Unité de recherche CNRS UMR 7062,
Centre d'Histoire des Sciences et des Philosophies Arabes et Médiévales,
7 rue Guy Moquet, BP 08, 94801 Villejuif, France.

Histoire de la théorie de l'élimination

Thèse de doctorat de l'Université Paris VII

Erwan Penchèvre

Directeur de thèse : Christian Houzel, Professeur émérite

Jury :

Massimo Galuzzi, Professeur à l'Université de Milan,
Jean-Pierre Jouanolou, Professeur à l'Université de Strasbourg,
Roshdi Rashed, Directeur de recherche émérite au C. N. R. S,
Jean-Jacques Sansuc, Professeur à l'Université Paris VII,
Ivahn Smadja, Maître de conférences à l'Université de Caen.

Je remercie vivement le Professeur Christian Houzel, sans la direction duquel cet ouvrage n'eût été possible, mais aussi le Professeur Roshdi Rashed, pour son soutien et ses nombreux encouragements, ainsi que tous les membres du jury, pour avoir accepter d'y participer.

Merci également à ceux qui m'ont aidé à comprendre leurs propres recherches dans les domaines dont j'ai écrit l'histoire : Professeur Jean-Pierre Jouanolou, Marc Giusti, Directeur de recherche, et Professeur Mireille Martin-Deschamps. Je dois aussi m'excuser de n'avoir peut-être pas donné une place suffisante à leurs travaux dans mon récit : ils écrivent l'avenir d'une discipline dont j'ai décrit le passé.

Merci aussi à Tsuyoshi Goto, Katia Asselah, Clémence Durvye et Sébastien Maronne, ainsi qu'aux membres du Centre d'Histoire des Sciences et des Philosophies Arabes et Médiévales de Villejuif, et aux participants au semestre de cours d'Histoire des Mathématiques de la Chaire U.N.E.S.C.O. Mathématiques et Développement à Tunis.

Merci enfin à mes parents, pour leur support matériel et intellectuel,

et à mon épouse, Françoise, à qui je dédie cet ouvrage commencé au lendemain de notre mariage.

Table des matières

Introduction **9**

I L'élimination algébrique au XVII^{e} siècle 17

1 L'œuvre algébrique de Johannes Faulhaber **19**
1.1 La « cubiccoß » 20
1.2 Thèmes représentés dans les 160 problèmes du *Cubicossischer Lustgarten* 24
1.3 L'œuvre algébrique de Faulhaber 30
1.4 Conclusion 38

2 Hudde, Newton, Leibniz **41**
2.1 Hudde et la méthode des coefficients indéterminés 41
2.2 Deux brouillons de Newton : l'élimination successive des plus hautes puissances . . . 47
2.3 Le théorème de Bézout chez Newton 49
2.4 Les manuscrits de Leibniz 51

3 Origine de l'élimination algébrique **55**
3.1 Abū Kāmil 55
3.2 La *syncrisis* de Viète 60
3.3 Le *Sìyuán yùjiàn* de Zhū Shìjié 62
3.4 La méthode de Fermat 65
3.5 Méthodes, règles, formules : l'élimination algébrique au XVII^{e} siècle 68
3.6 L'ordre adopté dans la résolution des problèmes 70

4 Les « arbres » de Rolle **73**

II Le théorème de Bézout 85

5 Maclaurin : géométrie et élimination **87**

6 Euler : une théorie de l'intersection **105**
6.1 Les trois méthodes d'Euler 105
6.2 Intersection et projection 108

7 Paradoxe de Cramer et théorème de Bézout **113**
7.1 Le théorème de Bézout démontré par Cramer et Poisson 115
7.2 Le paradoxe de Cramer 116

8 Les premières recherches de Bézout **119**
8.1 Cas de deux équations ; la matrice de Bézout . . . 121
8.2 Le cas de trois équations : échec . . . 123
8.3 La matrice de Bézout chez le mathématicien japonais Seki Takakazu . . . 124

9 Les *Réflexions sur la résolution des équations algébriques* de Lagrange **127**
9.1 Exposé par Lagrange de la méthode de Tschirnhaus pour le degré 3 . . . 127
9.2 Méthodes d'Euler et Bézout : «résolvantes de Lagrange» . . . 133
9.3 L'application de la théorie de l'élimination aux grands problèmes de l'algèbre . . . 135

10 Le Théorème de Bézout pour un nombre quelconque d'équations **139**
10.1 La théorie de l'élimination dans la première moitié du XIX$^{\text{ème}}$ siècle . . . 142
10.2 La *Théorie générale des équations algébriques* (1779) de Bézout . . . 148
10.3 Trois conceptions de l'élimination . . . 150
10.4 La démonstration de Bézout . . . 152
10.5 Des termes d'introduction fictive . . . 156
10.6 Le complexe de Koszul . . . 158
10.7 Les « équations incomplètes » . . . 161
10.8 Cas de deux coniques : problème des relations entre les coefficients . . . 167

11 Naissance d'une théorie **173**
11.1 Raisonnement *a priori* et calcul effectif . . . 177

III De Jacobi à Kronecker **181**

12 Jacobi et Sylvester **183**
12.1 Le bézoutiant de Sylvester . . . 183
12.2 Les facteurs superflus dans l'algorithme euclidien . . . 186
12.3 Le « cercle d'idées Sturmiennes » . . . 189
12.4 C. G. Jacobi . . . 192
12.5 Le bézoutiant est un combinant . . . 197

13 Courbes algébriques planes et polygone de Newton **199**
13.1 Exposé élémentaire : intersection de deux courbes algébriques planes . . . 199
13.2 Polygone de Newton et ramification . . . 202
13.3 Application : le genre (Riemann, Weierstrass, M. Noether) . . . 211

14 La multiplicité d'intersection **221**
14.1 Le concept de multiplicité d'intersection . . . 221
14.2 Le mémoire de Halphen sur la détermination des coniques (1873) . . . 229

15 Théorie arithmétique de la grandeur algébrique : Kronecker (1882) **233**
15.1 Théorie des corps . . . 233
15.2 Grandeur entière et discriminant . . . 235
15.3 Théorie de l'élimination . . . 237
15.4 Application : la théorie de Galois . . . 242
15.5 Théorie de la divisibilité dans les anneaux d'entiers algébriques . . . 245
15.6 Systèmes de diviseurs et codimension . . . 250

16 Elimination et géométrie dans la seconde moitié du XIX$^{\text{e}}$ siècle **259**
16.1 Cayley, Brill : le résultant réduit . . . 259
16.2 L'histoire de la théorie de l'élimination selon Brill . . . 265
16.3 Conclusion : géométrie algébrique et élimination de l'élimination . . . 269

IV Annexes **273**

A Le paragraphe § 62 de la *Théorie générale des équations algébriques* **275**

B Le *Mémoire sur l'élimination dans les équations algébriques* de Poisson **277**

C Etude des facteurs superflus dans les méthodes du XVII^{e} siècle avec Maple **281**

D Vérification expérimentale des formules de Bézout avec le logiciel Macaulay2 **285**

E Extrait de la préface des *Meditationes Algebraicae* de Waring (1782) **299**

Bibliographie **303**

Introduction

Le sujet de cet ouvrage est l'histoire de la théorie de l'élimination en algèbre, c'est-à-dire l'élimination des inconnues au sein d'un système d'équations polynomiales. Notre recherche est de nature locale, au sens où elle se repère au sein de l'histoire d'une discipline, l'algèbre. Nous avons exclu tout ce qui pourrait être reconnu comme technique ou méthode d'élimination en dehors de cette discipline. Mais l'algèbre, comme tout autre discipline, n'étant pas figée dans l'histoire, il a parfois fallu faire des choix. Ainsi la méthode de l'analyse, en géométrie classique, consistait, dans la résolution d'un problème, à déterminer les grandeurs géométriques inconnues (par exemple la position d'un point, un segment de droite ou un secteur angulaire) en travaillant sur elles comme si elles étaient connues, à la recherche d'une relation simple entre ces grandeurs et des grandeurs données. Cette méthode, traduite dans le langage de l'algèbre, consiste en la résolution d'un système d'équations à plusieurs inconnues et implique donc nécessairement l'idée d'élimination. Or les applications réciproques entre géométrie et algèbre ne sont pas rares, dès le X^{e} siècle en langue arabe, et les problèmes d'algèbre à plusieurs inconnues non plus. Pourtant la traduction explicite de problèmes de géométrie « à plusieurs inconnues » en des problèmes d'algèbre à plusieurs inconnues n'apparaît pas, à notre connaissance, avant le XVII^{e} siècle, et nous ne parlerons donc pas, par exemple, de l'histoire antérieure des « problèmes de constructions géométriques » [1].

De même nous n'avons pas étudié les aspects purement formels de l'élimination. L'élimination des inconnues peut certes être comprise comme cas particulier de l'élimination de quantificateurs en logique. On peut par exemple décrire ainsi la notion de résultant. Soit un système de n équations à $(n-1)$ inconnues, l'élimination consiste à transformer la proposition

$$\exists x_1 \exists x_2 ... \exists x_{n-1} \left([P_1(x_1, x_2, ..., x_{n-1}) = 0] \wedge [P_2(x_1, x_2, ..., x_{n-1}) = 0] \wedge ... \wedge [P_n(x_1, x_2, ..., x_{n-1}) = 0] \right)$$

en la suivante :

$$\mathrm{Res}(P_1, P_2, ..., P_n) = 0$$

qui ne contient plus aucun quantificateur. Nous n'avons pas fait l'histoire d'un formalisme ou d'une théorie logique, mais bien l'histoire d'un chapitre de l'algèbre. Mais l'on est bien conscient de l'influence réciproque possible entre algèbre et logique [2].

Malgré les restrictions qui s'imposaient d'elles-mêmes à notre sujet, la tâche restait vaste. L'histoire de l'élimination en algèbre, c'est l'histoire de toute l'algèbre. Jamais une algèbre pure, cependant : jusqu'à la fin du XVIII^{e} siècle, le noyau essentiel de l'algèbre se réduit à une théorie des équations à une seule inconnue, de laquelle l'élimination est absente. C'est dans les applications de l'algèbre qu'il fallait donc chercher l'histoire ancienne de l'élimination algébrique. Cela suffira, si besoin est, à expliquer les nombreuses pages que nous avons consacrées à des auteurs tels Abū Kāmil, Cardan, Faulhaber, Descartes, Fermat. On distinguera pourtant une histoire longue et une histoire courte. L'histoire courte commence seulement dans la seconde moitié du XVII^{e} et peut-être à partir de Hudde : notre chapitre 3 expliquera cela davantage.

Le but le plus élevé que nous nous proposons dans ce travail est de mettre à disposition du spécialiste – au rang duquel nous ne pouvons mériter accéder du fait de nos lectures d'une portion

1. Sur ce sujet, voir l'ouvrage récent de R. Rashed [263].
2. Voir l'ouvrage de H. Sinaceur [291]. Pour un exemple antérieur, *cf.* [207] sur l'œuvre de Boole.

encore congrue des travaux récents dans ce domaine [3] – les moyens de juger du présent et de l'avenir de sa discipline, par la connaissance du passé ; ce faisant, nous espèrons aussi combler une lacune dans la bibliothèque de l'historien des mathématiques.

Dans un premier chapitre, nous exposons l'œuvre algébrique du mathématicien allemand de la première moitié du XVII^{e} siècle, Johannes Faulhaber. Sans donner de méthode générale, Faulhaber avait recours à l'élimination dans l'usage qu'il faisait de la méthode des coefficients indéterminés, pour calculer certaines formules faisant intervenir des sommes de séries de puissances, ainsi que pour résoudre algébriquement des équations à une inconnue. Ce faisant, nous essayons de montrer la spécificité de l'œuvre algébrique de Faulhaber, en la considérant en elle-même, c'est-à-dire sans la comparer aux mathématiques de Descartes.

Dans le chapitre 2, nous analysons la contribution de Jan Hudde, puis les travaux de Newton et Leibniz sur l'élimination, qui ont été inspirés par ceux de Hudde. Les recherches de Hudde nous semblent marquer une étape importante dans un mouvement d'arithmétisation de l'algèbre ; mais il détourne aussi l'usage d'un algorithme arithmétique connu, en appliquant l'algorithme euclidien de calcul du plus grand commun diviseur de deux polynômes à l'élimination des inconnues. Il offre ainsi la première méthode générale d'élimination – si l'on excepte une méthode antérieure due à Fermat mais restée dans l'oubli et accompagnée d'une erreur grossière. Il utilise cette méthode dans la recherche des facteurs rationnels d'une équation à une inconnue, *via* la méthode des coefficients indéterminés. Cette approche ressemble beaucoup à la méthode de résolution de l'équation de degré 4 qu'utilisaient Faulhaber et Descartes, et Hudde est ainsi conduit à attribuer la méthode des coefficients indéterminés à Descartes. Nous discutons du bien-fondé de cette attribution. D'autre part, Hudde utilise aussi l'algorithme euclidien pour déterminer les racines multiples d'une équation à une inconnue. Newton s'intéressait à l'élimination algébrique dès la fin des années 1660 et il l'enseigna à Cambridge. Sa méthode d'élimination, que nous désignons par la locution « méthode d'élimination successive des plus hautes puissances de l'inconnue » est légèrement différente de l'algorithme euclidien. C'est, semble-t-il, cette méthode qui lui a permis de calculer les résultants de deux équations de degrés m et n pour les couples (m, n) suivants : $(2, 2), (3, 2), (4, 2), (3, 3)$. D'autre part, il est aussi le premier, à notre connaissance, à avoir énoncé le théorème de Bézout pour deux courbes algébriques planes. Les travaux de Leibniz sur l'élimination sont restés ignorés pendant longtemps et ils n'existent qu'à l'état de brouillons et notes manuscrites publiés récemment par E. Knobloch. Ils sont liés à la découverte par Leibniz du concept de déterminant. Leibniz connaissait ainsi l'expression du résultant (m, n) sous forme d'un déterminant de $(m+n)$ lignes. Mais il proposa aussi une autre méthode d'élimination, « par conjonction d'équations inverses », qui ressemble à celles de Hudde et Newton. Enfin, rebuté par les difficultés calculatoires de l'élimination, il entreprit une étude combinatoire de l'algorithme euclidien du plus grand commun diviseur, dans le but de trouver des règles syntaxiques décrivant la formule de l'équation finale à laquelle il conduit, règles qui aurait permis d'écrire l'équation finale sans avoir à effectuer l'algorithme. Il est ainsi amené à étudier les propriétés d'homogénéité du résultant.

Le chapitre 3 présente la première étape d'une réflexion globale sur le commencement historique et le devenir de l'élimination en algèbre. Nous y étudions d'abord quelques travaux antérieurs à 1650, sans toutefois le faire d'une manière aussi exhaustive qu'avec Faulhaber :

- les procédés qu'employait l'algèbre arabe, et en particulier Abū Kāmil (IX^{ème} siècle) pour résoudre des systèmes de plusieurs équations à plusieurs inconnues,
- la méthode de *syncrisis* de Viète que l'on peut interpréter comme résolution d'un système d'équations linéaires intervenant dans l'étude de la structure (*constitutio*) d'une équation à une inconnue,
- une méthode d'élimination chinoise, que l'on trouve dans un ouvrage du XIII^{e} siècle, appliquée à la résolution de systèmes à deux, trois ou quatre inconnues, mais limitée par l'absence d'une théorie arithmétique des polynômes (à en juger par cet ouvrage, les mathématiciens chinois de cette période ne pouvaient pas représenter un polynôme général à trois inconnues),

3. Nous reviendrons *infra* p. 272 sur les conséquences qu'aurait eues pour nous une prise en compte plus affirmée de ces travaux récents.

– une méthode d'élimination proposée par Fermat, mais que celui-ci ne semble pas avoir su appliquer en géométrie d'une manière aussi circonspecte que Hudde sept ans plus tard.

Nous comparons ensuite mathématiquement les méthodes de Fermat, Hudde, Newton et Leibniz. Nous remarquons qu'il ne peut encore être question d'une véritable *théorie* de l'élimination, bien que ces méthodes générales marquent une étape importante dans l'histoire de l'élimination et un changement dont nous essayons ensuite de déceler le principe dans les mathématiques du XVII$^{\text{ème}}$ siècle.

Bien que l'élimination naisse chez Hudde dans l'élaboration des mathématiques cartésiennes, Descartes est en quelque sorte absent de cette évolution ; la théorie se développe même contre Descartes, puisqu'elle met en échec, par le théorème de Bézout, sa classification des équations. Ce retournement s'accentue encore dans l'œuvre de Rolle, que nous commentons dans le chapitre 4. Les méthodes d'élimination antérieures procédaient en suivant un ordre séquentiel de transformation d'une ou plusieurs équations au sein d'un système unique et en quelque sorte statique, à l'image de l'analyse en géométrie classique. La méthode de Rolle adopte au contraire un ordre que l'on pourrait qualifier d'arborescent : étant donné un système d'équations initial, Rolle le transforme, au terme de la construction d'un « arbre de direction », en plusieurs systèmes, dont la réunion des solutions constitue l'ensemble des solutions du système proposé. Bien que Leibniz construisît aussi des arbres (pour étudier des algorithmes récursifs tels que la division euclidienne), ses arbres sont plutôt des arbres syntaxiques que l'on peut rattacher à son goût pour la combinatoire. L'idée de Rolle, peut-être moins élaborée que celle de Leibniz, semble cependant nouvelle. On en trouve l'écho au XIX$^{\text{e}}$ siècle dans les travaux d'Ossian Bonnet. Enfin, Rolle généralise le théorème de Bézout à un nombre quelconque d'équations, sans toutefois le démontrer.

A l'instar du théorème fondamental de l'algèbre, de nombreuses démonstrations du théorème de Bézout vont voir le jour, posant aussi plusieurs problèmes conceptuels. Ce théorème est au cœur des travaux commentés dans nos chapitres 5 à 11. Le chapitre 5 présente l'œuvre de Maclaurin en théorie des courbes algébriques. Dans l'application du théorème de Bézout, Maclaurin compte les points d'intersection avec multiplicité. Ceci suppose une classification (encore grossière) des singularités des courbes algébriques et un concept (pas encore nommé) de multiplicité d'intersection de deux courbes en un point. Le produit des degrés des équations n'est plus seulement conçu comme une borne au nombre de solutions, mais comme leur nombre exact. Le théorème de Bézout est alors perçu comme un théorème général. Malgré les nombreuses exceptions auxquelles il est encore soumis (dues à la présence de points imaginaires et de points à l'infini), Maclaurin en cherche en effet la démonstration : il prétend même l'avoir démontré dans un supplément non publié à la *Geometria Organica*, qui n'a pas été retrouvé. Ce théorème permet à Maclaurin un nouveau genre de preuves purement géométriques, plus tard adopté aussi par Poncelet dans son *Traité des propriétés projectives des figures*. On verra aussi Maclaurin s'intéresser à d'autres problèmes énumératifs : combien de points déterminent une courbe de degré donné ? Quel est le nombre maximal de points doubles sur une courbe de degré donné ? Un thème traverse toute son œuvre : la « description organique » des courbes par des instruments. Certains des instruments décrits par Maclaurin pour tracer des courbes ayant des singularités données au moyen de courbes de degrés inférieurs peuvent s'interpréter comme des transformations birationnelles. Son maître Newton faisait grand cas de la recherche d'un instrument pour tracer des cubiques de genre 1. La raison de la difficulté tient au fait que, dans les instruments connus alors, le tracé était entièrement guidé par le mouvement d'un point le long d'une droite, ou la rotation d'une droite autour d'un point ; ces instruments ne pouvaient donc tracer que des courbes rationnelles. Maclaurin parvient cependant à renverser cet obstacle, et offre un instrument de nature différente pour tracer les cubiques de genre 1 ; mais un autre ouvrage, publié de manière posthume en 1748, montre qu'il avait continué de chercher – et ce jusqu'à la fin de sa vie – un autre procédé pour tracer les cubiques.

Il faut attendre les ouvrages d'Euler (1748) et Cramer (1750) pour lire les premières démonstrations du théorème de Bézout. Les travaux d'Euler et Cramer font l'objet des chapitres 6 et 7. On distingue dès lors deux voies possibles pour l'élimination. La première, avait déjà été suivie par Leibniz. La seconde, illustrée par Euler dans un mémoire en 1748, puis par Cramer dans son traité *Introduction à l'analyse des lignes courbes algébriques* en 1750, ne semblait d'abord pouvoir

s'appliquer qu'au cas de deux équations $\begin{cases} P(x)=0 \\ Q(x)=0 \end{cases}$, et supposait qu'une de ces deux équations ait un nombre de racines (a_i) égal à son degré (théorème fondamental de l'algèbre). Les coefficients des deux équations peuvent renfermer une autre inconnue y. Ainsi pour deux équations à deux inconnues x, y, l'équation finale résultant de l'élimination de x est un polynôme en y, et l'on calcule son degré au moyen d'une étude combinatoire sur les fonctions symétriques des racines (a_i) (étude menée à bien en détails seulement par Cramer). Poisson généralise cette approche pour un nombre quelconque d'équations. Son court mémoire de 1802, que nous commentons au chapitre 7, est reproduit en annexe B. Toujours dans le cas de deux équations, les racines de l'équation finale correspondent géométriquement à la projection sur l'axe des ordonnées des points d'intersection des deux courbes d'équations données. On peut faire en sorte que ces racines soient distinctes au moyen d'un changement générique d'axes de coordonnées. Le lien conceptuel entre intersection, projection et élimination est particulièrement étudié, et étendu au cas de l'intersection de deux surfaces dans l'espace, par Euler. Mais l'absence du concept d'espace projectif se fait sentir, à cause du fait que la projection

$$\begin{array}{ccc} \mathbb{A}^3 & \longrightarrow & \mathbb{A}^2 \\ (x,y,z) & \longmapsto & (x,y) \end{array}$$

n'est pas une application fermée. Se posent aussi la question des points imaginaires, celle des intersections impropres (Euler remarque que deux courbes peuvent avoir des composantes communes, et donc une infinité de points communs) et celle des « contacts » (points de contact, « lignes de contact » entre deux sufaces). Enfin le « paradoxe de Cramer », qu'avait déjà rencontré Maclaurin, montre que, dans certaines situations pourtant assez générales en théorie des courbes, un système de N équations linéaires à N inconnues n'a pas une solution *unique*. Il y a certes rupture avec un principe communément admis au XVII$^{\text{e}}$ siècle. C'est Cramer qui utilisera, en 1750, le terme de « paradoxe » pour désigner ce fait. La théorie de l'élimination dans les systèmes d'équations linéaires n'avait pas encore atteint la perfection désirable, mais Euler, sollicité par Cramer, explique le « paradoxe » par l'existence de relations de dépendance linéaire entre les N équations ; il a même l'idée de compter le nombre de ces relations.

Au chapitre 8, nous abordons les premières recherches de Bézout. Elles s'inscrivent dans la continuité des travaux d'Euler et Cramer. L'idée mise en œuvre par Bézout en 1764 est de chercher à déterminer *a priori* le degré de l'équation finale résultant de l'élimination d'un nombre quelconque d'inconnues. A cet effet, l'ensemble des opérations conduisant à l'élimination d'*une* inconnue est réduit à la formation d'une somme de produits des équations initiales par des polynômes multiplicateurs indéterminés. La méthode de Bézout échoue sauf dans le cas de deux équations, cas auquel Euler l'avait déjà appliquée. Mais dans ce cas, à la recherche d'une méthode d'élimination qui abrège les calculs, il conçoit la « matrice de Bézout », matrice qu'un mathématicien japonais avait déjà utilisée à la fin du XVII$^{\text{ème}}$ siècle.

Dans les années 1770, le théorème de Bézout pour un nombre quelconque d'équations est à nouveau dans le vent, chez Waring et Lagrange, puis chez Bézout. A la suite de Cramer, Lagrange utilise la théorie des fonctions symétriques des racines pour l'élimination algébrique ; il étudie plus généralement les fonctions des racines invariantes par certaines permutations, dans ses *Réflexions sur la résolution algébrique des équations*. Ce mémoire est l'objet de notre chapitre 9. C'est d'ailleurs la lecture des magnifiques travaux d'Euler, Bézout et Lagrange sur la résolution algébrique des équations (à une inconnue) et sur le théorème fondamental de l'algèbre qui nous avaient d'abord conduits à étudier l'histoire de la théorie de l'élimination. Le théorème fondamental de l'algèbre, la recherche d'une méthode générale de résolution algébrique et le théorème de Bézout sont peut-être les trois grands problèmes de l'algèbre de la deuxième moitié du XVIII$^{\text{e}}$ siècle. Le théorème fondamental et le théorème de Bézout ne sont d'ailleurs pas sans lien. Les deux énoncés, l'un sur l'existence, le nombre et la forme des racines d'une équation à une inconnue, l'autre sur le nombre des solutions d'un système de n équations à n inconnues, se ressemblent et soulèvent des problèmes analogues : le nombre prédit n'est-il qu'une borne supérieure ou bien est-il possible, en comptant des racines complexes (resp. points imaginaires) et des racines multiples (resp. points d'intersection multiples), de faire de ces énoncés des énoncés universels ? D'autre part l'élimination des inconnues

dans un système de n équations à n inconnues le ramène au cas d'une unique équation à une inconnue : le problème de l'existence des solutions dans le théorème de Bézout est ainsi subordonné au théorème fondamental de l'algèbre. Enfin, certaines démonstrations du théorème fondamental de l'algèbre, dont celle de Lagrange, font abondamment recours à la théorie de l'élimination.

Le chapitre 10 est en majeure partie consacré au grand traité de Bézout, la *Théorie générale des équations algébriques*. Soient $f_1, ..., f_n$ des polynômes homogènes génériques à $(n+1)$ indéterminées de l'anneau gradué $C = k[X_0, ..., X_n]$ (où k est un corps), de degrés respectifs $k_1, ..., k_n$. Soit l'idéal $I = (f_1, ..., f_n)$. Eliminons $(n-1)$ parmi les n rapports indéterminés $\frac{X_i}{X_0}$ dans le système d'équations homogènes :

$$\begin{cases} f_1 = 0 \\ \vdots \\ f_n = 0 \end{cases}$$

La détermination du degré de l'équation finale par Bézout, faite par récurrence, et en comptant le nombre de « coefficients arbitraires » et le nombre de « coefficients inutiles » – notions correspondant respectivement au nombre d'inconnues et à la dimension de l'espace vectoriel des solutions lors de l'application de la méthode des coefficients indéterminés – le conduit en fait à calculer la dimension de certains espaces vectoriels, composantes homogènes de l'anneau C ou de ses quotients. Il aboutit à l'expression suivante du « degré de l'équation finale » (où r est un entier suffisamment grand, l'expression étant alors constante) :

$$(S) \quad \dim(C/I)_r = \dim C_r - \sum \dim C_{r-k_\alpha} + \sum \dim C_{r-k_\alpha-k_\beta} - ... + (-1)^n \dim C_{r-k_1-...-k_n}$$

Le raisonnement de Bézout repose sur l'exactitude de la suite :

$$0 \quad \to \quad (C/(f_2, ..., f_n))_{r-k_1} \quad \xrightarrow{f_1} \quad (C/(f_2, ..., f_n))_r \quad \to \quad (C/(f_1, f_2, ..., f_n))_r \quad \to \quad 0$$

où la première flèche désigne la multiplication par f_1, et la deuxième la projection canonique. Fait non trivial que Bézout ne peut justifier de manière rigoureuse. Ceci revient à démontrer que le « complexe de Koszul » suivant, associé à la suite $(f_1, ..., f_n)$, est exact :

$$0 \to C_{r-k_1-...-k_n} \to \ldots \to \bigoplus C_{r-k_\alpha-k_\beta} \to \bigoplus C_{r-k_\alpha} \to C_r \to (C/I)_r \to 0$$

Au moyen du calcul aux différences finies et de (S), Bézout montre finalement que $\dim(C/I)_r = k_1 k_2 ... k_n$. Nous expliquons en détail la démonstration de Bézout dans la section 10.4, après avoir élucidé la structure de son traité dans la section 10.3 (cette section, plus destinée au lecteur potentiel du traité de Bézout, n'est pas essentielle pour lire la section 10.4). La somme alternée (S) réapparaît dans des travaux de Cayley de la fin des années 1840. Cayley, peut-être le premier, pense que derrière cette somme alternée se cache en quelque sorte un objet réel, une suite d'espaces vectoriels et des applications de l'un dans l'autre. Et en effet, la formule (S) découle immédiatement de l'exactitude du complexe de Koszul. Cayley cherche une expression de l'équation finale au moyen des mineurs des matrices de ces applications, mais ne peut entièrement mener à bien son projet ; il fallait, pour construire le complexe de Koszul, des méthodes d'algèbre extérieure qui dépassent bien sûr les outils qu'employaient Cayley. Nous avons choisi de présenter ces travaux de Cayley, ainsi que ceux, contemporains, de Sylvester et Hesse, dans la section 10.1, avant d'exposer ceux de Bézout, dont nous croyons qu'ils sont en avance sur leur temps. De même, dans la section 10.6, nous présentons des travaux postérieurs de plus d'un siècle à l'œuvre de Bézout, tel un article de Hurwitz en 1913. Ayant ainsi répondu à la question de l'influence, directe ou indirecte, des travaux de Bézout sur ses contemporains et ses successeurs, nous reprenons, dans le court chapitre 11, la réflexion globale entamée au chapitre 3. Nous y survolons les caractéristiques de cette théorie de l'élimination qui achève de naître à la fin du XVIII$^{\text{e}}$ siècle.

Au chapitre 12, nous reprenons l'œuvre de Sylvester, au sein de laquelle s'articulent théorie de Sturm, théorie des invariants et théorie de l'élimination. A la suite de Cayley, Sylvester étudie la forme quadratique associée à la matrice de Bézout et la nomme « bézoutiant ». Il montre que le

nombre de racines réelles d'une équation $f = 0$ à coefficients réels est égal au nombre de racines positives plus un du polynôme caractéristique de la matrice de Bézout de f, f'. Il étudie de manière systématique les restes des divisions successives dans l'algorithme euclidien du plus grand commun diviseur, qu'il appelle « résidus complets » ; il détermine les facteurs superflus introduits par cet algorithme. Nous utilisons sa méthode pour déterminer aussi les facteurs superflus introduits par l'algorithme newtonien d'élimination des plus hautes puissances successives. Quant aux « résidus simplifiés », délivrés des facteurs superflus, Sylvester en donne des expressions en fonction des racines des équations initiales. Nous commentons aussi un article plus ancien de Jacobi (1835) qui anticipe certains des résultats de Sylvester. Enfin, les relations entre jacobien et bézoutiant inspirent à Sylvester des recherches sur les propriétés d'invariance du bézoutiant : le bézoutiant de f, ϕ est à la fois un covariant et un « combinant » de f, ϕ.

La théorie des courbes algébriques planes est un cas privilégié dans l'étude du phénomène de point d'intersection multiple. Le concept de multiplicité d'intersection de deux courbes en un point peut être défini au moyen de la formule du résultant due à Cramer, mais la signification géométrique d'une telle définition n'apparut que dans la deuxième moitié du XIX$^{\text{ème}}$ siècle, par l'étude locale des courbes au moyen du polygone de Newton ou de procédés analogues, à la suite de travaux de Puiseux et de Riemann. Nous décrivons ces travaux au chapitre 13. Tout ceci permit par exemple à Weierstrass de donner une étude exhaustive de l'intersection de deux courbes algébriques planes, dans un texte où règne cependant une certaine tension entre un mode de représentation propre à la « théorie des fonctions algébriques » (l'objet géométrique est alors le graphe dans $\mathbb{P}^1 \times \mathbb{P}^1$ d'une fonction multivaluée $\mathbb{P}^1 \to \mathbb{P}^1$) et un mode de représentation propre à la géométrie algébrique (l'objet est une courbe de $\mathbb{P}^2$, définie par une équation algébrique). Nous expliquons aussi dans ce chapitre le concept de « point infiniment voisin » introduit par M. Noether, ainsi que l'application de la théorie de l'élimination au calcul du genre d'une courbe.

Dans un article sur le « théorème $A\phi + B\psi$ », M. Noether détermine la dimension de l'espace vectoriel des relations linéaires que les coefficients d'une série entière $f(x, y)$ doivent vérifier si et seulement si $f(x, y) \in (\phi, \psi)$, où ϕ, ψ sont des polynômes donnés. Ce nombre, en général égal au produit des ordres de ϕ et ψ en $(0, 0)$, n'est autre que la multiplicité d'intersection de $\phi = 0$, $\psi = 0$ en $(0, 0)$ telle que la définira Macaulay en 1916. Au chapitre 14, nous comparons les principales définitions de la multiplicité d'intersection qui ont vu le jour au XIX$^{\text{ème}}$ siècle, et même des définitions plus récentes comme celles de Samuel et de Serre. On est ici guidé par le sentiment que le concept de point d'intersection multiple n'est ni un concept arbitraire, ni une généralisation évidente de la notion arithmético-algébrique de racine multiple d'une équation à une inconnue, ni l'objet d'aucune intuition géométrique naïve. Nous parlons *du* concept de point d'intersection multiple, car une certaine unité se dessine à travers ces diverses définitions – unité qui est même *démontrée* par A. Weil, après qu'il a proprement axiomatisé ce concept. Ce concept permet non seulement de faire de certains théorèmes *généraux* (mais sujets à exceptions) des théorèmes *universels*, tels le théorème de Bézout ou la formule du genre d'une courbe algébrique plane de degré et de singularités donnés ; il permet aussi de définir, dans certaines situations, un calcul des cycles. Le mémoire de Halphen sur la détermination du nombre de coniques vérifiant certaines conditions données en donne un exemple déjà célèbre dans l'histoire de la géométrie énumérative.

Le chapitre 15 se présente comme un commentaire des *Grundzüge einer arithmetischen Theorie der algebraischen Grössen* de Kronecker. S'il était besoin de justifier la tâche d'un tel commentaire, que l'on pense aux *Bunte Bemerkungen* [99] manuscrites de Dedekind, récemment publiées et témoignant des difficultés insurmontables qu'avait rencontrées un grand mathématicien à la lecture de cette œuvre. Nous ne pouvons d'ailleurs prétendre avoir entièrement élucidé le contenu des *Grundzüge*. Mais dans ce chapitre, ainsi que dans ceux sur Faulhaber et Maclaurin, nous avons choisi de mener une recherche plus large que notre sujet ne semblait l'imposer. Nous commentons les *Grundzüge* page après page, alors que la méthode d'élimination de Kronecker n'en occupe qu'une petite portion. Ce faisant, nous croyons avoir mieux compris le rôle de cette méthode dans la pensée kroneckerienne, pensée arithmétisante mais, de fait, n'excluant pas l'intuition géométrique. Dans une première partie des *Grundzüge*, Kronecker construit une théorie des corps et des entiers algébriques, au sein de laquelle la notion de discriminant joue un rôle privilégié. Son intérêt pour la théorie des nombres l'amène à construire une théorie de la divisibilité dans les anneaux d'entiers algébriques.

Ces anneaux ne sont pas toujours factoriels, et Kronecker cherche à rétablir leur caractère factoriel et même principal, par ce qu'il appelle lui-même l'« auxiliaire des coefficients indéterminés ». Il utilise ce même moyen en théorie de l'élimination, pour réduire le problème de la décomposition d'une variété en composantes irréductibles au simple problème de factorisation d'un polynôme. L'extension du concept de grandeur algébrique, racine d'une certaine équation qui la définit comme telle, appelle en effet immédiatement celui de variété algébrique, lieu des zéros d'un système d'équations. Dans la méthode d'élimination de Kronecker, le nombre d'inconnues devant être éliminées pour séparer une certaine composante irréductible est égal à le codimension, moins un, de cette composante. Ainsi se trouve défini le concept de codimension (*Stufe*). Kronecker applique cette méthode en théorie de Galois. Il faut, pour le comprendre, concevoir l'ensemble des racines de l'équation irréductible dont on étudie le groupe de Galois comme étant les coordonnées d'un point. L'orbite du point par le groupe de Galois de l'équation constitue alors une certaine variété de dimension 0. En théorie de Galois, Kronecker agit davantage en continuateur de l'œuvre d'Abel, s'intéressant plus « à la nature arithmétique des questions algébriques ». Enfin dans les dernières sections des *Grundzüge*, difficiles et moins connues, Kronecker tente, à la manière dont il avait transformé tout anneau d'entiers algébriques en un anneau principal, d'utiliser à nouveau l'auxiliaire des coefficients indéterminés pour transformer, en quelque sorte, toute variété en une intersection complète.

Dans le chapitre 16, après avoir brièvement évoqué le développement de la notion de résultant réduit à la suite d'un article de Cayley (1847), nous commentons une communication de A. Brill au congrès international de Heidelberg en 1904, sur l'« élimination et la géométrie dans les dernières décennies ». Brill jugeait que la méthode d'élimination de Kronecker servirait peut-être à démontrer de manière rigoureuse les principes de la géométrie énumérative. C'est alors pour nous l'occasion, en guise de conclusion, de réfléchir sur un apparent retrait de la théorie de l'élimination, à partir des années 1940–1950, justement dans des travaux de fondement de la géométrie algébrique, comme dans l'ouvrage d'A. Weil où celui-ci a prétendu « éliminer l'élimination ».

Première partie

L'élimination algébrique au XVII^{e} siècle

Chapitre 1

L'œuvre algébrique de Johannes Faulhaber

[1]L'historien des mathématiques Kästner décrivait Johannes Faulhaber comme étant « le plus célèbre des cossistes allemands du premier tiers du XVII^e siècle » [2]. L'intérêt de ce personnage est rehaussé pour nous par le fait qu'il se trouve à la charnière entre deux périodes de l'histoire des mathématiques, juste avant le renouveau causé par l'œuvre de Descartes. Cette position est aussi adoptée par le livre d'Ivo Schneider, *Johannes Faulhaber 1580–1635 : Rechenmeister in einer Welt des Umbruchs* (1993), qui décrit la confrontation, au début du XVII^e siècle, entre le milieu professionnel des *Rechenmeister* (auquel appartenait Faulhaber), et le groupe des « grands amateurs de mathématiques », tels Descartes et Viète. Ivo Schneider tente de déconstruire le mythe de la supériorité de l'un sur l'autre, mythe dont il cherche à voir la relation avec les récits d'une « rencontre » entre Faulhaber et Descartes, sous la plume des premiers biographes de Descartes dès le XVII^e siècle [3].

Dans ce chapitre, nous nous inspirons du travail d'Ivo Schneider tout en explorant un aspect de l'œuvre de Faulhaber peu mis en valeur par cet auteur. Faulhaber est célèbre pour ses travaux sur les sommes des séries de puissances, qui ont inspiré Jacques Bernoulli. Il est facile de rattacher cet intérêt pour l'arithmétique et l'analyse combinatoire aux spéculations mystiques sur les nombres pour lesquelles Faulhaber était aussi célèbre en son temps, puisqu'elles lui valurent un emprisonnement en 1606 (pour avoir prédit la fin du monde, comme l'avait fait Michael Stifel en 1533), puis un procès en 1618 (pour avoir prédit le passage d'une comète, qui pourtant eut bien lieu). Mais Faulhaber est moins connu pour son œuvre algébrique [4], qui n'est pas textuellement séparée de son œuvre arithmétique, et qui est obscurcie par la comparaison avec l'œuvre ultérieure de Descartes (celle-ci appartenant à une tradition bien différente, de géométrie algébrique). Nous reprenons donc dans ce chapitre l'analyse de trois de ses principaux ouvrages, dans le but d'éclaircir son rapport à l'algèbre. Le premier, l'*Arithmetischer Cubicossischer Lustgarten* (1604) est un livre constitué de 160 énoncés de problèmes, auquel un autre *Rechenmeister*, Peter Roth, jugea bon de répondre par une *Arithmetica Philosophica* publiée en 1608, principalement consacrée à la résolution de ces 160 problèmes [5]. Faulhaber fut mécontent de ce coup porté à ses intérêts financiers (il réclamait en effet un salaire des personnes désireuses d'apprendre de lui les résolutions de ses problèmes). Il

1. Ce chapitre reprend pour l'essentiel le contenu de notre article [248], rédigé à la suite d'un exposé au séminaire sur l'histoire de l'algèbre au Centre d'Histoire des Sciences et des Philosophies Arabes et Médiévales, dont le sujet nous avait été suggéré par Roshdi Rashed.

2. *Cf.* [185], vol. III, p. 29.

3. La possibilité d'une telle rencontre n'est cependant pas niée par Ivo Schneider, ni par les autres auteurs (Kurt Hawlitschek, Kenneth L. Manders) qui se sont penchés sur cette question. Nous aborderons dans les notes de bas de pages de ce chapitre quelques aspects de cette « rencontre ».

4. Voir toutefois l'article non encore publié de Kenneth L. Manders [225], dont le point de vue est assez différent du nôtre (*cf. infra* note 46).

5. Nous n'avons pas travaillé directement sur le *Lustgarten*. Les références ci-dessous renverront donc toujours à l'*Arithmetica Philosophica*, qui reproduit tous les énoncés des 160 problèmes.

tenta de récupérer le plus possible d'exemplaires imprimés de l'*Arithmetica Philosophica* pour éviter qu'ils ne circulent, et peut-être aussi pour s'en servir comme manuel d'enseignement dans son école à Ulm[6]. Faulhaber et Roth finirent par conclure un pacte : Faulhaber obtint un certain nombre d'exemplaires imprimés, et ils s'engagèrent tous deux à ne plus écrire d'algèbre pendant quelque temps. Roth mourut en 1617. A la suite de ce différend, Faulhaber se tourna davantage vers la « praxis mathématique », en particulier les techniques de fortification militaire, l'arpentage, la perspective, et les instruments de mathématiques. Cette spécialisation lui valut quelque succès pendant la Guerre de Trente Ans (1618–1648), mais ne l'empêcha pas de prolonger son œuvre arithmético-algébrique. C'est en 1619–1620 que se situe l'éventuelle rencontre avec Descartes. Les deux autres ouvrages dont il sera question ci-dessous sont les *Miracula Arithmetica* (1622) et l'*Academia Algebrae* (1631).

1.1 La « cubiccoß »

En 1631, dans son *Academia Algebrae*, son dernier traité d'algèbre, Faulhaber nous donne une vue rétrospective de son œuvre, sous la forme de recommandations suivant la préface et décrivant au lecteur les différentes étapes de l'initiation qu'il doit avoir reçue, en algèbre, avant de pouvoir aborder cet ouvrage :

> *Conseils rapides pour apprendre la Coß.*
> Si quelqu'un veut apprendre la Coß avec succès, rapidement et par les fondements, qu'il suive mon conseil. Qu'il acquière d'abord le Christoff Rudolff ou le Michael Stiffel, et y apprenne les principes et les espèces de la Coß, l'algorithme des binômes, des résidus et des nombres irrationnels[7], et ce qui s'y rapporte de semblable ; qu'il étudie ensuite dans de tels livres les première, seconde, troisième et quatrième règles de la Coß[8], c'est-à-dire la Coß linéaire ; maintenant, là où sont comparées simplement deux quantités l'une avec l'autre, il peut apprendre la Quadrat Coß, c'est-à-dire les 5ème, 6ème, 7ème et 8ème règles de Christoff Rudolph, dont j'ai résolu tous les exemples par la règle de fausse position. Maintenant, s'il veut passer à la Cubiccoß, il peut acquérir le Cardan, ou bien, puisqu'il préfèrerait un auteur allemand, qu'il s'essaie à mon *Cubicossischer Lustgarten* (dans lequel j'ai exposé les premiers exemples de toutes les règles successives - y compris de la Quadrat Coß - avec un soin particulier), tel que feu Petrus Roth l'a expliqué ; il y trouvera non seulement les 13 règles de Cardan, mais encore presque tout ce qui appartient à la Cubiccoß. J'avais certes voulu donner à imprimer avant lui [Petrus Roth] mon *Cubicossischer Lustgarten* résolu en abrégé. Mais parce que Petrus Roth a observé que cela pourrait porter atteinte à son propre ouvrage, l'*Arithmetica Philosophica*, il a passé contrat avec moi, par l'intermédiaire de mon ami fraternel, le mathématicien de Nuremberg Sebastian Kurz, pour que je suspende ce projet ; j'y ai alors consenti. Mais il pourrait encore peut-être arriver, si je publie entre-temps mon *General Opus Mathematicum* sur toutes les disciplines, que cela soit englobé dans cet ouvrage-ci ; si Dieu le veut.[9]

Si Faulhaber rappelle sa confrontation avec Roth, il ne semble pas en avoir gardé rancune, puisque c'est finalement à l'*Arithmetica Philosophica*, le traité de Roth, qu'il renvoie ici son lecteur

6. Le livre d'Ivo Schneider contient beaucoup de détails biographiques, rendus possibles par la correspondance entre Faulhaber et Kurz (un autre *Rechenmeister*, de Nürnberg), s'étendant sur la période 1604–1633, dont une grande partie est conservée à la Bibliothèque Nationale de France. Schneider dresse la liste des manuscrits en question (*cf.* [280], p. 239–245). Kurt Hawlitschek en donne des extraits importants dans [141].

7. Il s'agit de l'étude des irrationnelles euclidiennes et des formules visant à simplifier les expressions de la forme $\sqrt{a \pm \sqrt{b}}$ ou $\sqrt[3]{a \pm \sqrt{b}}$.

8. Les « huit règles de la Coß de Christoff Rudolff » enseignent respectivement la résolution des huit formes d'équations suivantes : (1) $ax = b$, (2) $ax^2 = b$, (3) $ax^3 = b$, (4) $ax^4 = b$, (5) $x^2 + bx = c$, (6) $x^2 + c = bx$, (7) $x^2 = bx + c$, (8) $x^4 + bx^2 = c$.

9. « Kurtz Bedencken von der Coß zu lehrnen », faisant suite à la préface de [111], et cité intégralement par Kästner, [185], vol. III, p. 143–144. Toutes les citations que nous ne citons pas dans leur langue originale sont traduites par nous, sauf indication contraire.

désireux d'apprendre la « Cubiccoß », c'est-à-dire la résolution des équations du troisième degré. C'est aussi par cet ouvrage qu'il nous faut aborder l'étude de l'œuvre algébrique de Faulhaber [10], ouvrage dont plus des trois quarts sont consacrés à la résolution des 160 problèmes du *Cubicossischer Lustgarten*. Roth et Faulhaber étaient tous deux *Rechenmeister* (l'un à Nürnberg, l'autre à Ulm), et l'*Arithmetica Philosophica* n'est postérieure que de 4 ans au *Cubicossischer Lustgarten* de Faulhaber. Ces deux ouvrages renvoient à une même période, à un même milieu professionnel et probablement aussi à des méthodes semblables.
Selon Ivo Schneider, l'*Arithmetica Philosophica* est l'un des tout premiers ouvrages à pleinement réaliser la « réception » des formules de Cardan en langue allemande [11].

La première partie de l'*Arithmetica Philosophica* comporte 13 chapitres, correspondant aux 13 formes canoniques de l'équation cubique. Bien que Roth ne le formule pas de manière générale, il se dégage facilement de son exposé que chaque forme canonique appartient à l'un des cas de figures suivants :
- une racine positive ;
- une racine positive et 0, 1 ou 2 racines négatives ;
- une racine négative et 0, 1 ou 2 racines positives ;
- 1, 2 ou 3 racines positives.

Pour les trois derniers cas, le nombre des racines dépend de la valeur du discriminant, que Roth indique pour les formes trinômes, auxquelles il ramène toutes les autres formes. Ces résultats étaient déjà sensibles dans l'œuvre de Cardan (*cf.* chapitre I de l'*Ars Magna*), et même, à quelques différences près, dans celle de 'Umar al-Khayyām ou de Sharaf al-Dīn al-Ṭūsī [12]). On peut résumer les résultats exposés par Roth en dressant le tableau suivant :

équation	racines positives	racines négatives
$x^3 + bx = c$	1	0
$x^3 + c = bx$	0, 1 ou 2	1
$x^3 = bx + c$	1	0, 1 ou 2
$x^3 + ax^2 = c$	1	0, 1 ou 2
$x^3 + c = ax^2$	0, 1 ou 2	1
$x^3 = ax^2 + c$	1	0
$x^3 + ax^2 + bx = c$	1	0, 1 ou 2
$x^3 + ax^2 + c = bx$	0, 1 ou 2	1
$x^3 + bx + c = ax^2$	0, 1 ou 2	1
$x^3 = ax^2 + bx + c$	1	0, 1 ou 2
$x^3 + ax^2 = bx + c$	1	0, 1 ou 2
$x^3 + bx = ax^2 + c$	3	0
$x^3 + c = ax^2 + bx$	0, 1 ou 2	1

Roth dit clairement que le nombre maximum de racines (réelles) d'une équation de degré n est n :

> Seynd in allen nachfolgenden Cossen auffs meinste so vil geltungen *radicis* zu finden / mit wieviel die höchste Quantitet der fürgegebenen Cossischen aequation / vermög der

10. L'*Arithmetica Philosophica* est composée de trois parties. La première (p. 1–18) est un exposé sur la résolution des équations cubiques. La deuxième (p. 19–174) consiste en la résolution détaillée des 160 problèmes de Faulhaber. La troisième (p. 175–192) propose 103 nouveaux problèmes, sans leurs résolutions, conduisant à des équations de degrés 4, 5, 6, 7. Cet ouvrage présente aussi un intérêt en rapport avec l'initiation mathématique du jeune Descartes. Celui-ci le mentionne en effet dans ses *Cogitationes Privatae*, *cf.* [90], vol. X, p. 242. D'ailleurs, les problèmes de la troisième partie ne sont pas sans rapport avec le contenu des *Progymnasta de solidorum elementis* ([91]) : Roth propose plusieurs problèmes mêlant des nombres figurés et le calcul des diamètres des sphères inscrites dans certains polyèdres irréguliers. Ce lien n'a pas encore été exploré, à notre connaissance.

11. *Cf.* [280], p. 61–66, en particulier p. 65 : « offenbar war man in Kreisen deutscher Rechenmeister erst um 1600 bereit, solche Radikalschachteln als sinnvolle Lösungen anzunehmen ». Ivo Schneider se base sur une comparaison des exposés respectifs de Cardan, Johann Jung (*Rechenmeister* à Lübeck, auteur d'un traité d'algèbre vers 1578 ; à ne pas confondre avec le savant plus connu Joachim Jung né à Lübeck en 1587 et mort à Hambourg en 1657), Michael Stifel et Peter Roth. Le « retard » dans la réception des formules de Cardan renvoie aussi à la question de la réception de l'algèbre elle-même, en langue allemande. Pour un article récent sur les origines de la tradition algébrique allemande, *cf.* Menso Folkerts [114].

12. *cf.* [315] pour Khayyām, et [157], p. 11–28 pour Ṭūsī. Ceux-ci ne prenaient cependant pas en compte les racines négatives, et n'indiquaient pas la possibilité de 3 racines positives.

Cossischen Progression / verzeichnet wird.

Il insiste particulièrement sur le fait qu'il ne s'agit que d'un maximum, pas toujours réalisé. Pour résoudre les équations cubiques, sauf dans le cas de trois racines réelles (cas « irréductible »), il applique les formules de Cardan. Pour le cas irréductible, il propose des méthodes numériques pour chercher les solutions entières. Il expose en détail ces méthodes pour les trois premières formes canoniques (équations trinômes sans terme en x^2). Pour les autres formes, il utilise des changements de variable afin de se ramener aux trois premières.

En fait, comme les problèmes traités dans l'*Arithmetica Philosophica* conduisent en général à des solutions rationnelles, le procédé numérique suffit dans beaucoup de cas, et est donc adopté de préférence à l'emploi des formules de Cardan. Roth enjoint en effet à rechercher d'abord les solutions entières, puis seulement, s'il n'en existe pas, à utiliser les formules de Cardan. Mais étant donné une équation à coefficients rationnels, Roth la transforme toujours en une équation à coefficients entiers, unitaire (c'est-à-dire dont le coefficient du terme de plus haut degré est 1) [13]. Les solutions rationnelles d'une telle équation sont entières, et se prêtent donc au procédé numérique en question. Décrivons rapidement l'ensemble des outils utilisés par Roth.

1.1.1 Les procédés numériques pour les trois premières formes canoniques

Les procédés numériques consistent simplement à tester l'équation pour toutes les valeurs entières de x dans un certain intervalle qui dépend de la forme canonique :

- pour $x^3+bx=c$, on teste toutes les valeurs entières de x dans l'intervalle $[1,\ \sqrt[3]{c}]$; le procédé donne l'unique racine positive, lorsqu'elle est rationnelle.
- pour $x^3=bx+c$, on teste toutes les valeurs entières de x dans l'intervalle $[\max(\sqrt[3]{c},\sqrt{b}),+\infty]$; le procédé donne l'unique racine positive, lorsqu'elle est rationnelle [14].
- pour $x^3+c=bx$, on teste toutes les valeurs entières de x dans l'intervalle $[1,\sqrt{b}]$; le procédé donne la ou les deux racines positives lorsqu'elles sont rationnelles [15].

Dans le premier et le troisième cas, Roth enseigne aussi par quelle extrémité de l'intervalle commencer, selon l'ordre de grandeur de $\frac{c}{b}$. Faulhaber connaissait un autre procédé numérique, attribué à Johann Jung [16], pour la recherche des solutions entières, généralisable aux équations de degrés plus élevés, et reposant sur un critère de divisibilité : ainsi dans la deuxième équation ci-dessus, toute solution x entière divise nécessairement c et $(b+\frac{c}{x})$. Roth ne parle pas de ce procédé.

1.1.2 Changements de variable

Pour ramener les formes quadrinômes et les formes trinômes contenant un terme en x^2 aux trois premières formes canoniques, Roth a recours à un changement de variable, que l'on verra ressurgir dans l'œuvre plus tardive de Faulhaber, de la forme $x=y-\frac{b}{3}$ (où b est le coefficient du terme en x^2 ; le signe variant suivant la forme canonique envisagée). Pour les formes canoniques $x^3+ax^2=c$ et $x^3+c=ax^2$, il propose, comme alternative, deux autres changements de variables, plus simples dans ces deux cas, de formes respectives $x=y^2-a$ et $x=\frac{(\sqrt[3]{c})^2}{y}$.

1.1.3 Trois règles de Stifel

Bien que les problèmes traités dans l'*Arithmetica Philosophica* conduisent en général à des solutions rationnelles, Roth utilise parfois les formules de Cardan. Ces formules contenant deux radicaux cubiques, il applique ensuite trois règles, attribuées à Michael Stifel, indiquant comment simplifier les radicaux cubiques [17]. Ces trois règles portent sur des radicaux des formes suivantes :

13. Il divise d'abord l'équation par le coefficient dominant pour obtenir une équation unitaire, puis fait un changement de variable de la forme $x=ay$ pour supprimer les fractions.
14. La valeur limite $x\geq\sqrt{b}$ s'obtient facilement à partir de la factorisation $(x^2-b)x=c$.
15. $x\leq\sqrt{b}$ s'obtient facilement à partir de la factorisation $(b-x^2)x=c$.
16. *Cf.* [280], pp. 65–66. Faulhaber oppose le procédé de Johann Jung (« Johann Jungen Invention welche Nicolaus Raimarus verbessern will ») à sa propre méthode de résolution algébrique dans [110], chap. 44, p. 70.
17. *Cf.* [151], pp. 22–24, pour un exposé de ces règles telles qu'elles figurent dans l'*Arithmetica Integra* de Stifel (1544).

- règle 1 : $\sqrt[3]{a+\sqrt{b}}$, $\sqrt[3]{a-\sqrt{b}}$
- règle 2 : $\sqrt[3]{\sqrt{a}+b}$, $\sqrt[3]{\sqrt{a}-b}$
- règle 3 : $\sqrt[3]{\sqrt{a}+\sqrt{b}}$, $\sqrt[3]{\sqrt{a}-\sqrt{b}}$

Ainsi, la première règle enseigne que lorsque $\sqrt[3]{a+\sqrt{b}} = p+\sqrt{q}$, q est un diviseur de b, tel que $\frac{b}{q}$ et $(\sqrt[3]{a^2-b}+q)$ soient des carrés. Alors $p = \sqrt{\sqrt[3]{a^2-b}+q}$.
La troisième règle enseigne que lorsque $\sqrt[3]{\sqrt{a}+\sqrt{b}} = \sqrt{p}+\sqrt{q}$, q est un diviseur de b, tel que $\frac{b}{q}$ soit un carré et que $\sqrt[3]{a-b}+q$ soit rationnel. Alors $p = \sqrt[3]{a-b}+q$.
En fait ces règles ne peuvent s'appliquer que lorsque les deux termes du binôme (ainsi $\sqrt{a}$ et $\sqrt{b}$ dans la troisième règle) sont vraiment incommensurables l'un par rapport à l'autre. A titre d'exemple, démontrons la règle 3 :

Démonstration de la règle 3. Si $\sqrt[3]{\sqrt{a}+\sqrt{b}} = \sqrt{p}+\sqrt{q}$, en mettant au cube on obtient $\sqrt{a}+\sqrt{b} = p\sqrt{p}+3p\sqrt{q}+3q\sqrt{p}+q\sqrt{q}$. Si l'on suppose alors les termes $\sqrt{a}$ et $\sqrt{b}$ incommensurables entre eux, on a (les lettres p et q étant interchangeables) :

$$\sqrt{a} = p\sqrt{p}+3q\sqrt{p}$$
$$\sqrt{b} = q\sqrt{q}+3p\sqrt{q}$$

D'où en mettant au carré :

$$a = (p+3q)^2 p$$
$$b = (q+3p)^2 q$$

d'où $a-b = (p-q)^3$ et $\sqrt[3]{a-b}+q = p$. D'où la formule ci-dessus, et les conditions nécessaires que q divise b et que $\frac{b}{q}$ soit un carré, *q. e. d.*

Dans ces trois règles, on peut procéder aussi bien avec des rationnels que des entiers, mais Roth s'efforce toujours d'éviter les fractions.

1.1.4 Relations entre les racines

Roth expose des procédés pour trouver les autres racines lorsque l'on en connaît déjà une, au moyen des coefficients de l'équation. Par exemple, pour la forme canonique $x^3+c = bx$, si x_1 est la racine négative, les deux racines positives sont $-\frac{x_1}{2} \pm \sqrt{b-3(\frac{x_1}{2})^2}$.

L'exposé de Roth a le mérite de l'exhaustivité. Les méthodes que l'on vient d'exposer lui suffisent pour résoudre tous les problèmes cubiques parmi les 160 problèmes posés par Faulhaber. Il cite l'*Ars Magna* de Cardan avec précision. D'autre part, il est intéressant de constater les choix différents faits par Roth et par Descartes (dans la *Géométrie*, 1637) pour contourner le même obstacle : le cas irréductible de l'équation cubique. Les deux auteurs enseignent en effet comment ramener toute équation cubique à l'une des trois premières formes canoniques, et ils renvoient pour sa solution aux formules de Cardan, sauf dans le cas irréductible [18]. Le moyen proposé par Descartes pour contourner l'obstacle est alors un moyen géométrique : la trisection de l'angle. Ainsi, pour les formes canoniques $z^3 = pz+q$ et $z^3 = pz-q$, lorsque $\frac{q^2}{4} \leq \frac{p^3}{27}$, Descartes ramène le problème à la trisection de l'angle sous-tendu par une corde de longueur $\frac{3q}{p}$ dans un cercle de rayon $\sqrt{\frac{p}{3}}$. Insistant sur le fait que, de son point de vue (le point de vue du géomètre, ici), ce procédé n'est pas moins légitime que l'utilisation des formules de Cardan, il propose même d'introduire une nouvelle notation algébrique pour le désigner : « quelque chiffre particulier, pour exprimer les subtendues (du tiers d'un angle), ainsi qu'on fait du chiffre $\sqrt{C}$ » [19]. Roth propose au contraire un procédé numérique, qui laisse

18. *Cf.* la *Géométrie*, [90], vol. VI, p. 471.
19. *Cf.* [90], vol. VI, p. 474. A l'occasion de ce rapprochement avec l'œuvre de Descartes, il convient de mentionner l'attention portée par Kenneth L. Manders à la remarque suivante, faite par un disciple de Faulhaber, Johannes Remmelin (1583–1632), dans un ouvrage publié en 1621 : « l'auteur [Faulhaber] a encore beaucoup d'autres secrets [...] en particulier quatre nouveaux compas proportionnels, dont l'un permet de trouver deux moyennes proportionnelles entre deux lignes. Comment d'autre part diviser tout arc de cercle en trois parties égales géométriquement. Comment

pressentir le lien étroit entre algèbre et arithmétique dans l'œuvre des *Rechenmeister*, et qui repose sur le fait que ceux-ci s'intéressent avant tout aux solutions entières, ou au moins rationnelles. On va voir à présent comment le lien entre ces deux disciplines se manifeste aussi dans le choix des 160 problèmes proposés par Faulhaber.

1.2 Thèmes représentés dans les 160 problèmes du *Cubicossischer Lustgarten*

1.2.1 Les problèmes traditionnels de la *Rechenkunst*

Parmi les 160 énoncés de problèmes du *Cubicossischer Lustgarten*, de nombreux sont déguisés sous forme de problèmes pratiques. On y trouve surtout des problèmes commerciaux, dont l'énoncé fait intervenir des crédits (par exemple le problème III), des intérêts composés (LXXXVIII), des parts de capital dans une société (CXIV), des salaires (I). La traduction de ces énoncés sous forme algébrique requiert tout d'abord le choix d'une ou plusieurs inconnues, puis une application (ou plusieurs) de la « règle de trois ». Roth est alors conduit à une équation. Le cœur du problème est la résolution de cette équation. Les solutions sont en général rationnelles. Parfois Roth utilise les formules de Cardan pour résoudre l'équation, et il doit donc traduire la solution sous forme rationnelle au moyen des trois règles de Stifel (par exemple, problème XCV). Parfois, il utilise son procédé numérique pour obtenir des solutions entières (ainsi dans le problème CXIV, où l'équation cubique obtenue appartient au cas irréductible).

Nous ne donnons ici que deux exemples, pour mettre en lumière l'utilisation de la règle de trois (une des nombreuses « règles » dont font état les *Rechenbücher*, à côté de la *regula falsi*, de la *regula alligationis*, etc.) et de la théorie des proportions dans la traduction des énoncés sous forme algébrique [20].

> Problème CXIV (*cf.* [271], p. 117v). Trois personnes forment une société. Ils placent ensemble 880 florins. A reste dans cette société pendant 8 mois, B pendant 6 mois, C pendant 5 mois. Ils gagnent au total 265 florins. Le capital placé et le gain de A s'élèvent à un total de 280 florins. Ceux de B à 390 florins. Ceux de C à 475 florins. Combien chacun avait-il placé, et combien chacun a-t-il gagné ? Réponse : A avait placé 200 florins, B 300 florins et C 380 florins. A a gagné 80, B 90 et C 95 florins.

Roth commence par simplifier l'énoncé en divisant toutes les quantités par 5. Ainsi capitaux et gains s'élèvent à : 56 pour A, 78 pour B, 95 pour C. Le gain total est 53. Il choisit comme inconnue le gain de A après 8 mois (on note cette inconnue x ci-dessous). Une première application de la règle

---

de même réaliser toutes les sections coniques et cylindriques au sujet desquelles d'autres auteurs ont écrit de gros livres. D'autre part il sait aussi démontrer, à partir du nombre 666, une Règle Générale pour une infinité de Coß [...] » (*cf.* [226], p. 2). Nous n'avons trouvé aucune autre trace de travaux sur les sections coniques dans les ouvrages de Faulhaber que nous avons consultés. D'autre part, la comparaison faite par Manders avec les compas décrits par Descartes dans ses *Cogitationes Privatae* souffre du fait que dans cette citation, les compas ne sont pas décrits comme des instruments permettant de résoudre des équations (comme ils le sont dans les *Cogitationes*), mais seulement comme des instruments de géométrie. La « Règle Générale pour une infinité de Coß » mentionnée ensuite par Remmelin est au contraire, on le verra ci-dessous, une méthode purement algébrique, proposée par Faulhaber en 1622 dans [110], pour résoudre les équations de tous degrés. Enfin, l'étude des compas par les géomètres et les algébristes arabes (*cf.* [263] et [246]), longtemps avant Descartes, rend inacceptable la remarque suivante de Manders : « the conic and cylindrical section compasses lack the degree of mathematical or practical interest which might otherwise account for independant discovery, or even for their retention in the set if its transmission from Descartes to Faulhaber was substantially indirect ».

