Université de Nancy I
Centre de Recherche en Informatique
de Nancy

ALGORITHMES DE FACTORISATION DES POLYNOMES A COEFFICIENTS ENTIERS

thèse de Doctorat de l'Université de Nancy I
soutenue en 15 Septembre 1989
par

Guy VIRY
B.P. 10310 **NIAMEY (Niger)** ou
CRIN B.P. 239 54306 **Vandoeuvre-les-Nancy (France)**

Composition du Jury :

Président	**Daniel**	**Lazard**
Rapporteurs	**Maurice**	**Mignotte**
	Pierre	**Marchand**
Examinateurs	**Claude**	**Pair**
	René	**Schott**

Université de Nancy I
Centre de Recherche en Informatique
de Nancy

ALGORITHMES DE FACTORISATION DES POLYNOMES A COEFFICIENTS ENTIERS

thèse de Doctorat de l'Université de Nancy I
soutenue en Septembre 1989
par

Guy VIRY
B.P. 10310 **NIAMEY (Niger)** ou
CRIN B.P. 239 54306 **Vandoeuvre-les-Nancy (France)**

Composition du Jury :

Président	**Daniel**	**Lazard**
Rapporteurs	**Maurice**	**Mignotte**
	Pierre	**Marchand**
Examinateurs	**Claude**	**Pair**
	René	**Schott**

Université de Nancy I
Centre de Recherche en Informatique
de Nancy

ALGORITHMES DE FACTORISATION DES POLYNOMES A COEFFICIENTS ENTIERS

thèse de Doctorat de l ' Université de Nancy I
soutenue en Septembre 1989
par

Guy VIRY
B.P. 10310 **NIAMEY (Niger)** ou
CRIN B.P. 239 54306 **Vandoeuvre-les-Nancy (France)**

Composition du Jury :

Président	**Daniel**	**Lazard**
Rapporteurs	**Maurice**	**Mignotte**
	Pierre	**Marchand**
Examinateurs	**Claude**	**Pair**
	René	**Schott**

Remerciements

Je profite de la rédaction de cette thèse pour remercier les différentes personnes qui m'ont aidé dans mon travail de recherche.

La première personne qui m'a fait débuter dans ce travail de recherche est le Professeur Claude Pair. Je le remercie d'avoir lu et corrigé mes premiers projets d'articles, de m'avoir conseillé et de m'avoir aidé à trouver une nouvelle orientation en me faisant connaître Maurice Mignotte. J'ai été content d'apprendre qu'il acceptait de faire partie de mon jury de thèse.

Je tiens à remercier particulièrement le Professeur Maurice Mignotte pour ses nombreux conseils et pour tout l'intérêt qu'il m'a accordé. Chacun des articles qui forment les parties essentielles de cette thèse a été corrigé et amélioré grâce à ses conseils. Ses remarques pour la rédaction de cette thèse ont également été très précieuses.

Je remercie également le Professeur René Schott qui m'a acceuilli dans son équipe en 1986. Malheureusement, j'ai quitté Nancy un an plus tard et je n'ai pas pu continuer à travailler avec lui comme je l'aurais souhaité. Je lui suis reconnaissant d'avoir néanmoins accepté d'être mon directeur de recherche.

Je remercie enfin le Professeur Daniel Lazard d'avoir accepté de présider le Jury de cette thèse et le Professeur Pierre Marchand d'avoir accepté d'être l'un des rapporteurs.

Résumé de la thèse.

Cette thèse présente les algorithmes classiques de factorisation des polynômes à une et plusieurs variables. Dans le cas des polynômes à une variable, deux nouvelles méthodes sont proposées. Dans la première, on calcule un facteur linéaire modulo un nombre premier p, puis on définit unh multiple de ce facteur qui divise P sur Z[X]. Dans la seconde méthode, le calcul des produits des facteurs de P modulo p^n est remplacé par le calcul des sommes des images de ces facteurs de P.

Dans le cas des polynômes à plusieurs variables, on donne deux méthodes pour diminuer les calculs de la dernière étape de la factorisation avec les algorithmes classiques. On utilise la représentation polyédrale et la notion de polynômes "normalisés".

La dernière partie de la thèse donne une méthode pour diminuer le degré total du polynôme donné, lorsque c'est possible. Cette méthode utilise la programmation linéaire entière.

Mots-Clés.

Polynômes à une variable, Polynômes à plusieurs variables, Factorisation, Méthode de Berlekamp, Algorithme de Lenstra, Polyèdre, Programmation linéaire entière.

Présentation

Introduction.

La factorisation des polynômes est l'un des problèmes importants que doivent résoudre les grands systèmes de calcul formel, comme Macsyma ou Reduce. C'est un problème difficile qui n'a été résolu de façon satisfaisante que récemment, grâce à l'algorithme de Berlekamp (1967). Le cas le plus complexe est celui des polynômes à plusieurs variables, notamment lorsque le nombre de variables devient grand. Le calcul formel en général et la factorisation en particulier, trouvent de nombreuses applications en robotique. Quelques exemples en sont donnés dans le livre récent (1987), intitulé "Calcul Formel", de J. Davenport, Y. Siret et E. Tournier, pages 126–127. En effet, les trajectoires et les contraintes du robot sont définies par des systèmes d'équations polynômiales. L'un des problèmes est traité par l'équipe du Professeur René Schott; la question posée est, d'une part de savoir si on peut ou non déplacer un objet dans un environnement donné et d'autre part, si la réponse est positive, de trouver le déplacement le plus rapide. Les réponses à ces questions nécessitent la résolution de systèmes semi–algèbriques, qui utilise les bases de Gröbner et les techniques de factorisation des polynômes à plusieurs variables. L'étude même de ces problèmes serait presque impossible sans le secours des grands systèmes comme Macsyma ou Reduce. Les fondements mathématiques du calcul formel sont empruntés à l'algèbre et à l'arithmétique; une présentation de ces mathématiques a fait l'objet d'un livre récent (1989) de M. Mignotte, intitulé "Mathématiques pour le Calcul Formel".

Plan de la thèse.

La thèse comprend trois parties. La partie A traite le cas des polynômes à une variable, qui est le plus simple. On commence par présenter en détail les méthodes classiques, à savoir l'algorithme de Berlekamp, suivi de l'algorithme de raffinement de la factorisation basé sur le lemme de Hensel et enfin la méthode de Lenstra dont le coût est polynômial. Cette partie A traite ensuite deux méthodes originales. La première consiste à rechercher une racine simple α du polynôme donné P, puis à utiliser la méthode de Lenstra pour trouver le facteur de P admettant α comme racine. Cette méthode est exposée dans **[V4]**. La deuxième méthode proposée a pour but de diminuer le coût

théorique de la dernière phase de la méthode classique de factorisation. Dans le pire cas, ce coût est exponentiel suivant le degré du polynôme. La nouvelle méthode consiste à transformer P et ses facteurs à l'aide d'un monomorphisme; cette transformation nous donne un critère pour choisir les facteurs de P, et diminue ainsi le coût dans le pire cas. Cette deuxième méthode est exposée dans **[V5]** et **[V6]**.

La partie B traite le cas des polynômes à m+1 variables, à coefficients entiers. On présente d'abord la méthode classique avec notamment l'algorithme amélioré de Wang. Ensuite trois méthodes originales sont données. La première est fondée sur la décomposition des entiers dans une base symétrisée; cette décomposition permet de "supprimer" une variable, et ainsi de se ramener à la factorisation d'un polynôme à une variable. Cette méthode est exposée dans **[V1]**.

La deuxième méthode de B utilise un changement de variable transformant le polynôme donné ainsi que ses diviseurs en polynômes "normalisés" ayant des propriétés caractéristiques. Cette notion de polynôme "normalisé" est fondée sur une représentation géométrique du polynôme sous forme d'un polyèdre. Elle permet de donner un critère pour choisir les facteurs de P; elle diminue ainsi le coût dans le pire cas. Cette méthode est exposée dans **[V2]**.

La troisième méthode proposée dans B consiste à généraliser aux polynômes à plusieurs variables, le monomorphisme défini dans la partie A. On obtient comme dans la méthode précédente un critère pour choisir les facteurs de P.

La partie C présente une méthode originale pour "simplifier" un polynôme à plusieurs variables. Cette "simplification" peut être définie suivant plusieurs sens: diminution du degré suivant l'une des variables, diminution du nombre des variables dans certains cas favorables, transformation du polynôme en polynôme unitaire, ... Ces "simplifications" sont fondées sur la représentation polyèdrale des polynômes et elles sont obtenues par la résolution d'un programme linéaire en nombres entiers.

Les algorithmes de factorisation de la partie A pour les polynômes à une variable et ceux de la partie B pour les polynômes à plusieurs variables ont une approche semblable. On présentera donc le processus général de factorisation en traitant simultanément les deux cas. Pour cela, on considère un polynôme P de $\mathbf{Z}[X,X_1,..,X_m]$ comme un polynôme en X à coefficients dans l'anneau **K**, où **K** est soit l'anneau **Z**, soit l'anneau $\mathbf{Z}[X_1,...,X_m]$.

Table des matières.

Partie B : Polynômes à plusieurs variables.

INTRODUCTION GENERALE

Les problèmes traités concernent les polynômes à une ou plusieurs variables à coefficients entiers. L'objectif est la factorisation de ces polynômes.

L'idée essentielle des nouvelles méthodes proposées est de donner des outils, permettant de transformer le problème en le simplifiant. Parmi ces outils, il y a la représentation des polynômes à plusieurs variables par des polyèdres, déjà utilisée par Ostrowski dans **[Os]**, la programmation linéaire en nombres entiers, la décomposition des fractions rationnelles en éléments simples, le développement logarithmique, qui définit un monomorphisme transformant les produits en sommes.

Les algorithmes classiques de factorisation ont comme origine l'algorithme de Berlekamp **[Be]**, puis l'adaptation du lemme de Hensel à la factorisation des polynômes à coefficients entiers par Zassenhaus **[Za]**. Ensuite l'algorithme a été adapté au cas des polynômes à plusieurs variables par Wang et Rothschield dans **[WR]**, et par Musser dans **[Mu]**. Enfin une amélioration a été donnée par Wang dans **[Wa]**, pour le cas des polynômes à plusieurs variables.

Ces algorithmes ne sont valables que si le polynôme donné P n'a pas de facteur carré; on étudiera plus loin comment on se ramène à ce cas. On suppose donc dans la suite que P n'a pas de facteur carré, et on le note sous la forme

$$P = a_o X^d + a_1 X^{d-1} + \ldots + a_{d-1} X + a_d.$$

On adopte dans la suite les notations suivantes pour les logarithmes: Log désigne le logarithme dans la base 2, log le logarithme népérien et Log_b le logarithme dans la base b.

Méthode classique de factorisation.

Le processus général de factorisation comporte trois étapes et ce processus est le même dans le cas de polynômes à une variable et dans le cas de polynômes à plusieurs variables.

On se propose de présenter l'algorithme sur un exemple, de façon à mettre en évidence les étapes de l'algorithme. On traite en parallèle un polynôme de $\mathbf{Z}[X]$ et un polynôme de $\mathbf{Z}[X,Y]$, pour montrer la similitude des calculs dans le cas d'une variable et de plusieurs variables.

Exemple.

Considérons les deux polynômes suivants $P = X^6 + 2X^4 + 4X^2 + 3$ et $Q = X^5 - 2X^4 - X + 2 + Y(2X^3 - 3X^2 + 2X) - Z(2X^4 + 2X^3 + X^2 + 5X - 2) - 2XZ^2 - 2X^2YZ$.

Etape 1.

On projette respectivement P sur $\mathbf{Z}/(5)[X]$ et Q sur $\mathbf{Z}[X,0,0]$. Notons $P_o = X^6 + 2X^4 - X^2 - 2$ et $Q_o = X^5 - 2X^4 + X + 2$, ces projections. Le polynôme P_o est factorisé sur $\mathbf{Z}/(5)[X]$ grâce à l'algorithme de Berlekamp. On trouve $P_o = (X+1)(X-1)(X+2)(X-2)(X^2 + 1)$. Le polynôme Q_o est factorisé sur $\mathbf{Z}[X,0,0]$ grâce à l'algorithme de factorisation des polynômes à une variable. On trouve $Q_o = (X + 1)(X - 1)(X - 2)(X^2 + 1)$.

Etape 2.

On note P_1, P_2,... les projections de P sur $\mathbf{Z}/(5^2)[X]$, $\mathbf{Z}/(5^4)[X]$,... On peut factoriser P_{i+1} à partir de la factorisation de P_i grâce au Lemme de Hensel. On obtient ainsi $P_1 = (X - 4)(X + 4)(X + 7)(X - 7)(X^2 - 8)$, puis $P_2 = (X - 29)(X + 29)(X + 182)(X - 182)(X^2 + 217)$. On arrête ce raffinement de la factorisation de P à P_2, car ses coefficients sont écrits dans l'intervalle $]-½5^4, ½5^4]$. Si les coefficients des diviseurs de P sont compris entre $-½5^4$ et $½5^4$, alors la projection de $\mathbf{Z}[X]$ sur $\mathbf{Z}/(5^4)[X]$ laisse invariants les diviseurs de P. Par suite, à l'étape 3, la factorisation de P_2 sur $\mathbf{Z}/(5^4)[X]$ permettra de trouver celle de P sur $\mathbf{Z}[X]$.

Dans le cas de plusieurs variables, on note Q_1, Q_2,... les projections de Q sur $\mathbf{Z}[Y,Z]/\Delta_2\,[X]$, $\mathbf{Z}[Y,Z]/\Delta_3\,[X]$, ... où Δ_2, Δ_3, ... sont les idéaux de $\mathbf{Z}[Y,Z]$ engendrés par les monômes de degré total 2, 3, ... La factorisation initiale de Q_o sur $\mathbf{Z}[X,0,0] = \mathbf{Z}[Y,Z]/\Delta_1\,[X]$ a été définie à l'étape 1. On peut factoriser Q_{i+1} à partir de Q_i grâce à une généralisation du Lemme de Hensel. On obtient successivement

$Q_1 = (X + 1 + \frac{1}{2}Z - \frac{1}{2}Y)\ (X - 1 + 2Z)\ (X - 2 - 4Z)(X^2 + 1 - \frac{1}{2}(X-1)Z + \frac{1}{2}(X+1)Y)$,

$Q_2 = (X + 1 + \frac{1}{2}Z - \frac{1}{2}Y - \frac{1}{4}Z^2 + \frac{1}{4}YZ)\ (X - 1 + 2Z - 8Y^2)\ (X - 2 - 4Z + 8Y^2)$

$(X^2 + 1 - \frac{1}{2}(X-1)Z + \frac{1}{2}(X+1)Y + \frac{1}{4}Z^2X + \frac{1}{4}Y^2 - \frac{1}{4}(X+1)YZ)$.

On s'arrête à la factorisation de Q sur $\mathbf{Z}[Y,Z]/\Delta_3\ [X]$, car si Q est réductible, ses diviseurs ont un degré en Y et Z majoré par celui de Q, à savoir 2. La projection de $\mathbf{Z}[X,Y,Z]$ sur $\mathbf{Z}[Y,Z]/\Delta_3\ [X]$ laisse donc invariants Q et ses diviseurs. A l'étape 3, la factorisation de Q_2 permettra de trouver celle de Q sur $\mathbf{Z}[X,Y,Z]$.

Etape 3.

Pour trouver les diviseurs de P sur $\mathbf{Z}[X]$, on calcule tous les produits possibles des facteurs de P_2 sur $\mathbf{Z}/(5^4)[X]$, c'est-à-dire tous les sous-produits de $(X - 29)(X + 29)(X + 182)(X - 182)(X^2 + 217)$. Si P est réductible sur $\mathbf{Z}[X]$, alors chacun de ses diviseurs s'identifie à l'un de ces sous-produits sur $\mathbf{Z}/(5^4)[X]$. On trouve ainsi que $(X - 182)(X + 182) = X^2 + 1$ modulo 5^4, et on vérifie que $X^2 + 1$ est bien un diviseur de P sur $\mathbf{Z}[X]$.

Pour trouver les diviseurs de Q sur $\mathbf{Z}[X,Y,Z]$, on calcule tous les produits obtenus à partir des facteurs de Q_2 sur $\mathbf{Z}[Y,Z]/\Delta_3\ [X]$. On trouve ainsi

$(X + 1 + \frac{1}{2}Z - \frac{1}{2}Y - \frac{1}{4}Z^2 + \frac{1}{4}YZ)$

$(X^2 + 1 - \frac{1}{2}(X-1)Z + \frac{1}{2}(X+1)Y + \frac{1}{4}Z^2X + \frac{1}{4}Y^2 - \frac{1}{4}(X+1)YZ)$

$= X^3 + X^2 + X + 1 + Z + XY$ modulo Δ_3,

$(X - 1 + 2Z - 8Y^2)(X - 2 - 4Z + 8Y^2) = X^2 - 3X + 2 - 2XZ$ modulo Δ_3. On obtient ainsi les diviseurs de Q sur $\mathbf{Z}[X,Y,Z]$.

On a supposé pour simplifier l'exemple que les polynômes P et Q sont unitaires. Supposons maintenant que ce ne soit pas le cas. Prenons par exemple

$P = 64\ X^6 + 32\ X^4 + 16\ X^2 + 3$ et

$Q = (1-2Z)X^5 - 2X^4 - X + 2 + (2X^3 - 3X^2 + 2X)Y - (2X^4 + X^3 + 5X^2 - 2X)Z - 2X^3Z^2 - 2X^3YZ$.

On écrit P sur $\mathbf{Z}/(5^4)[X]$ sous la forme

$P = 64(X^6 - 312\ X^4 - 156\ X^2 + 127) = 64\ P^*$. On factorise P^* comme on l'a fait plus haut sur $\mathbf{Z}/(5^4)[X]$; on trouve

$P^* = (X + 298)(X - 298)(X + 91)(X - 91)(X^2 - 102)$ modulo 5^4.

A l'étape 3 on calcule tous les produits possibles des facteurs ci-dessus de P^*, en leur adjoignant le facteur 64. Puis on vérifie si certains de ces produits divisent $64\ P$ sur

$Z[X]$. On obtient ainsi $64(X + 91)(X - 91) = 64\,X^2 + 16$ modulo 5^4; on peut vérifier que $64\,X^2 + 16$ divise $64\,P$. En rendant primitif $64\,X^2 + 16$ on obtient $4\,X^2 + 1$ qui est le diviseur correspondant de P.

Dans le cas de plusieurs variables, le polynôme Q donné ci-dessus peut s'écrire sous la forme suivante sur $Z[Y,Z]/\Delta_3\,[X]$ (en remarquant que l'inverse de $1 - 2\,Z$ est égal à $1 + 2Z + 4Z^2$),

$Q = (1-2Z)(\,X^5 - 2X^4 - X + 2 + 2XY + 4Z + 4XYZ + 8Z^2\,) = (1-2Z)\,Q^*$.

Comme dans le cas d'une variable, on factorise Q^* sur $(Z[Y,Z]/\Delta_3)[X]$, puis on calcule tous les produits possibles de facteurs en leur adjoignant le facteur $1-2Z$. On obtient ainsi le produit

$(1 - 2Z)(X - 1 + 2XZ - 8XZ^2)(X - 2 - 4XZ + 16XZ^2)$ qui est égal à

$X^2 - 3X + 2 - 2X^2Z$ sur $Z[Y,Z]/\Delta_3\,[X]$ et qui divise Q sur $Z[X,Y,Z]$.

Algorithme général de factorisation.

L'algorithme général présenté ci-dessous s'applique à un polynôme P à coefficients dans un anneau K. On étudiera en détails les deux cas les plus intéressants, à savoir $K = Z$ et $K = Z[X_1,...,X_m]$. Mais l'algorithme est valable pour tout anneau K, à condition de savoir factoriser la projection de P sur $K_o[X]$ et de définir des homomorphismes permettant de "remonter" de la factorisation de la projection de P sur $K_o[X]$ à la factorisation de P sur $K[X]$.

L'algorithme comporte les trois étapes suivantes :

.Etape 1.

On considère la surjection canonique π_o de K dans un anneau. On choisit K_o tel que a_o soit inversible sur K_o. La surjection canonique π_o induit un épimorphisme de $K[X]$ dans $K_o[X]$. Notons $a_o P_o$ l'image du polynôme P dans $K_o[X]$, P étant supposé sans facteurs carrés. De plus, π_o et K_o sont choisis de façon que P_o soit encore sans facteurs carrés sur $K_o[X]$, et de façon qu'on sache factoriser P_o sur $K_o[X]$. Dans la suite on notera Q_1^o, ..., Q_r^o la factorisation de P_o; remarquons que les facteurs Q_j^o sont unitaires, alors que les facteurs de P sur $K[X]$ n'ont aucune raison d'être unitaires. La factorisation de P sur $K[X]$ exigera donc de calculer les coefficients

dominants des facteurs Q_j^o. Pour cela on remarque que $a_o Q_j^o$ doit diviser $a_o P$; on obtient donc un diviseur de P en rendant primitif le polynôme $a_o Q_j^o$.
Cependant, on verra plus loin que Wang propose une autre méthode dans le cas des polynômes à plusieurs variables.

Polynôme à une variable.

On a $\mathbf{K} = \mathbf{Z}$, et on prend $\mathbf{K}_o = \mathbf{Z}/(p)$ où p est un nombre premier non diviseur de a_o et tel que P_o soit sans facteurs carrés sur $\mathbf{Z}/(p)$; la méthode de factorisation de P_o sur $\mathbf{Z}/(p)[X]$ est dûe à Berlekamp.

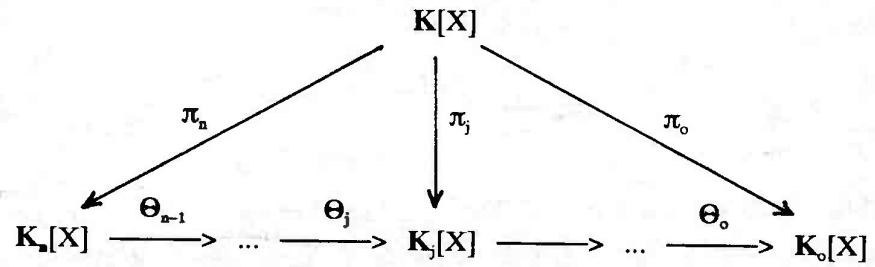
Polynôme à m variables.

On a $\mathbf{K} = \mathbf{Z}[X_1,...,X_m]$ et on choisit $\mathbf{K}_o = \mathbf{Z}$; alors $P_o = P(X,\alpha_1,..,\alpha_m)$ appartient à $\mathbf{Z}[X]$. Les nombres $\alpha_1, ..,\alpha_m$ sont choisis de façon que P_o reste sans facteur carré sur $\mathbf{Z}[X]$, et que le coefficient dominant de P_o, à savoir $a_o(\alpha_1,..,\alpha_m)$ ne s'annulle pas. Ensuite, P_o est factorisé sur $\mathbf{Z}[X]$ comme un polynôme à une variable.

.Etape 2.

On considère une suite de surjections canoniques π_j de $\mathbf{K}$ dans $\mathbf{K}_j = \mathbf{K}/\mathfrak{R}_j$ où $\mathfrak{R}_j$ est une relation d'équivalence telle que $a\ \mathfrak{R}_{j+1}\ b \Rightarrow a\ \mathfrak{R}_j\ b$. On note θ_j la surjection canonique de $\mathbf{K}_{j+1}$ dans $\mathbf{K}_j$; elle induit un épimorphisme Θ_j de $\mathbf{K}_{j+1}[X]$ dans $\mathbf{K}_j[X]$. Cette suite d'homomorphismes permet de "remonter" de la factorisation $Q_1^o \cdots Q_r^o$ de P_o sur $\mathbf{K}_o[X]$ à la factorisation $Q_1^n \cdots Q_r^n$ de P sur $\mathbf{K}_n[X]$ comme dans le schéma suivant :

(a)



Notons Ω_P le sous-ensemble de $\mathbf{K}[X]$ formé de P et de ses diviseurs $R_1,...,R_u$ sur $\mathbf{K}[X]$. On arrête la remontée dès qu'on a la propriété suivante:

(b) La surjection canonique $\pi_n: \mathbf{K}[X] \longrightarrow \mathbf{K}_n[X]$ laisse Ω_P invariant.

Dans le schéma **(a)** ci-dessus notons P_j la projection de P dans $\mathbf{K}_j[X]$. A la factorisation de P_o sur $\mathbf{K}_o[X]$ donnée à l'étape 1, on associe successivement la factorisation de P_1 sur $\mathbf{K}_1[X]$, puis la factorisation de P_2 sur $\mathbf{K}_2[X]$,... jusqu'à la factorisation de P_n sur $\mathbf{K}_n$. Cette dernière factorisation permettra à l'étape 3 de trouver la factorisation de P sur $\mathbf{K}[X]$.

Polynôme à une variable.

On a $\mathbf{K} = \mathbf{Z}$ et $\mathbf{K}_o = \mathbf{Z}/(p)$; on prend $\mathbf{K}_j = \mathbf{Z}/(p^N)$ où $N = 2^j$ avec $j = 1,..,n$; les coefficients de P_j et de ses diviseurs sont alors écrits comme des entiers compris entre $-\frac{1}{2}p^N$ et $\frac{1}{2}p^N$.

Si B_o désigne une borne des coefficients des diviseurs $R_1,..., R_u$ de P, on arrête la remontée dès que p^n devient supérieur à $2 \cdot B_o$. Alors la condition **(b)** est satisfaite.

Polynôme à m variables.

On a $\mathbf{K} = \mathbf{Z}[X_1,...,X_m]$ et $\mathbf{K}_o = \mathbf{Z}$; on prend $\mathbf{K}_j = \mathbf{Z}[X_1-\alpha_1,...,X_m-\alpha_m]/\Delta_j$ où Δ_j est l'idéal engendré par p^n et par les monômes de $\mathbf{Z}[X_1-\alpha_1,...,X_m-\alpha_m]$ de degré strictement supérieur à j, p^n étant un entier supérieur à $2 \cdot B_o$ (B_o désigne toujours une borne des coefficients des diviseurs de P). On peut représenter le polynôme P_j à l'aide des variables $Y_1 = X_1 - \alpha_1$, ..., $Y_m = X_m - \alpha_m$. Alors P_j s'écrit comme un polynôme en X dont les coefficients sont des polynômes en $Y_1,..,Y_m$ de degré total majoré par j et dont les coefficients sont compris entre $-\frac{1}{2}p^N$ et $\frac{1}{2}p^N$.

On arrête la remontée dès que j devient supérieur au degré de P. Alors la condition **(b)** est satisfaite.

.Etape 3.

Quand la condition **(b)** est satisfaite, on cherche à identifier tout diviseur Q_j de P_j sur $\mathbf{K}_j[X]$ à un polynôme de $\mathbf{K}[X]$. Les diviseurs de P_j sont obtenus en considérant tous les produits extraits du produit $Q^j_1,..., Q^j_r$ de P_j en facteurs irréductibles.

Mais on ne connait pas le coefficient dominant de Q_j; on sait seulement qu'il divise a_o; par suite, on cherchera un diviseur de $a_o P$ parmi tous les produits $a_o Q_j = a_o Q^j_u \cdot Q^j_v \cdots Q^j_w$. On vérifiera si $a_o Q_j$ divise $a_o P$ sur $\mathbf{K}[X]$ en effectuant la division

de $a_o P$ par $a_o Q_j$. Cependant, on doit considérer tous les produits extraits de $Q^j_1 \cdots Q^j_r$, à savoir 2^r produits, r étant majoré par le degré d de P. Si aucun de ces 2^r produits $a_o Q_j$ n'est un diviseur de $a_o P$ sur $\mathbf{K}[X]$, alors P est irréductible.
Le coût de cette troisième étape est exponentiel en r; comme r peut être égal à d, ce coût est donc exponentiel en d.

Généralisation de l'algorithme à un anneau factoriel K.

L'algorithme de factorisation présenté ci-dessus peut se généraliser à un anneau factoriel quelconque, à condition de trouver un sous-anneau $\mathbf{K_o}$ de $\mathbf{K}$ sur lequel on sache factoriser le polynôme donné P. On doit supposer que $\mathbf{K}$ est factoriel pour que P ait une factorisation unique sur $\mathbf{K}[X]$ de la forme $\alpha Q_1 \cdots Q_r$ où α est une unité de $\mathbf{K}$ et où $Q_1, \ldots, Q_r$ sont unitaires.
L'étape 1 consiste à projeter P sur un sous-anneau $\mathbf{K_o}$, puis à factoriser P sur $\mathbf{K_o}[X]$. L'étape 2 consiste à définir une suite d'anneaux $\mathbf{K_j}$ et une suite d'épimorphismes Θ_j de $\mathbf{K_{j+1}}[X]$ dans $\mathbf{K_j}[X]$ permettant de passer de la factorisation sur $\mathbf{K_j}[X]$ à la factorisation sur $\mathbf{K_{j+1}}[X]$.
Ce type de généralisation pourrait par exemple être appliqué à la factorisation des polynômes sur $\mathbf{C}[X,X_1,\ldots,X_m]$. L'étape 1 consiste à choisir un m-uple $(a_1,..,a_m)$, puis à factoriser $P(X,a_1,\ldots,a_m)$ sur $\mathbf{C}[X]$ en calculant ses racines à l'aide d'une méthode de calcul numérique. Ensuite les calculs sont identiques à ceux de la factorisation sur $\mathbf{Z}[X,X_1,..,X_m]$.

Variantes et améliorations classiques.

Polynômes à une variable de petits coefficients.

Signalons que dans le cas des polynômes à une variable, on peut s'arranger pour qu'aucune remontée ne soit nécessaire. Pour cela $\mathbf{K}_o$ doit vérifier la condition **(b)**, c'est-à-dire que les coefficients des diviseurs de P doivent être compris entre $-\frac{1}{2}p$ et $\frac{1}{2}p$.

Algorithme de Lenstra.

L'algorithme moins classique de Lenstra donné dans **[L3]**, permet de supprimer l'étape 3 et ainsi d'éviter la complexité exponentielle, mais il exige que les calculs de l'étape 2 soient effectués avec une valeur de p^N beaucoup plus grande que dans l'algorithme classique. Ce dernier algorithme de Lenstra a une complexité polynômiale en d, mais son exécution pratique est le plus souvent plus lente que celle de l'algorithme classique.

Récursivité sur la suite des variables $X_1,...,X_m$.

Pour les polynômes à m variables, Wang a proposé dans **[Wa]**, une amélioration de son algorithme initial. Cette amélioration consiste à effectuer la factorisation de façon récursive sur la suite $X_1, .., X_m$ des variables. Plus précisément, on définit la factorisation de

$P(X, X_1, ..,X_q, \alpha_{q+1}, .., \alpha_m)$ sur $\mathbf{Z}[X, X_1, .., X_q]$ à partir de la factorisation de $P(X, X_1, ..,X_{q-1}, \alpha_q, .., \alpha_m)$ sur $\mathbf{Z}[X, X_1, .., X_{q-1}]$.

Le processus est initialisé en partant de la factorisation de $P(X, \alpha_1, .., \alpha_m)$ sur $\mathbf{Z}[X]$, à l'aide de l'algorithme de factorisation des polynômes à une variable. Le principe de l'algorithme est donc le suivant :

A l'étape 1, on choisit $\mathbf{K}_o = \mathbf{Z}$ et $P_o = P(X,X_1,..,X_{q-1},\alpha_q,..,\alpha_m)$. P_o est factorisé sur $\mathbf{K}_o[X, X_1, .., X_{q-1}]$ en utilisant la récursivité de l'algorithme.

A l'étape 2, on prend $\mathbf{K}_j = \mathbf{K}_o[X_q-a_q]/\Delta_j$ où Δ_j est l'idéal engendré par p^N et par les monômes de $\mathbf{K}_o[X_q-\alpha_q]$ de degré supérieur ou égal à j. Cette étape consiste à

factoriser $P_j = P(X, X_1, .., X_q, \alpha_{q+1}, .., \alpha_m)$ sur $\mathbf{K}_j[X, X_1, .., X_{q-1}]$, P_j étant considéré comme un polynôme de variables $X, X_1, ..., X_{q-1}$ à coefficients dans $\mathbf{K}_j$.

A l'étape 3, on détermine les facteurs de $P(X, X_1, .., X_q, \alpha_{q+1}, .., \alpha_m)$ sur $\mathbf{Z}[X, X_1, ..., X_q]$.

Elimination des facteurs parasites.

On appelle facteurs parasites, les diviseurs de P sur $\mathbf{K}_n[X]$ qui ne sont pas des diviseurs sur $\mathbf{K}[X]$. Pour diminuer la probabilité d'avoir des facteurs parasites, on peut factoriser P sur une dizaine d'ensembles $\mathbf{K}_o[X]$ et ne conserver que la factorisation ayant donné le nombre minimum de facteurs.

Dans le cas de polynômes à plusieurs variables, on a $\mathbf{K}_o = \mathbf{Z}$ et la probabilité d'avoir des facteurs parasites est plus faible qu'avec $\mathbf{K}_o = \mathbf{Z}/(p)$.

Ensuite, on factorise $P_o(X,X_1,..,X_{q-1}, \alpha_q, .., \alpha_m)$ sur $\mathbf{K}_o[X, X_1,..,X_{q-1}]$, en utilisant la récursivité sur la suite des variables. La probabilité d'avoir des facteurs parasites est encore plus faible, dès que $q > 1$, puis elle diminue jusqu'à $q = m-1$.

Calcul des coefficients dominants des diviseurs à m variables.

Ce calcul proposé par Wang dans **[Wa]**, n'est valable que si la factorisation de l'étape 1 sur $\mathbf{K}_o[X]$ ne donne aucun facteur parasite.

Supposons que ce soit le cas et notons $b_o(X_1,...,X_m)$ le coefficient dominant d'un diviseur Q de P. Considérons la factorisation en facteurs irréductibles du coefficient dominant de P,

$a_o(X_1,...,X_m) = c_1(X_1,...,Xm)^{N1}\cdots c_t(X_1,...,X_m)^{Nt}$.

Alors Wang donne un critère simple qui permet de dire si un facteur de la forme $c_j(X_1,...,X_m)^N$ divise $b_o(X_1,...,X_m)$. On obtient ainsi les coefficients dominants des différents diviseurs de P. La connaissance de ces coefficients améliore beaucoup le temps de calcul. L'étude détaillée de cet algorithme est faite en **(B.I)**.

Variantes et améliorations personnelles.

Les différents algorithmes présentés dans les pages qui suivent proposent des modifications originales et personnelles de l'algorithme classique. Certains des résultats proposés ont déjà été publiés et sont référencés sous la forme **[V1,...,V6]**. Dans le cas des polynômes à une variable, on propose deux variantes:

(A_1). Suppression de l'étape 1 en donnant sans calculs une racine simple de P sur $\mathbf{Z}/(p)[X]$. On en déduit une racine de P sur $\mathbf{Z}/(p^N)[X]$ par une méthode analogue à celle de Newton. L'algorithme de Lenstra permet d'en déduire un facteur de P sur $\mathbf{Z}[X]$. Cette méthode est présentée dans **[V4]**. L'étude détaillée est faite ci-dessous en **(A.IV)**.

(A_2). On transforme $P(X)$ en un polynôme $P^*(X)$ représentant la dérivée logarithmique de $P(X)$. On obtient ainsi des majorations qui doivent être vérifiées par les diviseurs de P sur $\mathbf{Z}[X]$. En déterminant la factorisation de P sur $\mathbf{Z}/(p^N)$ avec N suffisamment grand, tous les calculs de produits de polynômes de l'étape 3 sont remplacés par des calculs de sommes.

En fait, on propose une solution intermédiaire qui n'augmente pas la complexité de l'étape 2, mais qui diminue le nombre de produits à calculer. La complexité reste donc exponentielle en d. Cette méthode est présentée dans **[V5]** et **[V6]**. L'étude détaillée est faite ci-dessous en **(A.V)**.
Dans le cas des polynômes à plusieurs variables on propose trois variantes:

(B_1). Le changement de variables de la forme $Y_j = X_j - \alpha_j$ de l'étape 1 dans l'algorithme classique peut être remplacé par le changement $Y_j = X_j - \alpha_j X_1$. Alors $P(X,X_1,..,X_n)$ est transformé en un polynôme unitaire ayant une propriété caractéristique qui doit aussi être vérifiée par ses diviseurs.

Cette transformation de P permet sans beaucoup changer le coût des calculs, d'obtenir un critère de reconnaissance des facteurs de P sur $\mathbf{K}[X]$, ce qui accélère les calculs de l'étape 3. Cette méthode est présentée dans **[V2]**.

(B_2). Rappelons que dans le cas des polynômes à une variable, on peut s'arranger pour qu'aucune remontée ne soit nécessaire (Voir le paragraphe 1 de la page 8). On peut aussi supprimer cette remontée pour les polynômes à plusieurs variables. Pour cela il suffit de vérifier la condition **(b')** suivante qui est analogue à **(b)**:

(b') $\left\{ \begin{array}{l} \text{Il existe un homomorphisme } \Phi: \mathbf{K}[X] \longrightarrow \mathbf{K}_o[X] \\ \text{dont la restriction à } \Omega_P \text{ est injective.} \end{array} \right.$

Cette condition **(b')** est vérifiée si on substitue aux variables $X_1,..,X_m$ des entiers $\alpha_1,...,\alpha_m$ suffisamment grands. On verra qu'alors la factorisation du polynôme à une variable $P(X,\alpha_1,..,\alpha_m)$ permet de retrouver les facteurs de $P(X,X_1,..,X_m)$ sur $\mathbf{Z}[X,X_1,..,X_m]$. Cette méthode est présentée dans **[V1]**.

(B_3). On applique la technique de **(A_2)** au cas d'un polynôme P(X) à coefficients dans $\mathbf{Z}[X_1,..,X_m]$. On obtient une relation que doivent vérifier les diviseurs de P sur $\mathbf{K}[X]$. Ce critère permet de diminuer le nombre de produits à calculer dans l'étape 3 de l'algorithme.

Algorithme de simplification des polynômes à plusieurs variables.

Le but de cet algorithme est de transformer le ou les polynômes étudiés en vue de faciliter les calculs ultérieurs, qui peuvent être par exemple, leur factorisation ou le calcul de leur PGCD.

(C). Notons le polynôme donné $P = \sum a_j X_1^{d1} \ldots X_m^{dm}$; on transforme P en un polynôme $P_o = \sum a_j X_1^{D1} \ldots X_m^{Dm}$ tel que l'application

ϕ: $(d1,..,dn) \longrightarrow (D1,..,Dm)$ soit une application linéaire. Une telle transformation permet de diminuer le degré du polynôme générique ou de le rendre unitaire. Elle permet aussi chaque fois que c'est possible de diminuer le nombre des variables. Cet algorithme est présenté dans **[V3]**.

Partie : A.

POLYNOMES A UNE VARIABLE.

Avant d'aborder le problème de la factorisation d'un polynôme P à coefficients entiers, on doit étudier plusieurs questions annexes comme le calcul du plus grand diviseur commun, l'évaluation d'une borne des coefficients des diviseurs du polynôme P, ainsi que la transformation de P en un polynôme unitiaire.

Donnons quelques définitions utiles pour la suite :

Le polynôme P est noté $a_o X^d + ... + a_d$. On dit que P est **primitif** si ses coefficients a_o, ..., a_d sont premiers entre eux.

Supposons P non primitif; on peut rendre P primitif, en le divisant par $PGCD(a_o, a_1, ..., a_d)$. Le polynôme primitif ainsi obtenu sera noté **prim(P)**.

Le coefficient a_o de P s'appelle le **coefficient dominant**; si $a_o = 1$, on dit que P est **unitaire**.

I. PROBLEMES PRELIMINAIRES A LA FACTORISATION.

1°) Calcul du plus grand commun diviseur de deux polynômes.

Toutes les propriétés et définitions de ce paragraphe sont empruntées à l'ouvrage de Knuth **[Kn]**.

Calcul du PGCD de deux polynômes de K[X] où K est un corps.

Soient deux polynômes A et B unitaires de K[X]. L'algorithme d'Euclide pour le calcul du PGCD de A et B consiste à construire une suite de couples (A_n, B_n) définis comme suit :

1) $(A_o, B_o) = (A, B)$;

2) on effectue la division de A_n par B_n; soit R le reste; alors on prend (A_{n+1}, B_{n+1}) égal à (B_n, R);

3) le dernier couple (A_n, B_n) est tel que $B_n = 0$; alors $PGCD(A, B) = A_n$.

Le coût de la division de A_k par B_k est celui du produit de $(d°B_n)\cdot(d°A_n - d°B_n)$ coefficients (en négligeant les additions). Si on note d_{An} et d_{Bn} les degrés de A_n et B_n, le calcul du PGCD de A et B revient à effectuer N produits, avec :

$$N = d_B(d_A - d_B) + d_{B1}(d_B - d_{B1}) + \ldots + d_{Bn}(d_{An} - d_{Bn}) + \ldots$$
$$= d_A\cdot d_B - d_B(d_B - d_{B1}) - \ldots - d_{An}(d_{An} - d_{Bn}) - \ldots$$
$$\leq d_A\cdot d_B.$$

Si on suppose que la taille des coefficients est majorée par la taille du mot machine de l'ordinateur utilisé, le coût du calcul du PGCD est donc donné par : $d_A\cdot d_B$.

Division de deux polynômes de Z[X].

Soient A et B deux polynômes de Z[X]. Si B n'est pas unitaire et s'écrit $B = b_o X^q + ... + b_q$, et si b_o ne divise pas le coefficient dominant de A, alors on ne peut pas définir la division de A par B. On peut cependant définir une autre opération appelée pseudo-division de A par B. Pour cela, on rend le polynôme B primitif, puis on calcule deux polynômes Q et R tels que :

(A.I.2) $(b_o)^{d°A-d°B+1} A = B·Q + R$ avec $d°R \leq d°B - 1$.

On peut vérifier facilement l'existence et l'unicité des deux polynômes Q et R.

Calcul du plus grand diviseur commun de A et B noté PGCD(A,B).

On peut opérer comme dans l'algorithme classique d'Euclide. Pour cela, on remarque que tout diviseur de A et B divise aussi R.

Réciproquement tout diviseur Δ de B et R divise $(b_o)^{d°A-d°B+1} A$, mais puisque B est primitif, ce diviseur Δ ne peut être que primitif, donc il divise A. On a ainsi l'égalité PGCD(A, B) = PGCD(B, R), à condition que B soit primitif. Pour pouvoir continuer le processus, il faut donc rendre R primitif. En conséquence, on obtiendra PGCD(A, B), en calculant la suite des couples (A_n, B_n) définis comme suit :

1) $(A_o, B_o) = (A, B)$;

2) on effectue la pseudo-division de A_n par B_n; soit R le reste;
alors on prend (A_{n+1}, B_{n+1}) égal à $(B_n, prim(R))$;

3) le dernier couple (A_n, B_n) est tel que $B_n = 0$.

Remarquons que $d°B_{n+1} \leq d°B_n - 1$; par suite on aura $B_n = 0$ au plus tard pour $n = d°B - 1$. Lorsque $B_n = 0$, il est clair que
$PGCD(A, B) = PGCD(A_{n-1}, B_{n-1}) = B_{n-1} = A_n$.

L'opération la plus coûteuse de cet algorithme est celle qui doit rendre primitif le reste R de la division de A_n par B_n. Les améliorations de l'algorithme consistent à supprimer ce calcul.

Examinons l'amélioration la plus simple; elle a été proposée par Collins dans **[C1]**. Avant de définir ce nouvel algorithme, étudions deux pseudo-divisions successives

$(b_o)^{N1} A_n = B_n·Q + R$ où $N1 = d°A_n - d°B_n + 1$,

$(r_o)^{N2} B_n = R· Q1 + R1$ où $N2 = d°B_n - d°R + 1$.

On peut démontrer que R1 est divisible par $\alpha = (b_o)^{N1}$ et que:

$PGCD(A_n, B_n) = PGCD(B_n, R) = PGCD(R, R1/(b_o)^{N1})$.

On en déduit le nouvel algorithme de Collins :

1) $(A_o, B_o) = (A, B)$; on prend $\alpha = 1$;

2) on effectue la pseudo-division de A_n par B_n, notée $(b_o)^N A_n = B_n \cdot Q + R$ où $N = d°A_n - d°B_n + 1$;
alors on prend (A_{n+1}, B_{n+1}) égal à $(B_n, R/\alpha)$, puis
α égal à $(b_o)^N$;

3) le dernier couple (A_n, B_n) est tel que $B_n = 0$.

Les algorithmes usuels de calcul du PGCD de deux polynômes ont profité d'autres améliorations; Collins propose un algorithme utilisant les "sous-résultants" des couples (A_n, B_n), dans **[C2]**; Brown propose dans **[Bo]** un algorithme fondé sur l'arithmétique modulaire.

Les coûts de ces algorithmes pour des polynômes de degré d sont de l'ordre de

(A.I.3) $d^2 \, Log(d)$.

Factorisation d'un polynôme P ayant au moins un facteur carré.

Supposons que P ait un facteur carré noté Q^2; alors $P = Q^2 \cdot R$. Par suite la dérivée P' de P s'écrit $P' = 2\, Q \cdot Q' \cdot R + Q^2 \cdot R$. Donc P et P' ont un PGCD de degré minoré par d°Q. Il en résulte que le discriminant Δ de P, à savoir $\Delta = PGCD(P,P')$, est un diviseur non trivial de P. D'autre part, on peut montrer que P/Δ n'a plus de facteur carré. Pour factoriser Δ, on peut lui appliquer le même traitement qu'à P. On peut donc toujours se ramener au cas suivant :

(A.I.4) P n'a pas de facteur au carré.

Dans la suite, on suppose que P vérifie **(A.I.4)**. Le coût de ces calculs est celui du calcul du PGCD de deux polynômes de degré d et d-1; ce coût est donc de l'ordre de $d^2 \, Log(d)$.

2°) Majoration des coefficients des diviseurs de P.

La plupart des résultats de ce paragraphe sont dûs à Maurice Mignotte. Ils figurent notamment dans le Chapitre IV, paragraphe 4, de son ouvrage **[M1]**. On peut les trouver également dans **[M2]**, **[Gu]**, **[Du]** et **[CMP]**.
Le polynôme donné P est écrit sous la forme

(A.I.5) $P = a_o X^d + ... + a_d = a_o (X - z_1) ... (X - z_d)$,

et l'un de ses diviseurs Q sur $\mathbf{Z}[X]$ est écrit sous la forme

(A.I.6) $Q = b_o X^n + b_1 X^{n-1} + ... + b_n = b_o (X - z_{N1}) ... (X - z_{Nn})$.

Le polynôme P s'écrit $Q \cdot R$ où le degré de R vaut $q = d - n$. Donnons quelques définitions, pour présenter les différentes majorations des coefficients b_j.

Définitions.

On appelle mesure de P le nombre M(P) égal au produit de $|a_o|$ et des nombres $\max(1, |z_k|)$, les nombres z_k étant les racines de P.
On appelle hauteur de P le nombre $H(P) = \max(|a_o|, ..., |a_d|)$. On appelle norme de P le nombre $\|P\| = \sqrt{|a_o|^2 + ... + |a_d|^2}$.

On appelle longueur de P le nombre $L(P) = |a_o| + ... + |a_d|$.
Enfin, on note $|P|$, le nombre $\max_{|z|=1} |P(z)|$.

Propriétés.

1) $H(P) \leq \|P\| \leq |P| \leq L(P) \leq (d+1) \cdot H(P)$.
2) $M(P) \leq \|P\|$.

Les inégalités de **1)** sont évidentes, sauf $\|P\| \leq |P|$ qui résulte de la formule de Parseval : $\|P\|^2 = \int_0^1 |P(e^{2i\pi t})|^2 \, dt$.

L'inégalité **2)** est dûe à Landau; elle a été redémontrée entre autre par Mignotte dans **[M2]**.

Dans la suite du paragraphe, le polynôme P est égal au produit $Q \cdot R$ et Q s'écrit $b_o X^n + b_1 X^{n-1} + \ldots + b_n$.

Inégalités utiles.

(A.I.7) $|b_j| \leq \binom{n}{j} M(P)$ (Mignotte **[M1]**)

(A.I.8) $L(Q) \leq 2^n M(P)$ (Mignotte **[M1]**).

(A.I.9) $H(Q) \leq \binom{[\frac{1}{2}n]}{n} \cdot \| P \|$ (Landau–Mignotte **[M1]**).

(A.I.10) $L(Q) \cdot L(R) \leq 2^d M(P)$ (Mignotte **[M1]**).

(A.I.11) $L(Q) \cdot M(R) \leq \dfrac{d!}{n!} L(P)$ pour $d \leq n + 3$ (Güting **[Gu]**).

(A.I.12) $|Q| \cdot L(R) \leq \dfrac{d!}{n!} |P|$ (Durand **[Du]**).

(A.I.13) $\| Q \| \cdot M(R) \leq \prod_{k=n}^{d-1} \dfrac{1}{\sin(\pi/(k+2))} \| P \|$ (Donaldson et Rahman **[DR]**).

(A.I.14) $|Q| \cdot M(R) \leq \dfrac{d^d}{n^n q^q} \| P \|$ (Mignotte **[M4]**).

Démonstrations des inégalités (A.I.7) à (A.I.10).

Si on note $z_1, \ldots, z_d$ les racines de P, alors on a

$|b_j| \leq \sum |z_{u1}| \cdot |z_{u2}| \cdots |z_{uj}|$ où on fait la somme de tous les produits $|z_{u1}| \cdots |z_{uj}|$ de j facteurs. Comme le nombre de ces produits est égal à $\binom{n}{j}$,

et que chaque produit $|z_{u1}|\cdot\cdot|z_{uj}|$ est majoré par la mesure $M(Q)$ du polynôme Q, alors on obtient l'inégalité $|b_j| \le \binom{n}{j} M(Q)$. Il est clair que $M(Q) \le M(P)$. On en déduit donc la première inégalité.

La deuxième inégalité se déduit aussitôt de la première. On obtient facilement l'inégalité suivante, grâce à l'inégalité $M(P) \le \|P\|$, donnée ci-dessus page 17. En combinant l'inégalité $|b_j| \le \binom{n}{j} M(Q)$, démontrée ci-dessus, à l'inégalité correspondante pour le diviseur R, on obtient la quatrième relation.

