divise H_1^2 donc p divise H_1 et p^2 divise H_1^2. Mais puisque p^2 ne divise pas ω_1 et que p^2 divise H_1^2, c'est que p divise $(\omega^2 - 2x^2(x^2-3)\omega\psi^2 + x^2(x-1)^2(x+1)^2(x+2)(x-2)\psi^4)$

p divisant déjà ω c'est donc que p divise $x^2(x-1)^2(x+1)^2(x-2)(x+2)\psi^4$

Mais p divise ω donc p ne divise pas ψ puisque PGCD $(\omega, \psi) = 1$.

Donc finalement : si p premier divise ω_1 alors p divise $x^2(x-1)^2(x+1)^2(x-2)(x+2)$. ω_1 étant sans facteur carré on en déduit que

$$\omega_1 = \lambda\ x^{\alpha}(x-1)^{\beta}(x+1)^{\gamma}(x-2)^{\delta}(x+2)^{\varepsilon}$$

où $\alpha, \beta, \gamma, \delta, \varepsilon$ sont égaux à 0 ou 1.

Revenons à $\xi = \frac{\omega}{\psi^2} = \frac{\omega_1\ \theta^2}{\psi^2}$ puisque ξ est toujours positif sur $\mathbb{R}$ et que $\frac{\theta^2}{\psi^2}$ l'est aussi, il faut que ω_1 soit toujours positif sur $\mathbb{R}$. Ceci entraîne $\alpha = \beta = \gamma = \delta = \varepsilon = 0$, $\lambda > 0$, donc finalement $\xi = \frac{\lambda\ \theta^2}{\psi^2} = (\frac{\theta\sqrt{\lambda}}{\psi})^2$ et ξ est un carré dans $\mathbb{R}(X)$.

Indiquons quelques notations avant de donner deux autres lemmes. Nous désignerons par $\mathcal{C}_K^{-1}$ le groupe des points de la courbe elliptique $\mathcal{C}^{-1}$ définis sur K. Nous désignerons par $\mathcal{C}_K$ le groupe des points définis sur K de la courbe elliptique.

$$\mathcal{C} : \eta^2 = \xi(\xi - x^2(x^2 - 3) - 2x)(\xi - x^2(x^2 - 3) + 2x)$$

Lemme 3 : $\mathcal{C}_{\mathbb{C}(x)} = \mathcal{C}_{\mathbb{Q}(x)}$

Lemme 4 : $\mathcal{C}_{\mathbb{C}(x)} = \mathcal{C}_{\mathbb{R}(x)}$ sauf s'il existe un point (ξ, η) de $\mathcal{C}^{-1}_{\mathbb{R}(x)}$ tel que $\xi \in \mathbb{R}(x)^{*2}$.

Pour les démonstrations des lemmes 3 et 4 qui font appel à la théorie des courbes elliptiques, voir [27] (Lemme 5-7 et Théorème 7-1).

2 - Démonstration du théorème

Remarquons que $k \subset K$ entraîne $\mathcal{C}_k \subset \mathcal{C}_K$. Donc si $\mathcal{C}_{\mathbb{Q}(x)} = \mathcal{C}_{\mathbb{C}(x)}$ puisque $\mathbb{Q}(x) \subset \mathbb{R}(x) \subset \mathbb{C}(x)$. Nous en déduisons que

$$\mathcal{C}_{\mathbb{C}(x)} = \mathcal{C}_{\mathbb{R}(x)}$$

En appliquant le lemme 4, nous obtenons donc qu'il n'existe pas (ξ, η) dans $\mathcal{C}^{-1}_{\mathbb{R}(x)}$ tel que $\xi \in \mathbb{R}(x)^{*2}$.

Puisque si nous avons $(\xi, \eta) \in \mathcal{C}^{-1}_{\mathbb{R}(x)}$ vérifiant les hypothèses du corollaire du lemme 1, $\xi \neq 0$ et ξ est somme de deux carrés dans $\mathbb{R}(x)$ et que d'après le lemme 2 un tel ξ est finalement un carré dans $\mathbb{R}(x)$. Nous obtenons donc qu'il n'existe pas de point $(\xi, \eta) \in \mathcal{C}^{-1}_{\mathbb{R}(x)}$ et vérifiant les hypothèses du corollaire du lemme 1 et donc que le polynôme $f(x, y) = 1 + x^2(x^2 - 3)y^2 + x^2 y^4$ n'est pas somme de trois carrés dans $R(x, y)$. D'où le théorème 11 énoncé au début de ce chapitre.

CHAPITRE 6 PROBLEMES ANALOGUES

Nous citerons ici quelques résultats récents concernant le 17è problème de Hilbert posé cette fois dans le cas de corps autres que $\mathbb{R}$.

1 . Cas de $\mathbb{Q}(X)$

En 1906 dans [20] Landau avait montré que dans le cas de $\mathbb{Q}(X)$ 8 carrés de polynômes suffisaient pour pouvoir décomposer tout polynôme de $\mathbb{Q}[X]$ ne prenant sur $\mathbb{Q}$ que des valeurs positives ou réelles.

Dans son article [28] paru en 1971 Yves Pourchet montre que ce nombre peut être ramené à 5 sans pouvoir être abaissé davantage.

Nous avons donc le théorème suivant :

Théorème 12

Soit $p \in \mathbb{Q}[X]$ *tel que* $\forall a \in \mathbb{Q} \quad p(a) \geqslant 0$ *alors* p *s'écrit sous la forme*

$$p = \sum_{i=1}^{5} g_i \qquad g_i \in \mathbb{Q}[X]$$

et il existe de tels polynômes p *qui ne sont pas somme de 4 carrés dans* $\mathbb{Q}[X]$.

Notons que l'auteur donne également dans cet article une caractérisation des sommes de quatre carrés dans $\mathbb{Q}[X]$.

Notons également ici que dans le cas de $\mathbb{Q}(X, Y)$ on ignore si on peut trouver une borne supérieure pour le nombre de carrés nécessaires à la décomposition des fonctions de $\mathbb{Q}(X, Y)$ qui sont positives sur tout élément $(a, b) \in \mathbb{Q}^2$.

2 - Cas de $Q^{\frac{1}{2}}(X)$

Dans un exposé fait à Toulouse en Mai 1971, W.J. Ellison s'intéresse au 17è problème de Hilbert sur le corps $\mathbb{Q}^{\frac{1}{2}}$. Par $\mathbb{Q}^{\frac{1}{2}}$ on désigne la clôture quadratique réelle de $\mathbb{Q}$.