20. L'importance de la règle de trois et de la théorie des proportions comme « trait d'union » entre arithmétique pratique et algèbre en Allemagne aux XV^{e} et XVI^{e} siècles est soulignée par Wolfgang Kaunzner dans [171], p. 164 et 166.

de trois lui permet d'établir le gain de A après les 6 premiers mois [21] :

$$\begin{array}{ccc} 8 \text{ mois} & x & 6 \text{ mois} \\ & \frac{3}{4}x & \end{array}$$

Le capital de A est alors : $56 - x + \frac{3}{4}x = 56 - \frac{1}{4}x$. Une deuxième application de la règle de trois lui permet d'établir le gain de B après les 6 premiers mois :

$$\begin{array}{ccc} 56 - \frac{1}{4}x & \frac{3}{4}x & 78 \\ & \frac{234x}{224-x} & \end{array}$$

Deux autres applications de la règle de trois lui permettent de même d'établir le gain de C après les 5 premiers mois :

$$\begin{array}{ccc} 8 \text{ mois} & x & 5 \text{ mois} \\ & \frac{5}{8}x & \end{array}$$

$$\begin{array}{ccc} 56 - \frac{3}{8}x & \frac{5}{8}x & 95 \\ & \frac{475x}{448-3x} & \end{array}$$

En égalant le total des trois gains à 53 florins, il obtient une équation cubique, appartenant au cas irréductible :

$$311584x - 2297x^2 + 3x^3 = 5318656 - 59360x + 159x^2$$

Il trouve une solution entière, $x = 16$, par le procédé numérique. *q. e. i.* [22]

> Problème LXXXVIII (p. 96v). Quelqu'un fait un emprunt de 1000 florins sur 3 ans, avec des intérêts composés. Si l'on multiplie les intérêts de la première année par les intérêts composés de la deuxième année, et que l'on multiplie ce produit par les intérêts composés de la troisième année, on obtient $144703\frac{1}{8}$ florins. Quels étaient les intérêts de la première année, pour ces 1000 florins ? Réponse : 50 florins.

Roth remarque d'abord que dans les problèmes d'intérêts composés, les intérêts successifs « croissent chaque fois proportionnellement ». Il énonce ensuite, verbalement, pour trois termes proportionnels, une formule que l'on écrirait :

$$\sqrt[3]{a \times ab \times ab^2} = ab$$

Il en déduit immédiatement que les intérêts composés de la deuxième année sont $\sqrt[3]{144703\frac{1}{8}} = 52\frac{1}{2}$. Il choisit alors pour inconnue x les intérêts de la première année, et applique la règle de trois pour en déduire les intérêts composés de la deuxième année en fonction de x :

$$\begin{array}{ccc} 1000 & x & 1000 + x \\ & \frac{1000x+x^2}{1000} & \end{array}$$

D'où l'équation :

$$52\frac{1}{2} = \frac{1000x + x^2}{1000}$$

qui a bien pour solution $x = 50$. *q. e. i.*

21. La règle de trois exprime que si $\frac{b}{a} = \frac{d}{c}$, alors $d = c \times \frac{b}{a}$. Roth dispose alors ces termes sous la forme :

$$\begin{array}{ccc} a & b & c \\ & c \times \frac{b}{a} & \end{array}$$

22. *Quod erit inveniendum.*

1.2.2 Arithmétique théorique

Un nombre important de problèmes font intervenir des nombres figurés [23]. Les problèmes VIII, IX, XXIII, XCVI à CXIII, CLVII et CLVIII portent spécifiquement sur les nombres polygonaux, et les problèmes CXXI à CXLVI, CLI à CLVII, et CLIX sur les nombres pyramidaux. Dans le problème VIII, Roth donne la formule suivante pour calculer un nombre polygonal à d arêtes de longueur x :

$$\begin{aligned} O_d(x) &= (\text{« premier terme »} + \text{« dernier terme »})\tfrac{x}{2} \\ &= [1 + (x(d-2) - d + 3)]\,\tfrac{x}{2} \end{aligned}$$

Le nombre polygonal $O_d(x)$ s'obtient ainsi comme somme d'une série arithmétique de raison $(d-2)$, dont le premier et le dernier termes sont respectivement 1 et $(x(d-2)-d+3)$ (dans la représentation géométrique du nombre polygonal, les termes de cette série correspondent à des « gnomons » ajoutés successivement à l'unité). De même, dans le problème CXXI, il donne une formule pour les nombres pyramidaux de paramètres d, x :

$$\sum_{k=1}^{x} O_d(k) = \frac{2 + (x(d-2) - d + 3)}{3} \cdot \frac{x(x+1)}{2}$$

Il donne ensuite une table de ces expressions algébriques [24], pour les valeurs de d de 3 à 100, accompagnée chacune du nom du nombre pyramidal en question, formé à partir de racines grecques (par exemple les nombres $O_{14}(x)$ sont dits « tétradécagonaux »). Le dernier terme de la série arithmétique servant à former un nombre polygonal, $(x(d-2) - d + 3)$, est appelé « racine d-gonale » (ainsi « racine tétradécagonale » pour $d = 14$). Le paramètre x est appelé « racine carrée » du nombre polygonal (« *quadrat Wurtzel* »).
Chacun de ces problèmes consiste en une équation faisant intervenir plusieurs nombres polygonaux et pyramidaux dont les paramètres (nombre d'arêtes, longueur d'une arête) présentent certaines relations mutuelles. Ainsi le CLVI consiste en l'équation :

$$\left(\sum O_{20}(x+9) - \sum O_{20}(x)\right) \times (O_3(x+9) - O_3(x)) = 1555281$$

Quelques autres problèmes attirent l'attention du lecteur sur des sujets variés d'arithmétique théorique. Les problèmes VI et CXVIII portent sur les nombres « proniques », de la forme $x(x+1)$ (un nombre pronique est le double d'un nombre triangulaire). Les problèmes LXIII et LXXI présentent des exemples de triangles numériques : le LXIII porte sur le triangle pythagoricien (15,20,25), et le LXXI sur le triangle (13,14,15), qui n'est pas pythagoricien mais qui a la particularité d'avoir une de ses hauteurs, et donc son aire, entière. Les problèmes CXII, CXIII et CLX proposent de chercher les deux paramètres d et x entiers d'un nombre polygonal donné N. Il s'agit donc de trouver une solution entière (d, x) de l'équation $O_d(x) = N$. Roth n'explique pas comment trouver cette solution, il indique seulement qu'il a utilisé des « tables » à cet effet. Il dit qu'il apprécie particulièrement les « tables » utilisées par Johann Jung. Il s'agissait peut-être de tables permettant la décomposition du nombre N en facteurs premiers [25]. Comme nous l'avons mentionné plus haut, un procédé pour trouver les solutions entières d'une équation cubique à une inconnue, reposant sur un critère de divisibilité, est aussi attribué à Johann Jung. Enfin, les problèmes CXLVII à CXLIX portent sur les sommes des séries de puissances $\sum x^2$ et $\sum x^3$. Roth renvoie en particulier, pour la formule $\sum x^3 = (\sum x)^2$, aux travaux de ses prédécesseurs Simon Jacob et Michael Stifel. On verra que cette formule a joué un rôle dans l'œuvre plus tardive de Faulhaber.

Le problème CXX demande la construction d'un carré magique 15×15. Il s'agit là d'un sujet classique, abordé par exemple par Stifel dans son *Arithmetica Integra* [26].

Quelques problèmes renvoient aussi spécifiquement à l'étude des irrationnelles, tel le problème LXVI : il s'agit de retrouver N dans l'équation $x^3 = 12x^2 - N$, sachant que la racine x est un

23. Pour plus de détail sur les nombres figurés chez Faulhaber, *cf.* Schneider [280], p. 72–80 et 109–122.
24. *Cf.* [271], p. 126r–129v.
25. N vaut, dans chacun des trois problèmes mentionnés, respectivement, 7^7, 41^5, 19^{11} ; une fois ces décompositions en facteurs premiers effectuées, le problème se résout facilement.
26. *Cf.* l'exposé détaillé de Hofmann, [151], p. 13–16.

« binôme » ou « résidu », dont la partie rationnelle est 5 (c'est-à-dire que x est de la forme $5 \pm \sqrt{b}$). Ce problème, avec les mêmes données numériques, était dans l'*Ars Magna* de Cardan [27].

1.2.3 Des problèmes d'algèbre à plusieurs inconnues

Un grand nombre de problèmes du *Cubicossischer Lustgarten* font intervenir deux inconnues, voire plus. L'énoncé de ces problèmes renvoie souvent directement à leur nature algébrique : ainsi des problèmes à deux inconnues qui commencent par « *Es seynd zwo Zahlen* » (« soient deux nombres... »), puis formulent une ou plusieurs conditions entre ces nombres. Parfois, l'une des conditions est donnée par le fait que les inconnues ont pour somme un nombre donné : « *mach auß 10 drey Theil* » (« fais trois parties de 10... »). Parfois aussi, les inconnues doivent former une suite géométrique : « *es seynd 4 Zahlen continuè proportionales* » (« soient 4 nombres continuellement proportionnels... »).

Les problèmes posés par Cardan dans le chapitre X de l'*Ars Magna* ont pu servir de modèles à Faulhaber : ainsi parmi les premiers problèmes de ce type dans le *Cubicossischer Lustgarten*, les problèmes XXXII, XXXIII, XXXIV sont identiques aux *quaestiones* II, III, IV du chapitre X de l'*Ars Magna* ([50], p. 29r).

Pour résoudre tous ces problèmes, Roth pose plusieurs inconnues symboliques. Il note l'une d'elles par les symboles cossiques (on la notera ci-dessous par x), et les autres par des majuscules latines A, B, C, etc. Dans les problèmes à deux inconnues, Roth pose souvent la première égale à $x + A$, et la deuxième égale à $x - A$, choix judicieux pour les problèmes dont l'équation résultant de l'élimination d'une inconnue est de degré 2 et qui conduit donc à une solution irrationnelle de la forme $a \pm \sqrt{b}$. Ce procédé remonte peut-être aux *Arithmétiques* de Diophante, et il apparaît pour la première fois en algèbre chez Abū Kāmil [28]. Lorsqu'un problème demande de partager un nombre N en trois parties, Roth pose souvent ces trois parties respectivement égales à x, A, $N - x - A$.

Dans le chapitre X de l'*Ars Magna*, Cardan distingue plusieurs formes canoniques d'équations à deux inconnues, de degré 2 :

$$\begin{aligned} x^2 &= ax + by \\ ax &= x^2 + by \\ by &= ax + x^2 \end{aligned}$$

$$\begin{aligned} xy &= ax + by \\ ax &= xy + by \end{aligned}$$

$$\begin{aligned} x^2 &= ax + cxy \\ xy &= ax + cx^2 \\ ax &= bxy + x^2 \end{aligned}$$

$$\begin{aligned} x^2 &= axy + by \\ xy &= ay + bx^2 \\ by &= ax^2 + xy \end{aligned}$$

De plus, pour chacune de ces formes, il suppose donné un des trois termes monômes qu'elle contient (ainsi, pour l'équation $x^2 = ax + cxy$, il suppose alternativement comme donné x^2, x, xy). Il donne ensuite, par une formule, la valeur de chacune des inconnues en fonction des coefficients et de la donnée supplémentaire. Cela suppose en fait l'élimination d'une des deux inconnues au sein d'un système de deux équations. Ainsi, pour l'équation $x^2 = ax + cxy$, lorsque xy est donné, il s'agit d'éliminer x ou y dans le système

$$\begin{cases} x^2 = ax + cxy \\ xy = d \end{cases}$$

Dans les huit premières formes canoniques [29], l'équation résultante est au plus de degré deux. Cardan

27. *Cf.* [50], chap. VI, § 2, p. 15r-15v.
28. D'après Roshdi Rashed, *cf.* [259].
29. *Cf.* [50], chap. X, § 1–8, p. 23r–26r.

« démontre » alors les formules qu'il énonce, à la manière dont les algébristes arabes démontraient les formules de résolution de l'équation de second degré à une inconnue, par « application d'aires ». Mais dans les trois dernières formes canoniques [30], l'équation résultante peut être du troisième degré. Dans ce cas, Cardan ne donne pas la formule (qui est parfois inutilisable, à cause du « cas irréductible »), mais seulement l'équation résultante.

Ce chapitre de l'*Ars magna* sera cité, d'abord par Stévin [31]. Alors que celui-ci s'apprête à traiter quelques problèmes à plusieurs inconnues, il écrit :

> Les exemples suiuans seront ceux la, ausquels se rencontrent postposees quantitez [=d'autres inconnues] : Mais il faut sçauoir que toute operation qui se faict par icelles, se peult aussi faire par positiues [=une seule inconnue] ; mais parce que la raison des nombres requis est aucunefois fort obscure, de sorte que pour l'absoluer par positiues quantitez, l'on auroit mestier de quelques theoremes, ou autres inductions, lesquelles souuentesfois ne nous viennent à la memoire, pourtant on les despesche pour le plus commode, par les postposees. [32]

Albert Girard sera lui aussi conscient de l'enjeu de faire progresser les méthodes de résolution des systèmes d'équations à deux inconnues présentes dans ce chapitre de l'*Ars Magna* lorsque, en 1629, il s'exclamera : « c'est ici où ils se sont arrestez [=Cardan et Stévin], mais nous achèverons et passerons au travers de ce nuage » [33]. Quant à Roth, il reproche à Cardan la facture théorique compliquée de ce chapitre :

> Sur cet exemple, et sur d'autres exemples avec les mêmes inconnues, Cardan a écrit, dans le dixième chapitre de son dixième livre, de nombreuses règles très difficiles et inintelligibles. Mais je veux te montrer ici, le mieux possible, que tous ces procédés peuvent s'accomplir au moyen de l'unique règle de trois, non seulement plaisamment et joliment, mais aussi de manière courte et judicieuse. [34]

C'est certainement aux « démonstrations » de Cardan que Roth adresse ce reproche. D'une part, Cardan ne proposait qu'une seule démonstration pour chaque forme canonique (quelle que soit la donnée supplémentaire formant la deuxième équation du système). Le lien entre la formule énoncée et sa démonstration n'était donc pas toujours très clair [35]. D'autre part, Roth n'utilise pas l'« application d'aires » [36]. Cette préférence apportée à la règle de trois est la manifestation d'une influence réciproque entre *Rechenkunst* (arithmétique pratique) et algèbre cossiste : la règle de trois, outil de la *Rechenkunst*, remplace la démonstration géométrique de l'algébriste italien Cardan.
Nous transcrivons ci-dessous deux exemples de problèmes à deux inconnues : le problème LXX (qui contient la critique à Cardan), et le problème LXXV (qui contient une équation à deux inconnues de degré 3, forme canonique absente du chapitre X de l'*Ars Magna*).

Problème LXX (p. 85r). Soit un exemple cossique conduisant à ces équations [37] :

$$x^2 = 2xA + 3A$$
$$xA = 8$$

Quelles sont les valeurs de x et A ? Réponse : $x = 1 + \sqrt{13}$ et $A = \sqrt{5\frac{7}{9}} - \frac{2}{3}$.

30. $x^2 = axy + by$, $xy = ay + bx^2$ et $by = ax^2 + xy$; *cf.* [50], chap. X, § 9–11, p. 26r–28r.
31. *Cf.* [293], p. 385–397
32. *Cf.* [293] p. 422 ; voir aussi déf. XXVIII p. 23–24, p. 379–380, p. 385–397 et p. 422–430 du même ouvrage, ainsi que notre citation d'Albert Girard *supra* p. 28.
33. [122], f. F4r et Gr.
34. *Cf.* [271], p. 85r.
35. Ainsi dans le cas de $x^2 = axy + by$, lorsque la donnée supplémentaire est xy. Le problème LXX est d'ailleurs identique à l'un des exemples numériques donnés par Cardan pour cette forme canonique.
36. Sauf dans l'unique problème XVIII, où il propose une première solution par application d'aires, conduisant à une solution irrationnelle compliquée, puis une deuxième solution, sans application d'aires, conduisant à la même solution, mais sous une forme plus simple.
37. Dans l'énoncé, Faulhaber note ces équations sous la forme : « 1z aequat. : 2 rad. A pl. 3 A. Facit 1 rad. A thut allhie 8 ».

Roth applique une première fois la règle de trois :

$$\begin{array}{ccc} xA & 8 & 2xA \\ & 16 & \end{array}$$

Il obtient $2xA = 16$. Il applique une deuxième fois cette règle :

$$\begin{array}{ccc} xA & 8 & 3A \\ & \frac{24}{x} & \end{array}$$

et obtient ainsi $3A = \frac{24}{x}$. D'où, en reportant ces valeurs dans la première équation de l'énoncé :

$$x^2 = \frac{24}{x} + 16$$

et finalement :

$$x^3 = 24 + 16x$$

Roth résout l'équation et trouve x, puis cherche la valeur de A en remplaçant x par sa valeur dans l'équation $xA = 8$. *q. e. i.*

> Problème LXXV (p. 88r). Cherche deux nombres, tels que le produit de leur différence par la différence de leur carrés soit 792, et que le produit de leur somme par la somme de leurs carrés soit 5720. Réponse : le plus grand est 14 et le plus petit 8.

Roth note ces deux nombres $A + x$ et $A - x$. En développant les produits décrits dans l'énoncé, il est conduit au système :

$$\begin{aligned} 8x^2A &= 792 \\ A^3 + x^2A &= 1430 \end{aligned}$$

c'est-à-dire :

$$\begin{aligned} x^2A &= 99 \\ A^3 + x^2A &= 1430 \end{aligned}$$

Roth écrit alors : « de ces deux équations, fais-en une seule, en les soustrayant ». D'où l'équation : $A^3 = 1331$, qui conduit directement à la solution. *q. e. i.*

Les 160 problèmes du *Cubicossischer Lustgarten* s'organisent donc autour de trois axes disciplinaires :
- un axe spécifiquement algébrique, centré sur la résolution des équations cubiques, mais incluant quelques exemples d'équations de degrés supérieurs [38], et aussi de nombreux exemples de systèmes d'équations à deux inconnues ou plus ;
- un rapport mutuel de l'algèbre à la *Rechenkunst*, qui se manifeste non seulement à travers le déguisement de problèmes algébriques sous forme d'énoncés de portée pratique (problèmes commerciaux, problèmes d'alliages, etc.), mais aussi à travers l'usage universel de la règle de trois ;
- un axe arithmétique, double : d'une part, l'arithmétique théorique fournit un grand nombre d'énoncés de problèmes en particulier autour des nombres figurés mais reste à un niveau élémentaire et ne sert que de prétexte à des énoncés en fait algébriques [39] ; d'autre part, l'étude des irrationnelles, en particulier les trois règles de Stifel, est rendu nécessaire pour simplifier la forme des solutions [40].

38. Ainsi les problèmes X, LXV, XCIII, XCIV, CLVII, CLVIII, CLX contiennent des équations de degrés 4 et 5. On reviendra plus loin sur la méthode proposée par Roth pour résoudre les équations de degré 4. Pour le degré 5, le problème CLX présente un exemple non trivial. L'équation en question étant à coefficients entiers, Roth est parvenu à trouver un facteur quadratique à coefficients entiers, mais il n'explique pas comment.

39. Pour les nombres polygonaux par exemple, la seule propriété arithmétique que devait connaître le lecteur pour résoudre les problèmes en question, est la représentation algébrique de ces nombres (que Roth obtient comme somme d'une série arithmétique dans le problème VIII). De même pour les nombres pyramidaux. Un autre exemple : le problème LXIII a pour solution un triplet pythagoricien (15,20,25), mais n'a en fait aucun intérêt arithmétique. En voici l'énoncé : « Soit un triangle orthogonal ABC. Soit AD une des hauteurs. On suppose que AB+BD vaut 36 et que AC+CD vaut 24. Question : que valent les segments AB, BD, DC, CA, AD, ainsi que l'aire du triangle ? ». Il s'agit simplement d'une application du théorème de Pythagore, conduisant à une équation quadratique en l'inconnue BD.

40. L'étude des irrationnelles apparaît aussi dans un cadre différent avec les problèmes LXVI (voir ci-dessus), LXVII, et CXIX.

Dans l'œuvre tardive de Faulhaber étudiée ci-après, le rapport entre arithmétique et algèbre va devenir prédominant.

1.3 L'œuvre algébrique de Faulhaber

Revenons à la courte citation faite au début de ce chapitre, les « Conseils rapides pour apprendre la Coß », donnés par Faulhaber en 1631 dans son *Academia Algebrae*. Nous nous étions arrêtés à la « Cubiccoß », et au récit de la confrontation avec Peter Roth. Faulhaber continue :

> Après cela, l'on peut maintenant apprendre la Zenßdezenßcoß, et d'autres équations plus élevées de Coß différentes ; ce que j'enseigne au moyen d'une voie générale dans mes *Arithmetica Miracula*, et que j'explique avec des exemples réguliers. Une fois que l'on a compris ces *Arithmetica miracula*, ainsi que d'autres écrits semblables, alors finalement l'on peut passer à cette *Academia Algebrae*, et y étudier le Procédé Général selon lequel doit être formée une infinité d'exemples de toutes les Coß successives ; tant qu'aucun mortel ne peut achever de les apprendre dans cette vie ; tant qu'il ne reste que Dieu, l'artisan le plus haut, qui les connaisse complètement. C'est aussi pour cela que je veux conduire le lecteur bénévole à cette Académie Algébrique, afin que, lorsqu'il aura tout compris suffisamment, il finisse pourtant par reconnaître avec moi que cette connaissance n'est qu'un recueil fragmentaire et que nous tous, hommes, dans cet art nous devons encore rester des écoliers, jusque dans notre cercueil, car plus l'on invente, plus il y a à apprendre, et plus à découvrir. Mais là-bas, dans cette vie-là, dans la véritable Académie céleste, nous parviendrons à une connaissance complète de cet art et d'autres arts. Que Dieu le Père, le Fils et le Saint Esprit nous aide. Amen. [41]

Faulhaber explique ici le rapport entre ses deux traités d'algèbre les plus importants : dans les *Miracula Arithmetica*, il prétend enseigner une méthode « générale » pour résoudre les équations algébriques de degré quelconque ; dans l'*Academia Algebrae*, il s'agit de donner à l'algèbre un champ d'application élargi, en proposant des exemples de problèmes conduisant à des équations de degré aussi élevé qu'on le désire.

1.3.1 Les *Miracula Arithmetica* : une méthode « générale et régulière » pour résoudre les équations de degré supérieur à 4. L'élimination chez Faulhaber.

Ce traité contient 50 « chapitres ». Les 30 premiers exposent et « démontrent » empiriquement une formule de récurrence sur les sommes itérées de puissances [42] :

$$ {}^{i+1}\sum n^k = \frac{n\left({}^{i}\sum n^k\right) - {}^{i}\sum n^{k+1}}{i} + {}^{i}\sum n^k $$

41. « Kurtz Bedencken von der Coß zu lehrnen », faisant suite à la préface de [111], et cité intégralement par Kästner, [185], vol. III, p. 143–144.

42. On notera dans la suite :

$$ {}^{i}\sum a_n = \sum_{k_i=1}^{n} \dots \sum_{k_1=1}^{k_2} a_{k_1} $$

La formule de récurrence énoncée est un cas particulier de la formule suivante, que l'on peut démontrer par récurrence sur $(n+i)$:

$$ {}^{i+1}\sum a_n = \frac{(n+i)\left({}^{i}\sum a_n\right) - {}^{i}\sum n a_n}{i} $$

Les chapitres 31 à 35 fournissent diverses propriétés des nombres icosaédraux, dodécaédraux[43], et « pyrgoïdaux »[44], et donnent le calcul des différences d'ordre quelconque de ces suites de nombres, et de leurs sommes itérées. Faulhaber utilise pour cela une formule que l'on pourrait écrire :

$$O(n) = a\binom{n+1}{3} + b\binom{n}{2} + n$$

où n est la longueur de l'arête, $O(n)$ le nombre polyédral. Les sommes itérées se déduisent alors aisément de la formule $\sum_{k=m}^{n} \binom{k}{m} = \binom{n+1}{m+1}$.

C'est au chapitre 36 que commence l'exposition de la « méthode générale » de résolution algébrique. Faulhaber se propose d'abord de représenter le nombre 666 (« *heilige Zahl* », nombre sacré) comme nombre polygonal. Il doit pour cela résoudre une équation quadratique de la forme $x^2 + bx = c$. Au moyen du changement de variable $x = y - \frac{b}{2}$, il se ramène à une équation de la forme $y^2 = c$. Au chapitre 37, il cherche à représenter 666 comme nombre polyédral, d'où une équation cubique cette fois. Au moyen d'un changement de variable $x = y - \frac{b}{3}$, il se ramène à une équation cubique sans terme en y^2, dont il donne immédiatement une solution (sans plus expliquer). Il remarque :

> Certes Cardan et Peter Roth ont déjà montré comment résoudre la Cubiccoß : c'est aussi ce qui provient de mon *Cubicossischer Lustgarten*. Mais ils ne l'ont pas tiré de la source inépuisable du nombre 666. Ainsi, ils ont séparé la Cubiccoß en 13 règles différentes. C'est pourquoi la découverte du procédé général leur est restée cachée.[45]

Ne laissons pas notre regard être obscurci par la spéculation mystique sur le nombre 666. Faulhaber entend bien substituer un procédé unique de résolution à la multiplicité des règles correspondant aux diverses formes canoniques. Certes Peter Roth utilisait déjà le changement de variable $x = y - \frac{b}{3}$ pour ramener les formes canoniques quadrinômes aux trois premières formes trinômes. Mais Faulhaber ne s'arrête pas à la « Cubiccoß ». Dans les chapitres 38 et 39, il utilise de même les changements de variables $x = y - \frac{b}{4}$, $x = y - \frac{b}{5}$, $x = y - \frac{b}{6}$, $x = y - \frac{b}{10}$ pour supprimer les termes en x^3, x^4, x^5, x^9 dans des équations de degrés respectifs 4, 5, 6, 10. Pour chacune, il donne la solution immédiatement après, sans plus expliquer comment il l'obtient.

Il désigne le procédé de changement de variable par le terme de « transfert » (*Transferierung*). Comme ce procédé fait intervenir le développement de binômes de la forme $(y - \frac{b}{n})^t$, il fournit à cette occasion une table contenant les coefficients binômiaux jusqu'à $t = 7$ (chapitre 36). Enfin, il remarque que les rares problèmes de degré 4 énoncés dans son *Cubicossischer Lustgarten* ne contiennent pas de terme en x^3, et il suggère que Peter Roth n'aurait pas su les résoudre, le cas échéant !

Dans le chapitre 40, Faulhaber énonce la règle (dite aujourd'hui « de Descartes ») déterminant le nombre de racines négatives ou positives d'une équation au moyen du nombre d'alternances des signes des coefficients successifs. Indiquons seulement un exemple :

$$\begin{array}{rccccccc} & & (0) & (0) & (1) & (2) & (3) & (0) \\ x^{10} = & & -5x^9 & -\frac{15}{2}x^8 & +7x^6 & -5x^4 & +\frac{3}{2}x^2 & +15743049850 \end{array}$$

Les nombres parenthésés indiquent le nombre d'alternances des signes, en partant de la gauche. Cette équation a donc au plus 3 racines positives, et 7 négatives.

Au chapitre 41, il explique comment, une racine a étant connue, l'on peut déterminer les autres racines en divisant l'équation par le binôme linéaire $(x - a)$; et réciproquement, comment obtenir une équation en multipliant les binômes linéaires correspondant à toutes ses racines[46].

43. Les nombres icosaédraux et dodécaédraux envisagés par Faulhaber sont des nombres figurés construits par « gnomons », bien que Faulhaber ne décrive pas explicitement ce procédé de construction. C'est la même conception des nombres polyédraux que l'on trouve dans les *Progymnasta de solidorum elementis* de Descartes (*cf.* [91]). D'autres auteurs utilisent une conception légèrement différente des nombres polyédraux (Maurolico).

44. Les nombres pyrgoïdaux sont des nombres polyédraux que l'on peut se représenter sous la forme d'un prisme surmonté d'une pyramide.

45. *Cf.* [110], p. 46.

46. Selon Ivo Schneider (*cf.* [280], p. 98), il pourrait s'agir aussi bien d'une découverte indépendante que d'une reprise des résultats de Viète sur ce sujet. D'autre part, à la fin de ce 41e chapitre, Faulhaber indique que certaines

Dans les chapitres 42 et 43, Faulhaber revient enfin sur la résolution d'une des équations de degré 4 dont il a expliqué le « *Transferierung* » dans le chapitre 38. Il fournit à cette occasion une méthode de résolution algébrique des équations de degré 4, identique à celle que l'on retrouve en 1637 dans la *Géométrie* de Descartes. Nous allons la comparer à la méthode utilisée par Peter Roth pour résoudre les quelques problèmes de degré 4 qu'il rencontre dans le *Cubicossischer Lustgarten*. Dans la description des deux méthodes, pour rendre l'exposé plus évident, nous choisirons des équations à coefficients littéraux, bien que tous les exemples traités par Roth et Faulhaber portent des coefficients numériques.

Méthode de Ferrari-Roth. Soit l'équation $x^4+bx+c=ax^2$ (ce sont des équations de cette forme que rencontre Roth dans les problèmes X et LXV du *Cubicossischer Lustgarten*). Roth introduit une nouvelle inconnue, que l'on notera y (il utilise cependant les mêmes notations, cossiques, que pour l'inconnue x, ce qui rend son exposé difficilement lisible). Il ajoute à chaque membre les termes $2yx^2+y^2$, ce qui donne l'équation :

$$x^4+2yx^2+y^2=2yx^2+ax^2-bx+y^2-c$$

Le membre de gauche est alors identique à un carré $(x^2+y)^2$. Pour faire de même à droite, Roth demande de faire « trois parts proportionnelles » avec les termes du membre de droite :

$$2yx^2+ax^2,\quad \tfrac{-bx}{2},\quad y^2-c$$

Le produit des extrêmes doit être rendu égal au terme moyen. Cela conduit à l'équation :

$$(2y+a)(y^2-c)=\left(\frac{-b}{2}\right)^2$$

Il s'agit d'une équation cubique en y. Roth cherche ses trois racines. Pour chacune de ses racines y_i, il remarque que l'équation initiale peut s'écrire :

$$(x^2+y_i)^2=(x\sqrt{2y_i+a}-\sqrt{y_i^2-c})^2$$

Le problème revient donc à résoudre les équations quadratiques [47] :

$$\pm(x^2+y_i)=x\sqrt{2y_i+a}-\sqrt{y_i^2-c}$$

équations peuvent être résolues au moyen des principes qu'il vient d'énoncer, mais que d'autres requièrent l'utilisation de certaines « tables ». Il ajoute qu'il ne veut pas publier ces tables parce qu'un de ses amis, « Herr Carolus Zolindius (Polybius) » prévoit de les publier bientôt à Venise ou à Paris. Ivo Schneider, Kurt Hawlitschek et Kenneth L. Manders fournissent plusieurs indices visant à identifier ce personnage à Descartes. Voici en effet un extrait des *Cogitationes Privatae* de Descartes : « Trésor mathématique de Polybe le Cosmopolite. On y donne les vrais moyens de résoudre toutes les difficultés de cette science ; on y démontre que, sur ces difficultés, l'esprit humain ne peut rien trouver de plus ; ceci pour secouer la paresse et condamner la témérité de certains, qui promettent de montrer, dans toutes les sciences, de nouveaux miracles ; et aussi pour alléger les travaux torturants de beaucoup, qui, pris nuit et jour dans certains nœuds gordiens de cette science, consument inutilement l'huile de leur esprit. L'ouvrage est offert de nouveau aux érudits du monde entier, et spécialement aux Frères Rose-Croix, très célèbres en Allemagne. » (*cf.* [90], vol. X, p. 214 ; on donne ici la traduction de Foucher de Careil). Les circonstances dans lesquelles Descartes a noté ces lignes ne sont pas connues. Cela pourrait faire l'office d'un titre de livre. Si l'on regarde à présent les titres des ouvrages de Faulhaber [109], [110], [111] (ces titres forment une suite), on est tenté de voir dans le texte de Descartes une certaine dérision à l'égard de ces titres. Toutefois, les auteurs mentionnés n'ont donné aucune preuve certaine de l'identification Polybius–Descartes. Kenneth L. Manders [225] voit dans les *Miracula Arithmetica* le fruit d'un travail commun entre Faulhaber et le jeune Descartes. Il analyse de manière détaillée la théorie des équations présente dans les *Miracula Arithmetica* de Faulhaber ; il montre en particulier que dans cet ouvrage et dans la théorie des équations de la troisième partie de la *Géométrie* de Descartes est présente une même « conception multiplicative des polynômes et des équations », dont il trouve l'origine dans les listes de problèmes de degrés supérieurs à 3 de l'*Arithmetica Philosophica* de Roth. Mais ce ne sont pas seulement les doutes possibles quand à l'authenticité de la rencontre Faulhaber–Descartes qui nous démarquent du point de vue adopté par Manders : nous essayons en effet de montrer la spécificité de l'œuvre algébrique de Faulhaber.

47. Roth néglige l'indétermination de signe $\pm$ dans le problème X, mais il la fait apparaître dans le problème LXV. Par contre, dans le problème LXV, il ne considère qu'une seule racine de la réduite en y.

Parmi les racines de ces équations, certaines sont identiques.

Méthode de Faulhaber–Descartes. Partons de la même équation initiale $x^4 + bx + c = ax^2$. Faulhaber introduit trois nouvelles inconnues, que nous noterons y, A, B (comme lui, sauf pour l'inconnue y qu'il note avec les mêmes symboles cossiques que la première inconnue x). Il propose d'identifier cette équation au produit $(-x^2 + xy + A)(x^2 + xy - B)$. Le développement de ce produit est $-x^4 + (y^2 + A + B)x^2 + (yA - yB)x - AB$. En l'identifiant terme à terme avec $-x^4 - bx - c + ax^2$, il obtient les trois équations suivantes :

$$(*) \qquad \begin{cases} y^2 + A + B = a \\ yB - yA = b \\ AB = c \end{cases}$$

On reconnaît ici la méthode dite aujourd'hui « méthode des coefficients indéterminés », que l'on attribue souvent à Descartes. Faulhaber résout le système linéaire en A, B formé par les deux premières équations et il obtient :

$$\begin{cases} 2B = \frac{b}{y} - y^2 + a \\ 2A = -y^2 + a - \frac{b}{y} \end{cases}$$

Il reporte ces valeurs dans $AB = c$ et obtient ainsi une équation de degré 6 en une seule inconnue y, qui ne contient que des termes carrés, et qu'il résout donc comme une équation cubique, au moyen du changement de variable $z = y^2$. Il reporte la valeur d'une de ses racines dans les expressions de A et B ci-dessus, et obtient ainsi une solution (y, A, B) du système ci-dessus, et donc la factorisation souhaitée $-x^4 + (y^2 + A + B)x^2 + (yA - yB)x - AB = (-x^2 + xy + A)(x^2 + xy - B)$. Le problème revient donc à égaler à zéro un facteur puis l'autre. Faulhaber utilise les mots « *partes aliquotae* » pour désigner ces deux facteurs quadratiques, mettant ainsi à profit un terme appartenant à la tradition arithmétique et désignant les facteurs premiers d'un nombre entier.

En comparant ces deux méthodes, on comprend l'intérêt apporté par Faulhaber au changement de variable $x = y - \frac{b}{4}$, et sa remarque quant à l'incapacité de Roth de traiter des équations de degré 4 contenant un terme en x^3. Leurs deux méthodes supposent en effet l'absence de terme en x^3. Dans la méthode de Faulhaber, on pourrait certes contourner la difficulté en introduisant une inconnue supplémentaire et en prenant comme « *partes aliquotae* » $(-x^2 + xy + A)$ et $(x^2 + xC - B)$ si l'équation avait un terme en x^3 ; mais la réduite de degré 6 obtenue contiendrait alors éventuellement des termes de degrés impairs et ne se ramènerait pas aussi facilement à une équation cubique [48].

Au chapitre 44, couronnement de ce traité, Faulhaber envisage la généralisation de sa méthode de recherche des « partes aliquotae » à des équations de degré plus élevé :

> Par ce procédé, il faut reconnaître que pour des Coß plus élevées, les *partes aliquotae* pourront encore être trouvées, puis résolues de manière régulière, bien qu'une telle chose puisse paraître impossible à la raison humaine ; car le fait que nous, êtres humains mortels, ne puissions résoudre toutes les Coß de manière régulière (certes, je ne parle pas là de la découverte de Johann Jung que Nicolaus Raimarus a amélioré), cela n'est pas de la faute de l'art [*die Kunst*], mais plutôt de notre ignorance et de notre faiblesse. Toutefois, si l'on veut procéder de la manière ci-dessus pour des Coß plus élevées, on devra poser y, A, B, C, D et ainsi de suite. Si l'on vient me dire : comment est-ce possible de résoudre de telles équations ? Je répondrai que cet art a beaucoup progressé ces derniers temps en langue allemande, plus qu'en toute autre langue. Qu'on consulte l'écrit de Sutor *Exemplum Arithmeticum*. [49]

48. Il faudrait alors, pour se débarasser des termes de degrés impairs, appliquer justement à la réduite l'opération « *Transferierung* ».
49. *cf.* [110], p. 70.

On voit apparaître ici un thème récurrent dans l'histoire de l'élimination algébrique : la difficulté des calculs d'élimination. Le fait que l'on ne soit pas encore parvenu à résoudre les équations de degrés supérieurs à 4 ne serait pas dû à l'algèbre (« l'art », *die Kunst*). L'algèbre ne manque pas de moyens : Faulhaber expose, à la suite de cela, une méthode d'élimination. C'est donc à la faiblesse humaine qu'il faut s'en prendre, incapable de mener le calcul. Pourtant la méthode proposée par Faulhaber est encore imparfaite. Si l'on en croit la citation ci-dessus, il s'agit d'une méthode nouvelle, ou du moins, récemment découverte[50]. L'*Exemplum Arithmeticum* contient simplement l'énoncé d'un problème[51], conduisant à un système de deux équations à deux inconnues que Faulhaber se propose de résoudre :

$$\left\{\begin{array}{ll}(L_1) & x^7 = -15x^5a^2 - 15x^3a^4 - xa^6 + 6240\\ (L_2) & 3x^5a^2 = -10x^3a^4 - 3xa^6 + 1008\end{array}\right.$$

Il faut bien remarquer en effet que toute la difficulté d'une tentative de généralisation de la méthode de recherche des « *partes aliquotae* » tient à deux choses : il faut d'une part savoir résoudre le système $(*)$, c'est-à-dire trouver la réduite en y, et d'autre part, espérer que cette réduite puisse se résoudre plus facilement que l'équation initiale (pour le degré 4, on se ramène à une réduite de degré 3). Faulhaber ne fait pas mention de cette deuxième difficulté. Mais pour la première, il propose sur l'unique exemple ci-dessus une méthode de résolution des systèmes de plusieurs équations à plusieurs inconnues, qui se veut assez générale. Pour les puissances de x, il utilise les notations cossiques. Pour les puissances de a, il utilise les notations a, aa, aaa, etc. Dans l'*Academia Algebrae*, en 1631, Faulhaber précise que les deux équations (L_1) et (L_2) proviennent en fait des deux suivantes :

$$\left\{\begin{array}{l}((x+a)^6 + (x-a)^6)((x+a) + (x-a)) = 24960\\ ((x+a)^6 - (x-a)^6)((x+a) - (x-a)) = 8064\end{array}\right.$$

Cet exemple ressemble beaucoup au problème LXXV du *Cubicossischer Lustgarten* (dont nous avons transcrit la résolution, *supra* p. 29) qui commence par les équations :

$$\left\{\begin{array}{l}((A+x)^2 + (A-x)^2)((A+x) + (A-x)) = 5720\\ ((A+x)^2 - (A-x)^2)((A+x) - (A-x)) = 792\end{array}\right.$$

Ce problème-ci semble avoir eu quelque succès : on le retrouve dans le commentaire de la *Géométrie* de Descartes par Schooten (*cf.* [89], *Commentarii in Librum I*, GGG). Quant au système d'équations (L_1), (L_2), Faulhaber passe par les étapes suivantes pour le résoudre :

$$\begin{array}{ll}L_3 \longleftarrow & L_1 - 5L_2\\ L_4 \longleftarrow & 3L_3\\ L_5 \longleftarrow & 14L_2 + L_4\\ (L_5) & 42x^5a^2 + 35x^3a^4 = 17712 - 3x^7\\ L_6 \longleftarrow & 5x^2L_3 - 2a^2L_5\\ L_7 \longleftarrow & \frac{13}{32}x^2L_5 - \frac{5}{32}L_6\\ (L_7) & 18x^7a^2 - 5535a^2 = 8133x^2 - 2x^9\end{array}$$

(L_7) donne a^2 et a^4 :

$$\begin{array}{l}a^2 = \frac{8133x^2 - 2x^9}{18x^7 - 5535}\\ a^4 = \left(\frac{8133x^2 - 2x^9}{18x^7 - 5535}\right)^2\end{array}$$

Dans (L_5), substituons à a^2 et a^4 leur valeur :

$$x^{21} = 10114056x^7 - 2154x^{14} - 1356572043$$

Cette méthode, qui se veut générale, est imparfaite : Faulhaber ne nous dit pas, à chaque étape, quelles équations choisir, ni quels monômes. Le caractère itératif de cette méthode reviendra dans les

50. « dise Kunst zu unserer letsten zeit in Teutscher Sprach vil höher gestigen / weder man sonsten in keiner Sprach zeigen kan. »

51. D'après Kenneth L. Manders ; nous n'avons pu consulté directement ce texte.

méthodes ultérieures (comme celle de Fermat, *cf. infra*). Faulhaber procède par étapes, en formant à chaque étape une nouvelle équation comme somme des produits de deux équations antérieures par des monômes. Il se montre satisfait, étant parvenu à l'équation L_7, car, dit-il, il a à présent deux « petites équations », L_5 et L_7, au lieu des deux équations initiales. Bien plus, L_7 permet d'obtenir directement la valeur de a^2 en fonction de x. En substituant cette valeur dans L_5, il obtient une équation en x, qu'il résout, puis il reporte la valeur de x obtenue dans l'équation L_5, qui lui donne une équation en a, qu'il résout.

Les derniers chapitres du traité sont moins importants pour notre propos : le chapitre 45 contient une généralisation à l'espace du théorème de Pythagore et de la formule d'Héron pour l'aire du triangle [52] ; les chapitres 46 à 50 présentent divers problèmes faisant intervenir des nombres figurés à paramètres non entiers [53].

1.3.2 L'*Academia Algebrae* : les formules des séries de puissances comme exemples fondamentaux de formules algébriques

L'*Academia Algebrae* a pour objet principal des formules du type :

$$\begin{array}{ll}(1) & \sum n^{2k} = (\sum n^2).P(\sum n) \\ (2) & \sum n^{2k+1} = (\sum n^3).P(\sum n)\end{array}$$

et plus généralement :

$$\begin{array}{ll}(3) & {}^t\sum n^{2k} = ({}^t\sum n^2).P(\frac{n(n+t)}{2}) \\ (4) & {}^t\sum n^{2k+1} = ({}^t\sum n).P(\frac{n(n+t)}{2})\end{array}$$

où P désigne un polynôme, différent dans chacune des formules ci-dessus. Faulhaber ne démontre pas ces formules. Il commence par poser dix « questions » [54]. Chacune porte sur la somme (ou la somme double) d'une série de puissances : il s'agit de trouver le polynôme P dans l'une des formules ci-dessus, puis de développer cette formule pour obtenir une expression $Q(n)$, polynomiale en n, de

52. *Généralisation du théorème de Pythagore* : si les aires des faces d'un tétraèdre rectangle sont notées A, B, C et D, D étant l'aire de la face opposée à l'angle droit, alors $A^2 + B^2 + C^2 = D^2$.
Quant à la généralisation de la formule d'Héron, il s'agit de trouver le volume d'un tétraèdre de côtés donnés. Faulhaber n'énonce pas explicitement la formule, mais il prétend la connaître et donne quelques indications, insuffisantes, sur la méthode pour y parvenir. Ivo Schneider tente de reconstituer sa démarche (*cf.* [280], pp. 125–127.).
Descartes énonce lui aussi la généralisation du théorème de Pythagore et la démontre dans ses *Cogitationes Privatae* ; il propose même de l'étendre à 4 dimensions. Il y énonce aussi une formule d'Héron pour les tétraèdres rectangles.
Les exemples numériques choisis par Descartes diffèrent de celui choisi par Faulhaber (celui-ci fait une fois de plus intervenir le nombre 666). Costabel remarque certes que Descartes applique la formule d'Héron plane à un exemple numérique extrait de la *Geometria Practica* de Clavius (*cf.* [80], p. 642 ; remarquons en passant que dans l'article [80], Costabel attribue improprement à Faulhaber l'ouvrage anonyme *Numerus Figuratus, sive Arithmetica Analytica Arte Mirabili Inaudita Nova Constans...* qui est en fait de son disciple Johannes Remmelin, comme le montre Ivo Schneider, [280], p. 113). Mais les notations utilisées par Descartes pour démontrer la généralisation de la formule de Pythagore ne sont pas les notations cossiques utilisées dans le reste de ses *Cogitationes*. Il utilise en effet le signe *qq* pour le carré de l'inconnue. Comme le remarque Charles Adam (*cf.* [90], vol. X, p. 247), cette notation figure aussi, conjointement aux notations cossiques, dans l'*Algebra* de Clavius ; nous n'avons pas vu la notation *qq* dans l'œuvre de Faulhaber. Ces passages des *Cogitationes* confirment donc plus l'influence de Clavius dans l'initiation mathématique de Descartes qu'ils ne nous renseignent de manière précise sur l'influence de Faulhaber.
Par contre une transmission de Descartes à Faulhaber n'est pas à exclure ici : le caractère déplacé de ces deux résultats géométriques au sein du traité arithmético-algébrique de Faulhaber renvoie à la possibilité d'une influence extérieure sur celui-ci. S'ils s'étaient rencontrés dans les années 1619–1620, le jeune Descartes pourrait très bien avoir interrogé Faulhaber sur ces résultats. Celui-ci (qui les ignorait peut-être), inspiré par leur élégance, les aurait inclus ensuite dans son traité de 1622.
Il nous faut encore critiquer le rapprochement fait par Kenneth L. Manders entre ces passages des *Cogitationes* et les *Progymnasta de solidorum elementis*. Manders cite un manuscrit de Matthäus Beger (1588–1661, disciple de Faulhaber), sur la notion d'angle solide trièdre. Costabel a pourtant montré (*cf.* [91]) que le mérite majeur des *Progymnasta de solidorum elementis* est la généralisation de la dualité courbure–capacité au cas d'un angle solide d'un nombre quelconque de faces (la cas trièdre étant particulièrement simple).

53. Par exemple, Faulhaber montre que le nombre sacré 1335 est un nombre pyramidal ayant pour nombre de côtés $2 + \frac{9}{20}$ et pour longueur d'arête 24. Bien entendu, un tel nombre figuré, n'ayant pas un nombre entier de côtés, ne se prête plus à aucune représentation géométrique.

54. *Cf.* [111], pp. *Bi v* à *Ciii r*.

la somme de la série en question, et enfin de résoudre une équation de la forme $Q(n) = c$ où c est une constante entière donnée.

Après avoir posé ces dix questions, Faulhaber décrit un procédé [55] permettant de calculer les coefficients du polynôme P pour $\sum n^{2k+1}$ lorsque sont connus ceux pour $\sum n^{2k}$. Ivo Schneider en déduit que Faulhaber devait avoir une formule de récurrence lui permettant d'obtenir directement les polynômes P pour tous les exposants $2k$ pairs, et une autre pour les exposants impairs. Nous rappelons ici brièvement la démarche proposée par Ivo Schneider pour l'établissement de telles formules de récurrence, en prenant pour modèle la démarche menant à la formule classique [56] :

$$\sum n^3 = \sum_{\nu=1}^{n} \left[\binom{\nu+1}{2}^2 - \binom{\nu}{2}^2 \right] = \binom{n+1}{2}^2$$

Au lieu de la somme indicée par ν dans la formule ci-dessus, Ivo Schneider considère plus généralement :

$$\sum_{\nu=1}^{n} \left[\binom{\nu+1}{2}^s - \binom{\nu}{2}^s \right]$$

pour s quelconque, afin d'obtenir une formule reliant les $\sum n^{2k+1}$ pour $(2k+1) \leq (s+1)$. Pour obtenir une formule reliant les $\sum n^{2k}$, il utilise plutôt la somme :

$$\sum_{\nu=1}^{n} \left\{ \left[\binom{\nu+s}{2s+1} + \binom{\nu+s+1}{2s+1} \right] - \left[\binom{\nu+s-1}{2s+1} + \binom{\nu+s}{2s+1} \right] \right\}$$

Enfin, pour obtenir une formule reliant les sommes itérées ${}^t\sum n^{2k}$, il propose d'utiliser :

$$\sum_{\nu=1}^{n} \left\{ \left[\binom{\nu+s+t-1}{2s+t} + \binom{\nu+s+t}{2s+t} \right] - \left[\binom{\nu+s+t-2}{2s+t-1} + \binom{\nu+s+t-1}{2s+t-1} \right] \right\}$$

Il renvoie l'idée de ces formules de récurrence à un article de L. Tits [57] publié en 1923.

Mais revenons à l'*Academia Algebrae*. Faulhaber propose lui-même une méthode pour calculer les polynômes P dans les formules (1) et (2). Il indique d'abord comment calculer directement une expression polynomiale en n pour les sommes et les sommes itérées des séries de puissances $\sum n^k$. Il remarque en effet [58] que la suite des puissances n^k a pour différence k-ième une suite arithmétique de différence constante $(k!)$. Alors on peut trouver, pour la suite n^k et pour ses sommes itérées, des formules analogues à celles données pour les nombres icosaédraux et dodécaédraux dans la *Miracula Arithmetica*, de la forme :

$$n^k = k! \binom{n+k-1}{k} + c_{k-1} \binom{n+k-2}{k-1} + \ldots + c_2 \binom{n+1}{2} + c_1 n$$

Faulhaber rappelle d'autre part que les expressions algébriques des sommes et des sommes itérées des séries $\sum n^k$ peuvent encore s'obtenir par récurrence au moyen de la formule énoncée au tout début des *Miracula Arithmetica*. Il propose alors, pour obtenir les polynômes P intervenant dans les formules énoncées plus haut, d'appliquer la méthode des coefficients indéterminés, déjà formulée et utilisée dans les *Miracula Arithmetica* pour la recherche des *partes aliquotae*. Certes il n'entre pas beaucoup dans les détails :

> De la même manière que l'on peut non seulement chercher les *partes aliquotas* et la valeur des racines dans les équations cossiques élevées, mais aussi trouver l'équation lorsque l'on connaît toutes ses racines, ici aussi, on peut non seulement trouver les

55. *Cf.* [111], p. *Ciiii v.*
56. La formule $\sum n^3 = (\sum n)^2$ était déjà connue d'al-Karajī.
57. *Cf.* [309].
58. *Cf.* [111], p. *Dii v.*

> quantités cossiques [=coefficients des puissances de x] à partir des multiplicateurs [=coefficients du polynôme P], mais aussi, inversement, trouver les multiplicateurs à partir de ces quantités algébriques, pour des quantités paires et impaires [59].

Il énonce dans les pages suivantes, sans les démontrer, les formules (3) et (4) pour des valeurs particulières de t et de k : ${}^{2}\sum n^4$, ${}^{3}\sum n^4$, ${}^{4}\sum n^4$, ${}^{6}\sum n^4$, ${}^{2}\sum n^5$, ${}^{3}\sum n^5$, ${}^{2}\sum n^6$, ${}^{3}\sum n^6$, ${}^{4}\sum n^6$, ${}^{2}\sum n^7$, ${}^{2}\sum n^8$.

La partie arithmétique de l'*Academia Algebrae* s'arrête là. Il s'agit maintenant pour Faulhaber de revenir sur sa méthode de recherche des *partes aliquotae*. Les dix questions posées au début du traité contenaient en effet chacune une équation algébrique à résoudre. Faulhaber compose alors un *Exposé approfondi sur les résolutions régulières et véritables d'un nombre infini de coss* [60]. Il y insiste particulièrement sur le changement de variable $y = x - \frac{b}{n}$ pour faire disparaître le terme en x^{n-1} dans une équation de degré n, procédé que selon lui, ni Cardan, ni Stifel, ni Johann Jung ne connaissaient. Il donne de nouveau la table des coefficients binomiaux [61], cette fois jusqu'à $n = 18$, et il indique que cette table peut s'étendre facilement, soit par multiplications, soit par additions « à partir de l'angle supérieur ». Il ajoute :

> L'utilisation et l'intérêt de cette table pour résoudre un nombre infini de coss sont expliqués et rendus suffisamment évidents dans mes *Miracula Arithmetica*, puisque les équations cossiques élevées, par cette nouvelle découverte inédite, subissent une préparation, afin d'être disposées à la Résolution Régulière, et afin que l'on puisse trouver et observer les *Partes Aliquotas*. L'art entier est livré et révélé, à travers divers exemples, avec beaucoup de zèle, dans le traité déjà cité. [62]

Il rappelle ensuite la théorie de l'élimination ébauchée dans les *Miracula Arithmetica* autour du problème de Sutor, et il annonce :

> Celui à qui il reste du temps, il pourrait former des centaines d'exemples semblables ; il pourrait aussi en diriger pour 3 nombres, 4 nombres, ou plus (dont les différences ou les rapports ne sont pas connus) [63] on doit cependant donner comme connues autant d'équations différentes qu'il y a de nombres, si l'on veut ne pas avoir plusieurs *Facit* [64].Mais de ces questions, de celles de Peter Roth, et aussi de nombreuses autres découvertes, peut-être qu'il en sera dit plus, en un autre lieu, si Dieu le veut. [65]

Après cet exposé sur la résolution algébrique, Faulhaber revient rapidement sur les nombres polyédraux [66]. Il donne des formules algébriques pour les sommes des nombres polyédraux correspondant aux cinq polyèdres réguliers (comme il l'avait fait pour l'icosaèdre et le dodécaèdre en 1622).

Il énonce ensuite rapidement un problème géométrique : trouver le diamètre de la sphère circonscrite à un assemblage de quatre sphères se touchant mutuellement, de diamètres respectifs 24, 23, 29 et 22.

Puis il propose une « *Wortrechnung* » : problème algébrique dont les solutions, entières, indiquent des lettres dans un alphabet, et forment ainsi un mot (en l'occurence, le nom du jour où Faulhaber aurait terminé son *Academia Algebrae*). Faulhaber montre comment l'énoncé de ce problème se ramène à une équation de degré 5, qu'il laisse le soin au lecteur de résoudre « suivant l'art de la Sursolit Coß » qu'il a « suffisamment enseigné dans les *Miracula Arithmetica* ».

Enfin vient un « appendice » [67], proposant une autre *Wortrechnung*, plus difficile. Les nombres indiquant les lettres du mot cherché font cette fois intervenir les coefficients de certaines puissances

59. *Cf.* [111], p. *Diii* ; « pour des quantités paires et impaires », c'est-à-dire dans le cas des sommes de puissances d'exposants pairs $\sum n^{2k}$ ou d'exposants impairs $\sum n^{2k+1}$.
60. *Cf.* [111], p. *Ei v.*
61. *Cf.* [111], p. *Eii v.*
62. *Cf.* [111], p. *Eii v.*
63. C'est-à-dire 3 inconnues, 4 inconnues, ou plus.
64. C'est-à-dire si l'on veut qu'il n'y ait qu'une seule solution, ou en d'autres termes, que le problème soit « déterminé ».
65. *Cf.* [111], p. *Eiii r.*
66. *Cf.* [111], p. *Eiii r.*
67. *Cf.* [111], p. *Eiii v.*

de x dans les expressions des sommes ${}^9\sum x^8$ et $\sum x^{25}$, ainsi que les coefficients des polynômes P dans les formules donnant $\sum x^{22}$, $\sum x^{23}$, $\sum x^{24}$ et $\sum x^{25}$.
Faulhaber propose ensuite un problème qu'il a déjà traité dans un de ses ouvrages de *praxis*, la division d'une corde d'un instrument de musique en 12 parties[68]. Il ne résout pas le problème mais affirme qu'il conduit à une équation de degré 12, résoluble au moyen des logarithmes de Neper (il s'agit d'extraire une racine douzième).
Finalement, il propose trois dernières questions[69], sans les résoudre, faisant intervenir les sommes ${}^t\sum x^5$ pour $t = 1$, 2, 4, 6, 8, 10 et 12 ; ${}^t\sum x^{13}$, ${}^t\sum x^{14}$ et ${}^t\sum x^{15}$ pour $t = 1$ et 2 ; $\sum x^{18}$, $\sum x^{24}$ et $\sum x^{27}$.

1.4 Conclusion

La méthode de résolution algébrique des équations proposée par Faulhaber sera attribuée par la suite à Descartes (par Hudde, *cf. infra* p. 43). Il aurait donc fallu que ce soit un géomètre qui découvre un procédé plus proche de la nature (ici, arithmétique) des équations ? C'est bien en effet une inspiration arithmétique qui semble gouverner cette méthode, comme l'indique le nom que lui donne Faulhaber, dès 1622, dans ses *Miracula Arithmetica* : la « recherche des *partes aliquotae* ». D'ailleurs, quelques années après la publication de sa *Géométrie*, Descartes semble lui-même avoir oublié cette méthode. Lorsque Dozem l'interroge à ce sujet, Descartes répond[70]en exposant la méthode de Ferrari, et non la sienne.

Faulhaber propose une règle unique pour résoudre une équation de degré n quelconque : supprimer le terme en x^{n-1}, puis chercher les *partes aliquotae* au moyen de techniques d'élimination algébrique dans un système d'équations à plusieurs inconnues. Dans certains cas, cette méthode ne peut remplacer des procédés numériques (comme l'indique ses allusions à l'usage de « tables »), mais ce n'est pas faute de moyens algébriques, c'est « à cause de notre ignorance ». Il s'agit là d'un schéma qui se répétera jusqu'au début du XIX$^{\text{e}}$ siècle. Ainsi Leibniz croyait pouvoir ramener la résolution de l'équation générale de degré 5 à celle d'un système linéaire de 284 équations : « Le calcul n'est ni laborieux, ni difficile, ni prolixe. »[71]. Tschirnhaus propose en 1683 de remplacer, dans la résolution de l'équation cubique, le changement de variable $y = x + a$ (pour supprimer le terme en x^2) par un autre de la forme $y = x^2 + bx + a$, afin de faire disparaître d'un seul coup les termes en x et en x^2. Il pense que sa méthode peut s'étendre à tous les degrés[72]. En 1749, c'est le théorème fondamental de l'algèbre qu'Euler tente de démontrer par la recherche des *partes aliquotae*, c'est-à-dire en décomposant toute équation en produit de facteurs quadratiques[73]. Lagrange remettra sur pied cette démonstration en 1772, en mettant à profit les techniques qu'il développe dans un mémoire de 1770–1771, les *Réflexions sur la résolution algébrique des équations*[74], dans lequel il tente de rendre apparente une continuité conceptuelle entre les différentes méthodes de résolution algébrique de ses prédécesseurs, tout en élaborant des théorèmes puissants qui seront repris au XIX$^{\text{e}}$ siècle par Gauss, Abel, Galois.

Pour mener à bien son projet, Faulhaber fut obligé d'envisager une méthode systématique de résolution des systèmes d'équations polynomiales à plusieurs inconnues, qu'il illustra au moyen de l'« exemple de Sutor ». Une idée semblable se trouve à la base des travaux les plus importants sur l'élimination algébrique. L'idée de Faulhaber se résume à numéroter les équations successives L_i qu'il obtient par le calcul (en commençant par les deux premières, données), et à appliquer à chaque étape une opération de la forme :

$$L_{i+1} = (\text{monôme})L_i + (\text{monôme})L_j$$

68. *Cf.* [111], p. *F i r.*
69. *Cf.* [111], p. *F ii.*
70. Lettre du 25 mars 1642, *cf.* [90], vol. III, p. 553–554.
71. Leibniz, mai 1678, cité par Knobloch [175], p. 148.
72. *Cf. infra* p. 75.
73. *Cf. infra* note 17 p. 77.
74. *Cf. infra* chapitre 9.

Bézout proposera, en 1779, de représenter l'équation résultante d'un système d'équations polynomiales $\{f_1 = 0, f_2 = 0, f_3 = 0, ...\}$ sous la forme $f_1\phi_1 + f_2\phi_2 + f_3\phi_3 + ... = 0$, où ϕ_1, ϕ_2, ϕ_3,... sont des polynômes adéquats [75].

Malgré les allusions de Descartes à la « méthode des coefficients indéterminés » [76] où à l'élimination algébrique des inconnues [77], la méthode de Faulhaber, si éloignée de toute intuition géométrique (comme le sera celle de Bézout, qui, dans un traité de 400 pages sur les équations, fait à peine allusion à leur représentation géométrique), est le fruit d'une tradition bien différente de celle que poursuit Descartes dans sa *Géométrie* [78].

Il faut aussi rattacher l'œuvre de Faulhaber aux travaux ultérieurs de Pascal (sur le triangle arithmétique et les séries de puissances), Bernoulli [79] (sur les nombres de Bernoulli), Newton [80] (sur les sommes itérées de séries arithmétiques). Mais ce qui a pu obscurcir l'appréciation des chapitres d'algèbre (36 à 44) des *Miracula Arithmetica* ce n'est pas seulement la comparaison avec Descartes, c'est aussi la prédominance, au sein du même ouvrage, de formules arithmétiques sur les séries de puissances, donc de considérations qui n'ont pas de rapport nécessaire à l'algèbre [81], puis de leur « démonstration » par induction empirique ; c'est la juxtaposition dans un même texte de deux rationalités pour nous différentes. C'est pourtant cette proximité disciplinaire, de fait, entre arithmétique et algèbre, qui permet à Faulhaber, dans son *Academia Algebrae*, de justifier l'extension de l'algèbre à des équations de degré supérieur à 3 par l'existence de problèmes arithmétiques conduisant à de telles équations (ainsi les formules de sommation des séries de puissances). Cependant ce rapport particulier, qui n'apparaît peut-être que comme justification *a posteriori*, et de manière explicite seulement dans son dernier traité (1631), ne doit pas détourner notre regard de la tendance générale qui a poussé de nombreux algébristes, à appliquer leur science à d'autres disciplines au sein de leurs traités d'algèbre [82].

75. *Cf. infra* p. 139.

76. *Cf.* [90], vol. VI, p. 423 : « l'invention de supposer deux équations de même forme, pour comparer tous les termes de l'une à ceux de l'autre,..., n'est pas l'une des moindres de la méthode dont je me sers ».

77. *Cf.* [90], vol. VI, p. 373 : « il se faut servir par ordre de chacune des équations... et faire ainsi, en les démêlant, qu'il n'en demeure qu'une seule ».

78. Pour deux analyses sur la corrélation entre équations algébriques et courbes géométriques au sein de la *Géométrie*, irréductible à un terme unique, *cf.* [37] et la préface à [315].

79. *Cf.* [18].

80. *Cf.* Newton, *Methodus Differentialis*, [238], vol. VIII, p. 236–257.

81. Nous ne parlons pas ici des nombres triangulaires, qui présentent un lien naturel évident entre arithmétique et algèbre puisqu'ils fournissent le développement du binôme.

82. Par exemple, Roshdi Rashed suggère que ceci a permis aux algébristes, à partir de l'œuvre d'Abū Kāmil, de « fonder l'analyse diophantienne rationnelle comme partie de l'algèbre », *cf.* [260], p. 39 ; les meilleurs traités d'algèbres, d'Al-Karajī à Euler, contiennent un chapitre d'analyse diophantienne rationnelle.