Plusieurs questions peuvent se poser pour choisir la meilleure majoration. Rappelons les comparaisons suivantes effectuées dans **[M4]** :

Les deux premières inégalités de Mignotte (c'est-à-dire $|b_j| \le \binom{n}{j} M(P)$ et $L(Q) \le 2^n M(P)$) sont les meilleures pour $d \ge n + \log(n) := M$; par contre les inégalités de Güting et Durand sont meilleures que la deuxième inégalité de Mignotte pour $d \le M$, mais dans l'intervalle l'inégalité de Donaldson et Rahman est meilleure que les inégalités de Güting et Durand. Enfin la dernière inégalité de Mignotte, à savoir

$$|Q| \cdot M(R) \le \frac{d^d}{n^n q^q} \|P\|$$

est meilleure que celle de Donaldson et Rahman pour $d \ge n + 10$.

D'autre part, on peut se demander si ces majorations ne sont pas trop restrictives. Une première réponse est que la deuxième majoration de Mignotte $L(Q) \le 2^n M(P)$ est la meilleure majoration possible écrite sous cette forme. En effet, dans **[M4]**, Mignotte montre qu'on ne peut pas remplacer 2 par $2 - \alpha$, quel que soit $\alpha > 0$. Cependant lorsque la mesure $M(P)$ n'est pas grande, on peut obtenir une meilleure borne. Mignotte donne dans **[M4]**, la majoration suivante :

(A.I.15) $\|Q\| \le e^{do} (d + 2\sqrt{d} + 2)^{do+1} M(P)^{do+1}$ où $do = \sqrt{d}$.

En comparant grossièrement cette majoration avec la majoration $L(Q) \le 2^n M(P)$, notée ci-dessus **(A.I.8)**, on peut dire que le coefficient 2^d de **(A.I.8)** est remplacé dans **(A.I.15)** par $2^{do(\text{Log } d)\cdot(\text{Log } M(P))/(\log 2)}$. Comme $do = \sqrt{d}$ la majoration est meilleure que **(A.I.8)** lorsque $(\text{Log } d)\cdot(\text{Log } M(P))/(\log 2) \le \sqrt{d}$.

D'autre part, il est presque toujours possible de donner une meilleure majoration de $H(Q)$ en combinant **(A.I.8)** avec la méthode de Graeffe. La méthode de Graeffe consiste à considérer le polynôme :

$$P(X)\cdot P(-X) = (a_o)^2(X - z_1)\cdots(X - z_d)\cdot(-X - z_1)\cdots(-X - z_d)$$

$= (a_o)^2(z_1^2 - X^2)\cdots(z_d^2 - X^2)$.

En notant $P_1(Y)$ le polynôme $(a_o)^2(z_1^2 - Y)\cdots(z_d^2 - Y)$, il est clair que $M(P_1) = M(P)^2$. On calcule les polynômes suivants P_2, ..., P_k tels que

$M(P_k) = M(P_{k-1})^2 = \ldots = M(P)^N$ où $N = 2^{k-1}$.

On obtient donc les majorations $M(P) \leq \| P_1 \|^{1/2}$, ..., $M(P) \leq \| P_k \|^{1/N}$. Cet algorithme est présenté dans **[CMP]**; on y démontre en particulier, que $\| P_k \|^{1/N}$ tend vers $M(P)$ lorsque k devient infini (Proposition 1).

En calculant la suite des polynômes P_1, ..., P_k, on obtient donc un algorithme qui calcule un majorant de $M(P)$ qui tend vers $M(P)$.

Considérons par exemple le polynôme écrit en REDUCE sous la forme

P(X) := 6 X**5 - 8 X**4 + 6 X**3 + 9 X**2 - 5 X - 3.

Si on définit une procédure REDUCE notée Mesure(P,m), qui calcule la norme du polynôme P_m, alors en tapant Mesure(P,1), Mesure(P,2), Mesure(P,3) et Mesure(P,4), on obtient successivement Mes ≤ 15,84, Mes ≤ 15,5, Mes ≤ 13,14 et Mes ≤ 13,01.

3°) Factorisation d'un polynôme non unitaire.

Les méthodes de factorisation étudiées dans la suite ont comme point de départ, la factorisation du polynôme P sur l'anneau $\mathbf{Z}/p[X] = \mathbf{F}_p[X]$. Le nombre p est un nombre premier choisi de façon que a_o soit inversible sur $\mathbf{F}_p$ (soit $\neq 0$). Alors l'image de $P = a_o X^d + \ldots + a_d$ sur $\mathbf{F}_p[X]$ peut s'écrire $a_o(X^d + a_1\cdot(a_o)^{-1}\cdot X^{d-1} + \ldots + a_d\cdot(a_o)^{-1}) = a_o\cdot P_o$, où P_o est unitaire. On est donc ramené à la factorisation sur $\mathbf{F}_p[X]$ d'un polynôme unitaire P_o. Notons cette factorisation $Q_1^o\cdots Q_r^o$. A partir de cette factorisation sur $\mathbf{F}_p[X]$, on déduit la factorisation $Q_1\cdots Q_r$ de P_o sur $\mathbf{Z}/(p^m)[X]$, soit $P = a_o\cdot Q_1\cdots Q_r$ modulo p^m. Il en résulte que si le produit $Q_{j\,1}\cdots Q_{j\,k}$ correspond à un diviseur de P sur $\mathbf{Z}[X]$, alors $Q = a_o\cdot Q_{j\,1}\cdots Q_{j\,k}$ divise $a_o\cdot P$ sur $\mathbf{Z}[X]$. On peut donc obtenir un diviseur de P sur $\mathbf{Z}[X]$, en rendant primitif le polynôme Q.

Ainsi, la factorisation du polynôme unitaire P_o permet de trouver les facteurs de P sur $\mathbf{Z}[X]$.

Dans la suite on déterminera donc la factorisation d'un polynôme unitaire. Puis dans la phase finale de la factorisation, on multiplie chaque diviseur potentiel Q par a_o, puis on effectue la division de $a_o\cdot P$ par $a_o\cdot Q$. On trouvera ainsi tout diviseur de P.

II) METHODES CLASSIQUES DE FACTORISATION.

1°) Factorisation de P sur $F_p[X]$ par la méthode de Berlekamp.

Introduction.

L'algorithme de Berlekamp est d'abord présenté dans l'ouvrage de Berlekamp, Chapitre 6 dans **[Be]**, puis par Knuth dans **[Kn]**, au paragraphe 4.6.2. La présentation adoptée ici est voisine de celle de l'ouvrage de Mignotte **[M1]**.

La première étape consiste à choisir un nombre premier p tel que le polynôme P reste sans facteur carré sur $\mathbf{Z}/(p)[X]$ et conserve son degré d.

Tous les calculs sont ensuite effectués sur le corps $\mathbf{F}_p = \mathbf{Z}/(p)$. L'étape principale de l'algorithme est la recherche d'un polynôme A défini modulo P et vérifiant la relation $A^p - A = 0$. Cela revient à résoudre un système homogène de d équations à d inconnues. On obtient ensuite les facteurs de P en calculant tous les PGCD de P et des polynômes $A - \alpha$ pour $0 \leq \alpha \leq p-1$.

Cette dernière étape est la plus longue.

Choix du nombre premier p.

Rappelons que P est supposé sans facteur carré sur $\mathbf{Z}[X]$. Cette propriété est essentielle dans l'algorithme de Berlekamp et elle doit rester vraie sur $\mathbf{F}_p[X]$. Comme P est sans facteur carré, il est premier avec sa dérivée P'. D'après l'identité de Bezout, on peut donc trouver deux polynômes primitifs uniques U et V de $\mathbf{Z}[X]$ de degrés respectifs majorés par $d°P - 1$ et $d°P$ tels que :

(A.II.1) $U.P + V.P' = \delta \in \mathbf{Z}$

où δ est le discriminant de P. Il est clair que P et P' auront un diviseur commun sur F_p si et seulement si δ est nul sur F_p, c'est-à-dire si p divise δ. On devra donc choisir un nombre premier p qui ne soit pas un facteur de δ.
De plus, on veut que a_o soit inversible sur $F_p[X]$; le nombre p devra donc être choisi en dehors des facteurs premiers du coefficient dominant a_o de P.

Etude de l'anneau $\Omega_P = F_p[X]/(P)$.

L'algorithme de Berlekamp utilise l'algèbre $\Omega_P = F_p[X]/(P)$ où P est égal au produit $Q_1 \cdot Q_2 \cdots Q_r$ de polynômes irréductibles. Quelques propriétés de cet anneau Ω_P seront utiles dans la suite.

Théorème chinois.

Soient $Q_1,\ldots,Q_r$ des polynômes de $F_p[X]$ deux à deux premiers entre eux, de produit P. L'homomorphisme naturel de $\Omega_P = F_p[X]/(Q_1 \cdots Q_r)$ dans $F_p[X]/(Q_1) \times \cdots \times F_p[X]/(Q_r)$ est un isomorphisme.

Démonstration.

Notons f l'homorphisme canonique de $F_p[X]$ dans $F_p[X]/(Q_1) \times \cdots \times F_p[X]/(Q_r)$. Il est clair que son noyau est égal à l'idéal $(Q_1 \cdots Q_r)$. Donc l'homorphisme canonique de l'énoncé est bien injectif. Pour montrer que f est surjective considérons des polynômes $P_1, \ldots, P_r$ de $F_p[X]/(Q_1), \ldots, F_p[X]/(Q_r)$. Comme $Q_1,\ldots,Q_r$ sont deux à deux premiers entre eux, alors on peut décomposer la fraction rationnelle 1/P en éléments simples; on obtient :

$$\textbf{(A.II.2)} \qquad \frac{1}{Q_1 \cdots Q_r} = \frac{S_1}{Q_1} + \ldots + \frac{S_r}{Q_r}.$$

Posons $A = \left(P_1 \frac{S_1}{Q_1} + \ldots + P_r \frac{S_r}{Q_r} \right) Q_1 \cdots Q_r$. Alors A est congru modulo Q_j à $P_j \frac{S_j}{Q_j} Q_1 \cdots Q_r$, donc d'après (A.II.2), A est congru à P_j modulo Q_j. Le théorème est donc démontré.

Dans la suite, on suppose que $P = Q_1 \cdot Q_2 \cdots Q_r$ est un produit de facteurs irréductibles et deux à deux distincts. Il résulte du théorème chinois que tout polynôme A de Ω_r peut s'écrire comme un r-uple $(A_1, ..., A_r)$ où $A_j \in \mathbf{F}_p[X]/(Q_j)$.

D'autre part, il est clair que tout élément de $\mathbf{F}_p[X]/(Q_j)$ est défini par un polynôme de degré $q_j - 1$, où q_j est le degré de Q_j. Par suite, $\mathbf{F}_p[X]/(Q_j)$ est isomorphe à $(\mathbf{F}_p)^{q_j}$, et il en résulte que Ω_r est isomorphe au produit $(\mathbf{F}_p)^{q_1} \times (\mathbf{F}_p)^{q_2} \times \cdots \times (\mathbf{F}_p)^{q_r}$.

On considère l'application f de $\Omega_r = \mathbf{F}_p[X]/(P)$ dans lui-même, définie par $f(Q) = Q^p - Q$; on peut écrire f sous la forme $(f_1,..., f_j,..., f_r)$ où f_j est définie sur $\mathbf{F}_p[X]/(Q_j)$. On peut vérifier facilement que f et f_j sont des endomorphismes. En effet, si a et b sont deux éléments de $\mathbf{F}_p$, alors

$$(a X + b Y)^p = a^p X^p + b^p Y^p = a X^p + b Y^p.$$

Par suite $f_j(a X + b Y) = a f_j(X) + b f_j(Y)$ et f_j est bien un endomorphisme. On en déduit que f est également un endomorphisme.

Proposition 1.

1°) On a pour tout X de $\mathbf{F}_p[X]/(P)$:

(A.II.3) $$X^p - X = X \cdot (X - 1) \cdots (X - \alpha) \cdots (X - p - 1).$$

2°) Le noyau de l'endomorphisme f est isomorphe à $(\mathbf{F}_p)^r$, où r est le nombre de facteurs irréductibles de P.

3°) Le nombre de facteurs irréductibles de P est égal à la dimension du noyau de Φ.

Démonstration.

Tout élément α de $\mathbf{F}_p$ vérifie l'équation $X^p - X = 0$. On peut donc mettre $X - \alpha$ en facteur pour $\alpha \in [0, p-1]$, et la relation **(A.II.3)** en résulte. D'après cette relation **(A.II.3)** le noyau de f_j est égal à $\mathbf{F}_p$, donc le noyau de f est égal à $(\mathbf{F}_p)^r$.

Recherche des facteurs irréductibles de P.

Notons A un polynôme du noyau de f de degré > 0; comme il est défini modulo P, A s'écrit $a_0 + a_1 X + ... + a_{d-1} X^{d-1}$.

D'après **(A.II.3)**, $A \cdot (A - 1) \cdots (A - p - 1)$ est un multiple de P, donc tout facteur irréductible Q_j de P divise le produit $A \cdot (A-1) \cdots (A-p-1)$, donc l'un des polynômes $A - \alpha$ pour $0 \le \alpha \le p-1$. Le calcul de PGCD(P, A−α) pour $\alpha = 0,..., p-1$

fournira donc des diviseurs non triviaux de P; malheureusement deux facteurs irréductibles Q_j et Q_k peuvent diviser le même polynôme $A - \alpha$, donc les diviseurs de P ainsi obtenus ne sont pas nécessairement irréductibles.

La connaissance d'un polynôme A de degré > 0 du noyau de f permet donc de donner une factorisation partielle de P. Pour trouver un tel polynôme A, on calcule $A^p - A$ modulo P, où A s'écrit sous la forme $a_o + a_1 X^p + ... + a_{d-1} X^{(d-1)p}$.

On calcule ensuite $A^p - A$, en remplaçant dans A^p, X^p,...,$X^{(d-1)p}$ par les restes de leurs divisions par P. On écrira donc X^{jp} sous la forme $A_{j\,0} + A_{j\,1} X + ... + A_{j\,d-1} X^{d-1}$.

Le coefficient de X^j dans le développement du polynôme $A^p - A$ s'écrit

$$a_1 A_{1\,j} + a_2 A_{2\,j} + ... + a_j(A_{j\,j} - 1) + ... + a_{d-1} A_{d-1\,j}.$$

En écrivant que les d coefficients de X^0, X^1, ..., X^{d-1} sont nuls on obtient un système homogène de d équations linéaires à d inconnues a_o, a_1, .. , a_{d-1}. Notons ce système (S). Si ce système est de rang t, alors la solution générale s'écrit :

(A.II.4) $A = u_1 A_1 + ... + u_t A_t$ avec $0 \le u_j \le p-1$.

Rappelons que la recherche des diviseurs de P se fait ensuite en calculant le PGCD de P et de $A_1 - \alpha$ pour tout nombre α de $[0, p-1]$.

Coût de l'algorithme.

La méthode de Berlekamp comporte plusieurs étapes. La première étape consiste à calculer les monômes X^{jp} modulo P. Le calcul de X^p peut être obtenu en effectuant $\mathrm{Log}(p)$ élévations au carré, ayant chacune un coût de d^2 produits de coefficients. Le produit de deux coefficients a un coût égal à $\mathrm{Log}^2 p$. Le coût du calcul de X^p vaut donc $d^2 \mathrm{Log}^3 p$.

Le calcul de chacune des autres puissances X^{jp} exige seulement le calcul d'un nouveau produit de polynômes, donc il a un coût égal à $d^2 \mathrm{Log}^2 p$. Cette première étape a donc un coût total égal à $d^2 \mathrm{Log}^3 p + d^3 \mathrm{Log}^2 p$.

La deuxième étape exige de résoudre un système de d équations à d inconnues. En utilisant la méthode du pivot, on doit effectuer $\frac{1}{2}d^3$ opérations élémentaires sur les coefficients. Comme les coefficients sont définis modulo p, le coût est donné par $\frac{1}{2}d^3 \mathrm{Log}^2 p$.

Dans le cas où le polynôme P n'est pas irréductible, on doit calculer jusqu'à p PGCD de polynômes de degré d. D'après l'étude faite plus haut, le calcul du PGCD par l'algorithme d'Euclide a un coût de $d^2 \text{Log}^2(p) \text{Log}(d)$.
La factorisation partielle de P a donc un coût total égal à

(A.II.5) $p\, d^2 \text{Log}^2 p \text{ Log}(d) + d^3 \text{Log}^2 p$.

Pour obtenir une factorisation complète, on peut considérer plusieurs polynômes A, puis en effectuant pour chacun d'eux les p calculs de PGCD. On peut espérer ainsi trouver rapidement les r facteurs irréductibles de P. En effet on sait que tout nouveau polynôme A donne une factorisation différente. Le coût est alors de l'ordre de

(A.II.6) $p\, d^3 \text{Log}^2 p$.

Une autre possibilité est d'appliquer l'algorithme à chacun des facteurs de la factorisation partielle obtenue. Le coût est du même ordre que si on prend plusieurs polynômes A, à savoir $p\, d^3 \text{Log}^2 p$.

Exemple.

Considérons le polynôme $P = X^5 + 3X^3 - X^2 + 2X - 1$.
Choisissons $p = 5$. Calculons $A(X) = a_o + a_1 X + a_2 X^2 + a_3 X^3 + a_4 X^4$.
$A^p - A = a_4 X^{20} + a_3 X^{15} + a_2 X^{10} + a_3 X^{15} - a_4 X^4 - a_3 X^3 - a_2 X^2 - a_1 X$.On calcule X^5, X^{10}, X^{15} et X^{20} modulo P. On obtient :

$$
\begin{aligned}
X^5 &= 1 - 2X + X^2 - 3X^3 , \\
X^{10} &= 2X + 2X^2 + 2X^3 + X^4 , \\
X^{15} &= -2 + X - 2X^2 + 2X^3 , \\
X^{20} &= 1 - 2X - X^2 - 2X^3 - X^4 .
\end{aligned}
$$

On en déduit :

$$
\begin{aligned}
A^5 - A = {} & (a_1 - 2a_3 + a_4) + \\
& X\,(-3a_1 + 2a_2 + a_3 - 2a_4) + \\
& X^2\,(a_1 + a_2 - 2a_3 - a_4) + \\
& X^3\,(2a_1 + 2a_2 + a_3 - 2a_4) + \\
& X^4\,(a_2 - 2a_4) .
\end{aligned}
$$

Le tableau des coefficients du système d'équations qu'on doit résoudre s'écrit :

a_o	a_1	a_2	a_3	a_4
0	1	0	−2	1
0	−3	2	1	−2
0	1	1	−2	−1
0	−3	2	1	−2
0	0	1	0	−2

Après la résolution on obtient le tableau suivant :

a_o	a_1	a_2	a_3	a_4
0	1	0	−2	1
0	0	1	1	−2
0	0	0	0	0
0	0	0	0	0
0	0	0	0	0

Le système est donc de rang 3 et ses solutions s'écrivent:

$$S = u \begin{vmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{vmatrix} + v \begin{vmatrix} 0 \\ 2 \\ 0 \\ 1 \\ 0 \end{vmatrix} + w \begin{vmatrix} 0 \\ -1 \\ 2 \\ 0 \\ 1 \end{vmatrix}$$

Le polynôme donné P admet donc 3 facteurs irréductibles sur $\mathbf{F}_5$. Si on prend la solution $A = 2X + X^3$ le calcul des PGCD donne

$$\begin{aligned} \text{PGCD}(P , X^3 + 2X\ \ ) &= 1 \\ \text{PGCD}(P , X^3 + 2X + 1) &= 2 \\ \text{PGCD}(P , X^3 + 2X + 2) &= X + 2 \\ \text{PGCD}(P , X^3 + 2X + 3) &= X - 2 \\ \text{PGCD}(P , X^3 + 2X + 4) &= X^3 + 2X - 1. \end{aligned}$$

Donc $P = (X^3 + 2X - 1)(X + 2)(X - 2)$ sur $\mathbf{F}_5$.

2°) Etude d'une autre méthode de factorisation de P sur $F_p[X]$.

Plusieurs autres méthodes sont fondées sur le même principe que la méthode de Berlekamp, mais cherchent à diminuer les calculs de PGCD de la dernière étape. L'une des plus intéressantes de ces méthodes est celle qui a été proposée indépendamment par Cantor et Zassenhaus dans **[CZ]**, et par Camion dans **[Ca]**. La présentation adoptée ici est celle de l'ouvrage de Mignotte **[M1]**.

Présentation de l'algorithme.

On note comme ci-dessus, P le polynôme à factoriser et on suppose que p est un nombre premier impair. On considère comme plus haut, l'endomorphisme f de $F_p[X]/(P)$, tel que $f(A) = A^p - A$ pour tout $A \in F_p[X]/(P)$; d'après l'isomorphisme du théorème chinois, A s'écrit $(A_1,..,A_j,...,A_r)$ où $A_j \in F_p[X]/(Q_j)$. Dans la suite on note A_o un élément du noyau de f. D'après le 2°) de la Proposition 1, A_o s'écrit $(a_1,...,a_r)$ où $a_j \in F_p$. La méthode de Berlekamp consiste à prendre un polynôme A_o du noyau de f, puis à calculer tous les PGCD de P et de $A_o - \alpha$ avec $\alpha = 0,..., p-1$. Il en résulte que le coût de ce calcul est proportionnel à p, ce qui oblige à choisir une valeur de p qui ne soit pas trop grande. Le lemme suivant permettra d'éviter ces nombreux calculs de PGCD.

Lemme.

Soit p un nombre premier impair, f l'application définie sur $F_p[X]/(P)$ par $f(Q) = Q^p - Q$ et A_o un polynôme du noyau de f.
Alors tout diviseur irréductible de P divise l'un des trois polynômes A_o, $A - 1$ et $A + 1$, où $A = (A_o)^{p-1}$.

Démonstration.

Comme A_o appartient au noyau de f, alors d'après le **2°)** de la Proposition 1, il s'écrit $A_o = (a_1,..,a_j,..,a_r) \in (F_p)^r$. Il en résulte que A^2 s'écrit sous la forme $(a_1^{p-1},..,a_j^{p-1},..,a_r^{p-1})$; comme $a_j^p = a_j$, alors $a_j^{p-1} = 1$ (ou 0 si $a_j = 0$). Par suite $A = (\delta_1,..,\delta_j,..,\delta_r)$ où $\delta_j = \pm 1$ (ou 0 si $a_j = 0$).

Comme A_o appartient au noyau de f, alors $A_o^p - A_o$ est nul modulo P. Par suite $A_o^p - A_o = A_o \cdot (A - 1) \cdot (A + 1)$ est un multiple de P. Le lemme est donc vérifié.
L'algorithme de factorisation consiste à résoudre le système d'équations définissant le noyau de f comme dans la méthode de Berlekamp. Ensuite on calcule les trois PGCD suivants PGCD(P,A_o), PGCD(P,A−1) et PGCD(P,A+1).
On obtiendra donc un diviseur non trivial de P, sauf si A−1 ou A+1 sont nuls modulo P. On obtient ainsi une factorisation partielle de P, à condition que la condition suivante soit vérifiée

(A.II.7) $A \neq \pm 1$ modulo P.

Calculons la probabilité pour que **(A.II.7)** soit vérifié. Considérons un polynôme $A_o = (a_1,...,a_r)$ du noyau de f; calculons $A = (A_o)^{(p-1)/2} = (\delta_1,..,\delta_j,..,\delta_r)$ où $\delta_j = \pm 1$. La probabilité pour que A_o soit tel que $A = (1,...,1)$ ou pour que $A = (-1,...,-1)$ est égale à $2 \cdot \alpha^r$ où α est la probabilité pour que $\delta_j = 1$. Il est clair que $\alpha = (p-1)/2 \cdot p$. Il en résulte que la probabilité pour que **(A.II.7)** soit vérifié est supérieure à ½.

Etude du coût.

Le calcul de A_o exige les mêmes calculs que dans l'algorithme de Berlekamp, à savoir $d^2 \, Log^3 p + d^3 \, Log^2 p$.
Le calcul de $A_o^{(p-1)/2}$ exige d'effectuer $Log(p/2)$ produits de polynômes, à savoir $d^2 (Log^2 p) \cdot (Log p - 1)$ produits élémentaires. En ajoutant les 3 calculs de PGCD on obtient un coût de $d^2 (Log^2 p) \cdot (Log p + 2)$. Comme la probabilité d'obtenir ainsi une factorisation non triviale de P est supérieure à ½, le coût moyen d'une factorisation partielle de P est de l'ordre de

(A.II.8) $d^2 \, Log^3 p + d^3 \, Log^2 p$,

ce qui est meilleur que le coût de l'algorithme de Berlekamp surtout lorsque p devient grand.

3°) Factorisation de P sur Z/(p^q)[X]. (Méthode de Zassenhaus)

Cette méthode est notamment présentée par Zassenhaus dans **[Za]** et par Wang et Rothschild dans **[WR]**. On suppose que P est unitaire comme dans le paragraphe précédent, et on reprend les notations de l'étape 2 de l'introduction.

Dans cette deuxième étape on considère la suite d'anneaux $\mathbf{K}_j = \mathbf{Z}/(p^{Nj})$ où $Nj = 2^j$. Cette suite permet de remonter de $\mathbf{K}_o = \mathbf{F}_p$ à $\mathbf{K}_{n+1}$.

Notons P_j la projection de P sur $\mathbf{K}_j[X]$. Pour chacune des surjections canoniques Θ_j: $\mathbf{K}_{j+1}[X]$ —> $\mathbf{K}_j[X]$, on doit construire l'image réciproque P_{j+1} de P_j. Pour simplifier les notations, posons $q = p^{Nj}$, $R^* = P_{j+1}$; de plus notons $R_1...R_r$ la factorisation de $R = P_j$ sur $\mathbf{Z}/(q)$ obtenue à partir de la factorisation de P sur $\mathbf{Z}/(p)$ trouvée à l'étape 1. Notons $R_1^*...R_r^*$ la factorisation correspondante de $R^* = P_{j+1}$ sur $\mathbf{Z}/(q^2)$.

On peut écrire $R^* = R + q\,S$ et $R_j^* = R_j + q\,S_j$ où S, R, S_j et R_j appartiennent à $\mathbf{Z}/(q)$. Les polynômes S et S_j sont appelés les coefficients de q du polynôme P et de ses diviseurs sur $\mathbf{Z}/(q^n)[X]$.

On connait P et les projections de ses diviseurs sur $\mathbf{Z}/(q)[X]$, à savoir $R_1,..., R_r$. On veut calculer les coefficients de q des diviseurs de P. Le lemme ci-dessous nous donne un moyen d'obtenir ces coefficients.

Lemme.

La factorisation de P est notée $R_1 \cdot R_2 \cdots R_r$ sur $\mathbf{Z}/(q)[X]$ et $R_1^* \cdot R_2^* \cdots R_r^*$ sur $\mathbf{Z}/(q^2)[X]$ où $R_j^* = R_j + q\,S_j$.

Les polynômes S_j sont égaux aux numérateurs de la décomposition en éléments simples de la fraction rationnelle $\frac{S}{P}$, où S est la projection de P sur $\mathbf{Z}/(q)[X]$.

Démonstration.

Comme R^* et ses diviseurs R_j^* sont unitaires, alors on a **(A.II.9)** deg S < deg R; deg S_1 < deg R_1;...;deg S_r < deg R_r. On peut écrire R^* sous la forme

(A.II.10) $R + q\,S = (R_1 + qS_1)(R_2 + qS_2)....(R_r + qS_r)$ sur $\mathbf{Z}/(q^2)$.

Il en résulte les relations

$$\textbf{(A.II.11)} \quad \begin{cases} R = R_1 \cdot R_2 \cdots R_r \\ S = S_1 \cdot R_2 \cdots R_r + R_1 \cdot S_2 \cdot R_3 \cdots R_r + \cdots + R_1 \cdots R_{r-1} \cdot S_r \end{cases} \quad \text{sur } \mathbf{Z}/(q)$$

Si on note F la fraction $\frac{S}{R}$, les relations **(A.II.11)** ci-dessus impliquent :

$$\textbf{(A.II.12)} \qquad F = \frac{S_1}{R_1} + \frac{S_2}{R_2} + \ldots + \frac{S_r}{R_r}.$$

D'après les majorations de degrés **(A.II.9)**, on peut obtenir $S_1,..,S_r$ en décomposant la fraction rationnelle F en éléments simples. Le lemme est donc démontré.

En fait l'algorithme classique résoud directement **(A.II.11)** sans passer par les fractions rationnelles. Cette technique a été présentée pour la première fois dans **[V2]**, puis elle a été reprise par Lugiez dans **[Lu1]**. Il semble qu'en utilisant l'algorithme de décomposition des fractions rationnelles très efficace donné par Kung et Tong dans **[KT]**, on obtienne grâce à cette technique une méthode de factorisation plus rapide que la méthode classique.

Le processus général de remontée quadratique est dû à Zassenhaus et figure dans **[Za]**; un algorithme analogue, mais linéaire avait été donné par Hensel. L'algorithme de remontée permet ainsi de définir la factorisation de P sur l'anneau $\mathbf{Z}_p$ des nombres p-adiques.

Notons n le plus petit entier tel que les coefficients des diviseurs de P sur $\mathbf{Z}[X]$ soient compris entre $-\frac{1}{2}p^N$ et $\frac{1}{2}p^N$, sachant que $N = 2^n$. On arrête la "remontée" à l'anneau $\mathbf{K}_n = \mathbf{Z}/(p^N)$; alors les diviseurs de P sur $\mathbf{Z}[X]$ sont définis par leur projection sur $\mathbf{K}_n$.

Pour majorer les coefficients des diviseurs, utilisons l'une des bornes de Mignotte données dans **II.2**, par exemple $B_o = 2^d M(P)$ où $M(P)$ est la mesure de P. On choisira donc $p^N \geq B_o$, soit

$$\textbf{(A.II.13)} \qquad N \geq (d + \mathrm{Log}M(P))/\mathrm{Log}p.$$

Etude du coût.

A chaque étape on doit d'abord effectuer le produit des diviseurs qui ont été définis modulo p^N à l'étape précédente. Ensuite on doit décomposer une fraction rationnelle en éléments simples.

Le calcul du produit de r polynômes de degrés d_1, ..., d_r exige de calculer $d_1 \cdots d_r$ produits de coefficients, donc le coût de ce calcul vaut

$[d_1 \cdot d_2 + (d_1+d_2) \cdot d_3 + \cdots + (d_1+\cdots+d_{r-1}) \cdot d_r] \cdot N^2 \mathrm{Log}^2 p$.

Comme $d_1 + \cdots + d_r = d$, ce coût est maximum lorsque $d_1 = \cdots = d_r = d/r$; il est donc égal à $d^2[1 + 2 + \cdots + (r-1)]/r^2 \mathrm{Log}^p = \frac{1}{2} d^2 N^2 \mathrm{Log}^2 p$. La décomposition des fractions rationnelles exige $d \cdot \mathrm{Log}^2 d$ produits de coefficients d'après Kung et Tong **[KT]**. Le coût de cette décomposition est donc donnée par $d \cdot \mathrm{Log}^2 d \cdot N^2 \cdot \mathrm{Log}^2 p$. Au cours des étapes successives, N prend les valeurs $2, 2^2, \ldots, 2^n$.

Le coût total pour obtenir la factorisation de P sur $\mathbf{Z}/(p^N)[X]$, à partir de la factorisation de P sur $\mathbf{Z}/(p)[X]$ est donc majoré par $d^2 \cdot \mathrm{Log}^2 p \cdot (4 + 4^2 + \ldots + 4^{n-1})$ dont l'ordre de grandeur correspond aux calculs de la dernière étape, à savoir

(A.II.14) $\quad d^2 \cdot \mathrm{Log}^2(p^N)$.

Pour que les diviseurs définis modulo p^N puissent être idenifiés aux diviseurs de P sur $\mathbf{Z}[X]$, on doit choisir un entier N vérifiant **(A.II.13)**, à savoir
$N \geq d \cdot \mathrm{Log} p \cdot \mathrm{LogM}(P)$. Alors le coût total de cette deuxième étape est égal à

(A.II.15) $\quad d^2 (d + \mathrm{LogM}(P))^2$.

Remarquons que si on cherche seulement une factorisation partielle de P, alors on peut ne considérer que les diviseurs de P de degré majoré par ½d. Les coefficients de ces diviseurs sont majorés par $2^{d/2}$ M(P). Alors le coût total de la deuxième étape de l'algorithme de factorisation partielle est égal à

(A.II.16) $\quad d^2 (\frac{1}{2}d + \mathrm{LogM}(P))^2$.

Ces coûts dépendent donc très directement de la borne prise pour majorer les coefficients des diviseurs de P. Il est donc utile d'utiliser les majorations les plus fines proposées par Mignotte et analysées en **A,I.2°)**.

3°) Factorisation de P sur Z[X].

Méthode classique pour une factorisation partielle ou complète.

Notons $Q_1...Q_j...Q_r$ la factorisation de P sur $\mathbf{Z}/(p^N)[X]$, obtenue au paragraphe précédent, avec $p^N > 2^d M(P)$.

La troisième étape de la factorisation consiste à calculer tous les produits $Q_{j1}...Q_{jt}$ extraits du produit $Q_1...Q_r$, puis à vérifier si l'un d'eux divise P sur Z[X]. Une méthode consiste à construire la liste L des polynômes égaux aux produits $Q_{j1}...Q_{jt}$ qui ont un degré majoré par ½ d et qui sont formés uniquement des facteurs $Q_1, ...,Q_j$.

On peut utiliser l'algorithme suivant qui construit la liste L:

Initialisation: L := {1}, Diviseur := ϕ;

Itération:

Pour j de 1 jusqu'à Fin faire

{ L* := { $Q_j \cdot Q$ / Q ∈ L et d°Q ≤ ½ d − d°Q_j };

{ L := L ∪ L*;

Pour $Q_j \cdot Q$ parcourant L* faire

{ R := Reste de la division de P par $Q_j \cdot Q$;

{ Si R = 0 alors Diviseur := Diviseur ∪ {$Q_j \cdot Q$};

Fin := (j = r) ou (Diviseur non vide);

Résultat:

La liste L de tous les produits $Q_{j1}...Q_{jt}$ calculés.

L'ensemble Diviseur contient des diviseurs de P, mais reste vide si P est irréductible sur Z[X]. Remarquons que si l'ensemble Diviseur n'est pas vide, les polynômes qu'il contient sont des produits $Q_{j1}...Q_{jt}$ minimaux; ce sont donc des diviseurs irréductibles de P. Pour trouver les autres diviseurs de P, on peut donc appliquer à nouveau

l'algorithme ci-dessus au quotient de P par le produit des polynômes de l'ensemble Diviseur. On peut ainsi obtenir la factorisation complète de P.

Etude du Coût.

Pour la recherche de tous les diviseurs de P sur Z[X], le coût est dominé par le calcul des sous-produits de $Q_1 \cdots Q_j \cdots Q_r$ sur $\mathbf{Z}/(p^N)[X]$ avec $p^N > 2^d M(P)$. Dans le cas où le polynôme est irréductible, il y a 2^r produits à calculer, ainsi que 2^r divisions; au total, le coût est de l'ordre de

(A.II.17) $2^r d^2 [d + \text{Log } M(P)]^2$.

Dans le cas d'une factorisation partielle, il suffit de prendre $p^N > 2^{½ d} M(P)$, et il n'y a que 2^{r-1} produits et 2^{r-1} divisions à calculer. Le coût est donc égal à

(A.II.18) $2^r d^2 [½ d + \text{Log } M(P)]^2$.

Le coût est donc exponentiel. Mais le nombre r est "en général" très petit devant d, bien qu'il puisse atteindre la valeur d. Berlekamp signale dans **[Be]**, page 86, que pour d "suffisamment grand", une valeur approximative de r est donnée par log(d). Un résultat précis est donné par M. Mignotte et J.L. Nicolas dans **[MN]**. Pour cela on majore le nombre de polynômes P vérifiant

$$| r - \log(d) | > \lambda \sqrt{\log(d)}.$$

Un majorant vaut $C \lambda^{-2} p^d$ pour tout $\lambda \geq 0$. D'après Mignotte, on peut prendre $C = 4$. Faisons donc cette hypothèse. Remarquons que le nombre total de polynômes P de $\mathbf{Z}/(p)[X]$ ayant un degré majoré par d, est égal à p^d. Si on prend donc $\lambda = \sqrt{8}$, on peut affirmer qu'un polynôme sur deux au moins, vérifie la relation

$$| r - \log(d) | \leq \sqrt{8} \log d.$$

On a donc au moins une fois sur deux $r \leq \log(d) + \sqrt{8} \log d$. Avec cette approximation, le coût usuel est égal à

(I) $d^{2+2\cdot\log(2)} [d + \text{Log } M(P)]^2$.

Essais d'amélioration de la méthode classique.

Signalons aussi une autre approche de l'algorithme classique, visant à diminuer la complexité de la dernière étape. Pour cela on se donne une suite d'entiers premiers $S_p = \{ p_1, .., p_j, ...\}$. Ensuite, on exécute l'algorithme de Berlekamp successivement avec $p = p_1, ..., p_j, ...$; on arrête dès qu'on trouve un nombre de facteurs r "petit" ou lorsque j est supérieur à une certaine valeur. Dans ce dernier cas, on continue les calculs des étapes suivantes en choisissant celui des entiers p_j qui a donné un nombre r minimum de facteurs. Alors r est très souvent égal au nombre exact de diviseurs de P sur $\mathbf{Z}[X]$ et si c'est le cas la dernière étape disparait.

Cependant, Musser signale dans **[Mu]**, p. 301, que pour tout d qui est une puissance de 2, il existe un polynôme irréductible de degré d qui a au moins ½d facteurs modulo p pour tout nombre premier p. La factorisation de ces polynômes a donc un coût qui est réellement exponentiel.

Remarquons aussi que dans le cas général, pour un degré élevé le nombre r de facteurs modulo p est au moins une fois sur deux voisin de log(d). Lorsque d est grand, les différents choix de p dans $S_p = \{ p_1, .., p_j, ...\}$ donnent donc le plus souvent des valeurs de r voisines, même pour un polynôme irréductible sur $\mathbf{Z}[X]$.

Exemple.

$P = X^6 + 2\,X^4 + 4\,X^2 + 3$.

Etape 1:

On choisit $p = 5$. La méthode de Berlekamp donne

$P_o = (X + 1)\,(X - 1)\,(X + 2)\,(X - 2)\,(X^2 + 2)$ sur $\mathbf{F}_5[X]$.

Posons $R = P_o = X^6 - 3\,X^4 - 6\,X^2 + 8$, $R_1 = X + 1$, $R_2 = X - 1$, $R_3 = X + 2$, $R_4 = X - 2$ et $R_5 = X^2 + 1$. On en déduit: $5.S = P - P_o = 5\,X^4 + 10\,X^2 - 5$.

On calcule la décomposition en éléments simples:

$$\frac{X^4 + 2X^2 - 1}{(X+1)(X-1)(X+2)(X-2)(X^2+1)} = \frac{S_1}{X+1} + \frac{S_2}{X-1} + \frac{S_3}{X+2} + \frac{S_4}{X-2} + \frac{S_5}{X^2+2}$$

On obtient:

$R_1 + 5.S_1 = X+1 + 5(-1) = X - 4$;

$R_2 + 5.S_2 = X-1 + 5(+1) = X + 4$;

$R_3 + 5.S_3 = X+2 + 5(+1) = X + 7$;

$R_4 + 5.S_4 = X-2 + 5(-1) = X - 7;$

$R_5 + 5.S_5 = X^2+2 + 5(-2) = X^2 - 8;$

Posons $U = P_1 = (X - 4)(X + 4)(X + 7)(X - 7)(X^2 - 8)$

$= X^6 - 73 X^4 + 54 X^2 - 23,$

$U_1 = X - 4,\ U_2 = X + 4,\ U_3 = X + 7,\ U_4 = X - 7,\ U_5 = X^2 - 8.$

On en déduit: $5.V = P - P_1 = 5^2(3 X^4 - 2 X^2 + 1).$

On calcule la décomposition en éléments simples:

$$\frac{3\ X^4 - 2X^2 + 1}{(X+4)(X-4)(X+7)(X-7)(X^2-8)} = \frac{V_1}{X+4} + \frac{V_2}{X-4} + \frac{V_3}{X+7} + \frac{V_4}{X-7} + \frac{V_5}{X^2-8}$$

On obtient:

$U_1 + 5.V_1 = X+1 + 5(-1) + 5^2(-1) = X - 29;$

$U_2 + 5.V_2 = X-1 + 5(+1) + 5^2(+1) = X + 29;$

$U_3 + 5.V_3 = X+2 + 5(+1) + 5^2(+2) + 5^3(+1) = X + 182;$

$U_4 + 5.V_4 = X-2 + 5(-1) + 5^2(-2) + 5^3(-1) = X - 182;$

$U_5 + 5.V_5 = X^2+2 + 5(-2) + 5^2(-1) + 5^3(+2) = X^2 + 217;$

Doit-on continuer encore la remontée? Supposons qu'on ne cherche qu'une factorisation partielle de P, c'est-à-dire qu'on ne cherche qu'un diviseur de degré majoré par ½d. Les coefficients d'un tel diviseur de P sont majorés par : $B = 2^{d/2} \| P \|$. Dans cet exemple B = 44. On doit donc avoir $p^N > 2 \cdot 44 = 88$, avec $N = 2^n$, ce qui est vérifié avec n = 2.

On construit la liste de tous les sous-produits de

$(X - 29)(X + 29)(X + 182)(X - 182)(X^2 + 217)$ de degrés inférieurs ou égaux à 3. On construit la liste:

$L = \{1, X - 29, X + 29, X + 182, X - 182, X^2 + 217, X^2 - 216, X^2 - 278, X^2 + 278, X^3 - 29X^2 - 43, X^2 + 1, \ldots\}.$

Seul $X^2 + 1$ égal à $(X + 182)(X - 182)$ modulo 5^4, divise P. On a donc trouvé un diviseur de P sur Z[X].

III. ALGORITHME DE LENSTRA.

1°) Présentation de l'algorithme de Lenstra.

L'algorithme de Lenstra tel qu'il est présenté dans **[L3]** se décompose en trois étapes:

Etape 1.

Factorisation du polynôme donné P dans $\mathbf{Z}/(p)$ par la méthode de Berlekamp sous la forme $Q_1^\circ \cdot Q_2^\circ$ où Q_1° est irréductible sur $\mathbf{Z}/(p)$.

Etape 2.

Remontée de la factorisation $Q_1^\circ \cdot Q_2^\circ$ de $\mathbf{Z}/(p)$ jusqu'à $\mathbf{Z}/(p^k)$ avec $p^k \geq B$; la borne B étant nettement supérieure à la borne de la méthode classique, soit $2^d \| P \|$. En effet on doit prendre ici $B = 2^{\frac{1}{2}dq} \| P \|^{d+q}$ si on cherche un diviseur de P sur $\mathbf{Z}[X]$ ayant un degré q. Cette étape 2 est donc nettement plus longue que l'étape 2 de l'algorithme classique.

Etape 3.

Recherche d'un facteur Q de P sur $\mathbf{Z}[X]$, dont la projection Q* sur $\mathbf{F}_p[X]$ soit un multiple de Q_1°. Pour cela, on étudie l'ensemble $\Im$ des multiples de Q_1° sur $\mathbf{Z}/(p^k)$ dont le degré est inférieur à d. Si P n'est pas irréductible sur $\mathbf{Z}[X]$, alors l'un de ses diviseurs P_1 appartient à cet ensemble $\Im$. En fait si $\Im$ contient un tel diviseur P_1 de P, alors P_1 est l'élément de $\Im$ dont la norme $\| P_1 \|$ est minimum. L'étape 3 consiste donc à rechercher un élément de $\Im$ de longueur suffisamment petite. Le calcul de cet élément petit de $\Im$ a un coût polynômial, alors que dans l'algorithme classique, l'étape 3 a un coût exponentiel.

2°) Définition des bases réduites.

Une définition géométrique et très complète des bases réduites est donnée dans **[Va]**. Ce qui suit a pour but d'introduire la notion de bases réduites pour la recherche d'un diviseur de P sur $\mathbf{Z}[X]$, puis de donner les grandes lignes de l'algorithme de

construction de Lenstra et enfin de rappeler les résultats sur les coûts. Les polynômes de $\mathbf{Z}[X]$ de degré majoré par d = deg P, forment un module sur $\mathbf{Z}$. En particulier l'ensemble noté $\Im$ des multiples de Q définis modulo p^k est un sous-module de dimension d. Une base évidente de ce sous- module est :

$B = \{Q, X\cdot Q, .., X^{d-q-1}Q, p^k, p^kX, .., p^kX^{q-1}\}$, avec q = deg Q.

Le problème posé est de trouver un élément de ce sous-module dont la norme est "petite".

L'idée est de transformer cette base en une base formée de "petits" vecteurs. Pour cela considérons une base quelconque $B = \{V_1,...,V_d\}$; le volume engendré par le parallélépipède défini avec les vecteurs de cette base est égal au déterminant $\Delta(\Im) = | V_1 \cdots V_d |$. Si la base est orthogonale, alors ce volume est égal au produit $\| V_1 \| \cdots \| V_d \|$. Dans le cas d'une base quelconque non orthogonale, on a la relation suivante, appelée inégalité de Hadamard

(A.III.1) $$\| V_1 \| \cdots \| V_d \| \geq \Delta(\Im).$$

Il en résulte que dans une base orthogonale, le produit des normes des vecteurs de la base est minimum. Si de plus on trie les vecteurs de cette base de façon que leurs normes soient croissantes, alors on est assuré que le premier vecteur de la base orthogonale a une longueur "assez petite". Pour obtenir de "petits" vecteurs de $\Im$, on pourrait donc penser à construire une base orthogonale.

La méthode habituellement utilisée pour construire une base orthogonale est appelée méthode d'orthogonalisation de Schmidt; elle définit les vecteurs orthogonaux par récurrence à l'aide des formules :

(A.III.2) $$\begin{cases} T_1 = V_1 \\ T_j = V_j - \sum_{0\leq k\leq j-1} \lambda_{j\,k} T_k \quad \text{où} \quad \lambda_{j\,k} = \dfrac{\langle V_j, T_k \rangle}{\| T_k \|^2} \end{cases}$$

Mais les vecteurs T_j ainsi construits n'ont aucune raison d'appartenir à $\Im$. En général le sous-module $\Im$ n'a pas de base orthogonale. On peut cependant construire une base $\{R_j\}$ qui soit "proche" de la base orthogonale $\{T_j\}$. Une telle base est appelée base réduite. De plus on voudrait que les normes des vecteurs de la base réduite soient si

possible "croissantes" ou au moins pas trop décroissantes. Notons $\mathfrak{I}_j$ le sous-module engendré par les j premiers vecteurs $R_1, .., R_j$ de la base construite.

Définition.

Notons $\{T_k\}$ la base orthogonale définie ci-dessus par la méthode d'orthogonalisation de Schmidt à l'aide des formules **(A.III.2)**. On appelle **base réduite** de $\mathfrak{I}$ une base $\{R_j\}$ vérifiant les relations suivantes

(A.III.3) $$\lambda_{j\,k} = \frac{\langle R_j, T_k \rangle}{\| T_k \|^2} \leq \tfrac{1}{2} \quad \text{où} \quad 1 \leq k \leq j \leq d$$

(A.III.4) $$\| p^\circ_j \|^2 \geq \tfrac{3}{4} \| p_{j-1} \|^2 \quad \text{pour} \quad 1 \leq j \leq d,$$

où les vecteurs p°_j et p_{j-1} sont égaux respectivement à $T_j + \lambda_{j\,j-1} T_{j-1}$ et à T_{j-1}, et sont les projections de R_j et R_{j-1} sur l'espace orthogonal à $\mathfrak{I}_{j-2}$.
Dans la suite, on exprimera l'inégalité **(A.III.3)** en disant que les vecteurs R_j sont pseudo-orthogonaux et on exprimera **(A.III.4)** en disant que les vecteurs R_j sont pseudo-croissants.

Vérifions maintenant que la base réduite $\{R_1, R_2, ..., R_j, ...\}$ contient effectivement des vecteurs "petits". On a plus précisément la relation suivante :

(A.III.5) $\| R_1 \|^2 \leq 2^{d-1} \| U \|^2$ pour tout vecteur U de $\mathfrak{I}$.

Nous démontrons ci-dessous cette inégalité de façon mettre en évidence l'importance des deux propriétés de la base réduite, à savoir la pseudo-croissance (relation **(A.III.4)**) et la pseudo-orthogonalité (relation **(A.III.3)**). Cette démonstration figue dans **[L3]** sous une autre forme.
Comme les vecteurs T_j sont orthogonaux, les relations **(A.III.3)** et **(A.III.4)** impliquent que $\| T_j \|^2 \geq (\tfrac{3}{4} - (\lambda_{j\,j-1})^2) \| T_{j-1} \|^2 \geq \tfrac{1}{2} \| T_{j-1} \|^2$. On en déduit par récurrence que $\| T_k \|^2 \leq 2^{j-k} \| T_j \|^2$ pour $k \leq j$. Il résulte des formules d'orthogonalisation **(A.III.2)** que $\| R_j \|^2 = \| T_j \|^2 + \sum_{0 \leq k \leq j-1} (\lambda_{j\,k})^2 \| T_k \|^2$ est

majoré par $\| T_j \|^2 + \sum_{0 \le k \le j-1} ¼\, 2^{j-k} \| T_j \|^2$, donc par $2^{j-1} \| T_j \|^2$, ou encore par $2^{k-1} \| T_k \|^2$ pour $k \ge j$.

Si $j = 1$, on obtient $\| R_1 \|^2 \le 2^{k-1} \| T_k \|^2$ pour tout $k \ge 1$. D'autre part on a

$U = \sum_{0 \le k \le d} \lambda_k R_k = \sum_{0 \le k \le d} \mu_k T_k$ avec $\lambda_k, \mu_k \in \mathbf{Z}$.

D'après (A.III.2), si k est le plus grand entier tel que $\mu_k \neq 0$, alors $\mu_k = \sigma_k$; par suite $\| U \|^2 \ge \sigma_k^2 \| T_k \|^2 \ge \| T_k \|^2$ pour un entier $k \ge 1$. Comme $\| R_1 \|^2 \le 2^{k-1} \| T_k \|^2$, on obtient $\| R_1 \|^2 \le 2^{d-1} \| U \|^2$. L'inégalité est donc bien vérifiée.