Théorème :

Si $f(X) \in \mathbb{Q}^{\frac{1}{2}}(X)$ *est somme de carrés dans* $\mathbb{Q}^{\frac{1}{2}}(X)$ *alors* $f(X)$ *peut être écrit comme somme d'au plus quatre carrés dans* $\mathbb{Q}^{\frac{1}{2}}(X)$.

Notons qu'il existe des polynômes de $\mathbb{Q}^{\frac{1}{2}}(X)$ qui sont somme de carrés et qui ne sont pas somme de 2 carrés dans $\mathbb{Q}^{\frac{1}{2}}(X)$, mais qu'on ignore si trois carrés suffisent pour écrire dans $\mathbb{Q}^{\frac{1}{2}}(X)$ tout polynôme somme de carrés dans $\mathbb{Q}^{\frac{1}{2}}(X)$.

3 - Corps K non commutatif

Nous avons page 11 exclu le cas de K ordonné non commutatif pour poser le 17è problème de Hilbert sous sa forme classique. Pourtant si K est un corps non commutatif, les sommes de produits de carrés jouant un rôle semblable à celui des sommes de carrés dans le cas commutatif, il est raisonnable de se demander si on ne pourrait pas généraliser le 17è problème de Hilbert au cas non commutatif en cherchant à décomposer les fonctions en sommes de produits de carrés.

4 - Corps K commutatif ayant plus d'un ordre

Dans ce cas on pourrait généraliser le $17^{\text{ème}}$ problème de Hilbert dans $K(X_1, \ldots, X_n)$ en cherchant à décomposer en somme de carrés dans $K(X_1, \ldots, X_n)$ les fonctions $f \in K(X_1, \ldots, X_n)$ qui sont telles que pour tout $(a_1, \ldots, a_n) \in K^n$, $f(a_1, \ldots, a_n)$ est totalement positif dans K.

En conclusion à cette partie nous pouvons donner le résumé suivant :

Soit K un corps tel que toute fonction définie de $K(X_1, \ldots, X_n)$ soit décomposable en somme de carrés dans $K(X_1, \ldots, X_n)$. Soit alors $P_K(n)$ le nombre de carrés permettant de décomposer toutes ces fonctions (K et n étant fixés).

Nous pouvons alors donner sur $P_K(n)$ le tableau suivant :

corps K \ Nombre de variables n	0	1	2	$n \geqslant 3$
$\mathbb{R}$	1	2	4	$n+2 \leqslant P(n) \leqslant 2^n$
K ordonné maximal	1	2	$3 \leqslant P(2) \leqslant 4$	$n+1 \leqslant P(n) \leqslant 2^n$
$\mathbb{Q}$	4	5	?	?

3EME PARTIE

17È PROBLEME DE HILBERT DANS IR(V)

Nous nous intéresserons ici au problème suivant :

Soit V une variété algébrique réelle de $\mathbb{R}^n$ (c'est-à-dire une partie de $\mathbb{R}^n$ définie par des équations polynômes), soit $\mathcal{J}(V)$ l'idéal associé à V (c'est-à-dire l'ensemble des polynômes nuls sur V).

Nous supposerons la variété irréductible et donc l'idéal $\mathcal{J}(V)$ premier. Soit alors $\mathbb{R}[V] = \mathbb{R}[X_1, \ldots, X_n]/\mathcal{J}(V)$, qui est donc intègre, et soit enfin $\mathbb{R}(V)$ son corps des fractions.

Soit alors $F \in \mathbb{R}(V)$ (F s'interprète comme une fonction définie presque partout sur V), telle que F est positive ou nulle sur les éléments de V.

A-t-on alors dans $\mathbb{R}(V)$ $F = \sum_{i=1}^{p} G_i^2$? Et si oui, peut-on trouver p(n) tel que toute fonction F puisse se décomposer en au plus p(n) carrés ?

Au chapitre 1, nous essaierons d'obtenir un résultat en appliquant les résultats connus dans $\mathbb{R}(X_1, \ldots, X_n)$. Nous obtiendrons ainsi un résultat partiel mais à la fois qualitatif et quantitatif et nous montrerons l'impossibilité d'obtenir un résultat général par cette méthode.

Au chapitre 2, nous étudierons un résultat de A.R. Robinson qui permet d'obtenir un résultat qualitatif général mais ne donne pas d'indication quantitative.

Remarquons que si $V = \mathbb{R}^n$, alors $\mathcal{J}(V) = (0)$ et $\mathbb{R}[V] = \mathbb{R}[X_1, \ldots, X_n]$; Dans ce cas le problème a déjà été résolu.

De même si $V = \emptyset$, alors $\mathcal{J}(V) = \mathbb{R}[X_1, \ldots, X_n]$ et le seul élément de $\mathbb{R}[V]$ est 0 qui est bien un carré.

Dans la suite nous supposerons donc $V \neq \mathbb{R}^n$ et donc $\mathcal{J}(V) \neq (0)$, et $V \neq \emptyset$ et donc $\mathcal{J}(V) \neq \mathbb{R}[X_1, \ldots, X_n]$.

CHAPITRE 1 RESOLUTION QUALITATIVE ET QUANTITATIVE PAR APPLICATION DES RESULTATS DES PARTIES 1 ET 2

I - Réductions du problème (Les notations sont celles définies à la page précédente).

1°) Il suffit de faire l'étude pour $F \in \mathbb{R}[V]$.
En effet, soit $F = \frac{P}{Q}$ avec P et Q dans $\mathbb{R}[V]$.
Si le problème est résolu pour les éléments de $\mathbb{R}[V]$ alors $F = \frac{PQ}{Q^2}$,
donc PQ est un élément de $\mathbb{R}[V]$ positif sur V et donc

$$PQ = \sum_{i=1}^{p} G_i^2 \qquad G_i \in \mathbb{R}(V).$$

On en déduit aussitôt $F = \frac{\sum_{i=1}^{p} G_i^2}{Q^2} = \sum_{i=1}^{p} \left(\frac{G_i}{Q}\right)^2$

et puisque $\left(\frac{G_i}{Q}\right) \in \mathbb{R}(V)$, nous avons bien une solution du problème pour F.

2°) Il est possible de se ramener à un problème dans $\mathbb{R}[X_1, \ldots, X_n]$.
Soit $f \in \mathbb{R}[X_1, \ldots, X_n]$; si f est positive sur les éléments de V alors $g = f + h$ où $h \in \mathcal{J}(V)$ est aussi positive sur les éléments de V.
Si F appartient à $\mathbb{R}[V]$ et est positive ou nulle sur V, quelle que soit f de $\mathbb{R}[X_1, \ldots, X_n]$ telle que la classe $\bar{f}$ de f dans $\mathbb{R}[X_1, \ldots, X_n] / \mathcal{J}(V)$ soit $\bar{f} = F$ alors sur $(a_1, \ldots, a_n) \in V$ nous avons

$f(a_1, \ldots, a_n) = F(a_1, \ldots, a_n)$.