Chapitre 2

Hudde, Newton, Leibniz

2.1 Hudde et la méthode des coefficients indéterminés

Le *De reductione aequationum* de Jan Hudde (1628–1704), paru dans l'édition latine (1659) de la *Géométrie* de Descartes, est en fait la traduction latine par Frans van Schooten d'une lettre de Hudde, datée de juillet 1657. Dès le début de cette lettre, Hudde expose son objet : l'étude de la réduction des équations dans l'absolu (*absolute*) et non relativement (*relative*) à un problème dont elles tirent leur origine. Il se limitera aux équations polynomiales normalisées (dont le coefficient dominant est 1) à une seule inconnue [1]. La lettre est une suite de 22 « règles », qui donnent toutes soit des critères d'irréductibilité, soit des méthodes pour la recherche de facteurs dans une équation ou un polynôme donné. Pour chaque groupe de règles, Hudde précise à quel domaine doit appartenir le polynôme $P(x)$ considéré [2], que les signes radicaux et fractions numériques et littérales soient proscrites ($P \in \mathbb{Z}[a,b,c,...][x]$), ou que les fractions numériques soient autorisées ($P \in \mathbb{Q}[a,b,c,...][x]$) – voire que les fractions littérales ($P \in \mathbb{Q}(a,b,c,...)[x]$) le soient aussi, et même les radicaux ($P \in K[x]$, où K est une extension algébrique de $\mathbb{Q}(a,b,c,...)$). En fait, Hudde considère même des équations où les radicaux contiennent l'inconnue x dans certains exemples [3]. Pourtant, Hudde ne précise jamais relativement à quel domaine se rapportent ses notions de réductibilité ou d'irréductibilité. On chercherait en vain une définition de ces notions dans la lettre. Un polynôme est dit « réductible » quand il est divisible par un facteur de degré (en x) strictement inférieur, mais sans préciser à quel domaine appartient ce facteur. Ainsi p. 410, Hudde dit que $x^3-3x^2+10x-3$ « n'est pas réductible », alors que les formules de Cardan (que Hudde connaissait bien) permettent d'exhiber un facteur linéaire (certes contenant deux radicaux cubiques ; cette équation n'appartient pas au « cas irréductible »). Au contraire p. 427–428, Hudde enseigne comment « réduire » le polynôme $x^4-2ax^3+(2a^2-c^2)x^2-2a^3x+a^4$, « divisible » par $x^2-(a+\sqrt{a^2 \pm c^2})x+a^2$. En fait, la notion d'irréductibilité sera seulement précisée au début du XIX$^{\text{e}}$ siècle dans les travaux de Gauss [4], d'Abel [5], puis de Galois [6]. Enfin, Dedekind,

1. Hudde savait que ses règles s'appliquent aussi bien à des équations (polynomiales) qu'à des polynômes (*quantitates compositae*) : *cf.* par exemple [89], note 1, p. 411, ou bien p. 422 : « considero enim quantitates haud secus atque aequationes, supponendo sc : illas =0 ».

2. *Cf.* [89], p. 408, 429, 458 et 461.

3. *Cf.* [89], p. 411, $P \in \mathbb{Q}(a,b,\sqrt{3a^2+b^2})[x,\sqrt{x^2+a^2-b^2}]$; p .416, $P \in \mathbb{Q}(a,c,\sqrt{3c^2+a^2})[x,\sqrt{x^2+a^2}]$. Hudde propose de définir le degré (*dimensiones*) d'une équation contenant des radicaux enveloppant l'inconnue comme étant le degré de l'équation résultant de l'élimination des radicaux. Il donne d'ailleurs, p. 429–433, une méthode d'élimination des radicaux différente de celle de Fermat.

4. Gauss démontre rigoureusement que l'équation $x^{n-1}+x^{n-2}+...+x+1=0$ n'a pas de facteur rationnel dans [120], § 341.

5. *Cf.* [2], t. I, p. 479 : « Une équation $\phi x = 0$, dont les coefficiens sont des fonctions rationnelles d'un certain nombre de quantités connues a, b, c,... s'appelle *irréductible*, lorsqu'il est impossible d'exprimer aucune de ses racines par une équation moins élevée, dont les coefficiens sont également des fonctions rationnelles de a, b, c,... ».

6. Qui précise systématiquement, grâce à sa théorie de l'« adjonction », relativement à quel corps de base une équation est dite réductible ou non. *Cf.* la première définition de son « premier mémoire », [119] p. 45 : « Une équation est dite réductible quand elle admet des diviseurs rationnels ; irréductible dans le cas contraire. (...) on pourra convenir de regarder comme rationnelle toute fonction rationnelle d'un certain nombre de quantités déterminées, supposées

Weber, et Kronecker, en 1882, fonderont une théorie de la divisibilité pouvant aussi s'appliquer dans des domaines de « fonctions entières », telles que les expressions de Hudde contenant $\sqrt{x^2+a^2}$. Malgré le caractère peu rigoureux [7] de la lettre de Hudde causé par l'absence de définitions et, le plus souvent, de démonstrations, ce texte semble pourtant marquer une étape importante dans un mouvement d'arithmétisation de l'algèbre commencé à la fin du X$^{\text{e}}$ siècle avec Al-Karajī [8]. Dans la lignée d'Al-Karajī, on peut citer Al-Samaw'al (mort en 1175) qui appliqua l'algorithme de division euclidienne aux polynômes pour obtenir ce que nous appellerions aujourd'hui le « développement limité » d'une fraction rationnelle dans l'anneau $\mathbb{Q}[x, \frac{1}{x}]$. Au XVI$^{\text{e}}$ siècle, inspiré par une tentative de Pedro Nuñez, Simon Stévin avait proposé, dans son *Arithmétique* (1585), le premier à notre connaissance, d'appliquer aux polynômes (en une indéterminée) l'algorithme arithmétique euclidien du plus grand commun diviseur [9]. Stévin utilisait cet algorithme afin de simplifier une équation de la forme

$$P(x) = Q(x)$$

en divisant chaque membre par le plus grand commun diviseur de P et Q. Nous allons en voir, chez Hudde, un usage différent et nouveau.

2.1.1 L'algorithme euclidien du plus grand commun diviseur chez Hudde

Regardons la règle 5 du *De reductione aequationum* :

> *Règle 5, qui enseigne la manière de réduire toute équation, qui peut être produite par la multiplication de deux autres, dont l'une contient une lettre qui n'est pas contenue dans l'autre.*
> Supposons une lettre = 0 ; cherchons si l'équation qui en résulte a un diviseur commun avec l'équation proposée. Si elle n'en a pas, supposons une autre lettre = 0, cherchons si cette Résultante a un diviseur commun : et ainsi de suite, jusqu'à ce qu'il y ait un diviseur commun, ou bien qu'il n'y ait plus d'autre lettre, qui n'eût pas été supposée = 0. [10]

C'est à l'occasion de cette règle 5, qui requiert la recherche d'un diviseur commun à deux équations, que Hudde expose l'algorithme euclidien, sur les deux exemples suivants :

$$(1)\begin{cases} d^3 - ad^2 + 2a^2b - 2abd = 0 \\ d^4 - b^2d^2 + a^2b^2 - a^2d^2 = 0 \end{cases}$$

$$(2)\begin{cases} 12a^4 + 11a^2x^2 + x^4 - 4ax^3 - 20a^3x = 0 \\ 12a^2x^2 - 3ax^3 + 24a^4 - 16a^3x + x^4 = 0 \end{cases}$$

Dans l'exemple (2), les deux polynômes appartiennent à l'anneau euclidien $\mathbb{Q}(a)[x]$ et se prêtent donc sans problème à l'application de cet algorithme. Notons respectivement (L_1) et (L_2) les deux équations proposées. Hudde obtient une valeur de x^4 au moyen de (L_1), qu'il substitue dans (L_2), et obtient ainsi :

$$(L_3) \quad ax^3 + a^2x^2 + 4a^3x + 12a^4 = 0$$

Il en déduit une valeur de x^3, qu'il substitue dans (L_1), en commençant par les plus hautes puissances de x, pour obtenir :

$$(L_4) \quad x^2 = ax - 6a^2$$

Puis il substitue cette valeur de x^2 dans (L_3) et trouve une équation identiquement nulle. Il en conclut que $x^2 - ax + 6a^2$ est le plus grand commun diviseur des deux équations proposées.

---

connues a priori. (...) Nous dirons que ces quantités sont ADJOINTES à l'équation ».

7. renforcé par le fait que les règles 3, 4, 6 et 7 sont fausses ([89] p. 414, 418 et 421).
8. Sur les débuts de ce mouvement d'arithmétisation de l'algèbre, *cf.* [258], [7] et [16].
9. *Cf.* [293], problème LIII.
10. *Cf.* [89], p. 420. On peut rendre cette règle rigoureuse en se limitant à la divisibilité d'un polynôme P dans $\mathbb{Q}(a, b, c, ...)[x]$, à condition que les dénominateurs des coefficients de P ne s'annulent pas lorsque l'on spécialise les lettres a, b, c,... en les posant égales à 0. Hudde était attentif à ce problème des dénominateurs, comme l'indique sa note 2 p. 412.

Pour l'exemple (1) au contraire, Hudde doit s'éloigner de son projet initial (se limiter aux équations normalisées à une seule inconnue x). Il étudie la divisibilité dans l'anneau $\mathbb{Q}[a, b, d]$, qui n'est pas euclidien. Il choisit pour cela de réécrire les deux polynômes comme polynômes normalisés en l'une des lettres a, b ou d ; ainsi par exemple pour la lettre d, en travaillant dans $\mathbb{Q}(a, b)[d]$, il trouve $(d - a)$ comme plus grand commun diviseur. En travaillant ensuite dans $\mathbb{Q}(b, d)[a]$, il retrouve le même résultat. Par contre, il remarque que si l'on choisit la lettre b, en travaillant dans $\mathbb{Q}(a, d)[b]$, les deux équations normalisées n'ont pas de diviseur commun, mais que le facteur $(d - a)$ apparaît alors au dénominateur des coefficients de ces équations :

$$\begin{cases} b = \frac{d^3 - ad^2}{2ad - 2a^2} \\ b^2 = \frac{d^4 - a^2 d^2}{d^2 - a^2} \end{cases}$$

2.1.2 Méthode des coefficients indéterminés : l'analyse cartésienne selon Hudde

En fait, Hudde ne limite pas l'usage de l'algorithme euclidien à cette théorie de la divisibilité naissante. Il y voit aussi une méthode de résolution des systèmes d'équations polynomiales, et il est en cela le premier à notre connaissance. Ce thème est important dans la lettre car Hudde est conduit à utiliser la méthode dite aujourd'hui « des coefficients indéterminés », pour la recherche de facteurs d'une équation à une inconnue [11]. Or cette méthode requiert justement la résolution d'un système d'équations polynomiales.

Etudions plus particulièrement le cas de la règle 19 « qui enseigne la manière de réduire toute équation rationnelle sans fraction ni second terme, qui puisse être divisée par une autre dont le second terme est rationnel » [12]. Soit un polynôme de degré 6 sans second terme :

$$x^6 + qx^4 + rx^3 + sx^2 + tx + v$$

Supposons, comme Hudde, que ce polynôme est décomposable en produit d'un facteur de degré 4 et d'un facteur de degré 2, à coefficients indéterminés y, z, k, l, w :

$$x^6 + qx^4 + rx^3 + sx^2 + tx + v = (x^4 - yx^3 + zx^2 + kx + l)(x^2 + yx + w)$$

En développant le produit et en identifiant terme à terme les coefficients des différentes puissances de x (c'est précisément cela que nous désignons par l'expression moderne « méthode des coefficients indéterminés »), on obtient le système d'équations polynomiales :

$$\begin{cases} (1) & z - y^2 + w = q \\ (2) & k + yz - wy = r \\ (3) & l + yk + wz = s \\ (4) & yl + wk = t \\ (5) & wl = v \end{cases}$$

En éliminant successivement z, k, l et w, Hudde est conduit à une équation ne contenant plus que l'inconnue y, de degré 15.

Hudde utilise aussi ce procédé pour décomposer l'équation générale de degré 4 sans second terme en produit de deux facteurs quadratiques et il est conduit à une équation de degré 6 en y, résoluble comme une équation cubique. Il retrouve ainsi la règle de résolution algébrique de l'équation de degré 4 donnée par Descartes dans sa *Géométrie* (mais que Descartes ne démontre pas par la méthode des coefficients indéterminés). Après avoir ainsi « restitué » la démonstration de la règle de Descartes, Hudde écrit :

> Il faut remarquer que cette Règle, par laquelle peuvent être réduites toutes les équations réductibles quadrato-quadratiques, est tout-à-fait la même que celle que Descartes décrit

11. Dans les règles 9, 11, 17, 18, 19 et 20.
12. *Cf.* [89], p. 486.

> p. 79, 80 et 81 de sa *Géométrie*. Et je ne puis douter, qu'il ne l'ait trouvée de la même manière, ou en tout cas pas très différemment ; surtout si ce qu'il enseigne p. 84 de général sur la Réduction des équations, coïncide avec sa méthode des sécantes, et ce qu'il a exposé ensuite p. 49. Il ne paraît pas vraisemblable, à voir l'élégante cohérence entre ce qui précède et ce qui suit, qu'il l'ait prise à d'autres auteurs, comme certains le croient. [13]

Suit un long éloge de Descartes, accompagné d'une réflexion sur les Anciens et les Modernes, que Jean Prestet traduira en français dans la préface au tome 2 de ses *Nouveaux éléments de mathématiques* en 1689. Hudde conclut cet éloge ainsi :

> Je suis certain, qu'il ne pourra être vraisemblable, ni à toi ni à moi, qu'Il ait emprunté cette Règle de Réduction à des textes écrits par d'autres, plutôt qu'il ne l'ait extraite et découverte de ses propres germes de toutes les sciences, fondamentaux et très féconds.

Hudde affirme donc que c'est par cette méthode que Descartes a dû découvrir sa règle de résolution algébrique, et il attribue finalement la méthode des coefficients indéterminés à Descartes ; plus encore, il l'interprète comme étant une manifestation de l'unité de la méthode cartésienne. Le seul endroit de la *Géométrie* où Descartes fait intervenir explicitement la méthode des coefficients indéterminés est dans son exposé de la « méthode des sécantes » [14]. Il est donc naturel que Hudde justifie son interprétation par des références précises. Il renvoie à deux allusions faites par Descartes. La première suit immédiatement l'exposé de la méthode des sécantes :

> Mais je veux bien, en passant, vous avertir que l'invention de supposer deux equations de mesme forme, pour comparer separement tous les termes de l'une a ceux de l'autre, et ainsi en faire naistre plusieurs d'une seule, dont vous avés vû icy un exemple, peut servir a une infinité d'autres Problesmes et n'est pas l'une des moindres de la methode dont je me sers. [15]

La deuxième allusion suit immédiatement l'exposé de la règle de résolution algébrique des équations de degré 4 :

> Je pourrais aussy en adjouster d'autres [règles] pour les equations qui montent jusques au sursolide [$= x^5$], ou au quarré de cube [$= x^6$], ou au delà ; mais j'ayme mieux les comprendre toutes en une, & dire en general que, lorsqu'on a tasché de les reduire a mesme forme que celles, d'autant de dimensions, qui vienent de la multiplication de deux autres qui en ont moins, et qu'ayant dénombré tous les moyens par lesquels cete multiplication est possible, la chose n'a pû succeder par aucun, on doit s'assurer qu'elles ne sçauroient estre reduites a de plus simples. [16]

Le langage commun à ces deux allusions (« supposer deux équations de même forme » et « réduire une équation à même forme qu'une autre ») semble bien attester de la communauté des méthodes utilisées. L'analyse de Hudde est donc pertinente : la méthode des coefficients indéterminés joue un rôle important dans l'organisation de la *Géométrie* de Descartes [17].

---

13. Dans l'édition de la *Géométrie* de Descartes [90] t. VI, les pages auxquelles Hudde renvoie dans cette citation sont respectivement : 457–459, 463–464, 423.

14. Il s'agit, là aussi, de trouver un facteur quadratique (carré, de la forme $x^2 - 2ex + e^2$) à une équation ; *cf.* [90], t. VI, p. 420. Dans ce cas, l'élimination des inconnues dans le système d'équations polynomiales est plus simple que dans l'exemple ci-dessus, et se fait, pour chaque inconnue, seulement en reportant dans les autres équations la valeur de l'inconnue donnée par l'une d'elles : Descartes n'avait donc pas besoin, à ce stade, d'une méthode générale d'élimination.

15. *Cf.* [90], t. VI, p. 423.

16. *Cf.* [90], t. VI, p. 463–464.

17. Nous pourrions ajouter encore un argument pour le montrer : Descartes a aussi dû utiliser cette méthode dans sa construction géométrique des équations de degré 4 par intersection d'un cercle et d'une parabole. Comme il choisit en effet de laisser la parabole fixe, la position et le rayon du cercle adéquat sont difficiles à déterminer, à moins d'introduire des coefficients indéterminés a, b, c dans l'équation du cercle $(y-a)^2 + (z-b)^2 = c^2$ (on pourra comparer aux constructions utilisées par al-Khayyam, chez qui les deux sections coniques choisies sont au contraire variables et dépendent d'une manière plus simple des coefficients de l'équation à construire). La même méthode s'appliquerait aussi à la détermination des courbes utilisées par Descartes dans la construction de l'équation de degré 6.

Quant à la question de priorité que Hudde soulève, il ne connaissait peut-être pas l'œuvre de Faulhaber, qui avait publié dès 1622 la démonstration de la règle de résolution des équations de degré 4 par la recherche d'un facteur

Pourtant, la méthode des coefficients indéterminés introduit un détour dans l'ordre classique de l'analyse (par l'introduction d'inconnues auxiliaires, d'une part, et par la comparaison *formelle* de deux équations terme à terme, d'autre part) qui n'est pas conceptualisé de manière générale par Descartes : ainsi dans les *Regulae ad directionem ingenii*, quand Descartes parle de ramener plusieurs équations à une seule (règle XXI) [18], il s'agit simplement pour lui de l'étape finale de l'analyse d'un problème (la règle XXI appartient à la suite de règles sur les « questions parfaitement comprises », pour lesquelles « la difficulté proposée doit être directement parcourue »), les équations ne contenant pas d'autres inconnues que celles demandées dans la question à résoudre. Et il en est de même au début de la *Géométrie* [19]. Jules Vuillemin, qui montre le caractère synthétique de la méthode cartésienne, a remarqué la singularité de cette méthode de décomposition d'une équation en produit de facteurs de degrés inférieurs (telle qu'elle intervient dans la résolution algébrique de l'équation de degré 4) : il s'agit selon lui d'une « synthèse renversée ». Il voit aussi dans la méthode de décomposition une illustration de la seconde règle du *Discours de la méthode* : « diviser chacune des difficultés que j'examinerais, en autant de parcelles qu'il se pourrait et qu'il serait requis pour les mieux résoudre » [20]. Il ajoute cependant : « Dans sa *Géométrie*, Descartes ne nous donne d'ailleurs que le résultat, et non le principe de cette méthode qui a contribué à sa gloire » [21]. L'unique occurence explicite de la méthode des coefficients indéterminés dans la *Géométrie* reste son application à la méthode des sécantes. A la fin du siècle suivant, Lagrange [22] renverra encore au *De reductione aequationum* de Hudde pour la démonstration de la règle de résolution de l'équation de degré 4, tout en acceptant pleinement l'attribution à Descartes. Il écrira en effet, au sujet de la méthode de Ferrari pour la résolution des équations de degré 4 :

> Cette méthode, qu'on peut regarder comme la plus ingénieuse de toutes celles qui ont été inventées depuis pour le même objet, a été adoptée par tous les Analystes qui ont précédé Descartes ; *mais cet illustre Géomètre a cru devoir lui en substituer une autre moins simple à la vérité et moins directe, mais à quelques égards plus conforme à la nature des équations : c'est celle que la plupart des Auteurs suivent aujourd'hui.* [23]

2.1.3 L'élimination des inconnues chez Hudde

Hudde, comme ses prédécesseurs Descartes, Albert Girard et Faulhaber [24], réfléchit sur l'importance du nombre d'équations et d'inconnues lors de la résolution d'un système d'équations polynomiales, mais cette fois, ce discours est accompagné d'une méthode générale qui le fonde :

> Maintenant, revenons à la Règle [=la règle 19], et témoignons en peu de mots que par l'opération faite ici se montre une Méthode générale d'élimination par ordre de toutes les quantités inconnues [25] qui peuvent l'être, et de celles qui sont considérées comme inconnues ; ce qui, du moins selon moi, est d'une grande utilité, car j'ai croisé des questions qui, plus difficiles lorsque l'on suppose seulement une inconnue, ne peuvent alors être résolues, ou bien seulement avec beaucoup plus de travail, ou encore pour lesquelles il faille prendre des voies autres que celles auxquelles nous sommes habitués (...). Mais, j'ai dit que par cette Règle [26] s'éliminent les quantités, du moins celles qui peuvent l'être : c'est-à-dire qu'elles ne peuvent pas toujours toutes l'être, ni même toutes

quadratique au moyen de la méthode des coefficients indéterminés, bien que ses notations soient obscures (*cf. supra* section 1.3.1). Quant à l'idée même d'un « coefficient indéterminé », elle n'est pas nouvelle non plus : on peut en effet en voir déjà une utilisation chez les algébristes italiens au XVI^{e} siècle, ainsi dans l'*Ars magna* où Cardan démontre la règle de Ferrari pour résoudre l'équation de degré 4 en introduisant un coefficient indéterminé, par la phrase « Sit igitur numerus quadratorum 1 positio,... » (*cf.* [50], f. 74r).

18. *Cf.* [90], t. X, p. 469.
19. *Cf.* [90], t. VI, p. 372.
20. *Cf.* [90], t. VI, p. 18.
21. *Cf.* [322], p. 130.
22. *Cf.* [189], § 34.
23. Nous soulignons, *cf.* [189] § 26.
24. Pour Faulhaber, nous l'avons vu *supra* p. 37. Pour Descartes et Albert Girard, *cf. infra* p. 69.
25. *generalis Methodus tollendi quantitates incognitas.*
26. ici, lire plutôt « méthode » : c'est de sa « Méthode générale » qu'il est question, la « règle 19 » est seulement le lieu de la lettre où cette méthode est introduite.

> sauf une, ni toutes sauf deux, etc. En effet, si la question n'est pas un Théorème, elles ne peuvent pas toutes l'être ; et si elle est déterminée, toutes peuvent l'être sauf une ; si elle est en vérité déficiente d'une condition, en quoi elle est moins déterminée, toutes peuvent l'être excepté deux, et ainsi de suite, comme tu sais. [27]

Illustrons sa « méthode générale » sur l'exemple des cinq équations mentionnées ci-dessus :

$$\begin{cases} (1) & z - y^2 + w = q \\ (2) & k + yz - wy = r \\ (3) & l + yk + wz = s \\ (4) & yl + wk = t \\ (5) & wl = v \end{cases}$$

L'équation (1) donne une valeur de z que Hudde reporte dans les 4 équations suivantes. Puis l'équation (2) donne une valeur de k que Hudde reporte dans les 3 équations suivantes. Puis l'équation (3) donne une valeur de l que Hudde reporte dans les 2 équations suivantes. Ces deux dernières équations, ainsi transformées, ne contiennent plus que les deux inconnues w et y :

$$\begin{cases} (4) & sy - ry^2 + qy^3 + y^5 - 4wy^3 - 2qwy + 3w^2y + rw = t \\ (5) & sw - ryw + qy^2w + y^4w - 3w^2y^2 - qw^2 + w^3 = v \end{cases}$$

L'élimination de w entre ces deux équations n'est pas aussi triviale que les opérations précédentes, car les deux équations sont de degré strictement supérieur à 1 en w. C'est ici que Hudde utilise la technique qu'il a expliquée auparavant, l'algorithme euclidien : l'équation (4) permet de substituer à w^2 et w^3, dans l'équation (5), des expressions de degré 1 en w. L'équation (5), ainsi transformée, est de degré 1 en w, et donne donc une valeur rationnelle de w en fonction de y. En reportant cette valeur dans l'équation (4), Hudde obtient l'équation finale de degré 15 en y.

Hudde désigne sa méthode par une expression dont on retrouvera la trace sous la plume de Leibniz : *generalis Methodus tollendi ordine omnes quantitates incognitas*, « Méthode générale d'élimination par ordre de toutes les quantités inconnues » [28]. Hudde ne parle pas, dans sa lettre, de « questions abondantes » (c'est-à-dire lorsque le nombre d'équations est supérieur au nombre d'inconnues), mais c'est peut-être la conception même de cette méthode d'élimination (algorithme de recherche du plus grand commun diviseur) qui l'empêche de tomber dans l'écueil où s'était trouvé Fermat, dans le cas où deux équations dont il faut éliminer une inconnue ont un facteur commun de degré strictement supérieur à un (nous y reviendrons lorsque nous aurons exposé la méthode d'élimination de Fermat). Remarquons pour finir que l'on retrouvera au siècle suivant la réduite de degré 15 pour la recherche d'un facteur de degré 2 d'une équation de degré 6 dans les *Recherches sur les racines imaginaires des équations* d'Euler (1749), afin de démontrer le théorème fondamental de l'algèbre [29].

2.1.4 Dérivée formelle et discriminant d'un polynôme

Une autre notion présente dans le *De reductione aequationum* est celle de discriminant d'un polynôme, que Hudde introduit dans la règle 10. Plus généralement, il se propose de trouver une racine d'ordre de multiplicité donné d'un polynôme. Soit P un polynôme. Hudde formule une équation en général équivalente à la suivante :

$$\operatorname{pgcd}(P, P', ..., P^{(n-1)}) = 0$$

dont les racines sont les racines d'ordre de multiplicité n de P. Il définit à cet effet des dérivées formelles successives de P, dont notre dérivée usuelle est un cas particulier [30]. Hudde démontre cette

27. *Cf.* [89], p. 492.

28. *Cf.* [89], p. 492. Leibniz écrira par exemple : « generalis methodus tollendi literam duabus aequationibus communem », *cf.* [212], n°42, p. 173.

29. *Cf. infra* note 17 p. 77.

30. Les dérivées formelles successives Q_1,..., Q_{n-1} de P peuvent être définies, retrospectivement, au moyen de notre dérivée usuelle :

$$Q_{i+1} := -bxQ_i' + (a + b\deg(Q_i))Q_i$$

règle de manière purement algébrique, dans une deuxième lettre, comprise dans l'édition latine de la *Géométrie*, intitulée « de maximis et minimis ».

2.2 Deux brouillons de Newton : l'élimination successive des plus hautes puissances

Newton pratiquait l'élimination algébrique dans des systèmes d'équations de petits degrés, dès les années 1669–1770. Au cours de cette période, il révisa un traité d'algèbre hollandais écrit par un certain Kinckhuysen à la demande d'un de ses amis qui voulait en publier une traduction latine. Newton, peu satisfait du contenu de ce traité, finit par y adjoindre un chapitre entier intitulé « la transformation de deux ou plusieurs équations en une seule, pour exterminer les quantités inconnues » [31]. On y trouve en particulier le résultant de deux polynômes de degrés respectifs 2 et 3. L'impression du manuel de Kinckhuysen échoua cependant et Newton mit à profit son travail dans le cadre des cours d'algèbre qu'il donna à Cambridge dans les années 1670–1684. Il dut déposer à la Bibliothèque de Cambridge une version manuscrite de ces cours en 1684, les « *deposited Lucasian lectures on algebra* », texte qui fut publié par Whiteside avec toute l'œuvre mathématique de Newton [32]. Ils contiennent un chapitre important sur la théorie de l'élimination, qui fournit des formules des résultants de deux équations de degrés respectifs m et n, avec les couples (m, n) suivants : $(2, 2)$, $(3, 2)$, $(4, 2)$, $(3, 3)$. Les formules littérales qu'il a données sont bien délivrées de tout facteur superflu ; elles donnent bien le résultant tel que l'on pourrait le calculer par un déterminant, à la manière de Sylvester. En 1684, Newton révisa une partie de son manuscrit dans l'intention de la publier ultérieurement, sous le titre *Arithmeticae universalis liber primus* [33]. Le texte qui en résulta contient une liste de « règles » pour la réduction des équations, qui devaient constituer l'outillage employé par Newton pour calculer les résultants mentionnés ci-dessus. Nous reproduisons ici ces « règles » :

> *Règle 8 :* Quand deux équations ou plusieurs contiennent un nombre égal d'inconnues, il faut choisir une de ces inconnues, qui soit facilement exterminable, ainsi qu'une de ces équations, qui contienne cette inconnue de la manière la plus simple ; alors la valeur de cette inconnue doit être déterminée grâce à cette équation, puis substituée dans les autres équations. Ensuite doit être exterminée chaque autre inconnue dans les équations restantes, par la même méthode. Et la même procédure est répétée jusqu'à ce que toutes les inconnues soient exterminées sauf celle qui restera dans l'équation finale résultante.
> *Règle 9 :* Quand l'inconnue à exterminer s'élève à deux dimensions ou plus dans chaque équation, on choisit l'équation dans laquelle cette inconnue est à sa plus petite dimension, et la valeur de sa plus haute puissance dans cette équation est déterminée et substituée dans chaque autre ; puis, dans l'équation résultante, la valeur de la plus haute puissance est déterminée et substituée dans l'équation précédemment choisie ; après quoi, dans l'équation la plus récente, la valeur de la plus haute puissance est déterminée et substituée dans l'équation qui résultait précédemment ; et la procédure continue ainsi jusqu'à ce

où a et b sont des entiers relatifs. L'équation donnant les racines d'ordre de multiplicité n est alors :

$$\mathrm{pgcd}(P, Q_1, ..., Q_{n-1}) = 0$$

On voit facilement, si x ne divise pas P, que cette équation est bien équivalente à :

$$\mathrm{pgcd}(P, P', ..., P^{(n-1)}) = 0$$

Par exemple, soit $P = x^3 - 4x^2 + 5x - 2$. Hudde calcule Q en multipliant les termes successifs de P par les termes d'une suite arithmétique d'incrément $b = 1$, le terme dominant étant multiplié par le premier terme de cette suite, soit $a = 0$:

$$\begin{aligned} Q &= -4x^2 + 10x - 6 \\ &= -xP' - \deg(P)P \end{aligned}$$

Alors $\mathrm{pgcd}(P, Q) = x - 1$ indique bien la racine double de P.

31. *Cf.* [238], vol. II, p. 401–411.
32. *Cf.* [238] vol. V p. 54–491 ; la section sur l'élimination est aux p. 118–129.
33. *Cf.* [238], vol. V, p. 538–611.

qu'il en résulte une équation dans laquelle la quantité à éliminer est d'une dimension seulement, et alors, grâce à celle-ci, l'inconnue peut être éliminée par la règle 1 [= règle 8].
Règle 10 : Lorsque l'une des équations ne contient aucun terme impair [= où l'une des inconnues a un exposant impair], il faut mettre dans chaque autre équation les termes pairs d'un côté, et les impairs de l'autre. Quand chaque côté est élevé au carré, il en résulte une troisième équation sans terme impair : grâce à cela la réduction est souvent plus rapide.
Règle 11 : à l'occasion, une quantité est exterminée en ajoutant ou soustrayant les égaux des égaux, ou encore en multipliant ou divisant les égaux par les égaux, ou même en posant égales l'une à l'autre deux grandeurs égales à une même troisième.
Règle 12 : Au lieu de quantités données compliquées, on peut à volonté en supposer et écrire de plus simples. [34]

En 1707, Whiston, successeur de Newton à la chaire de mathématique de Cambridge, publia les conférences de Newton, sous le titre *Arithmetica universalis ; sive de compositione et resolutione arithmetica liber*. Newton n'était pas satisfait de cette édition, incomplète et erronée. Il la révisa en vue d'une seconde édition, qui parut en 1722 à Londres, sous le titre *Editio secunda. Arithmetica universalis ; sive de compositione et resolutione arithmetica liber. In qua multa immutantur et emendatur, nonnulla adduntur*. Il reste deux brouillons de cette dernière période (1720–1722) contenant une esquisse des calculs de résultants faits par Newton. L'*Arithmetica universalis* de Newton n'eut pas beaucoup de succès de son vivant, mais elle fut traduite et rééditée plusieurs fois après sa mort.

La « règle 9 » énoncée ci-dessus résume parfaitement la méthode de Newton pour l'élimination. Soient deux équations à une inconnue, de même degré (les coefficients peuvent bien sûr désigner des polynômes en une seconde inconnue, en vertu par exemple de la « règle 12 ») :

$$\begin{cases} (1) \quad Ay^m + By^{m-1} + ... + V = 0 \\ (2) \quad ay^m + by^{m-1} + ... + v = 0 \end{cases}$$

L'équation (2) fournit une valeur de y^m, que l'on peut substituer dans (1), d'où l'équation :

$$(3) \quad (B - \frac{Ab}{a})y^{m-1} + (C - \frac{Ac}{a})y^{m-2} + ... + (V - \frac{Av}{a}) = 0$$

L'équation (3) fournit une valeur de y^{m-1}, que l'on peut substituer dans (2), ce qui revient à considérer l'équation :

$$(4) \quad [ay^m + by^{m-1} + ... + v] - \frac{ay}{B - \frac{Ab}{a}} \left[(B - \frac{Ab}{a})y^{m-1} + (C - \frac{Ac}{a})y^{m-2} + ... + (V - \frac{Av}{a}) \right] = 0$$

qui est de degré $(m - 1)$ en y. L'équation (4) fournit une autre valeur de y^{m-1}, que l'on peut substituer dans (3), et ainsi de suite. Dans la suite, nous renverrons à cette méthode par la locution « méthode d'élimination successive des plus hautes puissances de l'inconnue ». On remarque que

34. *Cf.* [238], vol. V, p. 584–594. La règle 9 qui nous intéresse le plus est énoncée de manière moins précise dans les *Lucasian lectures* (*cf.* [238], vol. V, p. 122, 123) :

Cum in neutra aequatione tollenda quantitas unius tantum dimensionis existit valor maximae potestatis ejus in utraque quaerendus est ; deinde si potestates istae non sint eaedem, aequatio potestatis minoris multiplicanda est per tollendam quantitatem aut per ejus quadratum aut cubum etc. ut ea evadat ejusdem potestatis cum aequatione altera. Tum valores illarum potestatum ponendae sunt aequales et aequatio nova prodibit ubi maxima potestas sive dimensio tollendae quantitatis diminitur. Et hanc operationem iterando quantitas illa tandem auferetur.

Quand chacune des équations est de degré supérieur à un en la quantité à éliminer, la valeur de sa plus haute puissance doit être déterminée dans chacune ; ensuite, si ces puissances ne sont pas les mêmes, l'équation de plus bas degré doit être multipliée par la quantité à éliminer, ou par son carré, ou son cube, etc. pour qu'elle devienne de même degré que l'autre équation. Alors les valeurs de ces puissances doivent être posées égales l'une à l'autre, et il en résulte une nouvelle équation de degré inférieur en la quantité à éliminer. Et en itérant cette opération, cette quantité disparaîtra finalement complètement.

cette méthode diffère légèrement de l'algorithme euclidien du plus grand commun diviseur, qui aurait eu pour effet de produire directement une équation (4) de degré $(m-2)$.

En pratique, Newton remarque qu'il est parfois possible de « contracter » le travail décrit par cet algorithme : les règles 10 et 11 peuvent aider à cela. Les opérations licites sont d'ajouter ou de multiplier aux deux membres d'une équation des expressions égales, ou encore d'élever au carré ces deux membres ; il est même possible d'extraire des radicaux si l'on peut les faire disparaître par la suite. Newton ne prend pas garde au risque d'introduire des facteurs superflus dans l'équation finale : il s'agit seulement d'éliminer une inconnue. Il est donc d'autant plus surprenant que les formules données par Newton pour les résultants $(2,2)$, $(3,2)$, $(4,2)$ et $(3,3)$ soient délivrées de facteur superflu. Au moyen des brouillons mentionnés ci-dessus, Whiteside a tenté de restituer le calcul du résultant pour le couple $(3,3)$, mais cette restitution semble être erronée, l'équation finale ainsi obtenue (que Whiteside n'a pas calculée) étant identiquement nulle [35] ! Et quelle que soit la valeur de la restitution de Whiteside, il reste deux points inexpliqués :

- dans ces brouillons, Newton n'utilise pas la notation condensée $Ay^m + By^{m-1} + ... + V$, et il attribue au contraire un coefficient à chaque terme monôme en x, y. Cela semble étrange, si le but de Newton était seulement de recalculer ou de vérifier les formules publiées dans son *Arithmetica*.
- à chaque fois, Newton n'achève pas le calcul de l'équation finale et s'arrête à l'avant-dernière équation, linéaire en y. Dans un des deux brouillons, il reprend la procédure sur la page de droite en échangeant le rôle joué par les deux équations initiales (c'est possible parce que les deux équations sont de même degré). Il obtient ainsi une autre équation linéaire en y. Whiteside suggère que Newton obtenait l'équation finale en égalant les deux valeurs de y obtenues ainsi. Mais Newton ne développe même pas ces valeurs de y, et il semble s'intéresser seulement à la forme du résultat : pour deux cubiques, la valeur de y obtenue est de la forme $y = \frac{\xi+...+\pi x^5}{\rho+...+\sigma x^4}$, pour une cubique et une quadratique, $y = \frac{\xi+...+\pi x^3}{\rho+...+\sigma x^2}$, pour une quartique et une quadratique, $y = \frac{\xi+...+\pi x^4}{\rho+...+\sigma x^3}$.

Il est donc permis de douter que ces brouillons concernent vraiment le calcul du résultant. En fait, il est possible de calculer les résultants donnés par Newton en s'en tenant seulement à la règle 9 énoncée ci-dessus : une fois l'équation finale obtenue, il suffit de l'écrire sous forme de fraction irréductible, et de ne garder que le numérateur [36].

2.3 Le théorème de Bézout chez Newton

On a vu apparaître au XVII^{e} siècle un certain nombre de directions théoriques, qui serviront de socle à ce que les mathématiciens de la fin du XVIII^{e} siècle (Bézout, Lagrange) et du XIX^{e} siècle désigneront comme « théorie de l'élimination » :

- Des énoncés généraux sur la résolution des systèmes d'équations polynomiales qui relient le nombre de solutions d'un système au nombre des équations et des inconnues (ainsi chez Faulhaber, Albert Girard, Descartes, Hudde – nous y ajouterons Fermat, *cf. section* 3.4).
- Le lien entre l'élimination et une théorie de la divisibilité des polynômes, accueillant l'algorithme euclidien de recherche du plus grand commun diviseur comme « méthode générale d'élimination » d'une inconnue dans un système de deux équations (chez Hudde et Leibniz, bientôt chez Rolle).
- Un lien, encore discret, entre les méthodes d'élimination et la géométrie analytique ; l'idée qu'à tout système d'équations correspond un « lieu » est rarement formulée de manière générale [37], mais le domaine en plein essort de la « construction géométrique des équations » a nécessairement recours à des techniques d'élimination.

Précisons ce dernier point. Il nous semble en effet que, parallèlement à la recherche d'une méthode générale d'élimination, c'est dans la géométrie analytique du XVII^{e} siècle qu'il faille chercher l'origine

35. Premier brouillon : [238], vol. V, p. 518–519 ; deuxième brouillon vol. VIII, p. 461–466 ; restitution de Whiteside : vol. V, p. 519, note 7.

36. On le montrera *infra* p. 188. Voir aussi section 3.5.

37. cf. par exemple la phrase elliptique de Descartes dans sa *Géométrie* : « s'il manque deux conditions à la détermination de ce point, le lieu où il se trouve est une superficie ».

d'un théorème, le « théorème de Bézout », clef de voûte de toutes les recherches sur l'élimination algébrique aux XVIII^{e} et XIX^{e} siècles. L'extension, par Descartes, du domaine de la géométrie analytique à des courbes de degré élevé s'accompagnait de l'échec d'une classification des problèmes et des courbes par « genre ». Celle-ci était fondée d'une part sur un principe déjà admis des « anciens », les problèmes étant classés suivant les courbes qu'on emploie pour les construire [38], et d'autre part sur un principe original, les courbes étant classées suivant la « composition » des mouvements des outils qui permettent de les tracer. Cette classification devait de plus se prêter à une traduction algébrique, faisant intervenir le degré des équations (équations à une inconnue pour les « problèmes », équations à deux inconnues pour les courbes). Par exemple, les problèmes se traduisant par des équations de degré 3 ou 4 en l'inconnue cherchée appartiennent à une même classe, pour autant que la résolution des équations de degré 3 ou 4 puisse se faire au moyen de courbes planes de degré 2 (sections coniques, descriptibles au moyen d'un « compas » particulièrement simple). La critique du système cartésien, par Fermat d'abord, dans sa *Dissertatio tripartita* [39], ne sut produire que des contre-exemples particuliers. « Chez Descartes, nous dit Fermat, les problèmes de degrés 8 ou 7 nécessitent des courbes de degrés 5 ou 6, ceux de degrés 10 ou 9 des courbes de degré 7 ou 8,..., et ainsi de suite à l'infini ; que ces cartésiens voient comme c'est loin de la simplicité et de la vérité géométrique ». Fermat montre, par exemple, comment construire toute équation de degré $2n$ au moyen de deux courbes de degré n (et de l'extraction d'une racine carrée) ; puis il donne des exemples d'équations de degré n constructibles par deux courbes de degrés $\leq \sqrt{n}+1$: ainsi l'extraction de $m(m+1)$ moyennes proportionnelles (équation de degré $m(m+1)+1$) peut se faire par l'intersection de deux courbes de degré $(m+1)$; enfin il ajoute que la résolution des équations de la forme $x^{pq}=a$ se ramène à l'intersection de deux courbes de degrés p et q. Mais bientôt, en 1688, toujours dans le cadre de la critique du système cartésien, Jacques Bernoulli écrit :

> Je pense qu'il n'est pas difficile de démontrer que des courbes de n'importe quel genre sont aptes à la construction d'équations d'autant de dimensions, qu'en compte le carré du nombre de dimensions, auxquelles montent les équations exprimant la nature de ces courbes. [40]

Chaque racine de l'équation considérée devant être construite au moyen d'un des points d'intersection des deux courbes, il faut donc penser que deux courbes de degré n puissent se couper en n^2 points. Il s'agit d'un cas particulier du « théorème de Bézout ». Bernoulli donne l'exemple de deux cubiques planes, qui se coupent en cinq points réels, correspondant aux cinq racines réelles d'une certaine équation de degré 9. Dans un brouillon de Newton, datant des années 1666–1671, dans le cadre de ses recherches, d'une part, sur la classification des cubiques, et, d'autre part, sur la géométrie analytique (préparant le *Tractatus de methodis serierum et fluxionum* de 1670–1671 ; le théorème de Bézout est d'ailleurs le seul résultat, dans ce brouillon, que Newton ne reprend pas dans son traité), Newton explore cette même direction et formule, pour la première fois à notre connaissance, le théorème dit « de Bézout » pour deux courbes algébriques planes, tout en indiquant le lien naturel entre ce théorème et l'élimination algébrique :

> Etant donné deux courbes, trouver leurs points d'intersection. C'est plus un principe qu'un problème. Mais plutôt susceptible d'une solution algébrique que d'une solution géométrique, et on le fait en éliminant une des deux inconnues des équations. D'où il apparaîtra qu'il y a autant de points d'intersection que le rectangle des dimensions des courbes. [41]

38. *Cf.* [90], t. VI, p. 388.

39. *Cf.* [112], t. I, p. 118–131. Voir aussi la lettre de Fermat à Digby [113], p. 491–497, et [261] pour une analyse de la *Dissertatio tripartita*.

40. *Cf.* [17] : « Existimo namque, demonstratu haud difficile esse, quod cujuslibet generis curvae aptae sint ad construendas Aequationes tot dimensionum, quot indigat quadratum numeri dimensionum, ad quas ascendunt Aequationes, curvarum illarum naturam exprimentes ». Le marquis de L'Hospital précisera un peu cela dans un ouvrage qui paraîtra en 1707, *cf.* [87] p. 446 : « Il faut extraire la racine quarrée de la plus haute dimension de l'inconnüe. Si elle est exacte, chacun des deux lieux doit avoir autant de degrés que cette racine contient d'unité ; et si elle ne l'est pas, ou le reste est égal, ou moindre que la racine, et alors l'un des lieux aura pour degré le nombre de la racine, et l'autre ce même nombre augmenté de l'unité : ou le reste est plus grand que la racine, et alors chacun des deux lieux aura pour degré le nombre de la racine augmenté de l'unité ».

41. *Cf.* [238], t. II, p. 177 : « Datis duabus curvis invenire puncta intersectionis. this is rather a principle than a

Il utilise ce « principe » pour démontrer que le nombre de tangentes à une courbe de degré n, passant par un point donné, est égal, en général, à $n(n-1)$ (comme intersection de la courbe de degré n et d'une polaire de degré $(n-1)$) « sauf si quelque partie extraordinaire s'échappe à l'infini, ou que deux parties viennent ensemble », c'est-à-dire si la courbe proposée a des points à l'infini, ou des points singuliers.

2.4 Les manuscrits de Leibniz

Eberhard Knobloch a édité 66 manuscrits de Leibniz qui se rapportent de près ou de loin à la théorie de l'élimination, dans un livre intitulé *Der Beginn der Determinantentheorie*[42] (1980). La préface annonce un second volume de commentaires, qui n'a, semble-t-il, pas encore vu le jour. Mais il a publié d'autre part deux articles sur le sujet (1974, 2001), contenant une étude détaillée des nombreuses contributions de Leibniz à la théorie de l'élimination. Nous nous contenterons ici d'évoquer les plus significatives, en les situant par rapport aux travaux antérieurs.

Leibniz a rencontré Hudde, et il a certainement lu le *De reductione aequationum*, comme l'indique l'un de ses premiers manuscrits sur l'élimination, dressant un programme pour ses recherches à venir en algèbre :

> Nous avons des Tables de Nombres, personne n'a donné de Tables de lettres ni de Combinaisons. Il faut donc chercher la manière de construire des Tables de telle sorte, que leur utilisation soit évidente.
> 1ère partie. Des racines en nombre quelconque des Equations à une inconnue (...) par exemple en degré 3, $x^3 + bx^2 + acx + a^2d = 0$ dont les racines irrationnelles peuvent être exhibées de manière générale et par une formule unique (...)
> 2ème partie. Des opérations sur les équations, méthode de recherche des diviseurs rationnels et irrationnels de degré inférieur des équations, et tables comme celles de Hudde. (...)
> 3ème partie. De la réduction des Equations de plusieurs inconnues aux équations à une seule inconnue (...)[43]

La recherche d'une formule générale de résolution algébrique des équations à une inconnue est une motivation importante dans les recherches de Leibniz en théorie de l'élimination, qui semble avoir partiellement abouti, à la suite de la correspondance Leibniz–Tschirnhaus, avec la publication en 1683 d'un article, dans lequel Tschirnhaus expose une méthode de résolution algébrique : *Methodus auferendi omnes terminos intermedios ex data aequatione*[44]. Cette méthode a au moins le mérite de s'appliquer aussi bien aux équations de degré 4 qu'à celles de degré 3, même si elle ne s'étend pas, comme l'espérait Tschirnhaus, aux équations de degré supérieur. Dans cet article, Tschirnhaus est amené à éliminer des inconnues[45], ainsi qu'à utiliser la méthode des coefficients indéterminés. La méthode de Tschirnhaus, imparfaitement comprise par certains de ses contemporains, a suscité une controverse dont nous reparlerons, ainsi que de la genèse de cette méthode, *infra* p. 75.

Leibniz connaissait (vers 1678–1683) une méthode d'élimination analogue à celle que publierait Euler dans son *Introductio ad analysin infinitorum* (1748), et conduisant directement au résultant. Elle utilise la méthode des coefficients indéterminés pour réduire le problème de l'élimination d'une inconnue entre deux équations polynomiales à un simple système d'équations linéaires. On peut la

probleme. But rather propounded of the Algebraicall then geometricall solution and that is done by eliminating one of the two unknown quantitys out of the equations. From whence it will appear that there are so many cut points as the rectangle of the curves dimensions. »

42. *Cf.* [212].

43. *Cf.* [212] n°33, p. 146, janvier 1675. Les « tables de Hudde » en question sont certainement celles que l'on trouve dans le *De reductione aequationum*, p. 440–457, pour la recherche de facteurs d'une équation, pas nécessairement rationnels mais ayant certains coefficients nuls.

44. *Cf.* [311].

45. Sans expliquer *comment* il élimine. Ce qui montre que l'élimination n'était plus un problème nouveau à l'époque où écrivait Tschirnhaus.

résumer ainsi [46]. Soit deux équations de degrés m et n :

$$\begin{cases} 10x^m + 11x^{m-1} + 12x^{m-2} + ... = 0 \\ 20x^n + 21x^{n-1} + 22x^{n-2} + ... = 0 \end{cases}$$

Les notations 10, 11, 12, ..., 20, 21, 22, ... sont des notations « indicielles » pour les coefficients des équations (le chiffre de gauche indiquant de quelle équation il s'agit, et le chiffre de droite, de quel coefficient). Multiplions la première par $30x^{n-1} + 31x^{n-2} + 32x^{n-3} + ...$ et la deuxième par $40x^{m-1} + 41x^{m-2} + 42x^{m-3} + ...$ (où 30, 31, 32,..., 40, 41, 42,... sont les « coefficients indéterminés »). Soustrayons les produits l'un de l'autre. On obtient une équation de degré $(m+n-1)$ en x. Si l'on impose que tous les coefficients sauf celui du terme constant s'annulent, on obtient un système de $(m+n-1)$ équations linéaires homogènes en les $(m+n)$ inconnues 30, 31,..., 40, 41,...., dont la solution n'est déterminée qu'à une constante multiplicative près. La première équation linéaire étant $30.10 - 40.20 = 0$, Leibniz peut donc choisir, par exemple, $30 = 20$ et $40 = 10$. Une fois le système résolu, on peut remplacer les inconnues par leurs valeurs dans l'expression du terme constant, qui, une fois égalée à zéro, donne l'équation finale résultant de l'élimination. Leibniz résume toute cette suite d'opérations sous la forme [47] :

$$\left.\begin{array}{ccccc} 13.31 & +12.31x & +11.31x^2 & +10.31x^3 & \\ & 13.30 & +12.30 & +11.30 & +10.30x^4 \\ -22.42 & -22.41 & -22.40 & & \\ & -21.42 & -21.41 & -21.40 & \\ & & -20.42 & -20.41 & -20.40 \\ \| & \| & \| & \| & \| \\ 0 & 0 & 0 & 0 & 0 \end{array}\right\} = 0$$

Cette écriture est très proche de l'écriture déterminantielle du résultant qu'adoptera Sylvester au XIX$^{\text{ème}}$ siècle.

Leibniz donne une démonstration de cette méthode [48] semblable à celle que donnera Euler en 1764 [49], et reposant sur la notion de diviseur commun. Notons les deux équations proposées $P(x) = 0$ et $Q(x) = 0$. Si P et Q ont un diviseur commun $(x - f)$ (c'est-à-dire une racine commune), on a $P(x) = (x-f)R(x)$ et $Q(x) = (x-f)S(x)$. Alors $S(x)P(x) - R(x)Q(x) = 0$. Pour exprimer la condition d'existence d'une racine commune, il suffit donc de prendre deux polynômes R et S de degrés respectifs $(\deg P - 1)$, $(\deg Q - 1)$, à coefficients indéterminés, et d'imposer que $S(x)P(x) - R(x)Q(x)$ soit identiquement nul.

Enfin, vers 1693–1694, Leibniz invente une méthode originale d'élimination, qu'il qualifie lui-même de « Méthode par conjonctions d'équations inverses » [50], dans laquelle les deux équations jouent un rôle symétrique (contrairement à l'algorithme euclidien où l'une est dividende, l'autre diviseur). Soit deux équations $(n \leq m)$:

$$\begin{cases} 10x^m + 11x^{m-1} + 12x^{m-2} + ... + 1m = 0 \\ 20x^n + 21x^{n-1} + 22x^{n-2} + ... + 2n = 0 \end{cases}$$

Leibniz multiplie la première par 20, la deuxième par $10x^{m-n}$, et soustrait les produits l'un de l'autre :

$$(20.11 - 10.21)x^{m-1} + (20.12 - 10.22)x^{m-2} + ...(20.1m - 10.2m) = 0$$

(on pose $2m = 0$ si $n \neq m$). Puis il multiplie la première par $2n$, la deuxième par $1m$, soustrait les produits l'un de l'autre, et divise par x :

$$... + (2n.1\overline{m-n} - 1m.20)x^{n-1} + (2n.1\overline{m-n+1} - 1m.21)x^{n-2} + ... + (2n.1\overline{m-1} - 1m.2\overline{n-1}) = 0$$

46. *Cf.* par exemple [212] n°40, p. 168, l. 32–54. Nous adoptons ici les notations du manuscrit plus tardif (1684–1694) n°54, p. 238, l. 176–188.

47. *Cf.* [212], n°54, p. 238.

48. *Cf.* [212] n°63 (vers 1711–1713), p. 309, l. 176–199.

49. *Cf.* [108], série 1, vol. VI, p. 197–211 ; ici § 13, p. 203.

50. *Methodus per conjonctiones aequationum inter tollendum, quae sibi sunt inversae*, *cf.* [212] n°54, p. 262, l. 236. Cette méthode est décrite dans ce même manuscrit, l. 31–51.

Ces deux nouvelles équations sont de degré $(m-1)$; en réitérant ce procédé m fois, on aboutit à une expression de degré 0 en x. On verra cette méthode d'élimination revenir sous la plume d'Euler en 1748.

A la même époque, Leibniz réalise une étude très approfondie de l'algorithme euclidien du plus grand diviseur, et, plus généralement, de la division euclidienne (des polynômes) [51]. L'aboutissement de ces recherches est une méthode pour écrire directement le résultat de la division euclidienne de deux polynômes (sans avoir à *effectuer* la division) au moyen d'une description combinatoire des expressions algébriques formant les coefficients du quotient et du reste. En particulier, la « Formule absolue générale du quotient » est décrite très clairement par Leibniz dans le brouillon d'une lettre à Charles Reyneau ([212] n°66, p. 324–326, février 1712, en français). Les manuscrits [212] n°57 et [174] mettent plus l'accent sur la description des restes successifs de la division euclidienne ; seuls ces restes sont utiles pour l'algorithme du plus grand commun diviseur. Dans ces recherches, l'originalité de Leibniz réside peut-être dans une approche qu'il qualifie lui-même de « synthétique ». L'algorithme euclidien, itératif, réduit le calcul d'un reste de deux polynômes de degrés respectifs $(m+n)$ et m au calcul d'un reste de deux polynômes de degrés respectifs m et $(m-1)$. Mais Leibniz cherche surtout des « formules » (plutôt que des règles de calcul). Le caractère itératif de l'algorithme a pour conséquence le caractère récursif de la formule à laquelle il conduit. En remplaçant, dans la formule d'un reste de deux polynômes de degrés repectifs $(m-1)$ et m, les coefficients de ces deux polynômes par des expressions algébriques adéquates, on obtient la formule d'un reste de deux polynômes de degrés respectifs $(m+n)$ et m. Leibniz désigne ce procédé de remplacement (qui est proprement un procédé d'instanciation de variables dans une formule) par le mot *explicatio*. Ainsi l'ordre des opérations est renversé : au lieu de réduire successivement le calcul des restes à des équations de degrés décroissants, on commence par écrire la formule du reste pour des équations linéaires, puis on écrit graduellement les formules des restes d'équations de degrés croissants. C'est ce renversement de l'ordre qui justifie le qualificatif « synthétique » [52]. D'autre part, le caractère récursif des formules et les « explications » successives se développent aisément sous la forme graphique d'un arbre (une *tabula dichotomica*), dont les deux manuscrits donnent de beaux exemples. L'étude combinatoire d'une formule consiste alors à trouver les propriétés combinatoires de l'arbre qui permettront d'écrire la dernière ligne (les feuilles) sans avoir à reconstruire tout l'arbre en partant de la racine.

Enfin, combinatoire et « notations indicielles » ont aussi conduit Leibniz à étudier les propriétés d'homogénéité du résultant de deux polynômes. En 1692–1693, il considère deux équations à une inconnue, dont les coefficients sont affectés d'un poids. Il écrit :

> Lois de combinaisons : si l'on a deux équations en x par lesquelles doit être tiré x, l'une de dimension a, l'autre de dimension b,
> 1) Nécessairement dans la nouvelle équation sans la lettre x, les coefficients de l'équation de dimension a monteront (...) à la dimension b, et inversement.
> (...)
> 4) Chaque terme est de dimension $(a+b)$ et la somme des chiffres de droite fait ab. [53]

Le fait que le poids du résultant est égal au produit des degrés des deux équations est un résultat important, équivalent au théorème de Bézout pour deux équations. Leibniz, armé qu'il était d'une écriture déterminantielle du résultant et de ses notations indicielles, aurait même pu en concevoir une *démonstration* semblable à celle donnée par Bézout en 1764 [54]. Mais il semble s'être contenté d'*observer* cette propriété d'homogénéité.

51. *Cf.* [212] n°57, p. 254–267, et le manuscrit édité séparément dans [174], p. 160–173, et abondamment commenté par Knobloch.
52. *Cf.* [212] n°57, p. 266, l. 343.
53. *Cf.* [212], n°56, p. 248, l. 48–60. Les « chiffres de droite » sont les chiffres de droite des coefficients qui entrent dans l'expression du résultant, dans la notation indicielle décrite ci-dessus.
54. Dans [26], p. 298–300.

Chapitre 3

Origine de l'élimination algébrique

3.1 Abū Kāmil

La considération de problèmes à plusieurs inconnues n'est pas nouvelle, et commence peut-être avec la reprise de l'analyse diophantienne rationnelle au sein des traités d'algèbre, d'Abū Kāmil (IX$^{\text{ème}}$ siècle), avant même la traduction des *Arithmétiques* de Diophante en langue arabe, puis al-Karajī, jusqu'à Bombelli et Viète. Les problèmes diophantiens offrent une première classe, très vaste, d'équations à plusieurs inconnues et de systèmes d'équations. Ce n'est pourtant qu'une classe particulière, et les méthodes diophantiennes [1] ont une généralité relative à cette classe de problèmes. Les algébristes qui les utilisent ont recours à certaines transformations algébriques afin d'éliminer certaines puissances d'une des inconnues, voire afin d'abaisser le nombre des inconnues.

Abū Kāmil étudie des problèmes diophantiens dans la dernière partie de son traité d'algèbre. L'ordre dans lequel sont exposés ces problèmes laisse percevoir une distinction entre des problèmes dont chacun a pour ensemble de solutions, quand il en existe au moins une, un ensemble infini qui peut être paramétré (c'est le cas des problèmes 1 à 25), et des problèmes pour chacun desquels Abū Kāmil ne connaissait pas de tel paramétrage rationnel (c'est le cas des problèmes 26 à 38). En fait, pour les problèmes de cette seconde classe (Abū Kāmil ne devait certes pas le savoir) il n'existe pas de paramétrage rationnel [2]. On trouve par exemple, au sein de la première classe, les problèmes 7–9, 11, 15–18, 20, 22 et 23 qui consistent chacun en un système de deux équations de degré 2, dont l'ensemble des solutions est, en langage moderne, une courbe dans $\mathbb{A}^3$, rationnelle de degré 4 et ayant un point double. Au contraire les problèmes 26 à 38 consistent chacun en un système de deux équations de degré 2 dont l'ensemble des solutions est une courbe lisse de degré 4, de genre 1 et n'admettant donc pas de paramétrage rationnel.

Une autre classe particulière de problèmes est celle des systèmes d'équations linéaires, qui font aussi très tôt l'objet d'une étude théorique de la part des algébristes. Mentionnons ainsi Al-Samaw'al (XII$^{\text{e}}$ siècle) qui dans son *Al-Bāhir en algèbre* recherche les conditions de compatibilité d'un système de 210 équations linéaires à 10 inconnues, au sein d'un chapitre de philosophie mathématique, comme exemple d'une proposition faisant intervenir un grand nombre de « conditions » [3]. Les systèmes d'équations linéaires font aussi l'objet des problèmes 39 à 42 de la dernière partie du traité d'Abū Kāmil. Celui-ci utilise les mots « chose », « dinar », « dirham », « fals » pour désigner quatre inconnues différentes [4], et attire particulièrement l'attention du lecteur lorsqu'une des in-

1. Pour une interprétation moderne des méthodes diophantiennes, « dans les termes et à l'aide des notions de la géométrique algébrique », faisant d'ailleurs entrevoir la parenté conceptuelle avec la théorie de l'élimination, *cf.* [96], t. 3, p. LXXXV–CXXXVI, et surtout p. CXVI.

2. Voir les commentaires de R. Rashed [260] p. 34–36.

3. *Cf.* [7], p. 77–82. Les équations, au nombre de $\binom{10}{6} = 210$, sont formées au moyen des sommes de six inconnues parmi dix.

4. Dès la première partie du traité d'algèbre d'Abū Kāmil, le mot *dinar* désigne une seconde inconnue et le mot *fals* une troisième. *Cf.* l'édition critique et la traduction du traité d'algèbre d'Abū Kāmil par R. Rashed, à paraître [259]. Nous pouvons aussi renvoyer, pour cette première partie du traité, à l'édition par Martin Levey d'une version hébraïque du traité [5] (les problèmes 8, 14, 31, 42 de cette version contiennent par exemple le mot *dinar*, et les

connues vient d'être éliminée. Grâce à la traduction et aux notes manuscrites que nous a prêtées R. Rashed [5], nous allons à présent donner le commentaire des problèmes 14 et 38 (deux problèmes diophantiens) et des problèmes 39, 40 et 42 (systèmes d'équations linéaires).

Problème 14. Il s'agit du système d'équations suivant :

$$\begin{cases} x+y=10 \\ x+20=z^2 \\ 50+y=t^2 \end{cases}$$

Abū Kāmil indique explicitement qu'il s'agit là d'un problème « indéterminé ». Pour en trouver les solutions rationnelles, il élimine d'abord y (au moyen de la première équation), puis x (au moyen de la seconde) :

$$\begin{cases} x=z^2-20 \\ y=30-z^2 \\ 80-z^2=t^2 \end{cases}$$

Le problème revient donc à trouver les solutions rationnelles de l'équation $z^2+t^2=80$. Plus généralement, Abū Kāmil sait trouver les solutions rationnelles de l'équation suivante (problème 12) :

$$z^2+t^2=c$$

Si $c=a^2+b^2$, où a et b sont deux nombres rationnels, alors il existe une infinité de solutions rationnelles (Abū Kāmil insiste particulièrement sur ce fait). Pour les trouver, il applique une méthode que connaissait bien Diophante sept siècles plus tôt (la « méthode de la corde »), mais qu'il expose dans le langage de l'algèbre. Il pose $z=a+\xi$ et $t=b-u\xi$, où u est un nombre rationnel quelconque (un paramètre). L'équation devient alors $b^2+u^2\xi^2-2bu\xi=c-(\xi^2+a^2+2a\xi)$, d'où $\xi=\frac{2bu-2a}{u^2+1}$. Abū Kāmil calcule ainsi une solution rationnelle du problème 14 :

$$z=5+\frac{3}{17},\ t=7+\frac{5}{17},\ x=6+\frac{230}{289},\ y=3+\frac{59}{289}.$$

Problème 38. Il s'agit du système d'équations :

$$\begin{cases} x^2+x+1=y^2 \\ x^2+2x+2=z^2 \end{cases}$$

En langage moderne, on y reconnaît les équations d'une courbe lisse de degré 4 et de genre 1 dans $\mathbb{A}^3$, dont la complétion projective possède quatre points à l'infini $(1:1:1:0)$, $(1:-1:-1:0)$, $(1:1:-1:0)$, $(1:-1:1:0)$. Pour en trouver un point rationnel, Abū Kāmil pose $z=y+u$ (où u est une constante ; $z=y+u$ est l'équation d'un plan passant par les points infinis $(1:1:1:0)$ et $(1:-1:-1:0)$). On obtient ainsi le système d'équations :

$$\begin{cases} x^2+x+1=y^2 \\ x^2+2x+2=z^2 \\ z=y+u \end{cases}$$

Ce système possède en général deux solutions (finies) que l'on trouve en résolvant l'équation de degré 2 en x résultant de l'élimination de y et z. En choisissant la valeur de u de sorte que le terme de degré 2 disparaisse, on obtient une équation linéaire en x, qui nous donne une solution rationnelle. Pour éliminer y et z, Abū Kāmil procède selon les étapes suivantes. Il élimine d'abord z :

$$\begin{cases} x^2+x+1=y^2 \\ x^2+2x+2=(y+u)^2 \end{cases}$$

problèmes 29, 33, 34 le mot *fals* ; *cf.* aussi les problèmes 60 et 61 où il est question de trois inconnues, mais désignées autrement).

5. *Cf.* son livre à paraître [259].

Puis il se débarasse du terme en y^2 dans la deuxième équation au moyen de la première :

$$\begin{cases} x^2+x+1=y^2 \\ 2uy=x+1-u^2 \end{cases}$$

Puis il élimine y :

$$4u^2(x^2+x+1)=(x+1-u^2)^2$$

Le choix $u=\frac{1}{2}$ lui donne la solution $x=\frac{7}{8}$, $y=\frac{13}{8}$, $z=\frac{17}{8}$.