3°) Construction d'une base réduite.

Considérons une base $B = \{V_1,...,V_d\}$ de $\mathfrak{I}$, et notons $\{T_1,...,T_d\}$ la base orthogonale associée. On construit de façon récursive une base $B_r^j = \{R_1,...,R_j, V_{j+1},...,V_d\}$ telle que $\{R_1,...,R_j\}$ soit une sous-base réduite, c'est-à-dire pseudo-orthogonale et pseudo-croissante. Pour $j=1$, $B_r^1 = \{R_1\} = \{V_1\}$. Supposons que la base B_r^{j-1} soit déjà construite, et déterminons B_r^k à partir de B_r^j.

Rendre la base pseudo-croissante.

Considérons les deux vecteurs $p°_j = T_j + \lambda_{j\,j-1} T_{j-1} = V_j - \sum_{0 \le k \le j-2} \lambda_{j\,k} T_k$ et

$p_{j-1} = T_{j-1} = V_{j-1} - \sum_{0 \le k \le j-2} \lambda_{j\,k} T_k$. On a soit, $\| p°_j \|^2 \ge ¾ \| p_{j-1} \|^2$, soit

$\| p_{j-1} \|^2 > ¾ \| p°_j \|^2$. Si la première inégalité n'est pas satisfaite, on échange les vecteurs V_j et V_{j-1} dans la base donnée. Alors, les vecteurs T_j et T_{j-1} sont également échangés, et il en est de même des vecteurs $p°_j$ et p_{j-1}. Par suite, la sous-base $B_r^j = \{R_1,...,R_{j-2}, V_j, R_{j-1}\}$ est pseudo-croissante; on l'écrit comme ci-dessus sous la forme $B_r^j = \{R_1,...,R_{j-1}, V_j\}$.

Rendre la base pseudo-orthogonale.

Pour tout $k \in [1,j-2]$, transformons le vecteur V_j en $V_j - r R_k$ avec r égal à l'entier le plus proche de $| \lambda_{j\,k} |$. Les nouveaux coefficients $\lambda_{j\,m}$ seront égaux à

$\frac{\langle V_j - r R_k, T_m \rangle}{\| T_m \|^2} = \lambda_{j\,m} - r \lambda_{k\,m}$. Pour $m = k$, on a $\lambda_{k\,k} = 1$, donc le nouveau coefficient

$\lambda_{j\,k}$ sera égal à $\lambda_{j\,k} - r$, donc vérifiera $|\lambda_{j\,k}| \leq ½$. D'autre part, remarquons que les nouveaux vecteurs $V_j - r\,R_k$ ont une projection inchangée sur le sous-espace orthogonal à $\{V_1,...,V_{j-1}\}$, donc T_j reste inchangé et la base reste pseudo-croissante. Par contre, la transformation analogue pour $k = j-1$, rendant $|\lambda_{j\,j-1}| \leq ½$ modifie T_j et la base risque de ne plus être pseudo-croissante. La transformation correspondante de V_j en $V_j - r\,R_{j-1}$ sera donc exécutée en premier, avant de rendre la base pseudo-croissante.

Lenstra démontre dans **[L3]** le résultat suivant :

Lemme 1.

Si les vecteurs de la base donnée vérifie $|V_k| \leq B$, alors la complexité de l'algorithme de construction de la base réduite est de l'ordre de $d^4\,\mathrm{Log}(B)$ multiplications portant sur des entiers de grandeur $d\,\mathrm{Log}(B)$.

Programme REDUCE du calcul d'une base réduite.

Le programme récursif suit de très près l'algorithme de Lenstra. L'intérêt d'un tel programme est de mettre en évidence la structure de l'algorithme; la base donnée est formée des vecteurs R(0),...,R(n) qui seront transformés de façon à obtenir une base réduite. On note comme dans l'algorithme T(0),...,T(n) la base orthogonale associée. On utilise la procédure Affichebase(n) qui affiche les n premiers vecteurs de la base réduite et la procédure Reduis(n,a,j) qui transforme le vecteur R(n) en R(n) - r R(j) où r est la partie entière de a, de façon que la base soit pseudo-orthogonale.

```
Procedure Bred(n);
Begin  Scalar k, j, m, q, c, l;

            % Initialisation :
   If  n > 0  then   << T(0) := R(0);  Bred(n-1);  Affichebase(n);  T(n) := R(n);

            % Orthogonalisation de la base  R(0),...,R(n) :
            For j:= 0:n-1   do  << Alpha(j) := Prod(R(n),T(j))/Norme(T(j));
                                   Poj := T(n);  T(n) := Poj - Alpha(j)×T(j)  >>;

            % Vérification de la pseudo-orthogonalisation des deux derniers vecteurs:
            If Abs(Alpha(n-1)) > 1/2  then  << R(n) := Reduit(n,Alpha(n-1),n-1);  Bred(n) >>
```

```
                    else
% Vérification de la pseudo-croissance :
 << If Norme(Poj) < 3×Norme(T(n-1))/4  then   << Echange(n,n-1);  Bred(n) >>
                          else      << j := n-2;

    % Si la base est pseudo-croissante, on examine si le vecteur  R(n)  est
% pseudo-orthogonal aux vecteurs précédents :

    While j >= 0 and Abs(Alpha(j))<=1/2  do  j := j - 1;
    If   j >= 0 then  << R(n) := Reduit(n,Alpha(j),j); Bred(n) >> >>
     >>          >>    End;
```

Exemple:

Considérons la base suivante :

B(0) := 625; B(1) := 244 + X; B(2) := 244×X + X**2;

L'exécution du programme REDUCE ci-dessus donne les 24 bases suivantes :

Base N° 1	Base N° 2	Base N° 3	Base N° 4	Base N° 5
{625,0}	{244,1}	{ 244, 1}	{-107,-3}	{-107, -3}
{244,1}	{625,0}	{-107,-3}	{ 244, 1}	{ 30, -5}

Les calculs concernant les quatre premières bases, consistent en échanger les deux vecteurs pour les ordonner suivant les modules croissants, et à retrancher du deuxième vecteur, r fois le premier vecteur pour que les deux vecteurs soient pseudo-orthgonaux.

Base N° 6	Base N° 7	Base N° 8	Base N° 9	Base N°10
{ 30,-5}	{ 30, -5}	{-17,-18}	{-17,-18}	{-17,-18}
{-107,-3}	{-17,-18}	{ 30, -5}	{ 13,-23}	{ 13,-23}

Base N° 11	Base N° 12	Base N° 13	Base N° 14	Base N° 15
{-17,-18}	{-17,-18}	{-17,-18}	{-17,-18}	{6,-7, 1}
{ 13,-23}	{ 13,-23}	{91,83,1}	{6,-7, 1}	{-17,-18}
{ 91,83,1}			{ 13,-23}	

Base N° 16	Base N°17	Base N° 18	Base N° 19	Base N° 20
{6,-7, 1}	{ 6,-7, 1}	{6,-7,1}	{-5,-2,-3}	{-5,-2,-3}
{-17,-18}	{ 13, -23}	{-5,-2,-3}	{6,-7,1}	{6,-7,1}
				{-17,-18}

Base N° 21	Base N° 22	Base N° 23	Base N° 24
{−5,−2,−3}	{−5,−2,−3}	{−5,−2,−3}	{−5,−2,−3}
{6, −7, 1}	{ 6,−7, 1}	{ 6,−7, 1}	{ 6,−7, 1}
	{−23,−11,−1}		{−3,−3,11}

4°) Diviseur de P défini par la base réduite.

Soit Q un diviseur de P sur $\mathbf{Z}/(p^k)$. Notons $B_r^k = \{R_1,...,R_d\}$ une base réduite du module $\Im$ des multiples de Q de degré majoré par le degré d de P. En fait, si la norme de R_1 est suffisamment petite par rapport à p^k, alors R_1 n'est pas premier avec P sur Z[X]. Cela résulte du Lemme suivant:

Lemme 2.

Supposons que P ne soit pas divisible par X. Soit Q un diviseur de P sur $\mathbf{Z}/(p^k)[X]$ de degré q. Notons $\Im$ le module formé des polynômes de $\mathbf{Z}/(p^k)[X]$ multiples de Q de degré inférieur au degré d de P. Soit R_1 le premier polynôme d'une base réduite de $\Im$. Supposons qu'on ait

(A.III.6) $$p^{kq} \geq 2^{d(d-1)/2} \binom{2d}{d}^{d/2} \| P \|^{2d-1}.$$

Alors R_1 n'est pas premier avec P sur Z[X] si et seulement si

(A.III.7) $$\| R_1 \|^d \leq \frac{p^{kq}}{\| P \|^{d-1}}.$$

La démonstration de cette propriété est donnée dans **[L3]** où elle fait l'objet de plusieurs lemmes; bien que la propriété de R_1 donnée dans ce Lemme est plus faible (R_1 non premier avec P) que celle donnée dans **[L3]** (R_1 divise P), la démonstration est plus instructive et elle figure dans une version antérieure de **[L3]**. Cette démonstration est également intéressante, car elle est le fondement de la méthode de factorisation de Lenstra et elle montre l'origine de la complexité de la minoration **(A.III.6)**; or cette minoration **(A.III.6)** est la cause de la grandeur du coût de cette méthode de factorisation.

Démonstration.

Démonstration.

Supposons que les relations **(A.III.6)** et **(A.III.7)** soient vérifiées, mais que R_1 et P soient premiers entre eux sur $\mathbf{Z}[X]$.

Supposons qu'il existe une combinaison linéaire

$\lambda_o P + \lambda_1 X\cdot P + \ldots + \lambda_{d-1} X^{d-1} P - \mu_o X^d R_1 - \mu_1 X^{d+1} R_1 - \ldots - \mu_{d-1} X^{2d-2} R_1 = 0.$

Alors P diviserait $(\mu_o + \mu_1 X + \ldots - \mu_{d-1} X^{d-1}) X^d R^1$; donc P ne serait pas premier avec R_1. Par suite, la famille $B^\circ = \{P, X\cdot P, .., X^{d-1}P, X^d R_1, X^{d+1} R_1, .., X^{2d-2} R_1\}$ est libre. C'est donc une base du module $\Im^\circ$, formée des polynômes de degré $\leq 2d-2$, qui sont combinaisons linéaires de P et R_1. Il est clair que $\Im^\circ$ est un sous-module de $\Im$. Donc tout polynôme $X^j P$ ou $X^k R_1$ de B° s'écrit sous la forme $\sum_{n\geq 1} \lambda_{j\,n}\, b_n$ où $B = \{b_1, .., b_n, ..\}$ est une base de $\Im$. Par suite le déterminant $\Delta(B^\circ)$ de la base B° est égal à $\left(\sum_j \sum_n \lambda_{j\,n}\right) D(B)$ où $\left(\sum_j \sum_n \lambda_{j\,n}\right)$ est un entier non nuls; ainsi

$\Delta(B^\circ) \geq \Delta(B)$. D'autre part l'inégalité de Hadamard implique que

$\Delta(B^\circ) \leq \| P \|^d \| R_1 \|^{d-1}$. Enfin, on peut calculer facilement $\Delta(B)$ en rendant orthogonale la base $\{Q, X\cdot Q, .., X^{d-q-1} Q, p^k, p^k X, .., p^k X^{q-1}\}$; comme les polynômes $X^j Q$ sont unitaires on obtient la base $\{X^q, X^{q+1}, .., X^d, p^k, ..., p^k X^{q-1}\}$. Donc $\Delta(B) = p^{kq}$. Par suite $p^{kq} \leq \Delta(B^\circ) \leq \| P \|^d \| R^1 \|^{d-1}$, ce qui est contraire à **(A.III.7)**.

Il reste à démontrer la réciproque en supposant maintenant que R^1 ne soit pas premier avec P. Notons $Q_o = \sum_j b_j X^j$ un polynôme de degré m, qui soit un diviseur commun à R_1 et P sur $\mathbf{Z}[X]$. D'après la majoration de Landau-Mignotte, on a:

$|\, b_j \,| \leq \binom{m}{j} \| P \|$. Par suite $\| Q_o \|^2 \leq \binom{2m}{m} \| P \|^2$, donc

$\| Q_o \|^2 \leq \binom{2d}{d} \| P \|^2$. D'après le Lemme 1, $\| R_1 \|^2 \leq 2^{d-1} \| Q_o \|^2$. Il en résulte :

$\| R_1 \|^{2d} \leq 2^{d(d-1)} \binom{2d}{d}^d \| P \|^{2d}$; par hypothèse la minoration **(A.III.6)** est vérifiée, donc la majoration **(A.III.7)** est vérifiée; par suite le Lemme est démontré.

5°) Description de l'algorithme de Lenstra.

L'algorithme de factorisation résulte directement du Lemme 2. L'étape 2 de l'algorithme classique est effectuée avec p^k suffisamment grand pour que la minoration **(A.III.6)** ci-dessus soit vérifiée, c'est-à- dire avec

(A.III.8) $p^k \geq 2^{d(d-1)/2q} \binom{2d}{d}^{d/2q} \| P \|^{(2d-1)/q}$.

L'étape suivante consiste à choisir un diviseur Q de P modulo p^k, qui soit irréductible, puis à construire une base réduite du sous-module engendré par

$\{Q, X\cdot Q, .., X^{d-q-1} Q, p^k, p^k X, .., p^k X^{q-1}\}$.

Ensuite, on regarde si le premier polynôme R_1 de cette base réduite a sa norme qui est majorée par la relation **(A.III.7)**; si c'est le cas, alors R_1 n'est pas premier avec P, et on obtient un diviseur de P en calculant le PGCD de R_1 et P. Si la norme de R_1 ne vérifie pas la majoration **(A.III.7)**, alors P est irréductible sur Z[X].

6°) Coût de l'algorithme.

Les calculs doivent être effectués modulo p^N et d'après la relation **(A.III.8)**, on est obligé de choisir p^N de l'ordre de $2^{d(d-1)/2q} \binom{2d}{d}^{d/2q} \| P \|^{(2d-1)/q}$.

D'après le Lemme 1, la réduction de la base exige d'effectuer $d^4 \text{Log}(p^N)$ multiplications portant sur des entiers de taille $d\cdot\text{Log}(p^N)$. Le coût est donc de l'ordre de $d^6 \text{Log}^3(p^N)$. Or p^N est borné par l'inégalité $\text{Log}(p^N) \geq (d/q)^2 + (d/q)\text{Log} \| P \|$. On obtient donc comme coût :

(A.III.9) $d^9/q^3 (d^3/q^3 + \text{Log}^3 \| P \|)$.

Le coût est maximum lorsque $q = 1$ et on obtient alors

(A.III.10) $d^{12} + d^9\text{Log}^3 \| P \|$.

IV. NOUVELLES METHODES DE FACTORISATION.

a) Factorisation partielle obtenue à partir d'un facteur linéaire.

La méthode proposée ci-dessous est originale; elle figure dans **[V4]**, et c'est sans doute la seule méthode rapide de factorisation n'utilisant pas l'algorithme de Berlekamp. Cette méthode consiste à chercher un seul facteur sur $\mathbf{Z}/(p)[X]$, puis sur $\mathbf{Z}/(p^m)[X]$ à l'aide de la formule d'approximation des racines de Newton. Ainsi les deux premières étapes de l'algorithme classique se réduisent à des calculs très simples.
Cependant la majoration de p^m, est celle de l'algorithme de Lenstra. Et l'étape 3 exige le calcul d'une base réduite et le coût de cette dernière étape est celui de l'algorithme de Lenstra.

1°) Introduction.

L'un des inconvénients de l'algorithme de Lenstra est qu'il exige des calculs très longs au cours de l'étape 2, pour trouver la factorisation du polynôme donné P sur $\mathbf{Z}/(p^k)$. Or au cours de l'étape 3, un seul des diviseurs trouvé est utilisé.
La méthode proposée ici consiste à ne calculer qu'un seul diviseur Q de P sur $\mathbf{Z}/(p^k)$. Le diviseur calculé est de degré 1, ce qui rend les calculs beaucoup plus rapides que ceux de l'étape 2 de l'algorithme classique.

2°) Calcul d'une racine simple de P sur Z/(*p*)[X].

On considère un polynôme $P = a_0X^d + ... + a_{d-1}X + a_d$ et on veut trouver un diviseur de P sur Z[X].

Dans toute la suite, on suppose que P n'a pas de facteur carré; si ce n'était pas le cas on trouverait facilement un diviseur de P en calculant le PGCD de P et de sa dérivée P'. Considérons le résultant de P et P', noté Res(P,P') :

(A.IV.1) $\text{Res}(P,P') = U P + V P' = \delta \in Z.$

La méthode de Berlekamp consiste à choisir un nombre premier *p*, puis à calculer les diviseurs de P sur Z/(*p*)[X]. La méthode proposée ici consiste à choisir un nombre α_0 comme racine de P, puis à déterminer un nombre premier *p* pour que α_0 soit racine de P sur Z/(*p*)[X]. On obtient ainsi un facteur linéaire de P sur Z/(*p*)[X]. Mais comme dans la méthode de Berlekamp, le carré de ce facteur ne doit pas diviser P sur Z/(*p*)[X]. Il faut donc que α_0 soit une racine simple de P sur Z/(*p*)[X]. Le lemme ci-dessous montre qu'il existe toujours une racine simple, pour un certain entier premier *p*.

Lemme 1.

Considérons un polynome de degré d à coefficients dans Z, sans facteur carré, noté $P = a_0X^d + ... + a_{d-1}X + a_d$.

Il existe un nombre premier $p \geq 2$, ne divisant pas a_0 et tel que P admette sur Z/(*p*) une racine simple α_0.

Démonstration.

Considérons les nombres $\alpha = \lambda a_0 a_d \delta$ avec $\lambda \in Z$ et $\delta = \text{RES}(P,P')$. Alors $P(\alpha)$ est égal à

$A = a_d(\lambda^d\delta^d(a_0)^{d+1}(a_d)^{d-1} + \cdots + a_{d-1}\lambda a_0 a_d \delta + 1) = a_d(\lambda a_0 \delta B + 1).$

Remarquons que $\lambda a_0 \delta B + 1$ est premier avec δ et a_d pour toute valeur de λ. Comme P est de degré d, il y a au plus d valeurs de λ qui rendent $P(\alpha)/a_d$ égal à 1 ou à −1. Par suite, il existe une valeur λ_0 de λ comprise entre −d et d, telle

que $P(\alpha)/a_d$ contienne un facteur premier p supérieur ou égal à 2. Notons α_0 le produit $\lambda_0\, a_d\, a_0\, \delta$. Alors α_0 est une racine de P sur $Z/(p)$. Comme p divise $P(\alpha)/a_d = \lambda\, a_0\, \delta\, B + 1$, il est premier avec δ et avec a_0.

Rappelons que d'après **(A.IV.1)**, on a $U{\cdot}P(\alpha_0) + V{\cdot}P'(\alpha_0) = \delta$. Donc p ne peut pas diviser $P'(\alpha_0)$. D'autre part , il est premier avec a_d. Ainsi α_0 est une racine simple de P qui vérifie le Lemme.

Dans la méthode classique de factorisation, on doit trouver un nombre premier p qui ne divise pas δ, ni a_d. Mais on évite de calculer δ. On prend une suite de nombres premiers p non diviseurs de a_d et on calcule le PGCD de P et P' sur $Z/(p)[X]$. On choisit le premier nombre p pour lequel ce PGCD est un polynôme constant non nul. Comme le résultant n'a qu'un nombre limité de facteurs premiers, on est assuré de trouver un tel nombre p au bout d'un nombre fini d'essais.

A chaque essai le coût est celui du calcul du PGCD de deux polynômes; d'après **(A.I.3)** il est de l'ordre de $d^2 \operatorname{Log} p$.

De plus le coût de la factorisation de P par la méthode de Berlekamp est de l'ordre de $p\, d^3 \operatorname{Log}^2 p$.

Dans la méthode proposée ici, on recherche une racine simple de P modulo un nombre premier p. Pour cela on essaie différentes valeurs α_0 pour cette racine; on prend la suite des nombres 0, 1, −1, ...,10, −10, ... Pour chacun d'eux on calcule $P(\alpha_0)$ et $P'(\alpha_0)$, puis le PGCD δ de $P(\alpha_0)$, $P'(\alpha_0)$ et a_d sur Z. Dès qu'on trouve une valeur α_0 telle que $P(\alpha_0)/\delta$ soit distinct de 1 et −1, alors on choisit p égal à l'un des facteurs premiers de $P(\alpha_0)/\delta$.

D'après le lemme, on est assuré de trouver une valeur α_0 qui convienne en un nombre fini d'étapes.

Etudions le coût de l'une de ces étapes. Le coût est dominé par le calcul de $P(\alpha_0)$ et $P'(\alpha_0)$. Si on utilise la méthode de Hörner, les calculs reviennent à effectuer d multiplications de α_0 par des entiers de taille $H(P)\,\alpha_0$, $H(P)\,\alpha_0^2$, ...,$H(P)\,\alpha_0^d$. Le temps du calcul de $P(\alpha_0)$ est donc majoré par $d^2 \operatorname{Log}^2 \alpha_0 + d \operatorname{Log} \alpha_0 \operatorname{Log} H(P)$; (on suppose pour simplifier que α_0 est positif).

De même que dans l'algorithme classique de Berlekamp, le nombre fini d'étapes est très souvent réduit à un ou deux. Dans le cas où le nombre d'étapes serait important, la méthode proposée devrait être abandonnée au profit de la méthode de Berlekamp.

3°) Remontée de la racine α_0 de $Z/(p)[X]$ sur $Z/(p^n)[X]$.

a) Formule de Newton.

Soit α la racine de P sur l'anneau Z_p des entiers p–adiques, dont la projection sur $Z/(p)$ est α_0. Soit Q le quotient de P par $X - \alpha$ sur Z_p. Notons respectivement α_k, P_k et Q_k les projections de α, P et Q sur $Z/(p^q)$, où $q = 2^k$. On a donc la relation suivante sur $Z/(p^q)$:

(A.IV.4) $\quad P_k = (X - \alpha_k)\, Q_k$

On a les relations suivantes sur $Z/(p^{2q})$:

(A.IV.5) $\quad \begin{cases} P_{k+1} = P_k + q\, P^\circ_k \\ Q_{k+1} = Q_k + q\, Q^\circ_k \\ \alpha_{k+1} = \alpha_k + q\, \alpha^\circ_k \end{cases}$

Le but de ce paragraphe est de définir α_{k+1} en fonction de α_k. On peut écrire l'égalité suivante sur $Z/(p^{2q})$:

(A.IV.6) $\quad P_k + q\, P^\circ_k = (X - \alpha_k - q\alpha^\circ_k)\,(Q_k + q\, Q^\circ_k).$

On en déduit la relation suivante sur $Z/(p^q)$:

(A.IV.7) $\quad P^\circ_k = (X - \alpha_k)\, Q^\circ_k - \alpha^\circ_k\, Q_k.$

On remplace X par α^k dans **(A.IV.7)**, et on remarque que $Q_k(\alpha_k) = P'_k(\alpha_k)$ où P'_k est la dérivée de P_k. On obtient $-\alpha^\circ_k\, P'_k(\alpha_k) = P^\circ_k(\alpha_k) = P_k(\alpha_k)$.
On en déduit, en utilisant **(A.IV.5)** la formule de Newton :

(A.IV.8) $\quad \alpha_{k+1} = \alpha_k - \dfrac{P_k(\alpha_k)}{P'_k(\alpha_k)}\, p^q .$

où α_{k+1} est défini modulo p^{2q} avec $q = 2^{k+1}$. Remarquons que le quotient du second membre doit être défini modulo p^q.

b) Coût des calculs.

Le calcul de α_{k+1} en fonction de α_k exige le calcul de $P(\alpha_k)$ et de $P'(\alpha_k)$ modulo p^k, puis le calcul du quotient. Ce calcul revient à calculer 2 d multiplications et additions et une division.

La taille des coefficients est majorée par p^k; or à chaque étape, p^k est élevé au carré. Le coût est donc majoré par deux fois les calculs de la dernière étape. Cette dernière étape doit définir α_{k+1} modulo p^m, où p^m est majoré par la relation **(A.II.8)**, dans laquelle on a remplacé q par 1, à savoir $p^m \geq 2^{d(d-1)/2} \binom{2d}{d}^{d/2} \| P \|^{2d-1}$.

Pour obtenir α_{k+1} modulo p^m il suffit d'effectuer les calculs modulo $p^{m/2}$. Le nombre de chiffres des nombres que nous avons à multiplier est donc de l'ordre de $\mathrm{Log}(p^{m/2}) \geq \frac{1}{4} d(d-1) + \frac{1}{2} d^2 + (d - \frac{1}{2}) \mathrm{Log} \| P \|$, soit encore $d^2 + d \mathrm{Log} \| P \|$. Les multiplications correspondant au calcul des coefficients α_k ont donc un coût qui est de l'ordre de

(A.IV.9) $\quad \mathrm{Max}(d^4, d^2 \mathrm{Log}^2 \| P \|)$.

4°) Recherche d'un multiple de $X - \alpha_k$ qui divise P sur Z[X].

Dans la suite, on suppose que tous les coefficients sont écrits à l'aide de nombres compris entre $-\frac{1}{2}p^m$ et $\frac{1}{2}p^m$.

Notons $Q = (X - \alpha_k) Q_0 = c_0 X^q + .. + c_k X^{q-k} + .. + c_q$ un multiple de $X - \alpha_k$ de degré q. Puisque Q divise P sur Z[X], alors on peut utiliser la borne de Landau–Mignotte, à savoir :

(A.IV.10) $\quad | c_k | \leq \binom{q}{k} \| P \|$.

On en déduit que $\| Q \|^2 = \sum_k | c_k |^2 \leq \sum_k \binom{q}{k}^2 \| P \|^2 = \binom{2q}{q} \| P \|^2$, puisque $\sum_k \binom{d}{k} \binom{q}{q-k}$ est le coefficient de X^q dans le produit $(1 + X)^q (1 + X)^q$, donc il est

égal à $\binom{2q}{q}$. Ainsi

(A.IV.11) $\quad \| Q \| \leq \binom{2q}{q}^{1/2} \| P \| := B.$

Il en résulte que si nous choisissons $p^m > 2\,B$, alors les coefficients de Q définis modulo p^m seront égaux aux coefficients de Q sur Z. Cependant la condition **(A.IV.10)** est nécessaire, mais non suffisante pour que Q divise P sur Z.
Pour obtenir une condition suffisante, nous devons prendre une plus grande valeur pour p^m. Comme nous ne connaissons pas le degré du diviseur $Q = (X - \alpha_k)\,Q_0$ de P, nous devons prendre $q = d-1$. Ecrivons la relation de la division de Q par $X - \alpha_k$ sur Z[X] :

$$Q = (X - \alpha_k)\,(y_q X^{q-1} + \ldots + y_2 X + y_1) + Q(\alpha_k)$$

où $Q(\alpha_k) = y_0 p^m$ puisque α_k est une racine de Q modulo p^m. Par suite :

$$Q = y_q(X^q - \alpha_k X^{q-1}) + \ldots + y_1(X - \alpha_k) + y_0 p^m$$

appartient à l'idéal engendré par la base $\Pi = \{ X^q - \alpha_k X_{q-1}, \ldots, X - \alpha_k, p^m \}$. Notons $\Im_\Pi$ cet idéal. Dans l'algorithme de Lenstra défini en **II paragraphe 5**, on doit transformer la base Π en une base réduite notée

$$\Pi_r = \{ \pi_0, \pi_1, \ldots, \pi_q \}$$

formée de polynômes π_j relativement petits. Plus précisément, d'après **(A.II.5)**, on doit avoir

(A.IV.12) $\quad \| \pi_0 \| \leq 2^{q/2} \| Q \|$

pour tout polynôme Q de l'idéal $\Im_\Pi$. En fait, on ne cherche pas directement Q, mais un polynôme R qui n'est pas premier avec P sur Z[X]. Lorsqu'on obtient un tel polynôme R, il reste à calculer le PGCD de P et de R. On a vu dans **II, paragraphe 4, Lemme 3** que P et R ne sont pas premiers entre eux à condition que

(A.IV.13) $\quad \| R \|^d \| P \|^{d-1} \leq p^m.$

Si on choisit comme polynôme R le premier polynôme π_0 de la base réduite, alors on déduit de **(A.IV.11)** et de **(A.IV.12)** que

(A.IV.14) $\quad \| R \| \leq 2^{3(d-1)/2} \| P \|.$

Il en résulte d'après la relation **(A.IV.13)** qu'on doit choisir

(A.IV.15) $p^m \geq 2^{3d(d-1)/2} \| P \|^{2d-1}$.

Supposons donc que p^m soit supérieur à cette borne; alors la connaissance de π_0 nous permet de conclure :

Si $R = \pi_0$ vérifie la majoration (A.IV.14), alors R et P ne sont pas premiers sur Z[X] et le calcul du PGCD de P et R nous donnera un facteur de P sur Z[X];

Si au contraire $R = \pi_0$ ne vérifie pas la majoration (A.IV.14), alors P est irréductible.

Remarquons que rien ne permet d'affirmer que le facteur obtenu, à savoir le PGCD de R et P soit irréductible.

Etudions la complexité de cette dernière étape de l'algorithme. Le temps de calcul vient du calcul de la base réduite. D'après le **Lemme 2** de **A.II.4** ce temps de calcul vaut $(d^4 \text{ Log} \| Q \|)\ (d \text{ Log} \| Q \|)^2$ où d'après **(A.IV.15)**, $\text{Log} \| Q \|$ est majoré par $3d(d-1)/2 + (2d-1)\text{Log} \| P \|$. On obtient donc comme coût de cette étape $d^6 \text{ Log}^3 \| Q \| = d^6(3d(d-1)/2 + (2d-1)\text{Log} \| P \|)^3$, ce qui est de l'ordre de

(A.IV.16) $\text{Max}\ (d^{12}, d^9 \text{Log}^3 \| P \|)$.

5°) Exemple.

Considérons le polynôme $P = 6X^5 - 8X^4 + 6X^3 + 9X^2 - 5X - 3$.

a) <u>Recherche d'une racine simple *p*-adique de P.</u>

On calcule $P' = 30X^4 - 32X^3 + 18X^2 + 18X - 5$ et on trouve que $P(1) = 5$, $P'(1) = 29$. Donc on prend $\alpha_1 = 1$ modulo $p = 5$. On calcule à l'aide de la formule de Newton

$$\alpha_2 = \alpha_1 - P(\alpha_1)/P'(\alpha_1) = 1 - 10/29 = 6 \text{ modulo } 5^2,$$

$$\alpha_3 = \alpha_2 - P(\alpha_2)/P'(\alpha_2) = 6 - 250/219 = -244 \text{ modulo } 5^4.$$

Donc $\alpha_3 = -244$ est une racine simple 5-adique de P définie modulo $5^4 = 625$.

b) <u>Recherche d'un diviseur de P qui soit un multiple de X + 244.</u>

Cherchons un diviseur de degré 3; on part de la base Π ci-dessous :

p^m	X + 244	X^2 + 244 X	X^3 + 244 X^2
625	244	0	0
0	1	244	0
0	0	1	244
0	0	0	1

On calcule ensuite la base réduite Π_r qui s'écrit :

p_0	p_1	p_2	p_3
3	2	−3	1
2	0	3	−4
−2	5	4	0
2	−2	1	3

On peut vérifier que π_0 et P ne sont pas premiers entre eux sur Z[X]. On a plus précisément $P = (3 + 2X - 2X^2 + 2X^3)(-1 - X + 3X^2)$.

c) Programmation de l'algorithme en REDUCE.

L'algorithme a été programmé en REDUCE en utilisant le programme donné en **A.III.3** pour le calcul d'une base réduite. L'exécution de ce programme n'est pas rapide, à cause de la longueur du calcul de la base réduite. L'intérêt de cet algorithme est d'avoir un coût polynômial suivant le degré, et d'être plus simple que celui de Lenstra.

b. Factorisation utilisant un monomorphisme.

Introduction.

La méthode proposée ci-dessous est originale; elle a été d'abord présentée à l'EUROCAL 87 **[V5]**, puis elle a été reprise dans un article qui va paraître dans la RAIRO **[V6]**.
Cette méthode de factorisation utilise un monomorphisme Φ_n qui transforme un produit en une somme. Ce monomorphisme sert dans la troisième étape de la factorisation, à savoir dans la recherche d'un diviseur du polynôme P donné, à partir de la factorisation de P sur $\mathbf{Z}/(p^m)[X]$. Si on note $Q_1 \cdots Q_r$ cette factorisation, on doit calculer tous les sous-produits extraits de $Q_1 \cdots Q_r$, et vérifier si l'un d'eux divise P sur $\mathbf{Z}[X]$.
La méthode proposée ci-dessous consiste à calculer les images des facteurs Q_j de P par Φ_n. La recherche des sous-produits de $Q_1 \cdots Q_r$ de l'algorithme classique est remplacée ici par la recherche des sous-sommes de $\Phi_n(Q_1) + \cdots + \Phi_n(Q_r)$. On obtient ainsi un critère de reconnaisance des diviseurs de P sur $\mathbf{Z}[X]$.
Dans la méthode classique, le coût est dominé par le calcul des sous-produits de $Q_1 \cdots Q_j \cdots Q_r$ sur $\mathbf{Z}/(p^m)[X]$ avec $p^m > 2^d \mid a_0 \mid \cdot \| P \|$. Comme il y a 2^{r-1} produits à calculer, ainsi que 2^{r-1} divisions, le coût est donné par
$2^r d^2(d + \mathrm{Log} \| a_0 P \|)^2$.
Avec la méthode proposée le facteur $(d + \mathrm{Log} \| a_0 P \|)^2$ est remplacé par
$2 + \mathrm{Log} \| P \|$. Mais l'utilisation du monomorphisme Φ_n augmente les calculs de la factorisation sur $\mathbf{Z}/(p^m)[X]$, car les calculs doivent être faits modulo p^m avec
$p^m > (4 \| P \|)^d$.
Cependant, au total le coût du nouvel algorithme est meilleur que celui de l'algorithme classique, lorsque le degré d de P est suffisamment grand, par rapport à $\| P \|$.

Un tel monomorphisme Φ_n permet également de calculer des expressions algébriques, comme le quotient de produits de puissances de polynômes comme on le verra dans le paragraphe 2°).

Le monomorphisme transforme une telle expression en une combinaison linéaire de polynômes ce qui permet de diminuer le coût du calcul de ce quotient.

L'usage du monomorphisme Φ_n pourrait se généraliser pour calculer une expression algébrique plus complexe contenant des radicaux, ou même pour le calcul du développement limité d'une expression contenant des logarithmes et des exponentielles.

1°) Définition du monomorphisme Φ_n.

Considérons le polynôme suivant de $Z[X]$,

$$P(X) = a_0 X^d + a_1 X^{d-1} + \cdots + a_{d-1} X + a_d,$$

dont les coefficients $a_0, a_1, ..., a_d$ sont premiers entre eux; un tel polynôme est appelé **polynôme primitif**.

Remarquons qu'on peut facilement transformer P en polynôme unitaire, en remplaçant X par X/a_0 dans P, puis en multipliant l'expression obtenue par a_0^{d-1}; on obtient le polynôme

$$\begin{aligned} P_u(X) &= a_0^{d-1} P(X/a_0) \\ &= X^d + a_1 X^{d-1} + a_2 a_0 X^{d-2} + \cdots + a_d a_0^{d-1}. \\ &= X^d + \alpha_1 X^{d-1} + \cdots + \alpha_{d-1} X + \alpha_d. \end{aligned}$$

On peut également retrouver P à partir de P_u, ou trouver un diviseur Q de P à partir d'un diviseur Q_u de P_u, en remplaçant X par $a_0 X$ dans P_u ou dans Q_u, puis en rendant primitif le polynôme obtenu.

Signalons que dans la méthode classique de factorisation, le polynôme P est rendu unitaire sur $Z/(p^m)$ en mettant a_0 en facteur (p étant choisi de façon qu'il ne divise pas a_0). Alors P a ses coefficients majorés par p^m. Cette mise en facteur de a_0 n'est pas possible ici, car le calcul de Φ_n exige que le polynôme de départ soit unitaire sur $Z[X]$ pour qu'aucun dénominateur n'apparaisse dans le calcul de Φ_n. Evidemment, P_u a des coefficients beaucoup plus grands que ceux de P. Si on veut comparer les deux méthodes, on devra donc effectuer les calculs du coût en fonction des coefficients de P et non pas en fonction de ceux de P_u. Ces calculs de coût seront effectués plus loin.

Le monomorphisme Φ_n sera défini ci-dessous pour un polynôme P unitaire. On suppose donc dans la suite du paragraphe que P est unitaire.
Notons P' la dérivée de P et considérons la division euclidienne définie par

(A.IV.17) $X^{n+1} P' = P \cdot P^* + R$ où $\deg R < \deg P$.

On écrit P^* sous la forme $P^* = a_0^* X^n + a_1^* X^{n-1} + \cdots + a_n^*$, et on note $\Phi_n(P)$ le polynôme réciproque de P^*, noté $\Phi_n(P) = a_0^* + a_1^* X + \cdots + a_n^* X^n$.
Ce polynôme $\Phi_n(P)$ sera appelé **développement logarithmique** de P à l'ordre n, car il est obtenu à partir de la dérivée logarithmique de P. L'application que Φ_n vérifie la relation:

(A.IV.18) $\Phi_n(P \cdot Q) = \Phi_n(P) + \Phi_n(Q)$.

Dans la suite, on exprimera cette propriété en disant que Φ_n est un homomorphisme. On vérifie facilement cette relation. En effet, d'après la définition de $\Phi_n(P)$ et de $\Phi_n(Q)$, on obtient

$X^{n+1} P' = P \cdot P^* + R$ où $\deg R < \deg P$ et

$X^{n+1} Q' = Q \cdot Q^* + S$ où $\deg S < \deg Q$. Par suite

$$X^{n+1} (P \cdot Q)' = X^{n+1} (P' \cdot Q + P \cdot Q')$$
$$= (P \cdot Q) \cdot (P^* + Q^*) + R \cdot Q + S \cdot P,$$

où $\deg (R \cdot Q + S \cdot P) < \deg P \cdot Q$. Donc **(A.IV.18)** est satisfait.
De la relation **(A.IV.17)** définissant P^*, on déduit les relations suivantes :

(A.IV.19) $a_0^* = d$ pour $j=0$;
(A.IV.20) $a_j^* + a_{j-1}^* a_1 + \cdots + a_1^* a_{j-1} + j a_j = 0$ pour $1 \le j \le d$;
(A.IV.21) $a_j^* + a_{j-1}^* a_1 + \cdots + a_{j-d}^* a_d = 0$ pour $d \le j \le n$.

Notons $z_1, \ldots, z_d$ les racines de P (distinctes ou non). Alors
$P = (X - z_1) \cdots (X - z_d)$; par suite $P^* = (X - z_1)^* + \cdots + (X - z_d)^*$. On déduit des trois relations ci-dessus que $\Phi_n(X - z_k) = 1 + z_k X + \cdots + (z_k)^j X^j + \cdots$. Par suite les coefficients a^*_j de P sont égaux à $s_j = \sum_{1 \le k \le d} (z_k)^j$.
Les relations **(A.IV.20)** et **(A.IV.21)** définissent donc la somme a_j^* des puissances $p^{ième}$ des racines d'une équation en fonction de $a_0, a_1, \ldots, a_j$. Ces relations ne sont donc rien d'autre que les **formules de Newton**.

La deuxième de ces relations, à savoir **(A.IV.20)** définit aussi $j\,a_j$ comme une fonction linéaire de $a_1^*, a_2^*, \ldots, a_j^*$ pour $j \geq 1$. Par suite les n premiers coefficients de P, à savoir $a_1, \ldots, a_n$ sont définis par les n premiers coefficients de P^*. Donc si $n = d$, la connaissance de $\Phi_n(P)$ suffit à définir P de façon unique.

Si on note Π^d l'ensemble des polynômes unitaires dont le degré est majoré par d, alors Φ_n est un monomorphisme de Π^d dans lui-même, pour $n \geq d$. Evidemment, Φ_n n'est pas surjectif, car la relation **(A.IV.20)** définit $j\,a_j$ comme un entier qui n'a aucune raison d'être divisible par j.

Etudions le coût du calcul du développement logarithmique jusqu'à l'ordre n d'un polynôme de degré d. Dans les formules de Newton données ci-dessus le nombre total de multiplications de coefficients est égal à

$$S = 1 + 2 + \cdots + d + d + \cdots + d,$$

où le nombre de termes de S est égal à n. Le coût relatif au calcul des coefficients a_j^* est donc majoré par d n multiplications de coefficients, c'est-à-dire que le coût est donné par $d\,n\,\mathrm{Log}^2 p^m$ si les calculs sont effectués modulo p^m.

2°) Utilisation de Φ_n pour calculer une expression algébrique rationnelle.

Si le polynôme B divise le polynôme A, alors comme Φ_n est un homorphisme, on a la relation $\Phi_n(A/B) = \Phi_n(A) - \Phi_n(B)$. Pour obtenir $Q = A/B$ à partir de A et de B il suffit donc de connaître le développement logarithmique de Q jusqu'à l'ordre $n = \deg Q = \deg A - \deg B$.

On se propose de généraliser la relation $\Phi_n(A/B) = \Phi_n(A) - \Phi_n(B)$ à tout couple de polynômes (A, B) en désignant par [A/B] le quotient de la division euclidienne de A par B, à savoir

(A.IV.22) $\quad \Phi_n([A/B]) = \Phi_n(A) - \Phi_n(B).$

Pour montrer cette relation, utilisons la relation **(A.IV.17)** définissant $\Phi_n(P) = P^*$, pour les polynômes $P = A, B$ et $Q = [A/B]$; on obtient

$$X^{n+1} A' = A \cdot A^* + R_1 \quad \text{avec } \deg R_1 < \deg A ,$$
$$X^{n+1} B' = B \cdot B^* + R_2 \quad \text{avec } \deg R_2 < \deg B \text{ et}$$
$$X^{n+1} Q' = Q \cdot Q^* + R \quad \text{avec } \deg R < \deg Q.$$

On a d'autre part $A = B\cdot Q + S$ où $\deg S < \deg B$. Calculons $X^{n+1} A'$ qui vaut $X^{n+1} (B'\cdot Q + B\cdot Q' + S')$; en remplaçant $X^{n+1} Q'$ et $X^{n+1} B'$ par leurs valeurs :

$$\begin{aligned} X^{n+1} A' &= Q\cdot(B\cdot B^* + R_2) + B\cdot(Q\cdot Q^* + R) + X^{n+1} S' \\ &= Q\cdot B\cdot(B^* + Q^*) + Q\cdot R_2 + R\cdot B + X^{n+1} S' \\ &= (Q\cdot B + S) (B^* + Q^*) + Q\cdot R_2 + R\cdot B + X^{n+1} S' - S(B^* + Q^*). \end{aligned}$$

Comme les développements logarithmiques sont faits à l'ordre $n = \deg Q$, on a $\deg B^* = \deg Q^* = n$ et de plus $\deg S' = \deg S - 1 < \deg B - 1$. Par suite $\deg (Q\cdot R_2 + R\cdot B + X^{n+1} S' - S (B^* + Q^*) < \deg B + \deg Q$.

Il en résulte que $B^* + Q^*$ est le quotient de la division enclidienne de $X^{n+1} A'$ par $A = Q\cdot B + S$. La relation **(A.IV.22)** est donc bien vérifiée.

Exemple.

Calcul de la partie entière de la fraction rationnelle $F = \dfrac{(X - 7)^{10} (X + 3)^7}{(X - 2)^{13}}$.

Pour obtenir la partie entière de F, on a besoin de son développement logarithmique à l'ordre $n = 17 - 13 = 4$. On calcule donc successivement

$$\begin{aligned} \Phi_4(X - 7) &= X^4 + 7 X^3 + 49 X^2 + 343 X + 2401 \\ \Phi_4(X + 3) &= X^4 - 3 X^3 + 9 X^2 - 27 X + 81 \\ \Phi_4(X - 2) &= X^4 + 2 X^3 + 4 X^2 + 8 X + 16. \end{aligned}$$

On en déduit

$$\begin{aligned} 10\,\Phi_4(X - 7) + 7\,\Phi_4(X + 3) - 13\,\Phi_4(X - 2) &= 4 X^4 + 23 X^3 + 501 X^2 + 3137 X + 24369. \\ &= 4 X^4 + b_1^* X^3 + b_2^* X^2 + b_3^* X + b_4^* \end{aligned}$$

On utilise la formule de Newton **(A.IV.20)** pour calculer les coefficients b_1, b_2, b_3 et b_4 de F. On obtient

$$\begin{aligned} b_1 &= - b_1^* && = - 23 \\ 2 b_2 &= - b_2^* - b_1 b_1^* && = - 501 + (23)^2 = 28 \\ 3 b_3 &= - b_3^* - b_1 b_2^* - b_2 b_1^* && = - 3137 + 501\cdot 23 - 23\cdot 14 = 8064 \\ 4 b_4 &= - b_4^* - b_1 b_3^* - b_2 b_2^* - b_3 b_1^* && = - 24369 + 23\cdot 3137 - 14\cdot 501 - 23\cdot 2688 \\ & && = - 21056. \end{aligned}$$

On en déduit que la partie entière de F est égale à

$X^4 - 23 X^3 + 14 X^2 + 2688 X - 5264$.

3°) Factorisation des polynômes.

Considérons la factorisation de P sur $\mathbf{Z}/(p^m)[X]$ en facteurs irréductibles, qui s'écrit $P = a_0 Q_1 \cdots Q_r$ et supposons que P ne soit pas irréductible sur $\mathbf{Z}[X]$ et s'écrive $P = Q \cdot R$, où $Q = b_0 Q_1 \cdot Q_2 \cdots Q_h$ et $R = c_0 Q_{h+1} \cdots Q_r$.
Considérons les développements logarithmiques des polynômes $Q_1, \ldots, Q_r$ à l'aide du monomorphisme Φ_n avec $n = d$. Le développement de Q est donc égal à $F = \Phi_n(Q_1) + \cdots + \Phi_n(Q_h)$ et celui de R est égal à $G = \Phi_n(Q_{h+1}) + \cdots + \Phi_n(Q_r)$.
Factoriser P sur $\mathbf{Z}[X]$ revient donc à définir deux sommes F et G extraites de $\Phi_n(Q_1) + \cdots + \Phi_n(Q_r)$. Mais F et G doivent être les développements logarithmiques de polynômes Q et R de $\mathbf{Z}[X]$. Le théorème ci-dessous donne des conditions pour que de telles sommes F et G correspondent à des polynômes de $\mathbf{Z}[X]$.

Théorème.

Soit $P = a_0 X^d + \cdots + a_d$ un polynôme primitif de $\mathbf{Z}[X]$, tel que $\| P/a_0 \| \geq 5{,}3$ et soit $P_u = X^d + \alpha_1 X^{d-1} + \cdots + \alpha_d$ le polynôme unitaire associé.
Soient deux polynômes $Q = X^q + b_1 X^{q-1} + \ldots + b_q$ et $R = X^r + c_1 X^{r-1} + \ldots + c_r$ de degrés q et r tels que $P_u = Q \cdot R$ sur $\mathbf{Z}/(p^m)[X]$ avec $p^m > (4 \| P \|)^d$.
On suppose que les coefficients de Q et R sont calculés dans l'intervalle $]-\frac{1}{2}p^m, \frac{1}{2}p^m]$ et on note b_j^* et c_j^* les coefficients des développements logarithmiques de Q et de R. Alors les deux assertions suivantes sont équivalentes:

1) $P_u = Q \cdot R$ sur $\mathbf{Z}[X]$;

2) $| b_j^* | \leq n(| a_0 | + H(P))^j + (q - n) | a_0 |^j$ et

$| c_j^* | \leq n(| a_0 | + H(P))^j + (r - n) | a_0 |^j$ pour $1 \leq j \leq d$,

où n est la partie entière de $\dfrac{\operatorname{Log} \| P/a_0 \|}{\operatorname{Log}(1 + H(P/a_0))} - 1$.

Remarque:

L'hypothèse $\| P/a_0 \| \geq 5{,}3$ ne sert qu'à simplifier la borne de p^m. Si on la supprime on obtient pour p^m la borne suivante $(5{,}65 \| P \|)^d$.

La démonstration du théorème utilise les deux lemmes suivants:

Lemme 1.

Soit (u_j) la suite définie par $u_0 = 1$ et

$j\,u_j = (q\,A^j + S^j)\,u_0 + (q\,A^{j-1} + S^{j-1})\,u_1 + \cdots + (q\,A + S)\,u_{j-1}$, pour $j \geq 1$, les nombres q, A et S étant des entiers positifs. Alors:

(i) $u_j = S^j + \binom{q}{1} \cdot A \cdot S^{j-1} + \cdots + \binom{q+j-1}{j} A^j$,

(ii) $u_j \leq (A + S)^j \left(1 + \frac{A}{S}\right)^{q-1}$.

Démonstration.

Considérons le polynôme unitaire $A = \sum_{0 \leq k \leq j} v_k X^k$ dont le développement logarithmique est égal à $\Phi_n(A) = \sum_{0 \leq k \leq n} - q\,A^k X^k$. La formule de Newton **(A.IV.20)** permet de calculer les coefficients v_k. On obtient ainsi $v_0 = 1$ et

$k\,v_k = q\,(v_0 A^k + v_1 A^{k-1} + \cdots + v^{k-1} A)$. Par suite, $k\,v_k$ est égal à

$(k - 1)\,v_{k-1}A + q\,v_{k-1}\,A$, soit à $(q + k - 1)\,v_{k-1}\,A$. Il en résulte que

$v_k = \frac{(q+k-1)(q+k-2)\cdots q}{k\,(k-1)\cdots 2 \cdot 1}\,A^k = \binom{q+k-1}{k} A^k$.

Considérons maintenant, le polynôme $B = \sum_{0 \leq k \leq j} w_k X^k$ de développement logarithmique égal à $\sum_{0 \leq k \leq n} -S^k X^k$. La formule de Newton **(A.IV.20)** permet de calculer facilement les coefficients w_k; on trouve $w_k = S^k$.