Donc si $F \in \mathbb{R}[V]$ est positive sur V, il en est de même de toute fonction $f \in \mathbb{R}[X_1, \ldots, X_n]$ représentant de $F \in \mathbb{R}[V]$.

Nous allons donc chercher un élément $g = f + h$ dans $\mathbb{R}[X_1, \ldots, X_n]$, avec $h \in \mathcal{J}(V)$ qui soit tel que g soit non seulement positif ou nul sur V mais aussi sur $\mathbb{R}^n$ tout entier afin de pouvoir alors appliquer les résultats d'Artin et de Pfister.

II - Démonstration dans un cas particulier

Notons $N_f = \{x \in \mathbb{R}^n, f(x) < 0\}$.

(Remarquons que N_f change si l'on change de représentant pour $F = \bar{f}$)

Supposons que $\mathcal{J}(V)$ est principal.
Supposons que $\overline{N_f}$ est borné et vérifie $\overline{N_f} \cap V = \emptyset$.

A . Utilisation des résultats acquis

Soit $F \in \mathbb{R}[V]$ positif ou nul sur V, f telle que $\bar{f} = F$ et que les hypothèses sur N_f ci-dessus soient vérifiées. Puisque $\mathcal{J}(V)$ est supposé principal, soit h le polynôme engendrant $\mathcal{J}(V)$ (h est non nul car $\mathcal{J}(V)$ est supposé différent de (0)).Alors en tout point x de $\overline{N_f}$, h(x) est non nul puisque $\overline{N_f} \cap V = \emptyset$.

Posons alors $a = \inf_{x \in \overline{N_f}} f(x)$ $\qquad -\infty < a \leqslant 0$

et $b = \inf_{x \in \overline{N_f}} (h(x))^2$

$\overline{N_f}$ étant fermé et borné dans $\mathbb{R}^n$ est un compact de $\mathbb{R}^n$ et donc $\inf_{x \in \overline{N_f}} (h(x))^2$ est atteint en un point de $\overline{N_f}$.

Cet infimum est donc différent de zéro puisque h(x) est non nul en chaque point de $\overline{N_f}$.

Il est positif puisque h^2 ne peut prendre que des valeurs positives ou nulles d'après son expression.
Donc $0 < b < \infty$.

Cherchons maintenant $\lambda \in \mathbb{R}$ tel que $g = f + \lambda h^2$ soit positif ou nul sur $\mathbb{R}^n$ tout entier.
Posons $\lambda = -\frac{a}{b} \geqslant 0$ et $\lambda < \infty$

. pour $x \notin \overline{N_f}$

$g(x)$ est positive ou nulle puisque $f(x)$ l'était et que $\lambda h^2 = -\frac{a}{b} h^2$ est positif ou nul sur tout $\mathbb{R}^n$ par construction.

. pour $x \in \overline{N_f}$ nous avons

$g(x) = f(x) - \frac{a}{b} h^2(x) \geqslant a + -\frac{a}{b} b = 0$

Donc $g(x) \geqslant 0$.

Nous avons donc bien trouvé une fonction $g \in \mathbb{R}[X_1, \ldots, X_n]$ telle que $\bar{g} = \bar{f} = F$ et $g(x) \geqslant 0$ pour tout $x \in \mathbb{R}^n$.
Alors par le théorème d'Artin et un résultat de Pfister nous pouvons affirmer que

$$g = \sum_{i=1}^{2^n} \left(\frac{g_i}{h_i}\right)^2 \qquad g_i \text{ et } h_i \text{ dans } \mathbb{R}[X_1, \ldots, X_n]$$

B . Passage à $R[X_1, \ldots, X_n]/\mathcal{J}_{(V)}$

L'égalité précédente peut s'écrire en réduisant au même dénominateur :

$$g\,k^2 = \sum_{i=1}^{2^n} k_i^2 \qquad k_i \text{ et } k \in \mathbb{R}[X_1, \ldots, X_n]$$

Nous en déduisons l'égalité dans $\mathbb{R}[V]$

$$\bar{g}\,\bar{k}^2 = \sum_{i=1}^{2^n} \bar{k}_i^2$$

Si on démontre que l'on peut avoir une telle égalité avec $\bar{k}$ non nul (c'est-à-dire $k \notin \mathcal{J}(V)$), on pourra en déduire l'égalité dans $\mathbb{R}(V)$,

$$\bar{g} = \sum_{i=1}^{2^n} \left(\frac{\overline{k_i}}{\bar{k}}\right)^2$$ qui est le résultat cherché.

Nous allons montrer qu'on peut effectivement se ramener à une telle égalité avec $k \notin \mathcal{J}(V)$.

Puisque $\mathcal{J}(V)$ est principal et engendré par h, si $k \in \mathcal{J}(V)$ alors $k = h\,q$.

Mais si $k \in \mathcal{J}(V)$ alors chacun des $k_i \in \mathcal{J}(V)$. En effet si $x \in V$, $k(x) = 0$ et donc $\sum_{i=1}^{2^n} k_i^2(x) = 0$. Cette égalité dans $\mathbb{R}$ entraîne $k_i(x) = 0$ quel que soit i. Ceci étant vrai pour tout $x \in V$, on en déduit que $k_i \in \mathcal{J}(V)$; donc $k_i = h\,q_i$.

En remplaçant dans l'égalité ci-dessus on obtient $g\,h^2\,q^2 = \sum_{i=1}^{2^n} h^2\,q_i^2$ et on peut simplifier par h^2.

En itérant le procédé on peut donc se ramener à une égalité dans $\mathbb{R}[X_1, \ldots, X_n]$:

$$g\,k'^2 = \sum_{i=1}^{2^n} (k'_i)^2 \quad \text{avec } k' \notin \mathcal{J}(V).$$

Alors $\bar{k}' \neq 0$ et donc nous avons dans $\mathbb{R}(V)$

$$\bar{g} = \sum_{i=1}^{2^n} \left(\frac{\overline{k'_i}}{\bar{k}}\right)^2$$

Nous pouvons donc énoncer :

Théorème 13

Soit V *une variété algébrique irréductible, dont l'idéal* $\mathcal{J}(V)$ *est principal. Soit* $F \in \mathbb{R}[V] = \mathbb{R}[X_1, \ldots, X_n] / \mathcal{J}(V)$ *telle que* F *soit positive ou nulle sur* V *et telle qu'il existe* $f \in \mathbb{R}[X_1, \ldots, X_n]$ *telle que* $\bar{f} = F$ *et que si on note* $N_f = \{x \in \mathbb{R}^n,\ f(x) < 0\}$, *alors* $\overline{N_f}$ *est bornée et* $\overline{N_f} \cap V = \emptyset$.