Problème 39. Nous traduirions aujourd'hui l'énoncé de ce problème en le système d'équations linéaires suivant :

$$\begin{cases} u=x+\frac{1}{2}(y+z+t) \\ u=\frac{1}{3}(x+z+t)+y \\ u=\frac{1}{4}(x+y+t)+z \\ u=\frac{1}{5}(x+y+z)+t \end{cases}$$

Première méthode : Abū Kāmil considère quatre inconnues qu'il nomme chose, dinar, dirham, fals et qui sont les quantités que nous avons notées x, y, z, t ci-dessus. L'élimination de u (qu'il n'a pas besoin de « nommer », puisque cette inconnue n'apparaît jamais en tant que telle dans les équations qu'il considère) le conduit à considérer successivement les équations suivantes dans la résolution du problème :

$$\begin{cases} x+\frac{1}{2}(y+z+t)=\frac{1}{3}(x+z+t)+y \\ \frac{1}{4}(x+y+t)+z=\frac{1}{3}(x+z+t)+y \\ \frac{1}{5}(x+y+z)+t=\frac{1}{3}(x+z+t)+y \end{cases}$$

La première équation fournit une valeur de y qui, reportée dans les suivantes, permet d'éliminer y (ce qu'Abū Kāmil explique en disant qu'il a « ôté le nom du dinar ») :

$$y=\frac{4}{3}x+\frac{1}{3}z+\frac{1}{3}t$$

La deuxième équation fournit alors une valeur de z qui, reportée dans la troisième équation, permet d'éliminer z :

$$z=\frac{13}{5}x+\frac{4}{5}t$$

Abū Kāmil récapitule alors les résultats obtenus :

$$\begin{array}{l} u=\frac{51}{15}x+\frac{6}{5}t \\ y=\frac{11}{5}x+\frac{3}{5}t \\ z=\frac{13}{5}x+\frac{4}{5}t \end{array}$$

Et la troisième équation (après élimination de y et z) est :

$$t=28x$$

Comme le système d'équations est homogène, on peut, comme le fait Abū Kāmil, poser $x=1$ pour obtenir une solution :

$$x=1, \quad y=19, \quad z=25, \quad t=28, \quad u=37$$

On remarque que, dans cette méthode, toutes les inconnues présentes dans les équations considérées par Abū Kāmil sont nommées, et chacune avec un nom distinct ; à chaque étape, la considération d'une nouvelle équation permet d'exprimer la valeur d'une des inconnues en fonction de celles qui n'ont pas encore été éliminées, et ainsi d'éliminer cette inconnue. Ce procédé revient à mettre le système d'équations sous forme triangulaire, et il était connu des contemporains d'Abū Kāmil.
Deuxième méthode : dans le but d'alléger les calculs qui, sans l'aide d'un symbolisme, sont certes

« difficiles », Abū Kāmil propose une deuxième méthode, qui consiste d'abord à choisir les inconnues de sorte que chacune des équations initiales ne contienne en elle-même qu'un petit nombre d'inconnues. Posons donc $\zeta = x + y + z + t$. Le système s'écrit alors :

$$\begin{cases} u = x + \frac{1}{2}(\zeta - x) \\ u = y + \frac{1}{3}(\zeta - y) \\ u = z + \frac{1}{4}(\zeta - z) \\ u = t + \frac{1}{5}(\zeta - t) \\ \zeta = x + y + z + t \end{cases}$$

L'élimination de u conduit Abū Kāmil à considérer en fait les équations suivantes, dans cet ordre :

$$\begin{cases} x + \frac{1}{2}(\zeta - x) = y + \frac{1}{3}(\zeta - y) \\ x + \frac{1}{2}(\zeta - x) = z + \frac{1}{4}(\zeta - z) \\ \zeta = x + y + z + t \\ x + \frac{1}{2}(\zeta - x) = t + \frac{1}{5}(\zeta - t) \end{cases}$$

La première équation ne contient alors que trois inconnues, et Abū Kāmil ne donne de nom qu'à ces trois inconnues x, y, ζ : la chose, le dinar et le dirham. Cette équation détermine une valeur de y, et l'inconnue y est ainsi éliminée (on a « ôté le nom du dinar ») :

$$y = \frac{3}{4}x + \frac{1}{4}\zeta$$

Une fois l'inconnue y éliminée, Abū Kāmil peut utiliser le nom « dinar » pour désigner une autre inconnue, en l'occurence z, et cela lui permet de prendre en considération la deuxième équation, qui détermine à son tour une valeur de z :

$$z = \frac{2}{3}x + \frac{1}{3}\zeta$$

L'élimination de y et z dans l'équation $\zeta = x + y + z + t$ détermine une valeur de t en fonction de x et ζ (ici Abū Kāmil ne prend pas la peine de « nommer » l'inconnue t), qui, reportée dans la dernière équation, conduit finalement à :

$$z = 73x,$$

ce qui mène facilement au résultat final. Donc dans cette méthode, Abū Kāmil ne « nomme » d'abord que les inconnues qui apparaissent dans la première équation ; l'élimination de l'une d'elle permet de réutiliser son nom pour désigner une autre inconnue dans l'équation suivante, et ainsi de suite. Les opérations de l'algèbre ne se rapportent toujours qu'à une seule équation à la fois, et c'est l'usage de ces opérations qui requiert de « nommer » les éléments de l'équation. C'est pourquoi il n'est pas strictement nécessaire de nommer les inconnues tant qu'elles n'apparaissent pas dans l'équation considérée, et qu'il est même possible d'utiliser deux fois le même « nom » pour désigner des inconnues différentes.

Problème 40. L'énoncé conduit au système d'équations suivant :

$$\begin{cases} v = x + \frac{1}{2}(y + z) \\ v = y + \frac{1}{3}(z + t) \\ v = z + \frac{1}{4}(t + u) \\ v = t + \frac{1}{5}(u + x) \\ v = u + \frac{1}{6}(x + y) \end{cases}$$

L'élimination de v conduit Abū Kāmil à considérer en fait les équations suivantes, dans cet ordre :

$$\begin{cases} x + \frac{1}{2}(y + z) = y + \frac{1}{3}(z + t) \\ x + \frac{1}{2}(y + z) = z + \frac{1}{4}(t + u) \\ x + \frac{1}{2}(y + z) = t + \frac{1}{5}(u + x) \\ x + \frac{1}{2}(y + z) = u + \frac{1}{6}(x + y) \end{cases}$$

Première méthode : Abū Kāmil nomme seulement les trois inconnues x, y, z : la chose, le dinar et le dirham. L'inconnue t n'est pas nommée, mais la première équation en donne immédiatement une valeur qui, reportée dans la deuxième équation et dans la troisième, permet d'éliminer t :

$$t = 3x - \frac{3}{2}y + \frac{1}{2}z$$

La deuxième équation donne alors une valeur de u qui, reportée dans la troisième et dans la quatrième, permet d'éliminer u :

$$u = x + \frac{7}{2}y - \frac{5}{2}z$$

La troisième équation, dans laquelle on a éliminé t et u, ne contient plus que les trois inconnues x, y, z, et elle donne la valeur de z :

$$z = (4 + \frac{4}{5})x - (2 + \frac{3}{5})y$$

Abū Kāmil récapitule les résultats obtenus :

$$\begin{array}{l} v = (3 + \frac{2}{5})x - \frac{4}{5}y \\ u = -11x + 10y \\ t = (5 + \frac{2}{5})x - (2 + \frac{4}{5})y \\ z = (4 + \frac{4}{5})x - (2 + \frac{3}{5})y \end{array}$$

Finalement, la quatrième équation, après élimination de u et z, donne :

$$61x = 47y$$

D'où la solution donnée par Abū Kāmil, en posant $y = 61$:

$$x = 47, \quad y = 61, \quad z = 67, \quad t = 83, \quad u = 93, \quad v = 111$$

On remarque que dans cette méthode, Abū Kāmil ne nomme que les inconnues qui apparaissent dans la première équation, sauf celle qu'il choisit d'éliminer au moyen de cette équation. Et ainsi de suite, à chaque étape, il considère une équation dans laquelle toutes les inconnues sont nommées, sauf celle qu'il choisit d'éliminer.

Deuxième méthode : la deuxième méthode n'est qu'une variante de la première, consistant à nommer l'inconnue que l'on cherche à éliminer à chaque étape, en réutilisant toujours le même nom (ici, Abū Kāmil utilise le nom « fals »). Cette méthode coïncide avec la deuxième méthode employée dans le problème 39. Comme les calculs sont ici identiques avec ceux de la première méthode et qu'il ne s'agit que d'une différence formelle, Abū Kāmil ne mène pas la résolution jusqu'au bout. C'est, en revanche, cette seule méthode qu'il emploie pour résoudre les deux problèmes suivants, 41 et 42, dont les énoncés sont :

$$\begin{cases} v = x + \frac{1}{2}y \\ v = y + \frac{1}{3}z \\ v = z + \frac{1}{4}t \\ v = t + \frac{1}{5}u \\ v = u + \frac{1}{6}x \end{cases}$$

$$\begin{cases} x + 1 = 2(y - 1) \\ y + 2 = 3(z - 2) \\ z + 3 = 4(t - 3) \\ t + 4 = 5(x - 4) \end{cases}$$

Problème 43. Il faut comprendre ainsi l'énoncé peu clair du problème 43 : trois hommes ont des fortunes respectives x, y, z et décident de redistribuer leur fortune en formant d'abord une cagnote, le premier donnant $\frac{1}{2}x$, le deuxième $\frac{1}{3}y$, le troisième $\frac{1}{6}z$. Ils se partagent ensuite la cagnote en en

prenant un tiers chacun, et l'énoncé affirme qu'ils ont alors chacun respectivement la moitié, le tiers et le sixième de la fortune totale. Ce qui se traduit en le système suivant :

$$\begin{cases} \frac{1}{2}x + \frac{1}{3}(\frac{1}{2}x + \frac{1}{3}y + \frac{1}{6}z) = \frac{1}{2}(x+y+z) \\ \frac{2}{3}y + \frac{1}{3}(\frac{1}{2}x + \frac{1}{3}y + \frac{1}{6}z) = \frac{1}{3}(x+y+z) \\ \frac{5}{6}z + \frac{1}{3}(\frac{1}{2}x + \frac{1}{3}y + \frac{1}{6}z) = \frac{1}{6}(x+y+z) \end{cases}$$

Première méthode : en fait, ce système est redondant. La troisième équation est comprise dans les deux premières. Abū Kāmil n'utilise donc que les deux premières équations. Il nomme x, y, z respectivement la chose, le dinar et le dirham. La première équation donne :

$$x = (2 + \frac{1}{3})y + (2 + \frac{2}{3})z$$

Puis, en remplaçant x par sa valeur dans la deuxième équation :

$$y = 13z$$

D'où une solution :

$$x = 33, \quad y = 13, \quad z = 1$$

Si l'on veut que la fortune finale de chacun soit exprimée par un nombre entier, on préfèrera la solution suivante donnée par Abū Kāmil :

$$x = 188, \quad y = 78, \quad z = 6$$

Deuxième méthode. Posons $u = \frac{1}{3}(\frac{1}{2}x + \frac{1}{3}y + \frac{1}{6}z)$, et $t = x + y + z$. Il faut alors résoudre le système suivant :

$$\begin{cases} \frac{1}{2}x + u = \frac{1}{2}t \\ \frac{2}{3}y + u = \frac{1}{3}t \\ \frac{5}{6}z + u = \frac{1}{6}t \\ u = \frac{1}{3}(\frac{1}{2}x + \frac{1}{3}y + \frac{1}{6}z) \\ t = x + y + z \end{cases}$$

Abū Kāmil nomme seulement t et u : la chose, le dirham. Les trois premières équations donnent les valeurs de x, y, z (inconnues qu'Abū Kāmil ne nomme pas) en fonction de t et u. On reporte ces valeurs dans la quatrième équation, d'où :

$$t = (6 + \frac{5}{7})u$$

Abū Kāmil pose $u = 42$ pour avoir une solution en nombres entiers. Dans cette méthode, on remarque que le système des cinq équations ci-dessus est redondant, et qu'Abū Kāmil n'utilise pas la cinquième équation. Cette méthode ressemble à la première méthode du problème 40 et ne diffère de la deuxième méthode du problème 39 que par le fait qu'Abū Kāmil ne nomme pas les inconnues x, y, z successivement éliminées au moyen des trois premières équations.

3.2 La *syncrisis* de Viète

Viète s'intéresse, au chapitre XVI du *De aequationum recognitione*[6] (1615), aux équations trinômes à une inconnue, de la forme $bx^m - x^n = c$ (en notations modernes, avec b et c positifs, et $n > m$). Il écrit le système suivant :

$$\begin{cases} by^m - y^n = c \\ bz^m - z^n = c \end{cases}$$

6. *Cf.* [320], p. 82–126.

Bien que les inconnues, suivant la notation adoptée par Viète, soient y et z, la *syncrisis* consiste en fait à exprimer b et c en fonction de y et z :

$$b = \frac{y^n - z^n}{y^m - z^m} = \frac{\sum^{n-1} y^k z^{n-1-k}}{\sum^{m-1} y^k z^{m-1-k}}$$

$$c = \frac{y^n z^m - z^n y^m}{y^m - z^m}$$

Ainsi, les coefficients de l'équation trinôme proposée s'expriment en fonction de deux de ses racines. Viète écrit en préambule à cette méthode, dite *syncrisis*, qu'elle permet d'étudier la constitution (*constitutio*) des équations trinômes : « Syncrisis est duarum aequationum correlatarum mutua inter se ad deprehendum earum constitutionem collatio ;... aequationum constitutio, syncrisi cognoscetur probe » ([320], p. 104).

Nous ne donnerons pas ici une description plus détaillée de son exposé. Mentionnons pourtant qu'en supposant seulement la racine y connue, et l'autre z inconnue, la *syncrisis* peut aussi s'interpréter comme la division par un binôme $(z - y)$ (*cf.* le traité *De aequationum emendatione*[7] de Viète, ch. III). Le quotient de l'équation $bz^m - z^n = c$ par le binôme $(z - y)$ est en effet $\sum^{n-1} y^k z^{n-1-k} = b \sum^{m-1} y^k z^{m-1-k}$. Dans ses premiers travaux sur les maximas et les minimas, Fermat s'est inspiré de cette interprétation[8].

On pourrait, rétrospectivement, en généralisant la méthode de *syncrisis*, obtenir une démonstration des formules données par Viète à la fin du *De aequationum emendatione*[9] exprimant les coefficients d'une équation au moyen des fonctions symétriques des racines. Par exemple, pour une équation cubique $x^3 + a_1 x^2 + a_2 x + a_3 = 0$, notons ses trois racines y_1, y_2, y_3. Au lieu de comparer deux équations, comme le fait Viète, comparons les trois suivantes :

$$\begin{cases} y_1^3 + a_1 y_1^2 + a_2 y_1 + a_3 = 0 \\ y_2^3 + a_1 y_2^2 + a_2 y_2 + a_3 = 0 \\ y_3^3 + a_1 y_3^2 + a_2 y_3 + a_3 = 0 \end{cases}$$

En considérant à présent a_1, a_2, a_3 comme les inconnues, il s'agit d'un système de trois équations linéaires. On peut le résoudre avec les formules de Cramer, et obtenir ainsi les formules exprimant a_1, a_2, et a_3 au moyen des fonctions symétriques des racines.

Bien sûr, Viète a pu aussi trouver ces formules en développant simplement le produit des binômes $(x - y_1)(x - y_2)(x - y_3)$. D'autre part, il ne semble pas avoir suspecté l'importance qu'auraient plus tard les fonctions symétriques des racines. Selon lui, la « nature » des équations réside, plutôt que dans les fonctions symétriques des racines, dans les termes d'une suite proportionnelle unique. Après avoir en effet ([320], ch. XVII, p. 108) énoncé une série de théorèmes sur les suites proportionnelles, Viète réécrivait (ch. XVIII, p. 111) les résultats du chapitre XVI (expression des coefficients des équations trinômes en fonction de deux de leurs racines) au moyen des termes d'une suite proportionnelle. Par exemple (ch. XVIII, théorème II, que nous traduisons en langage moderne) : si $bx - x^3 = c$ a deux racines distinctes y et z, alors elles sont le premier et le troisième terme d'une suite de trois proportionnelles p, q, r, et $b = p^2 + q^2 + r^2$, $c = p(q^2 + r^2)$ (rappelons qu'il s'agit du cas « irréductible » de Cardan). Mais l'expression des coefficients d'une équation au moyen des termes d'une suite proportionnelle unique ne peut se généraliser au delà des équations trinômes, sauf pour des équations particulières : une suite proportionnelle d'un nombre donné de termes ne renferme en effet que deux paramètres arbitraires – les deux premiers termes déterminent toute la suite – alors qu'une équation de degré n possède en général n coefficients arbitraires.

En fait, Albert Girard, dans son *Invention nouvelle en l'algèbre*, montre qu'il a senti à la fois les limites et la généralisation possible du projet de Viète, et que la clef pour comprendre la « nature » des équations réside dans les fonctions symétriques des racines, et non, comme le suggère Viète,

7. *Cf.* [320], p. 127–158.
8. *Cf.* [112], t. II, lettre de Fermat à Mersenne, 1er avril 1640, p. 187–188, et [161], p. 238.
9. *Cf.* [320], p. 158.

dans les termes d'une suite proportionnelle unique. Mais il ne remet pas en cause la méthode de *syncrisis* elle-même :

> Touchant François Viète, qui surpasse tous ses devanciers en l'algebre ; on peut voir en son traité (...) où il dit que telle syncrisis est pour trouver ou colliger la mutuelle comparaison de deux equations correlatives : et il oubliait pour parler generalement de dire és plans, et pour les solides, de trois correlatives, etc. [10]

On peut donc penser qu'Albert Girard avait précisément à l'esprit la généralisation que nous avons mentionnée ci-dessus. Cette généralisation suppose le théorème fondamental de l'algèbre, énoncé par Albert Girard (« toutes les équations d'algèbre reçoivent autant de solutions, que la dénomination de la plus haute quantité le démonstre » [11]) à la suite de Peter Roth [12]. Roth comptait aussi bien les racines réelles positives que négatives, Girard y ajoute les racines « enveloppées, comme celles qui ont des $\sqrt{-}$, comme des $\sqrt{-3}$, ... » [13]. Mais c'est simplement sur cette idée, présente dans la *syncrisis* de Viète, et reprise par Albert Girard, de la « mutuelle comparaison des équations corrélatives » que nous voulions insister en exposant cette méthode.

3.3 Le *Sìyuán yùjiàn* de Zhū Shìjié

Notre but n'étant pas ici de nous interroger sur les divergences possibles entre une tradition algébrique d'aire méditerranéenne et la discipline qui se dessine à travers les méthodes sino-japonaises de résolution de problèmes, nous interpréterons ces dernières méthodes dans les termes de l'algèbre [14] ; nous suivrons aussi ce principe dans la section 8.3 où nous aborderons les travaux d'un auteur japonais. Mais plutôt qu'une algèbre fondée sur une classification *a priori* des équations à une inconnue, assortie de méthodes de résolution algébrique, c'est l'algorithme de Ruffini–Horner qui forme le noyau des techniques de résolution de problèmes dans la tradition sino-japonaise. Le mathématicien chinois Zhū Shìjié fut ainsi l'auteur d'un ouvrage intitulé *Sùanxué qǐméng* (1299) contenant une méthode de résolution numérique essentiellement équivalente à la méthode de Ruffini-Horner [15].

Un ouvrage ultérieur de Zhū Shìjié, le *Sìyuán yùjiàn* (1303), contient une méthode d'élimination, appliquée à l'élimination d'une inconnue parmi deux dans un système de deux équations, de deux inconnues parmi trois dans un système de trois équations, et de trois inconnues parmi quatre dans un système de quatre équations. Cet ouvrage a été traduit en français par J. Hoe (une partie de cette traduction est parue dans [150]). Hoe a choisi d'utiliser, dans sa traduction, un langage « semi-symbolique », afin de faire sentir la concision de la langue chinoise. Cette concision, alliée à une méthode de calcul au moyen de bâtonnets sur un échiquier, rend en effet caduque la transcription en notation algébrique moderne, même pour le lecteur habitué à cette notation. Mais elle ne possède pas toute la puissance d'une notation symbolique. Cette concision se manifeste surtout dans la suite d'opérations ramenant l'énoncé d'un problème à un système d'équations. L'énoncé même du problème fait intervenir certaines grandeurs nommées, sur lesquelles le mathématicien chinois opère avec une grande facilité, puisque leurs noms, aussi bien que les opérateurs arithmétiques (addition, soustraction, multiplication, division), sont représentés chacun par un unique caractère chinois, c'est-à-dire une seule syllabe. Les équations obtenues sont écrites sur un « échiquier ». Chaque case contient un coefficient écrit au moyen d'une notation numérique figurée représentant les bâtonnets tels qu'ils étaient utilisés par le calculateur. Un tel échiquier peut contenir les coefficients de tous les monômes en deux inconnues x, y, en plaçant dans la case de coordonnées (m, n) le coefficient de $x^m y^n$. Nous verrons plus tard les difficultés rencontrées dans la représentation d'équations à plus de deux inconnues. Comme il n'est pas question d'expliquer ici de telles notations figurée et « semi-symbolique », nous utiliserons bien sûr la notation moderne. J. Hoe force notre attention

10. *Cf.* [122], f. F1v.
11. *Cf.* [122], f. E3v.
12. *Cf.* notre citation p. 21.
13. *Cf.* [122], f. Fv.
14. Quant à ces divergences, *cf.* [153], p. 229–230.
15. Cette méthode est décrite par exemple dans [153], chapitre 4, surtout p. 94–109. Sur l'histoire de la méthode de Ruffini-Horner, *cf.* [157], p. 63–72.

sur l'existence de deux étapes distinctes dans la résolution d'un problème : la « mise en équations des hypothèses », que l'on vient de décrire, est selon lui une méthode de raisonnement logique, contrairement à une deuxième étape, la méthode d'élimination, méthode « mécanique » qu'il qualifie même d'« algorithme » [16]. Le mérite du *Sìyuán yùjiàn* serait alors, non seulement d'améliorer des méthodes d'élimination qui existaient *déjà* au temps de Zhū Shìjié [17], mais aussi et surtout un mérite pédagogique, celui d'enseigner avec talent une certaine forme de raisonnement logique [18], au moyen d'exemples de problèmes bien choisis et assortis d'indications judicieuses « orientant ses élèves vers la solution la plus élégante ». Zhū Shìjié ne donne pas la résolution détaillée de ses problèmes, mais il indique quelles grandeurs choisir comme inconnues dans la mise en équations des hypothèses, et quelles sont les équations ainsi obtenues. Puis il laisse au lecteur le soin d'éliminer les inconnues, sans jamais décrire sa méthode, et il indique l'équation finale à une inconnue, ainsi qu'une de ses solutions. Bien qu'il ne décrive pas sa méthode d'élimination, il indique certaines équations intermédiaires. C'est sûrement grâce à ces étapes intermédiaires, ainsi qu'au moyen d'ouvrages ultérieurs, que J. Hoe est parvenu à reconstituer cette méthode (il ne s'explique pas sur ce point).

Il nous semble en fait inapproprié d'appeler une telle méthode « algorithme ». A la manière des mathématiciens du XVII$^{\text{e}}$ s., on peut la décrire comme étant constituée de trois « règles ». La première seule pourrait mériter le nom d'« algorithme ». Elle s'appelle *hùyǐn tōngfēn* et permet l'élimination d'une inconnue parmi deux dans un système de deux équations à deux inconnues (voire, dans certains cas, à trois ou quatre inconnues). Etant donné deux équations ($n \leq m$)

$$\begin{cases} (1) & A_m y^m + A_{m-1} y^{m-1} + ... + A_0 = 0 \\ (2) & a_n y^n + a_{n-1} y^{n-1} + ... + a_0 = 0 \end{cases}$$

on se ramène à deux équations plus petites en éliminant les plus hautes puissances de y de la manière suivante :

$$(3) \quad a_n(A_m y^m + A_{m-1} y^{m-1} + ... + A_0) - A_m y^{m-n}(a_n y^n + a_{n-1} y^{n-1} + ... + a_0) = 0$$

Le procédé est réitéré jusqu'à ce que l'on parvienne à une équation sans l'inconnue y. On pourrait nommer cet algorithme « algorithme d'élimination successive des plus hautes puissances de l'inconnue, sans fraction ». Il ressemble en effet à celui décrit par Newton. Pourtant, cette suite d'opérations, décrite par Hoe de manière générale [150] p. 133–134, n'est pas toujours suivie de manière fidèle dans les résolutions détaillées qu'il donne des quatre premiers problèmes du *Sìyuán yùjiàn*. La règle *hùyǐn tōngfēn* y est utilisée au moins à trois reprises [19]. Dans le troisième problème, Zhū Shìjié indique les deux équations linéaires en l'inconnue à éliminer que l'on doit obtenir au terme de l'avant-dernière étape de l'algorithme. Ces deux équations coïncident avec celles obtenues par Hoe, qui suit ici l'algorithme que nous venons de décrire. Par contre, dans le second problème, les deux équations intermédiaires indiquées par Zhū Shìjié laissent penser qu'il a plutôt fait l'élimination successive des *plus basses* puissances de l'inconnue. Ce procédé ramène les deux équations (1) et (2) ci-dessus à l'équation suivante :

$$(3') \quad \frac{1}{y}\left(a_0(A_m y^m + A_{m-1} y^{m-1} + ... + A_0) - A_0(a_n y^n + a_{n-1} y^{n-1} + ... + a_0)\right) = 0$$

C'est en effet ainsi que Hoe procède dans sa résolution du second problème (sans l'expliquer) [150] p. 153–155. Nous allons décrire de manière plus détaillée le quatrième problème, qui consiste en un système de quatre équations à quatre inconnues :

$$\begin{cases} (L_1) & -2y + x + z = 0 \\ (L_2) & -xy^2 + 4y + 2x - x^2 + 4z + xz = 0 \\ (L_3) & x^2 + y^2 - z^2 = 0 \\ (L_4) & 2y + 2x - w = 0 \end{cases}$$

16. *Cf.* [150] p. 156–157.
17. *Cf.* [150] p. 115–118.
18. *Cf.* [150] p. 48–49 et 113–114.
19. Une fois dans le second problème, une fois dans le troisième, une fois dans le quatrième.

On élimine d'abord z entre L_1 et L_2, ce qui est très facile puisque ces deux équations sont linéaires en z, et l'on obtient :

$$(L_5) \quad 2x^2 + (y^2 - 2y + 2)x - 12y = 0$$

Pour éliminer z entre L_1 et L_3, on utilise une autre règle de la méthode, appelée *tīfēn* et consistant à substituer dans L_3 la valeur de z donnée par l'équation L_1 linéaire en z, en calculant :

$$\begin{aligned} z &= 2y - x \\ z^2 &= (2y - x)^2 \end{aligned}$$

L'équation obtenue est divisible par y, et l'on obtient après simplification :

$$(L_6) \quad -3y + 4x = 0$$

L'inconnue z ayant ainsi été éliminée, il reste à éliminer x et y dans les trois équations L_4, L_5, L_6. On pourrait utilier *hùyǐn tōngfēn* pour éliminer x entre L_4 et L_5, mais Hoe préfère ici utiliser à nouveau *tīfēn*[20]. On obtient en tout cas, comme l'indique Zhū Shìjié :

$$(L_7) \quad 2y^3 + (-w - 8)y^2 + (6w + 28)y - w^2 - 2w = 0$$

L'élimination de x entre L_4 et L_6 est évidente (équations linéaires) et elle conduit à :

$$(L_8) \quad -7y + 2w = 0$$

Il reste maintenant à éliminer y entre L_7 et L_8. Or, l'élimination successive des plus hautes puissances ne conduit pas à l'équation intermédiaire linéaire en y indiquée par Zhū Shìjié. J. Hoe le remarque et montre comment obtenir cette équation intermédiaire par une autre suite d'opérations. Cette suite d'opérations ([150] p. 242–245) nous semble assez arbitraire. Elle conduit dans un premier temps à une équation qui est différente de celle indiquée par Zhū Shìjié, mais plus simple. Hoe la transforme ensuite en l'équation indiquée par Zhū Shìjié, par une dernière opération dont il affirme seulement qu'elle « n'est pas strictement nécessaire ». Pourtant, comme dans le deuxième problème, ici aussi l'élimination successive des *plus basses* puissances de y entre L_7 et L_8 conduit (après simplification, l'équation obtenue étant divisible par w) à l'équation intermédiaire de Zhū Shìjié

$$(L_9) \quad 8wy - 4w^3 + 3w^2 + 294w = 0,$$

puis à l'équation finale en w :

$$(L_{10}) \quad 4w^2 - 7w - 686 = 0.$$

Zhū Shìjié en donne une racine, $w = 14$.

Il est donc difficile de voir dans cette méthode un algorithme entièrement mécanique. Même dans la règle *hùyǐn tōngfēn*, on est libre de choisir entre l'élimination successive des plus hautes puissances et celle des plus basses puissances ; de plus, dans un cas comme dans l'autre, il existe encore une certaine liberté quant à l'ordre adopté pour mener les opérations, et même le résultat n'est pas unique. Le troisième problème du *Sìyuán yùjiàn* en donne un bel exemple. Nous avons dit que Zhū Shìjié et J. Hoe y effectuent l'élimination successive des plus hautes puissances. J. Hoe procède de la manière suivante[21].

$$\begin{cases} (L_1) & (-x + 1)y^2 + (1 + x + x^2)y - 2x^2 - x - 2 = 0 \\ (L_2) & y^3 + (-2x - 2)y^2 + (x^2 + 4x + 2)y - 2x^2 - 2x = 0 \end{cases}$$

L'opération $(-x + 1)L_2 - yL_1$ donne

$$(L_3) \quad (x^2 - x - 3)y^2 + (-x^3 - x^2 + 3x + 4)y + 2x^3 - 2x = 0$$

20. *Cf.* [150] p. 237–241 ; attention au fait que les pages 236 et 237, ayant été mal numérotées, sont imprimées dans le désordre.

21. *Cf.* [150] p. 203–211.

L'opération $(-x+1)L_3 - (x^2-x-3)L_1$ donne

$$(L_4) \quad (-x^2+3x+7)y + x^3 - 3x^2 - 7x - 6 = 0$$

L'opération $(-x^2+3x+7)L_1 - (-x+1)yL_4$ donne

$$(L_5) \quad (-2x^3+5x^2+11x+13)y + 2x^4 - 5x^3 - 15x^2 - 13x - 14 = 0$$

Et enfin l'opération $(-x^2+3x+7)L_5 - (-2x^3+5x^2+11x+13)L_4$ donne l'équation finale

$$(L_6) \quad 4x^4 - 24x^3 + 16x^2 + 24x - 20 = 0$$

dont Zhū Shìjié donne une racine $x = 5$. Mais dans l'opération menant à L_5, on a utilisé l'équation L_1, plutôt que l'équation L_3 plus récente et de même degré en y. Certes, le coefficient dominant de L_3 étant de degré 2 en x, le calcul aurait été plus compliqué, et c'est peut-être la raison pour laquelle Zhū Shìjié et J. Hoe ont préféré procéder ainsi. D'ailleurs, les deux calculs ne mènent pas au même résultat, et il est facile de vérifier que cette alternative aurait conduit à une équation finale de degré 5 en x, équation qui contient donc un facteur superflu par rapport à L_6. Enfin, l'équation finale L_6 n'est pas non plus délivrée de tout facteur superflu : le résultant de L_1 et L_2 est en effet $4x^2 - 16x - 20$, de degré 2 (on remarque qu'il a bien une racine égale à 5).

Nous avons parlé de trois règles. Les deux premières sont *hùyǐn tōngfēn* et *tīfēn*, la troisième consiste en une simple permutation des inconnues x, y, z, w, afin que les deux équations auxquelles on applique *hùyǐn tōngfēn* soient toujours des équations en x, y dont on élimine y.

Revenons enfin aux problèmes posés par la représentation des équations polynomiales sur l'échiquier avec les bâtonnets. Ce mode de représentation ne permet pas de noter un polynôme ayant un nombre quelconque d'indéterminées. Les coefficients des monômes $x^m y^n$ $(m, n \geq 0)$ peuvent tous être écrits dans les cases de coordonnées (m, n), et il en résulte des algorithmes arithmétiques particulièrement efficaces pour calculer dans l'anneau $\mathbb{Q}[x, y]$. Les coefficients des monômes $x^m z^n$ sont alors écrits, de manière analogue, dans les cases de coordonnées $(m, -n)$; ceux des monômes $w^m y^n$ sont écrits dans les cases de coordonnées $(-m, n)$ et ceux des monômes $w^m z^n$ dans celles de coordonnées $(-m, -n)$. Des cases exceptionnelles de l'échiquier servent à représenter les monômes yz, xw ou xyz. Ainsi, de nombreux monômes n'ont pas leur place sur l'échiquier, et il est assez remarquable que tous les problèmes proposés par Zhū Shìjié puissent être résolus sans eux. Les mathématiciens japonais du XVII$^{\text{ème}}$ siècle utiliseront une notation qui n'a pas cet inconvénient, mais qui profite à nouveau (comme Zhū Shìjié pour l'étape préalable de « mise en équations des hypothèses ») de la concision d'une langue écrite au moyen d'idéogrammes, tout en étant plus proche des notations symboliques algébriques utilisées en Europe à partir du XVI$^{\text{ème}}$ siècle [22].

3.4 La méthode de Fermat

Fermat écrivit vers 1650 un texte [23] dans lequel il décrivait une méthode générale d'élimination d'une inconnue dans un système de deux équations. Rien n'indique que Fermat ait eu connaissance d'une quelconque méthode d'élimination antérieure à la sienne. Il affirme lui-même la nouveauté de sa méthode dans sa lettre [24] à Carcavi du 20 août 1650. Soient (en notation moderne) les deux équations polynomiales :

$$\begin{cases} P_1(x) = 0 \\ P_2(x) = 0 \end{cases}$$

et supposons que $\deg P_1 \geq \deg P_2$. Soient a_1 et a_2 les termes constants de ces deux équations. Fermat calcule

$$P_3(x) = \frac{a_2(P_1(x) - a_1) - a_1(P_2(x) - a_2)}{x}$$

22. *Cf.* [153] p. 174–176, « La notation des polynômes à coefficients nominaux ».
23. cf. [112], tome I, p. 181–184, *Novus secundarum et ulterioris ordinis radicum in analyticis usus.*
24. cf. [112], t. 2, p. 284–287, lettre de Fermat à Carcavi, 20 août 1650.

qui est de degré strictement inférieur à $\deg P_1$, puis il répète ce procédé en considérant à présent les deux équations $P_2(x) = 0$ et $P_3(x) = 0$. Et ainsi de suite jusqu'à obtenir une équation de degré 1 en x. Il ne remarque pas qu'il existe des cas particuliers dans lesquels cela n'arrivera pas (par exemple si P_1 et P_2 sont tous deux divisibles par un facteur quadratique...). Lorsque Fermat applique cette méthode à deux équations en x, y pour éliminer x, il faut concevoir P_1 et P_2 comme étant des polynômes en x dont les coefficients sont des polynômes en y. Après la dernière étape, la valeur de x fournie par l'équation de degré 1 en x est substituée dans une des équations précédentes afin d'obtenir une équation finale qui ne contienne plus l'inconnue x :

> Pour ramener la recherche des deux inconnues à celle d'une seule, il faut reprendre une quelconque des équations précédentes ; la moins élevée est plus convenable pour que le degré de l'équation finale ne monte pas trop haut [...]. Au lieu de e [$= x$] on substituera sa valeur trouvée.

Fermat explique aussi comment sa méthode peut servir à éliminer des inconnues dans des systèmes de plus de deux équations, à plusieurs inconnues.

On voit transparaître quelques réflexes du Fermat des *Arithmétiques* de Diophante dans ce texte. Fermat insiste sur le fait que l'équation de degré 1 en x fournit une valeur de x rationnelle [25]. Il désigne lui-même le fondement (*fundamentum*) de sa méthode comme étant « la proportion de la double égalité » (P. Tannery a traduit ainsi *duplicatae aequalitatis analogia*), peut-être parce qu'il écrit d'abord la proportion :

$$\frac{a_1}{P_1(x) - a_1} = \frac{a_2}{P_2(x) - a_2}$$

afin de calculer $P_3(x)$. Le terme *duplicata aequalitas* désigne alors simplement le système des deux équations initiales, mais il est aussi utilisé en analyse diophantienne rationnelle (dans la traduction latine des *Arithmétiques* de Diophante par Bachet de Méziriac, qu'annota Fermat), pour désigner des systèmes de deux équations de la forme

$$\begin{cases} y_1^2 = a_1x^2 + b_1x + c_1 \\ y_2^2 = a_2x^2 + b_2x + c_2 \end{cases}$$

Si le lien avec l'analyse diophantienne reste implicite dans ce texte, Fermat y fait cependant référence explicite à la géométrie analytique [26]. En fait, comme le révèle la lettre à Carcavi, sa motivation était double : sa motivation première était la recherche d'une méthode d'élimination des « asymétries », c'est-à-dire la transformation d'une équation contenant des radicaux en une équation polynomiale, afin d'étendre sa méthode des tangentes à un ensemble plus vaste de courbes, comprenant les courbes algébriques irrationnelles. Sa seconde motivation est moins évidente ; elle concerne les « questions abondantes », et Fermat nous dit :

> Une recherche plus approfondie a montré que l'on pouvait déduire de là une méthode très remarquable pour la pleine et parfaite connaissance des lieux en surface.

Fermat illustre ceci par le problème suivant : placer un cercle sur un cône donné ayant pour base une section conique plane. En fait, comme l'a remarqué Jean Itard [27], Fermat a commis une erreur

25. Il est remarquable que Rolle (*cf. infra* p. 75), puis Lagrange (*cf.* [189], § 11), eux aussi intéressés par l'analyse diophantienne rationnelle, aient plus tard insisté à nouveau sur l'importance de cette équation de degré 1 en x.

26. R. Rashed [261] a déjà montré l'influence de l'analyse diophantienne dans d'autres travaux de Fermat en géométrie analytique.

27. cf. [161], p. 220. Pour montrer en quoi consiste cette erreur, restituons son raisonnement sous une forme plus générale. Supposons un système de deux équations à deux inconnues :

$$\begin{cases} P_1(x, y) = 0 \\ P_2(x, y) = 0 \end{cases}$$

Ajoutons à ce système une troisième équation $P_3(x, y) = 0$, qui soit en accord avec les données du problème. Comme le système contiendra alors plus d'équations que d'inconnues, Fermat qualifie un tel problème d'« abondant ». En éliminant d'abord y entre P_1 et P_3 puis entre P_2 et P_3, on obtiendra deux équations $Q_1(x) = 0$ et $Q_2(x) = 0$. L'élimination de x devrait alors conduire à une équation de degré 1 en x. Ce raisonnement est faux, car l'équation P_3 est particulière, et dans le cas considéré par Fermat, Q_1 et Q_2 ont un facteur commun : on se trouve dans le cas mentionné plus haut, dans lequel il est impossible d'obtenir une équation de degré 1 en x par la méthode de Fermat. Fermat ne donne cependant aucun détail ni calcul pour le problème auquel il propose d'appliquer cette méthode, et il n'a donc pas vu son erreur.

en affirmant que ce problème pouvait se ramener à une équation de degré 1 en x. Cette erreur repose simplement sur le fait que, lorsque P_1 et P_2 ont un facteur commun de degré supérieur à 1, la méthode de Fermat exposée ci-dessus ne conduit pas au résultat escompté. L'objectif visé ici par Fermat était donc illusoire.

On comprendra le vain espoir caressé par Fermat si l'on pense qu'il put être inspiré par l'analyse diophantienne rationnelle. La « méthode de la double équation » en analyse diophantienne dut lui servir de modèle. Soit en effet la double équation [28] :

$$\begin{cases} r^2x^2 + bx + 1 = y^2 \\ s^2x^2 + ex + 1 = z^2 \end{cases}$$

Comme l'on a deux équations seulement, et trois inconnues, on peut imposer une condition supplémentaire. La « méthode de la double équation » utilisée par Fermat consiste à poser $y - z = \lambda x$. On parvient alors à trouver une nouvelle équation $y^2 = Ax^2 + Bx + C$, en combinant les trois précédentes. Le choix du paramètre λ permet de garantir que $A = r^2$, auquel cas, en comparant les deux valeurs obtenues pour y^2 :

$$r^2x^2 + bx + 1 = Ax^2 + Bx + C$$

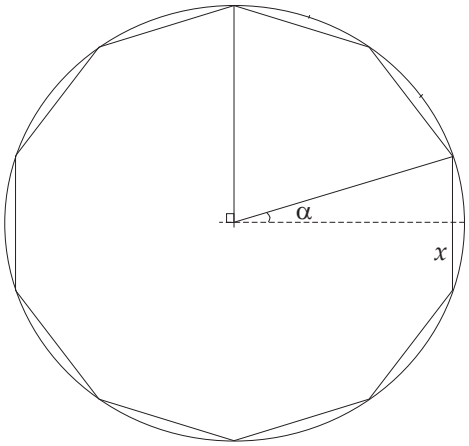
puis en simplifiant, on obtient :

$$bx + 1 = Bx + C$$

D'où une solution rationnelle. Résumons : cette méthode consiste, d'abord à imposer une condition supplémentaire au système pour le rendre « déterminé », puis à éliminer y et z pour obtenir une équation finale en x. Le choix arbitraire de la condition supplémentaire permet de plus de garantir l'élimination de la plus haute puissance de x (terme en x^2). Cette suite d'opérations est en tout point semblable à celle que nous avons retranscrite ci-dessus dans le problème 38 d'Abū Kāmil. Quant à la méthode générale d'élimination proposée par Fermat, répétons-le, dans le cas de l'élimination d'une inconnue x de deux équations, elle consiste à former une suite d'équations de degrés décroissants en x, en éliminant, à chaque étape, le terme constant de la dernière équation au moyen de l'équation précédente, jusqu'à l'obtention d'une équation linéaire en x, d'où une expression rationnelle de la valeur de l'inconnue x, qui, substituée à x dans une des équations précédentes, donne alors une équation finale libérée de l'inconnue x.

Rien ne semble indiquer que Hudde ait connu la méthode de Fermat. En fait, la méthode de Hudde coïncide avec celle de Fermat *via* le changement de variable $x \mapsto \frac{1}{x}$, à un facteur près dépendant des coefficients des équations initiales. Mais Hudde a évité l'écueil dans lequel Fermat était tombé. A cet égard, il sera instructif d'observer une autre application de la méthode de Hudde. Hudde mentionne, à la fin de sa lettre *Prima epistola de reductione aequationum* (p. 506), l'application de sa méthode à la construction des polygones réguliers, et il renvoie à ce sujet aux *Exercitationum mathematicarum libri quinque* de Frans van Schooten (publiés en 1657, année même de rédaction de la *Prima epistola*) [29]. Schooten y utilise en effet cette méthode d'élimination qu'il attribue à Hudde et l'applique à la construction des polygones réguliers. Mais ce n'est pas seulement la méthode d'élimination qu'il attribue à Hudde ; il laisse aussi entendre que Hudde lui-même a travaillé sur les polygones réguliers. L'équation donnant la longueur x du côté d'un polygone régulier inscrit dans un cercle de rayon 1 est obtenue en cherchant le plus grand commun diviseur de deux équations en x. On a donc là un exemple de « question abondante » (au sens de Fermat), dont Hudde et Schooten se gardent bien de chercher une solution rationnelle au moyen d'une équation finale de degré 1 (comme Fermat prétendait le faire). Par exemple, soit un décagone de côté x, inscrit dans un cercle de rayon 1 :

28. Pour un exposé plus détaillé de la méthode de la double équation chez Fermat, *cf.* [113], p. 21–22 et 162–194.
29. *Cf.* [281], section XXI, p. 464–475.



Soit $\alpha = \frac{\pi}{10}$. La corde de l'angle 2α est $x = 2\sin\alpha$. Les relations $4\alpha = \frac{\pi}{2} - \alpha$ et $6\alpha = \frac{\pi}{2} + \alpha$, traduites en termes de sinus, donnent alors les deux équations :

$$\begin{cases} x^4 - 4x^2 - x + 2 = 0 \\ x^6 - 6x^4 + 9x^2 - x - 2 = 0 \end{cases}$$

Hudde applique à ces deux équations sa méthode d'élimination, jusqu'à obtenir une équation de degré minimal, ici $x^2 + x - 1 = 0$, dont l'une des racines est la longueur du côté du décagone convexe. Schooten conclut cette section en écrivant : « ubi tandem advertandum, methodum hanc, quâ duae aut plures aequationes differentes, unam tantum eandemque quantitatem incognitam habentes, mutuâ reductione simplicissimam aequationem subministrant, ad aliorum quoque difficiliorum Problematum aequationes simplicissimas investigandas, inservire posse ».

3.5 Méthodes, règles, formules : l'élimination algébrique au XVII^{e} siècle

Méthodes, règles, formules, tables : c'est ainsi que ces auteurs (Fermat, Hudde, Newton, Leibniz) décrivent eux-mêmes l'objet de leurs recherches. Rétrospectivement, peut-on parler d'une « théorie de l'élimination », au sens où l'entendront par exemple Sylvester puis Kronecker au XIX^{e} siècle ? Ces méthodes, celle de Fermat, celle de Hudde, celle de Newton et celle dite « de conjonction des inverses » de Leibniz, algorithmes particulièrement simples, ne demandent aucune démonstration [30] : il est évident qu'elles conduisent au résultat cherché, une équation qui ne contienne plus l'inconnue à éliminer. Ces auteurs ne s'interrogent pas, par exemple, sur l'unicité de l'équation finale. En fait, les trois méthodes conduisent à des résultats différents. Comparons-les au résultant écrit sous forme d'un déterminant, dans le cas de deux polynômes de degré 3 :

$$\begin{cases} Ay^3 + By^2 + Cy + D = 0 \\ ay^3 + by^2 + cy + d = 0 \end{cases}$$

Posons :

$$R = \begin{vmatrix} A & B & C & D & 0 & 0 \\ 0 & A & B & C & D & 0 \\ 0 & 0 & A & B & C & D \\ a & b & c & d & 0 & 0 \\ 0 & a & b & c & d & 0 \\ 0 & 0 & a & b & c & d \end{vmatrix}$$

30. La méthode développée par Leibniz et conduisant directement au résultant est particulière : comme il réduit la question à un système d'équations linéaires, il faut, en arrière-plan, un calcul des déterminants. D'autre part, nous l'avons vu, Leibniz donne une *démonstration* de cette méthode au moyen de la notion de diviseur commun.

Alors la méthode de Fermat conduit au résultat suivant [31] :

$$d^2.\begin{vmatrix} C & D \\ c & d \end{vmatrix}.\begin{vmatrix} B & C & D \\ c & d & \\ b & c & d \end{vmatrix}.R = d^2(Dc - dC)(cCd - Dc^2 - Bd^2 + dDb)R$$

La méthode de conjonction des inverses de Leibniz au résultat :

$$(Da - dA)R$$

Et la méthode de Hudde au résultat suivant :

$$\frac{(Ab - aB)^2}{a.\begin{vmatrix} A & B & C & D \\ 0 & A & B & C \\ a & b & c & d \\ 0 & a & b & c \end{vmatrix}^2}R$$

La méthode d'élimination successive des plus hautes puissances de l'inconnue décrite par Newton n'introduit de facteur superflu qu'au dénominateur ; le numérateur du résultat obtenu est bien R, comme nous le montrerons plus loin pour deux équations de degré quelconque (*cf. infra* p. 188). Cependant la diversité des « règles » d'élimination décrites par Newton dans son *Arithmetica universalis* laisse dans le doute quant à la manière dont il a calculé les formules des résultants $(2, 2)$, $(3, 2)$, $(4, 2)$ et $(3, 3)$ données dans le même ouvrage. Les formules littérales qu'il a données sont bien délivrées de tout facteur superflu. Mais Newton n'a pas expliqué pas comment il les a trouvées. Est-ce, comme nous le pensons, en utilisant la méthode d'élimination successive des plus hautes puissances, et en ne gardant que le numérateur du résultat obtenu ? C'est peut-être ce doute qui conduit Euler, en 1748, à chercher une autre méthode d'élimination que celle de Hudde. Euler écrit en effet que sa nouvelle méthode est « celle dont Mr. Newton paraît s'être servi ». En fait, la nouvelle méthode d'Euler (*cf. infra* p. 105 méthode n°1) n'est autre que la méthode de conjonctions des inverses de Leibniz, qui présente en effet, en degré 3, un facteur superflu $(Da - dA)$ beaucoup plus simple que celle de Hudde (et Euler remarque bien ce facteur superflu).

Bien que des méthodes existent, l'on ne peut donc encore parler d'une théorie de l'élimination au sens strict. Les algébristes du XVIII[e] siècle qui fonderont la théorie de l'élimination (Cramer, Euler, Bézout) se poseront désormais le problème de touver une équation finale définie par certaines propriétés, en plus du fait qu'elle ne contienne plus l'inconnue à éliminer. Dans la méthode de Cramer-Euler [32], à chaque racine x_0 de l'équation finale résultant de l'élimination d'une inconnue y parmi deux, devra correspondre une solution (x_0, y_0) du système formé par les deux équations initiales. Dans les méthodes développées par Bézout, l'équation finale devra être une équation de degré minimum en les inconnues restantes (c'est-à-dire libérée de tout « facteur superflu »). Ces propriétés, qui garantissent facilement l'unicité de l'équation finale, devront être démontrées.

Pourtant, souvenons-nous des travaux antérieurs que nous avons mentionnés : l'élimination en analyse diophantienne, l'algèbre linéaire, le chapitre 10 de l'*Ars magna* de Cardan repris par Stévin et Albert Girard, la *syncrisis* de Viète, les recherches de Faulhaber, la méthode des coefficients indéterminées chez Descartes. A travers ces quelques exemples, on voit que, malgré l'absence d'une méthode générale pour résoudre des systèmes d'équations quelconques, un langage existait déjà, avant même le XVII[e] siècle, qui permettait de décrire de tels systèmes, ainsi que des méthodes pour résoudre certaines classes particulières de problèmes. Bien plus, existaient aussi certains énoncés supposant une classification générale des problèmes à plusieurs inconnues ; classification dont le principe n'était pourtant pas encore fondé démonstrativement. Aussi bien Descartes (dans sa *Géométrie* [33], en 1637) que Faulhaber [34], hors de tout contexte géométrique en 1631, affirment que lors de la mise

31. *Cf.* détail des calculs suivants en l'annexe C.
32. *Cf. infra* p. 106 méthode n°2.
33. *Cf.* [90], t. VI, p. 372–374.
34. *Cf. supra* p. 37.

en équations d'un problème, l'on doit trouver autant d'équations que l'on a posé d'inconnues, sans quoi le problème ne serait pas déterminé ; Albert Girard, en 1629, formule des idées semblables, sans pourtant parler de plusieurs « équations », mais seulement de plusieurs « conditions » conduisant à une équation unique, et en y ajoutant la considération de problèmes « excédents » :

> Les conditions d'une proposition achevées, on vient à une équation, que si il n'y a pas assez de conditions pour amener le tout à l'équation, et que les nombres Algébriques ayent entre eux les propriétés et conditions requises, alors la solution sera défaillante, et recevra autant de solutions qu'on voudra, (...).
> Que si on peut résoudre la proposition sans se servir de toutes les conditions, elle sera excédente, et faut retrancher la dernière condition, si elle répugne : Que si finalement la proposition peut faire parvenir à une équation, la proposition sera pleine et entière ; (...) (f. D1r)

On peut comparer de tels discours à la classification des problèmes jadis offerte par le géomètre Ibn Sinān (909–946), reliant le nombre de « solutions » au nombre des « hypothèses » d'un problème [35].

A ces énoncés généraux s'ajoutent même, au moins chez Faulhaber [36] et Descartes, une méthode d'élimination imparfaite. Descartes ne parle pas de la méthode d'élimination qu'il utilise dans la *Géométrie*, sinon par allusion [37]. Toutefois, dans un manuscrit, que Descartes recommandait à certains de ses correspondants, en 1638, comme une « introduction à sa Géométrie », mais dont l'auteur est resté anonyme (« un jeune Hollandais »), on trouve exposée une méthode qui suffit lorsque les équations sont de degré inférieur à 4 en l'inconnue à éliminer, et que la valeur de cette inconnue peut alors s'exprimer par des radicaux :

> on doit (...) par le moyen d'ycelles équations réduire toutes ces lettres en une seule, qui porte la solution du problème. Et pour venir à bout de ces réductions, il est besoin de considérer sy, par une équation, ou par la comparaison de deux ou plusieurs, en les adjoustant ou soustrayant l'une de l'autre, on ne pourra faire cognoistre une lettre. *Et sy cela ne se peut, il faut venir à l'extraction de la racine pour en trouver une ; puis après, on doit oster cette lettre de l'une des autres équations, et en son lieu mettre la valeur trouvée ; et ainsi on sera quitte d'une lettre inconnue.* (...) Et ainsi le problème sera résolu. [38]

C'est cette méthode que Descartes utilise effectivement dans la *Géométrie*, pour déterminer l'intersection d'une courbe algébrique quelconque et d'un cercle, dans sa méthode des sécantes [39].

Par comparaison, dans les années 1650, Hudde et avant lui Fermat, marquent un tournant décisif dans l'histoire de l'élimination algébrique, avec la découverte d'un algorithme pour l'élimination d'une inconnue dans un système de deux équations de degrés quelconques. C'est pourquoi nous avons fait allusion à une « histoire courte » de l'élimination en algèbre, qui commence dans la deuxième moitié du XVII^{e} siècle et prend seulement son envol avec les travaux de Newton, Leibniz et Michel Rolle. Nous allons à présent nous interroger sur les raisons de ce commencement.

3.6 L'ordre adopté dans la résolution des problèmes

Pour donner une réponse, au moins provisoire, à la question de l'origine historique de la théorie de l'élimination, il nous faut de nouveau porter notre attention sur les trois premières occurences,

35. *Cf.* [160], p. 24–30.
36. *Cf. supra* p. 33.
37. Ainsi p. 373–374 : « Et on peut tousjours reduire ainsi toutes les quantités inconnuës a une seule, lorsque le Problesme se peut construire par des cercles et des lignes droites, ou aussy par des sections coniques, ou mesme par quelque autre ligne qui ne soit que d'un ou deux degrés plus composée ». La méthode d'élimination à laquelle fait allusion Descartes semble donc se restreindre à des équations (à plusieures inconnues) de degrés inférieurs à 4. La lettre de Descartes à Carcavi du 17 août 1649 ([90], t. V, p. 392–393), dans laquelle il exprime son incompréhension de la méthode d'élimination des radicaux (*asymetria*) de Fermat, témoigne aussi du fait que Descartes n'avait pas de méthode *générale* d'élimination des inconnues (la méthode d'élimination des radicaux de Fermat repose en effet sur sa méthode générale d'élimination des inconnues ; plus généralement, toute méthode d'élimination des inconnues s'applique facilement à l'élimination des radicaux).
38. Nous soulignons, *cf.* [90], t. X, p. 673.
39. *Cf.* [90], t. VI, p. 414.

indépendantes, à notre connaissance, d'une méthode d'élimination qui se veut générale, dans les travaux des mathématiciens du XVII^{e} siècle.

Nous avons vu Faulhaber proposer une telle méthode – certes sans donner l'explication qui seule pourrait lui faire mériter le nom d'algorithme. Il utilisait cette méthode pour résoudre un système d'équations issu de l'application de la méthode des coefficients indéterminés à la résolution algébrique d'une équation de degré supérieur ou égal à quatre. Il l'utilisait aussi dans la recherche de formules sur les sommes de séries de puissances. Quant à la méthode de Fermat, on l'a vu naître d'un intérêt pour certains problèmes de géométrie analytique, sous l'influence d'un esprit versé dans les *Arithmétiques* de Diophante. Enfin, si la *Prima epistola de reductione aequationum* de Hudde semble émerger du cercle d'idées de Faulhaber et Descartes, sur l'application en algèbre de la méthode des coefficients indéterminés, c'est la même année que Frans van Schooten publie dans ses *Exercitationum mathematicarum libri quinque* l'application de cette méthode à la construction des polygones réguliers, laissant d'ailleurs entendre que Hudde lui-même a participé à l'étude de ce problème. Il nous est donc impossible de dire quelle fut l'inspiration originale de Hudde : la méthode des coefficients indéterminés, ou bien la résolution de problèmes géométriques telle la construction des polygones ?

Le contraste entre les recherches menées par ces trois auteurs, entre résolution algébrique des équations, analyse diophantienne rationnelle, géométrie analytique et construction de polygones réguliers, montre que la conception d'une méthode d'élimination ne peut s'expliquer par un rapport particulier de l'algèbre à une autre discipline. D'ailleurs, les algébristes arabes s'étaient déjà penchés sur l'analyse diophantienne rationnelle dès le IX^{e} siècle (Abū Kāmil), ainsi que sur la construction des polygones réguliers, et la méthode de résolution algébrique de l'équation de degré 4 de Faulhaber n'est qu'une variante de celle offerte par l'algébriste italien Ferrari dans la première moitié du XVI^{e} siècle. Bien que toute application de l'algèbre à un domaine renforce souvent celle-ci en retour, en lui prêtant le substrat intuitif qui lui manque, la théorie de l'élimination est proprement algébrique, dans le sens où elle se rapporte directement aux concepts de l'inconnue et de ses puissances, que l'on « élimine ». Pourtant, il ne faudra pas négliger ce renforcement, comme principe d'explication, pour rendre compte, par exemple, de l'histoire du « théorème de Bézout » : nous avons vu qu'il naît d'un rapport particulier de l'algèbre à la théorie des courbes, probablement au sein de la critique du système cartésien de classification des courbes par « genre ».

Serait-ce donc un obstacle technique, qui aurait retardé l'élaboration d'un algorithme d'élimination, jusqu'à ce que des mathématiciens aussi adroits qu'un Fermat l'aient surmonté (que l'on pense aux remarques précitées de Faulhaber ou A. Girard [40]) ? Non plus. En fait, l'algorithme d'élimination de Hudde coïncide précisément avec l'algorithme pour la recherche du plus grand commun diviseur de deux polynômes, déjà employé par Stévin en 1585 pour simplifier des équations de la forme $P(x) = Q(x)$ en divisant chaque membre par le plus grand commun diviseur de P et Q. D'ailleurs, il ne diffère pas essentiellement de l'algorithme euclidien de la recherche du plus grand commun diviseur de deux entiers. Pourtant il ne faut pas négliger la difficulté technique, comme principe d'explication, ici aussi, pour le développement ultérieur de la théorie de l'élimination : tant que la matière résiste, la recherche se poursuit. Nous le verrons dans le courant de recherches menant à la démonstration du théorème de Bézout, se heurtant sans cesse au problème des « facteurs superflus ». Que l'on pense aussi à la critique par Leibniz de l'approche de Hudde, trop « laborieuse ». L'algorithme euclidien du plus grand commun diviseur fit l'objet de recherches combinatoires profondes de la part de Leibniz, justement pour contourner les « calculs immenses », dont seuls les résultats sont consignés dans des « tables » dans l'*Epistola prima* de Hudde, pour étudier différents cas de l'équation générale de degré 6 au moyen d'une « équation réduite » de degré 15 (elle-même obtenue par la méthode d'élimination). Leibniz écrit en 1675 :

> Je conviens que Hudde ait mis au jour, avec beaucoup d'esprit, une adresse remarquable et un travail immense, une table à l'aide de laquelle, en la parcourant entièrement, puisse être connu si une équation rationnelle quelconque de degré 5 ou 6 a ou non des diviseurs ; mais qui tiendra tant à une quelconque équation qu'il mènera un tel examen au prix d'un travail incroyable, et qu'en sera-t-il, après, pour les équations plus élevées, quand la table

40. Faulhaber *supra* note 49 p. 33, Albert Girard p. 28.

ainsi construite sera d'une longueur immense... [41]

C'est, nous semble-t-il, pour résoudre cette difficulté calculatoire que Leibniz chercha, plutôt qu'un algorithme d'élimination, une *formule* pour l'équation finale résultant de l'élimination, et qu'il étudia à cet effet les propriétés d'homogénéité de l'équation finale [42].

Revenons à notre question initiale : pourquoi des méthodes d'élimination apparaissent-elles au XVII^{e} siècle ? L'expliquerait-on par l'avènement d'une nouvelle algèbre, assise sur le concept de système d'équations à plusieurs inconnues, voire d'une algèbre symbolique, qui ne se refuse plus à écrire données, inconnues, et coefficients indéterminés ou inconnues auxiliaires, au moyen de lettres et d'exposants ? Attention, l'explication risque d'être circulaire : comment caractériser une algèbre assise sur des concepts et symboles nouveaux, sinon par les méthodes dont rendent compte les symboles, et par la classification au sein de laquelle s'organisent les concepts ? D'ailleurs les algébristes arabes utilisaient parfois, dès le IX^{e} siècle, plusieurs inconnues qu'ils désignaient par des noms différents, et, au XVI^{e} siècle, l'allemand Stifel notait les inconnues au moyen des lettres de l'alphabet [43], en 1544. Stévin et Albert Girard parlaient de « postposées quantités » pour désigner plusieurs inconnues et leurs puissances. A partir du IX^{e} siècle, de nombreux traités d'algèbre contiennent un chapitre d'analyse diophantienne, et donc, en particulier, des systèmes d'équations. Enfin, Cardan utilisait un coefficient indéterminé pour décrire la méthode de résolution de l'équation de degré 4 due à Ferrari. Il faut donc rejeter encore une telle explication, à moins de préciser le sens d'une « nouvelle algèbre ».

Les problèmes traités par Faulhaber, Fermat et Hudde sont des problèmes déjà anciens. Ce qui est nouveau, c'est seulement l'ordre adopté dans la résolution des problèmes. Ainsi pour la construction du décagone, tout d'abord la mise en équations du problème, puis l'élimination, qui conduit à une équation à une inconnue, de degré minimal, résoluble par les algorithmes classiques. Fermat entend modeler la résolution de tout problème sur l'ordre adopté en analyse diophantienne rationnelle : tout d'abord, la recherche d'équations supplémentaires pour rendre le problème déterminé (voire « abondant », nous dit Fermat), puis l'élimination, qui conduit à une unique équation. La distinction d'étapes dans l'analyse, qui n'étaient pas conçues séparément par les anciens algébristes, s'illustre particulièrement dans la *Géométrie* de Descartes (livre I). Une première étape consiste en la mise en équations du problème. « Après cela, écrit Descartes, il se faut servir par ordre de chacune des équations qui restent, soit en la considérant toute seule, soit en la comparant avec les autres, pour expliquer chacune de ces lignes inconnues, et faire ainsi, en les démêlant, qu'il n'en demeure qu'une seule ». Puis vient la résolution de l'équation finale par les méthodes classiques. Les deux premières étapes de l'analyse – mise en équations, élimination – sont confondues dans l'algèbre classique (y compris chez Viète). Pendant longtemps encore (y compris chez Descartes, la *Géométrie*, livre III) le noyau d'un traité d'algèbre consistera en des algorithmes de résolution (algébriques, géométriques, arithmétiques, numériques) des équations à une inconnue, et leur démonstration. Mais la conception distincte d'un système d'équations issu d'une première étape de l'analyse, est le préalable de la mise en algorithme de cette deuxième étape qu'est l'élimination. [44] La lente évolution de l'algèbre ne montrera l'achèvement de cette prise de conscience qu'à la fin du XVIII^{e} siècle, quand l'élimination acquiert le statut de « théorie ». Au moins un auteur, Bézout, dans sa *Théorie générale des équations algébriques*, choisira alors de consacrer à l'élimination un traité général d'algèbre, à l'exclusion de toute méthode de résolution des équations à une inconnue.

41. *Cf.* [210], p. 698 : « Fateor, Huddenium summo ingenio ac miris artibus immensoque labore tabulam inde eruisse, cujus ope omnia pertentando sciri possit, an aequatio aliqua rationalis quinti sextique gradus sit divisibilis ; sed quis tanti putabit ullam aequationem, ut per tot examina incredibili labore ducat, et quid futurum putamus in altioribus, ubi tabula ipsa ad hunc constructa modum immensae magnitudinis futura esset... »

42. *Cf. supra* p. 53–53. Son approche, originale, est d'ailleurs semblable à celle qu'il adopta dans ses recherches sur la résolution algébrique des équations de degré supérieur à 4 : il s'agissait alors de faire une conjecture sur la forme des racines.