La formule du lemme définissant u_j par récurrence s'écrit sous la forme

$j\,u_j = -u^*_j - u^*_{j-1}\,u_1 - \cdots - u^*_1\,u_{j-1}$, avec $u^*_j = q\,A^j + S^j = v^*_j + w^*_j$. D'après la propriété d'homomorphisme de Φ_n, $(A \cdot B)^* = A^* + B^*$, donc les termes u_j sont les coefficients du polynôme $A \cdot B$; ils s'écrivent donc $u_j = \sum_{0 \leq k \leq j} v_k\,w_{j-k}$. Par suite :

$u_j = S^j + \binom{q}{1} A \cdot S^{j-1} + \ldots + \binom{q+j-1}{j} A^j$, et **(i)** est bien vérifié.

Il reste à majorer u_j. Pour cela remplaçons dans le second membre de (i), $\binom{q+k-1}{k}$ par $\binom{q+j-1}{k}$ pour $0 \leq k \leq j$. Puis multiplions les deux membres par S^{q-1}. On obtient $S^{q-1}\,u_j \leq (A+S)^{q+j-1}$, donc le lemme est bien vérifié.

Lemme 2.

Le plus petit commun multiple des nombres 2, 3, ..., n noté $\pi(n)$ est majoré par $(2,826)^n$.

Démonstration.

Si on note $\Psi(n)$ le logarithme népérien de $\pi(n)$, alors d'après **[RS]** on a $\Psi(n) \leq 1{,}03883\ n$. Donc, $\pi(n) \leq (e^{1,03883})^n \leq (2,826)^n$.

Démonstration du théorème.

1) ⟹ 2)

Le polynôme Q étant un diviseur de P_u, notons Q° le diviseur de P associé à Q et b°_j les coefficients du développement logarithmique de Q°. Notons $z_1, \ldots, z_d$ les racines de P, distinctes ou non et $z_1, \ldots, z_q$ celles de Q°. Le changement de variable substituant X à X/a_0 et rendant P unitaire transforme les racines z_j de P en $a_0 z_j$. On aura donc $| b^*_j | = | a_0 |^j | b^\circ_j |$.

Considérons la somme $f(z_1,\ldots,z_q) = | z_1 |^j + \ldots + | z_q |^j$, et notons $| z_u.z_v.\ldots.z_w |$ le produit des racines de Q° de module minoré par 1. Le produit $| a_0 z_u z_v \ldots z_w |$ noté $M(Q^\circ)$ est appelé la mesure de Q° et on a évidemment $M(Q^\circ) \leq M(P)$.

Rappelons le résultat suivant démontré par Mignotte dans **[M2]**, $M(P) \leq \| P \|$. La somme $f(z_1,\ldots,z_q)$ ci-dessus est majorée par la somme

$$| z_u |^j + | z_v |^j + \ldots + | z_w |^j + 1 + \ldots + 1$$

dans laquelle on a remplacé par 1, les racines $| z_j |$ majorées par 1. Comme le produit $| a_0 | \cdot | z_u | \cdot | z_v | \cdots | z_w |$ est majoré par $M(Q^\circ)$, f sera maximum lorsque ce produit sera égal à $M(Q^\circ)$. Or chaque racine a une valeur maximum donnée par l'inégalité de Cauchy $| z_k | < 1 + H(P/a_0)$, H(P) étant la hauteur de P (maximum de la valeur absolue des coefficients). Les sommes $f(z_1, \ldots, z_q)$ sont donc majorées par

$$B = f(1 + H(P/a_0), \ldots, 1 + H(P/a_0), 1, \ldots, 1)$$

où on a remplacé le plus grand nombre possible de racines $z_1, \ldots, z_q$ par leur valeur maximum $1 + H(P/a_0)$; notons n le plus grand entier tel que $(1 + H(P/a_0))^n < \| P/a_0 \|$; alors n est la partie entière de $\dfrac{\mathrm{Log} \| P/a_0 \|}{\mathrm{Log}(1 + H(P/a_0))} - 1$.

On a de plus

(A.IV.23) $\quad 1 + H(P/a_0) < \| P/a_0 \|^{1/n}$.

On obtient ainsi $| b^\circ_j | \leq n(1 + H(P/a_0))^j + q - n$. Il en résulte que

$| b_j^* | \leq n (| a_0 | + H(P))^j + (q - n) | a_0 |^j$ et
$| c_j^* | \leq n (| a_0 | + H(P))^j + (r - n) | a_0 |^j$.
L'assertion 2) du théorème est donc satisfaite.

Remarque.
On pourrait donner un meilleur majorant pour $| b_j^* |$ et $| c_j^* |$ en remplaçant n par la partie entière n_1 de $\dfrac{\text{Log}(M(Q°)/a_0)}{\text{Log}(1 + H(P/a_0))} - 1$ dans l'expression majorant $| b_j^* |$ et par la partie entière n_2 de $\dfrac{\text{Log}(M(R°)/a_0)}{\text{Log}(1 + H(P/a_0))} - 1$ dans l'expression majorant $| c_j^* |$. En remarquant que $M(Q°)\cdot M(R°) = M(P)$, on obtient $n = n_1 + n_2$; donc l'un des nombres n_1 ou n_2 est majoré par ½ n. Mais cette majoration est moins facile à utiliser.

2) ⟹ 1).
D'après 2), $| b_j^* | \leq (n(1 + H(P/a_0))^j + q - n) | a_0 |^j$; donc d'après la majoration **(A.IV.23)**, démontrée dans la première partie de la démonstration,
$| b_j^* | \leq (n \| P/a_0 \|^{j/n} + q - n) | a_0 |^j$. Or $n \leq \text{Log} \| P/a_0 \|$ et $n \| P/a_0 \|^{j/n} + q - n$ est une fonction décroissante de n pour ces valeurs de n. On obtient donc

$$| b_j^* | \leq \| P \|^j + (q - 1) | a_0 |^j.$$

Utilisons les formules de Newton **(A.IV.20)** et **(A.IV.21)** pour majorer les coefficients de Q. La première formule donne pour $j = 1$,
$| b_1 | = | b_1^* | \leq \| P \| + (q - 1) | a_0 |$, et elle donne pour $j > 1$,
$j | b_j | \leq | b_1^* | \cdot | b_{j-1} | + \cdots + | b_j^* |$.
Comme on a $| b_j^* | \leq \| P \|^j + (q - 1) | a_0 |^j$, on obtient :
$j | b_j | \leq (\| P \| + (q - 1) | a_0 |)\cdot | b_{j-1} | + \cdots + \| P \|^j + (q - 1) | a_0 |^j$.
Notons u_j la suite définie par les relations :
$u_1 = \| P \| + (q - 1) | a_0 |$ et
$ju_j = (\| P \| + (q - 1) | a_0 |)u_{j-1} + \ldots + \| P \|^j + (q - 1) | a_0 |^j$.
Il est clair que pour $1 \leq j \leq d$, on a $j | b_j | \leq j u_j$. Or, d'après le **lemme 1**, on a :
$u_j \leq (1 + | a_0 | / \| P \|)^{q-2}(| a_0 | + \| P \|)^j$. Il en résulte donc :

(A.IV.24) $j | b_j | \leq j (1 + | a_0 | / \| P \|)^{q-2}(| a_0 | + \| P \|)^j$.

On obtient évidemment une majoration analogue avec les coefficients c_j :

(A.IV.25) $j\ |c_j| \le j(1 + |a_0|/\|P\|)^{r-2}(|a_0| + \|P\|)^j$.

Malheureusement on ne peut pas "simplifier" par j, car on est sur $\mathbf{Z}/(p^m)$ et que $j\,b_j$ n'a aucune raison d'être divisible par j sur $\mathbf{Z}$.

Notons N le plus petit commun multiple de 2, 3, ..., q et M le plus petit commun multiple de 2, 3, ..., r. D'après les hypothèses du théorème on a

$N\cdot M\cdot P_u = N\cdot Q\cdot M\cdot R$ sur $\mathbf{Z}/(p^m)[X]$.

Le $k^{ième}$ coefficient de $N\cdot Q\cdot M\cdot R$ noté d_k s'écrit

$d_k = N\cdot M\cdot(b_k c_0 + b_{k-1} c_1 + \cdots + b_1 c_{k-1} + b_0 c_k)$, avec $b_j = 0$ (resp $c_j = 0$) pour $j > q$ (resp $j > r$). Majorons d_k en utilisant les majorations **(A.IV.24)** et **(A.IV.25)**, démontrées ci-dessus; on obtient

$$d_k \le N\cdot M(1+|a_0|/\|P\|)^{q-2}(1+|a_0|/\|P\|)^{r-2}\sum_{0\le j\le k} s_j,\ \text{où}\ S_k = \sum_{0\le j\le k} s_j \quad \text{avec}$$

$s_j = (|a_0| + \|P\|)^j\,(|a_0| + \|P\|)^{k-j}$ pour $j \le q$ et $k-j \le r$ et $s_j = 0$ pour $j > q$ et $k-j > r$. On vérifie facilement que S_k est une suite croissante, donc sa valeur maximum est atteinte pour $k = d$ et vaut $(|a_0| + \|P\|)^d$. Les coefficients d_k sont donc majorés par $N\cdot M\cdot(1 + |a_0|/\|P\|)^{d-4}(|a_0| + \|P\|)^d$.

D'après le **lemme 2**, on a $N \le (2{,}826)^q$ et $M \le (2{,}826)^r$. On a donc

$d_k \le (2{,}826)^d(1 + |a_0|/\|P\|)^{-4}((|a_0| + \|P\|)/\|P\|)^{2d}\,\|P\|^d$. Si on suppose $\|P/a_0\|$ supérieur ou égal à 5,3, alors $1 + |a_0|/\|P\|$ est majoré par 1,189, donc $(2{,}826)^d((|a_0| + \|P\|)/\|P\|)^{2d} \le 4^d$. D'autre part, la valeur maximum de $(1 + |a_0|/\|P\|)^{-4}((|a_0| + \|P\|)/\|P\|)^{2d}$ est obtenue pour $\|P/a_0\| = 5{,}3$ (en supposant $d \ge 2$), et dans ce cas $(1 + |a_0|/\|P\|)^{-4} < ½$. On obtient donc finalement $d_k \le ½(4\|P\|)^d \le ½p^m$.

Les polynômes $N\cdot M\cdot P_u$ et $N\cdot Q\cdot M\cdot R$ ont donc leurs coefficients compris entre $-½p^m$ et $½p^m$. Comme ils sont égaux modulo p^m, ils sont donc égaux sur $\mathbf{Z}[X]$ et le théorème est vérifié.

Remarque:

On pourrait améliorer la borne de p^m. Pour cela on ne calcule les développements logarithmiques de Q et R que jusqu'à l'ordre ½d, ce qui définit les ½d premiers coefficients de Q et R. Mais on calcule en plus les développements logarithmiques des

polynômes réciproques de Q et R toujours à l'ordre ½d; alors tous les coefficients de Q et R sont définis à partir de ces deux types de développements logarithmiques. Cependant il faut généraliser la définition de Φ_n au cas des polynômes non unitaires ce qui rend plus complexes les formules de Newton, **(A.IV.20)** et **(A.IV.21)**. Cette généralisation est faite dans **[V5]**; elle donne une borne de p^m de la forme $B^{d/2}$ au lieu de B^d.

4°) Etude du coût du nouvel algorithme.

La première étape de l'algorithme est la factorisation de P sur $\mathbf{Z}/(p)[X]$ par la méthode de Berlekamp. Elle exige $d^3 p \, \mathrm{Log}^2 p$ opérations élémentaires, comme on l'a vu dans **A.II**.

Les calculs de la deuxième étape sont dominés par le calcul de produit de polynômes de degré d, ce qui correspond à d^2 produits de coefficients. Comme les coefficients sont majorés par p^m le coût de la deuxième étape est donné par $d^2 \, \mathrm{Log}^2(p^m)$.

Dans l'algorithme classique on a $p^m \leq 2^d \, \| a_0 P \|$, donc le coût est égal à $d^2 (d + \mathrm{Log}(\| a_0 P \|))^2$.

Dans le nouvel algorithme, les coefficients sont bornés par $(4 \| P \|)^d$, donc le coût est égal à $d^4 (2 + \mathrm{Log} \| P \|)^2$. Le coût du calcul des images $\Phi_n(Q_1)$, ..., $\Phi_n(Q_r)$ a été étudié dans le premier paragraphe. Ce coût est égal à $d^2 \, \mathrm{Log}^2(p^m)$, soit encore à $d^4 (2 + \mathrm{Log} \| P \|)^2$. Le coût total du nouvel algorithme est donc donné par

$$2 \, d^4 (2 + \mathrm{Log} \| P \|)^2.$$

Pour la dernière étape de l'algorithme classique, on doit dans le cas où P est irréductible, calculer 2^{r-1} produits de polynômes sur $\mathbf{Z}/(p^m)[X]$, puis 2^{r-1} divisions; le nombre total de multiplications de coefficients est de l'ordre de $2^r d^2$. Le produit de deux coefficients définis modulo p^m a un coût égal à $\mathrm{Log}^2 p^m = (d + \mathrm{Log} \| a_0 P \|)^2$. On obtient donc pour la dernière étape un coût égal à $2^r d^2 (d + \mathrm{Log} \| a_0 P \|)^2$.

Dans le nouvel algorithme, la dernière étape exige le calcul de 2^r sommes de polynômes sur $\mathbf{Z}/(p^m)[X]$; le coût est donc donné par $2^r d^2 (2 + \mathrm{Log} \| P \|)$.

Pour comparer les deux algorithmes il faut donc comparer le coût de la dernière étape de l'algorithme classique avec le coût de l'étape précédente du nouvel algorithme. Le nouvel algorithme n'est donc intéressant que si

$$2 \, d^4 (2 + \mathrm{Log} \| P \|)^2 \leq 2^r d^2 (d + \mathrm{Log}(\| a_0 P \|))^2,$$

soit

$$d^2(2 + \mathrm{Log} \| P \|)^2 \leq 2^{r-1} (d + \mathrm{Log}(\| a_0 P \|))^2,$$

Supposons pour simplifier que $\mathrm{Log} \| a_0 P \|$ soit négligeable par rapport à d. On obtient $2 + \mathrm{Log} \| P \| \leq 2^{(r-1)/2}$, soit encore :

(A.IV.26) $\quad r > 2 \mathrm{Log}(2 + \mathrm{Log} \| P \|).$

Pour étudier cette inégalité, on doit comparer r et d. D'après M. Mignotte et J.L. Nicolas dans **[MN]**, le nombre de polynômes P vérifiant $| r - \log(d) | \geq \sqrt{\log d}$ est majoré par $C \lambda^{-2} p^d$ pour tout $\lambda \geq 0$. D'après M. Mignotte, on peut prendre $C = 4$. Faisons cette hypothèse pour la suite. Remarquons que le nombre total de polynômes unitaires de $\mathbf{Z}/(p)[X]$, ayant un degré $\leq$ d, est égal à p^d. Si on prend donc $\lambda = \sqrt{8}$, on peut affirmer qu'un polynôme sur deux au moins, vérifie la relation $| r - \log(d) | \leq \sqrt{8 \log d}$. On a donc une fois sur deux :

$$r \geq \log(d) - \sqrt{8 \log d}.$$

En utilisant, la relation **(A.IV.26)**, ci–dessus minorant r, on obtient la condition

$$\log(d) - \sqrt{8 \log d} - 2 \mathrm{Log}(2 + \mathrm{Log} \| P \|) > 0.$$

Il est clair que cette inégalité est vérifiée pour d suffisamment grand. Cependant les valeurs de d vérifiant cette inégalité sont de l'ordre du million. On est donc sûr que pour ces valeurs de d, la minoration de r, à savoir, **(A.IV.26)** est satisfaite au moins une fois sur deux et qu'alors le nouvel algorithme est le plus performant. Mais la minoration **(A.IV.26)** peut être satisfaite pour des degrés beaucoup moins élevés et si c'est le cas on a intérêt à utiliser le nouvel algorithme.

5°) Nouvel algorithme de factorisation.

Le nouvel algorithme comporte les étapes suivantes:

1. Factorisation de P sur $\mathbf{Z}/(p)[X]$ où p est un nombre premier choisi de façon que P n'ait pas de facteur carré;

2. Raffinement de la factorisation pour obtenir la factorisation de P sur $\mathbf{Z}/(p^m)[X]$ avec $p^m > 2^d \| a_0 P \|$; soit $Q_1 \cdots Q_r$ cette factorisation; si $r \leq 2 \text{Log}(2 + \text{Log} \| P \|)$, alors, on passe à l'étape 3, sinon on passe à l'étape 4;
3. Calcul de tous les sous-produits $Q_u \cdots Q_v$ et division de P par ces sous-produits; Fin de l'algorithme.
4. On vérifie si l'un des facteurs Q_j divise P sur Z[X]; si c'est le cas, Fin de l'algorithme, sinon on passe à l'étape 5.
5. Raffinement de la factorisation pour obtenir la factorisation de P sur $\mathbf{Z}/(p^m)[X]$ avec $p^m > (4 \| P \|)^d$;
6. Calcul des r+1 développements $\Phi_n(Q_1)$, ..., $\Phi_n(Q_r)$, $\Phi_n(P)$ jusqu'à l'ordre d;
7. Calcul des 2^r sommes F extraites de $S = \Phi_n(Q_1) + \ldots + \Phi_n(Q_r)$, puis calcul de $G = \Phi_n(P) - F$.
   Si les majorations 2) du théorème sont vérifiées, alors on calcule le produit $Q = Q_u \cdots Q_v$ correspondant à la somme F.

D'après le théorème, on est certain que Q est un diviseur de P sur Z[X]. Le nouvel algorithme n'exige des calculs supplémentaires à ceux de l'algorithme classique que dans le cas où d'une part le nombre de facteurs r modulo p est grand, à savoir supérieur à $2 \text{Log}(2 + \text{Log} \| P \|)$, et où d'autre part aucun facteur modulo p^m n'est un facteur de P sur Z[X].

Proposons une variante de l'algorithme ci-dessus, consistant à effectuer les calculs modulo $p^m > 2^d \| a_0 P \|$ comme dans la méthode classique, puis à considérer le plus grand entier ko tel que $n(|a_0| + H(P))^{ko} + (q - n)|a_0|^{ko} \leq p^m$. Cette variante comporte les étapes suivantes :

1'. Identique à l'étape 1. ci-dessus;
2'. Identique à l'étape 2. ci-dessus;
3'. Calcul des r+1 développements $\Phi_n(Q_1)$, ..., $\Phi_n(Q_r)$, $\Phi_n(P)$ jusqu'à l'ordre d modulo $p^m > 2^d \| a_0 P \|$;
4'. Calcul des 2^r sommes F extraites de $S = \Phi_n(Q_1) + \ldots + \Phi_n(Q_r)$, puis calcul de $G = \Phi_n(P) - F$.
5'. Calcul des produits $Q = Q_u \ldots Q_v$ correspondant aux sous-sommes de S qui vérifient les majorations 2) du théorème pour $1 \leq j \leq ko$.

Ainsi les étapes 2' et 3' de l'algorithme auront un coût du même ordre de grandeur que dans l'algorithme classique. Mais à l'étape 5' on ne calcule que les sous-produits

$Q_u...Q_v$ vérifiant les majorations du théorème. On diminue ainsi le coût de l'algorithme classique en diminuant le nombre de produits $Q_u...Q_v$ à calculer.

6°) Conclusion.

La borne prise pour p^m dans le théorème est très souvent excessive. On peut en fait prendre une borne beaucoup plus raisonnable. Par contre la borne donnée pour p^m dans l'algorithme classique est indispensable, quel que soit l'algorithme utilisé. Si on prenait une borne trop petite pour p^m, c'est-à-dire inférieure à $2^d \| a_0 P \|$, alors on pourrait passer à côté d'un facteur, qui aurait un coefficient supérieur à p^m. En effet dans ce cas on aurait $P = Q \cdot R$ modulo p^m avec les coefficients de Q et R réduits modulo p^m, par suite $P = Q \cdot R$ ne serait pas vrai sur $\mathbf{Z}[X]$.

Par contre, dans le nouvel algorithme on peut choisir p^m très inférieur à la borne théorique $(4 \| P \|)^d$, mais supérieur à la borne classique $2^d \| a_0 P \|$. Si on désigne, comme dans le paragraphe précédent, par ko le plus grand entier tel que $n(| a_0 | + H(P))^{ko} + (q-n) | a_0 |^{ko}$ soit majoré par p^m, alors tout produit $Q = Q_u...Q_v$ qui divise P sur $\mathbf{Z}[X]$, doit vérifier les majorations 2) du théorème pour $1 \le j \le ko$. On obtient ainsi un critère de reconnaissance des diviseurs de P sur $\mathbf{Z}[X]$, avec un coût analogue à celui de l'algorithme usuel.

Rien ne nous permet d'estimer le gain par rapport à l'algorithme classique, mais on peut espèrer qu'ainsi le calcul de nombreux produits $Q_u...Q_v$ sera évité, notamment dans le cas d'un polynôme irréductible.

Remarquons d'autre part que si $| a_d | < | a_0 |$, où a_d est le terme constant de P, on a intérêt à remplacer P par son polynôme réciproque. La borne donnée pour p^m devient alors $2^d \| a_d P \|$.

7°) Exemple.

On donne le polynôme $P = 3X^6 - 4X^4 - 8X^2 - 1$. On rend le polynôme unitaire en calculant $P_u(X) = 3^5 P(X/3) = X^6 - 12X^4 - 216X^2 - 243$.

Si on choisit $p = 5$, on doit faire les calculs modulo 5^m, avec, dans l'algorithme classique $5^m > 2^6 \| a_0 P \| = 1824$; on doit prendre $m \ge 8$ (pour $m = 4$, $p^m = 625$).

On factorise donc P_u sur $\mathbf{Z}/(5^8)[X]$, en utilisant la méthode classique. On obtient successivement les factorisations sur $\mathbf{Z}/(5)[X]$, ..., $\mathbf{Z}/(5^8)[X]$:

$P_u = (X - 1)(X + 1)(X - 3)(X + 3)(X^2 - 2)$ sur $\mathbf{Z}/(5)[X]$, puis

$P_u = (X + 4)(X - 4)(X - 8)(X + 8)(X^2 - 2)$ sur $\mathbf{Z}/(5^2)[X]$, puis

$P_u = (X + 167)(X - 167)(X + 79)(X - 79)(X^2 - 257)$ sur $\mathbf{Z}/(5^4)[X]$, puis

$P_u = (X + 155458)(X - 155458)(X - 59296)(X + 59296)(X^2 + 2243)$ sur $\mathbf{Z}/(5^8)[X]$.

On obtient donc les facteurs suivants :

$Q_1 = X + 155458$, $Q_2 = X - 155458$, $Q_3 = X - 59296$,

$Q_4 = X + 59296$, $Q_5 = X^2 + 2243$.

On calcule les développements logarithmiques des polynômes Q_j jusqu'à l'ordre deg P − 1 = 5. Les coefficients de $\Phi_n(Q_j)$ sont évalués grâce aux formules de Newton; on obtient :

$a_0^* = \deg Q_j$;

$a_1^* = - a_1$;

$a_2^* = (a_1)^2 - 2 a_0 a_2$;

$a_3^* = - (a_1)^3 + 3 a_0 a_1 a_2$;

$a_4^* = (a_1)^4 - 4 a_0(a_1)^2 a_2 + 2 (a_0)^2 (a_2)^2$;

Par suite :

$\Phi_5(Q_1) = 1 - 155458 X + 2264 X^2 - 3787 X^3 + 47571 X^4$;

$\Phi_5(Q_2) = 1 + 155458 X + 2264 X^2 + 3787 X^3 + 47571 X^4$;

$\Phi_5(Q_3) = 1 + 59296 X - 9 X^2 - 143039 X^3 + 81 X^4$;

$\Phi_5(Q_4) = 1 - 59296 X - 9 X^2 + 143039 X^3 + 81 X^4$;

$\Phi_5(Q_5) = 2 - 4486 X^2 - 94152 X^4$.

On calcule toutes les sous-sommes $F_1 = \Phi_5(Q_j) + \Phi_5(Q_k)$ et $F_2 = \Phi_5(Q_u) + \Phi_5(Q_v) + \Phi_5(Q_w)$ sous la forme $A^*_0 + A^*_1 X + A^*_2 X^2 + A^*_3 X^3 + A^*_4 X^4$, puis on examine les majorations 2) du théorème :

$$| A^*_j | \leq n(| a_0 | + H(P))^j + (q - n) | a_0 |^j$$

$$| B^*_j | \leq n(| a_0 | + H(P))^j + (r - n) | a_0 |^j,$$

avec ici q et r majorés par 4, $| a_0 | = 3$, $| a_0 | + H(P) = 11$ et n = 1. Ici ko est le plus grand entier tel que $11^j + (4-1)3^j$ soit majoré par 5^8, donc ko = 5.

Les seules sous-sommes de $\Phi_5(Q_1) + ... + \Phi_5(Q_5)$ qui vérifient les majorations 2) du théorème sont

$F_1 = \Phi_5(Q_3) + \Phi_5(Q_4) = 2 X^2 - 18 X^2 + 162$ et

$F_2 = \Phi_5(Q_1) + \Phi_5(Q_2) + \Phi_5(Q_5) = 4\,X^4 + 42\,X^2 + 990$

On ne calcule donc que les produits

$Q_1\,Q_2\,Q_5 = (X - 155458)(X + 155458)(X^2 + 2243) = X^4 - 21\,X^2 - 27;$

$Q_3\,Q_4 = (X + 59296)(X - 59296) = X^2 + 9.$

On vérifie que $P_n = (X^2 + 9)(X^4 - 21\,X^2 - 27)$.

En remplaçant X par $3X$, puis en rendant primitifs les polynômes obtenus, on obtient la factorisation $P = (X^2 + 1)(3\,X^4 - 7\,X^2 - 1)$.

8°) Application.

Le monomorphisme utilisé dans l'algorithme proposé pour améliorer la factorisation des polynômes peut aussi être utilisé pour effectuer d'autres calculs. On a vu au paragraphe 2, comment utiliser ce monomorphisme pour calculer le quotient de deux polynômes. On peut généraliser l'utilisation du monomorphisme à des expressions plus complexes contenant des produits, des quotients, des logarithmes et des exponentielles. On pourrait alors calculer par exemple, les développements limités de telles expressions.

Partie : B.

POLYNOMES A PLUSIEURS VARIABLES.

Avant d'aborder le problème de la factorisation d'un polynôme, on va étudier plusieurs questions annexes comme dans le cas des polynômes à une variable. Le polynôme donné P appartient à $Z[X, X_1, ..., X_m]$ et on le considère comme un polynôme de $K[X]$ avec $K = Z[X_1, ..., X_m]$.

Pour étudier le polynôme P, le premier travail est donc d'écrire P sous la forme $a_o X^d + ... + a_j X^{d-j} + ... + a_d$, où $a_j \in Z[X_1,...,X_m]$.

Dans les algorithmes utilisés, $a_o, ...,a_d$ doivent être des polynômes premiers entre eux. On devra donc calculer le PGCD δ de $a_o,..., a_d$, puis diviser chacun des coefficients $a_o,..,a_d$ par δ.

I) PROBLEMES PRELIMINAIRES A LA FACTORISATION.

1°) Calcul du plus grand commun diviseur.

On peut adapter l'algorithme d'Euclide, défini pour les polynômes à une variable dans la partie A. On peut en particulier, utiliser l'amélioration de Collins, étudiée dans **(A.I)**. Une autre technique a été proposée par Collins dans **[C2]**; elle utilise la notion de "sous-résultant" de deux polynômes. Un autre algorithme de type modulaire est proposé par Brown dans **[Bo]**. Les coûts de ces deux algorithmes pour des polynômes de degré d à m+1 variables sont de l'ordre de

$$(\mathrm{Log}^2 H(P) + m\,(d+1)\,\mathrm{Log}H(P))\cdot(d+1)^m.$$

2°) Factorisation d'un polynôme ayant au moins un facteur carré.

Comme on l'a vu dans **(A.I)**, on calcule le PGCD du polynôme donné P et de sa dérivée P' par rapport à la variable principale X. Si P a un facteur carré, alors le PGCD ainsi obtenu est un diviseur non trivial de P. On suppose donc dans toute la suite que

(B.I.1) P est sans facteurs carrés.

3°) Majoration des coefficients des diviseurs de P.

Les résultats de ce paragraphe sont dûs à Mignotte et figurent dans son ouvrage **[M1]**, au Chapitre IV. Un polynôme P de $Z[X,X_1,...,X_m]$, peut être écrit sous la forme

$$P = \sum_{j,\,j1,..\,jm} a_{j\,j1..jm}\, X^j X^{j1}...X^{jm},$$

la somme étant étendue à tous les indices j, j1, ..., jm tels que $a_{j\,j1\ldots jm}$ soit non nul.
On appelle **hauteur** de P le nombre $H(P) = \max_{j,\,j1,\ldots\,jm} |\, a_{j\,j1\ldots jm} \,|$.

On appelle **norme** de P le nombre $\| P \| = \sqrt{\Sigma \,|\, a_{j\,j1\ldots jm} \,|^2}$.

Les majorations des coefficients des diviseurs de P utilisent la notion de mesure, comme dans le cas des polynômes à une variable. Rappelons la définition de la mesure donnée en **(A.I.2)**.

On appelle **mesure** du polynôme $P = a_o X^d + \ldots + a_d$, le nombre

$$M(P) = |\, a_o \,| \cdot |\, z_j \,| \cdots |\, z_k \,| ,$$

où $z_j,\ldots, z_k$ sont les zéros de P de modules supérieurs à 1.

Malheureusement, cette définition concrète de la mesure ne se généralise pas aux polynômes à plusieurs variables. On va donc donner une autre formulation de M(P) qui pourra se généraliser. Cette nouvelle formulation de M(P) utilise la formule de Jensen, appliquée aux fonctions holomorphes.

Formule de Jensen.

Si f(z) est une fonction holomorphe à l'intérieur du disque unité, alors:

$$\textbf{(B.I.2)} \quad \log |\, f(0) \,| = \int_0^1 \log |\, f(e^{2i\pi t}) \,|\, dt + \Sigma \log |\, \zeta_j \,|$$

où les nombres ζ_j sont les zéros de f situés à l'intérieur du disque unité.

Appliquons cette formule au polynôme réciproque de P, qui est égal à $X^{d-1} P(1/X) = a_o + a_1 X + \ldots + a_d X^d$. Alors les zéros de ce polynôme réciproque de module inférieurs à 1, sont les zéros de P de module supérieur à 1. Notons $z_j,\ldots,z_k$ ces zéros. La formule de Jensen donne la relation

$$\log |\, a_o \,| = \int_0^1 \log |\, P(e^{-2i\pi t}) \,|\, dt + \log |\, 1/z_j \,| + \cdots + \log |\, 1/z_k \,| .$$

Il en résulte la relation :

$$\textbf{(B.I.3)} \quad \log M(P) = \int_0^1 \log |\, P(e^{2i\pi t}) \,|\, dt.$$

On en déduit la définition suivante, qui généralise la relation (B.I.3) ci-dessus.

Définition.

On considère le polynôme à m+1 variables $P(X,X_1,...,X_m)$, où $m \geq 1$. La mesure de P est définie à partir de la mesure du polynôme à m variables $P(X,X_1,...,X_{m-1},e^{2i\pi t})$ par la formule suivante

(B.I.4) $$M(P(X,X_1,...,X_m)) = \exp\left(\int_0^1 \log M(P(X,X_1,...,e^{2i\pi t}))dt. \right.$$

Pour m = 0, la mesure de P est donnée par la relation **(B.I.3)**.

A partir de cette définition de la mesure, on peut démontrer comme dans le cas des polynômes à une variable que :

(B.I.5) $$\begin{cases} M(P(X,X_1,...,X_m)) \leq \| P(X,X_1,...,X_m) \| \text{, et} \\ M(Q(X,..,X_m)\cdot R(X,..,X_m)) = M(Q(X,..,X_m))\cdot M(R(X,..,X_m)). \end{cases}$$

Dans la suite, on notera d, d1, ..., dm les degrés partiel d'un diviseur Q de P, par rapport aux variables $X, X_1, ..., X_m$; on notera q le degré total de Q, soit $q = d + d1 + ... + dm$. On peut vérifier la majoration suivante :

(B.I.6) $$H(Q) \leq \binom{d}{q}\cdot\binom{d1}{q}\cdots\binom{dm}{q}\cdot M(Q(X,X_1,..,X_m)).$$

On en déduit donc que :

(B.I.7) $$H(Q) \leq \binom{d}{q}\cdot\binom{d1}{q}\cdots\binom{dm}{q}\cdot \| P(X,X_1,..,X_m) \| .$$

On obtient en particulier :

(B.I.8) $$H(Q) \leq 2^q \| P(X,X_1,..,X_m) \| .$$

C'est cette dernière majoration qui sera utilisée dans la suite. Une autre majoration avait été donnée auparavant par Gel'fond dans **[Ge]**. Si $P = Q\cdot R$, alors :

(B.I.9) $$H(Q)\cdot H(R) \leq e^q H(P).$$

Cette majoration est moins bonne que celle de Mignotte, car $\| P \| \leq \sqrt{q}\, H(P)$ et $(e/2)^q \geq \sqrt{q}$ dès que $q \geq 2$.

4°) Factorisation initiale sur Z[X].

On a vu dans l'introduction, que la méthode classique de factorisation consiste à choisir des nombres $\alpha_1, ..., \alpha_m \in \mathbf{Z}$, puis à factoriser le polynôme $P_o = P(X, \alpha_1, ..., \alpha_m)$ sur Z[X]. On note cette factorisation $P(X, \alpha_1, .., \alpha_m) = Q^\circ_1(X) \cdots Q^\circ_r(X)$.
On en déduit que

(B.I.9) $P(X, X_1, .., X_m) = Q^\circ_1(X) \cdots Q^\circ_r(X)$ modulo Δ_1,

où Δ_1 est l'idéal engendré par les polynômes $X_1-\alpha_1, ..,X_m-\alpha_m$. Cette factorisation **(B.I.9)** est le point de départ de l'algorithme de factorisation des polynômes à plusieurs variables; comme on l'a vu dans l'Introduction, page 5, à l'Etape 1 de l'algorithme la factorisation initiale doit être sans facteurs carrés et le degré du polynôme doit être conservé dans la projection. Les deux propriétés suivantes doivent donc être vérifiées :

a) les diviseurs $Q^\circ_j(X)$ de $P(X, \alpha_1, .., \alpha_m)$ doivent être deux à deux distincts;

b) le degré de $P(X, \alpha_1, .., \alpha_m)$ doit être égal au degré d de $P(X, X_1, .., X_m)$ suivant X.

Les nombres $\alpha_1, .., \alpha_m$ devront donc être choisis de façon que **a)** et **b)** soient satisfaits. D'après le paragraphe **2°)** on suppose que P n'a pas de facteur carré. Donc P et P' sont premiers entre eux; par suite, on a la relation :

$$A(X)\cdot P(X, X_1, .., X_m) + B(X)\cdot P'(X, X_1, .., X_m) = \delta(X_1, .., X_m),$$

où δ est un élément non nul de **K**. Il suffit donc de choisir des nombres $\alpha_1, ..., \alpha_m$ qui n'annulent ni $\delta(X_1, .., X_m)$, ni le coefficient dominant de P qui s'écrit $a_0(X_1, .., X_m)$. Comme δ et a_0 sont des polynômes non nuls, ils ne peuvent pas s'annuler partout. Par suite, on pourra trouver des nombres $\alpha_1,.., \alpha_m$ vérifiant les propriétés **a)** et **b)**. Cependant, il est important de majorer le nombre de m-uples $(\alpha_1,.., \alpha m)$ ne vérifiant pas **a)** et **b)**, c'est-à-dire le nombre maximum d'essais infructueux. On peut le faire grâce au lemme suivant, qui se démontre par récurrence sur le nombre m de variables :

Lemme 1.

Soit **A** un anneau et $Q(X_1, .., X_m)$ un polynôme non nul de $A[X_1, .., X_m]$ de degré d_j suivant la variable X_j.

Soit E un pavé de A^m de la forme $E_1 \times .. \times E_i \times .. E_m$ où le cardinal de E_i est égal à (d_i+1). Alors, il existe un élément $(\alpha_1, .., \alpha_m)$ de E tel que $Q(\alpha_1, .., \alpha_m) \neq 0$.

Démonstration.

Le théorème est évident pour $m = 1$. Supposons-le vrai pour les polynômes à $m-1$ variables et écrivons Q comme un polynôme en X_m à coefficients dans $A[X_1, .., X_m]$. Supposons que le Lemme ne soit pas vérifié. Alors pour tout $(\alpha_1, .., \alpha_{m-1}, X) \in E$, le polynôme $Q(\alpha_1,...,\alpha_{m-1},X)$ est nul; ses coefficients doivent donc s'annuler. En utilisant l'hypothèse de récurrence, on en déduit que ce sont des polynômes nuls. Le polynôme Q est donc le polynôme nul, ce qui est absurde. Le Lemme est donc démontré.

Si on note $d, d_1,.., d_m$ les degrés de P suivant $X, X_1,.., X_m$ alors $\delta(X_1,...,X_m)$ est de degré maximum $(2d-1)d_1,...,(2d-1)d_m$ suivant $X_1, .., X_m$. Le produite $\delta.a_o$ a donc des degrés majorés par $2d.d_1, ..., 2d.d_m$. Le lemme permet donc d'affirmer qu'après avoir effectué au maximum

$$(2d\ d_1 + 1)\cdot(2d\ d_2 + 1)\cdots(2d\ d_m + 1)$$

essais avec des m-uples $(\alpha_1, .., \alpha_m)$, on trouve un m-uple qui n'annule ni δ, ni a_0. On suppose qu'on a trouvé un tel m-uple. Remarquons que pour la suite des calculs on a intérêt à avoir le maximum de nombres α_j nuls. On évite ainsi des calculs complexes de coefficients. La recherche de ces nombres α_j telle qu'elle est définie par Wang et Rothschield dans **[WR]**, consiste à effectuer tous les essais possibles en commençant par $\alpha_1 = \alpha_2 = ... = \alpha_m = 0$, puis à essayer de choisir l'un des α_j égal à 1 ou −1. Aucune stratégie particulière n'est adoptée, car très souvent les premiers m-uples choisis conviennent.

Exemple.

$$P = 6X^4 + 2X^3 + 5X^2 - 4 + 2Y(3X^2 - 2X) + Z(2X^4 + 2X^3 + 3X^2 + 5X - 2) + XZ^2 + 2X^2YZ.$$

Le coefficient dominant de P s'écrit $a_o(Y,Z) = 6 + 2Z$. On peut donc prendre ici $\alpha_1 = \alpha_2 = 0$. Alors on obtient $P(X,\alpha_1,\alpha_2) = P(X,0,0) = 6X^4 + 2X^3 + 5X^2 - 4$ qui est de degré 4 comme P et n'a pas de facteur au carré.

II) PROBLEMES RELATIFS AUX POLYNOMES NON UNITAIRES.

1°) Méthode classique.

On peut utiliser une technique analogue à celle des polynômes à une variable. Considérons le polynôme donné P de coefficient directeur $a_o(X_1,...,X_m)$. On a vu dans l'Introduction, que les calculs de factorisation sont effectués sur $\mathbf{K}_j[X]$, avec $\mathbf{K}_j = \mathbf{Z}[X_1-\alpha_1, ..., X_m-\alpha_m]/\Delta_j$. Or Δ_j est l'idéal engendré par les polynômes en $X_1-\alpha_1, ..,X_m-\alpha_m$ de degrés majorés par j et par p^n où p^n majore les coefficients des diviseurs de P sur $\mathbf{Z}[X,X_1,...,X_m]$. Comme pour les polynômes à une variable, p doit être choisi de façon que les diviseurs $Q°_1(X)$, ..., $Q°_r(X)$ de P sur $\mathbf{K}_j$ restent deux à deux distincts. De plus, on doit choisir un nombre premier p qui ne divise pas $a_o(\alpha_1,...,\alpha_m)$. D'autre part, comme on l'a vu au paragraphe précédent, α_1, .., α_m sont choisis de façon que $a_o(\alpha_1,...,\alpha_m) \neq 0$. Il en résulte que le polynôme $a_o(X_1,...,X_m)$ est inversible sur $\mathbf{K}_j = \mathbf{Z}[X_1-\alpha_1, ..., X_m-\alpha_m]/\Delta_j$.

Supposons donc que P ne soit pas unitaire. On peut alors procéder comme dans le cas des polynômes à une variable, c'est-à-dire écrire P sur $\mathbf{K}_j[X]$ sous la forme

$$P = a_o(X^d + a_1 (a_o)^{-1} X^{d-1} + ... + a_d (a_o)^{-1}) = a_o P_o, \text{ où } P_o \text{ est unitaire.}$$

On est donc ramené à la factorisation sur $\mathbf{K}_j[X]$ d'un polynôme unitaire P_o. Notons $Q_1....Q_r$ la factorisation obtenue après la dernière étape de "remontée", c'est-à-dire pour j = n; ainsi, $P = a_o\ Q_1....Q_r$ sur $\mathbf{K}_n[X]$. Par suite, si le produit $Q_{j1}...Q_{jk}$ correspond à un diviseur de P sur $\mathbf{K}[X]$, alors le produit $Q = a_o\ Q_{j1}...Q_{jk}$ calculé sur $\mathbf{K}_n[X]$, puis identifié à un polynôme de $\mathbf{K}[X]$, divise a_o P sur $\mathbf{K}[X]$. On peut donc obtenir un diviseur de P sur $\mathbf{K}[X]$ en rendant primitif le polynôme Q, c'est-à-dire en divisant Q par le polynôme δ qui est le PGCD des coefficients de Q.

Ainsi, la factorisation du polynôme unitaire P_o permet de trouver les facteurs de P sur $\mathbf{K}[X]$. On convient donc pour la suite, de factoriser un polynôme unitaire. Puis dans la phase finale de la factorisation, on multiplie chaque diviseur potentiel Q par a_o, puis on effectue la division de a_o P par a_o Q. On trouvera ainsi tout diviseur de P.

Mais ce procédé est plus lourd que dans le cas des polynômes à une variable. Il exige en particulier de multiplier chaque diviseur potentiel Q par le polynôme a_o^{-1}.
Un exemple d'un tel calcul a été donné dans l'Introduction page 3.

2°) Algorithme de Wang pour le calcul des coefficients dominants.

Elimination des facteurs parasites.

On appelle facteurs parasites, les diviseurs de P sur $\mathbf{K}_o[X]$, qui, après la "remontée" de $\mathbf{K}_o[X]$ dans $\mathbf{K}_n[X]$, ne corespondent à aucun diviseur de P sur $\mathbf{Z}[X,X_1,...,X_m]$. Pour diminuer la probabilité d'avoir des facteurs parasites, on peut factoriser P sur une dizaine d'ensembles $\mathbf{K}_o[X]$ et ne conserver que la factorisation ayant donné le nombre minimum de facteurs.
Dans le cas de polynômes à plusieurs variables, la factorisation initiale est faite sur $\mathbf{Z}[X,\alpha_1,...,\alpha_m]$. La probabilité d'avoir des facteurs parasites est plus faible que dans le cas des polynômes à une variable, avec $\mathbf{K}_o = \mathbf{Z}/(p)$.

Présentation de l'algorithme de Wang.

Wang propose dans **[Wa]**, une méthode de traitement des polynômes non unitaires, nettement plus efficace que la méthode classique exposée plus haut. Elle consiste à factoriser $a_o(X_1,...,X_m)$ sur $\mathbf{Z}[X_1,...,X_m]$, puis à répartir les facteurs trouvés entre les différents diviseurs de P. On obtient ainsi les coefficients dominants des diviseurs de P.

Rappelons qu'au départ on doit factoriser $P(X,\alpha_1,...,\alpha_m)$ sous forme d'un produit de facteurs irréductibles primitifs de $\mathbf{Z}[X]$; notons cette factorisation

$$P(X,\alpha_1,...,\alpha_m) = \delta\, Q°_1(X)...Q°_r(X),$$

où δ est le PGCD des coefficients du polynôme $P(X,\alpha_1,...,\alpha_m)$. Malheureusement la variante proposée par Wang exige que la factorisation ci-dessus ne contienne aucun facteur parasite. Supposons pour la suite du paragraphe, que ce soit le cas. Cette méthode exige aussi de choisir $\alpha_1,..,\alpha_m$ de façon qu'ils vérifient une condition supplémentaire, ce qui augmente le coût de la phase d'initialisation de l'algorithme de factorisation.
Dans tout le paragraphe, on utilisera les notations suivantes :

1) on appelle **coefficient dominant** de $P(X,X_1,...,X_m)$ et on note $a_o(X_1,...,X_m)$, le coefficient de la plus grande puissance de X;
2) ce coefficient $a_o(X_1,...,X_m)$ s'écrit d'une part sous la forme $a_o(X_1,...,X_m) = c_1(X_1,...,X_m)...c_r(X_1,...,X_m)$ où $c_1(X_1,...,X_m),...,c_r(X_1,...,X_m)$ désignent les coefficients dominants des diviseurs de P;
3) d'autre part, $a_o(X_1,...,X_m)$ s'écrit sous la forme $a_o(X_1,...,X_m) = ß\ b_1(X_1,...,X_m)^{N1}....b_t(X_1,...,X_m)^{Nt}$ où les polynômes $b_j(X_1,...,X_m)$ sont irréductibles et primitifs et où $ß \in Z$;

Condition préliminaire.

Comme on l'a dit plus haut, on doit choisir $\alpha_1,..,\alpha_m$ de façon plus restrictive que dans la méthode classique. Plus précisément $\alpha_1,...,\alpha_m$ doivent être choisis pour que la condition suivante soit vérifiée :

(B.II.W) la suite des nombres $b_j(\alpha_1,...,\alpha_m)$ est telle que, pour $1 \le j \le t$, $b_j(\alpha_1,...,\alpha_m)$ ait un facteur premier p_j qui ne divise ni δ, ni $ß$, ni aucun des nombres $b_k(\alpha_1,...,\alpha_m)$ pour $k < j$.

Dans la suite, cette condition sera appelée **condition de Wang**. On peut vérifier facilement qu'il existe une infinité de m−uples $(\alpha_1,...,\alpha_m)$ qui vérifient la condition de Wang. En effet, comme les polynômes $b_1(X_1,...,X_m)$, ..., $b_j(X_1,...,X_m)$ sont premiers entre eux, l'identité de Bezout s'écrit $ß_1\ b_1(X_1,...,X_m) + ... + ß_j\ b_j(X_1,...,X_m) = \sigma_j \in Z$.

Faisons varier $X_1 = \alpha_1$, ..., $X_m = \alpha_m$ dans Z. Supposons que le nombre $b_j(\alpha_1,...,\alpha_m)$ n'ait qu'un nombre fini de facteurs premiers. Dans ce cas, il existerait un facteur premier p tel que $b_j(\alpha_1,...,\alpha_m) = 0$ sur $Z/(p)$ pour $(\alpha_1,...,\alpha_m) \in E$, E étant un pavé vérifiant les conditions du Lemme 1. D'après ce Lemme le polynôme $b_j(X_1,...,X_m)$ serait nul sur $Z/(p)[X_1,...,X_m]$, il ne serait donc pas primitif sur $Z[X_1,...,X_m]$, ce qui serait contraire aux hypothèses. Il en résulte que $b_j(\alpha_1,...,\alpha_m)$ a un nombre infini de facteurs premiers lorsque $\alpha_1,...,\alpha_m$ varient dans Z.

En retirant de ces facteurs premiers les diviseurs de $\sigma_1,...,\sigma_{t-1}$, $ß$ et δ, il reste une infinité de facteurs ne divisant aucun des nombres $b_k(\alpha_1,...,\alpha_m)$ pour $k < t$. On a donc finalement une infinité le choix possibles pour le m−uple $(\alpha_1,...,\alpha_m)$.

L'algorithme de recherche d'un tel m−uple qui est donné par Wang dans **[Wa]**, est fondé sur des calculs de PGCD, de façon à éviter des factorisations d'entiers. Dans cet algorithme, qui est donné ci−dessous, la recherche des facteurs premiers p_j est

remplacée par celle de facteurs quelconques $d_j > 1$. Si une telle suite $(d_o, d_1, .., d_t)$ vérifie la condition de Wang, alors on obtient une suite $(p_o, p_1, .., p_t)$ en choisissant des facteurs premiers autres que 1, dans la suite $(d_o, d_1, .., d_t)$.

On donne ci-dessous les étapes de l'algorithme vérifiant si un m-uple donné $(\alpha_1, ..., \alpha_m)$ vérifie la condition de Wang :

Algorithme:

1) $d_o := ß\ \delta$;
2) pour $j := 1, 2, ..., t$, faire :
   a) $d_j := b_j(\alpha_1, ..., \alpha_m)$;
   b) on divise d_j par les facteurs communs à $d_o, .., d_j$;
      soit : pour $k := j-1, .., 0$:
         supprimer de d_j tout facteur figurant dans d_k
         pour cela :
         Tant que $\delta = PGCD(d_k, d_j) > 1$, diviser d_j par δ;
      Si d_j est réduit à 1, Aller à 4);
3) Fin avec la réponse VRAI et la suite $\{d_o, ..., d_t\}$;
4) Fin avec la réponse FAUX.

Détermination des coefficients dominants des diviseurs.

Cette détermination est fondée sur le lemme suivant donné par Wang.

Lemme 2.

Considérons le produit suivant $P(X, X_1, .., X_m) = Q_1(X, X_1, .., X_m) ... Q_r(X, X_1, .., X_m)$, et $\delta\ Q^\circ_1(X) ... Q^\circ_r(X)$, la factorisation de $P(X, \alpha_1, ..., \alpha_m)$ en facteurs irréductibles primitifs, avec $\delta \mid Z$. Notons $C(X_1, ..., X_m)$ le coefficient dominant de $P(X, X_1, .., X_m)$.

Alors, on a d'une part $C(\alpha_1, ..., \alpha_m) = \delta\ q_1 ... q_r$ où les nombres $q_1, ..., q_r$ sont les coefficients dominants de $Q^\circ_1, .., Q^\circ_r$, et d'autre part,

$$C(X_1, ..., X_m) = c_1(X_1, ..., X_m) .... c_r(X_1, ..., X_m)$$

où les $c_j(X_1, ..., X_m)$ sont les coefficients dominants des diviseurs $Q_j(X, X_1, ..., X_m)$ de P. Notons $ß\ b_1(X_1, ..., X_m)^{N1} .... b_t(X_1, ..., X_m)^{Nt}$, la factorisation de $C(X_1, ..., X_m)$ en facteurs irréductibles primitifs, avec $ß \in Z$.