Alors $F = \sum_{i=1}^{2^n} G_i^2$ *où* $G_i \in \mathbb{R}(V)$ *corps des fractions de* $\mathbb{R}[V]$.

Le résultat se généralise immédiatement aux fonctions $F \in \mathbb{R}(V)$ telles que $F = \frac{P}{Q}$ où $P \cdot Q$ vérifie les hypothèses du théorème ci-dessus d'après la remarque faite au début de ce chapitre.

Remarque

Les résultats de minoration du nombre de carrés nécessaires à la décomposition obtenus par MM. Cassels et Ellison ne peuvent donner de conséquences ici car certains des carrés peuvent s'annuler lorsqu'on passe de $\mathbb{R}[X_1, \ldots, X_n]$ à $\mathbb{R}[V]$.

III - Non possibilité d'une démonstration par application des résultats des parties 1 et 2 dans le cas général.

Soit V la courbe de $\mathbb{R}^2$ définie par $\begin{cases} X = t^2 \\ Y = t^3 \end{cases}$ (parabole semi-cubique).

L'ensemble des polynômes s'annulant sur V est l'idéal engendré par $Y^2 - X^3$ (*), $\mathcal{P}(V)$ est donc principal. Soit alors $f(X, Y) = X$. Le polynôme est bien positif sur V. Mais $\overline{N_f} \cap V \neq \emptyset$ puisque $\overline{N_f} \cap V = \{ (0, 0) \}$

On peut se ramener à un représentant g de $F = \bar{f}$ tel que N_g soit borné.

(*) Pour le vérifier soit $P(X, Y)$ s'annulant sur V ; on peut diviser $P(X, Y)$ par $Y^2 - X^3$ dans $\mathbb{R}[X]\,[Y]$ puisque le terme de plus haut degré en Y appartient à $\mathbb{R}$.

$$P(X, Y) = Q(X, Y)(Y^2 - X^3) + S(X)Y + T(X)$$

Pour $X = t^2$ $Y = t^3$ nous obtenons :

$$S(t^2)t^3 + T(t^2) = 0$$

$S(t^2)t^3$ est un polynôme en t dont les puissances sont impaires alors que celles du polynôme $T(t^2)$ sont paires. Ceci ayant lieu pour tout $t \in \mathbb{R}$ il faut que $S = T = 0$ et donc que $P(X, Y)$ soit multiple de $Y^2 - X^3$.

En effet considérons $g = X + (Y^2 - X^3)^2$, $\bar{g} = \bar{f} = F$

Montrons que N_g est bornée.
$X + (X^3 - Y^2)^2 < 0$ entraîne $X < -(X^3 - Y^2)^2 \leqslant 0$

D'où $X < -X^6 + 2X^3Y^2 - Y^4 \leqslant 0$

Tous les termes de la somme étant négatifs on en déduit :

(1) $X < -X^6$

(2) $X < -Y^4$

De (1) on déduit X étant négatif :

$1 > -X^5$ et donc $X^5 > -1$.

Nous avons donc $-1 < X^5 < 0$ d'où $-1 < X < 0$.

De (2) on déduit en utilisant le résultat ci-dessus
$-1 < X < -Y^4 \leqslant 0$

Donc $1 > Y^4$ et donc $Y \in]-1, +1[$

Finalement tout point (X, Y) de N_g est tel que
$-1 < X < 0$ et $-1 < Y < +1$. Donc N_g est borné.

La seule hypothèse non vérifiée est donc
$\overline{N_f} \cap V = \emptyset$ puisque $\overline{N_f} \cap V = \overline{N_g} \cap V = \{(0, 0)\}$

. Il est alors impossible de trouver un polynôme $p(X, Y)$ tel que
$g + p(X, Y)(Y^2 - X^3)$ soit positif ou nul sur $\mathbb{R}^2$.

En effet : $g + p(X, Y)(Y^2 - X^3) = X + [p(X, Y) + (Y^2 - X^3)](Y^2 - X^3)$

Or si l'on fait tendre un point vers l'origine sur l'axe OX, quel que soit le polynôme p choisi, l'expression sera équivalente à X et sera donc positive à droite et négative à gauche de l'origine.
Elle ne pourra donc pas rester positive ou nulle sur $\mathbb{R}^2$ tout entier.

Nous avons donc un exemple de fonction $F = \bar{X}$ telle que F est positive sur la variété V d'idéal $\mathcal{I}(V) = (Y^2 - X^3)$ principal. F est bien telle qu'il existe un représentant g tel que $\bar{g} = F$ avec $N_g = \{x \in \mathbb{R}^2,\ g(x) < 0\}$ borné, et du seul fait que $\overline{N_g} \cap V \neq \emptyset$, puisque $\overline{N_g} \cap V = \{(0, 0)\}$, il résulte qu'il est impossible de trouver un représentant de F tel que celui-ci soit positif ou nul sur $\mathbb{R}^2$ tout entier.

Pourtant on peut remarquer que dans ce cas nous avons sur V l'égalité $X\,X^2 = Y^2$ et puisque $X \notin \mathcal{I}(V)$, $\bar{X}(\bar{X})^2 = (\bar{Y})^2$ avec $\bar{X} \neq 0$ dans $\mathbb{R}[V]$.
D'où on déduit

$$\bar{X} = \left(\frac{\bar{Y}}{\bar{X}}\right)^2 \quad \text{dans} \quad \mathbb{R}(V)$$

$\bar{X} = F$ est donc bien un carré dans R(V) celui de $\frac{\bar{Y}}{\bar{X}}$.

Il convient donc de chercher une démonstration autre que celle par application des résultats connus car celle-ci ne pourra donner que des résultats partiels.

CHAPITRE 2 RESOLUTION QUALITATIVE PAR UNE METHODE DE LOGIQUE

Nous obtiendrons ici le théorème suivant :

Théorème 14 :

Soit $V \neq \emptyset$ *une variété algébrique irréductible de* $\mathbb{R}^n$ *et* $\mathcal{J}(V)$ *son idéal. Notons* $\mathbb{R}[V] = \mathbb{R}[X_1, \ldots, X_n] / \mathcal{J}(V)$ *et* $\mathbb{R}(V)$ *le corps des fractions de* $\mathbb{R}[V]$.