43. *Cf.* [151]. Il est probable que l'introduction d'un symbolisme pour noter les diverses inconnues ait accéléré le développement de l'algèbre à partir de la deuxième moitié du XVI^{e} siècle.

44. De plus de nouvelles méthodes algébriques, comme la méthode des coefficients indéterminés utilisée par Faulhaber, ainsi que la *syncrisis* de Viète, en introduisant des inconnues auxiliaires, causent un détour dans l'ordre de l'analyse. Autant d'occasions nouvelles d'avoir recours à des techniques d'élimination.

Chapitre 4

Les « arbres » de Rolle

Le *Traité d'algèbre* de Rolle (1690) compte quatre « livres ». Les deux premiers concernent la résolution des systèmes d'équations linéaires et la fameuse « méthode des cascades » de Rolle, méthode de séparation des racines réelles d'une équation algébrique à une inconnue. Le livre III est consacré à la théorie de l'élimination. Quant au livre IV, composite, il nous offre des renseignements sur les motivations de l'auteur ; c'est pourquoi nous commencerons par commenter ce dernier livre, d'ailleurs intitulé « dissertation sur l'algèbre ». Il joue un triple rôle : justifier l'ordre adopté dans la constitution du traité d'algèbre, défendre les prétentions de son auteur engagé dans une ou plusieurs controverses impliquant Tschirnhaus, Ozanam et surtout Prestet [1], et exposer ses recherches sur la résolution algébrique des équations.

Rolle distingue deux parties de l'algèbre : l'élimination algébrique dont « le succès n'est pas mis en doute » et la résolution algébrique des équations à une inconnue où, au contraire, « l'on y a rencontré des difficultés ». Cela suffit à expliquer la composition d'un traité d'algèbre dans lequel l'exposé des méthodes d'élimination (livres I et III) est entrecoupé par celui d'une méthode de séparation des racines (la « méthode des cascades » au livre II) dont le but ultime est bien sûr le calcul approché des racines d'une équation à une inconnue.

C'est aussi l'occasion d'un des reproches adressés par Rolle à Prestet. Celui-ci avait proposé, pour le calcul des racines d'une équation, la méthode d'approximation par dichotomie [2] ; Rolle lui reproche de n'avoir pas précisé comment choisir les bornes de l'intervalle initial devant contenir la racine cherchée [3]. Sans vouloir exposer en détail les problèmes posés par la « méthode des cascades » de Rolle, rappelons-en simplement le principe. Rolle calcule les dérivées successives du polynôme de degré n dont il cherche les racines (la dérivée i-ème définissant la $(n-i)$-ème « cascade »). La méthode des cascades est une méthode par récurrence : la connaissance des racines réelles de la dérivée $(i+1)$-ème permet la séparation des racines réelles de la dérivée i-ème, qui permet à son tour leur calcul approché (par dichotomie). La séparation des racines repose sur le fait qu'entre deux racines réelles distinctes de la dérivée i-ème existe au moins une racine de la dérivée $(i+1)$-ème. Les racines de la dérivée $(i+1)$-ème déterminent donc des intervalles contenant chacun au plus une racine de la dérivée i-ème, et c'est sur ces intervalles que l'on applique la méthode d'approximation par dichotomie. Il reste à déterminer la borne supérieure (la « grande hypothèse ») de l'intervalle contenant la racine la plus grande. Soit $P(x) = x^n + a_{n-1}x^{n-1} + ... + a_0$ un polynôme à coefficients entiers, Rolle affirme à cet effet que

$$A + 1 = \left(\max_{i|a_i<0} |a_i| \right) + 1$$

est un majorant des racines. On peut le vérifier en remarquant que, pour $x > A + 1$,

$$P(x) \geq x^n - Ax^{n-1} - ... - Ax - A = \frac{x^n(x - A - 1) + A}{x - 1} > 0$$

1. Sur l'œuvre de Prestet, *cf.* la thèse de Katia Asselah [8].
2. *Cf.* [8] p. 222-223.
3. *Cf.* [267] p. 220.

Bien sûr, la méthode des cascades pose un problème d'effectivité : le calcul approché des racines de la dérivée $(i+1)$-ème ne donne qu'une connaissance approchée des bornes des intervalles en question.

Rolle s'est aussi intéressé à la résolution algébrique des équations. Les recherches sur la résolution algébrique se heurtaient alors à deux obstacles. Les méthodes de résolution algébrique connues en degrés 3 et 4 semblaient ne pouvoir s'étendre aux degrés supérieurs (il faudra attendre Abel pour en comprendre la raison), et les méthodes connues, dès le degré 3, conduisaient parfois à l'expression de racines réelles par des formules contenant des quantités « imaginaires » (ainsi le cas dit « irréductible » de l'équation cubique à trois racines réelles [4]). Ces deux apories sont présentes dans la célèbre controverse sur la méthode de Tschirnhaus ainsi que dans les travaux de Leibniz sur la résolution algébrique (intimement liés à ceux de son ami et collègue Tschirnhaus).

Le problème des quantités imaginaires retiendra peu notre attention. Indiquons simplement que Leibniz, dès 1675, proclamait la légitimité d'un calcul sur les quantités imaginaires, au moins lorsqu'elles entrent dans l'expression d'une quantité réelle [5]. On pouvait même, dans le cas irréductible de l'équation cubique, utiliser les algorithmes d'extraction de la racine cubique d'un binôme pour chercher les racines rationnelles de l'équation [6] ; cela justifiait la validité des formules de Cardan dans le cas où les racines sont rationnelles. Restaient les équations cubiques ayant trois racines réelles irrationnelles : Leibniz semble convaincu que certaines quantités réelles ne peuvent s'exprimer par radicaux autrement qu'en ayant recours à des quantités imaginaires [7]. Tout comme Leibniz ou même Prestet, Rolle en 1690 ne fait pas d'opposition de principe à travailler sur des expressions contenant des imaginaires. Rappelons d'ailleurs que Prestet a donné, en 1689, l'expression des trois racines de l'équation cubique [8] $x^3+px+q=0$:

$$\begin{array}{l} A+B \\ j^2A+jB \\ jA+j^2B \end{array}$$

où $j=\frac{-1+i\sqrt{3}}{2}$, et

$$A=-\frac{1}{\sqrt{3}}\sqrt[3]{\frac{3\sqrt{3}}{2}q+3i\sqrt{3}\sqrt{-\left(\frac{p^3}{27}+\frac{q^2}{4}\right)}}$$
$$B=-\frac{1}{\sqrt{3}}\sqrt[3]{\frac{3\sqrt{3}}{2}q-3i\sqrt{3}\sqrt{-\left(\frac{p^3}{27}+\frac{q^2}{4}\right)}}$$

Nous intéresse plutôt ici le problème de la généralisation des méthodes de résolution algébrique aux équations de degrés supérieurs. C'est aussi l'occasion d'un premier reproche adressé par Rolle à Prestet. Celui-ci affirmait s'être heurté, dans le « cas irréductible » de l'équation cubique, à un « labyrinthe inexplicable de difficultés ». Il prétendait cependant avoir « reconnu même assez clairement par le secours des combinaisons, que la découverte légitime et naturelle était impossible ». En tout cas, il avait trouvé « une route assurée » pour « éviter la méprise et l'erreur ». Prestet donne l'exemple de quelqu'un qui serait justement tombé dans l'erreur : Tschirnhaus. Mais en fait de difficultés propres au « cas irréductible », c'est toute la méthode de Tschirnhaus, pensée par son auteur comme méthode générale de résolution algébrique, que Prestet remet en cause. Aux objections de Prestet, Rolle contre-objectera, précisément dans le but [9] de prévenir d'éventuelles

4. Un problème semblable existait pour les équations de degré 4, *cf.* [152] p. 94 et 167.

5. A cet égard, le texte qui nous renseigne le mieux est peut-être le manuscrit intitulé *De resolutionibus aequationum cubicarum triradicalium. De radicibus realibus, quae interventu imaginariarum exprimuntur. Deque sexta quadam operatione arithmetica* [210]. Dans *Leibniz in Paris* ([152] n. 23 p. 147), Hofmann laisse entendre que ce manuscrit, noté C 1031, aurait été envoyé à Huygens en été 1675. C'est aussi, semble-t-il, l'hypothèse retenue lors de l'édition du manuscrit C 1032 dans les *Sämtliche Schriften und Briefe* [213], 3ème série, tome 1, p. 623–640. Mais il a été démontré, lors de l'édition de C 1031, [213], 7ème série, tome 2, p. 678–700, que c'est en fait C 1032 qui fut envoyé en été 1675 à Huygens, et non C 1031, peut-être rédigé en octobre.

6. Algorithmes déjà présents chez Stifel au siècle précédent.

7. D'ailleurs, est-ce vrai ou faux ?

8. *Cf.* [8] § V.V.2.b) p. 213.

9. *Cf.* [267] p. 267.

objections à ses propres méthodes de résolution algébrique, méthodes fortement apparentées à celle de Tschirnhaus en tant qu'elles utilisent l'élimination. Quant au « cas irréductible », Rolle était, comme Leibniz, convaincu de l'impossibilité d'éviter les quantités imaginaires dans ce cas – il le « démontre » même [10]. Selon lui, ce n'est d'ailleurs pas tant la présence de quantités imaginaires qui rend incommode les formules exprimant les racines au moyen de radicaux, que les signes radicaux eux-mêmes, incommodes même s'ils ne contiennent aucune quantité imaginaire car ils ne donnent pas les solutions rationnelles sous forme rationnelle et car « il faut beaucoup de temps et de calcul pour les faire naître et pour les faire évanouir » [11]. Néanmoins l'idée de Tschirnhaus conduit à une méthode correcte pour résoudre les équations de degré 3, et les critiques de Prestet reposent sur une incompréhension levée par Rolle qui en donne une explication claire et concise :

> M. de Tschirnhaus se propose de résoudre l'égalité $y^3 - py - q \propto \theta$, et pour cela il suppose que $yy - ay - b - x \propto \theta$. Ainsi la question est réduite à deux égalités, et il est évident que si elles sont résolues, l'égalité proposée sera résolue aussi. Si l'on cherche un commun diviseur à ces deux égalités suivant l'origine y ; et si après la première division totale on détruit tous les termes du reste, chacun séparément, alors on introduira dans l'égalité supposée deux des racines de l'égalité proposée, et la lettre a deviendra égale à la somme de ces deux racines. Mais parce que M. de Tschirnhaus ne détruit le reste qu'après que l'inconnue a cessé de paraître, il ne force l'égalité $yy - ay - b - x \propto \theta$ qu'à recevoir une seule des racines de la proposée, et il ne fait d'ailleurs rien qui puisse en introduire davantage. [12]

Pour mieux comprendre l'origine de la méthode de Tschirnhaus, il faut remonter aux recherches que celui-ci menait de pair avec Leibniz à Paris, voire à leurs recherches indépendantes, antérieures à leur rencontre fin septembre 1675. Il semble difficile de distinguer la contribution propre de chacun des deux auteurs, bien que dans son livre *Leibniz in Paris*, J. E. Hofmann accorde la supériorité dans ce domaine à Tschirnhaus, au point d'y voir une des raisons pour lesquelles Leibniz, déçu par ses propres capacités à la suite de leur rencontre, se serait quelque peu détourné de la recherche d'une méthode de résolution algébrique générale. Quelques années plus tard, le 10 avril 1678, Tschirnhaus envoie un texte à Leibniz intitulé *Methodus generalis omnium aequationum radices exhibendi* [310] dans lequel il fait état de ses recherches. La « méthode générale » que propose Tschirnhaus repose sur une conjecture qu'Euler adoptera [13] à nouveau en 1732 (et qui, bien sûr, est fausse) : les racines de l'équation générale de degré n seraient de la forme

$$x = \underbrace{a + b + c + ...}_{(n-1) \text{ termes}}$$

où a, b, c,... sont les racines n-ièmes des $(n-1)$ racines d'une certaine équation de degré $(n-1)$ dont les coefficients sont rationnels en les coefficients de l'équation initiale. Tschirnhaus tente de démontrer cette conjecture au moyen des « formules de Girard » exprimant les sommes de puissances des

10. Son raisonnement, bien sûr peu satisfaisant, est à peu près le suivant ([267] p. 230–231 § 9 à 13). L'expression d'une racine au moyen de radicaux devrait les donner toutes ; chaque radical n-ième a n valeurs, dont une ou deux sont réelles suivant que n est impair ou pair. Supposons donc qu'il existe une expression d'une racine réelle de l'équation cubique dans le cas irréductible qui ne contienne pas de quantités imaginaires. Cette expression devrait aussi donner les deux autres racines en choisissant d'autres valeurs des radicaux qui la composent, mais il est impossible d'obtenir ainsi trois valeurs exactement (assertion non démontrée par Rolle – il pensait peut-être qu'en faisant varier les déterminations réelles des radicaux qui composent l'expression, on obtient 2^k valeurs réelles, où k est le nombre de radicaux n-ième pour n pair). Bien que la nécessité du recours aux imaginaires intervienne aussi dans le discours de Leibniz, celui-ci semble faire un pas de plus. Alors que Rolle prétend montrer la nécessité des quantités imaginaires dans l'expression par une formule *unique* des racines réelles dans le cas irréductible, Leibniz essaie de convaincre que, bien que dans certains cas, une racine (exprimée par la formule de Cardan dans le cas irréductible) soit rationnelle (et donc que les imaginaires ne sont pas utiles pour exprimer cette racine particulière), il est toutefois légitime de travailler sur la quantité représentée par la formule de Cardan (contenant des imaginaires) même quand la racine n'est pas par ailleurs exprimable rationnellement (*cf.* [210] p. 142 et 148). Il s'agit donc vraiment pour lui d'étendre le champ opératoire par adjonction des quantités imaginaires.

11. *Cf.* [267] p. 259.

12. *Cf.* [267] p. 223. On remarque l'usage du signe $\propto$ au lieu de $=$, et de θ pour désigner le zéro.

13. *Cf.* [157] p. 36–37. Abel affirmera que si une équation de degré premier n est résoluble algébriquement, alors ses racines sont de cette forme (dans un manuscrit de 1828, publié en 1839 ; *cf.* l'article de C. Houzel sur Abel [156]).

$(n-1)$ racines par leur fonctions symétriques élémentaires. Ainsi pour $n=3$, $x=a+b$ et $y=ab$, on a :

$$a^3+b^3=x^3-3xy$$

La méthode des coefficients indéterminés permet d'identifier terme à terme cette équation à l'équation générale de degré 3 sans terme en x^2, et cela conduit à une équation de degré 2 dont les racines sont a^3 et b^3. Pour $n=4$, $x=a+b+c$, $y=ab+ac+bc$, $z=abc$, on a :

$$a^4+b^4+c^4=x^4-4x^2y+4xz+2y^2$$

que Tschirnhaus pense pouvoir identifier terme à terme à l'équation générale de degré 4 sans terme en x^3. Posons avec lui $p=4y$, $q=4z$ et $r=2y^2-a^4-b^4-c^4$; la méthode des coefficients indéterminés conduit au système d'équations :

$$\begin{cases} a^4+b^4+c^4=\frac{p^2}{8}-r \\ ab+bc+ac=\frac{p}{4} \\ abc=\frac{q}{4} \end{cases}$$

Mais Tschirnhaus ne conduit pas le calcul jusqu'au bout. Toujours dans le cas $n=4$, Euler montrera que la conjecture est bonne en identifiant l'équation générale à une autre formule (qui n'est pas une formule de Girard car elle fait intervenir d'autres fonctions que les fonctions symétriques élémentaires) :

$$x^4=2x^2(a^2+b^2+c^2)+8xabc+4(a^2b^2+b^2c^2+a^2c^2)-(a^2+b^2+c^2)^2$$

Mais ce texte de 1678 nous montre aussi Tschirnhaus à la recherche de la loi de formation des coefficients numériques dans les formules de Girard, recherche qui n'est pas restée stérile puisque Leibniz, dans les années 1680–1682 au plus tard, est parvenu à découvrir cette loi [14]. Dans la table suivante, Tschirnhaus avait observé la loi de formation des coefficients dans les premières colonnes (sans les signes) [15] :

$$\begin{array}{lllll}
x &= (a+b+c) & & & \\
x^2 &= (a^2+b^2+c^2)+2y & & & \\
x^3 &= (a^3+b^3+c^3)+3xy & -3z & & \\
x^4 &= (a^4+b^4+c^4)+4x^2y-2y^2 & -4xz & & \\
x^5 &= (a^5+b^5+c^5)+5x^3y-5xy^2 & -5x^2z+5yz & & \\
x^6 &= (a^6+b^6+c^6)+6x^4y-9x^2y^2+2y^3 & -6x^3z+12xyz & -3z^2 & \\
x^7 &= (a^7+b^7+c^7)+7x^5y-14x^3y^2+7xy^3 & -7x^4z+21x^2yz-7y^2z & -7xz^2 & \\
x^8 &= (a^8+b^8+c^8)+8x^6y-20x^4y^2+16x^2y^3-2y^4 & -8x^5z+32x^3yz-24xy^2z & -12x^2z^2+8yz^2 & \\
x^9 &= (a^9+b^9+c^9)+9x^7y-27x^5y^2+30x^3y^3-9xy^4 & -9x^6z+45x^4yz-36x^2y^2z+9y^3z & -18x^3z^2+27xyz^2 & +3z^3
\end{array}$$

Prestet s'était lui aussi intéressé à la résolution algébrique générale des équations, dès 1675. A la suite des travaux de Faulhaber, Descartes et surtout Hudde, il ne semble avoir envisagé qu'une seule méthode générale de résolution algébrique, consistant à identifier l'équation de degré $(m+n)$ sans terme en x^{m+n-1}, au moyen de coefficients indéterminés, au produit de deux équations de la forme $x^m+yx^{m-1}+...=0$ et $x^n-yx^{n-1}+...=0$. La diversité des méthodes de résolution algébrique qui pouvaient à titre de conjecture s'offrir à l'investigation est donc loin d'avoir été épuisée par Prestet, et c'est donc judicieusement que Rolle tira occasion, dans son livre IV, du « secours des combinaisons » allégué par Prestet pour tourner celui-ci en dérision. Pourtant, avec le recul historique, on ne peut manquer d'admirer aujourd'hui l'usage d'un argument de type combinatoire

14. *Cf.* [173] p. 145.

15. *Cf.* [310], table 4, colonne B. Le manuscrit contient cette table (selon les éditeurs en [310] p. 373), mais ni [310] ni l'édition antérieure [211] par Gerhardt n'en ont repris le contenu. Nous reconstituons donc la table sans l'avoir vue. En particulier, Tschirnhaus avait observé comment former les coefficients de la colonne commençant par le terme en yz au moyen des colonnes commençant par y et par z, et comment former les coefficients de la colonne commençant par y^2z au moyen des colonnes commençant par y^2 et par yz.

chez Prestet [16] pour déterminer *a priori* le degré de la réduite en y : les racines de la réduite sont les sommes n à n des racines de la proposée, donc son degré est $\binom{n+m}{n}$. Nous croyons même qu'un tel argument ouvrait l'une des pistes qui ménerait, à la suite d'Euler et Lagrange, aux travaux de Galois. [17]

Nous n'exposerons pas ici les différents types de méthodes de résolution algébriques envisagés par Rolle, nous tournant plutôt vers une autre critique adressée par Rolle à Prestet et concernant l'élimination algébrique. L'on pourrait y appliquer la remarque faite par Rolle au début du livre IV : s'il critique Prestet, c'est qu'il doit le préférer « à bien d'autres, parce que de tous les algébristes c'est le dernier qui a écrit ». Le problème sur lequel Rolle force ici notre attention est en effet propre à toutes les techniques d'élimination en usage au XVII$^{\text{ème}}$: au fil du calcul, le système d'équations initial est transformé en d'autres systèmes, qui ne sont pas équivalents au premier, à cause des multiplications ou divisions effectuées avant chaque substitution. Rolle donne ici deux exemples, sans en détailler la résolution. Le premier est le système suivant :

$$\begin{cases} y^2 + yz + z^2 + 24 = 9y + 9z \\ y^2z^2 + 17y + 17z = 24yz \end{cases}$$

Eliminons successivement les plus hautes puissances de y. La première étape du calcul conduit au système suivant, équivalent au système initial :

$$\begin{cases} y^2 + yz + z^2 + 24 - 9y - 9z = 0 & (1) \\ y(-z^3 + 9z^2 - 24z + 17) = z^4 - 9z^3 + 24z^2 - 17z & (2) \end{cases}$$

L'on substitue ensuite, dans (1), y par sa valeur :

$$y = \frac{z^4 - 9z^3 + 24z^2 - 17z}{-z^3 + 9z^2 - 24z + 17}$$

Quelles que soient les variantes de cette méthode, ni Prestet, ni Newton ni Hudde avant lui, ne précisent comment tenir compte des dénominateurs dans la suite des opérations. Dans notre cas, la fraction une fois réduite [18] donne :

$$y = -z$$

Ayant substitué cette valeur dans (1), on obtient finalement le système

$$\begin{cases} z^2 + 24 = 0 \\ y = -z \end{cases}$$

qui *n'est pas équivalent* au système initial. En fait, les solutions du système initial sont précisément la réunion des solutions des deux systèmes suivants (en tout, huit solutions dans $\mathbb{C}^2$) :

$$\begin{cases} z^2 + 24 = 0 \\ y = -z \end{cases} \qquad \begin{cases} y^2 + yz + z^2 + 24 - 9y - 9z = 0 \\ -z^3 + 9z^2 - 24z + 17 = 0 \end{cases}$$

Le second exemple donné par Rolle est plus élémentaire – nous doutons que Prestet eût failli si ce problème lui avait été posé – mais intéressant car il s'agit d'un système de trois équations :

$$\begin{cases} vyz + 2z + 6v + 3y = 2vz + yz + 3vy + 6 \\ vyz + 20z + 30v + 24y = 5vz + 4yz + 6vy + 120 \\ vyz + 56z + 72v + 63y = 8vz + 7yz + 9vy + 504 \end{cases}$$

Rolle se contente d'écrire : [19]

16. *Cf.* [255] t. 2 p. 437, et la thèse de Katia Asselah, [8] p. 217–222.

17. Euler utilise le même raisonnement *a priori* dans une démonstration célèbre du théorème fondamental de l'algèbre, *cf.* [108], série 1, VI, p. 78–147, et en particulier théorème 8, p. 108.

18. En 1690, Rolle parle de fraction « abrégée ». En 1709 dans [268], il adopte un langage plus précis et distingue la « division actuelle », procédé par lequel numérateur et dénominateur d'une fraction sont libérés de leur commun diviseur, de la « division intentionnelle » qui consiste à garder la fraction écrite sous sa forme initiale sans effectuer l'opération.

19. *Cf.* [267] p. 260.

> Si l'on fait la division exactement, et si l'on pratique ce qui est porté par le Problème général de M. Prestet, l'on trouvera toutes ces contradictions : 1∞4. 1∞7. 4∞7. 2∞5. 2∞8. 5∞8. 3∞6. 3∞9. 6∞9.

En effet, factorisons chacune des équations ; le système s'écrit alors

$$\begin{cases} (v-1)(y-2)(z-3)=0 & (1)\\ (v-4)(y-5)(z-6)=0 & (2)\\ (v-7)(y-8)(z-9)=0 & (3) \end{cases}$$

Si l'on compare deux-à-deux les valeurs de v issues des équations (1) et (2), (1) et (3), (2) et (3), *après* réduction des fractions, afin d'éliminer v, on trouve donc $1=4$, $1=7$ et $4=7$, ce qui pourrait laisser croire que le système est sans solution, alors qu'il en possède six et est équivalent au système suivant :

$$\begin{cases} z=15-v-y\\ v^2+(y-13)v+(y^2-14y+52)=0\\ (y-2)(y-5)(y-8)=0 \end{cases}$$

Rolle enseigne une méthode générale d'élimination, au chapitre III du livre III du *Traité d'algèbre*, qui résout ces problèmes. Son exposé est complété par des « remarques » en fin de chapitre, assez confuses et probablement rédigées à la hâte avant l'impression [20]. Pourtant même le cas le plus simple, celui de deux équations, fait à nouveau l'objet d'un mémoire de Rolle en 1709 ; ce fait montre, à lui seul, le caractère imparfait de l'exposé de 1690. Mais ce qui nous semble ici remarquable, c'est la démarche opératoire de Rolle, conférant un statut d'objet mathématique à la notion de système d'équations. Tant que les méthodes procédant par équivalences ou inférences successives d'*un* système à *un* autre étaient jugées suffisantes, l'on pouvait encore se représenter chaque étape du calcul comme transformation d'*une* équation en *une* autre au sein d'un système en quelque sorte statique et faiblement conceptualisé [21]. Mais la situation que nous avons décrite dans le premier des deux exemples ci-dessus oblige à opérer sur un système engendrant deux nouveaux systèmes dont la réunion des solutions forme l'ensemble des solutions du système de départ. Rolle représente chaque système par un nœud d'un arbre (l'« arbre de direction ») donnant naissance à plusieurs branches portant chacune un autre nœud. A partir d'un nœud, il forme le nœud suivant en appliquant l'algorithme euclidien pour éliminer une inconnue dans deux des équations du système. Les deux équations sont remplacées, dans le nœud suivant, par un reste (de degré 0 en l'inconnue éliminée) et un avant-dernier reste (de degré non nul minimal en l'inconnue éliminée). Puis l'on réitère l'algorithme euclidien avec l'avant-dernier reste et une autre équation contenant l'inconnue à éliminer, jusqu'à ce qu'il n'y ait plus d'équation contenant cette inconnue. Rolle met alors en évidence le dernier des avant-derniers restes en l'écrivant, plutôt que dans le nœud lui-même, sur la branche qui mène à ce nœud. Puis on applique de nouveau l'ensemble des opérations pour éliminer une autre inconnue, jusqu'à ce qu'elles le soient toutes. Mais comme nous le disions, de certains nœuds naissent plusieurs branches : il faut pour le comprendre observer l'effet d'une division lors de l'exécution de l'algorithme euclidien. Soit P le dividende et Q le diviseur. Si l'on fait abstraction des autres équations présentes dans le système, on représentera le nœud en question ainsi :

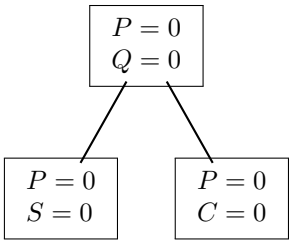
$$\boxed{\begin{array}{c} P=0\\ Q=0 \end{array}}$$

Avant d'effectuer la division, il faut chercher les diviseurs de Q de degré 0 en l'inconnue à éliminer [22]. S'il y a un tel diviseur C non trivial, notons $Q=CS$. Le nœud ci-dessus engendre alors deux autres nœuds :

20. *Cf.* [267] livre III, chapitre III, p. 187–199, et les « remarques » p. 200–217. Une erreur de mise en page [267] p. 184–185 pourrait-elle confirmer que les remarques des fins de chapitres n'ont pas été rédigées ni imprimées avec le même soin que le reste du traité ?

21. On remarque aussi chez Rolle l'usage, qui nous est resté, de l'accolade à gauche dans la notation des systèmes d'équations (certes pas de manière systématique) ; *cf.* par exemple [267] p. 259–260.

22. *Cf.* [267] p. 193–194, 3ème règle.



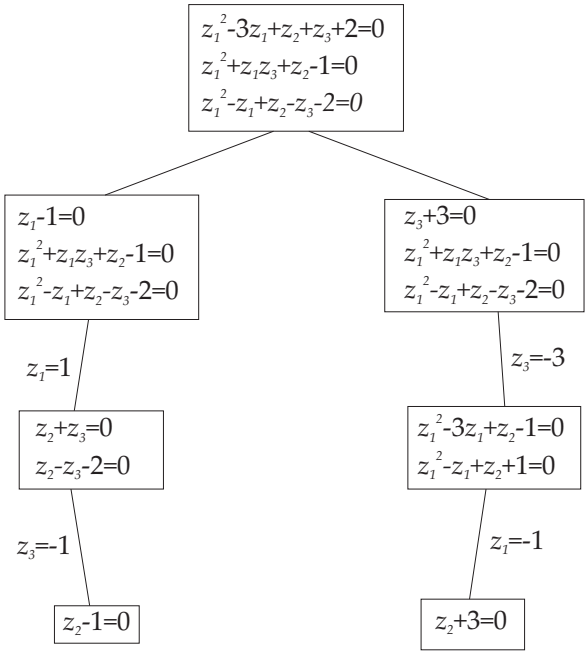
On se ramène ainsi au cas où il n'y a pas de diviseur non trivial de degré 0 en l'inconnue à éliminer avant d'effectuer la division euclidienne. Appliquons cette méthode, à titre d'exemple, au système suivant :

$$\begin{cases} z_1^2 - 3z_1 + z_2 + z_3 + 2 = 0 \\ z_1^2 + z_1 z_3 + z_2 - 1 = 0 \\ z_1^2 - z_1 + z_2 - z_3 - 2 = 0 \end{cases}$$

Cet exemple, dû à Netto [23], est intéressant car il rend compte des difficultés qu'avait probablement rencontrées Bézout au début de ses recherches. Mais le problème ne résiste pas à la méthode de Rolle. Nous n'en concluons pas pour autant que Bézout ignorait cette méthode : en fait le but qu'il poursuivait était bien différent de celui de Rolle. Bézout souhaitait réduire l'élimination en général à la résolution de systèmes d'équations linéaires, et connaître *a priori* les degrés des équations finales résultant de l'élimination. Mais l'application de la méthode de Rolle à l'exemple de Netto montre que, dès la fin du XVII$^{\text{ème}}$ siècle, existait une méthode générale d'élimination algébrique. La division euclidienne entre les deux premières équations conduit à une équation de degré 1 en z_1 qui possède un diviseur de degré 0 en z_1 :

$$(z_3 + 3)(z_1 - 1) = 0$$

Voici donc l'« arbre de direction » tel que l'aurait construit Rolle :



On a donc deux solutions, $(1, 1, -1)$ et $(-1, -3, -3)$. On remarque d'ailleurs que Netto s'est trompé puisqu'il ne trouve que la première de ces deux solutions [24].

Hélas, la méthode de Rolle, telle que nous venons de la décrire, n'échappe pas au problème des facteurs superflus qui arrêtera les savants plus tard dans le XVIII$^{\text{e}}$ siècle. Dans le mémoire de

23. *Cf.* [236] § 406.

24. Son erreur est une erreur de calcul. Il écrit en effet que le résultant en z_1 des deux dernières équations du système initial est $(z_2 - z_3 - 1)z_3(z_3 + 1)$ au lieu de $(z_3 + 1)^2(z_2 - z_3)$.

1709, Rolle résout une dizaine de systèmes de deux équations présentant chacun certaines « difficultés considérables », difficultés « auxquelles on n'avait pas fait assez attention » nous dit-il [25]. Le deuxième exemple de l'article I est sans doute le plus intéressant pour nous [26]. Notons :

$$A = y^2(x-g) - yx^2 + pl^2$$
$$B = y^2(x-n) - cly + x^3$$

Rolle se propose de résoudre le système $\begin{cases} A = 0 \\ B = 0 \end{cases}$ en les deux inconnues x, y (les autres lettres sont des constantes).

La division euclidienne de A par B a pour quotient $\frac{x-g}{x-n}$. On est donc conduit au système :

$$\begin{cases} B = 0 \\ A - \frac{x-g}{x-n}B = 0 \end{cases}$$

La seconde équation est linéaire en y :

$$A - \frac{x-g}{x-n}B = \frac{-x^3 + nx^2 + clx - clg}{x-n}y + \frac{-x^4 + gx^3 + pl^2x - pl^2n}{x-n}$$

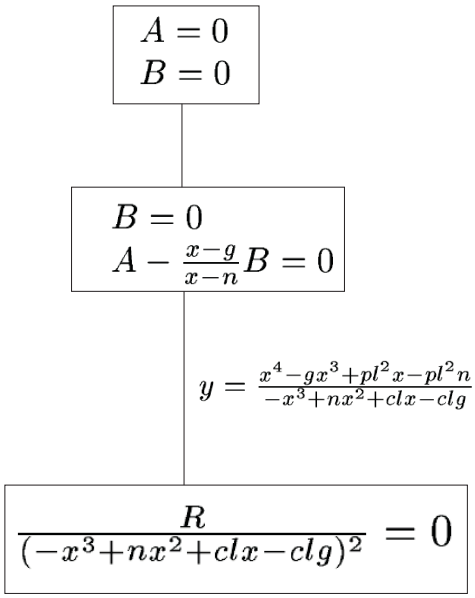
Elle nous donne une valeur de y :

$$y = \frac{x^4 - gx^3 + pl^2x - pl^2n}{-x^3 + nx^2 + clx - clg}$$

La division euclidienne de B par $A - \frac{x-g}{x-n}B$ revient en fait à substituer cette valeur de y dans B, comme le fait Rolle. D'où une équation finale de la forme :

$$\frac{R}{(-x^3 + nx^2 + clx - clg)^2} = 0$$

avec $\deg R = 9$. On pourrait résumer cette séquence d'opérations par l'arbre de direction :



Mais l'algorithme euclidien introduit le facteur superflu

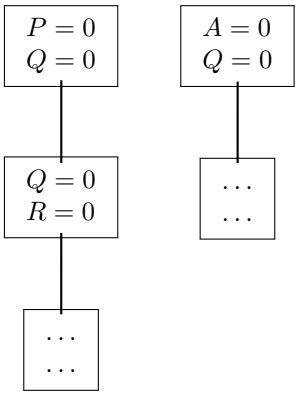
$$\frac{x-n}{(-x^3 + nx^2 + clx - clg)^2}$$

25. *Cf.* [268] p. 419.
26. *Cf.* [268] p. 422.

Nous verrons plus tard comment Sylvester déterminera le facteur superflu introduit par l'algorithme euclidien pour des équations de degrés quelconques (*infra* section 12.2). Rolle remarque en effet que $(x-n)$ divise R tandis qu'il n'existe pas de solution du système initial en $x = n$. Il se contente d'écrire qu'il faut, pour le voir, vérifier en substituant cette solution dans le système initial. Pourtant, certains indices laissent penser qu'il pressentait le rôle du coefficient dominant du diviseur dans l'apparition de solutions superflues. Cet exemple même en témoigne [27]. Dans le traité de 1690, on trouve aussi, parmi les « remarques » à la fin du chapitre III du livre III, la remarque suivante, certes assez obscure :

> En pratiquant la méthode à laquelle on s'est fixé ici, on prendra toutes les quantités particulières inconnues desquelles on s'est servi pour multiplier les diviseurs avant que faire la soustraction dans les divisions ; et si ces quantités n'ont point servi, comme produisans, à former de nouveaux nœuds, on les mettra à part à mesure qu'on les trouvera, jusques à ce qu'on soit arrivé aux extrémités de l'arbre de direction.
>
> On reprendra ces quantités mises à part avec les égalités des extrémités, et on y appliquera de nouveau la méthode ; et si l'on y trouve quelque résolution, l'on pourra s'en servir pour exclure de question plusieurs résolutions contradictoires, et pour simplifier les égalités principales. Et par là on peut voir que si les multiplicateurs sont connus, ces contradictions ne sont point à craindre. [28]

En fait, il est possible, en modifiant légèrement la méthode de Rolle, de parvenir à la connaissance des facteurs superflus. Observons l'effet d'une « division partielle » (c'est par ce terme que Rolle désigne le calcul d'un des termes du quotient dans la division euclidienne ; la division euclidienne elle-même étant la « division totale ») sur le système constitué de deux équations $\begin{cases} P = 0 \\ Q = 0 \end{cases}$. On va d'abord modifier l'algorithme euclidien afin d'éviter toute fraction. A cet effet, avant chaque division partielle, on multipliera le dividende par le coefficient dominant du diviseur. Nous désignerons l'algorithme euclidien ainsi modifié sous le nom d'« algorithme euclidien du plus grand commun diviseur sans fraction » pour le distinguer de l'algorithme classique. Notons A le coefficient dominant de Q. La première division partielle s'écrit alors $S = AP - MQ$ où S est le premier reste intermédiaire et M le premier terme du quotient de la division totale. Et ainsi des divisions partielles suivantes. Quant à la division totale, elle s'écrit donc $R = A^k P - TQ$ où R est le dernier reste, T le quotient, et k un entier positif. Le nouveau système $\begin{cases} Q = 0 \\ R = 0 \end{cases}$ auquel conduit cette division totale est alors équivalent au système $\begin{cases} A^k P = 0 \\ Q = 0 \end{cases}$, et non au système $\begin{cases} P = 0 \\ Q = 0 \end{cases}$. Il faut donc construire un autre arbre de direction, à partir du système $\begin{cases} A = 0 \\ Q = 0 \end{cases}$, qui nous renseignera sur les éventuelles solutions superflues introduites par cette division totale. On représenterait cela ainsi :



27. De même que l'exemple de [268] p. 424, où la situation est compliquée par la présence de solutions infinies.
28. *Cf.* [267] p.202. Les « égalités des extrémités » désignent les équations finales (c'est-à-dire les équations présentes dans les feuilles de l'arbre).

Les solutions fournies par le deuxième arbre de direction (l'arbre à droite) sont alors précisément les solutions superflues du premier. Se posent cependant un problème lié à la multiplicité des solutions. Supposons x solution de l'arbre à gauche avec multiplicité 3 et solution de l'arbre à droite avec multiplicité 1. La méthode ci-dessus nous fait croire que x est solution superflue de l'arbre à gauche, alors qu'elle en est en fait une véritable solution (avec multiplicité 2).

Cette modification de la méthode de Rolle, basée sur l'algorithme euclidien sans fraction, nous conduit naturellement à présenter les travaux d'Ossian Bonnet sur l'élimination, qui sont l'aboutissement des recherches de quelques savants français de la première moitié du XIX$^{\text{e}}$ siècle : Labatie, Sarrus, Lefébure de Fourcy, Bret. Bonnet se propose, comme Rolle, de ramener la résolution d'un système de deux équations à deux inconnues à celle de plusieurs systèmes de deux équations dont l'une ne contient plus qu'une seule inconnue. Il applique à cet effet l'algorithme euclidien sans fraction pour la recherche du plus grand commun diviseur des deux équations initiales $A = 0$, $B = 0$, algorithme qui conduit à déterminer des quotients successifs $q, q_1, ..., q_n$ et des restes successifs $Rr, R_1r_1, ..., R_{n-1}r_{n-1}, r_n$, où r_p désigne, dans chaque reste R_pr_p, un éventuel diviseur de degré 0 en y (en particulier le dernier reste r_n est de degré 0 en y) :

$$\begin{aligned}
&cA = Bq + Rr\\
&c_1B = Rq_1 + R_1r_1\\
&c_2R = R_1q_2 + R_2r_2\\
&\vdots\\
&c_nR_{n-2} = R_{n-1}q_n + r_n
\end{aligned}$$

Bonnet décrit bien la difficulté qu'il y a à interpréter le résultat de cet algorithme à cause des facteurs superflus introduits par les multiplicateurs successifs $c, c_1, ..., c_n$:

> (...) on pourra bien dire (...) que les solutions du système $cA = 0$, $B = 0$, sont les mêmes que celles du système $B = 0$, $Rr = 0$, que les solutions du système $c_1B = 0$, $R = 0$, sont les mêmes que celles du système $R = 0$, $R_1r_1 = 0$, etc. Mais comment, en général, les solutions du système $c_pR_{p-2} = 0$, $R_{p-1} = 0$, dépendent-elles des solutions des deux systèmes $c_p = 0$, $R_{p-1} = 0$, et $R_{p-2} = 0$, $R_{p-1} = 0$, et les solutions du système $R_{p-1} = 0$, $R_pr_p = 0$ des solutions des deux systèmes $R_{p-1} = 0$, $R_p = 0$, et $R_{p-1} = 0$, $r_p = 0$? Les opinions ont été partagées sur ce point ;... [29]

Bonnet introduit la notation $[A, B]$ pour désigner « l'ensemble des solutions finies » du système $\begin{cases} A = 0 \\ B = 0 \end{cases}$. Mais ce faisant, il tient compte des solutions multiples, dont il définit la multiplicité au moyen de la « formule de Poisson », comme nous le verrons *infra* section 14.1. Il est donc plus conforme à la pensée de Bonnet de se représenter $[A, B]$ non comme un ensemble mais comme une somme formelle de points, chacun affecté d'un coefficient numérique qui est la multiplicité d'intersection de $A = 0$ et $B = 0$ en ce point (on suppose que l'intersection est propre, c'est-à-dire que les courbes $A = 0$ et $B = 0$ n'ont pas de composante irréductible commune). Bonnet utilise d'ailleurs l'addition entre deux symboles $[A, B]$. Pour tous A, B, C, M, il démontre que :

$$\begin{aligned}
&[A, B] = [A + BM, B]\\
&[AB, C] = [A, C] + [B, C]
\end{aligned}$$

Il est ainsi à même de prouver un théorème de Bret [30] :

$$[A, B] = \begin{cases} [r, B] + [r_1, R] + [r_2, R_1] + ... + [r_n, R_{n-1}] \\ -[c, B] - [c_1, R] - [c_2, R_1] - ... - [c_n, R_{n-1}] \end{cases}$$

Il parvient ensuite à transformer cette expression de sorte qu'elle ne contienne plus que des additions, aucune soustraction [31]. Bonnet annonce qu'il donnera aussi les moyens de calculer le résultant de

29. *Cf.* [32] p. 58.
30. *Cf.* [32] p. 137–138.
31. Il retrouve ainsi un théorème de Labatie et Sarrus, *cf.* [32] p. 251–252.

A et B par cet algorithme [32]. On trouve en effet, dans une édition de 1849 du manuel d'algèbre de Choquet et Mayer, une démonstration attribuée à Bonnet :

> ...[Bonnet] annonçait qu'il s'occuperait aussi de l'élimination proprement dite, c'est-à-dire du cas où le procédé d'élimination est employé seulement pour obtenir une ou plusieurs équations indépendantes de l'une des inconnues, et qui donnent toutes les valeurs de l'autre inconnue. C'est ce sujet que je vais traiter : les résultats sont ceux auxquels M. Labatie était parvenu ; mais les démonstrations sont de M. Bonnet, qui ne les avait pas encore publiées, et qui a bien voulu me les communiquer. [33]

Le résultat est le suivant. Soient $m, n, p, p_1, ..., p_{n-1}$ les degrés de $A, B, R, R_1, ..., R_{n-1}$ par rapport à y ; et $\alpha, \beta, \rho, \rho_1, ..., \rho_{n-1}$ les coefficients dominants de ces polynômes en y. Alors le résultant est égal à :

$$\left(\frac{r}{c}\right)^n \left(\frac{r_1}{c_1}\right)^p \left(\frac{r_2}{c_2}\right)^{p_1} ... \left(\frac{r_n}{c_n}\right)^{p_{n-1}} \beta^{m-p}\rho^{n-p_1}\rho_1^{p-p_2}...\rho_{n-2}^{p_{n-3}-p_{n-1}}\rho_{n-1}^{p_{n-2}}.$$

Revenons au traité de Rolle, qui n'a pas fini de nous surprendre. Rolle y énonce le théorème de Bézout pour un nombre quelconque d'équations :

> Lorsque une question est déterminée, l'on saura combien il peut y avoir de résolutions au plus, si l'on multiplie mutuellement tous les nombres qui expriment la multitude des dimensions de chaque égalité, parce que la question peut avoir autant de résolutions différentes, qu'il y a d'unités dans le produit de cette multiplication. Fort souvent il y a moins de résolutions dans la question que d'unités dans ce produit ; mais il n'y en a jamais davantage. [34]

Il n'en donne certes aucune démonstration, mais la notion d'« arbre de retour », exposée par Rolle dans le même chapitre (avant même la notion d'« arbre de direction »), en est probablement le fondement heuristique. Remarquons d'abord que l'« arbre de direction » a pour effet de transformer un système d'un nombre quelconque d'équations en un ou plusieurs autres systèmes qui ont la particularité d'être en quelque sorte triangulaires, c'est-à-dire de la forme suivante :

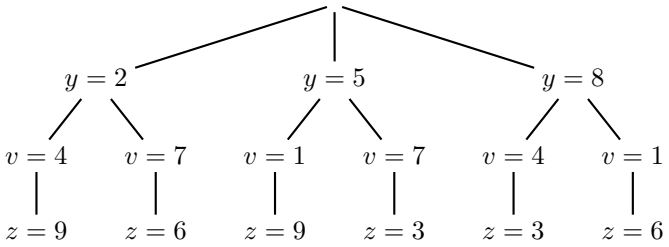
$$\begin{cases} P_1(x_1, x_2, ..., x_n) = 0 \\ P_2(x_2, ..., x_n) = 0 \\ \vdots \\ P_{n-1}(x_{n-1}, x_n) = 0 \\ P_n(x_n) = 0 \end{cases}$$

Quant à l'« arbre de retour », il permet de résoudre un tel système triangulaire. On construit un nœud pour chaque racine de l'équation $P_n(x_n) = 0$. Soit a_n l'une de ces racines, son nœud donne alors naissance à d'autres nœuds, chacun correspondant à l'une des racine de l'équation $P_{n-1}(x_{n-1}, a_n) = 0$, et ainsi de suite. Reprenons l'exemple du système suivant :

$$\begin{cases} z = 15 - v - y \\ v^2 + (y - 13)v + (y^2 - 14y + 52) = 0 \\ (y - 2)(y - 5)(y - 8) = 0 \end{cases}$$

Son arbre de retour est :

32. Le résultant, que nous avons déjà rencontré dans les travaux de Leibniz, s'exprime aussi par la formule de Poisson, *cf. infra* section 7.1.
33. § 1 p. 618.
34. *Cf.* [267] p. 187.



Il est alors évident que le nombre de feuilles de l'arbre est borné par le produit des degrés des équations proposées. Mais il est bien plus difficile de justifier ce fait dans le cas d'un système non triangulaire, et Rolle reste silencieux à ce sujet.

Deuxième partie

Le théorème de Bézout

Chapitre 5

Maclaurin : géométrie et élimination

Dans le cas de deux équations, le théorème de Bézout est souvent attribué à Maclaurin (1698–1746) [1], et dans le cas d'un nombre quelconque d'équations, à Bézout, ainsi que sa démonstration. Mais le théorème de Bézout, comme on l'a vu dans les chapitres précédents, était déjà là quand Maclaurin publia en 1720 (à l'âge de vingt ans, et avec l'aide de Newton que Maclaurin avait rencontré à Londres en 1719) sa *Geometria organica, sive descriptio linearum curvarum universalis*. Cet ouvrage fut pourtant crucial dans la compréhension du théorème de Bézout : on y voit apparaître le décompte des points d'intersection de deux courbes *avec leur multiplicité* dans l'application de ce théorème. Ceci suppose une classification (encore grossière) des singularités des courbes algébriques et un concept (pas encore nommé) de multiplicité d'intersection $I(\mathcal{C}, \mathcal{D}; P)$ de deux courbes $\mathcal{C}$ et $\mathcal{D}$ en un point P, avec, en filigrane, la formule :

$$(*) \quad I(\mathcal{C}, \mathcal{D}; P) \geq \mu_{\mathcal{C}}(P).\mu_{\mathcal{D}}(P)$$

où $\mu_{\mathcal{C}}(P)$ désigne l'ordre de multiplicité du point P sur la courbe $\mathcal{C}$. Le théorème de Bézout s'énonce alors :

$$\sum_{P \in \mathcal{C} \cap \mathcal{D}} I(\mathcal{C}, \mathcal{D}; P) = \deg(\mathcal{C}).\deg(\mathcal{D})$$

Nous allons donc exposer en détail l'œuvre de Maclaurin sur les courbes algébriques, pour mieux comprendre le contexte dans lequel ces idées ont vu le jour. Maclaurin avait déjà publié un article [219] aux *Philosophical Transactions* en 1719, dont le contenu est repris dans la *Geometria organica*. A partir de 1725, Maclaurin enseigne les mathématiques à Edinburgh (où il obtient, grâce au soutien de Newton, la chaire de mathématiques). Il a dû y enseigner, entre autres, la théorie des courbes, mais il reste peu de traces du contenu de ces cours. C'est aussi en disciple de Newton qu'il s'intéresse à l'analyse infinitésimale et publie un *Treatise of Fluxions*. Enfin, en 1748, ses collègues éditent de manière posthume un *Treatise of Algebra* sur la base des manuscrits qu'il leur a laissés, en appendice duquel est publié un texte important pour nous, rédigé par Maclaurin en latin, le *De linearum geometricarum proprietatibus generalibus*. La préface au *Treatise of Algebra* semble indiquer qu'il avait achevé ce traité juste avant sa mort, et la correspondance montre qu'il y travaillait en 1743–1744 à la suite de la communication par R. Smith d'un théorème de Cotes que nous citerons ci-après. Mis à part un article [221] paru en 1735 à la suite d'une controverse (nous y reviendrons), Maclaurin n'a donc fait paraître l'essentiel de ses recherches sur les courbes algébriques que dans deux ouvrages, la *Geometria organica* et le traité posthume, mais son intérêt pour ce sujet s'est étendu de ses premiers travaux mathématiques jusqu'à sa mort.

Les recherches décrites dans la *Geometria organica* s'articulent autour de trois thèmes. Tout d'abord, comme son titre l'indique, cet ouvrage s'intéresse à la « description organique » des courbes,

1. *Cf.* [235], § 6 : « Für den Fall $m = 2$ wurde das Theorem zuerst von C. Maclaurin ausgesprochen ; G. Cramer und Euler versuchten, Beweise für diesen Fall zu geben ». Voir aussi [234] I-9 § 51 et [49] p. 130.

c'est-à-dire aux procédés mécaniques pour tracer les courbes algébriques. Les mouvements utilisés dans les procédés décrits par Maclaurin sont : la rotation d'une droite mobile autour d'un point fixe, la translation d'un point mobile sur une droite fixe, le mouvement d'un point mobile le long d'une courbe fixe de degré quelconque, le mouvement d'une droite mobile autour d'une courbe fixe de degré quelconque (la droite devant rester tangente à la courbe), et des systèmes articulés quelconques de tels points et droites mobiles (en imposant parfois à deux droites mobiles de garder l'une par rapport à l'autre un angle constant [2]). Idéalement Maclaurin souhaite fournir des instruments pour tracer une courbe donnée par son degré et par un certain nombre de conditions, un certain nombre de points donnés devant appartenir à la courbe, éventuellement des points singuliers d'ordre de multiplicité donné. Mais ce projet est loin d'être atteint, et l'ouvrage s'achève sur un constat d'échec : « puisque les constructions de courbes par des points donnés que nous avons pu atteindre jusqu'à présent ne sont pas suffisamment générales, nous ne nous attarderons pas davantage. » [3]

Un deuxième thème est celui de la classification des courbes algébriques. La description organique doit obéir à un critère de rationalité : elle doit, idéalement, offrir un instrument pour construire toutes les courbes d'une certaine classe. Maclaurin classe les courbes algébriques (sur le modèle de l'*Enumeratio* de Newton pour les cubiques) au sein de chaque degré, suivant la nature des branches infinies, la position des asymptotes, et le nombre et la nature des points singuliers. Maclaurin distingue trois types de points singuliers : le nœud (*nodum*), le point de rebroussement (*cuspis*), et le *punctum conjugatum* (point réel isolé, en lequel deux branches imaginaires de la courbe se rencontrent) [4]. Le nœud est un point où se coupent plusieurs branches de la courbe, en nombre quelconque, d'où la notion de point double, triple, etc. (*punctum duplex, triplex*, etc.). Maclaurin est conscient de la complication qui survient lorsque deux des branches d'un tel *punctum multiplex* sont tangentes l'une à l'autre, mais il évite en général de considérer ce cas. Il envisage enfin des combinaisons quelconques de ces trois types en un même point. Pour illustrer la cohérence voulue par Maclaurin entre classification et description organique, nous allons citer intégralement un passage qui renseigne aussi sur les auteurs qu'il considérait comme ses prédécesseurs :

> Au moyen de ces propositions nous avons présenté une théorie de la description des courbes par le mouvement d'angles à travers des lignes droites, aussi simple et universelle que la nature de la chose semble l'admettre. La droite est la première de toutes les lignes et elle est exigée dans nos constructions ; par l'usage d'une seule construction nous avons décrit le système le plus simple après les droites, celui des lignes du second ordre ; et l'on concluait presque toute la géométrie des anciens au moyen de ces deux premiers ordres de lignes ; et ils s'étaient rarement élevés au delà, sauf à en employer quelqu'une d'ordre plus élevé, afin d'obtenir la solution de problèmes particulièrement difficiles.
>
> Les modernes ont découvert de nouvelles carrières de courbes, étendues à l'infini, et ils ont démontré des théorèmes tout-à-fait difficiles sur leurs mesures et leurs propriétés variées, puisqu'en effet ils possèdent l'éminente méthode de fluxion. Mais la plupart, lorsqu'ils fondent leurs propres théories, appliquent sans cesse une manière différente de réduire les courbes en ordres, presque sans songer à cette manière-là, la meilleure. Ainsi parmi les hommes illustres qui ont bien mérité de la géométrie, plusieurs ont construit de nouveaux ordres hyperboliques et paraboliques de courbes, conduits par l'exemple desquels Fermat et Varignon ont composé des classes semblables de spirales. De même que les anciens construisaient la conchoïde à partir de la ligne droite [5], la Hire a enseigné à construire quelques espèces plus complexes à partir de plus simples ; Nicole a fait tourner les courbes sur des bases qui leur sont semblables pour dessiner des courbes géométriques plus complexes [6] et a augmenté la Géométrie d'un nouvel artifice : ceux-là, de fait, et

2. Ceci a pour conséquence que les théorèmes énoncés dans la *Geometria organica* ne sont pas des théorèmes projectifs.

3. *Cf.* [220], p. 139 : « Cum curvarum constructiones per puncta quaevis proposita, quas adhuc assequi potuimus, non sint satis generales, his ulterius prosequendis non immorabimur. »

4. Maclaurin n'envisage bien entendu que des points singuliers réels. Mais remarquons rétrospectivement que cela suffit pour une étude exhaustive des cubiques planes à coefficients réels, puisqu'elles ne peuvent avoir de point singulier qui ne soit réel.

5. C'est la méthode d'« inclinaison » ou νεῦσις.

6. Les épicycloïdes.

> d'autres (dont il est inutile de rappeler ici qu'il faut louer leur travail), promurent les mathématiques avec un zèle fécond ; tous pourtant, avant Newton, ont dédié leur œuvre seulement à des courbes particulières, ou bien n'ont jugé digne que d'une médiation superficielle la distribution nécessaire des courbes que nous avons utilisée. En vérité la méthode de Newton [7] n'est ni étendue par cet auteur à toutes les lignes de n'importe quel ordre, ni distinctement à quelques-unes de chaque ordre [8] ; quelque méthode générale serait donc désirable, par l'usage de laquelle les lignes géométriques, distribuées en ordres d'après leurs intersections avec une droite [9], puissent être ainsi décrites qu'à chaque ordre convienne sa propre description, pas moins que sa propre équation algébrique générale ; et après tant de découvertes nouvelles sur la quadrature des courbes, desquelles la géométrie s'est enrichie au plus haut point, peu d'autres choses semblent à désirer pour obtenir en la très haute géométrie toute la perfection qu'il est possible d'espérer. Nous, en vérité, ne prétendons nullement être parvenu à cette fin dans ce qui précède ; au contraire divers problèmes sur la description des courbes peuvent être posés, à la construction desquels ces choses ne suffisent pas ; mais pour obtenir cette perfection en géométrie, elles préparent le chemin, ces choses que nous avons déjà démontrées, et pour marcher vers cette fin, elles pourront nous conduire, ces choses que nous allons démontrer dans la partie suivante.
>
> Nous avons montré tout-à-fait abondamment jusqu'à quel point les lignes droites peuvent conduire à la description des lignes de tout ordre ; mais pour que notre théorie soit rendue encore plus universelle, nous examinerons prochainement quelles lignes pourront être décrites par d'autres lignes quelconques. [10]

Nous ne ferons que mentionner le troisième thème présent dans la *Geometria organica* : les applications à la mécanique. Ainsi, au commencement d'un chapitre sur les courbes que l'on appellera plus tard « podaires », Maclaurin écrit :

> Comme la géométrie des courbes relativement à un centre donné peut s'appliquer facilement en Philosophie naturelle à l'explication des mouvements des corps et des forces, (...) nous considérerons dans cette section les courbes relativement à un point quelconque donné, duquel sont dirigés des rayons vers chaque point de la circonférence, ainsi que des perpendiculaires aux tangentes à la courbe en ces points. [11]

Le pied de la perpendiculaire décrit alors la podaire. Ce thème ne nuit pas pour autant à l'unité de l'ouvrage car on peut aussi interpréter cette définition de la podaire comme instrument de description organique. Il ne nuit pas non plus au caractère théorique de l'ouvrage. En effet, l'auteur écrit un autre chapitre dans lequel il étudie le mouvement d'un point soumis à une force ; il n'y est plus question de description organique et il montre ainsi l'intention théorique de ce traité. Pas plus qu'il ne s'agissait, pour bien des auteurs ayant écrit sur la description organique avant Maclaurin, de construire effectivement un instrument de tracé, il ne s'agit certainement pour Maclaurin ni d'imiter le réel ni de modéliser un phénomène au moyen d'un instrument : bien qu'il soit possible d'appliquer la théorie des courbes à la philosophie naturelle, c'est la courbe, et non l'instrument, qui représente les lois physiques. Ainsi pouvons-nous dire que l'objet premier dans la *Geometria organica* est la courbe. Maclaurin écrit à ce propos dans la préface :

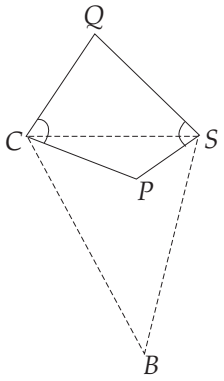
> Celui qui s'abandonne aux spéculations abstraites de la Géométrie, soit qu'il cherche à découvrir les diverses propriétés d'une ligne ou d'une figure particulière, soit qu'il étudie celles d'une famille entière de lignes, ou de tout le groupe qui réunit les familles distinctes, soit enfin qu'il s'attache aux caractères plus généraux qui sont communs à tous ces groupes, celui-là, disons-nous, peut suivre dans ses développements et dans son infinie variété toute idée préconçue de proportion et par conséquent de beauté. *Car parmi*

7. Nous verrons plus loin en quoi consiste la « méthode de Newton ». Il s'agit, pour l'essentiel, d'un instrument permettant de construire une courbe de degré $2n$ à partir d'une courbe de degré n.
8. *...neque ad omnes Lineas cujusvis ordinis, nec expresse ad aliquas omnium ordinum...*
9. C'est-à-dire selon leurs degrés.
10. *Cf.* [220], *pars* I *sectio* IV, *scholium* p. 77.
11. *Cf.* [220], *pars* II *sectio* III, p. 95.

> *toutes les lois existantes ou simplement possibles d'harmonie et de proportion, il n'en est pas une seule qui ne trouve son application dans quelque courbe, et qui n'ait ainsi sa représentation sensible. C'est à la sagacité du géomètre qu'il appartient de découvrir, dans chaque cas, cette courbe spéciale qui concrétise sa conception abstraite*, et ce travail qui la confirme est en même temps merveilleusement propre à accroître les forces et la pénétration de l'esprit. Mais cela même nous conduit à des considérations d'un ordre plus élevé. En effet, soit que la Géométrie pure nous dévoile les mystères du monde réel, soit qu'elle nous permette de prévoir et de préciser les phénomènes d'un monde imaginaire qui serait gouverné par des lois différentes de celles que nous connaissons, elle nous pénètre sans cesse de la sagesse profonde du grand Architecte de l'univers, en nous faisant reconnaître partout, pendant la durée limitée de notre vie et dans l'espace borné que nous embrassons, des ouvrages dignes de sa providence et de sa toute-puissance. [12]

L'instrument pour tracer la courbe n'est que l'instrument du géomètre à la recherche d'une classification, et nous verrons que cet instrument est plus fin que l'équation cartésienne.

Après ces généralités, nous allons décrire les principaux instruments étudiés par Maclaurin. Le premier remonte à Newton [13]. Il s'agit d'un mécanisme [14] constitué de deux points fixes S et C, autour desquels pivotent respectivement deux angles (de mesures constantes) $\widehat{PSQ}$ et $\widehat{PCQ}$. Soit B la position que prend le point P lorsque $Q \in (CS)$. Lorsque l'on fait parcourir au point Q d'intersection des droites mobiles (SQ) et (CQ) une courbe $\mathcal{C}$ de degré n, le point P trace en général une courbe de degré $2n$, ayant trois singularités d'ordre n, aux points S,C,B. Mais si $\mathcal{C}$ a une singularité d'ordre $(n-1)$ en C, alors la courbe tracée par P dégénère en une courbe de degré $(n+1)$ ayant une singularité d'ordre n en C, et S et B pour points simples.

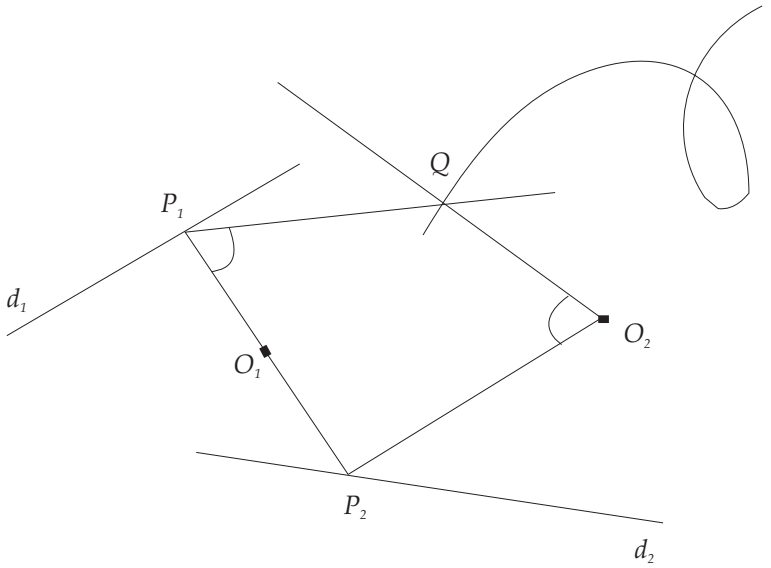


Maclaurin utilise d'abord cet instrument en prenant pour $\mathcal{C}$ une droite, afin de tracer une section conique [15] (qui peut être dégénérée). En prenant pour $\mathcal{C}$ une conique, on obtient une quartique, qui peut dégénérer en une cubique à point double [16] ; ce cas était déjà connu de Newton. On peut interpréter ce mécanisme comme une transformation birationnelle quadratique du plan $\phi_{S,C,B} : Q \mapsto P$, d'inverse $\psi_{S,C,B}$ (transformation dite « de Cremona »). D'autre part, cette application donne un paramétrage de la courbe tracée par les points de la courbe donnée. Ainsi, par composition, si la première courbe donnée est une droite, les courbes tracées après un certain nombre d'applications successives de cet instrument seront toujours *rationnelles* (c'est-à-dire paramétrées rationnellement par les points d'une droite). Maclaurin en donne un bel exemple [17] : il part d'une droite, applique n fois l'instrument de manière à obtenir à la k-ième étape ($1 \leq k \leq n$) une courbe de degré k ayant un point d'ordre $(k-1)$, puis une dernière fois pour obtenir une courbe de degré $(2n)$ ayant 3 points

12. Nous soulignons, *cf.* [220], préface. De Jonquières a traduit un extrait de la préface dans [86], p. 162–165 : c'est cette traduction que nous citons.
13. Au livre I des *Principia*, lemmes 20 et 21.
14. Le mécanisme est décrit par Maclaurin de manière générale en [220] partie II, section I, proposition III et ses corollaires I et IV.
15. *Cf.* [220], *pars* I, *sectio* I.
16. *Cf.* [220], *pars* II, *sectio* I.
17. *Cf.* [220], partie II, section V, prop. XXV et XXVII, p. 137–139.

d'ordre n et un point d'ordre $(n-1)$. Il montre que les positions des instruments successifs peuvent être déterminées de manière à ce que la courbe finale ait ses quatre points singuliers donnés, ainsi que $(2n)$ points simples donnés. Si l'on interprète ce procédé en le « lisant à l'envers », à la manière des géomètres de la fin du XIX$^{\text{ème}}$ siècle, dont le but n'était plus de tracer des courbes compliquées à partir de courbes simples, mais était au contraire de transformer des courbes compliquées en des courbes simples (en particulier pour « résoudre les singularités » au moyen de transformations birationnelles), on a transformé une courbe de degré $2n$ donnée (déterminée de manière unique par ses 4 singularités et ses $2n$ points simples donnés) par des applications successives $\psi_{S,C,B}$ pour des points S,C,B bien choisis, d'abord en une courbe de degré n ayant seulement une singularité d'ordre $(n-1)$, puis en une courbe de degré $(n-1)$ ayant une singularité d'ordre $(n-2)$, etc., jusqu'à obtenir une droite. Quelque soit l'interprétation choisie, cette construction donne un paramétrage rationnel de $\mathcal{C}$, au moyen d'une succession de transformations quadratiques « éclatant » les points singuliers.