Soit $(\alpha_1, ..., \alpha_m)$ un m-uple tel que, pour $1 \le j \le t$, $b_j(\alpha_1, ..., \alpha_m)$ ait un facteur premier p_j qui ne divise ni δ, ni ß, ni aucun des nombres $b_k(\alpha_1, ..., \alpha_m)$ pour $k < j$.

Alors le dernier facteur de C élevé à la puissance N, à savoir $b_t(X_1, ..., X_m)^N$, divise $c_j(X_1, ..., X_m)$, si et seulement si $b_t(\alpha_1, ..., \alpha_m)^N$ divise $\delta\ q_j$.

Démonstration.

Il est clair que si $b_t(X_1,...,X_m)^N$ divise $c_j(X_1,...,X_m)$, alors $b_t(\alpha_1,...,\alpha_m)^N$ divise $c_j(\alpha_1,...,\alpha_m)$, qui lui-même divise $\delta\, q_j$.

Supposons qu'au contraire, $b_t(X_1,...,X_m)^N$ ne divise pas $c_j(X_1,...,X_m)$. Alors $c_j(\alpha_1,...,\alpha_m)$ est de la forme $ß_j\, b_1(\alpha_1,...,\alpha_m)^{U1}...b_t(\alpha_1,...,\alpha_m)^{Ut}$ avec $ß_j$ qui divise $ß$ et $Ut < N$. Par suite, le facteur premier p_t défini dans le lemme ne figure dans $c_j(\alpha_1,...,\alpha_m)$ qu'avec la puissance Ut. Comme δ ne contient pas le facteur p_t, alors p_t ne peut pas non plus figurer dans $\delta\, q_j$ avec la puissance N. Par suite, $b_t(\alpha_1,...,\alpha_m)^N$ ne peut pas diviser $\delta\, q_j$. Le lemme est donc vérifié.

Le lemme 2 permet de répartir les puissances de $b_t(X_1,...,X_m)$ entre les différents coefficients dominants $c_j(X_1,...,X_m)$ des diviseurs de P. Lorsque cette répartition est terminée, on applique à nouveau le lemme 2 au produit des coefficients dominants, divisé par $b_t(X_1,...,X_m)^{Nt}$. On obtient ainsi la répartition des puissances de $b_{t-1}(X_1,...,X_m)$. En continuant ce processus, on obtient la répartition des puissances de tous les $b_j(X_1,...,X_m)$.

Par suite on peut écrire chaque coefficient $c_j(X_1,...,X_m)$ sous la forme

$c_j(X_1,...,X_m) = \mu_j\, b_1(X_1,...,X_m)^{U1}...b_t(X_1,...,X_m)^{Ut} = \mu_j\, b(X_1,...,X_m)$,

où les puissances U1, .., Ut sont déterminées. Il reste à définir les coefficients numériques μ_j.

Supposons que $\delta = 1$. Alors $C(\alpha_1,...,\alpha_m) = q_1...q_r = c_1(\alpha_1,...,\alpha_m)....c_r(\alpha_1,...,\alpha_m)$,

donc $c_j(\alpha_1,...,\alpha_m) = q_j$. On obtient donc $\mu_j = q_j / b(\alpha_1,...,\alpha_m)$.

Supposons maintenant que $\delta > 1$. Alors $\delta\, q_1...q_r = c_1(\alpha_1,...,\alpha_m)....c_r(\alpha_1,...,\alpha_m)$,

donc $c_j(\alpha_1,...,\alpha_m) = \delta_j\, q_j = \mu_j\, b(\alpha_1,...,\alpha_m)$ est le PPCM de q_j et $b(\alpha_1,...,\alpha_m)$. Par suite, $\mu_j = q_j / d_j$, où d_j est le PGCD de q_j et de $b(\alpha_1,...,\alpha_m)$.

Remarquons que s'il existe des facteurs parasites, alors le lemme ne permet pas toujours de répartir toutes les puissances des $b_j(X_1,...,X_m)$. Dans ce cas, on ne peut pas trouver les coefficients dominants $c_j(X_1,...,X_m)$ des diviseurs de P. On devra alors choisir un autre m-uple $(\alpha_1,...,\alpha_m)$, puis recommencer jusqu'à ce qu'on puisse trouver les coefficients $c_j(X_1,...,X_m)$ à l'aide du lemme. Dans ce dernier cas, on n'est toujours pas certain qu'il n'y a pas de facteurs parasites.

Exemple.

Reprenons l'exemple donné par Wang, dans [Wa].

$$P(X,Y,Z) = (4Y^4Z^2 + 4Y^3Z^3 - 4Y^2Z^4 - 4YZ^5)\,X^6 +$$
$$(Y^4Z^3 + 12Z^3 + 12Y^2Z^2 - Y^2Z^5 - 12YZ^3 - 12Z^4)\,X^5 +$$
$$(8Y^4 + 6Y^3Z^2 + 8Y^3Z - 4Y^2Z^4 + 4Y^2Z^3 - 8Y^2Z^2 - 4YZ^5 - 2YZ^4 - 8YZ^3)\,X^4 +$$
$$(2Y^4Z + Y^3Z^3 - Y^2Z^5 + 9Y^2Z - 12YZ^3 + 12YZ^2 - 12Z^4 + 3Z^3)\,X^3 +$$
$$(6Y^3 - 6Y^2Z^2 + 8Y^2Z - 2YZ^4 - 8YZ^3 + 2YZ^2)\,X^2 +$$
$$(2Y^3Z - 2Y^2Z^3 - 3YZ + 3Z^3)\,X + 2YZ^2 - 2Y^2.$$

Le coefficient dominant $C(Y,Z)$ se factorise sous la forme

$$C(Y,Z) = 4YZ^2(Y+Z)^2(Y-Z) = 4\,b_1\,b_{22}\,b_{32}\,b_4.$$

Par suite $C(-14,3) = 4\,(-14)\,3^2\,(-11)^2\,(-17)$.

La recherche de couples (α_1, α_2) satisfaisant la condition de Wang donne par exemple $(-14, 3)$, $(5,-12)$ et $(-23, 3)$. On utilise trois couples pour vérifier que $P(X,\alpha_1,\alpha_2)$ donne chaque fois le même nombre de facteurs, ce qui est bien le cas. On en conclut qu'il ne doit pas y avoir de facteurs parasites. Avec le premier couple $\alpha_1 = -14$, $\alpha_2 = 3$, on obtient

$$P(X,-14,3) = 1036728\,X^6 + 915552\,X^5 + 55748\,X^4 + 105621\,X^3 - 17304\,X^2 - 26841\,X - 644$$
$$= (187\,X^2 - 23)(44\,X^2 + 42\,X + 1)(126\,X^2 - 9\,X + 28)$$
$$= Q^\circ_1(X).Q^\circ_2(X).Q^\circ_3(X).$$

On a donc $C(Y,Z) = c_1(Y,Z)\,c_2(Y,Z)\,c_3(Y,Z)$. D'autre part, $C(\alpha_1,\alpha_2) = \delta\,q_1\,q_2\,q_3 = 1\cdot187\cdot44\cdot126$. On place d'abord le facteur $b_4 = Y - Z$ qui vaut -17 en (α_1,α_2). Comme 17 divise 187, b_4 divise $c_1(Y,Z)$. On place ensuite $b_3 = Y + Z$ qui vaut -11 en (α_1,α_2). Comme 11 divise 187 et 44, b_3 divise c_2 et c_3. On obtient finalement :

$c_1(Y,Z) = (Y+Z)(Y-Z)$, $c_2(Y,Z) = -4(Y+Z)$ et $c_3(Y,Z) = -YZ^2$,

qui sont les coefficients dominants de $Q_1(X,Y,Z)$, $Q_2(X,Y,Z)$ et $Q_3(X,Y,Z)$.

3°) Transformation de P en polynôme unitaire.

On propose ci-dessous une alternative à l'amélioration de Wang, qui ne dépend pas de l'existence ou non de facteurs parasites.

La solution consiste à transformer P en polynôme unitaire; alors les diviseurs de P sont également unitaires. Tous les problèmes posés par les coefficients dominants sont donc résolus. Introduisons quelques notations. Si Q est un polynôme de $Z[X,X_1,...,X_m]$, son degré suivant X est noté δQ. D'autre part, le degré suivant X du polynôme donné P est noté d, son degré total suivant $X_1, ..., X_m$ est noté do et son degré total suivant $X, X_1, .., X_m$ est noté Do. Le polynôme P peut s'écrire sous la forme $P = P^{(o)} + P^{(1)} + ... + P^{(j)} + ... + P^{(do)}$ où $P^{(j)}$ est une forme de degré j de $K[X_1,...,X_m]$, où $K = Z[X]$. Alors $j \leq \delta P^{(j)} \leq j + d$.

Rappelons que l'algorithme classique de factorisation défini dans l'Introduction consiste à factoriser successivement la suite des polynômes P_j, où les P_j sont les projections de P sur $K_j = Z[X_1-\alpha_1,...,X_m-\alpha_m]/\Delta_j$ où Δ_j est l'idéal engendré par p^N et par les monômes de $Z[X_1-\alpha_1,...,X_m-\alpha_m]$ de degré strictement supérieur à j; p^N étant supérieur aux coefficients des diviseurs de P. Pour simplifier l'écriture de P_j on utilise habituellement le changement de variables suivant $Y_1 = X_1 - \alpha_1, ..., Y_m = X_m - \alpha_m$.

Alors P_j s'écrit comme un polynôme en X dont les coefficients sont des polynômes en $Y_1,..,Y_m$ de degré total majoré par j et dont les coefficients sont compris entre $-\frac{1}{2}p^N$ et $\frac{1}{2}p^N$.

Remarquons que le polynôme $P(X,X_1,..,X_m)/\Delta_1$, qui est noté comme auparavant $P_o(X,X_1,...,X_m)$ s'écrit $P(X,\alpha_1,...,\alpha_m)$. On a vu ci-dessus que $\alpha_1, .., \alpha_m$ doivent être choisis de façon que la propriété suivante soit vérifiée :

(B.II.1) $P_o(X,X_1,..,X_m)$ n'a pas de facteur carré et son degré est égal à d.

Considérons un autre changement de variables, à savoir :

$$Y_1 = X_1 - \alpha_1 X, ..., Y_m = X_m - \alpha_m X.$$

Alors $P(X,X_1,...,X_m)$ est le polynôme $P(X,Y_1+\alpha_1 X,...,Y_m+\alpha_m X)$ noté pour la suite $P_N(X,Y_1,...,Y_m)$ (car un tel polynôme sera dit **normalisé**). La projection P_j de P sur $K_j = Z[X_1-\alpha_1,...,X_m-\alpha_m]/\Delta_j$ donne pour j = 0, le polynôme $P_o(X,X_1,...,X_m)$ qui est égal à $P(X,\alpha_1 X,..,\alpha_m X)$. La propriété **(B.II.1)** signifie donc que $P(X,\alpha_1 X,..,\alpha_m X)$ n'a pas de facteur carré et que son degré est égal à Do.

Vérifions le lemme suivant :

Lemme 3.

Considérons le polynôme $P(X,X_1,...,X_m)$ et le changement de variables suivant :

$$Y_1 = X_1 - \alpha_1 X, \ldots, Y_m = X_m - \alpha_m X.$$

1) Il n'existe qu'un nombre fini de m-uples $(\alpha_1,\ldots,\alpha_m)$ tel que **(B.II.1)** ne soit pas vérifié.

2) Notons $P_N(X,X_1,\ldots,X_m)$ le polynôme unitaire transformé de $P(X,X_1,\ldots,X_m)$ par un tel changement de variable vérifiant **(B.II.1)**. Alors P_N vérifie l'inégalité suivante:

(B.II.2) $\delta P_N^{(k)} \leq \delta P_N^{(o)} - k$ pour $k \geq 0$.

En particulier $P_N^{(k)} = 0$ pour $k \geq \delta P_N^{(o)}$.

Dans la suite, un polynôme Q de $Z[X,X_1,\ldots,X_m]$ vérifiant l'inégalité $\delta Q^{(k)} \leq \delta Q^{(o)} - k$ pour $k \geq 0$ et $\delta Q^{(k)} = 0$ pour $k \geq \delta Q^{(o)}$ est appelé polynôme **normalisé.**

Démonstration.

1) On a $\delta P(X,\alpha_1 X,\ldots,\alpha_m X) = Do$, à condition que les monômes de degré Do ne s'annulent pas entre eux. Notons $A_o(\alpha_1,\ldots,\alpha_m)$ la somme des coefficients de $P(X,\alpha_1 X,\ldots,\alpha_m X)$ de degré Do. Alors $P(X,\alpha_1 X,\ldots,\alpha_m X)$ a le même degré en X que $P(X,X_1+\alpha_1 X,\ldots,X_m+\alpha_m X)$ à condition que $A_o(\alpha_1,..,\alpha_m) \neq 0$.

D'autre part, notons $\delta_o(\alpha_1,\ldots,\alpha_m)$ le discriminant du polynôme $P(X,\alpha_1 X,\ldots,\alpha_m X)$. La propriété **(B.II.1)** sera donc vérifiée si le m-uple $(\alpha_1,\ldots,\alpha_m)$ est choisi de façon que le produit $A_o(\alpha_1,\ldots,\alpha_m).\delta_o(\alpha_1,\ldots,\alpha_m)$ soit non nul. D'après le lemme 1 il n'y a donc qu'un nombre fini de m-uples $(\alpha_1,\ldots,\alpha_m)$ tels que **(B.II.1)** ne soit pas vérifié.

2) Supposons maintenant que $A_o(\alpha_1,\ldots,\alpha_m)$ ne soit pas nul.

Considérons le polynôme $P_N(X,X_1,\ldots,X_m)$ égal à $P(X, X_1+\alpha_1 X,\ldots,X_m+\alpha_m X)$; alors $\delta P_N^{(o)} = \delta P(X, \alpha_1 X,\ldots, \alpha_m X) = Do$; d'autre part, P_N est égal à une somme de polynômes homogènes de la forme $m_j = \mu_j X^a (X_1+\alpha_1 X)^{b1}) \ldots (X_m+\alpha_m X)^{bm})$.

Notons $m_j^{(k)}$ la forme de m_j de degré k suivant $X_1,\ldots,X_m$. On a donc la relation $\delta m_j = \delta m_j^{(k)} + k$. Alors $\delta P_N^{(o)} = \delta(> m_j) \geq \delta(> m_j^{(k)}) + k$. De plus il n'y a dans P_N aucun monôme de degré supérieur à $\delta P_N^{(o)}$ suivant $X_1,\ldots,X_m$. Donc P_N est normalisé et le lemme est vérifié.

Coût des calculs pour rendre P unitaire.

Remarquons que le nombre d'essais pour trouver un m-uple $(\alpha_1,\ldots,\alpha_m)$ vérifiant le Lemme 3 est du même ordre de grandeur que celui des essais de la recherche classique

d'un m−uple tel que $P(X,\alpha_1,...,\alpha_m)$ soit sans facteur carré et ait un coefficient dominant $a_o(\alpha_1,..,\alpha_m)$ non nul.

Rappelons que dans le recherche classique du m−uple $(\alpha_1,...,\alpha_m)$, on fait des essais en prenant le maximum de nombres α_j nuls. On procède donc de la même manière ici, c'est−à−dire qu'on prend d'abord $\alpha_1 = ... = \alpha_m = 0$, puis on prend l'un des nombres α_j non nul, puis deux nombres α_j non nuls, ... Le coût de chaque essai est celui du calcul des coefficients du nouveau polynôme

$P_N(X,X_1,...,X_m) = P(X, X_1+\alpha_1 X,...,X_m+\alpha_m X)$, c'est−à−dire du calcul des produits $(X_1 + \alpha_1 X)...(X_m + \alpha_m X)$. Le nombre de ces produits, c'est−à−dire le nombre des coefficients de P est de l'ordre de

(B.II.3) $$\frac{(d+1)^{2m}}{(2m)!}.$$

Elimination du coefficient constant a_o.

Supposons dans la suite que $P(X,X_1,...,X_m)$ ait été normalisé par un changement de variable de la même forme que dans le lemme 3. Alors le coefficient dominant a_o de $P(X,X_1,...,X_m)$ est constant.

Si on choisit p non diviseur de a_o, on peut mettre a_o en facteur dans P et déterminer les diviseurs unitaires de P/a_o.

On trouvera ensuite le coefficient dominant de tout diviseur Q de P de la même manière que pour les polynômes à une variable, c'est−à−dire en rendant primitif le polynôme a_o Q.

On peut aussi rendre P unitaire par un nouveau changement de variable, en remplaçant par exemple X par X/a_o, puis en multipliant P par a_o^{d-1}. Si d et a_o sont petits, alors ce nouveau changement de variable ne fait pas beaucoup croître la taille des coefficients de P.

Dans la suite, on supposera que P est normal et unitaire.

Exemple.

$$P = 6X^4 + 2X^3 + 5X^2 - 4 + 2X_1(3X^2 - 2X) + X_2(2X^4 + 2X^3 + 3X^2 + 5X - 2) + X X_2^2 + 2X^2 X_1 X_2.$$

Calculons la somme des monômes de degré total maximum, à savoir

$A_o(X,X_1,X_2) = 2\ X^4\ X_2$. On peut donc prendre ici $\alpha_1 = 0$ et $\alpha_2 = \pm 1$. Le nouveau polynôme P_N s'écrit donc par exemple

$$P_N(X,X_1,X_2) = P(X,X_1,X_2 - X) =$$
$$-2X^5 + 4X^4 + 2X - 4 + 2X_1(-2X^3 + 3X^2 - 2X)$$
$$+ X_2(2X^4 + 2X^3 + X^2 + 5X - 2) + X X_2^2 + 2X^2 X_1 X_2.$$

On peut dans cet exemple, rendre P_N unitaire sans augmenter réellement la taille des coefficients, en changeant X_2 en $2X_2$, puis en divisant le polynôme obtenu par -2. On obtient le polynôme :

$$P^* = X^5 - 2X^4 - X + 2 + X_1(2X^3 - 3X^2 + 2X)$$
$$- X_2(2X^4 + 2X^3 + X^2 + 5X - 2) - 2X X_2^2 - 2X^2 X_1 X_2.$$

La factorisation de P revient donc à celle de P^*. Notons Φ l'application de $\mathbf{Z}[X,X_1,X_2]$ dans $\mathbf{Z}[X,X_1,X_2]$, définie par les changements de variables ci-dessus, qui transforme P en P^*. Après avoir factorisé P^* sur $\mathbf{Z}[X,X_1,X_2]$, on devra transformer chacun des diviseurs Q^* de P^*, à l'aide de l'application réciproque de Φ, notée Φ^{-1}. Dans cet exemple, Φ^{-1} est définie par le changement de la variable X_2 en $\frac{1}{2}X_2$, suivi d'une multiplication du polynôme obtenu par 2, suivi du changement de X_2 en $X_2 + X$.

4°) Calcul des coefficients d'un produit de polynômes.

Au cours des algorithmes de factorisation on doit souvent calculer les coefficients d'un polynôme défini comme un produit. Pour connaître les coefficients, on peut évidemment développer ce produit, mais pour les polynômes à plusieurs variables les calculs sont très longs. Le développement du produit donne en fait, tous les coefficients du polynôme, alors que souvent on n'a besoin que des coefficients de degrés donnés.

La méthode utilisée, qui a été proposée par Wang dans **[Wa]**, repose sur la formule de Taylor. D'après cette formule le coefficient du polynôme P de degrés j, j1,..., jm suivant $X, X_1,..., X_m$ est obtenu en dérivant P, j fois suivant X, j1 fois suivant X_1,..., jm fois suivant X_m, puis en remplaçant $X_1, .., X_m$ par 0 dans la dérivée ainsi obtenue, et enfin en divisant le nombre obtenu par le produit j! j1! ...jm!.

Cette méthode est intéressante car les calculs relatifs à la dérivation et à l'évaluation d'une expression en un point sont plus rapides que les calculs de développement d'un produit.

Exemple.

On considère le polynôme défini par le produit suivant :

$$P(X,Y) = (X^2\,Y + X^2 - 4\,X + 4)\,(X + X\,Y + 1).$$

On veut calculer le coefficient de $X^2\,Y$ dans P, sans effectuer le produit. Pour cela on calcule la dérivée :

$$\frac{1}{2!}\,\frac{\partial^3 P}{\partial X^2\,\partial Y}(0,\,0) = \tfrac{1}{2}\,Q(0,0) \quad \text{où} \quad Q(X,Y) = 6X + 12XY - 6 + 6X.$$

Le coefficient cherché vaut donc -3.

5°) Algorithme d'Euclide généralisé.

On se place sur l'anneau $\mathbf{K}[X]$ où $\mathbf{K} = \mathbf{Z}[X_1,...,X_m]$. Considérons des polynômes de $\mathbf{K}[X]$, notés $Q_1,...,Q_r$ dont le PGCD vaut 1 et soit P un autre polynôme de $\mathbf{K}[X]$.
On note dans ce paragraphe, $d°Q$ le degré de Q suivant $X_1,...,X_m$ et $\partial_j Q$ le degré de Q suivant X_j.
On se propose de déterminer des polynômes $A_1,..,A_r$ de $\mathbf{K}$ tels que

$$\textbf{(B.II.4)} \qquad \begin{cases} A_1 \cdot Q_2 \cdots Q_r + A_2 \cdot Q_1 \cdot Q_3 \cdots Q_r + \ldots + A_r \cdot Q_1 \cdots Q_{r-1} = P, \\ \text{avec } d°\,A_j < d°\,Q_j \text{ pour } 1 \le j \le r. \end{cases}$$

Le problème est possible, car pour tout couple de polynômes A et B de $\mathbf{K}[X]$, il existe deux polynômes R et S de degrés majorés par $d°B$ et $d°A$, tels que $A \cdot S + B \cdot T = \mathrm{res}(A, B)$, où $\mathrm{res}(A, B)$ est le résultant de A et B (Voir par exemple, Collins **[C2]**, p. 516).
La méthode présentée ci-dessous est celle donnée par Wang dans son article **[Wa]**. C'est une méthode qui est récursive, à la fois sur le degré des variables $X_1,...,X_m$ et sur le nombre r de polynômes Q_j.
On initialise l'algorithme en considérant les deux membres de **(B.II.4)** comme des polynômes de $\mathbf{K}_o[X]$, où $\mathbf{K}_o = \mathbf{K}/\Delta_{(0\,0..0)}$, $\Delta_{(0\,0..0)}$ étant l'idéal engendré par les éléments de $\mathbf{K}$, qui sont des monômes en $X_1, .., X_m$ de degrés strictement positifs. Les deux

membres de l'équation **(B.II.4)** se réduisent donc à des polynômes de $Z[X]$. Les projections des polynômes A_k, Q_k, P de (B.II.4) sur $Z[X]$ sont notés $A°_k$, $Q°_k$ et $P°$.

On peut alors résoudre cette équation **(B.II.14)** en utilisant l'algorithme d'Euclide généralisé. Plus précisément, on définit une suite de couples de polynômes $(A°_1, B°_1)$, $(A°_2, B°_2)$, ..., $(A°_{r-1}, A°_r)$ vérifiant respectivement les équations suivantes

(B.II.5)
$$\begin{cases} A°_1 \cdot Q°_2 \cdots Q°_r + B°_1 \cdot Q°_1 = P°, \\ A°_2 \cdot Q°_3 \cdots Q°_r + B°_2 \cdot Q°_2 = B°_1, \\ \cdots\cdots\cdots\cdots \quad \cdots\cdots\cdots\cdots \\ A°_{r-1} \cdot Q°_r + A°_r \cdot Q°_{r-1} = B°_{r-1}. \end{cases}$$

Chacune des équations ci-dessus peut être résolue avec l'algorithme classique d'Euclide; on peut par exemple utiliser l'algorithme de Knuth défini dans **[Kn]**. Les polynômes $A°_1$, ..., $A°_r$ obtenus vérifient **(B.II.4)** sur $K_1[X]$. Supposons qu'on ait construit les polynômes A^j_1,.., A^j_r qui vérifient **(B.II.4)** sur $K_j[X]$, où $K_j = K/\Delta_{(u\ v..w)}$, l'idéal $\Delta_{(u\ v..w)}$ étant engendré par les monômes de K de degrés suivants X_1, .., X_m respectivement supérieurs à u, v, .., w, avec $u + v + .. + w = j$. On a donc la relation

(B.II.6) $$A^j_1 \cdot Q_2 \cdots Q_r + A^j_2 \cdot Q_1 \cdot Q_3 \cdots Q_r + \ldots + A^j_r \cdot Q_1 \cdots Q_{r-1} = P.$$

qui est vérifiée sur $K_j[X]$. On veut définir les polynômes A^{j+1}_1,..., A^{j+1}_r, qui vérifient **(B.II.4)** sur $K_{j+1}[X]$, où par exemple, $K_{j+1} = K/\Delta_{(u+1\ v..w)}$. On a donc pour tout k, $A^{j+1}{}_k = A^j_k + a_k X_1^u X_2^{v-1} \ldots X_m^{w-1}$. En reportant ces valeurs dans **(B.II.6)**, puis en réduisant on obtient la relation

(B.II.7) $$a_1 Q_2 \cdots Q_r + a_2 Q_1 \cdot Q_3 \cdots Q_r + \ldots + a_r Q_1 \cdots Q_{r-1} = b$$

où b est le coefficient dans K_o du monôme $X_1^u X_2^{v-1} \ldots X_m^{w-1}$ du polynôme $D = P - A^j_1 \cdot Q_2 \cdots Q_r + A^j_2 \cdot Q_1 \cdot Q_3 \cdots Q_r + \ldots + A^j_r \cdot Q_1 \cdots Q_{r-1}$.

Pour obtenir les coefficients a_k il suffit d'utiliser **(B.II.7)** sur K_o. On obtient donc les coefficients a_k en résolvant les équations **(B.II.5)** pour $P° = b$. Pour le calcul de b, aucun développement n'est nécessaire; on utilise la méthode du paragraphe précédent pour obtenir le coefficient du monôme $X_1^u X_2^{v-1} \ldots X_m^{w-1}$ dans D en calculant la dérivée de D à l'ordre u+1 suivant X_1,... et à l'ordre w suivant X_m.

Le coût de cet algorithme est donc proportionnel au nombre de monômes

$\mu = b\, X_1^{u} X_2^{v-1} \ldots X_m^{w-1}$ de D. Comme les degrés des monômes μ sont égaux à j, alors le nombre total de monômes μ est majoré par $\binom{j+m}{m}$, qui est majoré par 2^{m+j} ou par $(j + m)^m$, m étant en général petit par rapport à j. Si d est le degré moyen de P suivant $X_1,\ldots,X_m$, alors le coût total est proportionnel à $m\, d\, (d + m)^m$. Chaque dérivation coûte la division d'un coefficient, donc le coût maximum est de $(d + 1)^m$. Le coût général est donc

(B.II.8) $m\, d\, (d + 1)^{2m}$.

III) METHODE CLASSIQUE DE FACTORISATION.

L'algorithme originel donné en même temps par Wang et Rothschield dans **[WR]** et par Musser dans **[Mu]** a bénéficié d'améliorations importantes qui ont été données par Wang dans **[Wa]**.

On a étudié en **(B.I)** l'une des améliorations qui consiste à calculer les coefficients dominants des diviseurs. Ce nouvel algorithme dû à Wang permet en pratique d'éliminer les problèmes posés par la factorisation des polynômes non unitaires.

La méthode originelle qui consistait à mettre en facteur le coefficient dominant conduisait à des calculs importants pour retrouver les vrais diviseurs de P à partir des diviseurs trouvés qui sont définis modulo Δ_j.

On proposera plus loin, dans **B.IV** une alternative à la méthode originelle de Wang, consistant à rendre le polynôme unitaire, ce qui évite les calculs complexes de cette première méthode.

Dans ce paragraphe, on présente la méthode améliorée de Wang, en supposant que les coefficients dominants des diviseurs ont déjà été déterminés.

Rappelons que pour calculer ces coefficients dominants, on a dû définir un m-uple $(\alpha_1,...,\alpha_m)$ tel que le polynôme $P(X,\alpha_1,..,\alpha_m)$ soit sans facteur carré et de même degré d que P suivant X. On considère ensuite le changement de variable

$$Y_1 = X_1 - \alpha_1, \ ..., \ Y_m = X_m - \alpha_m,$$

et dans la suite on utilise les nouvelles variables $Y_1,...,Y_m$ pour simplifier les notations, mais, à aucun moment, on n'aura besoin de développer le polynôme $P(X,\alpha_1+Y_1,...,\alpha_m+Y_m)$. Les valeurs de coefficients dont on aura besoin seront obtenues par des dérivations successives, comme on l'a montré au **I.4°)**.

La processus général de la méthode consiste à itérer pour j variant de 1 à m, l'algorithme suivant :

1°) On suppose connue la factorisation du polynôme

$P_{j-1} = P(X,\alpha_1+Y_1,...,\alpha_{j-1}+Y_{j-1},\alpha_j,\alpha_{j+1},...,\alpha_m)$ sur $Z[X,Y_1,...,Y_{j-1}]$;
cette factorisation ayant été obtenue en utilisant la récursivité de l'algorithme;
au départ, pour $j = 1$, P_0 s'écrit $P(X,\alpha_1,...,\alpha_m)$ et il est factorisé comme un polynôme à une variable;

2°) La factorisation obtenue à l'étape **1°)** est considérée comme la factorisation de P_j sur $Z[X,Y_1,...,Y_j]/(Y_j)$, (Y_j) désignant l'idéal engendré par Y_j;
On détermine ensuite la factorisation de P_j modulo Y_j^k, pour $k = 2, .., dk+1$, dk étant le degré de P suivant Y_j ;

3°) A partir de la factorisation de P_j modulo Y_j^{dk+1}, obtenue au 2°), et notée $Q_1 \cdots Q_r$, on déduit la factorisation de P_j sur $Z[X,Y_1,...,Y_j]$.

Dans tout le paragraphe, on note $K[X_j]$ l'anneau des polynômes en X_j à coefficients dans $Z[X,X_1,...,X_{j-1}]$, et le polynôme P_j considéré comme un polynôme de $K[X_j]$ sera noté simplement $B(X_j)$ et s'écrira $B(X_j) = b_0 + b_1 X_j + b_2 X_j^2 + ... + b_\delta X_j^\delta$.
Les diviseurs de B sur $K[X_j]$ seront notés de manière analogue, à savoir $A_1(X_j)$, ..., $A_r(X_j)$. Les coefficients du diviseur $A_k(X_j)$ seront notés $a_{k0},...,a_{kn}$. La factorisation cherchée de $B(X_j)$ s'écrit

(B.III.1)
$$\begin{cases} b_0 + b_1 X_j + b_2 X_j^2 + ... + b_\delta X_j^\delta = \\ (a_{10} + a_{11} X_j + a_{12} X_j^2 + ... + a_{1u} X_j^u)... \\ ... (a_{r0} + a_{r1} X_j + a_{r2} X_j^2 + ... + a_{rv} X_j^v). \end{cases}$$

Cette égalité est équivalente au système de relations suivantes:

(B.III.2)
$$\begin{cases} a_{10}\, a_{20}....a_{r0} = b_0 \\ a_{11}\, a_{20}...a_{r0} + a_{10}\, a_{21}\, a_{30}...a_{r0} + ... = b_1 \\ \quad ..... \\ a_{1k}\, a_{20}...a_{r0} + a_{10}\, a_{2k}\, a_{30}...a_{r0} + ... = b^\circ_k \\ \quad ..... \end{cases}$$

où

(R) $$b^\circ_k = b_k - \sum_{k1+...+kr=k} a_{1k1}\, a_{2k2}....a_{rkr}$$

La première égalité représente la factorisation de b_0 sur $K = Z[X,X_1,...,X_{j-1}]$. Cette factorisation est réalisée grâce au processus récursif de l'algorithme. C'est l'étape initiale

qui définit les termes $a_{1\,0},\ldots,\ a_{r\,0}$ qui sont les projections des diviseurs $Q_1,\ldots,Q_r$ sur **K**.
A la $k^{ième}$ étape on détermine les termes $a_{1\,k},\ldots,\ a_{r\,k}$, grâce à l'équation

$$a_{1\,k}\,a_{2\,0}\ldots a_{r\,0} + a_{1\,0}\,a_{2\,k}\,a_{3\,0}\ldots a_{r\,0} + \ldots = b^{\circ}_{k}$$

où b°_{k} est défini en fonction des termes des diviseurs précédemment déterminés, mais le polynôme b°_{k} n'est pas développé.
Les éléments $a_{1\,0},\ldots,\ a_{r\,0}$ de l'anneau **K** sont définis par un algorithme dérivé de celui d'Euclide. Cet algorithme a été défini en **B.I.5°**. C'est sans doute l'étape la plus longue de l'algorithme.
Le coût de cette étape est dominé par les calculs des dérivées successives. Si le nombre de dérivations atteint la taille maximum possible du polynôme, à savoir $(d+1)^m$, alors le coût total est donné par $d\,(d+1)^{2m+2}$. Comme l'algorithme est récursif, on doit factoriser P, m fois en introduisant chaque fois une variable supplémentaire. La complexité totale de la factorisation sur $Z[X,X_1,..,X_m]$ est donc donnée par

(B.III.3) $\quad m\,d\,(d+1)^{2m+2}$.

L'algorithme originel était plus intéressant, pour un polynôme dense, c'est-à-dire comportant $(d+1)^m$ monômes. En effet, les coûts théoriques sont à peu près les mêmes dans les deux cas et l'algorithme originel était beaucoup plus simple.
Cependant, très souvent les polynômes donnés ne sont pas denses; c'est notamment le cas de tous les exemples qui sont habituellement donnés. Le nouvel algorithme est donc très souvent meilleur que l'algorithme originel.

Recherche des diviseurs de B sur $K[X_j]$.

Considérons la factorisation de B sur $\mathbf{K}[X_j]/(X_j)^{dk+1}$, notée $Q_1\ldots Q_r$ et supposons qu'on ait supprimé de ces diviseurs tous les monômes en X_j de degré supérieur à dk.
S'il n'y a aucun facteur parasite, alors $Q_1\ldots Q_r$ est la factorisation cherchée de B sur $Z[X,X_1,\ldots,X_j]$. Sinon, tout diviseur Q de B sur $Z[X,X_1,\ldots,X_j]$, est égal à un produit de un ou plusieurs des polynômes $Q_1,\ldots,Q_r$.
La phase finale de l'algorithme peut être présentée comme suit

1°) On effectue la division de P par chacun des polynômes $Q_1, \ldots, Q_r$; si les restes des divisions sont nuls, on peut affirmer qu'il n'y a aucun facteur parasite et la factorisation est terminée; si au contraire il y a des facteurs parasites, on considère l'ensemble D des polynômes Q_j n'ayant pas donné un reste nul.

2°) Tant que D n'est pas vide, on calcule tous les produits $Q_k\ldots Q_n$ où

$Q_k,...,Q_n \in D$, ainsi que les divisions de B par chacun de ces produits. Si un nouveau diviseur $Q = Q_k...Q_n$ est trouvé, on retranche de D les polynômes $Q_k,...,Q_n$ de ce produit.

Le coût des calculs dépend beaucoup du nombre des facteurs parasites. Pour cette raison, on considère plusieurs m-uples $(\alpha_1,...,\alpha_m)$ vérifiant la condition de Wang **(B.II.W)**, puis on garde le m-uple tel que $P(X,\alpha_1,...,\alpha_m)$ ait le moins de facteurs possibles sur $\mathbf{Z}[X]$. Remarquons cependant que la condition de Wang est assez restrictive, donc d'une part, la recherche systématique de plusieurs m-uples aura un coût non négligeable, d'autre part elle sera inutile la plupart du temps.

Dans l'algorithme proposé ci-dessous, on ne détermine qu'un seul m-uple avec une propriété moins restrictive; le m-uple sera donc plus facile à obtenir et les calculs de la dernière étape seront très réduits grâce à un critère permettant de regrouper les facteurs parasites, correspondant à un même diviseur de P.

IV) NOUVELLES METHODES DE FACTORISATION.

Trois nouveaux algorithmes sont proposés ci-dessous pour améliorer l'algorithme classique. Le premier algorithme apporte une amélioration aux étapes 2 et 3, en utilisant la décomposition des fractions rationnelles en éléments simples pour accélérer l'étape 2, et en donnant un critère pour regrouper les diviseurs trouvés à l'étape 2 et obtenir ainsi les facteurs de P sur $\mathbf{Z}[X,X_1,...,X_m]$. Ce critère est obtenu en utilisant la notion de polynôme normalisé introduite plus haut dans II.3°).

Le deuxième algorithme détermine la factorisation de $P(X,X_1,...,X_m)$ à partir de celle de $P(X,\alpha_1,...,\alpha_m)$; les étapes 2 et 3 de l'algorithme classique sont donc supprimées et le nouvel algorithme se réduit à la factorisation d'un polynôme à une variable, ayant de grands coefficients.

Le troisième algorithme a pour but de diminuer le coût de l'étape 3 dans le cas où il y a beaucoup de facteurs parasites à l'issue de l'étape 2. Les facteurs obtenus à l'étape 2 sont transformés ce qui exige un calcul supplémentaire, dont le coût est du même ordre que celui de l'étape 2 classique. Mais ensuite, l'étape 3 n'exige plus le calcul d'aucun produit de polynômes.

a) Polynômes normalisés et décomposition de fractions rationnelles.

L'algorithme présenté ci-dessous utilise le changement de variable défini plus haut en **B.I.3°)**, qui rend le polynôme P unitaire. Mais le principe de l'algorithme est le même que celui de l'algorithme classique. La différence essentielle est que l'algorithme présenté utilise la décomposition des fractions rationnelles en éléments simples. Cet algorithme a été présenté pour la première fois dans **[V2]**, puis a été repris par Lugiez dans **[Lu1]**. Rappelons certaines notations qui seront utilisées dans tout le paragraphe IV. On suppose que le polynôme donné P a été transformé en un polynôme normal et unitaire, comme on l'a vu dans **II.3°**, à l'aide du changement de variable qui substitue $X_j + \alpha_j \cdot X$ à X_j. Dans ces conditions, $P(X,0,...,0)$ est un polynôme de degré d sans facteur carré et de plus $\delta P^{(k)} \leq d - k$ pour $k \geq 0$.

On note comme plus haut P sous la forme $P = P^{(o)} + P^{(1)} + ... + P^{(j)} + ... + P^{(\delta)}$, où $P^{(j)}$ est une forme de degré j de $Z[X_1,...,X_m]$. Les diviseurs de P sur $Z[X,X_1,...,X_m]/\Delta_j$ sont notés $Q_1, ..., Q_r$, où Δ_j est l'idéal engendré par p^N et les monômes de $Z[X_1,...,X_m]$ de degré total supérieur à j. Chaque diviseur Q_k s'écrit sous la même forme que P à savoir

$$Q_k = Q_k^{(o)} + Q_k^{(1)} + ... + Q_k^{(j)} + ...$$

1°) Factorisation de P sur $Z[X,X_1,..,X_m]/\Delta_q$.

Considérons la factorisation de P sur $Z[X,X_1,...,X_m]/\Delta_j$, qui s'écrit sous la forme suivante

(B.IV.1) $$P^{(o)} + .. + P^{(j)} + ... = (Q_1^{(o)} + .. + Q_1^{(j)} + ..) ... (Q_r^{(o)} + ... + Q_r^{(j)} + ..).$$

Cette égalité est équivalente au système des relations suivantes

$$(S)\quad \begin{cases} Q_1^{(o)}\cdot Q_2^{(o)}\cdots Q_r^{(o)} = P^{(o)} \\ Q_1^{(1)}\cdot Q_2^{(o)}\cdots Q_r^{(o)} + Q_1^{(o)}\cdot Q_2^{(1)}\cdot Q_3^{(o)}\cdots Q_r^{(o)} + \ldots = P^{(1)} \\ \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ Q_1^{(j)}\cdot Q_2^{(o)}\cdots Q_r^{(o)} + Q_1^{(o)}\cdot Q_2^{(j)}\cdot Q_3^{(o)}\cdots Q_r^{(o)} + \ldots = P_o^{(j)} \\ \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \end{cases}$$

où

$$(R)\qquad P_o^{(j)} = P^{(j)} - \sum_{j1+\ldots+jr=j} Q_1^{(j1)}\cdot Q_2^{(j2)}\cdots Q_r^{(jr)}$$

La première égalité $Q_1^{(o)}\cdot Q_2^{(o)}\cdots Q_r^{(o)} = P^{(o)}$ représente la factorisation de $P(X,0,\ldots,0)$ sur $Z[X]$. Cette factorisation sera donc la première étape de l'algorithme et elle est réalisée grâce à l'algorithme de factorisation des polynômes à une variable. La $j^{ième}$ étape consiste calculer $Q_1^{(j)}$, $Q_2^{(j)}$,..., $Q_r^{(j)}$, à l'aide de l'équation

$$(E)\qquad Q_1^{(j)}\cdot Q_2^{(o)}\cdots Q_r^{(o)} + Q_1^{(o)}\cdot Q_2^{(j)}\cdot Q_3^{(o)}\cdots Q_r^{(o)} + \ldots = P_o^{(j)},$$

où $P_o^{(j)}$ est défini par la relation **(R)** en fonction des polynômes $Q_k^{(jk)}$ où $0 \le jk \le j-1$. En résolvant les équations **(E)** successivement pour $j = 1, 2,\ldots, q$, on obtiendra pour tout k la suite $Q_k^{(o)}, Q_k^{(1)}, \ldots, Q_k^{(j)}, \ldots, Q_k^{(q)}$.

On va donc étudier la résolution des équations **(E)**. Remarquons que les polynômes cherchés $Q_k^{(j)}$ s'écrivent sous la forme d'une somme de monômes

$b_{k\,N1\,N2..Nm}\, X_1^{N1}\, X_2^{N2} \ldots X_m^{Nm}$.

L'équation **(E)** se décompose donc en une suite d'équations de la forme

$$(F)\qquad b_1\cdot Q_2^{(o)}\cdots Q_r^{(o)} + Q_1^{(o)}\cdot b_2\cdot Q_3^{(o)}\cdots Q_r^{(o)} + \ldots = c,$$

où les deux membres sont des polynômes en X. Divisons les deux membres de **(F)** par $P^{(o)}$. On obtient la relation

$$(B.IV.2)\qquad \frac{c}{P^{(o)}} = \frac{b_1}{Q_1^{(o)}} + \frac{b_2}{Q_2^{(o)}} + \ldots + \frac{b_r}{Q_r^{(o)}}.$$

Comme $\delta c < \delta P^{(o)}$ et $\delta b_k < \delta Q_k^{(o)}$, le second membre de **(B.IV.2)** est la décomposition en éléments simples de la fraction rationnelle $c/P^{(o)}$. Appelons $b_1,\ldots, b_r$ les composantes de c. Il est évident que $b_1,\ldots, b_r$ sont des fonctions linéaires de c. En conséquence, on calculera les composantes $b_1^k,\ldots, b_r^k$ de X^k pour $k \ge 0$. On obtiendra donc les polynômes $Q_1^{(j)}, \ldots, Q_r^{(j)}$, en remplaçant X^k par $b_1^k, \ldots, b_r^k$ dans $P_o^{(j)}$.

Par suite, les calculs essentiels sont les décompositions en éléments simples des fractions rationnelles $X^k/P^{(o)}$. Pour effectuer ces décompositions on peut utiliser l'algorithme proposé par Kung et Tong dans **[KT]**.

On obtient ainsi à partir des diviseurs $Q_k^{(o)}$ de $P^{(o)}$, les sommes $Q_k^{(o)} + Q_k^{(1)} + ... + Q_k^{(q)}$, qui représentent les diviseurs de P sur $\mathbf{Z}[X,X_1,..,X_m]/\Delta_q$.

Evaluation du coût.

La décomposition des fractions rationnelles $X^k/P^{(o)}$ avec l'algorithme de Kung et Tong exige $d\,\mathrm{Log}^2 d$ opérations élémentaires. Comme il y a au plus d décompositions, le coût total des décompositions est donc égal à $d^2\,\mathrm{Log}^2 d$.

La résolution des équations **(E)** exige aussi le calcul des polynômes $P_o^{(j)}$, c'est-à-dire le calcul des sommes (R). Pour ce calcul on peut effectuer les $r-1$ produits

$$(Q_1^{(o)} + .. + Q_1^{(j)} + ..)...(Q_r^{(o)} + .. + Q_r^{(j)} + ..),$$

en ne gardant que les monômes de degré k suivant $X_1,...,X_m$. Mais les calculs sont très longs, de l'ordre de $(d+1)^{2m+2}$. On peut donc procéder comme dans l'algorithme de Wang et calculer la dérivée partielle correspondant à chaque monôme $X_1^u \cdots X_m^v$ désiré. Les calculs sont alors du même ordre que ceux de l'algorithme amélioré de Wang.

Cependant l'algorithme de Kung et Tong pour la décomposition des fractions rationnelles, qui a été testé par Lugiez **[Lu1]**, paraît plus rapide que l'algorithme utilisé par Wang et présenté en **B.I.5°)**.

Exemple.

On continue l'exemple du paragraphe précédent, à savoir la factorisation du polynôme suivant

$$6X^4 + 2X^3 + 5X^2 - 4 + 2X_1(3X^2 - 2X) + X_2(2X^4 + 2X^3 + 3X^2 + 5X - 2) + X X_2^2 + 2X^2 X_1 X_2,$$

qu'on a transformé en un polynôme unitaire, et qui est noté P pour la suite :

$$P = X^5 - 2X^4 - X + 2 + X_1(2X^4 - 3X^2 + 2X) - X_2(2X^4 + 2X^3 + X^2 + 5X - 2) - 2X X_2^2 - 2X^2 X_1 X_2.$$

La première étape est la factorisation de $P^{(o)} = P(X,0,0) = X^5 - 2X^4 - X + 2$ à l'aide de l'algorithme classique de factorisation des polynômes à une variable, ce qui donne

$$P^{(o)} = Q_1^{(o)} \cdot Q_2^{(o)} \cdot Q_3^{(o)} \cdot Q_4^{(o)} = (X+1)(X-1)(X-2)(X^2+1).$$

La deuxième étape consiste à calculer les décompositions des fractions suivantes en éléments simples

$$\frac{1}{X^5-2X^4-X+2} = \frac{1}{12(X+1)} - \frac{1}{4(X-1)} + \frac{1}{15(X-2)} + \frac{X+2}{10(X^2+1)}$$

$$\frac{X}{X^5-2X^4-X+2} = \frac{-1}{12(X+1)} - \frac{1}{4(X-1)} + \frac{2}{15(X-2)} + \frac{2X-1}{10(X^2+1)}$$

$$\frac{X^2}{X^5-2X^4-X+2} = \frac{1}{12(X+1)} - \frac{1}{4(X-1)} + \frac{4}{15(X-2)} - \frac{X+2}{10(X^2+1)}$$

$$\frac{X^3}{X^5-2X^4-X+2} = \frac{-1}{12(X+1)} - \frac{1}{4(X-1)} + \frac{8}{15(X-2)} + \frac{-2X+1}{10(X^2+1)}$$

$$\frac{X^4}{X^5-2X^4-X+2} = \frac{1}{12(X+1)} - \frac{1}{4(X-1)} + \frac{16}{15(X-2)} + \frac{X+2}{10(X^2+1)}$$

On détermine les monômes en X_1 et X_2 des diviseurs Q_1, Q_2, Q_3 et Q_4 en remplaçant dans $P^{(1)} = X_1 (2 X^3 - 3 X^2 + 2 X) - X_2 (2 X^4 + 2 X^3 + X^2 + 5 X - 2)$, X^j par le numérateur A_k de la fraction $A_k/Q_k^{(o)}$ du développement de $X^j/P^{(o)}$ en éléments simples calculé ci-dessus. On obtient

$$Q_1^{(o)} + Q_1^{(1)} = X + 1 + \tfrac{1}{2} X_2 - \tfrac{1}{2} X_1,$$

$$Q_2^{(o)} + Q_2^{(1)} = X - 1 + 2 X_2,$$

$$Q_3^{(o)} + Q_3^{(1)} = X - 2 - 4 X_2,$$

$$Q_4^{(o)} + Q_4^{(1)} = X^2 + 1 + \tfrac{1}{2}(-X+1)X_2 + \tfrac{1}{2}(X+1)X_1.$$

On calcule $P_o^{(2)} = P^{(2)}$ – monômes en X_2^2, $X_2 \cdot X_3$ et X_3^2 de $Q_1 \cdot Q_2 \cdot Q_3 \cdot Q_4$.

On obtient

$$P_o^{(2)} = - 2XX_2^2 - 2X^2X_1X_2 - X_2^2(33X^3/4 + 7X^2 + 37X/4 + 15/2 + 2X) - X_1X_2(- X^3/2 + 3X^2/2 - 2X^2 - X) - X_1^2(X^3/4 - X^2/2 - X/4 + 1/2)$$

$$= X_2^2(33X^3/4 + 7X^2 + 37X/4 + 15/2) + X_1X_2(- X^3/2 + 3X^2/2 - X) + X_1^2(X^3/4 - X^2/2 - X/4 + 1/2).$$

On détermine les monômes en X_1^2, $X_1 \cdot X_2$ et X_2^2 des diviseurs Q_1, Q_2, Q_3 et Q_4 en remplaçant dans $P_o^{(2)}$, X^j par le numérateur A_k de la fraction $A_k/Q_k^{(o)}$ du développement de $X^j/P^{(o)}$ en éléments simples. On obtient :

$$Q_1 = Q_1^{(o)} + Q_1^{(1)} + Q_1^{(2)} = X + 1 + \tfrac{1}{2} X_2 - \tfrac{1}{2} X_1 - X_2^2/4 + X_1X_2/4,$$

$$Q_2 = Q_2^{(o)} + Q_2^{(1)} + Q_2^{(2)} = X - 1 + 2 X_2 - 8 X_2^2,$$

$$Q_3 = Q_3^{(o)} + Q_3^{(1)} + Q_3^{(2)} = X - 2 - 4 X_2 + 8 X_2^2,$$

$$Q_4 = Q_4^{(o)} + Q_4^{(1)} + Q_4^{(2)} = X^2 + 1 + \tfrac{1}{2}(-X+1)X_2 + \tfrac{1}{2}(X+1)X_1 + XX_2^2/4 + X_1^2/4 - (X+1)X_1X_2/4.$$

Les polynômes Q_1, Q_2, Q_3 et Q_4 sont les diviseurs cherchés de P sur $Z[X,X_1,X_2]/\Delta_3$.