Alors pour tout élément $F \in \mathbb{R}(V)$ *tel que* F *est positive ou nulle sur* V *on peut décomposer* F *dans* $\mathbb{R}(V)$ *sous la forme*

$$F = \sum_{1}^{k} G_i^2$$

Ce théorème résulte de la proposition suivante dûe à A. Robinson en 1956 dans [29] :

Proposition

Soit V une variété algébrique irréductible de $\mathbb{R}^n$ et $\mathcal{J}(V)$ l'idéal correspondant de $\mathbb{R}[X_1, \ldots, X_n]$. (V non vide). Soit $f(X_1, \ldots, X_n)$ un polynôme de $\mathbb{R}(X_1, \ldots, X_n)$ tel que $f(x_1, \ldots, x_n) \geqslant 0$ pour tout $(x_1, \ldots, x_n) \in V$.

Alors il existe des polynômes $h_o(X_1, \ldots, X_n), \ldots, h_k(X_1, \ldots, X_n)$ de $\mathbb{R}[X_1, \ldots, X_n]$ tels que

$$(h_o(X_1, \ldots, X_n))^2 f(X_1, \ldots, X_n) \equiv (h_1(X_1, \ldots, X_n))^2 + \ldots + (h_k(X_1, \ldots, X_n))^2 (\mathcal{J}(V))$$

avec $h_o(X_1, \ldots, X_n) \not\equiv 0 \quad (\mathcal{J}(V))$

Cette proposition permet d'obtenir facilement le théorème 14 car d'une telle congruence, on déduit en passant à $\mathbb{R}[V]$:

$$\bar{h}_o^2 \bar{f} = \bar{h}_1^2 + \ldots + \bar{h}_k^2 \qquad \bar{h}_o \neq 0$$

d'où dans $\mathbb{R}(V)$

$$\bar{f} = \left(\frac{\bar{h}_1}{\bar{h}_o}\right)^2 + \ldots + \left(\frac{\bar{h}_k}{\bar{h}_o}\right)^2$$

La résolution du cas f polynôme permet de passer aisément au cas de $f \in \mathbb{R}(V)$ comme nous l'avons déjà fait au chapitre 1.

Ce résultat est en fait obtenu comme corollaire de la proposition suivante démontrée par A. Robinson :

Théorème de Robinson

Soit K *un corps ordonné maximal et* V *une variété algébrique non vide irréductible de* K^n. *Soit* $\mathcal{J}(V)$ *l'idéal de* V *dans* $K[X_1, \ldots, X_n]$.
Soient alors $f(X_1, \ldots, X_n)$, $g_1(X_1, \ldots, X_n)$, $\ldots$, $g_m(X_1, \ldots, X_n)$ $(m \geq 1)$ *un ensemble de polynômes de* $K[X_1, \ldots, X_n]$ *qui n'appartiennent pas à* $\mathcal{J}(V)$ *et qui sont tels que pour tout* $(x_1, \ldots, x_n) \in V$ *le système d'inégalités* $g_i(x_1, \ldots, x_n) > 0$ *pour tout* i *de* 1 *à* n *entraîne* $f(x_1, \ldots, x_n) \geq 0$.
Alors il existe dans $K[X_1, \ldots, X_n]$ *une congruence modulo* $\mathcal{J}(V)$:

$$(h_o(X_1, \ldots, X_n))^2 f(X_1, \ldots, X_n) \equiv \sum_{i=1}^{k} (g_1(X_1, \ldots, X_n))^{\lambda_1^i} \ldots$$

$$\ldots (g_m(X_1, \ldots, X_n))^{\lambda_m^i} (h_i(X_1, \ldots, X_n))^2$$

où λ_j^i *est soit* 0, *soit* 1 *et où les* $h_i(X_1, \ldots, X_n) \in K[X_1, \ldots, X_n]$ *avec de plus* $h_o(X_1, \ldots, X_n) \not\equiv 0$ $(\mathcal{J}(V))$

Ce résultat peut s'exprimer dans K(V), corps des quotients de $\dot{K}[V] = K[X_1, \ldots, X_n] / \mathcal{J}(V)$, par

$$\bar{f} = \sum_{i=1}^{k} \bar{g}_1^{\lambda_1^i} \ldots \bar{g}_m^{\lambda_m^i} G_i^2 \quad \text{où}$$

$\lambda_j^i \in \{0, 1\}$, $G_i \in K(V)$ et $\bar{f}$ et $\bar{g}_j$ représentent les classes de f et des g_j dans $K[V]$.

Pour trouver la proposition énoncée au début il suffit alors de choisir, $K = \mathbb{R}$, $m = 1$, $g_1(X_1, \ldots, X_n) = 1$ qui étant positif strictement sur tout V entraîne bien $f(X_1, \ldots, X_n)$ positif ou nul sur V qui est pris pour hypothèse.
En fait en prenant K ordonné maximal, $m = 1$, $g_1(X_1, \ldots, X_n) = 1$, on peut obtenir cette proposition pour tout corps ordonné maximal et nous avons donc :

Théorème 15 :

Soit K *un corps ordonné maximal ;* $V \neq \emptyset$ *une variété algébrique irréductible de* K^n. *Soit* $\mathfrak{J}(V)$ *son idéal , notons* $K[V] = K[X_1, \ldots, X_n] / \mathfrak{J}(V)$ *et* K(V) *le corps des fractions de* K[V].
Alors pour tout élément $F \in K(V)$ *tel que* F *prend une valeur positive ou nulle sur les éléments de* V *on peut décomposer* F *dans* K(V) *sous la forme :*

$$F = \sum_{1}^{k} G_i^2$$

I - Notations et résultats utilisés pour la démonstration (*)

Définitions d'un ensemble saturé

Soit K un corps. $C \subset K$ sera dite saturée si $0 \in C$, $1 \in C$ et si C contient l'ensemble de tous les produits finis d'éléments de C.

Définition d'un ensemble formellement positif :

Soit K un corps ordonnable ; un sous-ensemble F de K sera dit formellement positif si $0 \in F$ et $1 \in F$, et si il existe un ordre sur K tel que F soit contenu dans l'ensemble des éléments positifs pour cet ordre.

(*) Les notations et résultats utilisés ne sont pas exactement ceux de l'article de Robinson qui fait appel lors de sa démonstration à des théorèmes démontrés dans un de ses précédents articles [30].

Nous avons préféré adapter la démonstration à la terminologie du livre récent de Monsieur Ribenboim sur l'Arithmétique des corps.