Un autre instrument présenté par Maclaurin [18] consiste en deux points fixes O_1 et O_2 et deux couples de droites articulées suivant des angles constants, le sommet de l'un étant situé au point O_2, l'un des côtés de l'autre passant par le point O_1, et le point d'intersection P_2 du côté passant par O_1 et d'un des côtés de l'angle en O_2 glissant le long d'une droite fixe d_2.



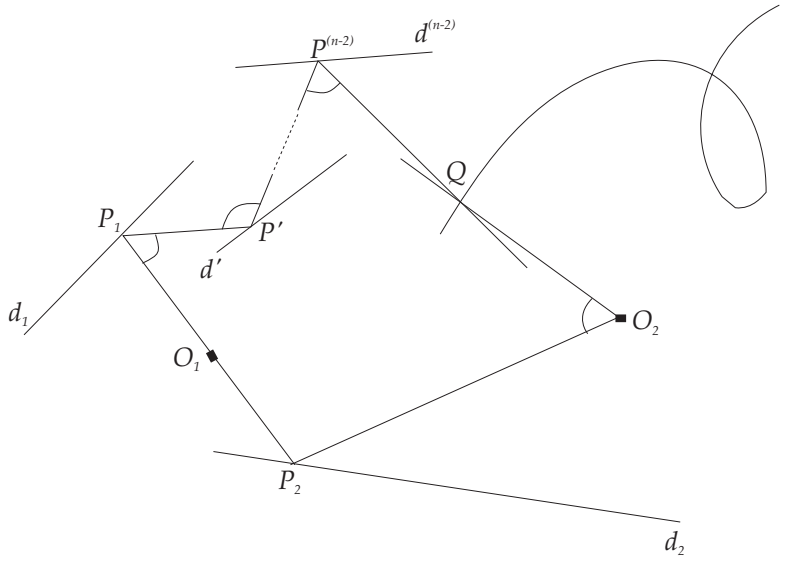
Alors il existe une transformation birationnelle du plan faisant correspondre au point d'intersection Q des deux autres côtés le sommet P_1 de l'autre angle. Si l'on fait parcourir au point P_1 une droite d_1, le point Q trace une cubique avec point double en O_2, et il s'agit de nouveau d'un paramétrage rationnel. Maclaurin envisage aussi le cas où O_2 est à l'infini. Il détermine la nature du point double et des branches infinies de la courbe en fonction des données de l'instrument et essaie de montrer que l'on obtient ainsi toutes les cubiques à point double. Newton avait classé les cubiques en « genres », selon la nature du point singulier et la position des branches infinies. Maclaurin se contente en fait de montrer que l'on peut tracer des cubiques de chaque genre (mais pas forcément toutes dans chaque genre) pour les genres suivants :

– l'hyperbole redondante (point double fini, trois asymptotes de directions distinctes réelles finies),
– l'hyperbole parabolique (point double fini, une asymptote réelle finie et des branches paraboliques dans une autre direction),
– la parabole divergente (point double fini, branches paraboliques avec un point d'inflexion à l'infini),
– l'hyperbole défective (point double fini, une asymptote finie réelle seulement),
– l'hyperbolisme d'hyperbole (point double infini, deux asymptotes réelles finies distinctes parallèles et une autre asymptote réelle finie dans une autre direction),
– l'hyperbolisme de parabole (point double infini, deux asymptotes réelles finies confondues et une

18. *Cf.* [220], *pars* I, *sectio* II, proposition V

autre asymptote réelle finie dans une autre direction),
– le trident cartésien (point double infini, une asymptote réelle finie et des branches paraboliques dans la même direction).
Parmi les cubiques à point double qu'obtient ainsi Maclaurin, il manque donc deux genres de la classification newtonienne : l'hyperbolisme d'ellipse (ayant un point isolé réel à l'infini) et la cubique de Wallis (ayant un point de rebroussement à l'infini, à tangente infinie). Ce n'est sûrement pas une inadvertance : Maclaurin a peut-être essayé de décrire ces courbes avec cet instrument, et il a dû rencontrer un obstacle. En fait, on peut démontrer qu'il est impossible de les obtenir avec un tel instrument. Sans le démontrer ici en détail, remarquons que le principal inconvénient des instruments envisagés par Maclaurin est qu'ils ne sont pas invariants projectivement. L'instrument contenant des angles constants, l'on ne peut pas obtenir (sauf si la droite à l'infini est conservée) l'image d'une courbe par une transformation projective en appliquant la transformation aux données de l'instrument que sont les points fixes et les droites fixes. Ajoutons que le premier instrument envisagé (que connaissait déjà Newton) permet pourtant d'obtenir toutes les cubiques à point double ; mais ni Newton ni Maclaurin ne l'ont démontré.

Maclaurin généralise le dernier instrument pour tracer des courbes de degrés élevés [19], en remplaçant le côté P_1Q par une ligne brisée dont les côtés forment des angles constants de sommets P', P'',..., $P^{(n-2)}$ se déplaçant respectivement sur des droites fixes d', d'',..., $d^{(n-2)}$. Il existe à nouveau une transformation birationnelle du plan entre les points P_1 et Q, transformant la droite fixe d_1 parcourue par le point P_1 en une courbe tracée par Q, rationnelle de degré $(n+1)$ avec un point singulier d'ordre n en O_2.

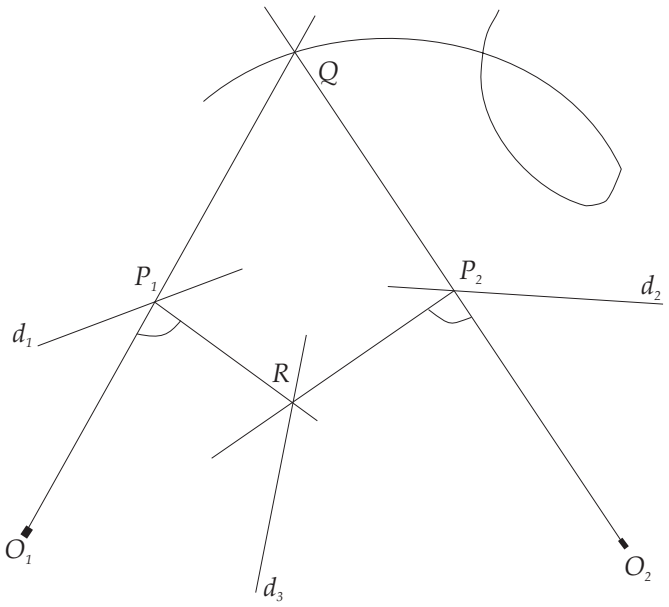


Plus loin, Maclaurin propose une autre généralisation de l'instrument, en remplaçant les droites fixes par des courbes fixes données [20].

Newton estimait la description des cubiques sans point double comme étant l'un des problèmes les plus difficiles [21]. Et en effet les deux instruments que nous venons de décrire, ainsi que leurs variantes, ne permettent de construire que des courbes rationnelles puisque le mouvement de l'instrument est entièrement déterminé et guidé par le mouvement d'un point le long d'une droite, mouvement qui réalise le paramétrage rationnel de la courbe tracée. Pour décrire les cubiques sans point double, il faut donc des instruments qualitativement différents, dont le mouvement ne puisse pas s'interpréter comme paramétrage rationnel de la courbe tracée. Maclaurin y parvient [22]. Le problème posé en ces termes, n'est-ce pas un défi pour l'imagination ? Pourtant, à voir l'instrument proposé par Maclaurin,

19. *Cf.* [220], *pars* I, proposition XXI.
20. *Cf.* [220], *pars* II, *sectio* II.
21. *Cf.* Newton, *Enumeratio*, § XXXIII.
22. *Cf.* [220], *pars* I *sectio* III, proposition XIV.

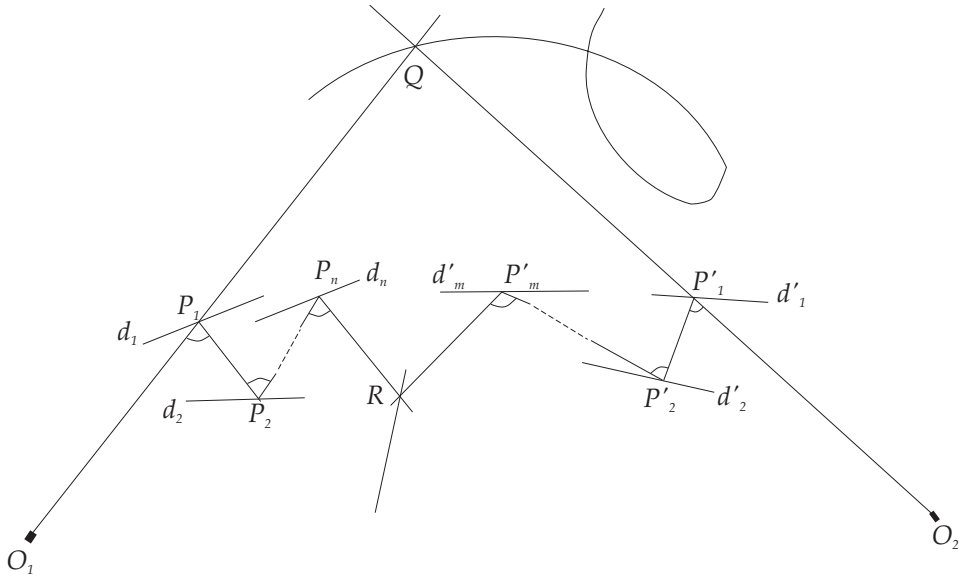
on manque facilement d'observer la différence qualitative avec les instruments précédents. Il est possible que Maclaurin lui-même n'ait pas vu aussi clairement que nous cette différence subtile. Toujours est-il qu'un instrument est conçu (peut-être pour la première fois ?) pour tracer des courbes non rationnelles. Soit deux points fixes O_1, O_2, et deux couples de droites articulées suivant des angles constants de sommets P_1 et P_2 parcourant deux droites fixes d_1 et d_2, un des côtés de chaque angle passant par chacun des points O_1, O_2 et les deux autres se coupant en un point R parcourant une droite fixe d_3.



Alors le point $Q = (O_1P_1) \cap (O_2P_2)$ décrit une quartique avec deux points doubles, en O_1 et O_2. S'il arrive que (O_1P_1) et (O_2P_2) se confondent lors du mouvement, elle dégénère en une cubique sans point double. Mais il n'y a pas de point privilégié qui guide le mouvement et détermine entièrement la position de l'instrument. Si on libère le point P_1, on a bien un transformation rationnelle du plan $Q \mapsto P_1$ qui éclate O_1 et O_2. Mais un seul passage de P_1 le long de la droite d_1 ne suffira pas à décrire toute la courbe (qui possède plusieurs composantes connexes dans $\mathbb{R}^2$, par exemple une composante non bornée et une « ovale »). A chaque position de P_1 sur d_1 correspond une position de R, et donc deux positions distinctes de P_2 sur d_2, ainsi que deux points de la courbe tracée, sauf dans le cas où le cercle RP_2O_2 est tangent à la droite d_2 (on peut montrer que cela se produit exactement 4 fois), auquel cas les deux points sont confondus. L'application $Q \mapsto P_1$ est donc un revêtement fini de degré 2 de la droite d_1, avec 4 points de ramification. Le théorème de Hurwitz affirme que la courbe est alors de genre 1 : il s'agit donc d'une quartique avec exactement deux points doubles en O_1 et O_2, ou bien d'une cubique sans point double.

Maclaurin généralise cet instrument pour tracer des courbes de degrés élevés [23], en remplaçant les côtés (P_1R) et (P_2R) par des lignes brisées $P_1P_2...P_mR$ et $P'_1P'_2...P'_nR$ à angles constants dont les sommets P_1, P_2,..., P_m parcourent des droites fixes d_1, d_2,..., d_m et les sommets P'_1, P'_2,..., P'_n des droites fixes d'_1, d'_2,..., d'_n.

23. *Cf.* [220], *pars* I, proposition XXIV.



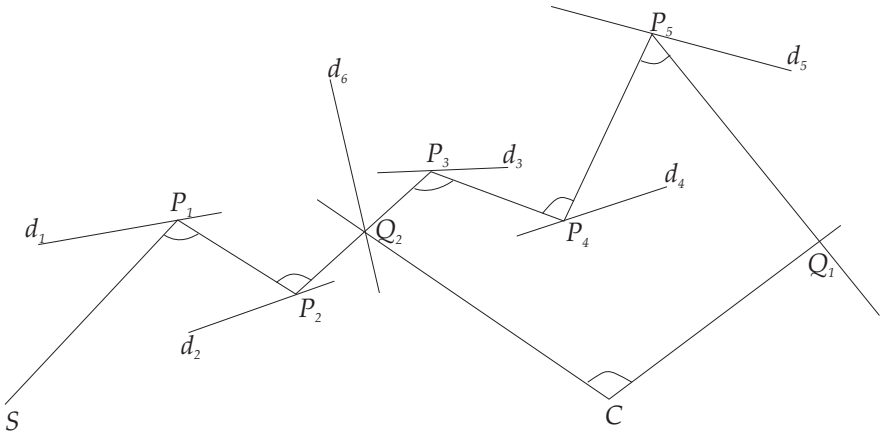
Le point Q décrit alors une courbe de degré $(n+m+2)$ ayant un point d'ordre $(m+1)$ en O_1 et un point d'ordre $(n+1)$ en O_2. En libérant le point P_1, on obtient à nouveau une transformation rationnelle $Q \mapsto P_1$ qui éclate O_1 et O_2. Puis il généralise encore cet instrument en faisant parcourir aux points des courbes fixes au lieu des droites fixes [24].

Nous avons décrit trois des instruments conçus par Maclaurin. Nous devons à présent décrire ses méthodes démonstratives. Pour démontrer quel est le degré d'une courbe tracée et quels sont ses points singuliers, Maclaurin fait appel tantôt à l'analyse (il calcule l'équation de la courbe décrite), tantôt à une méthode purement géométrique. Cette méthode fait intervenir théorème de Bézout et multiplicité d'intersection. Ajoutons qu'il n'a jamais recours au discriminant pour déterminer les points singuliers, et il ne peut donc pas les énumérer exhaustivement : il se contente de regarder comment se comporte la courbe en certains points particuliers de l'instrument (par exemple les points fixes). Pour mieux comprendre la méthode géométrique de Maclaurin, nous en donnerons deux exemples. Commençons par un exemple particulièrement simple, l'instrument attribué à Newton [25]. On sait que $\phi_{S,C,B}$ transforme les droites en sections coniques. Pour déterminer le degré de l'image $\phi(\mathcal{C}_n)$ d'une courbe $\mathcal{C}_n$ de degré n, on regarde l'intersection de $\phi(\mathcal{C}_n)$ avec une droite *générique* d. Or l'image inverse de la droite d, soit $\psi_{S,C,B}(d)$, obtenue en faisant parcourir la droite d au point P, est une section conique $\mathcal{C}_2$. A chaque point d'intersection de d et de $\phi(\mathcal{C}_n)$ correspond un point d'intersection de $\mathcal{C}_2$ et de $\mathcal{C}_n$, il y en a $2n$, donc $\deg \phi(\mathcal{C}_n) = 2n$. Pour déterminer la nature des points singuliers, Maclaurin remarque que quand $(SP) = (SC)$, la courbe $\phi(\mathcal{C}_n)$ passe par C. A cette position correspond une position de la droite (SQ), qui coupe $\mathcal{C}_n$ en n points. Donc n arcs de la courbe tracée passent au point C.

Il arrive aussi à Maclaurin d'appliquer le théorème de Bézout dans des situations plus compliquées, comme celle de l'exemple suivant [26] que nous détaillerons davantage. Soient deux points fixes S et C. Soit un premier système articulé de n droites mobiles (SP_1), (P_1P_2),...,(P_{n-2}, P_{n-1}), (P_{n-1}, Q_1), la première droite étant en rotation autour du point S, les points P_1,...,P_{n-1} devant parcourir des droites fixes (d_1),...,(d_{n-1}), et les angles $\widehat{SP_1P_2}$,...,$P_{n-2}\widehat{P_{n-1}}Q_1$ restant constants pendant le mouvement. Soit Q_2 l'intersection de la r-ième droite mobile $(P_{r-1}P_r)$ avec une n-ième droite fixe (d_n). Soit un second système articulé (en C) de deux droites mobiles (Q_2C),(CQ_1) dont la première passe par le point Q_2, et la deuxième forme un angle constant $\widehat{Q_2CQ_1}$ avec la première. La position des deux systèmes mobiles, et donc la position du point Q_1 d'intersection des droites $(P_{n-1}Q_1)$,(CQ_1) sont entièrement déterminées par la direction de la droite (SP_1). La propo-

24. *Cf.* [220], *pars* II, *sectio* II.
25. *Cf.* [220], *pars* II, *sectio* I, proposition III.
26. *Cf.* [220], *pars* I, proposition XXIII.

sition XXIII affirme que la courbe $\mathcal{C}$ décrite par le point Q_1 pendant le mouvement est une courbe algébrique de degré $(n+r)$.



Pour le démontrer, Maclaurin cherche en combien de points cette courbe coupe une droite générique (D) du plan [27]. A cet effet, il introduit un autre point mobile Q_3, intersection des droites mobiles (SP_1),(Q_2C). Dans une proposition antérieure (proposition XXI), il a montré que le point Q_3 ainsi défini décrit une courbe $\mathcal{C}'$, de degré $(r+1)$, passant par S et C, avec $\mu_{\mathcal{C}'}(S) = r$ et $\mu_{\mathcal{C}'}(C) = 1$. Ensuite, il libère le point Q_2 (autrement dit, il n'impose plus aux droites mobiles $(P_{r-1}P_r)$ et (CQ_2) de concourir avec la droite fixe (d_n)), et impose au point Q_1 de parcourir la droite (D). La même proposition antérieure indique que le point Q_3 décrit alors une courbe $\mathcal{C}''$ de degré $(n+1)$, avec $\mu_{\mathcal{C}''}(S) = n$ et $\mu_{\mathcal{C}''}(C) = 1$. Il est alors facile de voir que lorsque Q_3 parvient aux points d'intersection de $\mathcal{C}'$ et $\mathcal{C}''$, Q_1 est à l'intersection de (D) et $\mathcal{C}$, et réciproquement. Il y a cependant une exception, lorsque Q_3 parvient au point S ou au point C, car dans ce cas, la position du point Q_1 n'est pas déterminée de manière unique. Il ne faut donc pas compter les points S et C (*non sunt considerandi*). Or le théorème de Bézout affirme que les deux courbes se coupent en $(n+1)(r+1)$ points, comptés avec multiplicité, dont les points C et S. On a donc :

$$\begin{aligned}
&\mathrm{Card}((D)\cap\mathcal{C})\\
&= \mathrm{Card}(\mathcal{C}'\cap\mathcal{C}'' - \{S,C\})\\
&= \deg\mathcal{C}'.\deg\mathcal{C}'' - I(\mathcal{C}',\mathcal{C}'';S) - I(\mathcal{C}',\mathcal{C}'';C)\\
&= (n+1)(r+1) - nr - 1\\
&= n+r
\end{aligned}$$

Donc $\deg\mathcal{C} = n+r$.

Pour justifier de manière moderne cette démonstration, remarquons d'abord que lorsque Q_2 est libre, le mécanisme décrit en fait deux transformations rationnelles du plan (non inversibles) $\phi_1 : Q_3 \mapsto Q_1$ et $\phi_2 : Q_3 \mapsto Q_2$; ϕ_1 n'est définie ni en S ni en C (ce sont des « points singuliers » de la transformation ϕ_1). Alors $\mathcal{C}'' = \overline{\phi_1^{-1}((D))}$ et $\mathcal{C}' = \overline{\phi_2^{-1}((d_n)) \cap \phi_1^{-1}(\mathcal{C})}$. De plus $\phi_1{}_{|\mathcal{C}'} : \mathcal{C}' \to \mathcal{C}$ est inversible (c'est-à-dire birationnelle). L'application ϕ_1 « éclate » [28] les points S et C ; ainsi, bien que $C \in \mathcal{C}'\cap\mathcal{C}''$, il ne correspondrait à C un point de $(D)\cap\mathcal{C}$ que si les deux tangentes à $\mathcal{C}'$ et $\mathcal{C}''$ en C coïncidaient, ce qui, en général, n'arrive pas. De même pour S. Donc en général, ϕ_1 induit une bijection $\mathcal{C}'\cap\mathcal{C}'' - \{S,C\} \longrightarrow (D)\cap\mathcal{C}$ (en toute rigueur, il faudrait montrer que $\mathcal{C}'$ ne passe pas par d'autres points singuliers de la transformation ϕ_1). Ajoutons que la courbe $\mathcal{C}$ peut être paramétrée par le coefficient directeur de la première droite mobile (SP_1) et est donc rationnelle.

L'exemple de la proposition XXIII nous a permis d'expliquer un argument démonstratif, utilisé systématiquement par Maclaurin dans des démonstrations purement géométriques, pour déterminer le degré d'une courbe algébrique. Il utilise aussi la formule $(*)$ (*supra* p. 87) de manière réciproque,

27. *ducatur recta quaevis... et investigandum sit quoties curva occurrere potest huic rectae.*
28. Sur la notion d'« éclatement » ou *blowing up*, *cf.* [140], I.4, p. 28–30.

pour déterminer l'ordre de multiplicité d'un point singulier (dans le corollaire I de la proposition XXIII, il montre ainsi que $\mu_{\mathcal{C}}(C) = n$). Nous avons vu ci-dessus qu'il utilisait cette formule, non pas avec le signe $\geq$, mais avec le signe $=$. Il sait cependant que lorsqu'une branche de la première courbe est tangente à une branche de la seconde au point d'intersection, la multiplicité d'intersection augmente [29]. Pour obtenir une formule exacte, et dotée d'une interprétation géométrique, de la multiplicité d'intersection de deux courbes planes, il faudra attendre la notion de « points infiniment voisins » introduite par Max Noether à la fin du XIX$^{\text{e}}$:

$$I(\mathcal{C}, \mathcal{D}; P) = \sum \mu_{\mathcal{C}}(P_i).\mu_{\mathcal{D}}(P_i)$$

où la somme s'étend à tous les points P_i « infiniment voisins » du point P sur les courbes $\mathcal{C}$ et $\mathcal{D}$ (nous y reviendrons, *cf. infra* p. 208).

Ajoutons que Maclaurin semble accorder un rôle différent aux deux types de démonstrations (démonstration purement géométrique *versus* usage des coodonnées cartésiennes sans avoir recours au théorème de Bézout). Il donne ainsi, pour certaines propositions, deux démonstrations distinctes [30].

Nous l'avons dit ci-dessus, l'un des objectifs de Maclaurin était de fournir des procédés pour décrire, par le mouvement, des courbes de degré donné passant par des points donnés. C'est ainsi qu'il fut conduit à se poser la question du nombre de points qui déterminent de manière unique une courbe de degré donné. Et c'est ainsi qu'il se heurta au « paradoxe de Cramer ». Soit une courbe d'équation $P(x, y) = 0$, de degré n. Le nombre de termes d'un polynôme $P(x, y)$ de degré n est

$$\frac{(n+2)(n+1)}{2} = \frac{n(n+3)}{2} + 1$$

On peut supposer que l'un des coefficients (par exemple le terme constant) est égal à 1 ; l'équation d'une courbe de degré n dépend donc de $N = \frac{n(n+3)}{2}$ coefficients indéterminés. Imposer à cette courbe de passer par le point (x_0, y_0) revient à écrire l'équation $P(x_0, y_0) = 0$, linéaire en les coefficients du polynôme P. En demandant qu'une telle courbe passe par N points distincts donnés, on obtient donc un système de N équations linéaires entre ces N coefficients. Or un système de N équations linéaires à N inconnues a *en général* une solution unique. Ce principe était communément admis au XVII$^{\text{e}}$ siècle [31]. D'autre part, le théorème de Bézout affirme que deux courbes de degré n se coupent en n^2 points. Ces deux assertions semblent se contredire pour $n \geq 3$. Pour $n = 2$, il n'y a pas de problème particulier : deux coniques se coupent en quatre points (donc quatre points ne suffisent pas à déterminer une conique), et une conique est déterminée, en général, de manière unique par cinq de ses points. Mais dans le cas $n = 3$, $N = n^2 = 9$, la première assertion dit que 9 points suffisent à déterminer une cubique de manière unique, et la deuxième assertion soutient au contraire que *deux* cubiques se coupent en 9 points. Pour $n > 3$, la situation empire, car $\frac{n(n+3)}{2} < n^2$. Mais Maclaurin n'exprime pas davantage sa surprise qu'en ces termes : « neuf points ne suffisent pas à déterminer une cubique, alors que dix points sont trop » [32]. C'est Cramer qui utilisera, en 1750, le terme de « paradoxe » pour désigner ce phénomène (*cf. infra* section 7.2).

C'est encore à la recherche des conditions que l'on peut imposer à une courbe à tracer, de degré donné, que Maclaurin pose la question du nombre maximal de singularités d'une telle courbe. Comme le « paradoxe de Cramer », les résultats auxquels il parvient dans ce domaine feront l'objet

29. Ainsi dans la proposition XXVI de la première partie : *nisi forte Lineae se mutuo contingant..., quo in casu duae intersectiones abeunt in unum punctum contactus, et altera earum intersectio erit necessario aliqua earum quas non supposuimus esse in punctis C, S.*

30. C'est le cas de la proposition XXIII dont nous avons seulement rapporté ci-dessus la démonstration géométrique. D'autre part, la proposition XXI de la première partie, qui, nous l'avons vu, est utilisée dans la démonstration géométrique de la proposition XXIII, est démontrée seulement par l'analyse.

31. Comme nous l'avons dit, l'on rencontre au XVII$^{\text{e}}$ siècle des énoncés généraux sur les systèmes d'équations polynomiales, qui relient le nombre de solutions d'un système au nombre des équations et des inconnues, sans toutefois donner de critère rigoureux permettant de repérer les cas particuliers. Citons par exemple Descartes ([90], t. VI, p. 372) : « Et on doit trouver autant de telles Equations qu'on a supposé de lignes qui estoient inconnuës... Après cela,... il se faut servir par ordre de chascune des Equations..., pour expliquer chascune de ces lignes inconnuës, et faire ainsi, en les demeslant, qu'il n'en demeure qu'une seule... ».

32. *Cf.* [220], *pars* II, *sectio* V, colorollaire II au lemme III.

de nombreuses recherches au XIX^{e} siècle. On a vu que certains des procédés mécaniques envisagés par Maclaurin offrent un paramètrage rationnel de la courbe construite, mais pas tous : certains d'entre eux tracent des cubiques sans point double, courbes de genre 1, pour lesquelles n'existe aucun paramétrage rationnel. Le mécanisme envisagé est, par conséquent, d'une nature différente. Maclaurin était conscient de l'importance du nombre de singularités dans la description organique des courbes, comme l'indique la division en sections de la première partie de son ouvrage, la section II portant sur les cubiques avec point double, et la section III sur les cubiques sans point double et les quartiques : le degré ne suffit plus à classer les courbes algébriques. Pourtant, Maclaurin parvient rarement à une étude exhaustive des singularités de la courbe construite. La raison en est simple : il n'avait pas les moyens de déterminer l'ensemble des singularités d'une courbe donnée (ce qu'eût rendu possible la connaissance du discriminant de la courbe). Il se contentait donc d'étudier le comportement de la courbe au voisinage de certains points (dont les points fixes du mécanisme utilisé pour tracer), et les points singuliers imaginaires sont bien sûr exclus (remarquons cependant que le point singulier d'une cubique est toujours réel). Ainsi, dans la proposition XXIII, il montre que $\mu_{\mathcal{C}}(C) = n$, mais ne remarque pas l'existence d'autres points singuliers [33].

Dans ce contexte, la connaissance du nombre maximal de singularités d'une courbe de degré donné constitue un progrès important. On sait aujourd'hui qu'une courbe algébrique plane possède un paramètrage rationnel si et seulement si elle a le nombre maximal de points singuliers que peut avoir une courbe de son degré. Maclaurin savait qu'une courbe de degré $2n$, ayant trois singularités d'ordre n et une quatrième d'ordre $(n-1)$, ne peut en avoir d'autre [34]. Nous savons aujourd'hui qu'une telle courbe est donc susceptible d'un paramètrage rationnel. Et Maclaurin montre en effet, dans la dernière proposition de l'ouvrage (proposition XXVII de la seconde partie) comment tracer une telle courbe, lorsque ses 4 points singuliers sont donnés, ainsi que $2n$ points simples ; la méthode utilisée, que nous avons déjà décrite et que Maclaurin a tenté de généraliser sans succès (il le dit lui-même), donne un paramètrage rationnel.

Finalement, Maclaurin est le premier, à notre connaissance, à avoir montré qu'une courbe $\mathcal{C}_n$ de degré n a au plus $\frac{(n-1)(n-2)}{2}$ points doubles [35] : supposons avec lui qu'il y en ait $\frac{(n-1)(n-2)}{2}+1$, il existerait alors une courbe $\mathcal{C}_{n-2}$ de degré $(n-2)$ qui passe par tous ces points doubles et par $(n-3)$ points simples choisis arbitrairement sur $\mathcal{C}_n$ (car l'équation générale de degré $(n-2)$ a $\frac{(n-1)(n-2)}{2}+1+(n-3) = \frac{(n-2)(n+1)}{2}$ coefficients indéterminés). On compterait alors le nombre suivant de points d'intersection :

$$n-3+2\left(\frac{(n-1)(n-2)}{2}+1\right) = n(n-2)+1$$

Mais le théorème de Bézout en prédit au maximum $n(n-2)$.

On sait aujourd'hui que l'on peut classer les courbes algébriques planes selon le genre g, un entier positif dépendant du degré n de la courbe et du nombre de ses points doubles, les singularités d'ordre k comptant chacune, en général, pour $\frac{k(k-1)}{2}$ points doubles [36] :

$$g = \frac{(n-1)(n-2)}{2} - \text{nombre de points doubles} \geq 0$$

Nous reparlerons du genre dans un chapitre ultérieur, section 13.3.

Si l'on voulait à présent se conformer à l'historiographie déjà existante sur l'élimination et le théorème de Bézout, il nous faudrait quitter Maclaurin pour aborder les contributions de Cramer,

33. Ce qui le conduit à affirmer, dans les corollaires IV et VI de la proposition XXIII, qu'une variante du mécanisme décrit ci-dessus permet de tracer une quartique ayant un point double en C, qui, dans un cas particulier, dégénère en une cubique sans point double. En fait, la quartique possède trois points doubles (puisqu'elle est rationnelle), et dégénère en une cubique avec point double. Il y a une erreur semblable dans la proposition XVII (de la première partie). Ces erreurs ont été commentées par C. Tweedie [314]. Ajoutons qu'elles figuraient déjà dans le premier article de Maclaurin sur ce sujet, en 1719 (*cf.* [219], figures 6 et 9).

34. *Cf.* [220], partie II, section V, lemme III, corollaire V.

35. *Cf.* [220], p. 137.

36. *Cf.* par exemple [140], IV, exercice 1.8 «p_a of a Singular Curve ».

Euler, puis Bézout. Ces auteurs ont donné plusieurs démonstrations et généralisations du « théorème de Bézout » que nous avons rencontré dans la *Geometria organica*. Maclaurin n'y donne en effet aucune démonstration générale du théorème. Il se contente de le démontrer dans le cas où l'une des deux courbes est de degré 2 ou 3, et ceci en calculant explicitement l'équation finale résultant de l'élimination d'une des deux inconnues. Il avoue son échec dans le cas général [37]. Comme nous le verrons au chapitre 6, Euler tentera de préciser les conditions de possibilité de l'interprétation géométrique du théorème de Bézout, en étudiant sur des exemples le rapport entre élimination algébrique et intersection de courbes (dans le plan) ou de surfaces (dans l'espace), et en insistant sur le besoin de compter, parmi les points d'intersection, des « points imaginaires ». Il exposera en particulier le phénomène des lignes de contact entre deux surfaces, et des points de contact, qui ne sont que des points réels de courbes d'intersection imaginaires de deux surfaces. Mais de Newton à Bézout, le théorème de Bézout garde la marque de son origine algébrique. Les questions que se posent ces auteurs sont celle du degré de l'équation finale résultant de l'élimination, et celle de la méthode d'élimination à employer. Maclaurin, utilisant ce théorème comme *principe* dans des démonstrations de nature purement géométrique, semble donc faire figure d'exception. Suivons encore un peu la trajectoire du théorème de Bézout en géométrie.

A la *Geometria organica* succédèrent d'autres recherches de Maclaurin sur la description des courbes. Maclaurin fut stimulé dans ces travaux par un théorème dû à Pappus, que lui communiqua en 1722 son collègue Robert Simson (1687–1768) qui tentait alors de restituer le « Traité des porismes » perdu d'Euclide [38]. Voici le théorème dû à Pappus :

> Si trois points situés sur une seule droite d'une figure convexe ou non convexe sont donnés, et si les points restants, à l'exception d'un seul, sont liés à des droites données de position, ce seul point est aussi lié à une droite donnée de position. [39]

Maclaurin exprime ce théorème sous une forme plus générale (voir figure suivante) :

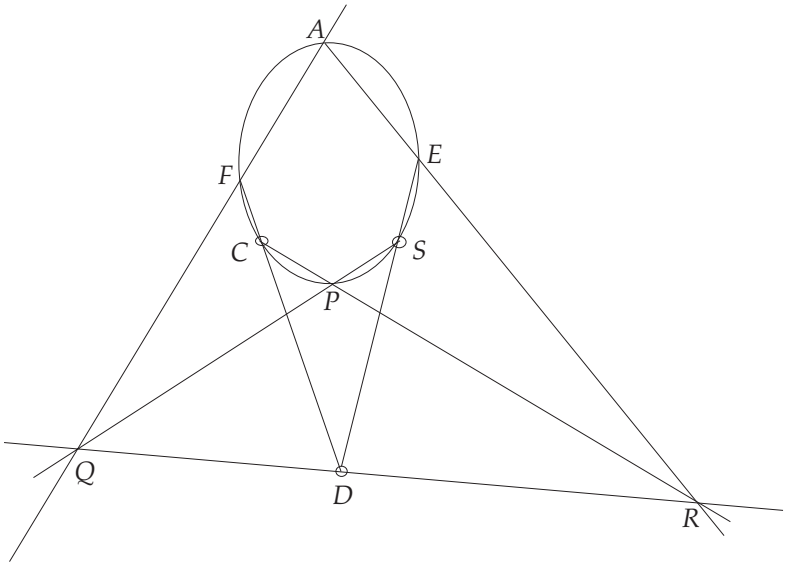
> Supposez que les trois pôles soient C, S et D, et que des lignes droites CR, SQ, QDR pivotent autour de ces pôles. La ligne qui pivote autour de D ne sert qu'à guider le mouvement des deux autres, de telle sorte que son intersection avec chacune d'elles parcourant un droite fixe, leur intersection mutuelle décrit un lieu, que l'on démontre être une section conique... L'on montre ensuite, que si C, S et D sont situés sur une même ligne droite, le point P décrit une ligne droite... [40]

37. *Cf.* [220], partie II, section V, lemme III, corollaire I : « Omnes altiorum curvarum casus quos adhuc tentavimus idem probant ; universalem vero hujus rei demonstrationem adhuc frustra quaesivimus propterea quod difficile sit divisores in arduis aequationibus invenire ». Quant au problème des « diviseurs » (*divisores*), peut-être Maclaurin fait-il allusion aux facteurs superflus que l'élimination introduit dans l'équation finale. C'est ce même problème qui retiendra longtemps Euler et Bézout dans des essais infructueux (*cf.* par exemple [26]).

38. Maclaurin écrit : « in June or July, 1722, upon a Hint I got from Mr. Simpson of Mr. Pappus's Porisms, I saw that what he has there ingeniously demonstrated, might be considered as a case of the above mentionned description of a Conick Section,... ».

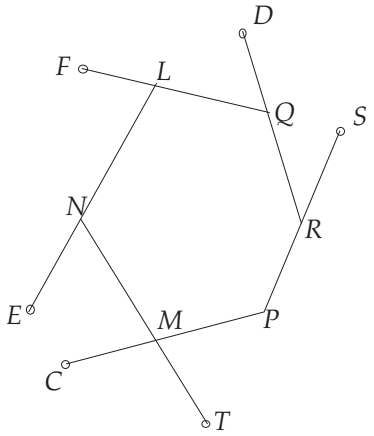
39. *Cf. La collection mathématique* de Pappus, [83], t. 2, p. 488.

40. *Cf.* [221], p. 151–153, ainsi que la figure 6.



On reconnaît dans cette figure le théorème de Pascal sur l'hexagramme inscrit dans une section conique. Au triangle PQR dont les côtés pivotent sur les pôles C,S,D, Maclaurin substitue ensuite un polygone d'un nombre quelconque de côtés, soit six pour fixer les idées (voir figure suivante), pivotant autour des pôles C,D,E,F,S,T, et il énonce le théorème général suivant :

> Supposez que L, Q, R, M, N parcourent des courbes de degrés respectifs m, n, r, s, t, alors le lieu de P est une courbe de degré au plus $2mnrst$. [41]



Maclaurin avait déjà, dans la *Geometria organica*, décrit des procédés pour tracer des courbes dont le degré dépende du produit des degrés d'un nombre quelconque de courbes données. Mais le théorème que l'on vient d'énoncer, par sa simplicité, sa parenté avec le théorème de Pappus, et son caractère projectif (contrairement aux procédés utilisant la rotation d'angles constants de la *Geometria organica*) était promis à un avenir meilleur. Ce théorème, parmi d'autres, fit l'objet d'une querelle de priorité entre Maclaurin et W. Braikenridge (1700–1762). Ce dernier publia en 1733, dans un ouvrage intitulé *Exercitatio geometrica de descriptione linearum curvarum*, de nombreux théorèmes sur le tracé des courbes, démontrés par le théorème de Bézout et la méthode géométrique qu'employaient Maclaurin dès 1721 comme nous l'avons vu. Nous n'entrerons pas dans les détails de cette controverse qui amena Maclaurin à publier son théorème en 1735 ; bien qu'elle suggère qu'il avait alors des résultats en réserve, non publiés mais enseignés, il faudrait pour le vérifier

41. *Cf.* [221], p. 156.

mieux connaître le contenu de ses cours, aujourd'hui perdus [42]. Indiquons simplement qu'à cette occasion, Braikenridge annonça qu'il connaissait une démonstration du théorème de Bézout, conçue par son collègue George Campbell (?–1766), mais pas encore rendue publique, à cause de la trop grande modestie de son auteur [43]. Dans sa riposte, publiée en 1735, Maclaurin affirma aussi avoir une démonstration, « générale », du théorème de Bézout [44], et même l'avoir déjà « imprimée » en 1722, dans un « supplément » à sa *Geometria organica*, supplément qu'il n'aurait cependant jamais publié, « ayant été engagé dans des travaux d'une nature différente, et dans des recherches sur d'autres sujets depuis lors » [45]. Il est d'ailleurs intéressant que dans ce texte de 1735, Maclaurin s'excuse de retarder encore la publication de son supplément, et s'en justifie par un goût pour les démonstrations géométriques, plus longues à produire que les démonstrations analytiques :

> As to the Demonstrations, it would take some time to put them in a proper Form to be published. You who have so nice a Taste of Demonstrations, will easily allow, that it ought not to be done in a hurry. I could send those that are Algebraick easily ; but do not care to send those that are Geometrical, till I have leisure. [46]

Nous n'avons pas trouvé d'autre trace des démonstrations de Campbell et de Maclaurin. Ajoutons toutefois que, peu avant la controverse avec Braikenridge, Campbell lui-même avait été engagé dans une querelle de priorité avec Maclaurin, au sujet de la « règle de Newton » sur le nombre de racines réelles d'une équation. Les deux mathématiciens montrèrent chacun à cette occasion un talent en algèbre qu'aurait pu illustrer une démonstration du théorème de Bézout [47].

Quant au théorème ci-dessus sur les polygones, il survécut à cette controverse et devint le coeur du dernier chapitre du *Traité des propriétés projectives des figures* de Poncelet, en 1822 (section IV, chapitre III). Poncelet utilise systématiquement, dans cette section, le théorème de Bézout et la méthode géométrique de Maclaurin pour déterminer le degré de courbes tracées mécaniquement [48]. Il fait profession de n'avoir plus recours à l'usage des coordonnées. Aussi donne-t-il une démonstration nouvelle, géométrique, du théorème de Bézout. Tout d'abord, son « principe de continuité » lui permet de définir le degré d'une courbe algébrique comme étant le nombre d'intersections de cette courbe avec une droite quelconque du plan (et non plus, comme chez Maclaurin, avec ce que nous voudrions aujourd'hui désigner par l'expression de droite *générique*) :

> Il est essentiel de remarquer qu'en vertu du principe de continuité, une droite, tracée arbitrairement dans le plan d'une courbe d'un certain degré, doit toujours être censée rencontrer cette courbe en un nombre de points *réels*, *imaginaires*, *multiples* ou *situés à l'infini*, égal au nombre qui exprime le degré de cette courbe... [49]

Voici donc la démonstration du théorème de Bézout, pour deux courbes planes de degré 2 (elle se généralise immédiatement à des courbes de degrés quelconques) :

---

42. Sur cette controverse, *cf.* l'article [229] de S. Mills, qui conclut que Maclaurin ne se sentait pas concerné par une question de priorité (Braikenridge ayant de toute façon *publié* le premier), mais se défendait plutôt contre une accusation injuste d'ignorance : « Here then is the reason for Maclaurin's having taken issue with Braikenridge : not that of priority as such, but the insult that Braikenridge implied that Maclaurin did not know of the theorems, when in fact he had been teaching them since 1725 » ([229] p. 237).

43. *Cf.* [38], préface : « *Cl. Geometra D. Georgius Campbel* inter plurima alia eximia inventa, quae ob summam ejus modestiam apud eum latent, hujus theorematis praeclaram habet demonstrationem, in lucem brevi ut spero prodituram ». Sur la personne de George Campbell, *cf.* [325].

44. *Cf.* [221] p. 148.

45. *Cf.* [221] p. 144. Maclaurin décrit alors ainsi le théorème de Bézout : « In the first part of the supplement, there is a general demonstration given of the theorem, that if two lines of the orders or dimensions expressed by numbers m and n, be described in the same plane, the greatest number of points in which these lines can intersect each other, will be mn, or the product of the numbers which express the dimensions of the lines, or the orders to which they belong ». Ce supplément n'a pas été retrouvé.

46. *Cf.* [221], p. 147.

47. Il y a toutefois un témoignage de R. Simson, qui n'en dit pas beaucoup plus que ce que nous savons : « Just now I was looking some Theorems yow wrote down in the first pages of a book of mine that is dated in that year [1720], which yow told me were Lemmas to that proposition concerning the Number of points 2 curves can intersect one Another... », lettre de R. Simson à Maclaurin, 27 juillet 1736, cité par [229].

48. *Cf.* [254], t. 1, section IV, par exemple § 475.

49. *Cf.* [254], t. 1, § 539.

> Si l'on suppose que l'une des deux courbes change de forme, par succession insensible, de façon qu'en conservant toujours le même degré ou la même nature, elle tende sans cesse à dégénérer en deux lignes droites ; par suite de la loi de continuité, elle devra aussi sans cesse conserver le même nombre de points d'intersection réels, imaginaires, etc., avec l'autre ; or il est évident que, à la limite, ce nombre ne saurait surpasser quatre, d'après la définition même adoptée pour les lignes du second ordre ; donc pareillement les courbes proposées ne peuvent avoir plus de quatre points communs, et, si elles en ont réellement cinq, elles se confondent nécessairement en une seule et même courbe. [50]

Au § 542, Poncelet dresse un bilan des conséquences du théorème de Bézout et du théorème de Maclaurin sur les polygones : ils conduisent « immédiatement à la propriété de l'hexagramme mystique de Pascal, et par suite à toute la théorie des pôles et polaires des sections coniques, théorie d'où dérive aussi l'élégant théorème de M. Brianchon sur l'hexagone circonscrit... », et permettent d'établir « sur-le-champ toute la théorie des centres et axes d'homologie des sections coniques » [51]. Des pans entiers de la géométrie projective trouvent ainsi un fondement unique. Poncelet en tire un argument de poids en faveur de son « principe de continuité » :

> Telle est donc l'influence que peut exercer l'admission de la loi de continuité dans les recherches géométriques, influence qui peut très bien être comparée, ce me semble, à celle qu'y exerce elle-même l'Analyse algébrique, par sa grande généralité.

Enfin, Poncelet propose plusieurs généralisations du théorème de Maclaurin sur les polygones, ainsi en remplaçant les pôles par des courbes fixes autour desquelles doivent rouler les droites mobiles (en restant tangentes à ces courbes). Dans le cas où toutes les courbes que parcourent les sommets d'une part, et toutes les courbes sur lesquelles roulent les droites mobiles d'autre part, sont confondues, il obtient un théorème sur les polygones simultanément inscrits et circonscrits à deux sections coniques :

> Quand un polygone quelconque est à la fois inscrit à une section conique et circonscrit à une autre, il en existe une infinité de semblables qui jouissent de la même propriété à l'égard des deux courbes ; ou plutôt tous ceux qu'on essaierait de décrire, à volonté, d'après ces conditions, se fermeraient d'eux-mêmes sur ces courbes. Et réciproquement, s'il arrive qu'en essayant d'inscrire à volonté, à une section conique, un polygone dont les côtés en touchent une autre, ce polygône ne se ferme pas de lui-même, il ne saurait nécessairement y en avoir d'autres qui jouissent de cette propriété. [52]

La controverse avec Braikenridge laisse penser que Maclaurin avait des résultats en réserve. Le traité *De linearum curvarum proprietatibus* rédigé par Maclaurin à la fin de sa vie n'a pas pour thème la description organique, ni la classification des courbes. Son sujet est ce que les géomètres du XIX$^{\text{ème}}$ siècle appelaient l'« analyse des transversales », et son application aux courbes de degré 2 et 3. Pourtant, on voit toujours la préoccupation pour la description organique en arrière-plan.

Maclaurin expose trois théorèmes sur les transversales, qui se rapportent chacun à une courbe algébrique plane de degré quelconque. Chacun de ces théorèmes fait intervenir les mesures de segments découpés sur des droites (les « transversales »). Ces mesures doivent être comprises comme étant des mesures algébriques (c'est-à-dire qu'elles peuvent être négatives, suivant la position relative des points). Le premier théorème suppose données deux directions. Alors pour tout couple de droites ayant chacune l'une des deux directions prescrites et se coupant en un point P variable, notons M, m, μ,... les points où la première droite rencontre la courbe, et I, K, L,... les points où la seconde droite rencontre la courbe (voir fig. 1 ci-dessous). On a [53] :

$$\frac{PM \times Pm \times P\mu \times ...}{PI \times PK \times PL \times ...} = \text{constante}$$

50. *Cf.* [254], t. I, § 598. Dieudonné a indiqué comment rendre cette démonstration rigoureuse dans [93], § VII,20.

51. Ailleurs, Poncelet ajoute que le théorème de Bézout « pourrait encore servir immédiatement à prouver que toutes les lignes du second ordre sont nécessairement des sections coniques ».

52. *Cf.* [254], t. 1, § 566.

53. *Cf.* [218] § 5. Ce théorème généralise un énoncé sur les coniques qui était déjà connu de Ibn Sahl au X$^{\text{e}}$ siècle, *cf.* [263] p. 299 et 302.

Ce théorème est une généralisation du concept de « puissance d'un point par rapport à un cercle ». Maclaurin applique ce théorème à l'étude de la courbure en un point. Il fixe la transversale (PI) et fait tendre le point P vers I : le rayon de courbure en I est alors la valeur limite d'une expression dépendant de PI et PM, expression qu'il transforme au moyen du théorème [54].

Dans le second théorème, on considère deux droites quelconques passant par P. Notons K, L, M,... les points où la première rencontre la courbe, et k, l, m,... les points où la première rencontre les tangentes à la courbe en ses points d'intersection avec la seconde droite (voir fig. 2 ci-dessous). Alors on a [55] :

$$\frac{1}{PK} + \frac{1}{PL} + \frac{1}{PM} + ... = \frac{1}{Pk} + \frac{1}{Pl} + \frac{1}{Pm} + ...$$

Ce théorème, dû à Cotes, fut communiqué à Maclaurin par un collègue.

Dans le troisième théorème, on fixe un point P. Pour toute droite mobile passant par P, notons D, E, I,... les points où elle rencontre la courbe, et M le point obéissant à la relation :

$$\frac{n}{PM} = \frac{1}{PD} + \frac{1}{PE} + \frac{1}{PI} + ...$$

où n est le degré de la courbe. Alors M parcourt une droite fixe [56]. Maclaurin utilise pour le démontrer la notion de faisceau harmonique de quatre droites (qu'il a pu apprendre dans le traité de La Hire sur les sections coniques). On appelle PM « moyenne harmonique » de PD, PE, PI,... Pour $n = 2$, la droite parcourue par le point M est la polaire de P par rapport à la courbe. Pour n quelconque, le cas particulier où P est à l'infini était connu de Newton, qui avait ainsi élaboré la notion de « diamètre » d'une courbe algébrique.

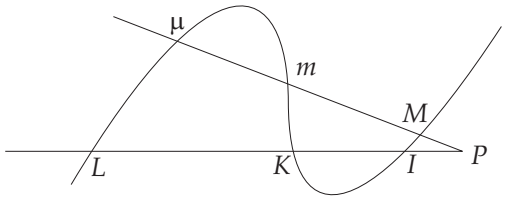


Fig. 1

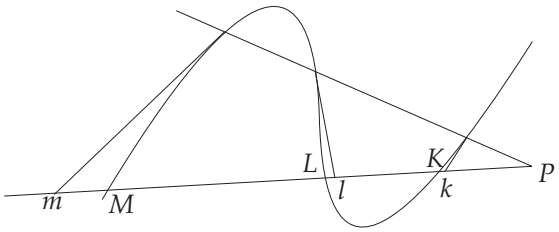


Fig. 2

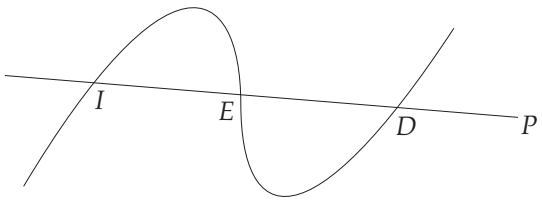


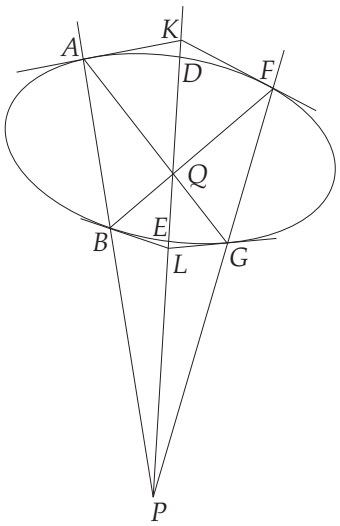
Fig. 3

Maclaurin applique ensuite ses théorèmes aux sections coniques. Il ne prétend pas alors énoncer de résultats nouveaux, mais justifie ainsi sa démarche :

> Ce qui concerne les sections coniques peut être déduit au mieux des propriétés du cercle, qui est la base du cône. Mais pour rendre plus clair l'usage des théorèmes précédents [les théorèmes sur l'analyse des transversales], cela vaudra la peine de déduire aussi ces propriétés de ces théorèmes. [57]

54. *Cf.* [218] § 15.
55. *Cf. ibid.* § 9.
56. *Cf. ibid.* § 27.
57. *Cf. ibid.* § 33.

Par exemple, Maclaurin démontre une proposition importante en géométrie projective des sections coniques : les pôles de droites concourantes sont alignés. Ce théorème avait déjà été démontré par La Hire, mais de manière significativement différente : celui-ci le démontrait d'abord pour le cercle, puis se ramenait au cas du cercle par projection. Maclaurin conçoit une démonstration unique pour toutes les sections coniques [58], au moyen de son second théorème sur les transversales. De même il redémontre le théorème sur l'hexagramme inscrit [59]. A cela s'ajoute un autre théorème, faisant intervenir un quadrilatère inscrit dans une section conique, et que l'on peut aussi interpréter comme procédé de description organique (ce théorème est parfois attribué à Maclaurin) [60]. Soient A, B et F trois points donnés sur une section conique donnée, et soient (AK) et (FK) deux tangentes données à la section conique se coupant au point K.



Faisons tourner une droite d autour du point K et notons $P = d \cap (AB)$, $Q = d \cap (BF)$, et $G = (PF) \cap (AQ)$. Alors le point G décrit la section conique. Maclaurin démontre ce théorème au moyen du théorème sur la transversale d :

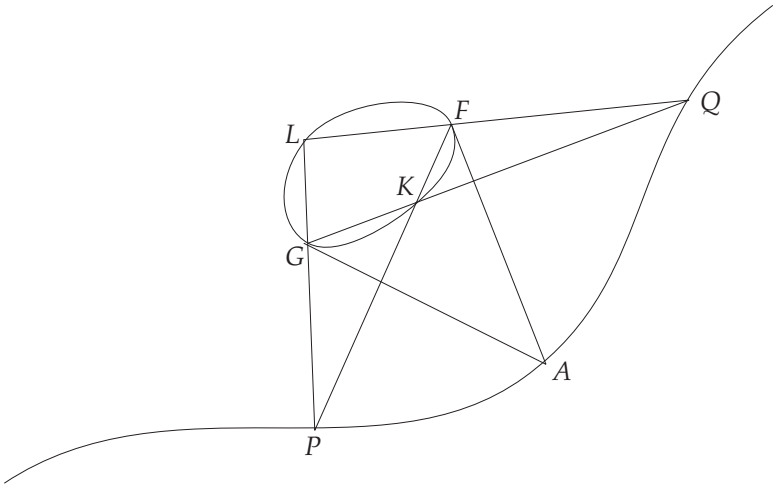
$$\frac{1}{PK} + \frac{1}{PL} = \frac{1}{PD} + \frac{1}{PE}$$

où D et E sont les points d'intersection de d avec la section conique, et L est le point d'intersection des tangentes en B et G.

Enfin, l'application de sa théorie des transversales à l'étude des cubiques conduit Maclaurin à énoncer des résultats que l'on interprète aujourd'hui (depuis Clebsch) comme reflétant une structure de groupe. Par exemple, Maclaurin montre que les tangentiels de trois points alignés sont alignés [61]. Newton connaissait déjà un cas particulier de ce théorème : le cas où les premiers trois points sont sur la droite à l'infini. Maclaurin déduit de ce théorème le fait que les points d'inflexion d'une cubique sont alignés trois à trois [62]. Il est instructif de comparer la démonstration de Maclaurin à celle de J. P. de Gua de Malves. Celui-ci réduit le problème, par projection, à démontrer qu'une hyperbole redondante à deux diamètres en a nécessairement un troisième [63]. Enfin, Gua remarque que Newton n'avait pas conscience de cette propriété des points d'inflexion, car les figures de l'*Enumeratio* la mettent en défaut.

58. *Cf. ibid.* § 27.
59. *Cf. ibid.* § 44.
60. *Cf. ibid.* § 36.
61. Nous appelons « tangentiel » d'un point de la cubique le second point où la tangente en le premier rencontre la cubique ; *cf. ibid.* § 57.
62. *Cf. ibid.* § 68 : *Recta duo puncta flexus contrarii conjungens vel transit per 3^{um} punctum flexus contrarii vel dirigitur in eandem plagam cum crure infinito curvae.*
63. *Cf.* [85] p. 225. Gua donne aussi une démonstration par l'analyse, *ibid.* p. 307.

Une autre application notable par Maclaurin de sa théorie des transversales aux cubiques est un cas particulier de « polygone de Steiner » inscrit dans une cubique. Maclaurin avait déjà annoncé cette découverte dans son *Treatise of fluxions* en 1742 (au beau milieu du traité § 401, en un lieu apparemment hors contexte). Le théorème est démontré au § 84 du *De linearum geometricarum*. Soient F et G deux points ayant même tangentiel A (c'est-à-dire que les tangentes en F et G se rencontrent en un point A de la cubique). Soit P un autre point de la cubique. Alors Maclaurin montre que le quadrilatère $PLQK$ dont les côtés passent alternativement par F et G (voir figure) se referme.



Or dans le *Treatise of fluxions*, Maclaurin prétend aussi, en rapport avec ce théorème, avoir découvert un instrument pour tracer une cubique passant par sept points donnés et ayant deux tangentes données en deux de ces points. Aucune mention n'est faite d'un tel instrument dans le *De linearum geometricarum*. Nous pourrions peut-être expliquer cette situation de la manière suivante. Maclaurin, en 1742, était toujours à la recherche de procédés de description organique des cubiques. Newton avait montré comment tracer une cubique passant par sept points donnés dont l'un est un point double (ce procédé ne s'appliquait donc qu'aux cubiques de genre 0). Maclaurin, venant de découvrir la possibilité d'inscrire un quadrilatère dans une cubique (le quadrilatère de Steiner que nous venons de décrire), a dû conclure un peu trop vite, par analogie avec le théorème du quadrilatère inscrit dans une section conique (qui donne effectivement un procédé de tracé d'une section conique ayant deux tangentes données comme nous l'avons vu ci-dessus), qu'il serait facile d'en déduire un procédé de tracé d'une cubique. Voulant s'assurer au plus vite de la priorité intellectuelle sur un sujet si important, il ajoute un paragraphe à son *Treatise of fluxions* alors sur le point de paraître. Il pensait avoir trouvé ce qu'il avait cherché en vain dans la *Geometria organica* : un instrument unique pour construire toutes les cubiques (indifféremment de genre 0 ou 1), et très simple puisqu'il ne fait intervenir que des lignes droites. Mais il dut, entre 1742 et 1748, revenir de son enthousiasme et réaliser son erreur. Dans le *De linearum geometricarum*, il montre même qu'il n'y a pas, dans la figure ci-dessus, une cubique unique passant par les points F, G, P, Q, K, L, A et ayant (AG) et (AF) pour tangentes. Hélas, il affirme qu'il y en a un « certain nombre », un nombre fini donc (les cubiques sont *numero determinatas* nous dit-il). Le problème serait donc « déterminé » (*adeoque problema esse determinatum*). C'est faux : le fait que (AF) et (AG) soient tangentes à la cubique est une conséquence du fait que le quadrilatère de Steiner est fermé [64], et les conditions restantes énoncent simplement que la cubique doit passer par 7 points donnés, d'où un système de sept équations linéaires homogènes en les 10 coefficients de l'équation de la cubique, système indéterminé et ayant une infinité de solutions.

64. En termes de la loi de groupe sur la cubique, le théorème de Steiner affirme que $2(F-G) \equiv 0$ est une condition nécessaire et suffisante pour que le quadrilatère se referme. Cette condition revient à affirmer que F et G ont même tangentiel.

Chapitre 6

Euler : une théorie de l'intersection

Les recherches d'Euler sur l'élimination font l'objet de trois mémoires présentés à l'Académie des sciences de Berlin ([103], [105] en 1748, et [107] en 1764), ainsi que du chapitre XIX du second tome de l'*Introductio in analysin infinitorum* [104], et du chapitre VI du « Traité abrégé des surfaces » publié en appendice de [104].

6.1 Les trois méthodes d'Euler

La première des méthodes d'élimination exposée par Euler dans [103] est reprise dans [104] et [107].
Méthode n°1 d'Euler :
Soient deux équations de degré m :

$$\begin{cases} Ay^m + By^{m-1} + Cy^{m-2} + ...Uy + V = 0 \\ ay^m + by^{m-1} + cy^{m-2} + ...uy + v = 0 \end{cases}$$

Euler multiplie la première par a, la deuxième par A, et soustrait les produits l'un de l'autre :

$$(Ba - bA)y^{m-1} + (Ca - cA)y^{m-2} + ... + (Va - vA) = 0$$

Puis il multiplie la première par v, la deuxième par V, soustrait les produits l'un de l'autre, et divise par y :

$$(Av - aV)y^{m-1} + (Bv - bV)y^{m-2} + ... + (Uv - uV) = 0$$

Ces deux nouvelles équations sont de degré $(m-1)$; en réitérant ce procédé m fois, on aboutit à une expression de degré 0 en y.

Euler traite seulement avec cette méthode le cas de deux équations de petits degrés (jusqu'au degré 3 dans [103] et [107], jusqu'au degré 4 dans [104]), et n'en donne pas d'exposé général. Il indique aussi comment procéder dans le cas de deux équations de degrés distincts (en abaissant progressivement le degré d'une des deux équations pour se ramener au cas précédent). Mais il remarque que l'expression finale obtenue par cette méthode contient en général des « facteurs inutiles ». Dans le cas $m = 3$, il trouve par exemple un facteur $(Ad - Da)$:

> (...)cette méthode conduit souvent à des équations trop compliquées, qui renferment des facteurs tout à fait inutiles pour le dessein qu'on a en vue. Car dans les équations cubiques, il est évident que le facteur $Ad - Da$ ne satisfait point à la question, puisque l'élimination ne saurait conduire à cette équation $Ad - Da = 0$. Donc, tant que ce facteur est contenu dans l'équation finale, on ne la peut regarder comme juste. [1]

1. *Cf.* [107], p. 201.

Il remarque enfin (au moins pour $m = 3$ et $m = 4$) que l'expression finale délivrée des facteurs inutiles est une somme homogène de produits de $2m$ lettres, m majuscules et m minuscules [2]. Dans [103], il utilise cette méthode pour deux équations cubiques dont les coefficients A,B,C,D,a,b,c,d sont des polynômes en une autre inconnue x, tels que :

$$\begin{aligned}
&d^oA = \alpha,\\
&d^oa = \beta,\\
&d^oA + 3 = d^oB + 2 = d^oC + 1 = d^oD,\\
&d^oa + 3 = d^ob + 2 = d^oc + 1 = d^od
\end{aligned}$$

Le degré de l'équation finale en x délivrée des facteurs inutiles est alors $9 + 3(\alpha + \beta)$. Bézout généralisera ce résultat dans son premier mémoire, en 1764. Dans [107], Euler attribue l'origine de cette première méthode à la méthode évidente de « substitution », utilisable dans le cas de systèmes d'équations linéaires, consistant à réduire un système de n équations à n inconnues en un système de $(n - 1)$ équations à $(n - 1)$ inconnues, en substituant la valeur d'une des inconnues, donnée par une des équations, dans les autres équations. D'autre part, il écrit que cette première méthode d'élimination est « celle dont Mr. Newton paraît s'être servi » pour établir les formules que l'on trouve dans son *Arithmetica universalis*. Nous avons déjà expliqué la cause de cette attribution erronée (*cf. supra* section 3.5) et nous avons remarqué que cette méthode n'est autre que la « méthode de conjonction d'équations inverses » de Leibniz.

En 1748, la motivation d'Euler pour les méthodes d'élimination résidait dans la théorie de l'intersection des courbes et des surfaces algébriques. Ainsi l'objectif de son mémoire [103] était de montrer que deux courbes algébriques d'ordre m et n dans le plan se coupent en au plus mn points (Euler remarquait aussi que les deux courbes peuvent avoir exceptionnellement des composantes communes, et donc une infinité de points communs). Dans l'*Introductio in analysin infinitorum*, il ne formulait pas ce théorème, mais écrivait (sans le démontrer) que deux surfaces d'ordre m et n dans l'espace se coupent suivant une courbe d'ordre au plus mn. La démonstration du mémoire [103], première démonstration connue du théorème de Bézout, était le corollaire d'une méthode d'élimination faisant intervenir les fonctions symétriques des racines :
Méthode n°2 d'Euler :
Soit deux équations de degrés m et n en y :

$$\left\{\begin{array}{l} y^m - Py^{m-1} + Qy^{m-2} - ... = 0\\ y^n - py^{n-1} + qy^{n-2} - ... = 0 \end{array}\right.$$

où P,Q,...,p,q,... sont des polynômes en x tels que :

$$\begin{aligned}
&d^oP + m - 1 = d^oQ + m - 2 = ... = m\\
&d^op + n - 1 = d^oq + n - 2 = ... = n
\end{aligned}$$

Notons A,B,C,...,a,b,c,... les racines de ces deux équations. Alors l'équation résultant de l'élimination de y doit être :

$$\left.\begin{array}{ll} & (A-a)(A-b)(A-c)...\\ \times & (B-a)(B-b)(B-c)...\\ \times & (C-a)(C-b)(C-c)...\\ \times & ... \end{array}\right\} = 0$$

Il s'agit d'une expression homogène de degré mn en les lettres A,B,C,...,a,b,c,..., et symétrique par rapport à chacun des deux ensembles de racines. On peut donc l'obtenir à partir des fonctions symétriques élémentaires des racines, c'est-à-dire à partir des coefficients P,Q,...,p,q,... ; de plus (théorème de Bézout), comme le degré *en* x de chacun de ces coefficients coïncide (par hypothèse) avec son degré *en les racines*, le degré en x de l'expression obtenue est aussi mn.