2°) Factorisation de P sur $Z[X,X_1,..,X_m]$.

Comme on l'a vu dans la méthode classique de factorisation, tout diviseur de $P(X,X_1,...,X_m)$ sur $Z[X,X_1,...,X_m]$ est égal à un produit $Q_{j1}^{(d)}\cdot Q_{j2}^{(d)}\cdots Q_{jt}^{(d)}$ qui est un sous-produit quelconque de $Q_1^{(d)}\cdot Q_2^{(d)}\cdot\cdot Q_r^{(d)}$.
L'algorithme classique consiste simplement à calculer tous ces sous-produits. Très souvent l'un des diviseurs $Q_k^{(d)}$ de P sur $Z[X,X_1,...,X_m]/\Delta_d$ divise réellement P sur $Z[X,X_1,...,X_m]$. Dans ce cas la recherche des autres diviseurs est simplifiée. Dans le cas où il n'y a pas de diviseurs parasites, la factorisation $Q_1^{(d)}\cdot Q_2^{(d)}\cdot\cdot Q_r^{(d)}$ est la factorisation cherchée sur $Z[X,X_1,...,X_m]$, et aucun calcul de produit n'est nécessaire.
On propose ici un critère pour qu'un produit $Q_{j1}^{(d)}\cdot Q_{j2}^{(d)}\cdots Q_{jt}^{(d)}$ corresponde à un diviseur de P sur $Z[X,X_1,...,X_m]$ sans avoir à effectuer le produit. Ce critère repose sur la notion de polynôme normalisé introduite en **(B.I.3)**. Il utilise aussi la notion représentation polyèdrale des polynômes présentée dans la partie C. Notons $\pi(P)$ le polyèdre associé au polynôme P et $\pi(P)_V$ la face définie par l'hyperplan d'appui directement perpendiculaire à V. Ces polyèdres appartiennent à l'espace affine Z^{n+1} ayant comme repère les axes de coordonnées $(Ox, Ox_1, .., Ox_m)$.
Le résultat essentiel concerne la représentation du produit de deux polynômes Q et R

(B.IV.4) $\pi(Q\cdot R)_V = \pi(Q)_V + \pi(R)_V$.

La suite du paragraphe utilise aussi la notion de monôme extrêmal. Un monôme extrêmal est un monôme associé à un sommet du polyèdre, et on a le résultat suivant, démontré dans la partie C :

Lemme 1.

Soit P un polynôme égal au produit Q·R où P, Q et R appartiennent à $Z[X,X_1,...,X_m]$.

Tout monôme extrêmal de P est égal à un produit unique de monômes extrêmaux de Q et R.

On utilisera également la notion de polynôme normalisé introduite en **B.I**, qu'on généralise comme suit :

On dit que Q est normalisé jusqu'au rang n si on a :

(B.IV.4) $\delta Q^{(k)} \leq \delta Q^{(o)} - k$ pour $0 \leq k \leq n$.

La relation ci-dessus signifie que le polyèdre $\pi(Q)$ associé à Q est situé en dessous de l'hyperplan d'équation $x + x_1 + ... + x_m = \delta Q^{(o)}$.
On obtient ainsi une méthode de recherche des diviseurs, grâce au lemme suivant :

Lemme 2.

Soit P un polynôme normalisé de $Z[X,X_1,...,X_m]$ et $Q_1 \cdot Q_2 \cdots Q_r$ la factorisation du polynôme P sur $Z[X,X_1,...,X_m]/\Delta_d$.
Considérons le produit $Q = Q_{j1} \cdots Q_{jt}$ où $1 \leq j1, j2,..,jt \leq r$ et R le produit des diviseurs Q_{jn} ne figurant pas dans Q.

1) Si l'un des produits Q ou R est normalisé jusqu'au rang n, alors l'autre produit est également normalisé jusqu'au même rang n.

2) Les trois propriétés suivantes sont équivalentes :

- **(i)** le produit Q est un diviseur de P sur $Z[X,X_1,...,X_m]$;
- **(ii)** Q est normalisé jusqu'au rang $\delta P^{(o)}$;
- **(iii)** les monômes de Q et de R dont les degrés suivant $X_1,...,X_m$ sont compris respectivement entre $\delta Q^{(o)}$ et d et entre $\delta R^{(o)}$ et d sont nuls.

Démonstration.

1) Vérifions 1) en montrant que si Q n'est pas normalisé, alors R ne l'est pas non plus. Par hypothèse, il existe donc ko tel que

(B.IV.5) $\delta Q^{(k)} > \delta Q^{(o)} - k$,

pour $k = ko$, ko étant le plus petit entier vérifiant **(B.IV.5)**. Calculons $P^{(ko)}$ en effectuant le produit $(Q^{(o)} + .. + Q^{(ko)} + ..)\,(R^{(o)} + .. + R^{(ko)} + ..)$. On obtient $P^{(ko)} = Q^{(ko)} \cdot R^{(o)} + Q^{(ko-1)} \cdot R^{(1)} + ... + Q^{(o)} \cdot R^{(ko)}$. Les polynômes $P^{(ko)}$, $Q^{(ko-1)} \cdot R^{(1)}$, .., $Q^{(1)} \cdot R^{(ko-1)}$ ont tous un degré en X majoré par $\delta P^{(o)} - ko$. Donc puisque Q vérifie **(B.IV.5)**, alors R doit aussi vérifier $\delta R^{(ko)} > \delta R^{(o)} - ko$. Donc R n'est pas normalisé.
2) Supposons que Q divise P sur $Z[X,X_1,...,X_m]$ et ni Q, ni R ne soient normalisés. Notons respectivement μ_1 et μ_2 les valeurs maximum de $\delta Q^{(k1)} + k1$ et $\delta R^{(k2)} + k2$ pour
$1 \leq k1, k2 \leq d$. Considérons les hyperplans d'appui H_1, H_2 et H de $\pi(Q)$, $\pi(R)$ et $\pi(P)$ perpendiculaires au vecteur $V = (1, 1,...,1)$; ils ont comme équations

$$\begin{cases} x + x_1 + \ldots + x_m = \mu_1 \\ x + x_1 + \ldots + x_m = \mu_2 \\ x + x_1 + \ldots + x_m = \mu. \end{cases}$$

D'après **(B.IV.4)**, on a $\mu = \mu_1 + \mu_2$. Soit m un monôme extrêmal de P associé à un point de H. Alors d'après le **lemme 1,** m s'écrit de façon unique $m_1 \cdot m_2$ où m_1 et m_2 sont des
monômes associés à des points de H_1 et H_2. Comme on suppose que Q et R ne sont pas normalisés, $\mu_1 > \delta\, Q^{(o)}$ et $\mu_2 > \delta\, R^{(o)}$. Par suite,

$$\mu = \mu_1 + \mu_2 > \delta\, Q^{(o)} + \delta\, R^{(o)} = \delta\, P^{(o)}.$$

Si k désigne le degré de m suivant $x_1,\ldots,x_m$ et $\delta\, m$ le degré de m suivant x, alors on obtient $\mu = k + \delta\, m < k + \delta\, P^{(k)}$. Par suite $\delta\, P^{(o)} < k + \delta\, P^{(k)}$ ce qui est contraire à l'hypothèse du lemme qui considère P normalisé.

Supposons maintenant que Q soit normalisé jusqu'au rang $d = \delta\, P^{(o)}$. D'après **1)** le polynôme R est également normalisé jusqu'au même rang. On a donc $\delta Q^{(k)} \le \delta Q^{(o)} - k$ et $\delta R^{(k)} \le \delta R^{(o)} - k$ pour $0 \le k \le d$. En particulier, $Q^{(k)} = 0$ pour $\delta Q^{(o)} < k \le d$ et $R^{(k)} = 0$ pour $\delta R^{(o)} < k \le d$. Donc **(iii)** est vérifié.

Si **(iii)** est vérifié, alors le polynôme $P - Q \cdot R$ ne peut contenir que des monômes de degrés supérieurs à d suivant $X_1,\ldots,X_m$. Comme d'autre part on a $P - Q \cdot R = 0$ sur $Z[X,X_1,\ldots,X_m]/\Delta_d$, alors $P - Q \cdot R$ est nul sur $Z[X,X_1,\ldots,X_m]$.

<u>Algorithme de recherche des diviseurs de</u> $Z[X,X_1,\ldots,X_m]$.

Soit $Q = Q_{j\,1} \cdots Q_{j\,t}$ où $1 \le j1, j2,..,jt \le r$, un sous-produit de $Q_1 \cdot Q_2 \cdots Q_r$ qui correspond à un diviseur de P sur $Z[X,X_1,\ldots,X_m]$. D'après les relations du système (S) du **IV.1°)**, $Q^{(j)} = Q_{j\,1}^{(j)} \cdot Q_{j\,2}^{(o)} \cdots Q_{j\,t}^{(o)} + Q_{j\,1}^{(o)} \cdot Q_{j\,2}^{(j)} \cdots Q_{j\,t}^{(o)} + \ldots$

Les polynômes $Q_{j\,k}^{(j)}$ sont obtenus à l'aide de décompositions de fractions rationnelles en éléments simples; en notant q leur degré suivant $X_1,\ldots,X_m$, ils s'écrivent sous la forme suivante

$$Q_{j\,k}^{(j)} = \sum_{d^o\sigma = j} (A_k^{\sigma} X^{q-1} + B_k^{\sigma} X^{q-2} + \ldots + C_k^{\sigma})\, \sigma.$$

Comme Q doit être normalisé, $Q^{(j)} = 0$ pour $\delta\, Q^{(o)} < j \le d$, donc pour ces valeurs de j, la somme des monômes de $Q^{(j)}$ de degré maximum suivant X est nulle. Or ces monômes de degré maximum ne peuvent provenir que des produits

$$Q_{j\,1}^{(\omega)}\cdots Q_{j\,k-1}^{(\omega)}\cdot Q_{j\,k}^{(j)}\cdot Q_{j\,k+1}^{(\omega)}\cdots Q_{j\,t}^{(\omega)},$$

les autres produits étant de degré inférieur suivant X. Comme les polynômes $Q_{j\,1}^{(\omega)},\ldots,$ $Q_{j\,t}^{(\omega)}$ sont unitaires, les coefficients des monômes de $Q^{(j)}$ de degré maximum sont égaux à la somme des coefficients A^{σ}_{k}. On a donc la relation :

(B.IV.6) $$\sum_{1\le k\le t} A^{\sigma}_{k} = 0.$$

pour les monômes σ des polynômes $Q_{j\,k}^{(j)}$, tels que $\delta Q^{(\omega)} < j \le d$. Dans la suite cette relation sera appelée **condition d'annulation**. On obtient donc l'algorithme suivant de recherche des diviseurs de P en deux étapes

AI. On détermine les sous-ensembles $F = \{Q_{j\,1},\ldots,Q_{j\,t}\}$ de $E = \{Q_1,\ldots,Q_r\}$ vérifiant la condition d'annulation **(B.IV.6)** pour un produit Q de degré minimal suivant $X_1,\ldots,X_m$, c'est-à-dire majoré par ½d;

AII. On vérifie que Q est normalisé en effectuant le produit $Q_{j\,1}\cdots Q_{j\,t}$, uniquement pour les sous-ensembles F définis dans **AI.**

Dans la partie **AI**, le diviseur cherché Q a un degré suivant $X_1,\ldots,X_m$ majoré par ½d. Les monômes σ pour lesquels la condition d'annulation est vérifiée, ont un degré suivant $X_1,\ldots,X_m$ noté $d°\sigma$; ils sont donc caractérisés par la relation suivante

(B.IV.7) $$\tfrac{1}{2}\,d \le d°\sigma \le d.$$

Construisons la matrice M de termes A^{σ}_{k} à n lignes et r colonnes. Le nombre n de lignes est égal au nombre de monômes σ vérifiant **(B.IV.7)**. L'étape **AI** de l'algorithme revient donc à trouver les sous-ensembles d'indices {j1,...,jt} de {1,...,r} dont les colonnes $C_{j\,1},\ldots,C_{j\,t}$ dans la matrice M ont une somme $C_{j\,1} + \ldots + C_{j\,t}$ nulle. Pour cette recherche de colonnes dont la somme est nulle, on peut remplacer une ligne par une combinaison linéaire des autres; on peut donc utiliser la méthode du pivot dans le but de faire apparaître le maximum de 0 dans M. Lorsque les pivotages sont terminés on obtient une matrice M' de la forme suivante

$$M' = \begin{vmatrix} 0 & 0 & 1 & \ldots\ldots & a_{u+1} & a_{u+2} & \ldots\ldots & a_r \\ 1 & 0 & 0 & \ldots\ldots & b_{u+1} & b_{u+2} & \ldots\ldots & b_r \\ 0 & 1 & 0 & \ldots\ldots & c_{u+1} & c_{u+2} & \ldots\ldots & c_r \\ \ldots\ldots & & \ldots\ldots & & \ldots\ldots & \ldots\ldots & & .. \\ 0 & 0 & 0 & \ldots\ldots & g_{u+1} & g_{u+2} & \ldots\ldots & g_r \end{vmatrix}$$

Comme r est le plus souvent, très inférieur à n, le nombre u de lignes linéairement indépendantes est majoré par r et à partir d'un certain rang toutes les lignes de M' sont nulles. On ne fait donc des calculs de pivotages que pour r lignes au plus. Notons M° la matrice obtenue à partir de M' en supprimant les lignes nulles:

M° est une matrice carrée d'ordre r et les calculs de la transformation de M en M' sont de l'ordre de r^3.

L'algorithme de recherche des diviseurs de P se décompose donc comme suit.

Algorithme détaillé de recherche des diviseurs.

A.I.1. Transformation de la matrice M en M°;

A.I.2. Recherche d'indices j1,..., jt tels que la somme des colonnes $C_{j1} + \ldots + C_{jt}$ soit nulle.

A.II.1. Pour chaque ensemble {j1,...,jt} défini en **A.I.2**, calcul de chaque produit $Q_{j1} \cdots Q_{jt}$ en ne gardant que les monômes de degré suivant $X_1,\ldots,X_m$ inférieurs ou égaux à d;

A.II.2. On vérifie que le produit $Q = Q_{j1} \cdots Q_{jt}$ est normalisé; si c'est le cas Q est l'un des diviseurs cherchés.

Exemple.

On continue l'exemple du paragraphe précédent, à savoir la factorisation du polynôme suivant $P = X^5 - 2X^4 - X + 2 + X_1(2X^3 - 3X^2 + 2X)$

$$- X_2(2X^4 + 2X^3 + X^2 + 5X - 2) - 2X X_2^2 - 2X^2 X_1 X_2.$$

La factorisation de P sur $Z[X,X_1,X_2]/\Delta_3$ donne les diviseurs suivants

$Q_1 = X + 1 + \frac{1}{2}X_2 - \frac{1}{2}X_1 - \frac{1}{4}X_2^2 + \frac{1}{4}X_1X_2,$

$Q_2 = X - 1 + 2X_2 - 8X_2^2,$

$Q_3 = X - 2 - 4X_2 + 8X_2^2,$

$Q_4 = X^2 + 1 + \frac{1}{2}(-X+1)X_2 + \frac{1}{2}(X+1)X_1 + \frac{1}{4}XX_2^2 + \frac{1}{4}X_1^2 - \frac{1}{4}(X+1)X_1X_2.$

La matrice associée M s'écrit :

$$M = \begin{array}{r} \\ X_1^2 \\ X_1X_2 \\ X_2^2 \end{array} \left| \begin{array}{cccc} Q_1 & Q_2 & Q_3 & Q_4 \\ 0 & 0 & 0 & 0 \\ \frac{1}{4} & 0 & 0 & \frac{1}{4} \\ -\frac{1}{4} & -8 & 8 & \frac{1}{4} \end{array} \right|$$

Après les calculs de pivotages, on obtient la matrice simplifiée:

$$M' = \left| \begin{array}{cccc} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & -1 \\ 0 & -1 & 1 & 0 \end{array} \right|$$

La seule partition de $\{Q_1, Q_2, Q_3, Q_4\}$ vérifiant la condition d'annulation est $\{Q_1, Q_4\}$, $\{Q_2, Q_3\}$; donc les seuls diviseurs possibles de P sont donnés par les produits

$Q = Q_1 \cdot Q_4 = X^3 + X^2 + X + 1 + X_2 + XX_1$,

$R = Q_2 \cdot Q_3 = X^2 - 3X + 2 - 2XX_2$.

On peut vérifier que le produit $Q \cdot R$ est bien égal à P.

b) Substitution de $X_1, ..., X_m$ par des entiers $\alpha_1, ..., \alpha_n$.

L'algorithme proposé ici a été présenté pour la première fois dans [V1]. La même technique a été reprise plus tard pour la factorisation des polynômes à deux variables par Lugiez dans [Lu2] et pour le calcul du pgcd des polynômes à plusieurs variables par Geddes dans [CGG].

Rappelons d'abord qu'avant les méthodes modulaires basées sur le lemme de Hensel et sur l'algorithme de Berlekamp, il existait une méthode dûe à Kronecker et présentée par Van der Waerden dans [VW].

Cette méthode consiste à factoriser un polynôme de $\mathbf{K}[X]$, où $\mathbf{K}$ est un anneau factoriel, sur lequel on a déjà défini une algorithme de factorisation. Si le polynôme donné est de degré d, on cherche un diviseur Q de P de degré $\leq \frac{1}{2}d$. Pour déterminer Q, on considère r éléments de $\mathbf{K}$, où $r > \frac{1}{2}d$, notés $\alpha_1, ..., \alpha_r$. Puis on calcule $P(\alpha_1),..., P(\alpha_r)$ qu'on factorise sur $\mathbf{K}$. On considère ensuite tous les r-uples $(p_1,..., p_r)$ où $p_1, ..., p_r$ sont respectivement des facteurs de $P(\alpha_1),..., P(\alpha_r)$. Ensuite, pour chacun de ces r-uples, on calcule le polyôme d'interpolation, par exemple par la méthode de Lagrange. Si P admet un diviseur $Q \in \mathbf{K}[X]$, ayant un degré majoré par $\frac{1}{2}d$, alors ce diviseur sera obtenu par interpolation. Mais pour trouver Q on devra aussi diviser P par chacun des polynômes obtenu par interpolation.

On obtient ainsi un algorithme de factorisation sur $\mathbf{Z}[X,X_1,..,X_m]$. En effet, l'algorithme présenté ci-dessus fait passer de la factorisa-tion sur $\mathbf{Z}[X,X_1,...,X_j]$, à la factorisation sur $\mathbf{Z}[X,X_1,..,X_{j+1}]$. Rappelons qu'on peut factoriser tout élément n de $\mathbf{Z}$, en essayant de diviser n par tout nombre premier majoré par sa racine carrée. On pourra donc en déduire la factorisation des éléments de $\mathbf{Z}[X]$, puis de $\mathbf{Z}[X,X_1]$,... et enfin de $\mathbf{Z}[X,X_1,...,X_m]$.

Le coût d'un tel algorithme est énorme, car le nombre de polynômes à interpoler est très grand; en effet, si P est de degré d et si chaque élément $P(\alpha_j)$ admet au plus K facteurs, premiers ou non, alors le nombre de r-uples possibles est majoré par $K^{d/2}$. Rappelons l'idée générale de cet algorithme de Kronecker; elle repose sur le fait que, si on considère r projections de P sur **K**, on peut en déduire les projections de l'un des diviseurs Q de P, puis en déduire Q par interpolation, à partir de ses projections sur **K**.

Le nouvel algorithme proposé ici est basé sur un principe analogue, mais il propose deux idées nouvelles :

1°) Ne considérer qu'une seule projection de P, puis en déduire une projection de Q qui soit suffisante pour déterminer Q;

2°) Commencer le processus non pas avec **K** = **Z**, car le nombre de facteurs d'un entier est grand, mais avec **K** = Z[X], en utilisant l'algorithme de factorisation des polynômes à une variable.

Remarquons qu'une seule projection de Q sur **K** peut être suffisante pour déterminer Q, à condition de connaître une borne B des coefficients de Q. En effet, supposons pour simplifier que **K** = Z[X] et que P, Q $\in$ Z[X,X_1]. Le polynôme Q s'écrit donc

$$Q(X,X_1) = A_o(X_1) + A_1(X_1)\cdot X + \ldots + A_q(X_1)\cdot X^q.$$

Considérons $Q(X,\alpha) = A_o(\alpha) + A_1(\alpha)\, X + \ldots + A_q(\alpha)\, X^q$, avec $\alpha > 2\cdot B$. Le $k^{ième}$ coefficient de $Q(X,\alpha)$ s'écrit donc sous la forme

(B.IV.8) $$A_k(\alpha) = b_o + b_1\cdot\alpha + \ldots + b_n\cdot\alpha^n.$$

Si on suppose que $-\frac{1}{2}\alpha < b_o, b_1, \ldots, b_n \leq \frac{1}{2}\alpha$, alors le second membre de **(B.IV.8)** est l'écriture du nombre $A_k(\alpha)$ dans la base α, avec la convention de choisir les "chiffres" de la base dans l'intervalle $]-\frac{1}{2}\alpha, \frac{1}{2}\alpha]$ au lieu de l'intervalle habituel $[0, \alpha[$. Avec ce choix particulier des chiffres dans l'intervalle $]-\frac{1}{2}a, \frac{1}{2}p]$, on dit que **le nombre est écrit dans la base symétrisée** α.

Dans ces conditions, la connaissance du polynôme $A_k(\alpha)$ permet de déterminer A_k; il suffit d'écrire les coefficients de $A_k(\alpha)$ dans la base symétrisée α, puis de remplacer α par X_1. Ainsi, à tout diviseur $Q(X,\alpha)$ de $P(X,\alpha)$, on associe un polynôme $Q(X,X_1)$. Mais ce polynôme $Q(X,X_1)$ n'est pas toujours un diviseur de $P(X,X_1)$ sur Z[X,X_1].

Comme dans l'algorithme classique de factorisation, la phase finale consistera à considérer tous les produits possibles des polynômes $Q(X,X_1)$ ainsi obtenus, puis à effectuer les divisions de P par ces produits.

1°) Suppression des diviseurs ayant moins de m+1 variables.

Supposons que P ait un diviseur Q ne contenant pas la variable X_j. Le polynôme P s'écrit donc $P = Q \cdot R$, par suite $\partial_j P = Q \cdot \partial_j R$ où $\partial_j P$ et $\partial_j R$ désignent les dérivées partielles de P et R par rapport à X_j. Par suite, si on divise P par le polynôme $\Delta_j = PGCD(P, \partial_j P)$, alors tout diviseur de P/Δ_j contiendra la variable X_j.
En divisant P par $\Delta_1, .., \Delta_m$ on supprime tous les diviseurs de P contenant moins de $m+1$ variables. On supprime en même temps les diviseurs de la forme Q^n qui sont des puissances d'un diviseur Q.
Le coût du calcul du PGCD de deux polynômes à $m+1$ variables, de hauteur H et de degré dj suivant X_j a été donné en **B.I.1°)**; il est de l'ordre de $(Log^2H + m \cdot (dj + 1) \cdot Log\ H) \cdot (dj + 1)^m$. Donc, si d désigne le degré maximum de P suivant ses variables $X, X_1, ..., X_m$, alors le coût total des suppressions des diviseurs ayant moins de $m+1$ variables est majoré par $m^2\ Log^2H\ (d + 1)^{m+1}$.
Ce coût est inférieur au coût de recherche des autres diviseurs de P. En conséquence, les calculs de ce paragraphe permettent de trouver d'éventuels diviseurs de P, plus rapidement que par l'algorithme classique.
Dans la suite on supposera donc que tout diviseur Q de P contient chacune des $m+1$ variables $X, .., X_m$ et on notera $d,..,dm$ les degrés de P suivant les variables $X, .., X_m$. Donc tout diviseur Q aura des degrés suivant $X, .., X_m$ majorés respectivement par $d-1, .., dm-1$.

2°) Suppression des diviseurs homogènes.

Remarquons que si le polynôme donné P est homogène, alors ses diviseurs sont aussi homogènes. On peut alors simplifier P et ses diviseurs en supprimant une variable, n'importe laquelle. On a donc ainsi diminué le nombre des variables. Lorsque la factorisation du nouveau polynôme est terminée, on doit rajouter la variable supprimée en rendant les diviseurs obtenus homogènes.

Supposons donc que le polynôme P ne soit pas homogène, mais qu'il contienne un diviseur homogène. Alors P s'écrit $P = Q \cdot R$ où Q est homogène. Calculons les dérivées premières de P par rapport aux variables $X, X_1,...,X_m$; on note ces dérivées $\partial P, \partial_1 P,..., \partial_m P$. On a donc $\partial_j P = \partial_j Q \cdot R + Q \cdot \partial_j R$. Par suite le polynôme
$X\,\partial P + X_1\,\partial_1 P + ... + X_m\,\partial_m P$ est égal à
$(X\,\partial Q + X_1\,\partial_1 Q + ... + X_m\,\partial_m Q)\cdot R + Q\cdot(X\cdot\partial R + ... + X_m\cdot\partial_m R)$. D'après l'identité d'Euler, $X\,\partial Q + X_1\,\partial_1 Q + ... + X_m\,\partial_m Q = n\,Q$, donc le PGCD des polynômes P et $X\,\partial P + X_1\,\partial_1 P + ... + X_m\,\partial_m P$ contient le facteur Q. On obtient donc ainsi un diviseur non trivial de P, et en divisant P par ce PGCD, P ne contient plus de diviseurs homogènes.
On suppose donc que c'est le cas dans la suite.
Grâce au paragraphe précédent on peut même supposer que P ne contient plus de facteur linéaire. En effet, si c'était le cas, alors ce diviseur serait soit homogène, soit privé d'une des variables de P. Les diviseurs cherchés seront donc de degré au moins égal à 2.

3°) Isomorphisme Φ de $Z_b[X]$ dans Z.

Soit $b \in \mathbf{N}$ et considérons l'homomorphisme f de $\mathbf{K}[X]$ dans $\mathbf{K}$ défini par $f(P) = P(b)$. Il est clair que f est surjectif, mais non injectif. On peut rendre f injectif, en le composant avec la surjection canonique $\sigma: \mathbf{Z}[X]/\mathfrak{R} \longrightarrow \mathbf{Z}[X]$, $\mathfrak{R}$ étant la relation d'équivalence définie comme suit :

(B.IV.9) $P(X)\ \mathfrak{R}\ Q(X) \Leftrightarrow P(b) = Q(b) \Leftrightarrow \left\{ \begin{array}{l} \text{Il existe } S \in \mathbf{K}[X] \text{ tel que} \\ P(X) = Q(X) + (X - b)\cdot S \end{array} \right.$

Dans la suite on notera Φ l'isomorphisme de $\mathbf{Z}[X]/\mathfrak{R}$ dans $\mathbf{K}$ induit par f. Notons $C(P)$ la classe de P, et choisissons un représentant de $C(P)$ ayant les plus petits coefficients possible. Le lemme suivant nous permet de choisir ces coefficents inférieurs en valeur absolue à $\frac{1}{2}b$.

Lemme.
Considérons l'anneau $Z_b[X] = \mathbf{Z}[X]/\mathfrak{R}$ où $\mathfrak{R}$ est la relation d'équivalence définie par **(B.IV.9)**.

La classe d'équivalence de P possède un représentant unique, noté P° dont les coefficients a_j appartiennent à l'intervalle $]-½b, ½b]$.

Démonstration.

Soit $P(X) = a_o X^d + ... + a_j X^{d-j} + a^\circ_{j+1} X^{d-j-1} + ... + a^\circ_d$, où $\{a_o, ..., a_j\}$ est la liste des coefficients de tête n'appartenant pas à l'intervalle $]-½b, ½b]$. Si cette liste est vide le lemme est vérifié. Sinon, il existe une division de a_j par b qui s'écrit $a_j = b\cdot q_j + a^\circ_j$ où $a^\circ_j \in]-½b, ½b]$. Par suite, on a

$a_j X^{d-j} = a^\circ_j X^{d-j} - q_j\cdot(X- b)\cdot X^{q-j} + q_j X^{q-j+1}$.

En remplaçant $a_j X^{d-j}$ par $a^\circ_j X^{d-j} + q_j X^{q-j+1}$ dans $P(X)$, on obtient un nouveau polynôme équivalent à P dont le nombre des coefficients n'appartenant pas à $]-½b, ½b]$ a diminué de 1.

On a donc défini un algorithme permettant de passer du polynôme donné P au polynôme P° de coefficients a°_j majorés en valeur absolue par $½b$.

Supposons maintenant que P ait deux représentants P° et P^* de coefficients a°_j et a^*_j compris entre $-½b$ et $½b$, et supposons que $P^\circ \neq P^*$. Le polynôme $P^\circ - P^*$ a ses coefficients dans l'intervalle $]-½b, ½b]$, donc

$$(P^\circ - P^*)(b) = c_o b^n + c_1 b^{N-1} + ... + c_n = 0, \text{ avec } c_o \neq 0.$$

Or on a, d'une part $|c_o|\cdot b^n \geq b^n$, et d'autre part $|c_1 b^{N-1} + ... + c_n| \leq ½b(b^{N-1} + ... + 1) \leq b^n - 1$. On obtient donc une contradiction et le lemme est vérifié.

Le corollaire suivant est une conséquence directe du lemme :

Corollaire.

Soit $b \in \mathbf{N}$ et P un polynôme de $\mathbf{Z}[X]$. Alors P s'écrit de façon unique sous la forme suivante:

$$P = P^\circ + (X - b)\cdot P^1 + (X - b)^2\cdot P^2 + ... + (X - b)^n\cdot P^n$$

où $P^\circ, P^1, ..., P^n$ ont tous leurs coefficients dans l'intervalle $]-½b, ½b]$.

De plus, si on note Log_b le logarithme dans la base b, alors $d^\circ P^j \leq \mathrm{Log}_b P(b)$ et $N \leq \mathrm{Log}_{b-1} H(P)$.

Démonstration.

D'après le lemme P s'écrit de façon unique

(B.IV.10) $P = P^\circ + (X - b)Q^1$ avec $H(P^\circ) \leq \frac{1}{2}b$.

Comme on l'a vu dans le lemme les monômes a_jX^{d-j} de P s'écrivent sous la forme $a^\circ_j X^{d-j} - q_j\cdot(X- b)\cdot X^{q-j} + q_j X^{q-j+1}$ et les monômes de Q^1 sont de la forme $q_j\cdot X^{d-j}$, avec $q_j \leq a_j/b$. On en déduit que les coefficients de Q^1 sont majorés par $H(P)/b + H(P)/b^2 + ...$, donc par $H(P)/(b-1)$.

On peut écrire de manière analogue, Q^1 sous la forme $P^1 + (X - b)\cdot Q^2$, où $H(P^1) \leq \frac{1}{2}b$ et $H(Q^2) \leq H(P)/(b-1)^2$.

On arrête le développement lorsque $H(Q^n) \leq \frac{1}{2}b$, c'est-à-dire lorsque $H(P)/(b-1)^n \leq \frac{1}{2}b$. Si on suppose $b \geq 2$, ceci est réalisé si N est la partie entière de $\mathrm{Log}_{b-1}H(P)$.

D'autre part, si on note δ le degré de P° et d celui de P, on a $b^\delta \leq P^\circ(b)$, donc δ et d sont majorés par $\mathrm{Log}_bP(b)$. Il en est donc de même du degré de $(X - b)Q^1$. Donc le degré $d1$ de $d^\circ Q^1$ est majoré par $\mathrm{Log}_bP(b) - 1$ et par suite, $Q^1(b) \leq \frac{1}{2}b\cdot b^{d1} \leq \frac{1}{2} P(b)$. Le degré $d1^\circ$ de P^1 est donc tel que $b^{d1^\circ} \leq \frac{1}{2} P(b)$, donc $d1^\circ \leq \mathrm{Log}_bP(b)$.

Supposons que P ait deux développements distincts égaux :

$P^\circ + (X-b)\cdot P^1 + .. + (X-b)^nP^n = Q^\circ + (X-b)\cdot Q^1 + .. + (X-b)^m\cdot Q^m$.

Notons j l'indice du premier polynôme P^j distinct de Q^j. On a donc $(X-b)^jP^j + .. + (X-b)^nP^n = (X-b)^jQ^j + .. + (X-b)^mQ^m$. En calculant la dérivée $j^{ième}$ des deux membres on obtient $P^j + (X-b)R = Q^j + (X-b)S$, donc $P^j(b) = Q^j(b)$. Les coefficients des polynômes P^j et Q^j sont donc les "chiffres" du développement dans la base b du même nombre; ils sont donc égaux. Le corollaire est donc démontré.

4°) Diminution du nombre d'indéterminées.

Notons $\mathbf{Z}_b[X]$, l'ensemble des polynômes de $\mathbf{Z}[X]$ dont les coefficients appartiennent à $]-\frac{1}{2}b, \frac{1}{2}b]$. Le lemme nous permet de définir une bijection de $\mathbf{Z}_b[X]$ dans $\mathbf{Z}$. Tout polynôme de $\mathbf{Z}[X]$, dont les coefficients sont suffisamment petits peut donc être représenté par un nombre de $\mathbf{Z}$.

Considérons un polynôme P de $\mathbf{Z}[X,X_1,...,X_m] = A[X,X_1,...,X_{m-1}]$, où $A = \mathbf{Z}[X_m]$. On peut écrire P sous la forme

$$P = \sum_{j\,j1..jm} a_{j\,j1..jn}\, X^j\, X_1^{j1} \ldots X_n^{jn}, \text{ où } n = m-1 \text{ et } a_{j\,j1..jn} \in A.$$

Au lieu d'écrire $a_{j\,j1..jn}$ sous forme polynômiale, soit $a_{j\,j1..jn} = \alpha_o + ... + \alpha_{dm}\, X^{dm}$, on peut grâce au lemme, le représenter par le nombre $\alpha_o + ... + \alpha_{dm}\, b^{dm}$. Le polynôme P s'écrit donc comme un polynôme P_o de $Z[X,X_1,...,X_{m-1}]$. On a ainsi supprimé une variable, à savoir X_m, en augmentant la taille des coefficients. Si on note $\frac{1}{2}B$ une borne des coefficients du polynôme donné P, alors la taille des coefficients de P_o est majorée par $\frac{1}{2}B^{dm+1}$, où dm est le degé de P suivant X_m.

Comme cette représentation des polynômes est destinée à trouver les diviseurs de P, on devra choisir une base b de taille supérieure à deux fois les coefficients des diviseurs de P.

On utilise la borne de Mignotte donnée en **B.I.3°)**; si $P = Q \cdot R$, alors $L(Q) \cdot L(R) = 2^d \| P \|$, où D est le degré total de P et $\| P \|$ la norme euclidienne du vecteur dont les coordonnées sont les coefficients de P. L'un des diviseurs, Q ou R vérifie donc

(B.IV.11) $\quad 2 \cdot H(Q) \leq 2^{D/2+1} \| P \|^{1/2}$.

Notons pour la suite, B la partie entière du second membre de cette inégalité; alors $H(Q) \leq \frac{1}{2}B$. On doit donc choisir une base b_m telle que $b_m \geq B$; prenons pour simplifier $b_m = B$. D'après le paragraphe, tout diviseur a un degré suivant X_m majoré par dm−1. Les coefficients des diviseurs du nouveau polynôme P sont donc majorés par $\frac{1}{2}B^{dm}$, donc pour éliminer une nouvelle variable, soit X_{m-1}, on peut choisir une nouvelle base $b_{m-1} = B^{dm}$.

On peut ensuite continuer le processus et ainsi éliminer successivement toutes les variables X_m, X_{m-1}, ..., X_1. Les bases successives utilisées seront alors $b_m = B$, $b_{m-1} = B^{dm}$,..., $b_1 = B^{dm \cdots d2}$. Le polynôme obtenu, noté $P_o(X)$ a donc ses coefficients majorés par $\frac{1}{2}(b_1)^{d1}$, soit par $\frac{1}{2}B^{d1 \cdot d2 \cdots dm}$.

Le polynôme $P_o(X)$ est ensuite factorisé sur Z[X]. Puis, en partant de chacun des diviseurs $Q_o(X)$ obtenus, on doit trouver le polynôme $Q(X,X_1,...,X_m)$ qui lui correspond en utilisant l'application réciproque de Φ. Pour déterminer Q, on calcule d'abord $Q_1(X,b_1) = Q_o(X)$, obtenu en écrivant les coefficients de Q_o dans la base symétrisée b_1. On en déduit $Q_1(X,X_1)$, puis on continue les calculs jusqu'à $Q_m(X,X_1,...,X_m)$, qui est égal à $Q(X,X_1,...,X_m)$.

5°) Algorithme de calcul des images et des images réciproques.

Considérons comme ci-dessus l'anneau $A = Z[X_m]$, et notons $A[X,X_1,..,X_n]$ l'anneau $A[X,X_1,..,X_{m-1}]$. Soit P un polynôme de $A[X,X_1,..,X_n]$, B la borne définie ci-dessus par **(B.IV.11)** et $\{d_1,...,d_m\}$ la liste des degrés de P suivant $X,...,X_m$.

Algorithme du calcul de $P_o = \Phi(P)$.

On initialise l'algorithme en posant

$b_{mp+1} = B$, $d_{mp+1} = 1$ et $P_m(X,X_1,...,X_m) = P(X,X_1,...,X_m)$.

Pour j décroissant de m jusqu'à 1, faire :

- prendre $\delta = d_{j+1}$, puis poser $b_j = (b_{j+1})^\delta$;
- remplacer X_j par b_j dans le polynôme $P_j(X,...,X_j)$;
- noter $P_{j-1}(X,.. ,X_{j-1})$ le polynôme obtenu;

Le résultat est le polynôme $P_o(X)$.

Algorithme du calcul de $Q = \Phi^{-1}(Q_o)$.

Pour chaque coefficient $c \in Z$ de $Q_o(X)$,

- écrire c dans la base symétrisée B, sous la forme suivante $\alpha_o + \alpha_1 \cdot B + .. + \alpha_d \cdot B^d + ...$, où $-\frac{1}{2}B < \alpha_j \leq \frac{1}{2}B$, puis écrire chaque exposant D de B dans la base composée $\{d_m, d_m \cdot d_{m-1}, ..., d_m \cdots d_2\}$, c'est-à-dire sous la forme $D = \xi m + ... + \xi 2\, d_m \cdots d_3 + \xi 1\, d_m \cdots d_2$.
- Remplacer le monôme $\alpha_d B^d$ par $\alpha_d X_1^{\xi 1} X_2^{\xi 2} \cdots X_m^{\xi m}$.

La somme de tous les monômes ainsi obtenus à partir des monômes $\alpha_d B^d$ est le polynôme $Q(X,...,X_m)$.

6°) Exemple.

$$\begin{aligned} P(X,Y,Z) = {} & X^6Y^2 && + X^4(Y^3Z + Z^4 + 1) \\ & + X^3(Y^4 + Y^2Z^3 + Z) && + X^2Y(Z^5 + Z) \\ & + X((Z^4 + 1)Y^2 + Z^2Y + Z^7 + Z^3) + ZY^2 && + Z^4. \end{aligned}$$

Le degré total étant 8, on a $B = 2^5 \, \| P \|^{1/2} = 64$. Le degré de P suivant Y vaut $d_Y = 4$; donc il est majoré par 3 dans chaque diviseur. On remplacera donc Y par $64 = 2^6$, puis Z par $64^4 = 2^{24}$.

Si on programme en REDUCE, on obtient :

$P := X^6 Y^2 + X^4 Y^3 Z + X^4 Z^4 + X^4 + X^3 Y^4 + X^3 Y^2 Z^3 + X^3 Z + X^2 Y Z^5$
$+ X^2 Y Z + X Y^2 Z^4 + X Y^2 + X Y Z^2 + X Z^7 + X Z^3 + Y^2 Z + Z^4.$

P° := SUB(Y=64, Z=64**4, P);

$P^\circ := 4096\, X^6 + 79228162514264341991590461441\, X^4 +$
$19342813113834066828853248\, X^3 +$
$85070591730234615865843651859015794688\, X^2 +$
$37414441915671114738466187083851756431234290627788 8\, X +$
$7922816251426433766226342707 2.$

On factorise le polynôme Po en utilisant la méthode classique de factorisation des polynômes à une variable. On trouve

$P^\circ = Q^\circ \cdot R^\circ = (X^3 + 1073741824\, X + 4722366482869645217792)$
$(4096\, X + 79228162514264337593543950337\, X + 16777216);$

Chacun des cinq coefficients des deux diviseurs doit maintenant être écrit dans la base 64. On trouve

1073741824	$= 64^5$	$= 64^4 \cdot 64;$
4722366482869645217792	$= 64^12 + 64^2$	$= (64^4)^3 + 64^2;$
4096	$= 64^2$	$= 64^2;$
79228162514264337593543950337	$= 64^{16} + 1$	$= (64^4)^4 + 1$;
16777216	$= 64^4$	$= 64^4.$

Par suite $Q^\circ = X^3 + (64^4) \cdot 64 \cdot X + (64^4)^3 + 64^2$ et

$R^\circ = 64^4 \cdot X^3 + ((64^4)^4 + 1) \cdot X + 64^4.$

En remplaçant 64^4 par Z, puis les autres "64" par Y, on obtient

$Q = X^3 + Z \cdot Y \cdot X + Z^3 + Y^2$ et $R = Y^2 \cdot X^3 + (Z^4+1) \cdot X + Z.$

7°) Etude du coût.

Les calculs sont dominés par la factorisation du polynôme P° dont les coefficients sont très grands. La factorisation modulo un nombre premier p par l'algorithme de Berlekamp exige $p\ d^2\ (\ d + Log p\ )$ opérations, où d représente ici le degré de P suivant la variable X. Le coût de cette factorisation reste donc faible. Par contre son raffinement grâce au lemme de Hensel, exige comme on l'a vu en **(A.II.14)**, de faire $d^2\ Log^2\ (p^n)$

opérations, où p^n majore les coefficients des diviseurs de P^o. On a vu ci-dessus au paragraphe **4°)** que ces coefficients sont majorés par $\frac{1}{2}B^{d1 \cdot d2 \cdots dm}$, où
$B = 2^{D/2+1} \| P \|^{1/2}$.
Le coût total est donc de l'ordre de
$(d \cdot d1 \cdot d2 \cdots dm)^{2} (d + d1 + \cdots + dm + \text{Log} \| P \|)^2$.
Majorons $d \cdot d1 \cdot d2 \cdots dm$ par δ^{m+1} où δ est le degré moyen des variables. On obtient alors comme coût total

(B.IV.12) $\delta^{2m+2} ((m+1) \cdot \delta + \text{Log} \| P \|)^2$.

Ce coût est donc du même ordre que celui de l'algorithme originel de Wang et de Musser, qui exige pour le raffinement des diviseurs le calcul de produits de polynômes à $m+1$ variables, qui donnent des calculs de l'ordre de $\delta^{2m+2} (m \cdot \delta + \text{Log} \| P \|)^2$. Cependant l'algorithme proposé est moins complexe et il devient intéressant lorsque la borne $\frac{1}{2}B^{d1 \cdot d2 \cdots dm}$ des coefficients du polynôme à une variable qu'on doit factoriser est relativement petite. Le coût de la factorisation est alors celui de la factorisation d'un polynôme à une variable.

c) Factorisation utilisant un monomorphisme.

1°) Introduction.

On se propose d'utiliser un monomorphisme Φ_n analogue à celui qui a été défini dans **(A.V)** et qui transforme un produit en une somme. On utilise ce monomorphisme pour éviter le calcul des produits de polynômes de la forme $Q_{j\,1}\cdots Q_{j\,u}$ à la troisième étape de l'algorithme, lorsqu'il y a des facteurs parasites. Le calcul de ces produit est remplacé par le calcul des sommes des images de ces polynômes, à savoir $\Phi(Q_{j\,1}) + \cdots + \Phi(Q_{j\,u})$.

Cet algorithme ne sera donc utilisé que lorsqu'aucun diviseur de P sur $Z[X,X_1,...,X_m]/\Delta_j$, ne divise P sur $Z[X,X_1,..,X_m]$. De plus, le critère donné évitera d'avoir à effectuer la division de P par le diviseur présumé, et en premier lieu les divisions de P par $Q_1,..., Q_r$.

On utilise la notion de polynôme normalisé introduite en **II**, et on suppose que le polynôme donné P a été normalisé. L'algorithme présenté ici, s'applique après la deuxième étape de la factorisation; on connait alors la factorisation de P sous la forme $Q_1\cdots Q_j\cdots Q_m$ où les Q_j sont des polynômes normalisés de $Z[X,X_1,...,X_m]/\Delta_d$.

Le monomorphisme Φ doit donc être défini pour un polynôme normalisé de $Z[X,X_1,...,X_m]/\Delta_j$. Pour définir Φ, on reprend les mêmes notations que dans **(B.V)**, en notant le polynôme U dont on calcule l'image sous la forme $U(X,X_1,...,X_m) = X^n + b_1 X^{n-1} + ... + b_n$, avec $b_1, ..., b_d \in Z[X_1,...,X_m]/\Delta_d$.

2°) Définition du monomorphisme.

Dans la suite, on notera K l'anneau $Z[X_1,...,X_m]/\Delta_d$. Rappelons qu'on peut définir, comme on l'a vu en **I**, le quotient de tout polynôme de $K[X]$ par un polynôme unitaire

U de **K**[X]. Plus précisément, à tout polynôme V de **K**[X], on peut associer deux polynômes Q et R tels que

(B.IV.13) $V = U \cdot Q + R$ où $d^\circ R \leq n - 1$.

Notons d un entier fixé, qui dans la suite sera égal au degré suivant X du polynôme P qu'on veut factoriser.

Définition.

On considère le polynôme unitaire U(X) de **K**[X], et on note U'(X) sa dérivée.
On appelle développement logarithmique à l'ordre d de U(X), le quotient du polynôme $X^{d+1} \cdot U'(X)$ par le polynôme U(X).
Le développement de U(X) est un polynôme de degré d noté
$U^\circ(X) = b^\circ_0 X^d + b^\circ_1 X^{d-1} + \ldots + b^\circ_d$.

Par exemple, le développement à l'ordre 3 du polynôme $U(X,Y,Z) = X + Y + Z + 1$ est égal à $U^\circ(X,Y,Z) = 3X^3 - (Y + Z + 1) X^2 + (Y + Z + 1)^2 X - (Y + Z + 1)^3$.
On retrouve sans difficultés les mêmes propriétés que dans **(A.V)**, qu'on rappelle dans la proposition suivante :

Proposition 1.

Soient U et V deux polynômes unitaires. Alors $\Phi(U \cdot V) = \Phi(U) \cdot \Phi(V)$.
Les coefficients b°_j de $\Phi(U)$ sont définis par les relations:

$$b^\circ_0 = d;$$
$$b^\circ_1 = -b_1;$$
$$b^\circ_j = -j\, b_j - b_{j-1}\, b^\circ_1 - \ldots - b_1\, b^\circ_{j-1} \qquad 1 \leq j \leq q;$$
$$b^\circ_j = -b_d\, b^\circ_{j-d} - \ldots - b_1\, b^\circ_{j-1} \qquad q \leq j \leq d.$$

On peut déduire des relations ci-dessus que b_j est défini comme une suite récurrente linéaire en fonction de $\Phi(U)$. Par suite Φ est un monomorphisme défini sur l'ensemble des polynômes unitaires de **K**[X].
L'algorithme de recherche des diviseurs de P est basé sur la notion de polynôme normalisé. Commençons par donner un critère simple pour vérifier si un polynôme est normalisé.

Proposition 2.

Considérons le polynôme Q un polynôme de K[X], qui s'écrit sous la forme : $X^q + b_1 X^{q-1} + ... + b_k X^{q-k} + ... + b_q$, et notons $d°b_k$ le degré de b_k suivant $X_1,...,X_m$.

1°) Q est normalisé si et seulement si

(B.IV.14) $d° b_k \leq k$ pour $1 \leq k \leq q$.

2°) Si Q et R sont deux polynômes normalisés de K[X], alors Q + R et Q·R sont normalisés.

Démonstration.

1°) Dire que Q est normalisé, signifie d'après **(B.I.12)**, que $\delta Q_N^{(k)} \leq \delta Q_N^{(o)} - k$ pour $k \geq 0$, δ étant le degré suivant X. Cela signifie que le degré total de $Q_N^{(k)}$, à savoir $\delta Q_N^{(k)} + k$ est majoré par le degré q suivant X, ou encore que $d°b_k + q - k \leq q$, soit $d°b_k \leq k$.

2°) Notons R sous la forme $X^q + b_1 X^{q-1} + ... + b_q$, et supposons Q et R normalisés. Alors $d°b_k \leq k$ et $d°c_k \leq k$. Par suite le $k^{ième}$ coefficient de Q + R, à savoir $b_k - c_k$ vérifie **(B.IV.14)**. Donc Q + R est normalisé. Le $k^{ième}$ coefficient de Q·R s'écrit $c_k = > a_j \cdot b_n$ et il vérifie également **(B.IV.14)**; donc Q·R est bien normalisé.

Rappel:

Rappelons quelques résultats démontrés au **B.III.2°), Lemme 2** :

1°) Si P = Q·R est normalisé et si Q est normalisé, alors R est également normalisé.

2°) Le polynôme Q de K[X] divise P sur $Z[X,X_1,...,X_m]$ si et seulement si Q est normalisé.

La proposition suivante nous permettra de vérifier si Q divise P sur $Z[X,X_1,...,X_m]$, en vérifiant la normalisation de $\Phi(Q)$.

Proposition 3.

Soit P un polynôme normalisé de K[X], qui est factorisé sous la forme $Q_1 \cdots Q_r$ sur K[X] et Q un sous-produit $Q_{j_1} \cdot Q_{j_2} \cdots Q_{j_u}$ de $Q_1 \cdots Q_r$ de degré majoré par ½d.

Les deux propriétés suivantes sont équivalentes :

(i) Q est normalisé;

(ii) $\Phi(Q)$ est normalisé.