Lemme 1

Soit K un corps de caractéristique différente de 2. Soit F un sous-ensemble saturé de K (donc $0 \in F$ et $1 \in F$).

Les conditions suivantes sont équivalentes :

(1) K est ordonnable et F est formellement positif.

(2) $K \neq \Sigma\, F\, K^2$ (ensemble des sommes finies d'éléments a x^2 où $a \in F$ et $x \in K$).

(3) Si $\sum_{i=1}^{k} a_i\, x_i^2 = 0$ avec $a_i \in F$, $a_i \neq 0$, $x_i \in K$ et $n \geqslant 1$ alors

$$x_1 = x_2 = \ldots = x_k = 0$$

(4) Il existe $S \subset K$ tel que $K^2 \subset S$, $F \subset S$, $S + S \subset S$, $S.S \subset S$, $S \cap (-S) = \{0\}$

Définition d'un élément F-positif

F une partie formellement positive d'un corps ordonnable K. $a \in K$ sera dit F-positif si a est positif pour tout ordre sur K tel que F soit contenu dans l'ensemble des éléments positifs pour cet ordre.

Lemme 2

Si F est une partie saturée formellement positive du corps ordonnable K, a est F-positif si et seulement si $a \in \Sigma\, F\, K^2$.

Lemme 3

Soit $\mathcal{A}$ l'ensemble d'axiomes caractérisant les corps ordonnés maximaux (*), écrites dans le langage égalitaire $\mathcal{L}$ formé de 0 et 1 comme symboles de constantes, - symbole fonctionnel à 1 variable, + et $\times$ symboles fonctionnels à 2 variables et $>$ 0 symbole fonctionnel à 1 variable.
Alors $\mathcal{A}$ permet l'élimination des quantificateurs dans $\mathcal{L}$.

(*) L'ensemble des formules $\mathcal{A}$ a déjà été écrit au chapitre 2 de la première partie.

Lemme 4

Si $\mathcal{A}$ permet l'élimination des quantificateurs dans $\mathcal{L}$ et si $\mathcal{D}_{\mathcal{M}}$ est le diagramme d'un modèle $\mathcal{M}$ de $\mathcal{A}$, alors ($\mathcal{A}$, $\mathcal{D}_{\mathcal{M}}$) est saturé pour le langage $\mathcal{L}'$ obtenu en ajoutant aux symboles de constantes du langage $\mathcal{L}$ les éléments de l'ensemble de base $\mathcal{M}$.

Rappelons qu'un ensemble de formule F est dit saturé pour un langage $\mathcal{L}$ si pour toute formule close F de $\mathcal{L}$, F ou $\neg$ F est conséquence de $\mathcal{A}$.

Corollaire des lemmes 3 et 4

K est un modèle de $\mathcal{A}$; soit $\mathcal{D}_K$ son diagramme alors toute formule close vraie dans K est vraie dans tout modèle de ($\mathcal{A}$, $\mathcal{D}_K$) donc dans tout corps ordonné maximal contenant K et extension ordonnée de K.
Pour la démonstration des lemmes 3 et 4, déjà utilisés dans la première partie se reporter à [11].

II - Démonstration du théorème de Robinson

Soient $p_1(X_1, \ldots, X_n), \ldots, p_r(X_1, \ldots, X_n)$ une base de l'idéal $\mathcal{J}(V)$.
Notons que K(V) corps des quotients de $K[V] = \frac{K[X_1, \ldots, X_n]}{\mathcal{J}(V)}$ est une extension de K; puisque K est ordonné maximal, K est de caractéristique nulle, et donc K(V) est aussi de caractéristique nulle.

Soit C l'ensemble formé des produits dans K(V): $a\, \bar{g}_1^{\lambda_1} \ldots \bar{g}_m^{\lambda_m}$

où $\lambda_j \in \mathbb{N}$ et où $a \in K$ et $a \geqslant 0$.

L'ensemble C est un sous-ensemble de K(V) qui est saturé ($0 \in C$ et $1 \in C$ et les produits finis d'éléments de C appartiennent à C). K(V) est de caractéristique différente de 2. Nous pouvons donc appliquer les lemmes 1 et 2.

1er cas ======= S'il est faux que K(V) soit ordonnable et C formellement positif alors par le lemme 1 il est aussi faux que K(V) soit distinct de $\Sigma\ C\ K(V)^2$ donc $K(V) = \Sigma\ C(K(V))^2$. Ceci signifie que tout élément F de K(V) peut s'écrire :

$$F = \sum_{i=1}^{k} a_i\ \bar{g}_1^{\lambda_1^i} \ldots \bar{g}_m^{\lambda_m^i}\ H_i^2 \quad \text{où } \lambda_j^i \in \mathbb{N},\ a_i \in K,\ a_i \geqslant 0 \text{ et}$$

$H_i \in K(V)$.

Puisque $a_i \in K$ et $a_i \geqslant 0$ alors $a_i = (\sqrt{a_i})^2$ dans K. D'autre part on peut absorber les puissances paires des $\bar{g}_j^{\lambda_j^i}$ dans H_i^2 et donc on aura dans K(V)

$$F = \sum_{i=1}^{k} \bar{g}_1^{\lambda_1^i} \ldots \bar{g}_m^{\lambda_m^i}\ G_i^2 \qquad \lambda_j^i \in \{0, 1\}.$$

2è cas ====== Si K(V) est ordonnable et C formellement positif dans K(V) alors un élément F de K(V) pourra s'écrire comme un élément de $\Sigma\ C\ K(V)^2$ si et seulement si F est C-positif dans K(V) (Lemme 2).

Par hypothèse nous avons une fonction $f \in K[X_1, \ldots, X_n]$ telle que pour tout $(x_1, \ldots, x_n) \in V$ $\quad g_1(x_1, \ldots, x_n) > 0, \ldots g_m(x_1, \ldots, x_n) > 0$ entraîne $f(x_1, \ldots, x_n) \geqslant 0$.

$(x_1, \ldots, x_n) \in V$ c'est $p_1(x_1, \ldots, x_n) = 0, \ldots, p_r(x_1, \ldots, x_n) = 0$.

Donc nous avons la formule X :

$\wedge x_1 \ldots \wedge x_n (p_1(x_1, \ldots, x_n) = 0 \wedge \ldots \wedge p_r(x_1, \ldots, x_n) = 0 \wedge g_1(x_1, \ldots, x_n) > 0 \wedge \ldots \wedge g_m(x_1, \ldots, x_n) > 0 \rightarrow f(x_1, \ldots, x_n) \geqslant 0)$ qui est vraie dans K.