Cramer donna une démonstration plus détaillée par cette même méthode dans son traité de 1750 (voir chapitre suivant). Cette méthode pose un problème conceptuel : elle suppose l'existence des

2. *Cf.* [104], chap. XIX, § 482.

racines. Mais il s'agissait là d'un principe tacitement admis jusqu'à ce que Gauss se proposât de le démontrer en 1799 dans sa thèse sur le théorème fondamental de l'algèbre. Ainsi en 1746, Euler avait écrit une démonstration (incomplète) du théorème fondamental de l'algèbre, en faisant ce même présupposé de l'existence des racines [3]. D'autre part, comme le montre son *Introductio in analysin infinitorum* [104], il envisageait la possibilité d'« intersections imaginaires » entre deux courbes ou deux surfaces et connaissait des exemples où, à une abscisse réelle d'un point d'intersection, correspondent deux ordonnées complexes conjuguées.

Enfin, Euler introduisit [104] une troisième méthode d'élimination, qui forma la base de son mémoire de 1764 et de celui qu'écrivit Bézout en la même année :
Méthode n°3 d'Euler :
Soit deux équations de degrés m et n :

$$\begin{cases} Py^m + Qy^{m-1} + Ry^{m-2} + ... = 0 \\ py^n + qy^{n-1} + ry^{n-2} + ... = 0 \end{cases}$$

Multiplions la première par $py^{k-m} + ay^{k-m-1} + by^{k-m-2} + ...$ et la deuxième par $Py^{k-n} + Ay^{k-n-1} + By^{k-n-2} + ...$, où k est inconnu pour le moment. Soustrayons les produits l'un de l'autre. On obtient une équation de degré $(k-1)$ en y. Si l'on impose que tous les coefficients des termes non constants s'annulent, on obtient un système de $(k-1)$ équations linéaires en les $(2k-m-n)$ indéterminées A,B,...,a,b,... Afin de résoudre ce système, on impose que le nombre d'équations soit égal au nombre d'indéterminées. Alors $k = m + n - 1$.

Là encore, nous avons déjà vu cette méthode d'élimination sous la plume de Leibniz, dans des manuscrits restés pendant longtemps inédits (*cf. supra* section 2.4). On peut aussi y voir un écho du « théorème de Bézout arithmétique », qui affirme que le plus grand commun diviseur de deux entiers p et q peut s'écrire sous la forme $pr+qs$, et qui était déjà connu de Bachet de Méziriac, en 1612 [4].

L'orientation du mémoire [107] écrit par Euler en 1764 est plus algébrique. Il n'y est plus question de courbes ni de surfaces. Le problème général dont traite ce mémoire s'énoncerait ainsi : deux équations algébriques (à une inconnue) étant proposées, trouver les « déterminations » nécessaires pour que ces équations aient un nombre donné quelconque de racines communes. Par exemple, pour que les équations aient une racine commune, il faut que leurs coefficients vérifient l'équation résultant des méthodes d'élimination décrites plus haut. Euler expose à cet effet une variante de la méthode n°3 et donne une autre justification du fait qu'il faut prendre $k = m + n - 1$. Supposons que ω soit une racine commune. Alors on peut écrire les deux équations sous la forme :

$$\begin{aligned} z^m + Pz^{m-1} + Qz^{m-2} + ... &= (z-\omega)(z^{m-1} + Az^{m-2} + Bz^{m-3} + ...) = 0 \\ z^n + pz^{n-1} + qz^{n-2} + ... &= (z-\omega)(z^{n-1} + az^{n-2} + bz^{n-3} + ...) = 0 \end{aligned}$$

D'où il suit que :

$$\begin{aligned} &(z^{n-1} + az^{n-2} + bz^{n-3} + ...)(z^m + Pz^{m-1} + Qz^{m-2} + ...) \\ &-(z^{m-1} + Az^{m-2} + Bz^{m-3} + ...)(z^n + pz^{n-1} + qz^{n-2} + ...) \quad = \quad 0 \end{aligned}$$

En égalant à zéro les coefficients de chaque puissance de z, on obtient ainsi $(m+n-1)$ équations linéaires. Le nombre des inconnues A,B,...,a,b,... est $(m+n-2)$. Euler écrit :

> Ayant donc une équation de plus, on parviendra enfin à une équation, qui ne contiendra plus aucune de ces lettres A,B,etc. et a,b,etc., et comme z ne s'y trouvera pas non plus, ce sera l'équation cherchée, à laquelle l'élimination conduit... [5]

3. *Cf.* Leonhard Euler, *Recherches sur les racines imaginaires des équations*, communiqué à l'Académie des sciences de Berlin le 10 novembre 1746, publié en 1749.
4. *Cf.* [161], p. 22–23 et 193–194. R. Rashed a montré la forte probabilité qu'Ibn al-Haytham (965–1040) ait connu « de quelque manière » ce théorème (*cf.* [257] « Ibn al-Haytham et le théorème de Wilson », p. 227–243).
5. *Cf.* [107], § 16.

Encore une fois, nous avons vu Leibniz développer la même idée dans un manuscrit tardif, un demi-siècle auparavant. Euler généralise ensuite cette méthode pour traiter le cas d'un nombre $N > 1$ de racines communes, et il est conduit à un système de N relations entre les coefficients des deux équations initiales.

Mis à part la méthode n°2 et la démonstration du théorème de Bézout qui marquent une avancée certaine, les idées d'Euler sur l'élimination algébrique que nous venons d'exposer ne sont donc pas nouvelles, et il convient de remarquer qu'Euler avait certainement bien connaissance des manuscrits de Leibniz – il le montre dans d'autres secteurs de ses recherches [6].

6.2 Intersection et projection

Dans l'*Introductio in analysin infinitorum*, c'est au tome second, dédié aux courbes, qu'Euler expose la théorie de l'élimination (chapitre XIX). Il la destine à la « construction des équations de degré élevé », sujet classique, déjà systématisé aux XI^{e}–XII^{e} siècles par 'Umar al-Khayyām et Sharaf al-Dīn al-Ṭūsī pour les équations de degré 3. Au XVII^{e} siècle, Descartes semblait y voir une méthode générale de résolution géométrique des équations à une inconnue. Le sujet est abordé par Euler au chapitre XX de son traité. Il s'agit de résoudre géométriquement une équation polynomiale à une inconnue, en considérant deux courbes algébriques bien choisies, telles que les abscisses de leurs points d'intersection fournissent des racines de l'équation proposée. C'est donc le problème inverse du problème de l'élimination tel qu'il est traité dans le chapitre XIX. Si $P(x, y) = 0$ et $Q(x, y) = 0$ sont les équations algébriques de deux courbes dans le plan de coordonnées x, y, alors les abscisses x des points d'intersection des deux courbes sont racines de l'équation $R(x) = 0$ résultant de l'élimination de y. Euler affirme que la réciproque aussi est exacte : à chaque racine de l'équation $R(x) = 0$ correspond au moins un point de l'intersection, dont elle est l'abscisse. Il remarque cependant que parfois, à une racine réelle de $R(x) = 0$ correspondent deux « points d'intersections imaginaires », ayant des ordonnées complexes conjuguées, et aucun point d'intersection réel. Pour justifier cette affirmation, il a recours a sa méthode d'élimination n°1, et il remarque qu'avant l'élimination finale donnant $R(x) = 0$, elle conduit à une équation linéaire en y. Cette équation linéaire en y fournit alors immédiatement une ordonnée pour chaque abscisse x racine de $R(x)$. Dans les cas particuliers où l'équation linéaire en y est en fait de degré 0 en y, il faut remonter à l'étape précédente de l'élimination, donnant une équation de degré 2 en y. Cette équation peut avoir deux racines réelles (on a alors deux points d'intersection réels) ou bien deux racines complexes conjuguées (deux points d'intersection imaginaires). Euler s'intéresse aussi aux cas où l'équation linéaire en y possède un facteur de degré > 0 en y, et un facteur de degré 0 en y et de degré > 0 en x. Les abscisses x qui sont racines du second facteur doivent alors être traitées sans avoir recours à l'équation linéaire en y.

Au problème des points d'intersection imaginaires s'ajoute celui des points à l'infini, et celui des points d'intersection multiples. Bien que ces phénomènes ne soient pas expliqués dans l'*Introductio* d'Euler, il ne pouvait les ignorer, puisqu'ils influent sur la manière de compter les points d'intersection de deux courbes, ou d'une courbe et d'une droite. Par exemple, le décompte d'un point d'intersection multiple lui fut utile afin de démontrer qu'une cubique plane a au plus un point double. Curieusement, Euler commit alors une erreur en généralisant trop vite aux degrés supérieurs :

> Une ligne du troisième ordre, à moins qu'elle ne soit composée de trois lignes droites, ne peut avoir plus d'un point double ; car supposons qu'elle en pût avoir deux, et imaginons par ces points une droite ; elle couperait la courbe en quatre points, ce qui est contraire à la nature des lignes du troisième ordre. La ligne du quatrième ordre aura seulement deux points doubles, et une ligne du cinquième n'en pourra avoir plus de trois, ainsi de suite. [7]

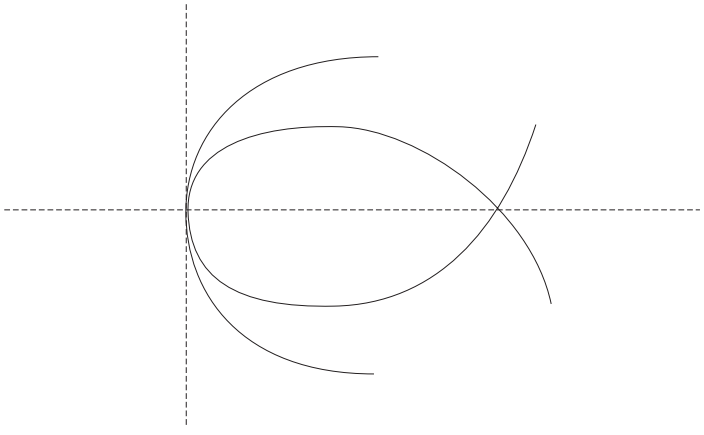
Un tel silence sur la notion de multiplicité d'intersection ne doit pas nous surprendre si l'on pense que cette notion est bien difficile à expliquer par l'algèbre (voir *infra* p. 211). L'objectif de l'*Introductio*

6. Costabel explique cela dans [91] (au sujet de la copie par Leibniz d'un manuscrit cartésien).
7. *Cf.* [104] t. 2 § 300.

in analysin infinitorum était justement [8] d'aborder l'analyse infinitésimale sous l'angle de l'algèbre – attitude parente de celle qui consisterait aujourd'hui à exposer une théorie des séries formelles comme prolégomènes à l'analyse. Expliquer la notion de multiplicité d'intersection dans ce cadre requiert des instruments qu'Euler ne maîtrisait pas suffisamment, tel le développement en série de Puiseux. A cet égard, la controverse sur l'existence des points de rebroussement de seconde espèce est instructive [9]. Un point de rebroussement est dit de seconde espèce lorsque les deux branches de la courbe sont situées du même côté de la tangente en ce point. Lors de la rédaction de l'*Introductio in analysin infinitorum* (en 1744), Euler refusait l'existence de telles singularités. Il étudiait en effet la forme des points singuliers exclusivement au moyen de courbes osculatrices d'équation $y^m = ax^n$, ce qui revient à ne garder que le premier terme du développement en série de Puiseux de y fonction de x. Il était cependant revenu assez vite de cette erreur, car il écrivit à Cramer le 15 décembre 1744 pour lui demander d'insérer une note dans l'ouvrage dont celui-ci devait superviser l'impression [10]. La note en question donne l'exemple d'un point de rebroussement de seconde espèce, exemple que l'on trouve expliqué de manière plus détaillée dans un discours prononcé semble-t-il en 1747 *Sur le point de rebroussement de la seconde espèce de M. Le Marquis de L'Hôpital.* Il s'agit d'une quartique dont deux branches forment un tel point de rebroussement en 0 et sont de la forme :

$$y = \sqrt{x} \pm \sqrt{x\sqrt{x}}$$

Pour montrer qu'il s'agit bien d'un point de rebroussement de seconde espèce, son argument repose sur le fait que tout développement de y dont l'un des coefficients est imaginaire doit être rejeté comme ne décrivant pas une branche de la courbe . En particulier, concevant l'expression ci-dessus comme étant l'expression algébrique d'une des racines de l'équation de la courbe, il semble d'abord possible de changer le signe du radical $\sqrt{x}$ pour obtenir l'expression d'une autre racine, et donc d'une autre branche de la courbe, mais cela est interdit par le fait que l'expression $\sqrt{x\sqrt{x}}$ deviendrait alors imaginaire (pour x positif). La quartique n'a donc que deux branches en zéro, entièrement situées dans le quadrant $(x \geq 0, y \geq 0)$. On comprendra mieux encore cet exemple au moyen d'un passage de la lettre de Euler à Cramer du 13 août 1746 [11]. Cramer avait, semble-t-il, douté de l'exemple donné par Euler. Il avait fourni à Euler le tracé mécanique d'une courbe contenant les branches $y = \sqrt{x} \pm \sqrt{x\sqrt{x}}$, de la forme :



Au point $(0, 0)$ cette figure présente une singularité plus compliquée qu'un point de rebroussement de seconde espèce, et qui ne ressemble plus beaucoup à un « point de rebroussement ». Mais Euler remarque que ce n'est pas la quartique $y^4 - 2xy^2 - 4x^2y - x^3 + x^2 = 0$ qu'a tracée Cramer (car celle-ci n'a pas de diamètre), mais plutôt la courbe $(y^4 - 2xy^2 - x^3 + x^2)^2 = (4x^2y)^2$: la quartique ne consiste

8. Comme l'a montré C. Houzel dans [154].
9. Sur cette controverse, *cf.* [108] série I vol. 27 p. XXX-XXXI.
10. C'est ce que laisse comprendre le résumé de cette lettre dans le catalogue de la correspondance d'Euler *cf.* [108] série 4A vol. 1 p. 92, lettre n°463. La lettre n'a pas encore été éditée. La note en question est au § 300 du tome 2 de l'*Introductio*.
11. *Cf.* lettre de Euler à Cramer du 13 août 1746 (British Library, Additional catalogue 23899, f. 17–18).

qu'en la moitié de la courbe tracée. Ajoutons que, dans le discours *Sur le point de rebroussement...*, Euler montre qu'une courbe contenant un tel point est de degré ≥ 4 par un argument qui fait à nouveau intervenir la multiplicité d'intersection. Une droite passant par le point de rebroussement de seconde espèce, écrit-il, coupe la courbe en 4 points au moins donc la courbe ne peut être de degré < 4 : Euler devait donc compter l'intersection d'une droite générique passant par le point de rebroussement avec multiplicité 2. Enfin, d'Alembert [12] donna lui aussi un exemple de point de rebroussement de seconde espèce, dans un ouvrage qui parut la même année que l'*Introductio* d'Euler (1748) :

$$y = x^2 \pm x^{\frac{5}{2}}$$

Quant aux points d'intersection à l'infini, Euler en parle aussi dans la lettre que nous venons de mentionner, au sujet d'un autre problème. Toutefois, il n'envisage pas l'existence de points qui soient à la fois imaginaires et à l'infini :

> Je vois aussi par votre lettre que j'ai avancé qu'il est impossible que deux lignes du troisième ordre ne se coupent qu'en huit points sans qu'il y soit la neuvième intersection. Or vous remarquez que huit intersections peuvent avoir lieu, lorsque deux branches deviennent parallèles entre elles. Sur quoi je suis parfaitement de votre sentiment : mais dans ce cas il y a en effet le neuvième intersection, quoiqu'elle soit éloignée à l'infini : car une intersection à l'infini est réelle, et point du tout imaginaire. Ces intersections de deux [lignes] du troisième ordre sont exprimées par une équation des neuf degrés et il y [aura] par conséquent autant d'intersections, que cette équation aura de [racines] réelles : or si elle a des racines imaginaires, leur nombre est toujours pair et partant le nombre des intersections doit être ou neuf ou sept ou cinq ou trois ou un : pourvu qu'on n'exclue les intersections à l'infini du nombre des intersections réelles : et après cette restriction, je ne doute pas, que vous ne soyez parfaitement de mon sentiment... [13]

Le dernier chapitre du *Traité abrégé des surfaces* (publié en appendice de l' *Introductio in analysin infinitorum*) est aussi dédié à la théorie de l'élimination. Euler y étudie l'intersection de deux surfaces données par des équations polynomiales $P(x, y, z) = 0$ et $Q(x, y, z) = 0$: il s'agit d'une courbe ou d'un ensemble de points, dont les projections sur le plan $\{z = 0\}$ vérifient l'équation $R(x, y) = 0$ obtenue en éliminant z entre les deux équations initiales. La fin du chapitre contient une application de cette théorie au calcul du plan tangent à une surface en un point : il suffit de trouver les tangentes à deux sections de cette surface par des plans distincts passant par ce point.

Pourtant, la projection de l'intersection sur le plan $\{z = 0\}$ n'est pas nécessairement égale à la courbe d'équation $R(x, y) = 0$. Euler en était conscient : il remarquait qu'à certains points réels de l'équation $R(x, y) = 0$ peuvent en effet correspondre des points d'intersection imaginaires de coordonnée z complexe. Mais aujourd'hui, on sait que même en se plaçant sur un corps algébriquement clos (c'est-à-dire en prenant en considération tous les points imaginaires), il n'y a pas toujours égalité. Ainsi la courbe $\mathcal{C}$ intersection du système $\begin{cases} xz - 1 = 0 \\ y + xz - 1 = 0 \end{cases}$ a pour projection les points de coordonnées $x \neq 0$ et $y = 0$, alors que l'équation résultant de l'élimination de z en soustrayant la première équation de la deuxième est simplement $y = 0$. Cependant la droite $\{y = 0\}$ est bien l'adhérence de la projection [14]. Dans le langage moderne de la géométrie algébrique, on expliquerait ce phénomène de la manière suivante. La projection

$$\begin{array}{ccc} \mathbb{A}^3 & \xrightarrow{p} & \mathbb{A}^2 \\ (x, y, z) & \longmapsto & (x, y) \end{array}$$

n'est pas une application fermée pour la topologie de Zariski. C'est-à-dire que l'image d'une courbe définie par un système d'équations algébriques (un « fermé de Zariski ») n'est pas elle-même, en

12. Sur ce mémoire de d'Alembert, voir [157] p. 29–33.
13. British Library, Additional catalogue 23899, f. 18.
14. En d'autres termes, il n'existe pas d'équation algébrique « plus fine », pour déterminer la projection, que celle résultant de l'élimination d'une des inconnues.

général, un ensemble défini par des équations algébriques. La situation ne s'améliore pas beaucoup en considérant cette projection comme étant induite par l'application :

$$\begin{array}{ccc} \mathbb{P}^3 - (0:0:1:0) & \longrightarrow & \mathbb{P}^2 \\ (X:Y:Z:T) & \longmapsto & (X:Y:T) \end{array}$$

Une courbe qui ne passe pas par le point $(0:0:1:0)$ a bien pour projeté une courbe de $\mathbb{P}^2$. Mais la courbe $\mathcal{C}$ ci-dessus passe justement par ce point. On peut au contraire considérer la projection comme étant induite par l'application :

$$\begin{array}{ccc} \mathbb{A}^2 \times \mathbb{P}^1 & \xrightarrow{\overline{p}} & \mathbb{A}^2 \\ (x, y, (Z:T)) & \longmapsto & (x, y) \end{array}$$

Notons $\overline{\mathcal{C}}$ l'adhérence de $\mathcal{C}$ dans $\mathbb{A}^2 \times \mathbb{P}^1$ pour la topologie de Zariski. L'image de $\overline{\mathcal{C}}$ est alors un fermé [15] et est donc définie par une équation dans $\mathbb{A}^2$ (sauf dans le cas trivial où $\overline{\mathcal{C}}$ est une droite se projetant en un unique point...). Cependant il n'est pas toujours aisé de déterminer l'adhérence de $\mathcal{C}$. Ainsi l'homogénéisation des équations du système ci-dessus par rapport à la variable z introduit une composante connexe superflue : le système d'équations

$$\begin{cases} xZ - T = 0 \\ yT + xZ - T = 0 \end{cases}$$

a pour solution $\mathcal{C} \cup \{x = 0,\ (Z:T) = (1:0)\}$ et non $\overline{\mathcal{C}} = \mathcal{C} \cup \{(0, 0, (1:0))\}$. Suivant la manière dont on élimine z dans le système non homogénéisé, l'équation finale va ou non contenir la composante superflue. On a vu que l'on pouvait obtenir $y = 0$ comme équation finale ; mais si l'on forme le résultant de ces deux équations en Z, T, qui dans ce cas élémentaire n'est autre que leur déterminant

$$\begin{vmatrix} x & -1 \\ x & y - 1 \end{vmatrix} = xy,$$

on retrouve l'image de la composante superflue $\{x = 0\}$.

Plus généralement, on a la situation suivante. Soit un système d'équations en les inconnues $x_1, x_2, ..., x_m, y_1, y_2, ..., y_{n+1}$ homogènes en $y_1, y_2, ..., y_{n+1}$, qui définit un fermé X de $\mathbb{A}^m \times \mathbb{P}^n$:

$$X \hookrightarrow \mathbb{A}^m \times \mathbb{P}^n \xrightarrow{p} \mathbb{A}^m$$

L'image de X par la projection p est un fermé de Zariski [16] et est donc à son tour défini par un système d'équations en $x_1, x_2, ..., x_m$. L'élimination algébrique consiste à déterminer ces équations à partir de celles de X. Le cadre naturel de la théorie de l'élimination est donc, plutôt qu'un espace projectif, le produit d'un espace quelconque (ici $\mathbb{A}^m$, mais on aurait pu aussi bien prendre $\mathbb{P}^m$) par un espace projectif $\mathbb{P}^n$ dont les coordonnées sont les variables à « éliminer ». C'est par exemple le point de départ de l'article fondamental de J.-P. Jouanolou (1991), qui insiste sur le caractère « élémentaire » d'un énoncé analogue à celui-ci, à condition d'avoir recours au langage moderne de la géométrie algébrique :

> L'énoncé précédent, *élémentaire* dans le cadre des schémas, est connu sous le nom de *théorème principal de l'élimination*. L'étape suivante consiste à expliciter l'idéal $\mathfrak{A}$ [ici, l'idéal de $p(X)$] et à en étudier les propriétés. [17]

15. En effet l'application $\overline{p}$ est fermée car elle s'obtient à partir du morphisme propre $\mathbb{P}^1 \longrightarrow \{0\}$ au moyen du diagramme cartésien suivant :

$$\begin{array}{ccc} \mathbb{A}^2 \times \mathbb{P}^1 & \xrightarrow{\overline{p}} & \mathbb{A}^2 \\ \downarrow & & \downarrow \\ \mathbb{P}^1 & \longrightarrow & \{0\} \end{array}$$

Le fait que $\mathbb{P}^1 \longrightarrow \{0\}$ est propre est démontré en [140] th. II.4.9 p. 103.

16. Une démonstration *ab initio* de ce fait, sans présupposer les notions de la géométrie algébrique, figure aussi dans le manuel d'algèbre [191] chap. IX § 3.

17. *Cf.* [168] § 1.

Ceci explique aussi que certains géomètres du XIX^{e} siècle travaillaient plus volontiers dans $\mathbb{P}^1 \times \mathbb{P}^1$ que dans $\mathbb{P}^2$ pour la théorie des courbes algébriques planes – nous y reviendrons en étudiant Weierstrass.

Revenons pour le moment à Euler, qui nous offre quelques exemples intéressants d'intersections de surfaces. Soient une sphère et un plan [18], d'équations :

$$\begin{cases} z^2 + y^2 + x^2 = a^2 \\ \alpha z + \beta y + \gamma x = f \end{cases}$$

Lorsque $f = a\sqrt{\alpha^2 + \beta^2 + \gamma^2}$, Euler obtient

$$y = \frac{\beta f - \beta\gamma x \pm \alpha\sqrt{-1}(\frac{f}{a}x + a\gamma)}{\alpha^2 + \beta^2}$$

en éliminant z. On dirait aujourd'hui que l'intersection est composée de deux lignes complexes, se coupant en un point réel. Euler ne s'intéresse qu'au point réel : il remarque que la seule solution réelle de cette équation suppose que $\frac{f}{a}x + a\gamma = 0$, et trouve alors

$$x = \frac{a\gamma}{\sqrt{\alpha^2 + \beta^2 + \gamma^2}}, y = \frac{a\beta}{\sqrt{\alpha^2 + \beta^2 + \gamma^2}}, z = \frac{a\alpha}{\sqrt{\alpha^2 + \beta^2 + \gamma^2}},$$

coordonnées du point de tangence entre le plan et la sphère. Il étudie ensuite l'intersection entre un cône scalène et une sphère [19], et trouve dans certains cas une « ligne de contact » circulaire (lorsque le cône est droit) correspondant à un carré dans l'équation de la projection, dans d'autres cas deux « points de contact » réels. Il formule une « règle générale » pour repérer les points et lignes de contact :

> On déduit donc de ce qui précède une règle générale pour reconnaître si une surface quelconque est touchée par un plan, ou par une autre surface. Car, après avoir éliminé des deux équations, une des variables, il faudra voir si l'équation qui en résulte peut être décomposée en facteurs simples, ou non. Si elle renferme deux facteurs simples imaginaires, il y aura un point de contact qu'on trouvera, en faisant l'un et l'autre facteur égaux à 0 ; et si elle contient deux facteurs simples, et tous deux égaux entre eux, les surfaces se toucheront l'une l'autre suivant une ligne droite... Mais, si cette équation a deux facteurs non simples, mais égaux, ou si elle est divisible par un carré, alors sa racine étant supposée égale à 0 fera connaître la projection de cette ligne qui résulte du contact. On voit aussi par-là que, si cette équation renfermait 4 facteurs imaginaires, les surfaces se toucheraient réciproquement en 2 points. [20]

Euler écrit encore, au sujet des lignes de contact :

> Ces sortes de contacts se connaîtraient par l'équation même, si on en trouve une pour la projection, qui contienne deux racines réelles, parce que ce contact n'est rien autre chose que le concours de deux intersections. [21]

Le fait que chaque facteur de l'équation de la projection correspond à une courbe appartenant à l'intersection, et que les « points de contact » ne sont que les sections réelles de courbes d'intersection imaginaires laisse pressentir un théorème moderne sur la dimension de l'intersection de deux variétés affines : si k est un corps algébriquement clos, et que Y et Z sont des variétés affines de dimensions r et s dans k^n, alors chaque composante irréductible de $Y \cap Z$ est de dimension $\geq r + s - n$ (*cf.* [140], chap. I, § 7.1). Ainsi, sur le corps des nombres complexes, dans l'espace ($n = 3$), deux surfaces ($r = s = 2$) ont une intersection de dimension ≥ 1 (c'est-à-dire une courbe, ou une surface si les deux surfaces proposées sont confondues). Le phénomène des « points de contact » isolés de deux surfaces n'existe donc pas sur le corps des complexes.

18. *Cf.* [104], §140–142.
19. *Cf.* [104], §143–146.
20. *Cf.* [104], §142.
21. *Cf.* [104], §139.

Chapitre 7

Paradoxe de Cramer et théorème de Bézout

Gabriel Cramer (1704–1752), professeur de mathématiques à Genève, fit le projet d'écrire un ouvrage sur les courbes algébriques dès 1735, et ce projet aboutit à la publication de son *Introduction à l'analyse des lignes courbes algébriques* en 1750. Il connaissait bien Euler, avec qui il s'était lié d'amitié (ainsi que Daniel Bernoulli) alors qu'ils étaient réunis dans l'entourage de Jean Bernoulli à Bâle en 1727 (plus tard, Jean Bernoulli confierait à Cramer la tâche de publier ses œuvres et celles, posthumes, de son frère Jacques (1654–1705)).

Dans l'introduction de son traité, Cramer mentionne surtout sa dette envers l'œuvre de Newton sur les courbes algébriques, ainsi qu'envers J. P. de Gua de Malves. Il remarque aussi qu'il aurait profité, dans sa rédaction, de l'*Introductio in analysin infinitorum* d'Euler, s'il avait connu cet ouvrage plus tôt, mais que son approche des courbes algébriques est bien différente de celle d'Euler. L'objet du traité est de donner un principe de classification des courbes algébriques, suivant leur degré (l'« ordre »), leurs branches infinies ou asymptotes (qui permettent de les répartir en « classes »), et leurs points singuliers (qui permettent de distinguer des « genres » et des « espèces » au sein de chaque « classe »). Nous parlerons ici seulement des trois premiers chapitres de ce traité. Cramer y définit la notion de courbe algébrique plane et étudie l'effet des transformations linéaires affines (changement de système d'axes de coordonnées) sur l'équation d'une telle courbe. Il démontre que le degré d'une équation est invariant par ces transformations affines :

> Je veux dire, qu'encore qu'on puisse exprimer la nature d'une ligne algébrique par une infinité d'équations différentes, selon le choix qu'on fait de l'origine, et la position qu'on donne aux axes ; cependant toutes ces équations sont d'un même ordre [= degré], auquel par conséquent la ligne proposée doit se rapporter. [1]

Peu après, Cramer calcule le nombre de termes monômes dans un polynôme général en x, y de degré n donné [2] :

$$1 + 2 + ... + (n + 1) = \binom{n+2}{2}$$

Au paragraphe § 38, Cramer se propose de déterminer l'équation $A+By+Cx+Dy^2+Eyx+x^2 = 0$ d'une courbe de degré 2 passant par 5 points donnés. Il lui faut pour cela résoudre le système linéaire

1. *Cf.* [81], § 34.
2. *Cf.* [81] § 37. Remarquons que Cramer connaissait la méthode dite « du polygone de Newton » pour le développement en série des racines y d'une équation algébrique en x et y, suivant les puissances de x (nous exposerons cette méthode au chapitre 13). Cela apparaît dès le début de son traité puisqu'il compte le nombre de termes monômes du polynôme général de degré n au moyen du « triangle de de Gua », variante de la figure qu'utilisait Newton dans cette méthode, figure sur laquelle sont représentés tous les monômes en x, y (cette figure est aussi parfois appelée « parallélogramme de Newton »).

suivant :

$$\begin{cases} A + Ba + C\alpha + Da^2 + Ea\alpha + \alpha^2 = 0 \\ A + Bb + C\beta + Db^2 + Eb\beta + \beta^2 = 0 \\ A + Bc + C\gamma + Dc^2 + Ec\gamma + \gamma^2 = 0 \\ A + Bd + C\delta + Dd^2 + Ed\delta + \delta^2 = 0 \\ A + Be + C\epsilon + De^2 + Ee\epsilon + \epsilon^2 = 0 \end{cases}$$

En choisissant un système de coordonnées adéquat afin que certaines des quantités a,α,b,β,... s'annulent, il résout ce système facilement, par substitutions, à la manière dont Newton l'avait d'ailleurs résolu au siècle précédent [3]. Il trouve une unique solution. Puis il renvoie à son appendice n°1, dans lequel il énonce des formules générales pour la solution d'un tel système linéaire, connues aujourd'hui sous le nom de « formules de Cramer » [4]. Pour un système à trois inconnues z,y,x, Cramer écrit par exemple :

$$z = \frac{A^1Y^2X^3 - A^1Y^3X^2 - A^2Y^1X^3 + A^2Y^3X^1 + A^3Y^1X^2 - A^3Y^2X^1}{Z^1Y^2X^3 - Z^1Y^3X^2 - Z^2Y^1X^3 + Z^2Y^3X^1 + Z^3Y^1X^2 - Z^3Y^2X^1}$$

que nous écririons aujourd'hui :

$$z = \frac{\begin{vmatrix} A^1 & Y^1 & X^1 \\ A^2 & Y^2 & X^2 \\ A^3 & Y^3 & X^3 \end{vmatrix}}{\begin{vmatrix} Z^1 & Y^1 & X^1 \\ Z^2 & Y^2 & X^2 \\ Z^3 & Y^3 & X^3 \end{vmatrix}}$$

Les expressions analogues pour x et y ont même dénominateur et Cramer indique que lorsque le dénominateur est nul, soit le problème est « impossible » (si certains des numérateurs sont non nuls), soit le problème est « indéterminé » (si tous les numérateurs sont nuls). Il justifie cette assertion par la seule remarque que dans ce cas, ses formules donnent $\frac{0}{0}$. Quant au problème de la détermination d'une courbe passant par un nombre donné de points, dans le cas d'indétermination $\frac{0}{0}$, Cramer conclut qu'« alors on peut faire passer par les points donnés une infinité de lignes du même ordre ».

Quant aux dites « formules de Cramer », Thomas Muir, dans son ouvrage *The theory of determinants in the historical order of its developpement* (1890) les attribuait en effet à Cramer, ainsi que la règle donnant le signe de chaque terme du déterminant. C. B. Boyer a cependant remarqué que ces formules figuraient déjà dans le *Treatise of algebra* posthume de Maclaurin [5]. Plus récemment, E. Knobloch [175] a remarqué que Leibniz connaissait déjà de semblables formules, il a étudié les diverses « règles des signes » présentes dans les manuscrits leibniziens, parmi lesquelles plusieurs sont fausses, et il a comparé les règles des signes exactes de Leibniz à celles utilisées par Maclaurin et Cramer.

Revenons à Cramer. Au paragraphe § 42, il remarque que le problème de trouver l'intersection de deux courbes se ramène au problème de l'élimination algébrique d'une inconnue y entre les équations des deux courbes. Il s'intéresse au nombre de points d'intersection de deux courbes d'ordres donnés. Mais il indique qu'à chaque racine de l'équation résultante en x peut correspondre parfois plusieurs points d'intersection d'ordonnées y distinctes, et parfois aucun, dans le cas où l'on aurait un « point d'intersection imaginaire ». Pour remédier à la première difficulté, Cramer propose d'utiliser les transformations affines introduites précédemment :

> Si l'on ne cherche pas tant à connaître en combien de points se rencontrent deux lignes dont les équations sont données, qu'à savoir (ce qui a aussi son usage) en combien de points une ligne d'un ordre donné peut rencontrer une autre ligne d'un ordre aussi donné ; on doit concevoir, ce qui est toujours évidemment possible, un axe des abscisses dont la

3. *Cf.* problèmes 55–57 des *Lucasian lectures on algebra* [237], p. 308–320 ; le n°57 correspond au n°61 de l'édition de 1722 de l'*Arithmetica universalis*, cité par Cramer. Le problème 55 est la construction d'une conique passant par 5 points ; le 56 celle d'une conique passant par 4 points, et tangente à une droite donnée en l'un de ces points ; le 57 celle d'une conique passant par 3 points, et tangente à des droites données en deux de ces points.

4. *Cf.* [81], p. 656–657.

5. *Cf.* Carl B. Boyer, « Colin Maclaurin and Cramer's Rule », *Scripta mathematica*, n°27 (1964), p. 377–379).

> position soit telle qu'à chaque point de rencontre il réponde une ordonnée et une abscisse différentes.

Cramer explique ensuite que le degré de l'équation résultant de l'élimination de y donnera alors une majoration du nombre de points d'intersection (il peut encore y avoir des « points imaginaires » parmi ces points). Puis il énonce le théorème suivant sur le degré de cette équation :

> (...) si l'on a deux variables, et deux équations indéterminées qui expriment le rapport de ces variables avec des constantes, desquelles l'une soit de l'ordre m et l'autre de l'ordre n, lorsque au moyen de ces équations on chasse une de ces variables, celle qui reste n'a, dans l'équation finale qui la détermine, que mn dimensions au plus. [6]

Cramer renvoie aussi à l'article de Maclaurin, dans lequel ce dernier énonçait cette proposition et prétendait l'avoir démontrée dès 1721 [7]. Il ajoute finalement :

> Ce principe, purement algébrique, devrait être démontré dans l'Algèbre. Comme je n'en connais aucune qui en donne la démonstration, j'ai cru devoir l'insérer dans l'Appendice n°2.

7.1 Le théorème de Bézout démontré par Cramer et Poisson

Soient deux équations en x, y, notées par Cramer :

$$\begin{array}{ll}(A) & x^n - [1]x^{n-1} + [1^2]x^{n-2} - ...[1^n] = 0\\ (B) & (0)x^0 + (1)x^1 + (2)x^2 + ... + (m)x^m = 0\end{array}$$

où $[1]$,$[1^2]$,...,(0),(1),... sont des polynômes en y avec $d^o[1^k] = k$ et $d^o(k) = m - k$. Alors l'équation résultant de l'élimination est :

$$\begin{array}{lll}(C) & ((0)a^0 + (1)a^1 + ... + (m)a^m) & \\ & \times((0)b^0 + (1)b^1 + ... + (m)b^m) & \\ & \times \dots & = \quad 0\end{array}$$

où a,b,... sont les racines de (A). Cette méthode est essentiellement équivalente à la méthode n°2 d'Euler, mais la détermination du degré de (C) est menée de manière plus rigoureuse par Cramer. Un lecteur moderne pourrait y reconnaître à bon droit la première vraie démonstration du théorème de Bézout. Cramer développe (C) en une somme de produits, chacun composé d'un « facteur-premier », produit de puissances des coefficients (0),(1),..., et d'un « facteur-second », fonction symétrique des racines a,b,.... Soit par exemple, pour $n = 4$ et $m = 3$, le facteur-premier $(1) \times (2)^2 \times (3)$, que Cramer note (1223). Le facteur-second qui lui est associé est la fonction symétrique $\sum ab^2c^2d^3$, que Cramer note [1223]. Il ordonne ensuite tous les facteurs-seconds dans un tableau et montre par récurrence que tous s'expriment rationnellement en fonction de la première ligne du tableau, constituée des facteurs $[0000] = 1$, $[0001] = [1]$, $[0011] = [1^2]$,... (ce sont les fonctions symétriques élémentaires, elles-mêmes égales aux coefficients de l'équation (A)). Le raisonnement par récurrence garantit de plus que $d^o[1223] = 1+2+2+3$. Comme l'on a aussi $d^o(1223) = d^o(1)+2d^o(2)+d^o(3) = 4.3-(1+2+2+3)$, on en déduit $d^o(1223)[1223] = 4.3$. Ceci vaut pour chaque terme de (C). On a donc montré, en général, $d^o(C) = mn$. Cramer a fait le calcul effectif de l'équation (C) dans le cas $m = n = 4$ (*cf.* [81], p. 672).

Poisson, en 1802, reprend les idées maîtresses de cette démonstration, pour la généraliser à un nombre quelconque d'équations (nous avons reproduit le mémoire de Poisson en annexe B). Le cas de deux équations lui sert d'étape initiale, pour un raisonnement par récurrence. Soit donc un système de n équations (non homogènes) à n inconnues, de degrés respectifs k_1,...,k_n :

$$\left\{\begin{array}{l}P_1(x, y, ..., u) = 0\\ \vdots\\ P_n(x, y, ..., u) = 0\end{array}\right.$$

6. *Cf.* [81], § 46.
7. *Cf. supra* p. 100.

En éliminant de toutes les manières possibles $(n-2)$ parmi les $(n-1)$ premières inconnues, au sein des $(n-1)$ premières équations, on obtient $(n-1)$ nouvelles équations :

$$\begin{cases} Q_1(x,u)=0 \\ Q_2(y,u)=0 \\ \vdots \end{cases}$$

On suppose le théorème de Bézout déjà démontré par récurrence pour les systèmes de $(n-1)$ équations : le degré en x,y,... des polynômes Q_1,Q_2,... est donc $p=k_1k_2...k_{n-1}$. On a donc p racines $x_1,...,x_p$ de $Q_1(x,u)$, qui sont, selon l'expression de Poisson, des « fonctions irrationnelles de u », et à chacune d'elles correspond une solution du système initial. On a de même p racines y_1,...,y_p de $Q_2(y,u)$, et ainsi de suite. On peut supposer que ces racines sont rangées dans le bon ordre, c'est-à-dire que pour tout i, $(x_i,y_i,...)$ est une solution des $(n-1)$ premières équations du système initial. Alors l'équation finale en u est :

$$\prod_{i=1}^{p} P_n(x_i,y_i,...,u)=0$$

Nous désignerons désormais le membre de gauche de cette équation par la locution « formule de Poisson ». Pour déterminer le degré de cette équation, la démarche de Poisson consiste à généraliser la notion de fonction symétrique des racines au cas de plusieurs variables. Une fonction de plusieurs ensembles de racines x_i,y_i,... est symétrique si elle est invariante par permutation des indices i. Il calcule ces fonctions symétriques au moyen d'un procédé qui était déjà, nous l'avons vu ci-dessus, cher à Cramer : un changement affine de coordonnées $t=x+gy+hz+....$ Mais Poisson choisit ici pour $g,h,...$ des « quantités arbitraires et indépendantes » des autres variables[8]. Il est alors possible, au moyen des coefficients des monômes en $g,h,...$ et t de l'équation en t et u résultant de l'élimination des autres inconnues dans les $(n-1)$ premières équations du système initial, de calculer certaines fonctions symétriques (dans le cas $n=2$, il s'agit des sommes de puissances des racines), à partir desquelles toutes les autres s'expriment aisément, ainsi que leurs degrés en u. Poisson vérifie ainsi que le degré de l'équation finale est égal à $pk_n=k_1k_2...k_n$.

7.2 Le paradoxe de Cramer

Le problème du paragraphe § 38 du traité de Cramer, mentionné plus haut, conduit au célèbre « paradoxe de Cramer ». Nous l'avons déjà rencontré dans la *Geometria organica* de Maclaurin (*cf. supra* p. 96). Rappelons-en la teneur. D'une part l'équation générale d'une courbe de degré n possède $N=\binom{n+2}{2}-1$ coefficients arbitraires. En demandant qu'une telle courbe passe par N points distincts donnés, on obtient un système de N équations entre ces N coefficients. D'autre part deux courbes de degré n peuvent se couper en n^2 points distincts[9]. Dans le cas $n=3$, Cramer écrivit à Euler le 30 septembre 1744 pour lui faire part du « paradoxe » : *deux* cubiques se coupent en $3^2=9$ points, et pourtant il semble suffire de $\binom{5}{2}-1=9$ points pour déterminer une cubique de manière *unique*. Pourtant, il connaissait lui-même la cause de ce phénomène : les « formules de Cramer » exprimant la solution d'un système d'équations linéaires par des déterminants peuvent prendre la forme « indéterminée » $\frac{0}{0}$. Aussi la solution d'un système de N équations linéaires à N inconnues n'est-elle pas toujours unique. Mais cela n'enlève rien au caractère paradoxal du résultat ; en effet « ce qui est fâcheux, écrit-il, c'est qu'on ne le puisse voir intuitivement que par un calcul immense »[10]. De même, Cramer conclut ainsi le chapitre III de son traité, après avoir exposé le « paradoxe » :

8. Ce procédé est comparable à celui utilisé par Laplace puis Gauss pour démontrer le théorème fondamental de l'algèbre, *cf. infra* note 26 p. 137.

9. Cramer envisageait pourtant l'existence de points multiples, et, comme on l'a vu, de « points imaginaires ».

10. *Cf.* lettre de Cramer à Euler, 11 novembre 1744, citée dans [108], I–26, p. XII. Seul un résumé de la correspondance Euler-Cramer sur le paradoxe de Cramer est publié à ce jour : [108], série IVA, tome 1, lettres 461 (30 sept. 1744) à 467 (13 août 1746). Il est annoncé que le prochain volume à paraître de la correspondance d'Euler contiendra la correspondance Euler-Cramer.

> Cette contradiction se lève par la remarque qui termine le § 38. C'est qu'encore qu'on ait autant d'équations qu'il en faut, généralement parlant, pour déterminer tous les coefficients, (...) il peut pourtant arriver que ces coefficients restent indéterminés. Alors l'équation prise reste indéterminée et représente une infinité de courbes de même ordre (...). Si deux lignes d'ordre 4 se coupent en 16 points, parmi lesquels on en choisit 14 quelconques, il y a une infinité de lignes du 4^{e} ordre qui peuvent passer par ces 14 points. Ce qui est un véritable paradoxe. [11]

La dernière phrase de cette citation montre l'obstacle, ou la surprise, sur le plan intuitif, qu'avait dû poser pour lui cette découverte. La théorie de l'élimination dans les systèmes d'équations linéaires n'avait donc pas encore atteint la perfection désirable. La compréhension du même phénomène par Euler semble avoir été plus fine. Sollicité par Cramer, Euler explique le « paradoxe » par l'existence de relations de dépendance linéaire entre les N équations. Il a même l'idée de compter le nombre de ces relations. Il écrit en 1747 un mémoire entier *Sur une contradiction apparente dans la doctrine des lignes courbes* (qui parut en 1748 dans le même volume des mémoires de l'Académie de Berlin que son premier mémoire sur la théorie de l'élimination). Ce mémoire se conclut ainsi :

> Quand deux lignes du 4^{e} ordre s'entrecoupent en 16 points, puisque 14 points, lorsqu'ils conduisent à des équations toutes différentes entre elles, sont suffisants pour déterminer une ligne de cet ordre, ces 16 points seront toujours tels que trois ou plusieurs des équations qui en résultent sont déjà comprises dans les autres. De sorte que ces 16 points ne déterminent plus que s'il n'y en avait que 13 ou 12 ou encore moins et partant pour déterminer la courbe entièrement on pourra encore à ces 16 points ajouter un ou deux points. La même chose arrivera, si deux lignes du 5^{e} ordre se coupent l'une l'autre en 25 points, qui n'étant pas suffisants à déterminer la courbe, ils ne vaudront plus que 19, ou même 18, de sorte que 6 ou 7 points sont superflus, et partant ces 25 points seront toujours tellement disposés que dès que la courbe passera par 19 de ces points, elle passera d'elle même aussi par les autres, ou il sera impossible qu'elle passe par 19 points, sans passer en même temps par tous les 25. [12]

On dirait aujourd'hui qu'un système linéaire de $N + k$ équations à N inconnues (pour $k \geq 0$), qui a au moins deux solutions distinctes, est de rang $r \leq N - 1$, et dans ce cas, les équations vérifient certaines relations entre elles (elles ne sont pas « linéairement indépendantes »). Ces relations de dépendance linéaire forment à leur tour un espace vectoriel, de dimension $N + k - r$. Quant au nombre d'équations qu'il faudrait ajouter pour rendre le système déterminé (c'est-à-dire ayant une solution unique), il est égal à $N - r$. Pour deux courbes de degré $n \geq 3$, on a $N + k = n^2$ et $N = \frac{n(n+3)}{2}$. Alors $r \leq \frac{n(n+3)}{2} - 1$, et $N + k - r \geq n^2 - \left[\frac{n(n+3)}{2} - 1\right] = \frac{(n-1)(n-2)}{2}$. Pour $n = 5$, on a donc $N + k - r \geq 6$, « de sorte que 6 ou 7 points sont superflus », comme le dit Euler. Bien sûr les notions de rang et d'indépendance linéaire ne sont pas encore définies à l'époque d'Euler. Ces compréhensions encore frêles des phénomènes de l'algèbre linéaire ne sont cependant pas isolées : Bézout, dans son traité de 1779, justement pour démontrer le théorème qui porte son nom, fait constamment usage d'une formule sur les systèmes d'équations linéaires homogènes linéairement indépendantes, que l'on pourrait écrire aujourd'hui :

$$\text{nombre d'équations} = \text{nombre d'inconnues} - \text{« nombre de coefficients inutiles »}$$

où le dernier terme désigne en fait le nombre de paramètres de la solution générale du système d'équations, c'est-à-dire la dimension de l'espace vectoriel des solutions [13].

11. *Cf.* [81], § 48.
12. *Cf.* [105], § 23.
13. *Cf. infra* p. 154. Frobenius n'introduira le mot *rang* qu'à la fin des années 1870, pour désigner le plus haut rang des mineurs non nuls d'une matrice (*der höchste Grad nicht verschwindender Unterdeterminanten*), mais l'invariance d'un tel nombre (pour la multiplication par des matrices inversibles) est démontrée par Sylvester en 1850.

Chapitre 8

Les premières recherches de Bézout

Le mémoire *Sur le degré des équations résultantes de l'évanouissement des inconnues* (1764) de Bézout est probablement inspiré par ses recherches sur la résolution algébrique des équations. Il avait publié un mémoire sur ce sujet en 1762, et il en publie un autre en 1765. Dans le mémoire de 1765 on rencontre par exemple, pour résoudre l'équation générale de degré 3, un système de trois équations à trois inconnues, de degrés respectifs 2, 3 et 4. L'élimination de deux de ces inconnues conduit à une « réduite » [1] de degré $2.3.4 = 24$, qui s'abaisse en fait au degré 6 par un changement de variable [2]. Dans le mémoire de 1764, Bézout introduit cependant les « méthodes d'élimination » comme un domaine à part entière, rendu nécessaire par le fait suivant :

> Si on a des méthodes pour résoudre, par approximation, les problèmes déterminés lorsqu'on n'a qu'une seule équation, on n'en a pas de même pour les résoudre par approximation lorsque les relations des inconnues, qui en font l'objet, restent, pour ainsi dire, dispersées dans plusieurs équations (...) [3]

Une autre application possible des méthodes d'élimination est mentionnée par Bézout : le calcul du plus grand commun diviseur d'un ensemble de polynômes. Dans ce mémoire, il applique la méthode d'élimination n°3 d'Euler (*cf. supra* p. 107) à la démonstration du théorème de Bézout pour deux équations à deux inconnues. Rappelons que, pour deux équations $\begin{cases} P(x) = 0 \\ Q(x) = 0 \end{cases}$, cette méthode consiste à postuler que l'équation finale est une équation de la forme $PR + QS = 0$ de degré minimal (en particulier, il n'est pas nécessaire de faire l'hypothèse que l'une des deux équations possède un nombre de racines x égal à son degré – contrairement à ce qui se passe dans la démonstration par Cramer du théorème de Bézout). On cherche alors à déterminer les polynômes R et S au moyen de la méthode des coefficients indéterminés – il faut au préalable déterminer leurs degrés – et l'on est ainsi ramené à un problème linéaire. En imposant la recherche d'une équation *de degré minimal*, cette approche soulève le problème des « facteurs superflus » que tout calcul d'élimination risque de faire apparaître dans l'équation finale. Le problème des facteurs superflus va arrêter Bézout pendant longtemps dans sa recherche d'une méthode générale d'élimination des inconnues pour un nombre quelconque d'équations. Cherchant à appliquer la méthode n°3 d'Euler à un nombre quelconque d'équations, il explique en 1764 que la méthode consistant à éliminer en comparant les n équations d'un système ($n > 2$)

$$\begin{cases} P_1 = 0 \\ \vdots \\ P_n = 0 \end{cases}$$

1. On emploie ici un terme introduit plus tard par Lagrange.
2. *Cf.* [27], p. 539.
3. *Cf.* [26], p. 288.

successivement deux à deux conduit à des facteurs superflus particulièrement difficiles à repérer, et il conclut : « Mais quel est le fil qui conduirait dans ce labyrinthe ? » [4]. Euler aussi, à la recherche d'une démonstration du théorème de Bézout, semble avoir été frappé de confusion : « le calcul conduit à des formules si compliquées que l'on en perd patience » [5], écrit-il à Cramer en 1750. Les propos de Maclaurin sur la démonstration du théorème ont un goût semblable (*cf. supra* n. 37 p. 98). Le problème des facteurs superflus concentre toute la difficulté technique et conceptuelle des recherches du XVIII$^{\text{ème}}$ siècle en théorie de l'élimination. La méthode adoptée par Bézout en 1764 consiste à chercher un nouveau système d'équations résultant de l'élimination d'*une* inconnue, et dont chacune est de la forme $P_1R_1 + ... + P_nR_n = 0$, de degré minimal. Il réitère ce procédé pour éliminer toutes les inconnues, l'une après l'autre [6]. Cette approche est encore imparfaite et ne permet pas d'éviter tout facteur superflu. C'est pourtant en y restant fidèle, et en en trouvant l'expression ultime, qu'il parviendra en 1779 à une méthode correcte.

Mais le mémoire de 1764 garde un certain intérêt, et nous allons en décrire le contenu [7]. Bézout commence par rappeler les travaux de Newton (sans en citer aucun en particulier), Cramer et Euler [8]. Il donne ensuite deux lemmes sur les systèmes d'équations linéaires. Il considère des systèmes d'équations linéaires à n équations et n inconnues, de la forme :

$$\begin{aligned} &ax + by + cz + ... = 0 \\ &a'x + b'y + c'z + ... = 0 \\ &a''x + b''y + c''z + ... = 0 \\ &... \end{aligned}$$

Le lemme I (p. 292–295) explique (sans démonstration) comment former l'« équation de condition » (c'est-à-dire le déterminant) correspondant à un tel système linéaire. Le lemme II (p. 295–298) est une propriété de l'« équation de condition ». Il conduit au corollaire combinatoire suivant, important dans la suite du mémoire :

> ...si ces coefficients [9] étaient des fonctions d'une ou de plusieurs quantités dont la plus haute dimension formerait, dans les coefficients d'une même inconnue, une progression arithmétique, et si en même temps ces progressions avaient toutes une différence commune, il s'ensuit, dis-je, que dans l'équation de condition, la plus haute dimension des quantités, dont les coefficients sont des fonctions, serait $S + n.\frac{n-1}{2}k$, S marquant la somme des plus hautes dimensions dans les coefficients de la première équation, (...) et k la différence des plus hautes dimensions d'une équation à l'autre. [10]

Les pages 298 à 317 sont ensuite consacrées à l'exposition d'une méthode générale pour calculer le degré de l'équation finale résultant de l'élimination de $N-1$ inconnues parmi N, dans N équations. C'est là l'objectif principal de Bézout, car une fois le degré de l'équation finale connu, on a, dit-il, « la pierre de touche à l'aide de laquelle on peut juger du mérite des méthodes qu'on se propose d'employer pour éliminer » [11].

4. *Cf.* [26], p. 290.
5. Extrait de la correspondance non encore éditée, mais cité par J. J. Gray [132], p. 190.
6. « Éliminer les inconnues dans les équations de différents degrés, l'une après l'autre », *cf.* [26] p. 291.
7. *Cf.* [26]. Ce mémoire contient : p. 288 introduction ; p. 292 deux lemmes d'algèbre linéaire ; p. 298 application à la recherche du degré de l'équation résultant de l'élimination, dans le cas des équations à deux inconnues ; p. 301 cas des équations à trois inconnues ; p. 309 cas des équations à quatre inconnues ; p. 313 cas des équations à cinq inconnues ; p. 316 cas des équations à N inconnues ; p. 317 procédé pour l'élimination et réflexions qui tendent à l'abréger.
8. *Cf.* [26], p. 288–291.
9. Il s'agit des coefficients des équations du système linéaire.
10. *Cf.* [26], p. 297.
11. *Cf.* [26], p. 288.

8.1 Cas de deux équations ; la matrice de Bézout

Soient donc deux équations :

$$\begin{cases} Ax^m + Bx^{m-1} + Cx^{m-2} + ... + V = 0 \\ A'x^{m'} + B'x^{m'-1} + C'x^{m'-2} + ... + V' = 0 \end{cases}$$

où A,B,C,...,A',B',C',... sont des polynômes en y, tels que $d^oA = d^oB - 1 = d^oC - 2 = ... = p$ et $d^oA' = d^oB' - 1 = d^oC' - 2 = ... = p'$.
Bézout indique que la question d'éliminer x « se réduit à trouver une fonction de x, par laquelle multipliant la première équation, et une autre fonction de x, par laquelle multipliant la seconde, la somme des deux produits soit telle que chaque puissance de x disparaisse ». Bézout multiplie la première équation par un polynôme en x à coefficients indéterminés de degré $(m'-1)$, et la deuxième par un polynôme de degré $(m-1)$, et fait la somme des deux produits. Il obtient un polynôme en x de degré $(m+m'-1)$, dont les coefficients sont des combinaisons linéaires de A,B,C,...,A',B',C',..., à coefficients indéterminés. En égalant à 0 les coefficients de toutes les puissances de x, il obtient un système de $(m+m')$ équations linéaires, dont les coefficients sont des polynômes en y :

$$\begin{aligned} AM + A'M' &= 0 \\ AN + A'N' + BM + B'M' &= 0 \\ AP + A'P' + BN + B'N' + CM + C'M' &= 0 \\ AQ + A'Q' + BP + B'P' + CN + C'N' + DM + D'M' &= 0 \\ &... \end{aligned}$$

L'« équation de condition » de ce système linéaire est l'équation finale recherchée. Jusque là, la méthode de Bézout est identique à la méthode n°3 d'Euler. On reconnaît, dans cette équation de condition, ce que l'on appelle aujourd'hui « résultant » des deux équations proposées, et que l'on écrit sous une forme que l'on appelle parfois « déterminant de Sylvester » ou « résultant de Sylvester » :

$$\begin{vmatrix} A & B & C & ... & V & & & & \\ & A & B & C & ... & V & & & \\ & & ... & & & & & & \\ & & & ... & & & & & \\ & & & & A & B & C & ... & V \\ A' & B' & C' & ... & V' & & & & \\ & A' & B' & C' & ... & V' & & & \\ & & ... & & & & & & \\ & & & ... & & & & & \\ & & & & A' & B' & C' & ... & V' \end{vmatrix} = 0$$

C'est une équation en y, et Bézout calcule son degré au moyen du corollaire au lemme II énoncé ci-dessus. Il trouve $(mm' + pm' + p'm)$.

La section VI du mémoire est intitulée « Procédé de la méthode pour l'élimination et réflexions qui tendent à l'abréger ». Le « procédé de la méthode » consiste, comme on l'a vu, à ramener le problème de l'élimination à un système d'équations linéaires. C'est par l'étude d'un tel système linéaire que Bézout est parvenu à calculer le degré de l'équation finale. Mais la résolution effective du système linéaire, conduisant au calcul effectif de l'équation finale, peut être lourde en calculs, et il convient de trouver des raccourcis : c'est là l'objet de la section VI. Bézout tente de réduire le nombre d'équations dans le système linéaire et parvient à une méthode qui gardera par la suite le nom de « méthode d'élimination abrégée de Bézout ». Soit donc :

$$\begin{cases} Ax^m + Bx^{m-1} + Cx^{m-2} + ... + T = 0 \\ A'x^{m'} + B'x^{m'-1} + C'x^{m'-2} + ... + T' = 0 \end{cases}$$

où $m' \leq m$. La méthode exposée plus haut consistait à « multiplier la première par un polynôme indéterminé du degré $m'-1$, et la seconde par un polynôme indéterminé du degré $m-1$, les ajouter

et égaler à zéro les coefficients de chaque puissance de x ». Elle conduisait à un système linéaire de taille $(m + m')$. Bézout propose une alternative conduisant à un système de taille m :

> 1° On multipliera la première par A', et la seconde par $Ax^{m-m'}$, et on retranchera le second produit du premier.
>
> 2° On multipliera la première par $A'x + B'$, et la seconde par $Ax^{m-m'+1} + Bx^{m-m'}$, et on retranchera le second produit du premier.
>
> 3° On multipliera la première par $A'x^2 + B'x + C'$, et la seconde par $Ax^{m-m'+2} + Bx^{m-m'+1} + Cx^{m-m'}$, et on continuera ainsi jusqu'à ce qu'on ait m' équations équations qui seront chacune du degré $m - 1$; alors on multipliera chacune de ces équations par un coefficient indéterminé, et l'équation $A'x^{m'} + ...$ par le polynôme indéterminé $M'x^{m-m'-1} + N'x^{m-m'-2} + P'x^{m-m'-3} + ...$, et on ajoutera ensemble tous ces produits, et on égalera à zéro la somme des coefficients de chaque puissance de x. [12]

Dans le cas $m = m'$, Sylvester a nommé la matrice du système d'équations linéaires ainsi construit *bezoutic square*, nous l'appellerons « matrice de Bézout ». Cette méthode a pu être inspirée à Bézout par les travaux d'Euler [13]. Observons en effet le cas où $m = m' = 3$. Bézout nous enjoint tout d'abord d'établir les trois équations suivantes :

$$(1)\begin{cases} (A'B - AB')x^2 + (A'C - AC')x + A'D - AD' &= 0 \\ (A'C - AC')x^2 + (A'D - AD' + B'C - BC')x + B'D - BD' &= 0 \\ (A'D - AD')x^2 + (B'D - BD')x + C'D - CD' &= 0 \end{cases}$$

Les opérations suivantes reviennent en fait à considérer le système linéaire :

$$(2)\begin{cases} (A'B - AB')M + (A'C - AC')N + (A'D - AD')P &= 0 \\ (A'C - AC')M + (A'D - AD' + B'C - BC')N + (B'D - BD')P &= 0 \\ (A'D - AD')M + (B'D - BD')N + (C'D - CD')P &= 0 \end{cases}$$

L'équation de condition de ce système est l'équation finale recherchée [14].

Observons à présent la méthode n°1 d'Euler (*cf.* p. 105 ci-dessus). Euler ramenait le cas de deux équations de degré 3 au cas de deux équations de degré 2, puis il ramenait ces deux équations de degré 2 à deux équations linéaires, par une méthode semblable. Utilisons les notations de Bézout pour désigner les deux équations de degré 3. Les deux équations de degré 2 sont construites par Euler de la manière suivante :

$$\begin{aligned} &A'(Ax^3 + Bx^2 + Cx + D) - A(A'x^3 + B'x^2 + C'x + D') = 0 \\ &[D'(Ax^3 + Bx^2 + Cx + D) - D(A'x^3 + B'x^2 + C'x + D')]\tfrac{1}{x} = 0 \end{aligned}$$

La première équation coïncide avec la première ligne du système (1), et la deuxième équation avec la dernière ligne (sauf le signe), comme le montre le calcul :

$$\begin{aligned} & D'(Ax^2 + Bx + C) - D(A'x^2 + B'x + C') \\ = \;& (A'x^3 + B'x^2 + C'x + D')(Ax^2 + Bx + C) - (Ax^3 + Bx^2 + Cx + D)(A'x^2 + B'x + C') \\ = \;& -[(A'D - AD')x^2 + (B'D - BD')x + C'D - CD'] \end{aligned}$$

Euler procédait de même à partir de ces deux équations de degré 2 pour obtenir deux équations de degré 1, puis encore de même pour obtenir une équation finale unique. Mais l'équation finale ainsi obtenue étant de degré 12, elle contient un facteur superflu de degré 3, qui est $AD' - A'D$. Bézout a donc pu s'inspirer de la méthode d'Euler, mais sa propre méthode a l'avantage d'éviter le facteur superflu.

Revenons au mémoire de Bézout. Dans le cas où $m' < m$, il propose une troisième possibilité, conduisant à un système linéaire de taille m' :

12. *Cf.* [26], p. 322.
13. Comme le pense Jacobi, [163] § 1.
14. Bézout remarque que la solution du système (2) revient à considérer chaque puissance de x comme une inconnue puis à résoudre le système (1) comme s'il s'agissait d'un système linéaire. Ceci est rendu possible par le fait que le déterminant d'une matrice est égal à celui de sa transposée.