Démonstration.

Notons R le polynôme de $\mathbf{K}[X]$, quotient de P par Q. Notons Q sous la forme $X^q + b_1 X^{q-1} + ... + b_q$ et son image par Φ s'écrit $Q^\circ = q \cdot X^d + b^\circ_1 X^{d-1} + ... + b^\circ_d$.

Supposons d'abord que (ii) soit vérifié.

D'après le lemme, on a $d^\circ(b_k) \leq k$, pour $1 \leq k \leq q$. Montrons qu'on a également $d^\circ(b^\circ_k) \leq k$ pour $1 \leq k \leq d$. L'inégalité est vraie pour $k = 1$. Or d'après la **Proposition 1**, $d^\circ(b^\circ_k) \leq \max(d^\circ(b_k), d^\circ(b_{k-1}) + d^\circ(b^\circ_1),...,d^\circ(b_1) + d^\circ(b^\circ_{k-1}))$, pour $1 \leq k \leq q$, et $d^\circ(b^\circ_k) = \max(d^\circ(b_d) + d^\circ(b^\circ_{k-d}),...,d^\circ(b_1) + d^\circ(b^\circ_{k-1}))$, pour $q \leq k \leq d$. Il en résulte donc que $d^\circ(b^\circ_k) \leq k$ pour $1 \leq k \leq d$. Par suite, $\Phi(Q)$ est normalisé.

Supposons maintenant que $d^\circ(b^\circ_k) \leq k$ pour $1 \leq k \leq d$. Alors on a $d^\circ(b_1) = 1$; d'autre part, pour $k > 1$, $d^\circ(b_k) \leq \max(d^\circ(b^\circ_k), d^\circ(b_1) + d^\circ(b^\circ_{k-1}),...,d^\circ(b_{k-1}) + d^\circ(b^\circ_1))$, pour $1 \leq k \leq q$, donc $d^\circ(b_k) \leq k$. Il en résulte que Q est normalisé.

L'intérêt essentiel de ce résultat est qu'on peut déterminer $\Phi(Q)$ en calculant la somme $\Phi(Q_{j\,1}) + ... + \Phi(Q_{j\,N})$. Il n'y a donc plus aucun produit de polynômes à calculer, mais uniquement des sommes.

3°) Algorithme utilisant le monomorphisme Φ.

Supposons qu'on écrive les coefficients b°_k de Q° sous la forme suivante $b^\circ_k = \sum_{k\,k1..km} \beta_{k\,k1..km} X^{k1}...X^{km}$. Alors Q est normalisé si et seulement si $\beta_{k\,k1..km} = 0$ pour tout m-uple (k1,...,km) tel que $k1 + ... + km > k$, pour $1 \leq k \leq d$. Si on écrit les coefficients des polynômes $Q^\circ_1,..,Q^\circ_r$ sous la forme $\beta^1_{k\,k1..km},..., \beta^r_{k\,k1..km}$, alors le produit $Q_{j\,1}\cdots Q_{j\,u}$ correspond à un polynôme normalisé, si et seulement si on a les relations suivantes :

(B.IV.15) $\beta^{j1}_{k\,k1..km} + ... + \beta^{ju}_{k\,k1..km} = 0$ pour $k1 +...+ km > k$.

La recherche des polynômes $Q_{j\,1},...,Q_{j\,u}$ qui satisfont **(B.IV.15)** peut se faire en résolvant le système suivant, dans lequel les inconnues $e_1,...,e_r$, valent 0 ou 1 :

(B.IV.16) $\begin{cases} \beta^{1}_{k\,k1..km}\, e_1 + ... + \beta^{r}_{k\,k1..km}\, e_r = 0 \text{ pour tout m-uple } (k1,...,km) \\ \text{tel que } k1 + ... + km > k, \text{ avec } 1 \le k \le d. \end{cases}$

Dans la suite, ce système sera appelé **système de normalisation**. Notons t le rang de ce système et soit $\sigma = r - t$. Notons E le sous-espace vectoriel de $\mathbf{Q}^r$ formé par les solutions du système de normalisation **(B.IV.16)**; en résolvant ce système par la méthode du pivot, on obtient une base de E de la forme :

$$B_1 = \begin{pmatrix} b_{11} \\ .. \\ b_{t1} \\ 1 \\ 0 \\ .. \\ 0 \end{pmatrix}, \ldots\ldots\ldots, B_i = \begin{pmatrix} b_{1i} \\ .. \\ b_{ti} \\ 0 \\ 0 \\ .. \\ 0 \end{pmatrix}, \ldots\ldots\ldots, B_\sigma = \begin{pmatrix} b_{1\sigma} \\ .. \\ b_{t\sigma} \\ 0 \\ 0 \\ .. \\ 1 \end{pmatrix}$$

Notons B la base $\{B_1, ..., B_\sigma\}$.

Considérons un diviseur Q de P sur $\mathbf{Z}[X,X_1,..,X_m]$ qui s'écrit $Q_u \cdot Q_v \cdots Q_w$ sur $\mathbf{Z}[X,X_1,..,X_m]/\Delta_d$. D'après le Rappel, 2°) ci-dessus, Q est normalisé. Par suite, le vecteur $V = (e_1,...,e_r)$ associé à Q, dont les coordonnées e_i vérifient $e_u = e_v = .. = e_w = 1$ et $e_i = 0$ pour les autres valeurs de i, doit être engendré par la base B. Donc Q doit contenir l'un au moins des facteurs $Q_{t+1},...,Q_r$. Comme P est sans facteur carré, le vecteur V est donc obtenu en faisant la somme de un ou plusieurs des vecteurs de B. L'une des solutions du système de normalisation est donnée par le vecteur $B_o = (1,1,...,1)$, qui est obtenu en faisant la somme de tous les vecteurs de B. Cette solution correspond au polynôme P lui-même.

Si certains des vecteurs B_i ont toutes leurs coordonnées égales à 0 ou 1, alors ces vecteurs B_i correspondent à des produits $Q_u \cdot Q_v \cdots Q_w$ qui sont normalisés, donc qui sont des diviseurs irréductibles de P sur $\mathbf{Z}[X,X_1,..,X_m]$. On dira dans la suite, qu'un tel vecteur ayant ses coordonnées égales à 0 ou 1 est **binaire** et on notera B° l'ensemble des vecteurs non binaires de B.

Tout autre diviseur irréductible de P (ne correspondant à aucun vecteur de B), est associé à une autre solution du système de normalisation; cette solution est obtenue en effectuant une somme de vecteurs de B°. Il est clair que le nombre des diviseurs de P sur $\mathbf{Z}[X,X_1,..,X_m]$, noté r° est majoré par σ. Si $r^\circ = \sigma$, alors tous les vecteurs de

B sont binaires et la factorisation de P est obtenue en calculant les produits $Q_u \cdot Q_v \cdots Q_w$, associés aux vecteurs de B.

Remarquons aussi que tout vecteur de B_i est associé à un diviseur irréductible de P égal à un produit $Q_u \cdot Q_v \cdots Q_w$ contenant un et un seul des diviseurs Q_{t+1}, .., Q_r. Donc si P ne contient aucun diviseur de ce type, alors aucun des vecteurs B_i ne correspond à un diviseur de P.

Le nombre d'équations du système de normalisation est très grand. Ce sont les équations qui vérifient $k1 + ... + km > k$, avec $1 \leq k \leq d$; leur nombre est donc majoré par $(d + 1)^m$. Mais parmi elles, t seulement sont linéairement indépendantes. Comme la plupart de ces équations sont inutiles, on pourrait n'en choisir qu'un nombre limité, par exemple d. Mais, alors les solutions du système pourraient correspondre à des produits $Q_u \cdot Q_v \cdots Q_w$ non normalisés, qui ne seraient donc pas des diviseurs de P.

Le nombre d'équations du système de normalisation, donne un coût de résolution de l'ordre de $(d + 1)^m \cdot r^2$.

Le nombre maximum de sommes de vecteurs à calculer est égal à $2^{s°}$, où $s°$ désigne le nombre des vecteurs de $B°$. Comme les vecteurs ont chacun t coordonnées, le nombres total d'additions d'entiers est donc majoré par le produit $t \cdot 2^{s°}$, avec $s° \leq r - t$.

Au lieu de limiter le nombre d'équations du système de normalisation à $\frac{1}{2}(d + 1)^m$, imaginons qu'on résolve ce système avec toutes les équations telles que $k > k1 + ... + km$ (le nombre de ces équations étant arbitrairement grand). On peut ensuite transformer la base B de façon à rendre ses coefficients entiers; il suffit de multiplier chacun des vecteurs B_1, ..., B_σ par le nombre δ égal au PPCM des dénominateurs des coefficients b_{ij}. On peut enfin rendre positifs les coefficients de chacun des vecteurs de base B_i en lui ajoutant le vecteur $n_i B_o$, n_i étant le plus grand coefficient négatif de $\delta \cdot B_i$. Notons B'_i les nouveaux vecteurs obtenus, qui forment une nouvelle base B', dont les coefficients sont dans N, l'un au moins de ces coefficients étant nul, dans chaque vecteur B'_i.

Notons (m1,...,mr) les coefficients de B'_i. Le polynôme associé à B'_i, à savoir $R_i = (Q_1)^{m1} \cdots (Q_r)^{mr}$ est normalisé, à un ordre quelconque, puisqu'on a supposé que le nombre d'équations du système de normalisation n'est pas limité. Si on pose, $mo = \max (m1,...,mr)$, alors R_i divise P^{mo} sur $Z[X,X_1,..,X_m]/\Delta_d$. D'après la Proposition 2 ci-dessus, P^{mo} est normalisé. Il en résulte, d'après le 2°) du Rappel donné

ci-dessus, que R_i divise P^{mo} sur $Z[X,X_1,..,X_m]$, par suite R_i n'est pas premier avec P. Le PGCD S de R_i et P doit être égal sur $Z[X,X_1,..,X_m]/\Delta_d$ à un produit $(Q_1)^{n1}\cdots(Q_r)^{nr}$ où n_1, .., n_r valent 0 ou 1. Par suite $R°_i = R_i/S$ divise toujours P^{mo}, alors qu'il est premier avec P; donc $R°_i = 1$. Par suite, $m1 = ... = mr = D + n_i$; donc $n_i = 0$, $D = 1$ et les coordonnées non nulles de B_i sont toutes égales à 1.

La conclusion est que si on prend suffisamment d'équations dans le système de normalisation, les vecteurs de base des solutions de ce système sont binaires et donnent directement les diviseurs irréductibles de P sur $Z[X,X_1,..,X_m]$.

Alors, la dimension de l'espace des solutions est égale au nombre de facteurs irréductibles de P. De plus, comme P n'a pas de facteurs carré, la base B est orthogonale.

Par contre, si on ne résoud le système de normalisation qu'avec un nombre limité d'équations (de l'ordre de $½(d + 1)^m$), la résolution du système peut donner d'avantage de solutions. Une solution supplémentaire non binaire correspond à une relation linéaire parasite entre les équations du système; chaque équation qu'on ajoute au système aura donc tendance à augmenter le rang t du système et ainsi à diminuer le nombre des vecteurs non binaires de la base B.

Rappelons qu'on doit effectuer toutes les sommes de vecteurs non binaires de B et déterminer celles qui sont égales à un vecteur binaire. Le coût est alors majoré par $t\cdot 2^{s*}$, s^* étant le nombre des vecteurs non binaires.

En fait, on peut diminuer le coût de ce calcul en partionnant l'ensemble B^* des vecteurs non binaires de B. Pour cela, on introduit la définition suivante concernant les vecteurs non nuls de Q^r :

Un vecteur V est dit **positif** si sa première coordonnée non nulle est positive; dans le cas contraire, le vecteur est dit **négatif**. On note B^+ et B^- les vecteurs respectivement positifs et négatifs de B^*. De plus, notons $B°$ l'ensemble obtenu en adjoignant à B^* l'ensemble des opposés des vecteurs unitaires de Q^r. Le calcul des sommes des vecteurs de B^* qui sont binaires revient alors au calcul des sommes des vecteurs de $B°$ qui sont nulles. L'algorithme de recherche de ces sommes consiste à partir d'un vecteur quelconque de B^*.

S'il est négatif, on lui ajoute un vecteur de B^+ et s'il est positif, sa $i^{ième}$ coordonnée étant positive, on lui ajoute soit un vecteur de B^-, soit l'opposé du $i^{ième}$ vecteur unité.

On désigne par n_1 et n_2 le nombre d'éléments de B^+ et B^-; comme $n_1 + n_2 = s^*$, on considère que $n_1 = n_2 = ½\ s^*$. Alors la recherche d'une solution qui est la somme de j vecteurs de B° conduit à effectuer un nombre de sommes approximativement égal à $½s^*\cdot½(s^*-1)\cdots½(s^*-j)/j!$, soit $(½)^j \binom{s^*}{j}$.

Le nombre total de sommes à calculer est donc de l'ordre de $(3/2)^{s^*}$.

Le coût total de la troisième étape de la factorisation du polynôme P, consistant à déterminer les produits $Q_u\cdot Q_v\cdots Q_w$ qui divise P sur $\mathbf{Z}[X,X_1,...,X_m]$, est obtenu en ajoutant le coût de la résolution du système de normalisation par la méthode du pivot, et le coût de la recherche des sommes des solutions du système de normalisation qui sont binaires. Ce coût total noté CT est donc égal à

(B.VI.5) $CT = (d + 1)^{m+2}$ multiplications d'entiers + $t\cdot(3/2)^{s^*}$ additions d'entiers.

Le coût reste donc exponentiel suivant le degré d, car on ne peut majorer s^* que par d; mais il est nettement meilleur que le coût de l'algorithme classique qui est de 2^r produits de polynômes à m variables.

Algorithme de recherche des diviseurs de P.

1ère étape:

On calcule la factorisation de P sur $\mathbf{K}[X]$, notée $Q_1\cdots Q_r$.

2ème étape:

On calcule des images $Q^\circ_1,...,Q^\circ_r$ de $Q_1,...,Q_r$ par Φ et on note $ß^j_{k\ k1..km}$ $X^{k1}...X^{km}$ les monômes de Q°_j, les polynômes $Q^\circ_1,..., Q^\circ_r$ étant de degré d.

3ème étape:

On résoud le système de normalisation par la méthode du pivot, puis on effectue toutes les sommes des vecteurs de base non binaires de la solution.

Chaque vecteur binaire obtenu correspond à un produit $Q_u\cdot Q_v\cdots Q_w$ qui est un diviseur de P sur $\mathbf{Z}[X,X_1,...,X_m]$.

Etude du coût.

Le coût de la première étape est celui de l'algorithme classique, à savoir $m\cdot d\cdot(d + 1)^{2m+2}$.

Le nombre de produits à calculer pour trouver les images $\Phi(Q_1),..,\Phi(Q_r)$ est majoré par $½(d+1)^2$. Comme il s'agit de produits de polynômes à m variables de degré d, le coût de ce calcul est majoré par $(d+1)^{2m+2}$.

L'ordre de grandeur du coût est donc du même ordre que celui de l'algorithme classique. Pratiquement, l'algorithme proposé devient intéressant s'il y a beaucoup de facteurs parasites et si le nombre des produits de polynômes $Q_{j\,1}\cdots Q_{j\,r}$ qu'on doit calculer est grand. Par contre, le coût théorique des calculs est devenu presque polynômial, puisque la partie exponentielle du coût, $t\cdot(3/2)^{r-1}$ représente des additions d'entiers.

4°) Exemple.

Considérons le polynôme suivant, qui est normalisé

$$P = X^4 + 2X^3 - X^2 - 2X + (9X^3 + 10X^2 + X - 2)Y + 9XY^2.$$

On factorise $P^{(o)} = P(X,0) = X^4 + 2X^3 - X^2 - 2X$; on obtient

$$P^{(o)} = X(X-1)(X+1)(X+2) = {}_{Q1}{}^{(o)}\cdot Q_2^{(o)}\cdot Q_3^{(o)}\cdot Q_4^{(o)}.$$

Calculons les remontées des diviseurs $Q_j^{(o)}$ jusqu'au degré 4 suivant Y. On obtient ainsi la factorisation de P sur $\mathbf{Z}[X,Y]/(Y^5)$:

$$Q_1 = X + Y + Y^2 + 2Y^3 + 5Y^4;$$
$$Q_2 = X - 1 + 3Y - 6Y^2 + 6Y^3 + 6Y^4;$$
$$Q_3 = X + 1 - Y - Y^2 - 2Y^3 - 5Y^4;$$
$$Q_4 = X + 2 + 6Y + 6Y^2 - 6Y^3 - 6Y^4.$$

On constate sans calcul qu'aucun de ces quatre polynômes n'est normalisé, donc aucun ne divise P sur $\mathbf{Z}[X,Y]$. Pour éviter de calculer les produits $Q_j\cdot Q_k$, déterminons les images $\Phi(Q_j)$ jusqu'à l'ordre 2, puisqu'on recherche un diviseur de degré maximum égal à 2. On obtient $Q^\circ_j = b^\circ_o X^2 + b^\circ_1 X + b^\circ_2$ où $b^\circ_o = d^\circ Q_j$, $b^\circ_1 = -b_1$ et $b^\circ_2 = (b_1)^2$. On trouve

$$Q^\circ_1 = X^2 - X(Y + Y^2 + 2Y^3 + 5Y^4) + Y^2 + 2Y^3 + 5Y^4;$$
$$Q^\circ_2 = X^2 + X(1-3Y+6Y^2-6Y^3-6Y^4) + 1 - 6Y + 21Y^2 - 48Y^3 + 60Y^4;$$
$$Q^\circ_3 = X^2 - X(1-Y-Y^2-2Y^3-5Y^4) + 1 - 2Y - Y^2 - 2Y^3 - 5Y^4;$$
$$Q^\circ_4 = X^2 - X(2+6Y+6Y^2-6Y^3-6Y^4) + 4 + 24Y + 60Y^2 + 48Y^3 - 60Y^4.$$

Les polynômes de la forme $Q^\circ_j + Q^\circ_k$ sont normalisés si et seulement si leurs coefficients des monômes XY^2, XY^3, XY^4, Y^3 et Y^4 s'annulent.

Les ensembles d'indices $\{j, k\}$ correspondant à des polynômes $Q^\circ_j + Q^\circ_k$ normalisés, sont donc solution du système suivant dans lequel e_1, e_2, e_3, e_4 sont égaux à 0 ou 1

$$
\begin{array}{rcrcrcrcl}
-e_1 & + & 6\,e_2 & + & e_3 & - & 6\,e_4 & = & 0 \\
-2\,e_1 & - & 6\,e_2 & + & 2\,e_3 & + & 6\,e_4 & = & 0 \\
-5\,e_1 & - & 6\,e_2 & + & 5\,e_3 & + & 6\,e_4 & = & 0 \\
2\,e_1 & - & 48\,e_2 & - & 2\,e_3 & + & 48\,e_4 & = & 0 \\
5\,e_1 & + & 60\,e_2 & - & 5\,e_3 & - & 60\,e_4 & = & 0 \\
e_1 & + & 60\,e_2 & - & 5\,e_3 & - & 60\,e_4 & = & 0
\end{array}
$$

Les solutions sont $\{e_1, e_2, e_3, e_4\} = \{1, 0, 1, 0\}, \{0, 1, 0, 1\}$ et $\{1, 1, 1, 1\}$. Il y a donc trois diviseurs de P, à savoir $Q_1 \cdot Q_3$, $Q_2 \cdot Q_4$ et $Q_1 \cdot Q_2 \cdot Q_3 \cdot Q_4$. En effectuant les deux premiers produits on trouve les deux diviseurs suivants

$Q = Q_1 \cdot Q_3 = X^2 + X + Y$ et $R = Q_2 \cdot Q_4 = X^2 + X - 2 + 9\,X\,Y$.

Partie : C

SIMPLIFICATION DES POLYNOMES A PLUSIEURS VARIABLES

INTRODUCTION.

Dans toute cette partie, les polynômes appartiennent à l'anneau $Z[X_1,..,X_n]$, des polynômes à n varaibles, à coefficients entiers.

On considère m polynômes de $Z[X_1,...,X_n]$, notés $P_1,...,P_m$ sur lesquels on veut effectuer un calcul ne faisant intervenir que des produits, par exemple le calcul du PGCD de $P_1,...,P_m$ ou la factorisation de l'un des polynômes P_i.

Le but de ce chapitre est la recherche d'une application ϕ qui transforme $P_1,...,P_m$ en polynômes plus "simples", suivant certains critères. Par exemple, l'un des transformés de $P_1, .., P_m$ aura un degré diminué par rapport à une des variables $X_1, .., X_n$.

Mais, l'application ϕ doit être injective, pour que les calculs qui sont effectués sur $\phi(P_1), .., \phi(P_m)$ et qui donnent le résultat V, permettent de trouver sans ambiguïté le résultat $\phi^{-1}(V)$, qui aurait été trouvé en effectuant les calculs sur $P_1,...,P_m$. Par contre, il n'est pas nécessaire de supposer ϕ surjective.

La première partie de ce chapitre reprend l'article **[V3]**. Comme les calculs envisagés (calcul de PGCD, factorisation) ne font intervenir que le produit des polynômes, l'application doit simplement conserver les produits, c'est-à-dire vérifier la relation

$$\phi(P \cdot Q) = \phi(P) \cdot \phi(Q).$$

Pour des raisons techniques, on considère dans la suite des polynômes généralisés, dont les exposants des monômes peuvent être négatifs. D'autre part, on utilise la représentation des monômes des polynômes par des points de $\mathbf{Z}^n$. On obtient les définitions suivantes :

Définition.

On appelle polynôme généralisé, un polynôme $P = a_1 \cdot m_1 + \ldots + a_k \cdot m_k$, où $a_1, .., a_k \in \mathbf{Z}$ et où les monômes $m_1, .., m_k$ ont leurs exposants dans $\mathbf{Z}$. On note $\mathbf{Z}\{x_1,\ldots,x_n\}$ l'anneau des polynômes généralisés.

Pour chaque monôme $a_j \cdot m_j$ de P, on note M_j le point de $\mathbf{Z}^n$ qui a comme coordonnées les exposants des variables de m_j. On note $\mu(M_j)$ le monôme m_j associé au point M_j de $\mathbf{Z}^n$.

On appelle polyèdre associé à P l'enveloppe convexe notée $\Pi(P)$ des points $M_j = \mu^{-1}(m_j)$ de l'espace affine $\mathbf{Z}^n$.

On appelle sommet du polyèdre convexe $\Pi(P)$ un point de $\Pi(P)$ qui n'appartient à aucun segment $[A, B]$ avec $A, B \in \Pi(P)$.

On appelle monôme extrêmal de P un monôme m_j dont le point associé dans $\mathbf{Z}^n$ est un sommet de $\Pi(P)$. Si Π' et Π'' sont des polyèdres convexes de $\mathbf{Z}^n$, on appelle somme de Π' et Π'' et on note $\Pi' + \Pi''$ l'enveloppe convexe des points $M'_j + M''_j$ où $M'_j \in \Pi'$ et $M''_j \in \Pi''$.

On appelle face du polyèdre Π l'intersection de Π avec un hyperplan d'appui (les points de Π sont tous situés du même côté de l'hyperplan). Un sommet est une face de dimension 0. Si V est un vecteur de $\mathbf{Z}^n$, on note Π_V la face définie par l'hyperplan d'appui H perpendiculaire à V, V étant dirigé vers l'extérieur du polyèdre.

Un résultat essentiel est la relation suivante :

(C.1) $(\Pi' + \Pi'')_V = \Pi'_V + \Pi''_V$.

Cette notion de polyèdre associé à un polynôme a été définie par Ostrowski dans [Os], puis utilisée pour donner des critères d'irréductibilité des polynômes. Le résultat essentiel est donné dans le lemme suivant :

Lemme.

Soient deux polynômes P et Q de $\mathbf{Z}[X_1,...,X_n]$. Alors:

1) $\Pi(P\cdot Q) = \Pi(P) + \Pi(Q)$.

2) Tout monôme extrêmal de $P\cdot Q$ s'écrit de façon unique sous la forme $m'\cdot m''$ où m' et m" sont des monômes de P et Q; de plus m' et m" sont nécessairement des monômes extrêmaux de P et Q.

Démonstration:

1) Si $M \in \Pi(P\cdot Q)$, alors $\mu(M) = m'\cdot m''$ où m' et m" sont des monômes de P et Q. Donc $M = \mu^{-1}(m') + \mu^{-1}(m'') \in \Pi(P) + \Pi(Q)$.

Montrons maintenant que tout sommet M de $\Pi(P) + \Pi(Q)$ appartient à $\Pi(P\cdot Q)$. On a $M = M' + M''$ où $M' \in \Pi(P)$ et $M'' \in \Pi(Q)$. Si $m'\cdot m''$ qui s'écrit $\mu(M')\cdot\mu(M'')$ n'était pas un monôme de $P\cdot Q$, il devrait s'annuler avec un autre produit $m'_1\cdot m''_1$. Alors M serait le milieu du segment d'extrèmités $M_1 = M' + \mu^{-1}(m''_1)$ et $M_2 = \mu^{-1}(m'_1) + M''$. C'est contraire au fait que M est un sommet, donc $m'\cdot m''$ est un monôme de $P\cdot Q$ et 1) est vérifié.

2) Soit m un monôme extrêmal de $P\cdot Q$ associé à un sommet M de $\Pi(P) + \Pi(Q)$. Alors $M = M' + M''$ où M' et M" sont des sommets de $\Pi(P)$ et $\Pi(Q)$ d'après **(C.1)**. Si on avait aussi $M = A' + A''$, où $A' \in \Pi(P)$ et $A'' \in \Pi(Q)$, alors M appartiendrait au segment $[M'+M''_1, M'_1+M'']$, donc il ne serait pas un sommet.

I. APPROCHE THEORIQUE.

1°) Définition de l'application ϕ.

Au départ les polynômes donnés appartiennent à $Z[X_1,...,X_n]$, mais au cours des calculs de simplification on obtient des polynômes généralisés appartenant à $Z\{X_1,...,X_n\}$. On est donc amené à définir une application notée ϕ_g de $Z\{X_1,...,X_n\}$ dans $Z\{X_1,...,X_n\}$. Remarquons que si $P \in Z\{X_1,...,X_n\}$, on peut l'écrire de façon unique sous la forme $m \cdot P_o$ où $m = (a,M)$ avec $a \in Z$ et $M \in Z^n$ et où P_o est un polynôme primitif de $Z[X_1,...,X_n]$. On dira que P_o est **réduit.**

Si on représente P_o par l'ensemble $\{(a_1,M_1),...,(a_k,M_k)\}$, alors $m \cdot P_o$ est représenté par $\{(a \cdot a_1,M+M_1),...,(a \cdot a_k,M+M_k)\}$. Donc les polyèdres associés à P et à P_o sont égaux à une translation près de vecteur AM où $A = (0,...,0)$.

On peut vérifier facilement que $\Pi(P_o)$ a au moins un point dans chacun des hyperplans de coordonnées d'équations $X_j = 0$ $(1 \leq j \leq n)$.

L'application ϕ_g définie sur $Z\{X_1,...,X_n\}$ doit vérifier les deux conditions suivantes :

(1) $\phi_g(P \cdot Q) = \phi_g(P) \cdot \phi_g(Q)$.

(2) $\phi_g(1) = 1$.

D'autre part on ne veut pas que ϕ_g modifie le nombre des monômes des polynômes, ce qui est le cas des changements de variables $X_i \rightarrow X_i + X_1$ $(2 \leq i \leq n)$ utilisés pour rendre un polynôme unitaire et qui conduisent à une explosion du nombre des monômes. On impose donc la condition supplémentaire suivante :

(3) $\phi_g(P)$ et P ont le même nombre de monômes.

Cette condition entraîne que ϕ_g définit une application ϕ_o sur l'ensemble des monômes à n variables

ϕ_o: $(a,M) \rightarrow (\rho(a),\sigma(M))$ où $\rho: Z \rightarrow Z$ et où $\sigma: Z^n \rightarrow Z^n$.

Pour simplifier ϕ_o, on prend pour ρ l'identité. D'après **(1)** on a

$\phi_g((1,M_1)\cdot(1,M_2)) = \phi_g(1,M_1)\cdot\phi_g(1,M_2)$. Par suite, $\sigma(M_1+M_2) = \sigma(M_1) + \sigma(M_2)$. D'après (2) on a $\sigma(-M) = -\sigma(M)$. Il en résulte que σ est une application linéaire du module Z^n dans lui-même. Notons la matrice de σ sous la forme

$$T = \begin{vmatrix} u_1 & u_2 & \ldots & u_n \\ \ldots & \ldots & \ldots & \ldots \\ v_1 & v_2 & \ldots & v_n \\ \ldots & \ldots & \ldots & \ldots \\ w_1 & w_2 & \ldots & w_n \end{vmatrix}$$

où le déterminant de T est non nul. L'image de $M = (d1,\ldots,dn)$ par σ s'écrit $M' = (d'1,\ldots,d'n)$ où $d'i = v1\ d1 + \ldots + vn\ dn$. L'application ϕ_g s'écrit donc

$$\phi_g: \sum_i a_i X_1^{d1} \ldots X_n^{dn} \rightarrow \sum_i a_i X_1^{d'1} \ldots X_n^{d'n}$$

où $d'1 = u1\ d1 + \ldots + un\ dn$, ..., $d'n = w1\ d1 + \ldots + wn\ dn$. En remplaçant d'1,..., d'n par leurs valeurs, les produits $X_1^{d'1} \ldots X_n^{d'n}$ s'écrivent $X_1^{u1\ d1 + \ldots + un\ dn} \ldots X_n^{w1\ d1 + \ldots + wn\ dn}$, soit encore $(X_1^{u1}\ X_2^{v1} \cdots X_n^{w1})^{d1} \cdots (X_1^{un}\ X_2^{vn} \cdots X_n^{wn})^{dn}$. Par suite, ϕ_g transforme $P(X_1, X_2, \ldots, X_n)$ en $P(X_1^{u1}\ X_2^{v1} \cdots X_n^{w1}, \ldots, X_1^{un}\ X_2^{vn} \cdots X_n^{wn})$. Comme $\phi_g(P)$ est un polynôme généralisé de $Z\{X_1,\ldots,X_n\}$, on doit le multiplier par un monôme μ pour obtenir un polynôme de $Z[X_1,\ldots,X_n]$. L'application de simplification, notée ϕ définie sur $Z[X_1,\ldots,X_n]$ s'écrit donc $\phi = N_\mu \cdot \phi_g$, où $N_\mu(Q) = \mu Q$.

Si ϕ n'est pas surjective, alors ϕ^{-1} n'est définie que sur $\phi(Z[X_1,\ldots,X_n])$. Supposons que dans ces conditions un polynôme P soit simplifié en $\phi(P)$ et que $\phi(P)$ soit factorisé en $\phi(P) = P_1 \cdot P_2$. Alors $\phi^{-1}(P_1)$ et $\phi^{-1}(P_2)$ ne sont pas nécessairement définis (les exposants des variables X_i sont fractionnaires). Si, par contre $\phi^{-1}(P_1) \in Z[X_1,..,X_n]$, alors on a aussi $\phi^{-1}(P_2) \in Z[X_1,..,X_n]$; par suite $\phi^{-1}(P_1)$ et $\phi^{-1}(P_2)$ sont des diviseurs de P.

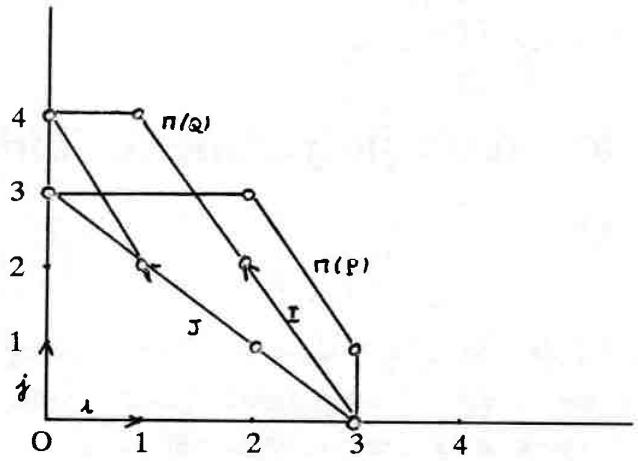
Exemple.

Considérons les deux polynômes suivants dont on veut calculer le PGCD:

$P = X\,Y^2 + X^2\,Y + X^3 + Y^3 + X^2\,Y^2 - X^3\,Y - X^2\,Y^3$ et

$Q = X\,Y^2 + X^3 - X\,Y^4 - Y^4$.

Représentons les polygones associés :



Cette représentation polygonale met en évidence le repère (I, J) dans lequel les polynômes seront simplifiés. La matrice de passage de (I, J) à (i, j) s'écrit

$$T = \begin{vmatrix} 1 & 1 \\ -2 & 1 \end{vmatrix}$$

L'application ϕ_g correspondante transforme $X^{d1}\,Y^{d2}$ en $X^{d1+d2}\,Y^{-2\,d1-d2}$. On obtient donc $\phi_g(P) = X^3Y^{-7}(\,Y^3 + Y^2 + Y + Y^4 + X\,Y - X - X^2\,)$ et

$\phi_g(Q) = X^3Y^{-6}(\,Y^2 + 1 - X^2 - X\,Y^2\,)$.

Les polynômes réduits correspondants s'écrivent

$\phi(P) = Y^4 + Y^3 + Y^2 + Y + X(Y - 1) - X^2$ et

$\phi(Q) = Y^2 + 1 - X\,Y^2 - X^2$.

Ces polynômes sont unitaires en X et la division de $\phi(P)$ par $\phi(Q)$ s'écrit

$$\phi(P) = \phi(Q) + X(Y^2 + Y - 1) + Y^4 + Y^3 + Y - 1.$$

Le reste s'écrit sous la forme $R = (X + Y^2 + 1)(Y^2 + Y - 1)$. Par suite $\Delta = X + Y^2 + 1$ est le PGCD des polynômes $\phi(P)$ et $\phi(Q)$. Calculons $\phi^{-1}(\Delta)$ en utilisant la matrice inverse de T, à savoir

$$T^{-1} = \begin{vmatrix} -1 & -1 \\ 2 & 1 \end{vmatrix}$$

On transforme donc $X^{d1}\ Y^{d2}$ en $X^{-d1-d2}\ Y^{2d1+d2}$; Δ est donc transformé en $X^{-2}(X\ Y^2 + Y^2 + X^2)$. Donc $\phi^{-1}(\Delta)$ est égal à $X\ Y^2 + Y^2 + X^2$; comme $\phi^{-1}(\Delta) \in \mathbf{Z}[X,Y]$, c'est bien le PGCD de P et Q.

2°) Simplification d'un polynôme à l'aide du polyèdre associé.

On a montré au **1°)** qu'une application ϕ définie sur $\mathbf{Z}[X_1,...,X_n]$, correspond à un changement de repère de l'espace affine $\mathbf{Z}^n$. La recherche de telles applications ϕ revient donc à définir des changements de repère de $\mathbf{Z}^n$. Le fait que P soit simplifié apparaît dans la représentation polyèdrale comme un rapprochement du nouveau repère et du polyèdre associé $\Pi(P)$. Le sens d'un rapprochement devra être précisé. Remarquons que les coordonnées des points de $\Pi(P)$ doivent être positives dans le nouveau repère $\mathfrak{R}$.
Examinons quelques exemples de simplications :

a) <u>Simplification de P faisant disparaître des variables.</u>
Si le polyèdre $\Pi(P)$ de $\mathbf{Z}^n$ est de dimension m inférieure à n, alors on peut choisir un nouveau repère $\mathfrak{R}$ tel que les coordonnées $X_{m+1},...,X_n$ des points de $\Pi(P)$ soient nulles. Ce rapprochement de $\mathfrak{R}$ et de $\Pi(P)$ correspond à une simplification de P diminuant le nombre des variables.

b) <u>Simplification de P faisant apparaître un monôme constant.</u>
Le rapprochement considéré ici consiste à prendre un point de $\Pi(P)$ comme origine A du nouveau repère $\mathfrak{R}$. Le monôme de P associé à ce point est transformé en monôme constant. Comme les coordonnées de $\Pi(P)$ doivent rester positives, $\Pi(P)$ doit être contenu dans le premier "quadrant" du repère $\mathfrak{R}$. L'origine du nouveau repère doit donc être l'un des sommets de $\Pi(P)$. Donc tout monôme extrêmal, c'est-à-dire associé à un sommet peut être tranformé en monôme constant par une application ϕ.

c) Transformation de P en un polynôme unitaire.

Le polynôme P est unitaire suivant la variable X_1 s'il contient un monôme m_o de la forme $a X_1^{d1}$ où $a \in Z$ et où les autres monômes de P ont un degré suivant X_1 inférieur à d1. Le point M_o de $\Pi(P)$ associé à m_o doit donc appartenir à l'axe de coordonnée AX_1 du nouveau repère, les premières coordonnées (suivant X_1) des autres points de $\Pi(P)$ étant inférieures à la première coordonnée de M_o.

Il est clair que tout sommet de $\Pi(P)$ peut être un tel point M_o dans un certain repère $\mathfrak{R}$. En effet il suffit de prendre l'axe AX_1 porté par une arête de $\Pi(P)$ passant par M_o et de choisir l'hyperplan H contenant les autres axes de coordonnées tel que $\Pi(P)$ soit contenu dans l'espace compris entre H et l'hyperplan parallèle à H et passant par M_o, ce dernier hyperplan ne devant pas contenir d'autre point de $\Pi(P)$ que M_o. Pour simplifier P, on doit choisir H de façon à ce qu'il contienne une face de dimension maximum; mais il est possible que H ne puisse rien contenir d'autre qu'un sommet de $\Pi(P)$.

C'est le cas par exemple, si $\Pi(P)$ est un cube.

d) Simplification de P diminuant le degré suivant la variable X_i

On veut que le polynôme P soit simplifié au maximum en tant que polynôme en X_i. Une telle simplification correspond à la recherche d'un hyperplan de coordonnée H du nouveau repère qui soit le plus proche posssible de $\Pi(P)$ ou plus précisément de l'ensemble A(P) formé des points de Z^n associés aux monômes de P. De plus les points de A(P) doivent être situés du même côté de H.

On peut exprimer que A(P) est proche de H en définissant une distance d dans Z^n. On choisit pour la suite une distance qui s'exprime simplement, à savoir la distance

$d_1^\circ(M,N) = \sum_{1 \le j \le n} | X_j - Y_j |$ où $M = (X_1,...,X_n)$ et $N = (Y_1,...,Y_n)$.

On peut exprimer la distance entre H et A(P) de deux manières :

$d_1(A(P), H) = \sum_{M \in A(P)} d_1^*(M,H)$ ou $d_\infty(A(P), H) = \max_{M \in A(P)} d_1^*(M,H)$

sachant que $d_1^*(M,H) = \min_{N \in H} d_1^\circ(M,N)$ est la nouvelle coordonnée X_i de M qu'on notera dans la suite $\delta_i(M)$.

L'avantage de la distance d_1 est qu'elle est linéaire suivant les coordonnées des points M, ce qui simplifie la résolution à l'aide de la programmation linéaire.

L'avantage de la distance d_∞ est qu'elle définit des polynômes de degré minimum. La résolution peut encore se faire à l'aide de la programmation linéaire, mais les calculs sont plus longs.

Notons V_i le vecteur de coordonnées $\delta_i(M_1),...,\delta_i(M_p)$. Le problème à résoudre est alors le suivant: trouver un nouvel hyperplan H tel que

$$\begin{cases} \| V_i \| & \text{soit minimum, avec les contraintes} \\ V_i \geq 0\,, & \text{c'est-à-dire } \delta_i(M_j) \geq 0 \text{ pour } 1 \leq j \leq p. \end{cases}$$

Les nombres $\delta_i(M_1),...,\delta_i(M_p)$ sont les mesures algèbriques des distances de $M_1,...,M_p$ à H mesurées le long de l'axe AX_i.

e) <u>Simplification de P diminuant le degré de suivant chacune des variables.</u>

On reprend l'algorithme précédent successivement pour chacune des variables $X_1,...,X_n$. A la $i^{ième}$ étape, on définit l'hyperplan H_i d'équation $X_i = 0$, en résolvant le programme linéaire suivant :

$$\begin{cases} \text{Min } \| V_i \| & \text{(c'est-à-dire que l'objectif est de rendre } V_i \text{ minimum)} \\ V_i \geq 0 & \\ V_i & \text{étant linéairement indépendant de } V_1,...,V_{i-1}. \end{cases}$$

La dernière contrainte oblige les hyperplans de coordonnées, d'équations $X_i = 0$, notés H_i à former un repère de Z^n. Si au cours de la résolution le programme devient irréalisable, cela signifie que la troisième contrainte ne peut pas être satisfaite, donc que $\Pi(P)$ est de dimension inférieure à n. Dans ce cas la variable X_i peut être supprimée.

f) <u>Simplification de P diminuant le nombre de variables apparaissant dans les monômes.</u>

On définit une nouvelle distance entre $M = (X_1,...,X_n)$ et $N = (Y_1,...,Y_n)$, notée $d_o(M,N)$ égale au nombre de coordonnées distinctes de M et N (ce qui pourrait donner un sens à la somme $\sum_{1 \leq j \leq n} | X_j - Y_j |^0$).

On notera donc $\| V \|_o$ le nombre de coordonnées non nulles du vecteur V. On verra au paragraphe suivant que pour tout polyèdre convexe Π, il existe un repère $\mathfrak{R}$ dont la distance d_o à Π est minimum; on dira dans ce cas que le repère $\mathfrak{R}$ est adhérent à Π.

La simplification de P consistant à diminuer le nombre de variables apparaissant dans les monômes revient donc à trouver un repère $\mathfrak{R}$ le plus proche possible de $\Pi(P)$ suivant la distance d_o, c'est-à-dire un repère $\mathfrak{R}$ adhérent à $\Pi(P)$. Le repère $\mathfrak{R}$ sera défini en déterminant successivement les hyperplans de coordonnées de $\mathfrak{R}$ notés $H_1,...,H_n$. A la $i^{ième}$ étape on cherche le $i^{ième}$ hyperplan H_i en résolvant le programme linéaire suivant :

$$\begin{cases} \text{Min } \| V_i \|_o \\ V_i \geq 0,\ C_1 \geq 0,\ ...,\ C_q \geq 0, \end{cases}$$

où $C_1 \geq 0,..., C_q \geq 0$ sont des contraintes supplémentaires pour que les hyperplans H_i forment un repère de $\mathbf{Z}^n$ adhérent à $\Pi(P)$.

3°) Repère adhérent à un polyèdre.

Le dernier programme linéaire proposé ci-dessus, pour diminuer le nombre de variables apparaissant dans les monômes, conduit à des calculs complexes. On propose dans ce paragraphe une autre approche du problème. Pour cela, on introduit quelques définitions sur les polyèdres, qu'on peut trouver notamment dans Grünbaum **[Gr]**.
L'étude théorique comprend un théorème d'existence (et d'unicité); la résolution effective sera étudiée dans la partie approche pratique qui suit ce paragraphe. Cette résolution pratique donne un algorithme détaillé pour le cas d'un polynôme à trois variables.

Soit Π un polyèdre convexe de $\mathbf{Z}^n$ où $n \geq 2$. On notera $S = \{M_o, M_1,..., M_{n-1}\}$ une suite de points de points de Π tels que $\{M_o, M_1,..., M_i\}$ définisse une face de dimension i de Π pour $0 \leq i \leq n-1$. Une telle suite de points sera appelée **une suite caractéristique de Π.**
Rappelons quelques définitions et quelques propriétés sur les faces d'un polyèdre convexe. On a vu dans l'introduction qu'une face est l'intersection du polyèdre avec un hyperplan d'appui. Une face est donc encore un polèdre convexe. On appelle **hyperface** une face de dimension $n-1$. Un sommet et une arête sont des faces de dimension respectives 0 et 1. Enfin le résultat suivant sera utilisé dans la suite :

Lemme.

Soit Π un polyèdre convexe.

I. Si F est une face de Π, alors toute face de F est une face de Π.

II. Si f est une face de dimension n−2 de Π, alors il existe exactement deux hyperfaces de Π qui contiennent f.

Donnons maintenant la définition d'un repère adhérent :

Définition.

Soit S une suite caractéristique de Π. Un repère $\mathfrak{R} = \{B_oX_1,...,B_oX_n\}$ de Z^n est appelé repère adhérent à Π suivant S si :

(i) l'angle polyèdre de $\mathfrak{R}$ contient Π;

(ii) pour tout $i \in [0, n-1]$ et tout $k \in [0,i]$, l'hyperface $\mathfrak{R}_{ik}$ du repère $\mathfrak{R}_{i+1} = \{B_oX_1,...,B_oX_{i+1}\}$ de Z^{i+1} contient une face Π_{ik} de dimension i de Π; de plus, la face $\mathfrak{R}_{io}$ contient les i+1 premiers points de S, à savoir $M_o, M_1,...,M_i$.

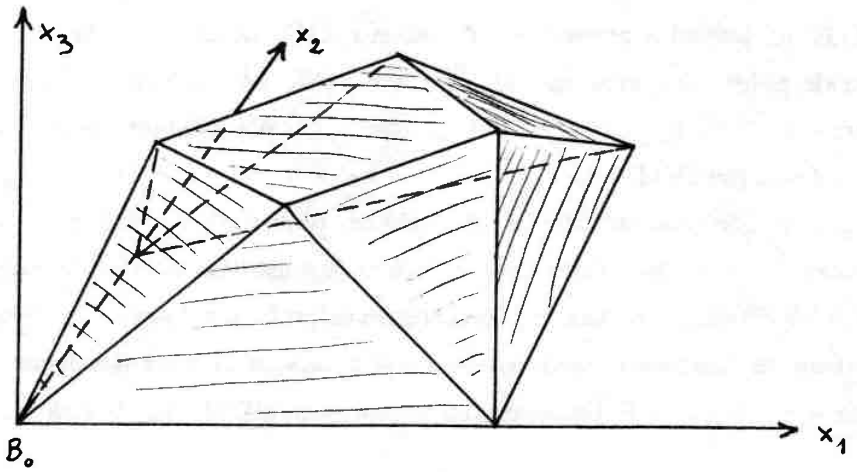
Etudions la propriété (ii) pour les premières valeurs de i. Cette propriété signifie que:

pour i = 0, l'origine B_o de $\mathfrak{R}$ est le point M_o;

pour i = 1, les axes B_oX_1 et B_oX_2 contiennent respectivement l'arête M_oM_1 et une autre arête de Π;

pour i = 2, d'une part $\mathfrak{R}_{20}$ contient la face définie par les points M_o, M_1 et M_2, et d'autre part $\mathfrak{R}_{12}$ et $\mathfrak{R}_{22}$ contiennent deux autres faces de Π de dimension 2.

Exemple. Considérons le polyèdre Π ci−dessous.



On veut construire un repère $\mathfrak{R}$ dont les faces, (c'est-à-dire ici les axes et les plans de coordonnées), contiennent le maximum de faces de Π. Pour cela, on choisit comme origine $B°$ un sommet de Π, puis on prend les axes de coordonnées $B°x_1$ et $B°x_2$ le long de deux arêtes d'une même face de Π, de façon que le plan de coordonnée $B°x_1x_2$ contienne une face de Π. On peut en plus imposer aux deux autres plans de coordonnées $B°x_1x_3$ et $B°x_2x_3$ de contenir deux autres faces de Π. Mais il est impossible d'imposer que le troisième axe $B°x_3$ soit porté par une arête de Π. Le repère défini sur le croquis ci-contre vérifie la définition d'un repère adhérent à Π. Il est dans un certain sens, le plus proche possible de Π.

Théorème.

Soit Π un polyèdre convexe de $\mathbf{Z}^n (n \geq 2)$ de dimension n et $S = \{M_o, M_1, ..., M_{n-1}\}$ une suite caractèristique de Π. Alors, il existe un repère $\mathfrak{R}$ unique de $\mathbf{Z}^n$ qui est adhérent à Π suivant S.

Démonstration.

Raisonnons par récurrence sur n.

Supposons $n = 2$ et notons $S = \{M_o, M_1\}$. D'après le (ii) de la définition ci-dessus, les axes $B°x_1$ et $B°x_2$ de $\mathfrak{R}$ doivent contenir deux arêtes de Π et la nouvelle origine $B°$ doit être confondue avec M_o. Les axes $B°x_1$ et $B°x_2$ sont donc définis de façon unique.

Supposons $n > 2$. Notons Π^* l'hyperface de Π définie par les points $\{M_o, M_1, .., M_{n-1}\}$ de S. La suite $S^* = \{M_o, M_1, .., M_{n-2}\}$ est une suite caractèristique de Π^*. D'après l'hypothèse de récurrence, il existe un repère unique $\mathfrak{R}^*$ de $\mathbf{Z}^{n-1}$ adhérent à Π^* suivant S^*. Notons $\mathfrak{R}^*_j$ les hyperfaces de $\mathfrak{R}^*$ pour $0 \leq j \leq n-1$. D'après le (ii) de la définition ci-dessus, $\mathfrak{R}^*_j$ contient une face Π^*_j de dimension $n-2$ de Π^*, donc de Π. D'après le **(II)** du Lemme ci-dessus, il existe deux hyperfaces de Π contenant Π^*_j; l'une est Π^*, notons Π_j l'autre face. Au total, chacune des $n-2$ hyperfaces $\mathfrak{R}^*_j$ de $\mathfrak{R}^*$ contient une hyperface Π_j de Π. Or d'après (ii), l'hyperface de $\mathfrak{R}$ autre que $\mathfrak{R}^*$, contenant $\mathfrak{R}^*_j$, doit contenir une telle hyperface Π_j de Π. Les hyperfaces de $\mathfrak{R}$ sont donc définies par les hyperfaces $\Pi^*, \Pi_1, ..., \Pi_{n-1}$ de Π.