Mais d'après le corollaire des lemmes 3 et 4 cette formule X est donc aussi vraie dans tout modèle de $(\mathcal{A}, \mathcal{D}_K)$ donc dans tout corps ordonné maximal extension ordonnée de K.

Appelons K' la clôture réelle de K(V). L'unique ordre de K' prolonge celui de K(V) qui prolongeait l'unique ordre de K. Donc K' est un corps ordonné maximal extension ordonnée de K. La formule X est donc vraie dans K' et donc aussi dans K(V) qui est contenu dans K' et ceci quel que soit l'ordre choisi sur K(V).
Choisissons dans K(V) les éléments $\bar{X}_1, \ldots, \bar{X}_n$
Alors $p_i(\bar{X}_1, \ldots, \bar{X}_n) = \bar{p}_i(X_1, \ldots, X_n) = 0$

De plus C étant formellement positif dans K(V), on peut considérer les seuls ordres qui sont tels que les éléments de C soient positifs. Les $\bar{g}_i = g_i(\bar{X}_1, \ldots, \bar{X}_n)$ sont donc positifs dans K(V) et $\bar{g}_i \neq 0$ car $g_i \notin \mathcal{J}(V)$.

Donc sur l'élément $(\bar{X}_1, \ldots, \bar{X}_n)$ la première partie de la formule est vérifiée pour tous les ordres tels que C soit contenu dans l'ensemble des éléments positifs.

Donc pour tous ces ordres $f(\bar{X}_1, \ldots, \bar{X}_n) = \bar{f}$ est positif ou nul.
Ceci signifie que l'élément $F = \bar{f}$ est C-positif dans K(V).

On en déduit (Lemme 2)

$F \in \Sigma\, C(K(V))^2$ et donc

$$F = \sum_{i=1}^{k} a_i \bar{g}_1^{\lambda_1^i} \ldots \bar{g}_m^{\lambda_m^i} H_i^2 \qquad a_i \in K \qquad a_i \geqslant 0$$

$$\lambda_j^i \in \mathbb{N} \qquad H_i \in K(V)$$

et donc par un raisonnement déjà fait

$$F = \sum_{i=1}^{k} \bar{g}_1^{\lambda_1^i} \ldots \bar{g}_m^{\lambda_m^i} G_i^2 \qquad \lambda_j^i \in \{0, 1\}, \text{ égalité dans } K(V)$$

. Dans les deux cas nous obtenons le résultat annoncé dans ce théorème.

III - Démonstration des résultats utilisés

1 . Démonstration du lemme 1

1 ==> 2 Soit un ordre sur K tel que $F \subset P$ ensemble des éléments positifs pour cet ordre. Puisqu'on a toujours $K^2 \subset P$, que $P + P \subset P$ et que $P \cdot P \subset P$ on en déduit $\Sigma\, F\, K^2 \subset P \neq K$.

2 ==> 3 Soit $\sum_{i=1}^{n} a_i x_i^2 = 0$; supposons $x_1 \neq 0$ par exemple,

alors $-a_1 = \sum_{i=2}^{n} a_i \left(\frac{x_i}{x_1}\right)^2$ $\quad -a_1 \in (-F)$ puisque $a_1 \in F$

et par cette égalité $-a_1 \in \Sigma\, F\, K^2$.

Montrons qu'alors $a_1 = 0$ ce qui est contraire à l'hypothèse de 3.

Pour cela montrons que $(-F) \cap (\Sigma\, F\, K^2) = \{0\}$

Supposons $a \in (-F) \cap (\Sigma\, F\, K^2)$ et $a \neq 0$

Soit y arbitraire dans K et $x = \frac{y - a}{2a}$ (x est défini puisque car $K \neq 2$). Alors $y = 2ax + a = a(x + 1)^2 - ax^2$

$a \in \Sigma\, F\, K^2$, $(x + 1)^2 \in K^2$ donc $a(x + 1)^2 \in \Sigma\, F\, K^2$;
$-a \in F$ et $x^2 \in K^2$ donc $-a x^2 \in F\, K^2$. D'où $y \in \Sigma\, F\, K^2$.
Ceci étant vrai pour tout $y \in K$ entraîne $K = \Sigma\, F\, K^2$ ce qui est contraire au résultat 2. Donc $(-F) \cap (\Sigma\, F\, K^2) = \{0\}$ ce qui entraîne $a_1 = 0$ contrairement à l'hypothèse. Donc il n'existe pas de x_i non nul.

3 ==> 4 Soit $S = \Sigma\, F\, K^2$. S possède les propriétés suivantes :
$K^2 \subset S$, $F \subset S$, $S + S \subset S$ (évident); de plus $S \cdot S \subset S$ car F est saturé par hypothèse. Enfin $S \cap (-S) = \{0\}$. En effet, supposons $\sum_{i=1}^{n} a_i x_i^2 = -\sum_{j=1}^{m} b_j y_j^2$ où a_i et b_j appartiennent à F et sont non nuls et où x_i et y_j appartiennent à K.

Alors $\sum_{i=1}^{n} a_i x_i^2 + \sum_{j=1}^{m} b_j y_j^2 = 0$ d'où d'après 3 : $x_i = 0$, $y_j = 0$

quels que soient i et j.

4 ==> 1 Considérons alors la famille S des sous-ensembles Q de K qui vérifient :

$S \subset Q$, $Q + Q \subset Q$, $Q \cdot Q \subset Q$, $Q \cap (-Q) = \{0\}$

S est non vide car elle contient S et S est une famille inductive. D'après le lemme de Zorn S possède un élément maximal que nous noterons Q.

Si nous montrons que $Q \cup (-Q) = K$ alors Q sera l'ensemble des éléments positifs pour un ordre sur K. Et puisque $F \subset S \subset Q$, F sera formellement positif.

Montrons que $Q \cup (-Q) = K$. Soit $x \notin Q$ nous allons montrer que $-x \in Q$.

Posons $Q' = Q - xQ = \{c - xd, c \in Q \text{ et } d \in Q\}$ $Q \subset Q'$, $Q' + Q' \subset Q'$, $Q'Q' \subset Q'$ (puisque K^2 contenu dans S est contenu dans Q). De plus $Q' \cap (-Q') = \{0\}$. En effet, supposons l'existence de deux éléments tels que $c - xd = -(c' - xd')$ où c, c', d, d' appartiennent à Q. On en déduit

$$c + c' = x(d + d')$$

a) Si $d + d' = 0$ alors $c + c' = 0$ donc $c = -c'$.
Or c et c' appartiennent à Q qui vérifie $Q \cap (-Q) = \{0\}$ donc $c = c' = 0$.
De même $d = d' = 0$.

b) Si $d + d' \neq 0$ alors

$$x = \frac{c + c'}{d + d'} = \frac{(c + c')(d + d')}{(d + d')^2}$$

$(c + c')(d + d') \in Q$ et $\frac{1}{(d + d')^2} \in K^2$ donc $x \in K^2$ $Q \subset Q$

ce qui est contraire à l'hypothèse;
Donc $Q' \cap (-Q') = \{0\}$.