Vérifions que ces n hyperfaces $\Pi^*, \Pi_1, ..., \Pi_{n-1}$ sont linéairement indépendantes, ce qui prouvera que $\mathfrak{R}$ est un repère de $\mathbf{Z}^n$. Soit $\mathfrak{R}°$ un repère de $\mathbf{Z}^n$ obtenu en

complètant $\mathfrak{R}^*$ avec un vecteur supplémentaire. Les équations des hyperfaces Π_j dans le repère $\mathfrak{R}^\circ$ s'écrivent $a_{1\,j}\,x_1 + \ldots + a_{n\,j}\,x_n = 0$. Les intersections $\mathfrak{R}^*_j$ de ces hyperfaces avec $\mathfrak{R}^*$ s'obtiennent en faisant $x_n = 0$. Or l'équation de $\mathfrak{R}^*_j$ dans $\mathfrak{R}^*$ est $x_j = 0$. L'équation de Π_j dans $\mathfrak{R}^\circ$ est donc $x_j + a_{n\,j}\,x_n = 0$. Ces hyperplans sont donc bien linéairement indépendants. L'axe $B^\circ x_n$ ajouté à $\mathfrak{R}^*$ pour obtenir $\mathfrak{R}$ est l'intersection $\Pi_1 \cap \ldots \cap \Pi_{n-1}$. On convient de diriger $B^\circ x_n$ vers le demi-espace contenant Π. Ainsi les coordonnées x_n des points de Π seront positives.

Il reste à vérifier que $\mathfrak{R}$ est adhérent à Π suivant S. Considérons une hyperface $\mathfrak{R}_j$ de $\mathfrak{R}$; elle contient une hyperface Π_j de Π, donc Π est situé d'un même côté de $\mathfrak{R}_j$: les coordonnées x_j des points de Π sont toutes positives ou toutes négatives. Or les points de Π^* vérifient $x_j \geq 0$ par hypothèse de récurrence. Donc $\mathfrak{R}$ vérifie la propriété (i) de la définition du repère adhérent. Par hypothèse de récurrence, la propriété (ii) est vérifiée pour $0 \leq i \leq n-2$, car $\mathfrak{R}_{i+1} = \{B^\circ x_1, \ldots, B^\circ x_{i+1}\}$ est contenu dans $\mathfrak{R}^*$. Pour $i = n-1$, les hyperfaces $\mathfrak{R}_{n-1\,j}$ de $\mathfrak{R}$ contiennent les hyperfaces Π_j de Π par construction de $\mathfrak{R}$. Donc la propriété (ii) est vérifiée.

II. APPROCHE PRATIQUE.

1°) Présentation générale des algorithmes.

Considérons un polynôme P de $Z[X_1,...,X_n]$ noté sous la forme $\sum_{1 \le i \le p} a_i X_1^{di1} ... X_n^{din}$.
On appelle matrice associée à P, la matrice suivante :

$$M(P) = \begin{pmatrix} d_{11} & ..... & d_{1p} \\ ... & ..... & ... \\ d_{n1} & ..... & d_{np} \end{pmatrix} \qquad \text{où} \quad d_{ij} \in N.$$

D'après le paragraphe sur la définition de l'application ϕ, page 125, une simplification correspond à un changement de variable qui est défini par une matrice de passage notée

$$T = \begin{pmatrix} x_{11} & ..... & x_{1n} \\ ... & ..... & ... \\ x_{n1} & ..... & x_{nn} \end{pmatrix} \qquad \text{où} \quad x_{ij} \in Z.$$

Rappelons la définition d'un polynôme réduit, vue page 124. Le polynôme P qui qui appartient a priori à $Z\{X_1,...,X_n\}$, est dit réduit si c'est un polynôme primitif de $Z[X_1,...,Xn]$. L'opération de réduction notée N_μ (définie page 125) consiste à multiplier P par un monôme μ de $Z\{X_1,...,X_n\}$.
Si P est réduit, alors chaque ligne de M(P) a au moins un coefficient nul. Dans ce cas, on dira que M(P) est **réduite**. On note ν l'application réduisant une matrice M(P); on a donc la relation $\nu(M(P)) = M(N_\mu(P))$. Le polynôme simplifié (suivant les différents sens donnés pages 127, 128 et 129) aura donc une matrice de la forme $\nu \cdot T \cdot M(P)$. En fait, l'opération de réduction peut se faire simplement, en ajoutant à M(P) une ligne de 1 notée ℓ_o. On obtient ainsi une nouvelle matrice associée à P de la forme

$$M°(P) = \begin{pmatrix} d_{o1} & ..... & d_{on} \\ d_{11} & ..... & d_{1p} \\ ... & ..... & ... \\ x_{n1} & ..... & x_{nn} \end{pmatrix}$$

où $d_{ij} \in N$ et où
$d_{o1} = ... = d_{op} = 1.$

De plus, on ajoute une colonne à la matrice T qui s'écrit sous la forme :

$$T^\circ = \begin{pmatrix} x_{10} & x_{11} & \ldots\ldots & x_{1n} \\ \ldots & \ldots & \ldots\ldots & \ldots \\ x_{n0} & x_{n1} & \ldots\ldots & x_{nn} \end{pmatrix} \qquad \text{où} \quad x_{ij} \in Z.$$

La simplification de P consiste donc à chercher une matrice T° telle que

$$\begin{cases} T^\circ \cdot M^\circ(P) & \text{soit petit} \\ T^\circ \cdot M^\circ(P) & \text{soit positif.} \end{cases}$$

En fait, la matrice T° sera définie ligne par ligne. On notera $\ell_o, \ell_1, \ldots, \ell_n$ les lignes de $M^\circ(P)$ et $x_{io}, .., x_{in}$ les coefficients de la ligne de T° qu'on veut définir. La simplification cherchée revient à transformer $\ell_i = (d_{i1}, \ldots, d_{ip})$ en $\ell_i' = (\delta_{i1}, \ldots, \delta_{ip})$ où i prendra successivement les valeurs 1, ..., n et où $\delta_{ij} = \sum_{o \le k \le n} d_{kj}\ x_{ik}$ avec comme objectif que $\| \ell_i' \|$ soit minimum, et avec comme contrainte $\ell_i' \ge 0$ (c'est-à-dire $\delta_{ij} \ge 0$ pour $1 \le j \le p$). On doit ajouter d'autres contraintes, en particulier pour exprimer que les nouvelles lignes ℓ_i' de M(P) soient linéairement indépendantes (pour que ϕ soit injective). Notons ces contraintes supplémentaires sous la forme

$$C_1 \ge 0, \ldots, C_q \ge 0.$$

La simplification du polynôme donné P s'exprime donc sous la forme d'un programme linéaire, qui s'écrit :

$$\text{Min} \ \| \ell_i' \|$$

$$\begin{cases} \delta_{i1} = \sum_{o \le j \le n} d_{i1}\ x_{ij} \ge 0 \\ \ldots\ldots\ldots\ldots \\ \delta_{ip} = \sum_{o \le j \le n} d_{ip}\ x_{ij} \ge 0 \\ C_1 \ge 0, \ldots, C_q \ge 0. \end{cases} \qquad \text{où} \quad x_{ij} \in Z$$

Si ce programme linéaire admet comme solution $\ell_i' = 0$, cela signifie qu'on peut supprimer la variable X_i dans le polynôme P.

Pour des raisons techniques relatives à la résolution du programme linéaire, on suppose que les variables x_{ij} sont minorées par un entier connu $-M$. En conséquence, on utilise pour les calculs les variables positives $x_{ij}' = x_{ij} + M$. On utilise, dans la suite la méthode de résolution exposée dans Simonnard **[Si]**. On adopte en particulier les notations de

cette méthode qui nous oblige à rechercher un maximum et à changer le sens des inégalités des contraintes. Avec ces nouvelles conventions le programme linéaire s'écrit sous la forme suivante

$$\text{Max} \ \| -\ell_i' \|$$

$$\begin{cases} -x_{ij}' \le 0 & \text{pour} \quad 0 \le j \le n \\ \delta_{i1} = \sum_{o \le j \le n} -d_{ik}\, x_{ij}' \le -M \sum_{o \le j \le n} d_{jk} & \text{pour} \quad 1 \le k \le p \\ -C_j \le 0 & \text{pour} \quad 1 \le j \le q \\ x_{ij}' \quad \text{entier.} \end{cases}$$

Pour résoudre une tel programme linéaire, on l'écrit sous forme d'un tableau, comme ci-dessous :

$\| -\ell_i' \|$	a_o	a_1	a_2	.....	a_n
x_{io}'	0	−1	0	.....	0
..	..	..	..	.....	..
x_{in}'	0	0	0	.....	−1
δ_{i1}	$-M \cdot S_1$	$-d_{11}$	$-d_{21}$	.....	$-d_{n1}$
...	...	..	..	.....	..
δ_{ip}	$-M \cdot S_p$	$-d_{1p}$	$-d_{2p}$	.....	$-d_{np}$
C_1	C_{10}	C_{11}	C_{12}	.....	C_{1n}
..	...	..	..	.....	..
C_q	C_{q0}	C_{q1}	C_{q2}	.....	C_{qn}

On verra que pour chacune des deux normes $\| \ell_i' \|_1$ et $\| \ell_i' \|_\infty$, les coefficients $a_1, ..., a_n$ sont positifs. Ce programme linéaire est donc dual−réalisable. La résolution telle qu'elle est définie dans Simonnard [Si], consiste à choisir une ligne d'indice i_o dont le premier coefficient est < 0. On déduit de cette ligne une nouvelle ligne obtenue en divisant chaque coefficient de la ligne i_o par un certain nombre λ, puis en prenant la partie entière. Cette nouvelle ligne a un coefficient égal à -1 qu'on choisit comme

pivot. Amprès le pivotage, les coefficients du tableau restent donc entiers. On sait qu'après un nombre fini de pivotages, on obtient un programme réalisable qui est optimal.

2°) Simplification suivant les degrés.

On procède par étapes successives en remplaçant progressivement les lignes $\ell_o,...,\ell_i,...$ de $M°(P)$ par $\ell_o',..., \ell_i',...$ où $\ell_o = \ell_o' = (1,...,1)$. A la $i^{ème}$ étape les lignes de $M°(P)$ sont égales à $\ell_o', ..., \ell_{i-1}', \ell_i,..., \ell_n$ et on calcule

$$\ell_i' = x_{io}\,\ell_o + ... + x_{i\,i-1}\,\ell_{i-1} + x_{i\,i}\,\ell_i + ... + x_{i\,n}\,\ell_n$$

qui sera substitué à ℓ_i.

On suppose qu'au départ, les lignes de $M°(P)$, à savoir $\ell_o,..., \ell_n$ sont linéairement indépendantes (si ce n'était pas le cas, on pourrait supprimer une variable de P). Pour que les nouvelles lignes de $M°(P)$ restent linéairement indépendantes, on impose au coefficient $x_{i\,i}$ de ℓ_i d'être non nul, c'est-à-dire $x_{i\,i} \geq 1$ ou $x_{i\,i}' \leq -1$. Le programme linéaire qu'on doit résoudre, s'écrit donc à la $i^{ème}$ étape :

$$\text{Min } \| \ell_i' \|$$

$$\begin{cases} \delta_{ik} = \displaystyle\sum_{o\leq j\leq n} d_{ik}\, x_{ij} \geq 0 & \text{pour } 1 \leq k \leq p \\ x_{i\,i} \geq 1 \quad \text{ou} \quad x_{i\,i} \leq -1, \quad x_{i\,i} \in Z & \text{pour } 0 \leq j \leq n. \end{cases}$$

Les calculs de la résoultion de ce programme linéaire vont dépendre de la norme $\| \ell_i' \|$ choisie, $\| \ell_i' \|_1$ ou $\| \ell_i' \|_\infty$. Examinons ces calculs dans chacun des deux cas.

A. Résolution du programme linéaire avec la norme $\| \ell_i' \|_1 = \delta_{i\,1} + ... + \delta_{i\,p}$.

La fonction économique du programme linéaire s'écrit :

$$Z = \| -\ell_i' \| = -\sum_{1\leq k\leq p} \delta_{i\,k} = -\sum_{1\leq k\leq p} d_{j\,k}\, x_{i\,j}$$

$$= \sum_{o\leq j\leq n} \Big(\sum_{1\leq k\leq p} d_{j\,k}\Big)\, x_{i\,j}' + M \sum_{o\leq j\leq n} \sum_{1\leq k\leq p} d_{j\,k}\,.$$

Les contraintes supplémentaires, xi $i \geq 1$ ou $x_{i\,i} \leq -1$ s'écrivent $-x_{i\,i}' \leq -M-1$ ou $x_{i\,i}' \leq M-1$. La résolution de ce programme linéaire sera composée de deux parties; au cours de la première partie on ne choisit aucune des deux dernières lignes comme ligne i_o pour engendrer une nouvelle contrainte. On aboutit ainsi au tableau suivant :

$\| -\ell_i' \|$	0	a_{01}	a_{02}	.....	a_{0n}
x_{io}'	M	a_{11}	a_{12}	.....	a_{1n}
..	..	..	..	.....	..
x_{in}'	M	a_{n1}	a_{n2}	.....	a_{nn}
δ_{i1}	0	$a_{n+1\ 1}$	$a_{n+1\ 1}$	.....	$a_{n+1\ n}$
...	...	..	..	.....	..
δ_{ip}	0	$a_{n+p\ 1}$	$a_{n+p\ 2}$	.....	$a_{n+p\ n}$
$x_{i\ i} \geq 1$	−1	$a_{n+p+1\ 1}$	$a_{n+p+1\ 2}$	.....	$a_{n+p+1\ n}$
$x_{i\ i} \leq -1$	−1	$a_{n+p+2\ 1}$	$a_{n+p+2\ 2}$	.....	$a_{n+p+2\ n}$

La deuxième partie de la résolution consiste à choisir la première des deux dernières lignes et à exécuter des pivotages pour rendre positifs tous les termes de la première colonne (en excluant du tableau, la dernière ligne). Ensuite, on recommence les calculs en échangeant les rôles des deux dernières lignes, et à la fin on choisit la meilleure des deux solutions obtenues.

Exemple.

Considérons le polynôme $P = Y^2 Z^2 + X Z + X Y^2 + X^2 Y^3$.

La matrice associée s'écrit :

$$M°(P) = \begin{array}{c} u_i \\ x_i \\ y_i \\ z_i \end{array} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 2 \\ 2 & 0 & 2 & 3 \\ 2 & 1 & 1 & 0 \end{pmatrix}$$

Admettons que u_i, x_i, y_i et z_i soient minorés par -10 et posons $u_i' = u_i + 10$, ... , $z_i' = z_i + 10$. La résolution du programme linéaire rendant minimum la somme des degrés $\| \ell_x' \|_1$ donne successivement les tableaux suivants :

$\| -\ell_X' \|$	190	4	4	7	4
u_1'	0	−1	0	0	0
x_1'	0	0	−1	0	0
y_1'	0	0	0	−1	0
z_1'	0	0	0	0	−1
δ_{x1}	−50	−1	0	−2	−2
δ_{x2}	−30	−1	−1	0	−1
δ_{x3}	−50	−1	−1	−2	−1
δ_{x4}	−60	−1	−2	−3	0
$x_1 \geq 1$	−11	−1	0	0	0
$x_1 \leq -1$	9	1	0	0	0

70	4	0	7	0
30	−1	1	0	1
0	0	−1	0	0
0	0	0	−1	0
0	0	0	0	−1
−20	−1	1	0	1
0	−1	0	0	0
−20	−1	0	−2	0
−30	−1	−1	−3	1
19	−1	1	0	1
−21	1	−1	0	−1
−10	−1	0	−1	0

$\| -\ell_X' \|$	30	4	0	3	0
u_1'	40	−1	1	1	1
x_1'	0	0	−1	0	0
y_1'	0	0	0	−1	0
z_1'	0	0	0	0	−1
δ_{x1}	−10	−1	1	−1	−1
δ_{x2}	10	−1	0	1	0
δ_{x3}	−10	−1	0	−1	0
δ_{x4}	−20	−1	−1	−2	1
$x_1 \geq 1$	29	−1	1	1	1
$x_1 \leq -1$	−31	1	−1	−1	−1

0	1	0	3	0
30	−2	1	1	1
0	0	−1	0	0
10	1	0	−1	0
0	0	0	0	−1
0	0	1	−1	1
0	−2	0	1	0
0	0	0	−1	0
0	1	−1	−2	1
19	−2	1	1	1
−21	21	−1	−1	−1

On obtient donc $\ell_X = 0$ pour $u_1 = 20$, $x_1 = -10$, $y_1 = 0$ et $z_1 = -10$. On peut donc supprimer la variable X ou la variable Z dans P. On obtient alors le polynôme $P^\circ = Y^2 + X + X Y^2 + X^2 Y^3$.

Reprenons le calcul de simplification avec ce polynôme P°, dont la matrice associée s'écrit :

$$M(P^\circ) = \begin{matrix} u_4 \\ x_4 \\ y_4 \end{matrix} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 2 \\ 2 & 0 & 2 & 3 \end{pmatrix}$$

La recherche de simplification de la ligne des x donne les tableaux successifs suivants

$\| -\ell_x' \|$	150	4	4	7
u_1'	0	−1	0	0
x_1'	0	0	−1	0
y_1'	0	0	0	−1
δ_{x1}	−30	−1	0	−2
δ_{x2}	−20	−1	−1	0
δ_{x3}	−40	−1	−1	−2
δ_{x4}	−60	−1	−2	−3
$x_1 \geq 1$	−11	0	−1	0
$x_1 \leq -1$	9	0	1	0

70	4	0	7
20	−1	1	0
0	0	−1	0
0	0	0	−1
−10	−1	1	−2
0	−1	0	0
−20	−1	0	−2
−40	−1	−1	−3
−11	0	−1	0
9	0	1	0
−10	−1	0	−1

$\| -\ell_x' \|$	30	4	0	3
u_1'	30	−1	0	0
x_1'	0	0	−1	0
y_1'	0	0	0	−1
δ_{x1}	0	−1	1	−1
δ_{x2}	10	−1	0	1
δ_{x3}	−10	−1	0	−1
δ_{x4}	−30	−1	−1	−2
$x_1 \geq 1$	−11	0	−1	0
$x_1 \leq -1$	9	0	1	0

0	1	0	3
20	−2	1	1
0	0	−1	0
0	0	0	−1
10	0	1	−1
0	−2	0	1
0	0	0	−1
−10	1	−1	−2
−11	0	−1	0
9	0	1	0

$\|-\ell_x'\|$	0	1	0	3
u_1'	10	−1	1	−1
x_1'	10	−1	−1	2
y_1'	10	1	0	−1
δ_{x1}	0	1	1	−3
δ_{x2}	0	−2	0	1
δ_{x3}	0	0	0	−1
δ_{x4}	0	0	−1	0
$x_1 \geq 1$	−1	−1	**−1**	2
$x_1 \leq -1$	−1	1	1	−2

0	1	0	3
9	−2	1	1
11	0	−1	0
10	1	0	−1
−1	0	1	**−1**
0	−2	0	1
0	0	0	−1
1	1	−1	−2
0	0	−1	0

$\|-\ell_x'\|$	−3	1	3	3
u_1'	8	−2	2	1
x_1'	11	0	−1	0
y_1'	11	1	−1	−1
δ_{x1}	0	0	0	−1
δ_{x2}	−1	**−2**	1	1
δ_{x3}	1	0	−1	−1
δ_{x4}	3	1	−3	−2
$x_1 \geq 1$	0	0	−1	0
	−1	**−1**	0	0

−4	1	3	3
10	−2	2	1
11	0	−1	0
10	1	−1	−1
0	0	0	−1
1	−2	1	1
1	0	−1	−1
2	1	−3	−2
0	0	−1	0

Le résultat $x_1 = 1$, $u_1 = 0$ et $y_1 = 0$ signifie qu'on doit conserver la ligne des x. Ensuite, on reprend le tableau de la fin de la première partie en choisissant cette fois la dernière ligne; on ajoute la contrainte déduite de la dernière ligne :

−1	0	0	**−1**

Les pivotages donnent les tableaux suivants :

$\|-\ell_X'\|$	−3	1	0	3
u_1'	11	−1	1	−1
x_1'	8	−1	−1	2
y_1'	11	1	0	−1
δ_{X1}	3	1	1	−3
δ_{X2}	−1	−2	0	1
δ_{X3}	1	0	0	−1
δ_{X4}	0	0	−1	0
$x_1 \geq 1$	0	0	0	−1
	−1	−1	0	0

−4	1	0	3
12	−1	1	−1
9	−1	−1	2
10	1	0	−1
2	1	1	−3
1	−2	0	1
1	0	0	−1
0	0	−1	0
0	−1	0	0

La conclusion de cette première étape est qu'on peut remplacer ℓ_X par $\ell_X' = 2\,\ell_o - \ell_X$, mais la somme des degrés n'est pas améliorée. On conserve donc le tableau de départ M(P°) pour la deuxième étape de la simplification. La première partie donne les mêmes tableaux qu'à la première étape, à part les deux dernières lignes qui sont modifiées. A la fin de la première partie, ces deux lignes s'écrivent :

−1	1	0	−1
−1	−1	0	1

En choisissant la première de ces deux lignes on obtient les tableaux suivants :

$\|-\ell_Y'\|$	−3	4	0	3
u_2'	11	−2	1	−1
x_2'	8	1	−1	2
y_2'	11	0	0	−1
δ_{Y1}	3	−2	1	−3
δ_{Y2}	−1	−1	0	1
δ_{Y3}	1	−1	0	−1
δ_{Y4}	0	0	−1	0
$y_2 \geq 1$	0	0	0	−1

−4	4	0	7
13	−2	1	−3
7	1	−1	3
11	0	0	−1
5	−2	1	−5
0	−1	0	0
2	−1	0	−2
0	0	−1	0
0	0	0	−1

La conclusion est que la ligne des y de M(P°) peut être remplacée par

| 5 0 2 0 |

mais la somme des degrés n'est pas améliorée. On reprend donc les calculs à la fin de la première partie en choisissant cette fois la dernière ligne. On obtient :

$\| -\ell_Y' \|$	−1	1	0	1	−5	1	0	4
u_2'	11	−1	1	−2	13	−1	1	−2
x_2'	11	−1		−11	10	−1	−1	1
y_2'	9	1	0	0	9	1	0	0
δ_{Y1}	−1	1	1	−2	1	1	1	−2
δ_{Y2}	2	−2	0	−1	3	−2	0	−1
δ_{Y3}	0	0	0	−1	1	0	0	1
δ_{Y4}	0	0	−1	0	0	0	−1	0
$y_2 \geq 1$	0	−1	0	0	0	−1	0	0
	−1	0	0	−1				

La conclusion de la deuxième étape est qu'on peut simplifier la deuxième ligne de M(P°); la nouvelle matrice de P s'écrit :

$$M(\phi(P)) = \begin{matrix} \ell_X' \\ \ell_Y' \end{matrix} \begin{pmatrix} 0 & 1 & 1 & 2 \\ 1 & 3 & 1 & 0 \end{pmatrix} \qquad \text{avec} \quad T^\circ = \begin{pmatrix} 0 & 1 & 0 \\ 3 & 0 & -1 \end{pmatrix}$$

Le polynôme simplifié correspondant s'écrit : $\phi(P) = Y + X Y^3 + X Y + X^2$.

B. Simplification avec la norme $\| \ell_i \|_\infty = \text{Max}\ (\delta_{i1}, ..., \delta_{ip})$.

La fonction économique s'écrit $Z = -\text{Max}(\delta_{i1}, ..., \delta_{ip}) = -m$. On en déduit les contraintes :

(I) $m \geq \delta_{i2}, ..., m \geq \delta_{ip}$

où m est considéré comme une nouvelle variable. Supposons qu'on connaisse un majorant D des degrés. On en déduit les contraintes :

(II) $m \leq \delta_{i1} + D(1 - e_1), ..., m \leq \delta_{ip} + D(1 - e_p)$

où $e_1, ..., e_p$ sont de nouvelles variables entières ≥ 0. Il résulte de (I) et (II) que $e_1 \leq 1, ..., e_p \leq 1$; d'autre part $\delta_{i\,1} \geq 0, ..., \delta_{i\,p} \geq 0$. On ajoute enfin la contrainte :

(III) $e_1 + ... + e_p \geq 1$

pour que m soit égal à l'un au moins des degrés $\delta_{i\,1}, .., \delta_{i\,p}$. Il est clair que (I), (II) et (III) impliquent que m est bien égal à $\text{Max}(\delta_{i\,1}, .., \delta_{i\,p})$. Le programme linéaire s'écrit donc :

$$\text{Max } Z = -m$$
$$\begin{cases} 0 \leq \delta_{i\,1} \leq m \leq \delta_{i\,1} + D(1 - e_1) \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ 0 \leq \delta_{i\,p} \leq m \leq \delta_{i\,p} + D(1 - e_p) \\ e_1 + ... + e_p \geq 1 \end{cases}$$

La simplification se fait en deux parties; dans la première partie on résoud le programme linéaire défini au paragraphe précédent A. avec la norme $\| \ell_i \|_1$. On obtient ainsi la valeur minimum de $\| \ell_i \|_1$. Dans la deuxième partie, on complète le tableau optimal de ce programme linéaire en ajoutant les colonnes correspondant aux nouvelles variables $e_1, .., e_p$, m et en ajoutant les lignes correspondant aux nouvelles contraintes (I), (II) et (III); puis on prend une nouvelle fonction économique, à savoir $Z = -m$.
En résolvant ce programme linéaire avec le polynôme P traité dans l'exemple précédent, on trouve le même polynôme simplifié $P^\circ = Y + XY^3 + XY + X^2$, qu'en utilisant la norme $\| \ell_i \|_1$. Cela signifie qu'il n'est pas possible de trouver un polynôme simplifié de P de degré inférieur à 2.

3°) Simplification faisant apparaître un monôme constant ou rendant le polynôme unitaire.

Remarquons qu'on peut passer facilement d'un polynôme ayant un monôme constant à un polynôme unitaire et vice–versa.
En effet, si P admet le monôme constant a, alors on peut rendre le polynôme homogène, puis supprimer l'une quelconque des variables. Alors le nouveau polynôme est unitaire.

Si le polynôme donné est unitaire et si son monôme de plus grand degré s'écrit $a X^d$, alors on rend le polynôme homogène, puis on supprime la variable X. Le polynôme obtenu admet le monôme constant a.

Dans chacun des deux cas, on peut vérifier facilement que l'application associant au polynôme son transformé est bien un homomorphisme ϕ.

Dans la suite, on ne traite donc que l'une des deux simplifications, à savoir celle qui fait apparaître un monôme constant.

On utilise la simplification suivant les degrés des variables et à partir de la deuxième étape (c'est-à-dire pour la résolution du deuxième programme linéaire), on ajoute les contraintes correspondant à la condition suivante

(C) { L'un au moins des degrés nul au cours de toutes les étapes précédentes est encore nul.

Il est clair, qu'après la dernière étape, l'un des monômes aura tous ses degrés nuls: ce sera un monôme constant.

Considérons la $i^{ième}$ étape de cette simplification. Notons $\delta_{i\,j1}, \ldots, \delta_{i\,jq}$ les degrés nuls au cours des étapes précédentes et e_{jk} $(1 \le k \le q)$ une nouvelle variable qui vaut 1 si $\delta_{i\,jk}$ reste encore nul et qui vaut 0 sinon. Soit D un majorant des degrés $\delta_{i\,j}$. Les variables e_{jk} doivent vérifier les contraintes suivantes :

(1) $1 - e_{jk} \le \delta_{i\,jk} \le D\,(1 - e_{jk})$ où $1 \le k \le q$

(2) $e_{j1} + \ldots + e_{jq} \ge 1$

(3) e_{jk} entier positif pour $1 \le k \le q$.

Remarquons que (1) et (3) impliquent que $e_{jk} = 0$ ou 1, donc que (1), (2) et (3) équivalent bien à (C). On obtient donc le programme linéaire suivant :

$$\text{Min} \quad \| \ell_i \|$$

$$\begin{cases} \ell_i \ge 0 \\ x_i \ge 1 \quad \text{ou} \quad x_i \le -1 \\ (1),\ (2),\ (3). \end{cases}$$

Exemple.

On considère le polynôme $P = Y^2 + X + X Y^2 + X^2 Y^3$ des exemples précédents, qui a été simplifié suivant les degrés.

Le programme linéaire faisant apparaître un monôme constant s'écrit :

$$\text{Min} \quad \| \ell_Y \|$$
$$\begin{cases} \delta_{y1} \geq 0, \ \delta_{y2} \geq 0, \ \delta_{y3} \geq 0, \ \delta_{y4} \geq 0 \\ 1 - e \leq \delta_{y1} \leq 10\,(1 - e) \\ e \geq 1 \\ y \geq 1 \quad \text{ou} \quad y \leq -1 \end{cases}$$

On obtient $\ell_y = (0, 3, 1, 1)$, par suite le polynôme simplifié correspondant s'écrit $P^1 = 1 + X Y^3 + X Y + X^2 Y$.

On obtient facilement le polynôme unitaire associé qui s'écrit :

$P^2 = T^4 + X + X T^2 + X^2 T$.

5°) Simplification obtenue à l'aide d'un repère adhérent.

Cette simplification est basée sur la recherche d'un repère adhérent et elle n'utilise pas la programmation linéaire. Les calculs se feront comme ci-dessus sur le tableau T des coordonnées des points M_i qui sont les images des monômes du polynôme donné.

Les calculs consisteront à effectuer un changement de repère pour que le nouveau repère soit adhérent à Π suivant une suite caractèristique qu'on devra déterminer. Ce changement de repère sera obtenu en effectuant une suite de transformations simples qui consistent remplacer une ligne de T par une combinaison linéaire des autres lignes. Une telle transformation sera notée $\ell_i \rightarrow a_1 \ell_1 + \ldots + a_i \ell_i + \ldots + a_n \ell_n$, où $a_1, .., a_n \in Z$ et où $a_i \neq 0$.

Il n'est pas nécessaire d'imposer $a_i = \pm 1$ car comme on l'a remarqué dans l'Introduction, l'application ϕ n'a pas besoin d'être surjective, mais uniquement injective. Il suffit donc que le déterminant de la matrice du changement de variable soit non nul.

Le tableau final donne les coordonnées des points $M_o, \ldots, M_p$ dans le nouveau repère. Supposons pour simplifier que la suite caractèristique soit formée de M_o (nouvelle origine), du point M_1 (situé sur le premier axe $B°x_1$), du point M_2 (situé sur le deuxième axe), puis des points M_3' et M_3'' (situés dans les plans $B°x_1x_2$ et $B°x_2x_3$), ... et enfin des points $M_n' M_n'', .., M_n^{(n-1)}$ (situés dans les hyperfaces contenant le dernier

axe $B°x_n$).

Il en résulte en particulier que dans le nouveau repère, les coordonnées de M_o sont nulles, que M_1 et M_2 n'ont qu'une coordonnée non nulle. Les propriétés caractèrisant le nouveau repère adhérent apparaissent donc sur le nouveau tableau $T°$ qui donnent les coordonnées des points $M_o, M_1, .., M_p$ dans le nouveau repère.

On a représenté ci-dessous ce tableau $T°$ avec les conventions suivantes :

la valeur 0 représente une coordonnée nulle, la valeur 1 une coordonnée non nulle et le symbole * représente une coordonnée qui est soit nulle, soit non nulle.

Tableau $T°$ dans le nouveau repère adhérent à Π.

	M_o	M_1	M_2	M_3'	M_3''	M_4'	M_4''	M_4'''		M_n'	M_n''		$M_n^{(n-1)}$
x_1	0	1	0	0	*	0	*	*	..	0	*	..	*
x_2	0	0	1	*	0	*	0	*	..	*	0	..	*
x_3	0	0	0	1	1	*	*	0	..	*	*	..	*
x_4	0	0	0	0	0	1	1	1	..	*	*	..	*
x_5	0	0	0	0	0	0	0	0	..	*	*	..	*
..	..	..	..	..	..	..	..	..	..	..	..	..	..
x_n	0	0	0	0	0	0	0	0	..	1	1	..	1

Remarquons que les lignes du nouveau tableau $T°$ sont linéairement indépendantes, ce qui n'est pas forcément le cas du tableau donné T. L'algorithme présenté ci-dessous construit $T°$ colonne par colonne. Si les lignes de T sont liées, alors à la $i^{ième}$ étape, on ne pourra pas trouver de colonne dont le $i^{ième}$ terme soit non nul; c'est-à-dire que la $i^{ième}$ ligne du tableau sera nulle. Dans ce cas, on supprime du tableau la ligne nulle, ce qui revient à supprimer une variable du polynôme donné.

Algorithme de recherche d'un repère adhérent dans Z^3.

L'ensemble des points $M_o,..., M_p$ est donné par le tableau suivant :

	M_o	M_1	M_2	M_3	M_4	M_5	M_6	M_7		M_{n-1}	M_n
ℓ_1	a_o	a_1	a_2	a_3	a_4	a_5	a_6	a_7	..	a_{n-1}	a_n
ℓ_2	b_o	b_1	b_2	b_3	b_4	b_5	b_6	b_7	..	b_{n-1}	b_n
ℓ_3	c_o	c_1	c_2	c_3	c_4	c_5	c_6	c_7	..	c_{n-1}	c_n
ℓ_o	1	1	1	1	1	1	1	1	..	1	1

On a adjoint au tableau une ligne de 1 qui permettra de translater le repère. Ainsi en retranchant un multiple de ℓ_o à chacune des trois lignes, on peut supposer que chacune d'elle contient un ou plusieurs termes nuls.

Supposons par exemple que $a_o = a_1 = .. = a_i = 0$, avec $b_o \leq b_1 \leq ... \leq b_i$. Alors, on prend comme nouvelle ligne ℓ_2, la ligne $\ell_2' = \ell_2 + \lambda\, \ell_1 - b_o\, \ell_o$, où λ est suffisamment grand pour que tous les termes de ℓ_2' soient positifs. On procède de la même manière pour annuler c_o. Notons toujours ℓ_1, ℓ_2 et ℓ_3 les trois lignes du nouveau tableau dans lequel la première colonne est nulle.

Pour annuler b_1, on peut remplacer ℓ_2 par $\ell_2' = a_1\, \ell_2 - b_1\, \ell_1$. Mais la nouvelle ligne ℓ_2' doit rester positive; il faut donc que a_1 / b_1 soit égal au plus grand quotient a_i / b_i; supposons que ce soit le cas.

Pour annuler c_1, on procède de la même manière, après avoir ajouter $\lambda\, \ell_2'$ à ℓ_1, avec λ suffisamment grand pour que le quotient a_1 / c_1 majore les quotients a_i / c_i.

On peut ensuite annuler, l'un des deux termes b_2 ou c_2 de la même manière que l'on a annulé b_1. Supposons donc que $c_2 = 0$. Pour pouvoir ensuite annuler a_2 de manière analogue, il faut que le quotient b_2 / c_2 majore les autres quotients b_i / c_i. Pour cela il suffit d'ajouter $\lambda\, \ell_3$ à la ligne ℓ_2 avec λ suffisamment grand.

Il reste enfin à annuler par exemple a_3 et b_4. On annule a_3 en supposant que le quotient c_3 / a_3 majore les quotients c_i / a_i, et on annule b_4 en supposant que le quotient c_4 / b_4 majore les quotients c_i / b_i.

Il peut arriver que les points M_3 et M_4 soient confondus et qu'il appartienne au troisième axe $B°x_3$ du repère adhérent. La colonne de M_3 a alors deux coordonnées nulles, mais les coordonnées des autres points M_4, ..., M_p sont quelconques.

On obtient ainsi le tableau suivant correspondant à un repère adhérent :

	M_o	M_1	M_2	M_3	M_4	M_5	M_6	M_7	..	M_{n-1}	M_n
ℓ_1	0	a_1	0	0	a_4	a_5	a_6	a_7	..	a_{n-1}	a_n
ℓ_2	0	0	b_2	b_3	0	b_5	b_6	b_7	..	b_{n-1}	b_n
ℓ_3	0	0	0	c_3	c_4	c_5	c_6	c_7	..	c_{n-1}	c_n

Exemple.

Considérons le polynôme suivant $P = Y^2Z^2 + X\,Z + X\,Y^2Z + X^2Y^3$, qui sera utilisé dans la suite pour les autres types de simplification. La tableau T de départ s'écrit :

	M_o	M_1	M_2	M_3
ℓ_1	1	0	1	2
ℓ_2	0	2	2	3
ℓ_3	1	2	1	0

En faisant apparaître des zéros dans la première colonne, la première ligne s'annule; cela signifie qu'un repère de Z^2 suffira pour représenter le polyèdre, qui est en réalité un polygone. On obtient ainsi les tableaux suivants :

	M_o	M_1	M_2	M_3
ℓ_1	0	2	2	3
ℓ_2	0	3	2	2

puis

	M_o	M_1	M_2	M_3
	0	0	4	10
	0	3	2	2

Le polynôme simplifié s'écrit donc $P = 1 + Y^{15} + X^4Y^2 + X^{10}$.

III. CONCLUSION.

On peut imaginer d'autres simplifications que celles qui ont été définies ci-dessus en utilisant les mêmes principes et en résolvant des programmes linéaires. Le principe général de ces algorithmes est de décomposer la simplification du polynôme, variable par variable: il y a n étapes pour un polynôme à n variables. Cette décomposiion ramène la simplification à la résolution de programmes linéaires de taille raisonnable. Mais les propriétés recherchées pour le polynôme simplifié: posséder un monôme constant, diminuer le degré suivant certaines variables,... doivent être elles-mêmes décomposables variable par variable. On pourrait envisager des simplifications plus générales, par exemple on pourrait rechercher un polynôme simplifié qui soit unitaire et qui possède un monôme constant (ce qui n'est possible que si le polyèdre associé au polynôme se projette à l'intérieur d'une de ses arêtes). Une telle simplification faciliterait beaucoup la factorisation du polynôme en facteurs irréductibles. Mais le programme linéaire correspondant à cette simplification serait très gros: il y aurait déjà n^2 variables et n^2 contraintes pour traduire que les coordonnées des points restent positives.

En conclusion, les algorithmes proposés peuvent être généralisés pour transformer un polynôme donné en un polynôme ayant des propriétés quelconques portant sur les degrés des monômes, mais la résolution à l'aide de programme linéaire n'est simple et rapide que si ces propriétés sont décomposables variable par variable.

CONCLUSION GENERALE.

En conclusion de cette thèse, on peut dégager deux idées nouvelles qui ont été utilisées, à savoir le développement logarithmique, et la représentation polyèdrale pour les polynômes à plusieurs variables. On a ainsi mis en évidence deux outils qui ont permis de simplifier partiellement les algorithmes de factorisation.

Mais ces outils pourraient certainement être utilisés pour résoudre d'autres problèmes.

Examinons d'abord le développement logarithmique. On a vu au B, Chapitre V qu'on pouvait généraliser la définition aux polynômes à plusieurs variables et en déduire des propriétés intéressantes sur les diviseurs du polynôme. On a pu ainsi donner un algorithme de factorisation de coût presque polynômial suivant le degré, ce coût étant du même ordre (à un coefficient d près), que celui de l'algorithme classique dans les autres étapes de l'algorithme.

On a vu également comment définir le développement logarithmique d'une fraction rationnelle au A, Chapitre V, 2° et on en a déduit un algorithme de calcul du quotient entier de deux polynômes écrits sous forme de puissances. On pourrait définir de façon analogue le développement logarithmique d'une exponentielle ou d'un logarithme. L'une des applications possibles serait le calcul du développement limité d'une telle expression comportant des produits, puissances, quotients, logarithmes et exponentielles.

Examinons maintenant la représentation polyèdrale des polynômes à plusieurs variables. Ce type de représentation des polynômes a permis de mettre en évidence des propriétés intéressantes sur les polynômes à plusieurs variables, notamment la notion de polynôme normalisé, qui caractèrise les diviseurs d'un polynôme. Cette représentation a montré aussi l'existence de monômes spéciaux, appelés monômes extrêmaux jouant un rôle intéressant dans le produit des polynômes: tout monôme extrêmal du produit est égal à un produit unique de monômes extrêmaux des diviseurs. La dernière partie C a montré une autre utilité de la représentation polyèdrale, celle de simplifier un ou plusieurs polynômes qu'on veut étudier. On peut simplifier dans un sens précis, par exemple pour obtenir un polynôme unitaire, abaisser le degré, diminuer le nombre de variables lorsque

c'est possible. Dans chaque cas, on doit résoudre un programme linéaire, et on est assuré de trouver la meilleure solution qui existe. On pourrait utiliser cette représentation polyèdrale dans le calcul du PGCD de plusieurs polynômes, pour simplifier les polynômes étudiés.

Ostrowski, qui a introduit cette notion dans [Os], l'a utilisée pour obtenir des critères d'irréductibilité de polynômes, et pour étudier certaines factorisations particulières. Ce type de représentation des polynômes à plusieurs variables peut donc être utile dans beaucoup de domaines.

Pour terminer remarquons une caractèristique commune à ces deux outils que sont le développement logarithmique et la représentation polyèdrale. Si on note $\Phi(P)$ le développement logarithmique d'un polynôme P et $\Pi(P)$ la représentation polyèdrale de P, alors $\Phi(P\cdot Q) = \Phi(P) + \Phi(Q)$ et $\Pi(P\cdot Q) = \Pi(P) + \Pi(Q)$. Chacun transforme donc un produit en somme; pour cette raison, ils permettent tous les deux de simplifier l'étude des produits, et par suite l'étude de la factorisation.

BIBLIOGRAPHIE.

[Be] **E.R.Berlekamp**, Algebraic Coding Theory, Mac Graw-Hill, New-York, 1968.

[Bo] **W.S. Brown**, On Euclid's Algorithm and the Computation of Polynomial Greatest Commun Divisors, J. Ass. Comp. Mach., Vol. 18, p. 478-504, 1971.

[BCL] **B. Buchberger, G.E. Collins, R. Loos**, Symbolic and Algebraic Computation, Springer-Verlag, Computing Supplementum 4, Wien-New-York, 1982.

[Ca] **P. Camion**, A deterministic Algorithm for Factoring Polynomials of $F_q[X]$, Annals of Discrete Math. 17, North-Holland Publishing Company, p.149-157, 1983.

[C1] **G.E. Collins**, Subresultants and reduced polynomial remainder sequences, J. Ass. Comp. Mach., vol. 14, N°1, p.128-142, 1967.

[C2] **G.E. Collins**, The calculation of Multivariate polynomial Resultants, J. Ass. Comp. Mach., vol. 18, N°1, p.515-534, 1971.

[CZ] **D.G. Cantor, H. Zassenhaus**, A new algorithms for factoring polynomials over finite fields, Math. of Comp., Vol 36, p. 587-592, 1981.

[DST] **J. Davenport, Y. Siret, E. Tournier**, Calcul Formel, Collection Etudes et Recherches en Informatique, Masson, Paris, 1987.

[CMP] **L. Cerlienco, M. Mignotte, F. Piras**, Computing the measure of a polynomial, J. Symbolic. Comp. 4, N° 1, p. 21-34, 1987.

[Du] **A. Durand**, A propos d'un théorème de S. Berstein sur la dérivée d'un polynôme, C.R. Acad. Sci. Paris Sér. I Math. 290, p. 523-525, 1980.

[DR] **J.D. Donaldson, Q.I. Rahman**, Inequalities for polynomials with a prescribed zero, Pacific J. Math. 41, p.375-378, 1983.

[CGG] **B.W. Char, K.O. Geddes, G.H. Gonnet**, Heuristic Polynomial GCD Algorithm Based on Integer GCD Computation, EUROSAM 84, Lecture Notes in Computer Science, N°174, p.285-296, 1984.

[Ge] **A. O. Gel'fond**, Transcendental and Algebraic Numbers, GITTL, Moscou, 1952, trad. anglaise, Dover, New-York, p. 135., 1960.

[Gr] **Grünbaum**, Convex Polytopes, John Wiley, New-York, 1967.

[Gu] **R. Güting**, Polynomials with multiple zeroes, Mathematika 14, p.181-196, 1967.

[Ka] **E. Kaltofen**, Factorisation of Polynomials, Symbolic and Algebraic Computation, Springer–Verlag, Computing Supplementum 4, Wien–New–York, 1982.

[Kn] **D.E. Knuth**, The Art of Computer Programming, vol. II, Addison–Wesley, 1969.

[KT] **H.T. Kung, D.M. Tong**, Fast algorithms for partial fraction decomposition, SIAM J. Comp., vol 6, n°3, p. 582–593, 1977.

[L3] **A.K. Lenstra, H.W. Lenstra, L. Lovasz**, Factoring Polynomials with rational coefficients, Math. Ann. 261, p. 515–534, 1982.

[Lu1] **D. Lugiez**, Algorithmes de factorisation de polynômes, Thèse de 3ième cycle, INPL, Grenoble, 1984.

[Lu2] **D. Lugiez**, Heuristic Bivariate Lifting, AAECC, Lecture Notes in Computer Science, N°229, p.384–391, 1986.

[Mo] **R.T. Moenck**, On the efficiency of algorithms for Polynomial factoring, Math. of Comp., vol. 31, n°137, p.235–250, 1977.

[M1] **M. Mignotte**, Mathématiques pour le Calcul Formel, Presses Universitaires de France, 1989.

[M2] **M. Mignotte**, An inequality about factors of Polynomials, Math. Comp., Vol. 28, p. 1153–1157, 1974.

[M3] **M. Mignotte**, Some useful Bounds, Symbolic and Algebraic Computation, Computing Supplementum 4, Springer–Verlag, Wien–New–York, p.259–263, 1982.

[M4] **M. Mignotte**, An inequality about Irreductible Factors of Integer Polynomials, J. Number Theory, Vol.30, N°2,p.156–166, 1988.

[MN] **M. Mignotte, J–L Nicolas**, Ann. Inst. Henri Poincaré, V.19,N°2, p. 113–121.

[MSP] **R. Mohr, R. Schott, C. Pair**, Construire les Algorithmes, Dunod–Informatique, 1988.

[Mu] **D.R. Musser**, Multivariate Polynomial Factorization, J. Ass. Comp. Mach., vol. 22, n°2, p. 291–308, 1975.

[Os] **A.M. Ostrowski**, On Multiplication Polynomial Factorization, Aequationes Math., vol. 13, p. 201–228, 1975.

[RS] **J.B. Rosser, L. Schoenfeld**, Approximate formulas for some functions of prime numbers, Illinois J. Math., V.6, 1962, pp. 64–94, Theorem 12.

[Si] **Simonnard**, Programmation linéaire, Tome 2, Dunod, Paris, 1973.

[Va] **B. Vallée**, Une approche géométrique de la réduction des réseaux en petite dimension, Thèse d'Université, Université de Caen, 1986.

[V1] **G. Viry**, Factorisation des polynômes à plusieurs variables à coefficients entiers, RAIRO Informatique Théorique, vol. 12, n°4, p. 305–318, 1978.

[V2] **G. Viry**, Factorisation des polynômes à plusieurs variables, RAIRO Informatique Théorique, vol. 14, n°2, p.209–223, 1980.

[V3] **G. Viry**, Simplification of Polynomials in n variables, EUROSAM 84, Lecture Notes in Computer Science, N°174, p.64–73, 1984.

[V4] **G. Viry**, Polynomial Factorization over Z[X], A.A.E.C.C.–3, Lecture Notes in Computer Science, N°229, p.326–332, 1986.

[V5] **G. Viry**, Multiplication of Polynomials. Application to the Factorization, over Z[X], EUROCAL 87, Leipzig, 1987.

[V6] **G. Viry**, Factorisation sur Z[X] des polynômes de degré élevé à l'aide d'un monomorphisme, à paraître dans RAIRO Informatique Théorique.

[VW] **B. L. van der Waerden** , Algebra, Edition anglaise, Ungar Publishing, New–York, 1970.

[Wa] **P.S. Wang**, An Improved Multivariate Polynomial Factoring Algorithm, Math. of Comp., Vol. 32, p. 1215–1231, 1978.

[WR] **P.S. Wang, L.P. Rothschield**, Factoring Multivariate Polynomials over the Integers, Math. of Comp., Vol. 29, p. 935–950, 1975.

[Za] **H. Zassenhaus**, On Hensel Factorization, J. Number Theory, Vol. 1, p.291–311, 1969.

NOM DE L'ETUDIANT : VIRY Guy

NATURE DE LA THESE : Doctorat de l'Université de NANCY I en Informatique

VU, APPROUVE ET PERMIS D'IMPRIMER

NANCY, le 11 SEP. 1989 n°1600

LE PRESIDENT DE L'UNIVERSITE DE NANCY I

UNIVERSITE DE NANCY I
LE PRESIDENT
M. BOULANGE

Résumé de la thèse.

Cette thèse présente les algorithmes classiques de factorisation des polynômes à une et plusieurs variables. Dans le cas des polynômes à une variable, deux nouvelles méthodes sont proposées. Dans la première, on calcule un facteur linéaire modulo un nombre premier p, puis on définit unh multiple de ce facteur qui divise P sur Z[X]. Dans la seconde méthode, le calcul des produits des facteurs de P modulo p^n est remplacé par le calcul des sommes des images de ces facteurs de P.

Dans le cas des polynômes à plusieurs variables, on donne deux méthodes pour diminuer les calculs de la dernière étape de la factorisation avec les algorithmes classiques. On utilise la représentation polyédrale et la notion de polynômes "normalisés".

La dernière partie de la thèse donne une méthode pour diminuer le degré total du polynôme donné, lorsque c'est possible. Cette méthode utilise la programmation linéaire entière.

Mots–Clés.

Polynômes à une variable, Polynômes à plusieurs variables, Factorisation, Méthode de Berlekamp, Algorithme de Lenstra, Polyèdre, Programmation linéaire entière.

Résumé de la thèse.

Cette thèse présente les algorithmes classiques de factorisation des polynômes à une et plusieurs variables. Dans le cas des polynômes à une variable, deux nouvelles méthodes sont proposées. Dans la première, on calcule un facteur linéaire modulo un nombre premier p, puis on définit unh multiple de ce facteur qui divise P sur Z[X]. Dans la seconde méthode, le calcul des produits des facteurs de P modulo p^n est remplacé par le calcul des sommes des images de ces facteurs de P.

Dans le cas des polynômes à plusieurs variables, on donne deux méthodes pour diminuer les calculs de la dernière étape de la factorisation avec les algorithmes classiques. On utilise la représentation polyédrale et la notion de polynômes "normalisés".

La dernière partie de la thèse donne une méthode pour diminuer le degré total du polynôme donné, lorsque c'est possible. Cette méthode utilise la programmation linéaire entière.

Mots-Clés.

Polynômes à une variable, Polynômes à plusieurs variables, Factorisation, Méthode de Berlekamp, Algorithme de Lenstra, Polyèdre, Programmation linéaire entière.