Q' a toutes les propriétés des éléments de la famille $\mathcal{S}$. Donc $\mathcal{S}$ contient Q'. Mais Q' $\supset$ Q et Q était maximal dans $\mathcal{S}$. On en déduit Q' = Q, donc Q = Q - x Q, et en particulier - x $\in$ Q.
On a donc bien montré que x $\notin$ Q entraine - x $\in$ Q donc que Q $\cup$ (- Q) = K.

2 . Démonstration du lemme 2

Supposons a $\in \Sigma$ F K^2 et un ordre sur K dont l'ensemble P des éléments positifs contient F. Alors F $\subset$ P et puisque $K^2 \subset$ P et P + P $\subset$ P nous avons Σ F $K^2 \subset$ P. Donc a $\in \Sigma$ F K^2 entraîne a $\in$ P , et a est bien F-positif.

Réciproquement :

Soit a un élément n'appartenant pas à Σ F K^2, alors a $\notin$ F.
Posons G = F $\cup \{-a\}$, alors $\tilde{G}$ saturé de G est l'ensemble :
$\tilde{G}$ = F $\cup \{-a\ F\} \cup \{a^2\ F\} \cup \{-a^3\ F\} \cup \ldots$

. Soit alors :
$S = \Sigma\ \tilde{G}\ K^2 = \Sigma$ F K^2 - a Σ F K^2.
Nous avons de façon évidente :
$\tilde{G} \subset S$, $K^2 \subset S$, S + S $\subset$ S et S . S $\subset$ S

. De plus S $\cap$ (- S) = {0}. En effet, suposons z $\in$ S $\cap$ (- S) alors z = c - a d = - (c' - a d') avec c, c', d et d' appartenant à ΣF K^2. On en déduit c + c' = a(d + d')

a) Si d + d' = 0 alors c + c' = 0. F étant formellement positif il existe un ordre sur K tel que Σ F $K^2 \subset$ P ensemble des éléments positifs pour cet ordre.
Or P $\cap$ (- P) = {0} ; donc c = - c' avec c et c' appartiennent à P entraîne c = c' = 0. De même nous avons d = d' = 0.

b) Si d + d' $\neq$ 0 alors $a = \dfrac{(c + c')}{(d + d')} = \dfrac{(c + c')(d + d')}{(d + d')^2}$

B I B L I O G R A P H I E

[1] D. HILBERT — Uber die Darstellung definiter Formen als Summe von Formenquadraten.
Math. Ann. 32 (1888) p. 342-350.

[2] D. HILBERT — Uber ternäre definite Formen.
Acta Math. 17 (1893) p. 169-197.

[3] D. HILBERT — Mathematische probleme.
Göttinger Nach. (1900) p. 284-285.

[4] D. HILBERT — Grundlagen der Geometrie.
Leipzig 1899, Kap VII, insbesondere § 38.

[5] E. LANDAU — Uber die Darstellung definiter binärer Formen durch Quadrate.
Math. Ann. Bd 57 (1907) p. 53-64.

[6] E. ARTIN — Uber die Zerlegung definiter Funktionen in Quadrate.
Abh. Math. Sem. Hamburg 5 (1927) p. 100-115.

[7] N. JACOBSON — Abstract Algebra.
Van Nostrand.

[8] P. RIBENBOIM — L'Arithmétique des corps.
Hermann. Paris 1972.

[9] N. BOURBAKI — Livre II, ch. 6. Groupes et corps ordonnés.
Hermann. Paris.

[10] A. ROBINSON — Model Theory.

[11] G. KREISEL et J.L. KRIVINE — Eléments de Logique Mathématique. Théorie des modèles. Dunod. Paris 1967.

[12] S. LANG — Algebra. Addison-Wesley.

[13] D.W. DUBOIS — Note on Artin's solution of Hilbert's 17^{th} problem. Bull. of Amer. Math. Soc. July 1967 vol. 73 nº 4 p. 540-541.

[14] L. FUCHS — Abelian groups. Akad. Kiadó Budapest 1958 § 42.

[15] G. BACHMAN — Introduction to p-adic numbers and valuation theory. Academic press - New-York 1964 p. 122.

[16] VAN DER WAERDEN — Algebra. Ungar - New-York.

[17] P. RIBENBOIM — Le Théorème des zéros pour les corps ordonnés. Séminaire d'Algèbre et Théorie des Nombres. Dubreil-Pisot 24è année 1970-1971 Exposé 17.

[18] T.S. MOTZKIN — The arithmetic geometric inequality (6, 2) Inequalities O-Shisha. Academic press, New-York 1967.

[19] R.M. ROBINSON — Some definite polynomials which are not sums of squares of real polynomials. Amer. Math. Soc. (Abstract) 1969.

[20] E. LANDAU — Uber die Darstellung definiter Funktionen durch Quadrate. Math. Ann. 62 (1906) p; 272-285

[21] J.W.S. CASSELS On the representation of rational functions as sums of squares.
Acta Arithmetica IX (1964) p. 79-82.

[22] H. DAVENPORT A problematic identity.
Mathematika 10 (1963) p. 10-12.

[23] A. PFISTER Zur Darstellung definiter Funktionen als Summe von Quadraten.
Invent. Math. 4, (1967) p. 229-237.

[24] A. PFISTER Multiplikative quadratische Formen.
Arch. Math. 16 (1965) p. 363-370.

[25] S. LANG On quasi algebraic closure.
Ann. of Math. 55 (1952) p. 373-390.

[26] J. AX On ternary definite rational functions.
(non publié).

[27] J.W.S. CASSELS, W.J. ELLISON, A. PFISTER
On sums of squares and on elliptic curves over function fields.
Journal of number theory (Mai 1971) vol. 3 n° 2

[28] Y. POURCHET Sur la représentation en sommes de carrés des polynômes à une indéterminée sur un corps de nombres algébriques.
Acta Arithmetica, Wars. Zawa, t. 19 p. 89-104 (1971)

[29] A. ROBINSON Further remarks on ordered fields and definite functions.
Math. Ann. Bd 130 (1956) p. 405-409.

[30] A. ROBINSON On ordered fields an definite functions.
Math. Ann. Bd 130 (1955) p. 257-271.