> On formera toujours, comme il vient d'être dit, les m' équations du degré $m-1$; mais au lieu de multiplier l'équation $A'x^{m'}+...$, par le polynôme indéterminé $M'x^{m-m'-1}+...$, on prendra dans cette équation la valeur de $x^{m'}$, avec laquelle on formera, en multipliant successivement par x, et substituant à mesure la valeur de $x^{m'}$, toutes les puissances de x supérieures à $x^{m'}$, jusqu'à x^{m-1} ; on substituera ces valeurs dans chacune des m' équations, et on n'aura plus que m' équations chacune du degré $m'-1$, (...) [15]

Ce passage (p. 323–327) met en lumière toute la force de la méthode imaginée par Bézout : cette troisième possibilité conduit en effet rapidement à l'équation finale, mais elle a l'inconvénient d'introduire un facteur superflu. Seulement, la connaissance préalable du degré de l'équation finale véritable, garantie par les premières sections du mémoire, permet à Bézout de calculer le degré du facteur superflu, puis, par un procédé que l'on nommerait aujourd'hui « spécialisation », de calculer effectivement ce facteur superflu. Considérons l'exemple $m = 9$, $m' = 3$. Notons $d^oA = d^oB - 1 = d^oC - 2 = ... = p$ et $d^oA' = d^oB' - 1 = d^oC' - 2 = ... = p'$. On est conduit à trois équations de degré 2. Bézout calcule le degré de leurs coefficients, puis le degré de l'équation finale résultante. Il trouve $3p + 21p' + 27$. Mais la valeur prédite par la section I du mémoire était $mm' + pm' + p'm = 3p + 9p' + 27$. Il y a donc un facteur superflu de degré $12p'$. Voici alors ce que propose Bézout :

> Il faut remarquer que la nouvelle méthode donne le même résultat littéral, quel que soit p' ; donc l'équation finale doit avoir le même diviseur littéral quand p' est zéro, que quand il ne l'est pas ; or quand p' est zéro, l'équation finale est à son véritable degré, puisqu'elle est alors du degré $3p + 27$; donc le diviseur qu'elle peut avoir, et qu'elle a en effet, ne peut être qu'une puissance de A', puisqu'il n'y a que A' qui soit de zéro dimensions ; donc puisque ce diviseur doit être le même dans tous les cas, on en doit conclure que l'équation finale dans le cas présent sera divisible par A'^{12}. [16]

En termes modernes, on pourrait décrire ce que fait Bézout en disant qu'il « spécialise » le coefficient A', en lui affectant une valeur scalaire (de degré zéro). Puisque la formule littérale de l'équation finale reste inchangée après « spécialisation », la connaissance du facteur superflu dans le cas « spécial » permet de remonter au cas général [17].

8.2 Le cas de trois équations : échec

Dans le cas de deux équations, une circonstance particulière facilitait grandement le calcul de l'équation finale : les deux polynômes multiplicateurs étant choisis de degrés $(m-1)$ et $(m'-1)$, l'un fournissait m coefficients indéterminés, et l'autre m'. La somme des produits étant de degré $(m + m' - 1)$, on obtenait immédiatement un système linéaire de $(m + m')$ équations à $(m + m')$ inconnues. L'élimination d'une inconnue parmi trois suit le même schéma. Toute la difficulté est alors de choisir les degrés des polynômes multiplicateurs, de telle sorte que l'on obtienne un système linéaire ayant autant d'équations que d'inconnues. La solution de ce problème n'est pas unique. Des choix différents des degrés conduisent à une équation finale différente. On retient parmi ces choix ceux qui conduisent à deux équations finales de degrés minimaux : on se ramène ainsi à l'élimination déjà étudiée d'une inconnue parmi les deux restantes, au moyen des deux équations.

Bézout donne quelques exemples d'application de sa méthode. Dans le cas de trois équations de même degré m, il élimine d'abord *une* inconnue en comparant les équations trois à trois de deux manières différentes. Il obtient ainsi deux équations à deux inconnues, et il est conduit, pour le degré de l'équation finale, aux valeurs $\frac{3}{4}m^2(\frac{3}{4}m^2+1)$ lorsque m est pair, et $(\frac{3m^2+1}{4})^2$ lorsque m est impair [18]. Ce résultat, strictement inférieur à m^4 (valeur obtenue lorsqu'on élimine les inconnues en prenant les équations sucessivement deux à deux), est cependant toujours bien supérieur à la

15. *Cf.* [26], p. 323.
16. *Cf.* [26], p. 326.
17. Sur la « spécialisation », voir nos remarques générales *infra* p. 173.
18. *Cf.* [26], p. 306–307.

valeur m^3 que donnera Bézout en 1779. Mais Bézout semble bien être déjà conscient, en 1764, de l'imperfection de sa méthode (*cf.* [26], p. 291). Seul le cas de deux équations est traité avec exactitude.

Sans entrer dans les détails du calcul du degré de l'équation finale, nous conclurons en rappelant les trois idées fondamentales qui se trouvent dans le mémoire de 1764, et qui resteront à la base des méthodes exposées ultérieurement dans le traité de 1779 :

- l'équation finale peut s'obtenir comme une somme dont chaque terme est le produit d'une des équations proposées et d'un certain polynôme multiplicateur ;
- comme Bézout écrit lui-même, « je réduis, dans ce mémoire, tout le travail de l'élimination, à quelque degré que montent les équations, je le réduis, dis-je, à éliminer des inconnues au premier degré » [19] ;
- Bézout distingue bien deux étapes dans l'élimination : le calcul du degré de l'équation finale [20], puis le calcul effectif de l'équation finale [21]. Il s'agit là d'un ordre important, car la connaissance préalable du degré de l'équation finale permet dans certains cas de repérer l'existence d'un facteur superflu, et de le supprimer.

8.3 La matrice de Bézout chez le mathématicien japonais Seki Takakazu

Le mathématicien japonais Seki Takakazu (?–1708) écrivit un traité intitulé « Méthode pour résoudre les problèmes dissimulés » (*Kaifukudai no hô*), daté de 1683. Ce traité, traduit en français dans la thèse de A. Horiuchi (annexe n° 4) s'inscrit dans un ensemble d'ouvrages concernant des techniques de résolution de problèmes apparentés à des systèmes d'équations polynomiales. Les recherches japonaises furent fécondées, dans la seconde moitié du dix-septième siècle, par la découverte partielle des travaux du mathématicien chinois de la fin du treizième siècle Zhū Shìjié [22]. Les japonais connurent le *Sùanxué qǐméng* (1299) de cet auteur, mais certainement n'eurent-ils pas accès à son *Sìyuán yùjiàn* (1303, nous avons parlé de cet ouvrage *supra* section 3.3). Cela n'empêcha pas Seki Takakazu d'exposer des techniques de transformations algébriques conduisant à l'élimination des inconnues, procédé certes nécessaire pour réduire un système d'équations algébriques à une unique équation résoluble numériquement. Le climat de rivalité entre écoles est illustré par la présence, dans de nombreux traités de mathématiques japonais de cette période, de « problèmes légués » (*idai*), énoncés de problèmes particulièrement difficiles et publiés sans leurs solutions. Par ces « problèmes légués », l'auteur offrait une publicité pour les techniques enseignées dans son école, supposées les résoudre. Cette compétition intense, nous rappelant le contexte des recherches des derniers grands *Rechenmeister* en Allemagne au début du XVII$^{\text{ème}}$ siècle et le fameux *Cubicossischer Lustgarten* de Faulhaber dont les problèmes furent résolus par son rival Peter Roth, mena Seki Takakazu à concevoir des méthodes d'élimination qui ont peu à envier aux méthodes européennes du XVIII$^{\text{ème}}$ siècle. Les historiens ont déjà fait de Seki le collègue de son contemporain Leibniz dans l'invention des déterminants, sauf en ce qui concerne la règle des signes des termes du déterminant que donneront ses successeurs [23], certes avant 1700. Dans le *Kaifukudai no hô*, Seki enseigne comment résoudre des systèmes de n équations polynomiales à n inconnues pouvant s'écrire sous la forme suivante, en

19. *Cf.* [26], p. 291.
20. Exposé dans les cinq premières sections de son mémoire de 1764, il le sera dans le livre I de son traité de 1779.
21. Exposé dans la section VI de son mémoire, et dans le livre II de son traité.
22. Sur la réception de la méthode chinoise de résolution numérique des équations par les japonais et en particulier Seki Takakazu, en dépit du peu d'informations qu'offrait le *Sùanxué qǐméng* de Zhū Shìjié, *cf.* [153] p. 92 et 202–207.
23. *Cf.* [153], p. 196.

quelque sorte triangulaire sauf pour les deux dernières équations [24] :

$$\begin{cases} P_1(x_1, x_2) = 0 \\ P_2(x_1, x_2, x_3) = 0 \\ \vdots \\ P_{n-2}(x_1, ..., x_{n-1}) \\ P_{n-1}(x_1, ..., x_n) \\ P_n(x_1, ..., x_n) \end{cases}$$

La méthode se résume à éliminer x_n entre les deux dernières équations, pour obtenir un nouveau système de $(n-1)$ équations à $(n-1)$ inconnues, et ainsi de suite. Pour l'élimination d'une inconnue parmi deux équations, Seki semble avoir eu recours à plusieurs procédés distincts, dont l'un est le calcul du déterminant de la matrice de Bézout, pas toujours infaillible puisqu'il ne connaissait pas la règle des signes des termes du déterminant. Soient deux polynômes de même degré m

$$\begin{cases} P(x) = a_0 + a_1 x + ... + a_m x^m \\ Q(x) = b_0 + b_1 x + ... + b_m x^m \end{cases}$$

Seki cherche alors le déterminant de la matrice des coefficients des m polynômes de degré $(m-1)$ suivants :

$$\begin{cases} R_1(x) = a_m Q(x) - b_m P(x) \\ R_2(x) = a_{m-1} Q(x) - b_{m-1} P(x) + x R_1(x) \\ \vdots \\ R_m(x) = a_1 Q(x) - b_1 P(x) + x R_{m-1}(x) \end{cases}$$

L'on se convainc facilement que ce déterminant coïncide avec celui décrit par Bézout en 1764. Il s'agit d'une formule bien efficace, qui aurait certainement réjoui Leibniz s'il l'avait connue. Ce sera encore dans un climat de rivalité que la matrice de Bézout (ainsi que la forme quadratique associée, le bézoutiant) connaîtra, vers le milieu de XIX$^{\text{ème}}$ siècle, un nouveau succès, auprès des tenants de la théorie des invariants, mais non plus comme technique de résolution de problèmes (voir chapitre 12).

24. *Cf.* [153], chapitre 6, § 1.3 et 2.1. Nous traduisons ici en termes algébriques.

Chapitre 9

Les *Réflexions sur la résolution des équations algébriques* de Lagrange

Lagrange publie ce mémoire en deux parties en 1770 et 1771 dans les Mémoires de l'Académie de Berlin (il succède à Euler, à l'Académie de Berlin, de 1766 à 1788). Le mémoire comprend quatres sections. Les trois premières concernent respectivement les équations (à une inconnue) de degrés 3, 4, puis 5 et supérieur à 5. Elles passent en revue les différentes méthodes de résolution algébrique existantes au temps de Lagrange. [1]

Ce sont les méthodes de Tschirnhaus et d'Euler/Bézout qui retiennent le plus l'attention de Lagrange, par leur caractère général [2]. La méthode de Tschirnhaus, issue de son article de 1683 *Methodus auferendi omnes terminos intermedios ex data aequatione*, est l'occasion pour Lagrange d'exposer plusieurs théorèmes liés à la théorie de l'élimination, ainsi que d'aborder un problème important : la décomposition d'une équation de degré élevé en facteurs de degrés inférieurs. Ce problème interviendra de nouveau dans le mémoire [190] *Sur la forme des racines imaginaires des équations* à l'occasion de la démonstration par Lagrange du théorème fondamental de l'algèbre, en 1772. Quant aux méthodes d'Euler et Bézout, Lagrange mentionne les textes [102], [106], [25], [27] des deux auteurs. Dans son mémoire [27] de 1765, Bézout présente deux variantes de sa méthode, la deuxième étant spécialement destinée aux équations de degré composé. La méthode de Bézout ressemble beaucoup à la méthode d'Euler, et Lagrange les présente donc ensemble. Elles font intervenir une conjecture sur la forme des solutions. Les calculs auxquels est conduit Lagrange dans ces deux méthodes sont moins compliqués que dans la méthode de Tschirnhaus, peut-être parce que, dans les méthodes d'Euler et Bézout, certaines expressions, les «résolvantes de Lagrange», sont *linéaires* en fonction des racines de l'équation proposée.

La quatrième et dernière section du mémoire de Lagrange contient la démonstration de théorèmes utilisés dans les sections précédentes, ainsi que des considérations générales sur la résolution algébrique.

9.1 Exposé par Lagrange de la méthode de Tschirnhaus pour le degré 3

Nous avons déjà rencontrée cette méthode *supra* p. 75. Cette méthode consiste à éliminer x dans les équations :

1. Le mémoire de Lagrange est déjà bien connu (*cf.* [321], et [157] p. 44–49).
2. Dans ce chapitre, nous utiliserons l'expression « méthode de Bézout », « méthode d'Euler », pour désigner, en général, leurs méthodes de résolution algébrique d'une équation ; et non, comme dans le reste de cet ouvrage, leurs méthodes d'élimination.

$$\begin{cases} x^3 + nx + p = 0 \\ x^2 = bx + a + y \end{cases}$$

On obtient ainsi un avant-dernier reste, linéaire en x, et un reste de la forme $y^3 + Ay^2 + By + C = 0$, où A est de degré 1 en a, b, et B est de degré 2. On pose ensuite $A = 0$ et $B = 0$; en éliminant a, on obtient une équation dite «réduite en b» du second degré.

Lagrange insiste sur un point qui n'est pas clair dans l'article de Tschirnhaus, mais qui, comme nous l'avons déjà mentionné, a été parfaitement compris par Rolle : une fois les valeurs de a et b calculées, il faut, afin de calculer les solutions, utiliser l'avant-dernier reste linéaire en x, et non l'équation $x^2 = bx + a + y$:

> On ne doit pas prendre indifféremment pour x une des racines de l'équation supposée[...] Ainsi que l'Auteur le fait ; car, pour que cela fût permis, il faudrait que cette équation renfermât deux des racines de la proposée, et par conséquent que b fût la somme de ces deux racines ; or, comme il n'y a pas plus de raison pour que b soit la somme de deux quelconques des trois racines de la proposée que de deux autres quelconques, il s'ensuit que b devrait avoir autant de valeurs différentes qu'il y a de manières de prendre les trois racines deux à deux, c'est-à-dire trois valeurs [...] ; au lieu que nous avons vu que la quantité b n'a que deux valeurs, puisqu'elle ne dépend que d'une équation du second degré. [3]

9.1.1 Théorie de l'élimination

Afin de justifier *a priori* que le degré de l'équation en y résultant de l'élimination dans la méthode de Tschirnhaus est bien 3, Lagrange aborde d'une manière générale le problème de l'élimination d'une inconnue entre deux équations polynomiales

$$\begin{cases} P(x) = 0 \\ Q(x) = 0 \end{cases}$$

Il décrit sa conception de l'élimination en quelques paragraphes, que nous reproduisons ci-dessous :

> (...) la méthode ordinaire d'élimination, suivant laquelle on fait disparaître successivement les plus hautes puissances de l'inconnue en déduisant, des deux équations données où la même inconnue se trouve élevée à des puissances quelconques, une suite d'autres équations où le plus haut degré de l'inconnue est successivement moindre, jusqu'à ce qu'on arrive à une équation où l'inconnue ne se trouve plus, et qui est le résultat de l'élimination ; cette méthode, dis-je, revient dans le fond à la même que celle qui sert à trouver le plus grand commun diviseur des deux quantités qui forment les premiers membres des deux équations données ; les restes que l'on aura par les divisions successives qu'il faudra faire donneront, étant égalés à zéro, les mêmes équations que celles qui proviennent de l'élimination ; le dernier reste où l'inconnue ne se trouve plus devra être égal à zéro pour que les deux quantités proposées aient un diviseur commun du premier degré, lequel sera par conséquent l'avant-dernier reste où l'inconnue ne sera que linéaire ; de sorte qu'en égalant aussi à zéro cet avant-dernier reste on aura une valeur de l'inconnue qui sera la racine commune des deux équations.
> (...) il est évident que si l'on veut qu'elles aient deux racines communes il faudra qu'elles soient divisibles exactement par un facteur du second degré ; par conséquent, en cherchant le plus grand diviseur des deux quantités qui forment les premiers membres des équations proposées, dès qu'on sera parvenu à un reste où l'inconnue se trouvera au second degré, il faudra, pour que ce reste soit un diviseur commun des deux équations, que le reste suivant soit nul de lui-même ; or ce dernier reste ne renfermera que deux termes, l'un où l'inconnue ne se trouvera pas, et l'autre où elle se trouvera à la première dimension ; c'est pourquoi il faudra faire chacun de ces termes en particulier égal à zéro,

3. *Cf.* [189], § 11.

> ce qui donnera deux équations contenant les conditions nécessaires pour qu'il y ait dans les proposées deux racines communes. [4]

La fin de cet extrait fait écho aux résultats déjà énoncés par Euler en 1764 (*cf.* [107]) sur le nombre des équations de conditions nécessaires pour faire que deux équations aient un nombre donné de racines communes. Dans les pages suivantes, Lagrange fait cependant appel à une autre interprétation de l'équation résultant de l'élimination. Si x', x'',... sont les racines du polynôme Q, cette équation sera le produit : $P(x')P(x'')\ldots$. Pour la démonstration, Lagrange renvoie à l'*Introduction à l'analyse des lignes courbes* de Cramer, ainsi qu'à son propre mémoire *Sur l'élimination des inconnues dans les équations* [5]. Il énonce ensuite le résultat général suivant sur le degré de l'équation résultante :

> ...il s'ensuit que cette équation doit être telle, que les coefficients de l'équation $P = 0$ y forment partout des produits d'autant de dimensions qu'il y a de quantités P', P'', P''',... ou de racines x',x'',x''',... dans l'équation $Q = 0$, c'est-à-dire autant qu'il y a d'unités dans le degré de cette dernière équation... [6]

Lagrange démontre encore le résultat suivant, qui précise celui d'Euler :

> ...si $r = 0$ est la condition nécessaire pour que les équations $P = 0$ et $Q = 0$ aient une racine commune, on aura, pour les conditions de deux racines communes,
>
> $$r = 0 \text{ et } \frac{dr}{d\rho} = 0 \text{ ;}$$
>
> pour trois racines communes,
>
> $$r = 0,\ \frac{dr}{d\rho} = 0 \text{ et } \frac{d^2r}{d\rho^2},$$
>
> et ainsi de suite, ρ étant le dernier terme de l'une des équations proposées [=terme constant]. [7]

Lagrange est d'ailleurs le premier, à notre connaissance, à utiliser la locution « théorie de l'élimination » pour désigner ce domaine de recherches [8].

9.1.2 Le coeur de la méthode de Lagrange : réduite et résolvante

Dans chacune des méthodes de résolution algébrique qu'il étudie, Lagrange cherche une expression des racines de la réduite, en fonction des racines de la proposée. Il appelle cette expression «résolvante». Pour la méthode de Tschirnhaus en degré 3, il établit par exemple que $b = \frac{x'^2+\alpha x''^2+\alpha^2 x'''^2}{x'+\alpha x''+\alpha^2 x'''}$ où α est une racine primitive cubique de l'unité. Les différentes valeurs obtenues à partir de cette expression de b en effectuant toutes les permutations possibles des racines x',x'',x''',... donnent l'ensemble des racines de la réduite, et le nombre de ces valeurs donne le degré de la réduite [9].

Cela découle d'un théorème plus général que Lagrange démontre dans la quatrième partie de son mémoire. Il introduit à cet effet la notation $f[(x')(x'')(x''')...]$ pour désigner une fonction rationnelle des μ racines de la proposée. En éliminant successivement les racines x',x'',... au moyen de l'équation

4. *Cf.* [189], § 11–12.
5. *Cf.* [188]. Dans ce court mémoire, il développe en série formelle l'expression

$$\log(P(x')P(x'')...) = \log P(x') + \log P(x'') + ...$$

et parvient à la calculer au moyen des sommes de puissances des racines de Q, qui, selon le théorème de Newton, sont rationnelles en les coefficients de Q.
6. *Cf.* [189], § 13.
7. *Cf.* [189], § 12.
8. *Cf. ibid.* § 15.
9. Lagrange calcule b en résolvant le système :

$$\begin{cases} x'^2 = bx' + a - \sqrt[3]{C} \\ x''^2 = bx'' + a - \alpha\sqrt[3]{C} \\ x'''^2 = bx''' + a - \alpha^2\sqrt[3]{C} \end{cases}$$

proposée, il remarque que cette expression est racine d'une certaine «réduite» $\Theta = 0$, dont les coefficients sont rationnels en les coefficients de la proposée, et il démontre que cette réduite est :

$$\Theta = \underbrace{(t - f[(x')(x'')(x''')...])(t - f[(x'')(x')(x''')...])...}_{toutes\ les\ permutations\ des\ racines\ x',x'',x''',...}$$

Cette réduite a des racines multiples lorsque certaines permutations laissent l'expression $f[(x')(x'')(x''')...]$ inchangée. Dans ce cas, Lagrange montre que le nombre ν de ces permutations divise le nombre total de permutations $\mu!$, que ν autres permutations laisseront fixe l'expression semblable $f[(x'')(x')(x''')...]$, et de même pour chaque autre expression semblable, de sorte que la réduite s'écrira $\Theta = \theta^\nu$ et pourra ainsi s'abaisser à une réduite de degré $\frac{\mu!}{\nu}$, à coefficients rationnels en les coefficients de la proposée [10].

En langage moderne, si l'on considère l'action du groupe des permutations des racines sur les expressions de la forme $f[(x')(x'')(x''')...]$, on dit que l'ordre du stabilisateur d'une telle expression divise l'ordre du groupe des permutations, ce qui découle simplement du fait que le stabilisateur est un sous-groupe, et que l'ordre d'un sous-groupe divise l'ordre du groupe, par un théorème bien connu, que l'on appelle parfois «théorème de Lagrange» (comme le faisait déjà Jordan dans son *Traité des substitutions*, Gauthier-Villars, Paris 1870, p. 25).

9.1.3 Interprétation moderne de la méthode de Tschirnhaus en degré 3

Soit $K = \mathbf{Q}(n, p, \alpha)$. Alors le corps de décomposition de l'équation proposée est $K(x', x'', x''')$ (où x',x'',x''' désignent les racines de la proposée), et $K(b)$ est une sous-extension correspondant (par la correspondance de Galois) à un sous-groupe d'indice 2 du groupe S_3, en l'occurence le groupe alterné A_3. On résume la situation dans le diagramme suivant :

$$\begin{array}{l} K(x', x'', x''') \\ \quad | \quad 3\ (\textit{équation}\ y^3 + C = 0) \\ K(b) \\ \quad | \quad 2\ (\textit{réduite en}\ b) \\ K \end{array}$$

9.1.4 Méthode de Tschirnhaus en degré quelconque ; théorème de Bézout

Il s'agit d'éliminer x entre les deux équations suivantes :

$$\left\{ \begin{array}{l} x^\mu + mx^{\mu-1} + ... = 0 \\ x^{\mu-1} + fx^{\mu-2} + gx^{\mu-3} + ... + l + y = 0 \end{array} \right.$$

On obtient une équation en y de la forme $y^\mu + Ay^{\mu-1} + ... + Uy + V = 0$, où A est de degré 1, B de degré 2,..., V de degré $\mu - 1$ en les coefficients f,..., l. En imposant que $A = 0$,..., $U = 0$, et en éliminant g,...,l entre ces équations, on obtient une réduite en f. Mais ici se pose le problème de l'élimination dans un système de plus de deux équations, à plusieurs inconnues. Lagrange ne résout dans son mémoire que le problème de deux équations, à *une* inconnue, tel que nous l'avons décrit ci-dessus. Il ne donne même pas (ni ici, ni dans son court mémoire [188]) l'énoncé général du résultat démontré par Cramer sur le degré de l'équation finale résultant de l'élimination d'une inconnue *parmi deux*, dans un système de deux équations. Pourtant, il connaissait ce dernier résultat, et il envisageait même sa généralisation à un nombre quelconque d'équations. C'est en tout cas ce que suggère l'extrait suivant du § 14, qui fait suite à son exposé sur l'élimination d'une inconnue, et qui renvoie à l'équation $y^\mu + Ay^{\mu-1} + ... + Uy + V = 0$ ci-dessus :

> En général, pour faire disparaître à la fois les termes $p^{\text{ième}}$, $q^{\text{ième}}$, $r^{\text{ième}}$,..., on aura à résoudre autant d'équations qu'il y aura de ces termes, avec un même nombre d'inconnues,

10. *Cf.* [189], § 97. Lagrange énonce au § 98, comme cas particulier de son théorème, le théorème de Newton sur les fonctions symétriques des racines : « ...théorème que nous avons déjà supposé dans les Sections précédentes comme évident par soi-même, mais dont la démonstration rigoureuse dépend des principes établis ci-dessus ».

> et ces équations seront des degrés $p-1$, $q-1$, $r-1$,..., en sorte que l'équation finale montera, en général, au degré $(p-1)(q-1)(r-1)....$ [11]

Il s'agit là du « théorème de Bézout » pour un nombre quelconque d'équations, théorème que Bézout, le premier, démontrera en effet, pour des équations générales, en 1779. Nous avons vu que Rolle avait déjà pressenti ce théorème au siècle précédent. Edward Waring l'a aussi rappelé (sans le démontrer) dans la première édition (1770) de ses *Meditationes Algebraicae*. Nous avons reproduit, traduit et annoté, en annexe E, un extrait de la préface à la seconde édition des *Meditationes algebraicae* (1782). Cet extrait décrit l'histoire de l'élimination, telle que la concevait Waring. En comparant les deux éditions des *Meditationes algebraicae* (1770 et 1782), on remarque l'écart qui les sépare. Entre ces deux dates est paru, en 1779, le traité de Bézout.

Si ce théorème de Bézout est vrai dans le cas où souhaite l'appliquer Lagrange, alors en faisant disparaître tous les termes sauf V, la réduite devrait être de degré $(\mu-1)!$. Lagrange démontre en effet ce résultat particulier, en étudiant le système suivant :

$$\left\{\begin{array}{l} x'^{\mu-1} + fx'^{\mu-2} + ... + l + u = 0 \\ x''^{\mu-1} + fx''^{\mu-2} + ... + l + \alpha u = 0 \\ x'''^{\mu-1} + fx'''^{\mu-2} + ... + l + \alpha^2 u = 0 \\ \vdots \end{array}\right. \tag{9.1}$$

où $u = \sqrt[\mu]{-V}$. Il remarque que les permutations circulaires des racines x',x'',x''',... ont le même effet que le changement de variable $u \to \alpha u$. Elles ne changent donc pas la valeur de f, et la réduite en f est donc bien de degré $(\mu-1)!$ d'après le théorème du § 97 mentionné plus haut.

9.1.5 Méthode de Tschirnhaus en degré 4

Pour $\mu = 4$, la réduite en f est donc de degré 6. Au moyen du système 9.1, Lagrange calcule aussi la résolvante en f, qui est de la forme $f = -\frac{M+N\sqrt{-1}}{P+Q\sqrt{-1}}$. Il détermine empiriquement les 6 valeurs prises par f en permutant les racines [12]. Parmi ces 6 valeurs, chaque couple de valeurs conjuguées est solution d'une équation de la forme $f^2 + tf + u$ où $t = \frac{2(MP+NQ)}{P^2+Q^2}$ et $u = \frac{M^2+N^2}{P^2+Q^2}$. Il cherche ensuite quelles sont les valeurs prises par t et u lorsque l'on permute les racines. Il indique que t et u restent inchangés par les permutations [13] : (12), (34), (13)(24). Il en déduit que les $4! = 24$ valeurs possibles de t et u sont égales 8 à 8 et qu'elles se réduisent donc à 3 valeurs différentes [14]. Il ajoute ensuite que la permutation (14)(23) ne change pas non plus les valeurs de t et u, mais qu'elle «ne doit pas entrer en ligne de compte, parce qu'il [cet échange] est déjà renfermé dans les précédents». On peut voir ici se profiler une des intuitions qui présideront à l'apparition de la notion de groupe [15].

---

11. *Cf.* [189], § 14.
12. *Cf.* [189], § 42.
13. On utilise ici les notations modernes pour désigner les permutations des 4 racines ; Lagrange n'a pas de notation particulière, il décrit les permutations verbalement.
14. *Cf.* [189], § 43.
15. Posons $K = \mathbf{Q}(m,n,p,q,\alpha)$. Si $K(a,b,c,d)$ est le corps de décomposition de l'équation de degré 4 proposée (a,b,c,d désignent les racines de la proposée), avec les notations précédentes, $K(f)$ est une sous-extension correspondant à un sous-groupe cyclique d'ordre 4 de S_4, et $K(t,u)$ est une sous-extension correspondant à un sous-groupe dihédral D_4, d'ordre 8, engendré par les permutations (12),(34),(13)(24), et contenant en particulier la permutation (14)(23)=(34) ∘ (13)(24) ∘ (34). On a le diagramme suivant :

$$\begin{array}{l} K(a,b,c,d) \\ \quad | \quad 4\ (\textit{corps de rupture de l'équation } y^4 + D = 0) \\ K(f) \\ \quad | \quad 2\ (\textit{corps de rupture de } f^2 + tf + u) \\ K(t,u) \\ \quad | \quad 3 \\ K \end{array}$$

9.1.6 Idée d'une méthode générale : décomposition de la réduite en facteurs de degrés inférieurs

La méthode qui vient d'être exposée pour $\mu = 4$ consiste en fait à décomposer la réduite de degré 6 en 3 équations de degré 2. Le problème général de la décomposition d'une équation en facteurs de degrés inférieurs joue un rôle théorique important dans le mémoire de 1770. Lagrange cherche à généraliser la méthode à la réduite générale de degré $(\mu - 1)!$. Penchons-nous sur le cas où μ est premier. A la suite de l'étude des propriétés des racines μ-ièmes de l'unité, Lagrange remarque qu'en changeant α en α^2,α^3,... dans la résolvante f, on obtient $(\mu-1)$ des valeurs prises par f en permutant les racines. D'où l'idée de décomposer la réduite en f en $(\mu-2)!$ équations de degré $(\mu-1)$, chacune de la forme $f^{\mu-1} + Ff^{\mu-2} + Gf^{\mu-3} + ...$; Lagrange propose deux méthodes différentes pour le faire :

Méthode 1 : on élimine α dans l'expression de la résolvante f au moyen de l'équation $\dfrac{\alpha^\mu - 1}{\alpha - 1} = 0$. Cette élimination peut se faire soit par l'élimination successive des plus hautes puissances de α, soit en calculant explicitement les $(\mu - 1)$ valeurs de la résolvante f obtenues en changeant α en α^2,α^3,..., puis en calculant explicitement les F, G,... Les expressions ainsi obtenues de F, G,... en fonction des racines de la proposée permettent finalement de déterminer les réduites en F, G,... (*cf.* [189] § 57, $2^{\text{ème}}$ alinéa, et § 58, $3^{\text{ème}}$ alinéa).

Méthode 2 : on pose que $f^{\mu-1} + Ff^{\mu-2} + Gf^{\mu-3} + ...$ est un diviseur de la réduite en f. Par la division euclidienne, on obtient un reste de degré $(\mu - 2)$, d'où $(\mu - 1)$ conditions conduisant par élimination à une équation en F de degré $(\mu-2)!$ (*cf.* [189] § 58, $2^{\text{ème}}$ alinéa). Cette méthode pose un problème : une fois F trouvé, comment déterminer les autres coefficients G, H,... ? On peut encore espérer que les avant-derniers restes fournis par le processus d'élimination donnent des expressions linéaires en G,H,... permettant d'exprimer ces quantités en fonction de F. Mais il reste, pour s'en convaincre, à démontrer le caractère général d'un tel procédé. C'est l'objectif visé par Lagrange dans la troisième partie de son mémoire, aux §§ 100–104. Certes, après avoir calculé la réduite en F, on peut toujours en calculer une autre en G, puis une en H, etc. Mais le problème reste entier : étant donné une racine de la réduite en F, comment savoir quelle racine de la réduite en G lui correspond ? Notons t', t'',... les racines de la réduite en F et y',y'',... celles de la réduite en G qui leur correspondent respectivement. Lagrange étudie le système suivant (*cf.* [189], § 58) :

$$\left\{\begin{array}{l} y' + y'' + y''' + ... = M \\ t'y' + t''y'' + t'''y''' + ... = M_1 \\ t'^2y' + t''^2y'' + t'''^2y''' + ... = M_2 \\ \vdots \end{array}\right. \tag{9.2}$$

Les racines t',t'',... de la réduite en F étant connues, et M, M_1,M_2,... étant des quantités rationnelles en les coefficients de la proposées, les valeurs de y',y'',... s'obtiennent en résolvant le systéme linéaire. C'est d'ailleurs ainsi que Lagrange démontre, au § 100, son théorème sur les « fonctions semblables » (nous reviendrons sur ce théorème dans la section suivante). Bien que ce procédé nécessite, à première vue, de connaître *toutes* les racines de la réduite en F, Lagrange parvient à s'en passer. Il aboutit à une expression de y rationnelle en t, de la forme :

$$y = -\frac{\frac{P}{t} + \frac{Q}{t^2} + \frac{R}{t^3} + \frac{S}{t^4} + ...}{A + 2Bt + 3Ct^2 + 4Dt^3 + ...}$$

où les lettres majuscules désignent des fonctions rationnelles des coefficients de la proposée.

Si au contraire μ n'est pas premier, Lagrange montre de même comment décomposer la réduite de degré $(\mu - 1)!$ en $\frac{(\mu-1)!}{\phi(\mu)}$ équations de degré $\phi(\mu)$ au moyen d'une équation de degré $\frac{(\mu-1)!}{\phi(\mu)}$. Abel affirmera plus tard que l'existence d'une racine rationnelle de cette réduite de degré $\frac{(\mu-1)!}{\phi(\mu)}$ est une condition nécessaire à la résolubilité algébrique de l'équation proposée (*cf.* [156] p. 63).

9.1.7 Théorème sur les fonctions semblables

Au paragraphe § 88 de son mémoire, Lagrange définit la notion de «fonctions semblables» :

> Si l'on a plusieurs fonctions des mêmes quantités, on appellera fonctions *semblables* celles qui varient en même temps ou demeurent les mêmes lorsqu'on y fait les mêmes permutations entre les quantités dont elles sont composées.

Lagrange introduit même de nouvelles notations pour désigner ces fonctions semblables. Au paragraphe § 100, il démontre un théorème sur les fonctions semblables, qu'il énonce plus loin dans les termes suivants :

> Si l'on a deux fonctions quelconques t et y des racines x',x'',x''',... de la proposée, et que ces fonctions soient telles, que toutes les permutations entre les racines x',x'',x''',... qui feront varier la fonction y, fassent varier aussi en même temps la fonction t, on pourra, généralement parlant, avoir la valeur de y en t et en m,n,p,..., par une expression rationnelle, de manière que connaissant une valeur de t on connaîtra aussi immédiatement la valeur correspondante de y ; nous disons *généralement parlant*, car s'il arrive que la valeur connue de t soit une racine double, ou triple, etc., de l'équation en t, alors la valeur correspondante de y dépendra d'une équation carrée, ou cubique, etc., dont tous les coefficients seront des fonctions rationnelles de t et de m,n,p,... [16]

Cet énoncé s'applique en particulier, comme on vient de le voir, aux quantités F et G dans le problème de la décomposition de la réduite de Tschirnhaus. De plus, Lagrange réutilisera ce théorème dans son mémoire de 1772 *Sur la forme des racines imaginaires des équations*, où il «remet sur pied» une démonstration du théorème fondamental de l'algèbre proposée par Euler en 1749, qui présentait justement le même problème de la décomposition d'une équation en facteurs de degrés inférieurs.

9.2 Méthodes d'Euler et Bézout : «résolvantes de Lagrange»

La méthode d'Euler/Bézout est exposée, pour les équations de degré quelconque, aux paragraphes § 66–68 du mémoire de Lagrange. Il s'agit d'éliminer y entre ces deux équations :

$$\begin{cases} x = a + by + ... + hy^{\mu-2} + ky^{\mu-1} \\ y^\mu + V = 0 \end{cases}$$

puis de comparer l'équation obtenue avec l'équation proposée $x^\mu + mx^{\mu-1} + nx^{\mu-2} + ... = 0$. On obtient un système de μ équations à $\mu + 1$ inconnues a,b,...,k et V. Bézout choisit $V = -1$. Par élimination, on obtient une réduite en b,c,... ou k. Les racines de la proposée sont alors données par l'expression $x = a + b\alpha + ... + h\alpha^{\mu-2} + k\alpha^{\mu-1}$, où α est racine μ-ième de l'unité [17].

A partir du système linéaire suivant (où α est racine primitive μ-ième de l'unité) :

$$\begin{cases} x' = a + b + ... + k \\ x'' = a + b\alpha + ... + k\alpha^{\mu-1} \\ x''' = a + b\alpha^2 + ... + k\alpha^{2(\mu-1)} \\ \vdots \end{cases}$$

16. *cf.* [189], § 104, alinéa n°1. Cet énoncé peut sembler contradictoire. En fait, Lagrange distingue l'expression de t en fonction des racines et des coefficients (formels) de la proposée, et sa valeur numérique lorsque l'on donne aux coefficients de la proposée des valeurs particulières. Deux expressions différentes de t peuvent avoir la même valeur numérique, et cette valeur sera alors racine multiple de l'équation en t. Il étudie séparément le cas où à une même expression de t répondent plusieurs expressions formelles différentes de y (*cf.* l'alinéa n°2 de l'énoncé du § 104, non reproduit ici).

17. Dans son exposé, Lagrange est soucieux de rendre apparente une certaine « continuité » entre les différentes méthodes de ses prédécesseurs. Ainsi, il fait le lien entre les méthodes de Tschirnhaus et d'Euler/Bézout : pour $\mu = 4$, dans la méthode de Tschirnhaus, l'avant-dernier reste est $x = -\dfrac{Q + Ry + (m - f)y^2}{S + Ty + y^2}$ (Q,R,S,T sont des fonctions de m, n, p, f, g, h). Lagrange trouve un inverse de $S + Ty + y^2$ modulo $y^4 + D$, ce qui lui permet d'écrire x sous la forme $a + by + cy^2 + dy^3$ (*cf.* [189], § 46) et de retrouver ainsi la seconde conjecture d'Euler, et la méthode proposée par Bézout en 1765.

on obtient les résolvantes :

$$\begin{cases} \mu k = x' + \alpha x'' + ... + \alpha^{\mu-1} x^{(\mu)} \\ \mu h = x' + \alpha^2 x'' + ... + \alpha^{2(\mu-1)} x^{(\mu)} \\ \vdots \\ \mu b = x' + \alpha^{\mu-1} x'' + ... + \alpha^{(\mu-1)(\mu-1)} x^{(\mu)} \\ \mu a = x' + x'' + ... + x^{(\mu)} \end{cases}$$

Au paragraphe § 69, Lagrange considère d'une manière générale les expressions de la forme

$$t = x' + \alpha x'' + \alpha^2 x''' + ...,$$

dites «résolvantes de Lagrange». Ces résolvantes, linéaires, déjà connues (indépendamment semble-t-il) de Vandermonde (dans son *Mémoire sur la résolution des équations* publié en 1770-1771) joueront un rôle important dans les travaux ultérieurs de Gauss [18], Abel et Galois. L'expression t^μ reste inchangée par une permutation circulaire des racines x',x'',... (cela découle des propriétés des racines μ-ièmes de l'unité). Donc on a une réduite en $\theta = t^\mu$ de degré $(\mu - 1)!$. On retrouve ici aussi le problème de décomposer cette réduite en facteurs de degrés inférieurs. Lagrange utilise à cet effet les mêmes outils théoriques que dans la méthode de Tschirnhaus.

Pour $\mu = 5$, on a donc une réduite en $\theta = b^5$ de degré $4! = 24$. Lagrange montre que cette réduite se décompose en 6 équations de degré 4 au moyen d'une réduite de degré 6, dont il calcule explicitement le premier coefficient. Malfatti, indépendamment de Lagrange, est parvenu lui aussi à une réduite de degré 6 de l'équation générale de degré 5, et l'a calculée entièrement, dans un article de 1771.

Le problème de la décomposition d'une équation en plusieurs facteurs est un problème théorique important. En voici un autre exemple. Dans son mémoire de 1765, Bézout introduit une variante de sa méthode de résolution, spécialement adaptée aux équations de degrés composés. Par exemple, il cherche à traiter l'équation générale de degré $6 = 3 \times 2$ en posant (on adopte ici les notations de Lagrange) :

$$\begin{cases} x^6 + mx^5 + ... + rx + s = 0 \\ x^3 - (a + by)x^2 + (a' + b'y)x - (a'' + b''y) = 0 \\ y^2 - 1 = 0 \end{cases}$$

Il s'agit d'éliminer y entre les deux dernières équations, d'identifier terme à terme avec la première l'équation résultante, puis d'éliminer a, a', b', a'', b'' pour obtenir une réduite en b. Cela revient en fait à décomposer l'équation de degré 6 en deux facteurs cubiques de la forme

$$x^3 - (a \pm b)x^2 + (a' \pm b')x - (a'' \pm b'')$$

La réduite en b est de degré 20 et n'a que des termes carrés. Bézout prétend que cette réduite peut être décomposée à son tour en deux équations de degré 10 (*cf.* [27], p. 549). Lagrange montre, au paragraphe n°85 de son mémoire, que cette décomposition est impossible [19].

18. Dans le cas où l'équation proposée est $\frac{x^n-1}{x-1} = 0$ avec n premier, g un générateur du groupe multiplicatif $\mathbf{F}_n^*$, r une racine de l'équation $\frac{x^n-1}{x-1} = 0$ et R une racine de l'équation $x^{n-1} - 1 = 0$, on retrouve les « sommes de Gauss » $\sum_{k=0}^{n-2} R^k r^{g^k}$ introduites par Gauss dans la section VII des *Disquistiones Arithmeticae* de 1801 afin de résoudre par radicaux cette équation.

19. Plus exactement, il montre qu'une telle décomposition en deux facteurs ne peut pas se faire au moyen d'une équation de degré 2 à coefficients rationnels en les coefficients de la proposée.

9.3 L'application de la théorie de l'élimination aux grands problèmes de l'algèbre

Les prédécesseurs et les contemporains de Lagrange (Tschirnhaus, Euler, Bézout, Vandermonde, Malfatti, etc.) ne semblent pas avoir douté de la possibilité de résoudre algébriquement l'équation générale de degré 5. La plupart de leurs méthodes de résolution (celles de Faulhaber-Descartes, de Tschirnhaus, d'Euler-Bézout) utilisent la théorie de l'élimination. C'est peut-être Lagrange, qui, le premier, s'est posé la question de la *possibilité* d'une méthode générale de résolution algébrique [20] en degré 5. Au siècle suivant, Abel et Galois montreront qu'une telle méthode n'existe pas. Lagrange montre que les méthodes particulières de ses prédécesseurs pour la résolution des équations de degré $\mu \leq 4$ ne peuvent pas s'appliquer aussi facilement aux équations de degré $\mu > 4$, puisqu'elles conduisent à des réduites de degré supérieur au degré de la proposée. Même lorsqu'il est possible de décomposer ces réduites en facteurs de degrés inférieurs, les équations auxiliaires sont de degré trop élevé à partir de $\mu = 5$.

Mais Lagrange a découvert un outil qui lui permet de juger *a priori* du succès d'une méthode donnée de résolution algébrique, et même de formuler *a priori* toutes les méthodes possibles de résolution algébrique pour un degré donné (il le fait effectivement pour les degrés 3 et 4 aux paragraphes n°105 à 109 de la section IV de son mémoire) : étudier l'effet des permutations sur les fonctions rationnelles des racines de la proposée. L'enthousiasme de Lagrange à l'égard de ce nouvel outil se fait sentir dans les termes qu'il emploie pour le décrire :

> ...je donnerai à cette occasion les vrais principes, et pour ainsi dire, la métaphysique de la résolution des équations du troisième et du quatrième degré[s]. [21]

> Voilà, si je ne me trompe, les vrais principes de la résolution des équations et l'analyse la plus propre à y conduire ; tout se réduit, comme on voit, à une espèce de calcul des combinaisons, par lequel on trouve *a priori* les résultats auxquels on doit s'attendre. Il serait à propos d'en faire l'application aux équations du cinquième degré et des degrés supérieurs, dont la résolution est jusqu'à présent inconnue ; mais cette application demande un trop grand nombre de recherches et de combinaisons, dont le succès est encore d'ailleurs fort douteux, pour que nous puissions quant à présent nous livrer à ce travail ; nous espérons cependant pouvoir y revenir dans un autre temps, et nous nous contenterons ici d'avoir posé les fondements d'une théorie qui nous paraît nouvelle et générale. [22]

Cette notion de «permutation» n'est encore entre les mains de Lagrange qu'un accessoire, et non un objet d'étude à part entière, comme elle le deviendra dans l'oeuvre de Ruffini, Cauchy puis Galois qui dégageront la notion de «groupe de permutations». Lagrange n'utilise les permutations que dans le but de déterminer *a priori* le degré des réduites, et de les calculer effectivement.

Etant donné le rôle important joué par la théorie de l'élimination dans les tentatives du dix-huitième siècle, il n'est pas surprenant de la retrouver dans l'œuvre même de Galois. Les exposés modernes de la théorie de Galois semblent pourtant s'être affranchis presque entièrement de la théorie de l'élimination. Dans son premier mémoire (1830), Galois indique qu'il existe une fonction linéaire V des racines de l'équation proposée, « telle qu'aucune des valeurs que l'on obtient en permutant dans cette fonction les racines de toutes manières ne soient égales ». Il démontre ensuite, dans le « lemme III », que toutes les racines de l'équation s'expriment rationnellement en fonction de

20. *Cf.* par exemple [189] § 86 :

> ...ces méthodes se réduisent toutes à un même principe général, savoir à trouver des fonctions des racines de l'équation proposée, lesquelles soient telles : 1° que l'équation ou les équations par lesquelles elles seront données, c'est-à-dire dont elles seront les racines [...], se trouvent d'un degré moindre que celui de la proposée, ou soient au moins décomposables en d'autres équations d'un degré moindre que celui-là ; 2° que l'on puisse en déduire aisément les valeurs des racines cherchées. [...] mais est-il toujours possible de trouver de telles fonctions, pour les équations d'un degré quelconque, c'est-à-dire pour tel nombre de racines qu'on voudra ? C'est sur quoi il paraît très difficile de pouvoir prononcer en général.

21. *Cf.* [189], § 87.
22. *Cf.* [189], § 109.

V. Il s'agit en quelque sorte du « théorème de l'élément primitif », pour le corps de décomposition de l'équation proposée. Il pose $V = \phi(a, b, c, d, ...)$ où a,b,c,d,... sont les racines, et considère l'équation :

$$F(V, a) = \underbrace{(V - \phi(a, b, c, d, ...))(V - \phi(a, c, b, d, ...)...}_{\text{pour toutes les permutations laissant fixe } a} = 0$$

Il remarque que cette équation n'a qu'une seule racine commune a avec l'équation proposée. La théorie de l'élimination, telle qu'exposée par exemple dans le mémoire de Lagrange ou dans ceux d'Euler, fournit alors une expression de cette racine commune comme fonction rationnelle de V (par exemple en considérant l'avant-dernier reste dans le calcul du p.g.c.d. des deux équations). Galois n'a pas besoin d'en dire plus. Abel avait d'ailleurs déjà utilisé ce procédé [23]. Pourtant Poisson, chargé d'examiner le mémoire de Galois, écrira dans la marge : « la démonstration de ce lemme n'est pas suffisante ; mais il est vrai, d'après le n°100 du mémoire de Lagrange ». Galois ne retiendra pas une telle remarque : sa propre démonstration est suffisante, et elle est certes plus simple que celle de Lagrange. Comme on l'a vu, Poisson était pourtant, lui aussi, familier avec la théorie de l'élimination, et son *Mémoire sur l'élimination dans les équations algébriques* (1802) est très digne d'intérêt.

Ces recherches en résolution algébrique des équations n'ont donc pu aboutir aux preuves d'impossibilité d'Abel et Galois, qu'après une lente maturation mettant en scène, parmi les premiers rôles, la théorie de l'élimination. C'est un destin semblable qu'a eu le théorème fondamental de l'algèbre ; ici aussi la théorie de l'élimination s'est trouvée appliquée à la décomposition d'une équation. L'approche d'Euler (1749), remise sur pied par Lagrange en 1772, fut critiquée par Gauss qui ne pouvait admettre sans démonstration qu'une équation de degré n ait forcément n racines (ce que supposaient toujours Euler et Lagrange, qui ne faisaient que s'interroger sur la forme des racines, et non leur nombre) ; Gauss écrivit en 1799, à la suite d'une liste d'objections portant sur les démonstrations du théorème fondamental faites par ses prédécesseurs :

> Ce grand mathématicien [=Lagrange] s'est efforcé avant tout de combler les lacunes dans la première preuve d'Euler, et en particulier, il a vraiment achevé si profondément ce qui constituait mes seconde et quatrième objections ci-dessus, qu'il n'y reste rien à désirer ; sauf qu'il reste peut-être quelques points douteux dans son traitement antérieur de la théorie de l'élimination, sur lequel reposent toutes ses investigations. Au contraire, il ne résout pas du tout ma troisième objection ; toutes ses recherches sont en effet basées sur la supposition que chaque équation de degré n a vraiment n racines.

Outre cette objection, remarquons qu'il règne une certaine ambiguïté dans la notion de «fonction rationnelle des racines» dans le mémoire de 1770 : pour justifier l'action du groupe symétrique S_n sur les racines d'une équation de degré n, il faut faire l'*hypothèse* que les seules relations algébriques entre les racines sont les relations symétriques prévues par le théorème de Newton. Cela garantit le fait que les permutations des racines sont des automorphismes du corps contenant toutes les racines (le corps de décomposition de l'équation proposée) et que le groupe de Galois de ce corps est bien le groupe symétrique S_n. Lagrange ne formule pas cette hypothèse explicitement [24]. Il fait toutefois la différence entre la «forme» des fonctions rationnelles des racines et leur «quantité absolue» (par «quantité absolue», il faut ici entendre la valeur numérique prise par ces fonctions lorsque l'on donne aux coefficients de la proposée des valeurs particulières ; voir par exemple le paragraphe n°103 où Lagrange oppose explicitement ces deux notions). D'ailleurs la fin de la section IV du mémoire est consacrée à l'étude des équations dont les racines sont liées par des relations algébriques (autres que les relations symétriques).

---

23. *cf.* [2] t. 2, *Précis d'une théorie des fonctions elliptiques* (1829), chapitre II, § 1, p. 547. Galois connaissait probablement ce mémoire.

24. Par exemple, il s'exprime ainsi au paragraphe n°94 :

> ...l'équation qui donnera la valeur d'une fonction quelconque des mêmes quantités [les racines x',x'',...] devra donner également les autres fonctions qui viendront de toutes les permutations possibles entre elles. Cette proposition paraît même assez évidente par elle-même pour n'avoir pas besoin de démonstration...

Gauss écrivit une démonstration algébrique du théorème fondamental de l'algèbre en 1815 (ce fut sa deuxième démonstration ; il en donna quatre [25]) dans laquelle la théorie de l'élimination joue un rôle central. Soit $P(z) = z^n + a_1 z^{n-1} + ... + a_n$ un polynôme réel de degré n. Gauss cherche à montrer qu'il possède un facteur réel de degré 2. En identifiant terme à terme $P(z)$ au polynôme $(z^2 - \phi_{n-1} z + \phi_n)(z^{n-2} + \phi_1 z^{n-3} + ... + \phi_{n-2})$, on obtient n équations à n indéterminées ϕ_1,ϕ_2,...,ϕ_n. Si l'on y ajoute l'indéterminée U et l'équation $U - \phi_{n-1} X + \phi_n = 0$, où X est une quantité réelle arbitraire pour le moment, on obtient $(n+1)$ équations à $(n+1)$ inconnues [26]. Eliminons ϕ_1,...,ϕ_n pour obtenir une équation finale $f(U, X, a_1, ..., a_n)$. Si l'on spécialise les coefficients a_1,...,a_n en leur donnant pour valeurs les fonctions symétriques élémentaires λ_1,...,λ_n de n racines z_1,...,z_n, on peut montrer que $f(U, X, \lambda_1, ..., \lambda_n) = \prod_{j<k}(U - (z_j + z_k)X + z_j z_k)$, et ainsi calculer $f(U, X, \lambda_1, ..., \lambda_n)$ en fonction des coefficients λ_1,...,λ_n (car le membre de droite est symétrique en les racines). D'autre part, on peut aussi obtenir une expression de $f(U, X, \lambda_1, ..., \lambda_n)$ en spécialisant l'expression générale de $f(U, X, a_1, ..., a_n)$ obtenue auparavant. Gauss montre que les fonctions symétriques élémentaires λ_1,..., λ_n sont algébriquement indépendantes sur $\mathbb{Q}$, et il s'ensuit que les deux expressions doivent être égales, littéralement. On peut alors montrer facilement que X peut être choisi de telle sorte que l'équation $f(U, X, a_1, ..., a_n) = 0$ n'ait pas de racine double. Dans ce cas, les autres inconnues ϕ_1,...,ϕ_n peuvent s'exprimer rationnellement en U et donner effectivement un facteur de degré 2 du polynôme $P(z)$.

25. *Cf. Abrégé d'histoire des mathématiques*, Dieudonné, chap. II, § III.A sur les démonstrations du théorème fondamental de l'algèbre par Gauss. Nous nous sommes inspirés de cet exposé pour ce qui va suivre.

26. Remarquer que l'idée d'exprimer $U = \phi_{n-1} X + \phi_n$ avec un X arbitraire ressemble au changement de coordonnées $t = x + gy + hz + ...$ utilisé par Poisson dans sa démonstration du théorème de Bézout en 1802 (*cf. supra* section 7.1). Une idée semblable avait aussi été utilisée par Laplace pour démontrer le théorème fondamental de l'algèbre dans ses leçons de mathématiques à l'Ecole Normale (*cf.* [192] p. 79–80).

Chapitre 10

Le Théorème de Bézout pour un nombre quelconque d'équations

Nous avons vu dès les premières démonstrations du théorème de Bézout deux voies possibles. L'une, utilisée avec succès par Cramer, consiste à utiliser la théorie des fonctions symétriques des racines. Bézout semble être resté étranger à cette approche. Mais c'est en étudiant cette théorie, et plus généralement les fonctions des racines invariantes par certaines permutations, que Lagrange, dans son mémoire de 1770–1771, démontre le théorème de Bézout pour des systèmes particuliers de plus de deux équations ; la même année 1770, Waring énonce aussi le théorème pour plus de deux équations dans ses *Meditationes algebraicae* (*cf. supra* section 9.1.4). Bézout connut certainement assez vite ces travaux de Lagrange et de Waring. En effet, le mémoire de Lagrange le concerne directement puisque Lagrange y critique les méthodes de résolution algébrique des équations à une inconnue qu'il avait conçues dans les années 1760. Quant à Waring, il affirme dans la préface à la seconde édition de ses *Meditationes algebraicae* avoir envoyé un exemplaire de sa première édition dès 1770 à certains savants, dont Lagrange et Bézout [1]. C'est donc peut-être ces collègues qui firent découvrir à Bézout le théorème qui porte aujourd'hui son nom. En effet, bien que ses travaux des années 1762–1765 sur la résolution des équations algébriques montrent plusieurs exemples où ce théorème est vérifié dans des systèmes de plus de deux équations, nous avons vu que la méthode d'élimination adoptée par Bézout en 1764 le conduisait, à cause d'un facteur superflu, à une équation finale de degré bien plus élevé que le produit des degrés des équations initiales (*cf. supra* section 8.2). Or, Lagrange et Waring seront aussi parmi les premiers lecteurs du traité de Bézout de 1779, et ils répondront rapidement à cet ouvrage, Lagrange par lettre, Waring dans la seconde édition de ses *Meditationes algebraicae*.

L'autre voie possible pour démontrer le théorème de Bézout était justement celle qu'avait poursuivie Bézout en 1764. Si sa méthode, par laquelle il éliminait les inconnues *l'une après l'autre*, ne permettait pas d'éviter tout facteur superflu, c'est pourtant en la modifiant légèrement qu'il parvint en 1779 à une méthode correcte, fondée sur le concept d'« Equation-somme » :

> Nous concevons qu'on multiplie chacune des équations données, par un polynôme particulier, et qu'on ajoute tous ces produits. Le résultat sera ce que nous appellerons l'*Equation-somme*, laquelle deviendra l'équation finale par l'anéantissement de tous les termes affectés des inconnues qu'il s'agit d'éliminer. [2]

Autrement dit, Bézout postule que l'équation finale résultant de l'élimination *de toutes les inconnues sauf une* dans un système de n équations à n inconnues est une équation de degré minimal, de la forme $P_1R_1 + ... + P_nR_n = 0$. Cette méthode n'a plus le caractère itératif des méthodes précédentes (élimination successive des plus hautes puissances de l'inconnue ; comparaison des équations deux à deux ; élimination des inconnues l'une après l'autre ;...). Mais le problème est toujours de déterminer, d'abord, le degré des « polynômes multiplicateurs » R_i, ainsi que le degré de cette équation finale.

1. *Cf.* p. *xxi* de l'édition de 1782 des *Meditationes algebraicae*.
2. *Cf.* [29], § 224.

C'est là le « nœud de la difficulté » nous dit Bézout. Et c'est là qu'intervient le théorème général de Bézout, affirmant que l'équation finale (résultant de l'élimination de $(n-1)$ inconnues dans un système de n équations *génériques*) a pour degré le produit des degrés des équations proposées.

Nous exposerons la démonstration de Bézout seulement après avoir présenté les idées de Cayley (1847, 1848), qui suivent de près celles de Bézout et les éclairent, bien que Cayley ne le cite pas. Pour l'instant, notons que la structure du traité de Bézout est compliquée. Il n'introduit en fait son concept d'équation-somme que dans la deuxième partie (« livre second »), restant proche dans la première de la méthode d'élimination des plus hautes puissances de l'inconnue héritée du XVII$^{\text{e}}$ siècle. Peut-être était-il conscient que la *nouveauté* de sa propre méthode allait poser problème. Il avertit haut et fort des défauts des méthodes employées par ses prédecesseurs :

> Nous avons longtemps travaillé sans succès tant que nous n'avons attaqué quelques-unes des matières contenues dans cet ouvrage, qu'à l'aide des méthodes connues jusque là.

Pour appliquer ces méthodes, il faut se ramener au cas de deux équations. Par exemple, pour trois équations à trois inconnues x, y, z de degré m, il s'agit d'éliminer y entre les deux premières équations, puis entre la première et la troisième. On obtient ainsi deux équations à deux inconnues x, z de degré m^2. Si l'on élimine z entre ces deux équations, on obtient une équation finale en x de degré m^4. En fait, le degré ne devrait pas dépasser m^3. Cet abaissement du degré peut tenir à deux causes. D'une part, les coefficients des deux équations de degré m^2 peuvent présenter certaines relations qui font que l'équation finale à laquelle elles conduisent aura certains coefficients nuls. La méthode d'Euler et de Cramer prédit en effet le degré de l'équation finale seulement dans le cas d'équations générales, dont les coefficients sont algébriquement indépendants. D'autre part, il peut y avoir un « facteur superflu » causé par l'élimination successive des inconnues, et en général difficile à reconnaître et à supprimer. Nous y avons déjà fait allusion (*cf. supra* p. 120).

Penchons-nous davantage sur ce type de « facteurs superflus ». Eugen Netto, dans ses *Vorlesungen über Algebra*, en explique l'occurence en ces termes :

> Dans le système
>
> $$f_1(z_1,z_2,z_3)=0, f_2(z_1,z_2,z_3)=0, f_3(z_1,z_2,z_3)=0,$$
> $$R_{f_1,f_2}=g_3(z_2,z_3)=0, R_{f_1,f_3}=g_2(z_2,z_3)=0,$$
> $$R_{g_3,g_2}(z_3)=0,$$
>
> à chaque racine ζ_3 de $R_{g_3,g_2}(z_3)=0$ correspond certes une valeur de $z_2=\zeta_2$ telle que les deux équations $g_3(z_2,z_3)=0$, $g_2(z_2,z_3)=0$ soient vérifiées ; et à cette valeur correspond deux valeurs ζ_1', ζ_1'' telles que $f_1(\zeta_1',\zeta_2,\zeta_3)=0$, $f_2(\zeta_1',\zeta_2,\zeta_3)=0$, et $f_1(\zeta_1'',\zeta_2,\zeta_3)=0$, $f_3(\zeta_1'',\zeta_2,\zeta_3)=0$; mais $(\zeta_1',\zeta_2,\zeta_3)$ n'est peut-être pas racine de $f_3=0$ (...) [3]

Netto donne ensuite un exemple numérique. Nous préférons l'exemple suivant, plus élémentaire (d'ailleurs nous avons vu que Netto a fait une erreur dans son exemple, *cf. supra* p. 79) :

$$\begin{array}{ll}(1) & (z-x-y-1)(z+x-y+1)=0\\ (2) & z+x+y-1=0\\ (3) & z-x+y+1=0\end{array}$$

En éliminant z entre (1) et (2), on obtient $(-2x-2y)(-2y+2)=0$. En éliminant z entre (1) et (3), on obtient $(-2y-2)(2x-2y)=0$. En éliminant x entre ces deux équations, on obtient finalement $y(y+1)(y-1)=0$, équation dont la racine $y=0$ ne correspond à aucune solution du système ci-dessus (pas même une solution infinie, dans $\mathbb{P}^3$). Dès 1764, Bézout avait bien compris la situation :

> (...) quand, ayant plus de deux équations, on élimine en les comparant deux à deux ; quand on ne ferait monter chaque équation résultante de l'évanouissement d'une inconnue qu'au degré précis où elle doit monter, en vain chercherait-on dans chacune le diviseur qui, en les abaissant, ferait que l'équation finale serait d'un moindre degré ; aucune n'aura de diviseur ; ce ne pourra être qu'en les comparant entre elles qu'on trouvera une équation qui aura en effet un diviseur ; mais quel est le fil qui conduirait dans ce labyrinthe ? [4]

3. *Cf.* [236], vol. II, §406.
4. *Cf.* [26], p. 290. Nous avons déjà cité la fin de cette phrase de Bézout, *supra* p. 120.

Quand il écrivait ces lignes, il n'était pas encore sorti du labyrinthe. Quinze ans plus tard, dans son traité, il écrit en effet :

> (...) quoique par les moyens proposés dans ce Mémoire [le mémoire de 1764] on arrive en effet, toujours, à une équation finale beaucoup moins composée, que par les méthodes qu'on avait jusque là, néanmoins on n'arrive pas à l'équation finale la plus simple ; et quoique le facteur qui complique le résultat soit bien moins élevé que par les autres procédures, il est en général d'autant plus composé, que les équations proposées le sont elles-mêmes. (...) Il y avait longtemps que je soupçonnais que la cause générale des vices des méthodes employées pour cet objet, était la nécessité de n'éliminer les inconnues que successivement. [5]

De fait, l'élimination des inconnues à partir d'un nombre d'équations supérieur ou égal à trois ne réussissait donc que pour les équations linéaires. Même dans le cas de deux équations de degrés quelconques, les méthodes connues ne réussissaient pas toujours ; même pour les systèmes d'équations linéaires, Bézout prétend pouvoir améliorer les méthodes connues. Enfin il ne se limite pas à la considération des équations génériques (dites par Bézout « équations complètes », et qu'il traite en premier, aux § 42–48), mais il étudie des classes d'équations de plus en plus vastes, obtenues en imposant que certains coefficients soient nuls, ou vérifient certaines conditions. Pour ces classes d'équations, le théorème général énoncé ci-dessus ne donne qu'une majoration du degré de l'équation finale, degré que Bézout tente de calculer exactement dans chaque cas. Il en résulte un travail aux proportions gigantesques, ordonné suivant un gradient de généralité croissante, et laissant comprendre que le cas des équations génériques n'est qu'un cas très particulier (*cf. infra* section 10.7). Bézout nous prévient :

> Quelque idée que nos lecteurs aient pu se faire déjà de l'étendue de la matière que nous entreprenons de traiter, celle qu'il en prendra par la suite, surpassera probablement la première. [6]

L'œuvre de Bézout offre une solution concrète au problème technique de l'élimination, toutes les étapes du calcul étant représentées par une unique « équation-somme » et ainsi réduites à l'application de la méthode des coefficients indéterminés et à la résolution d'un unique système d'équations *linéaires*, mais elle offre aussi une avancée conceptuelle nous rapprochant de la théorie des idéaux. Pourtant l'exposé de Bézout manque de rigueur, et il fut perçu comme tel par Poisson, qui en 1802, plutôt que de corriger Bézout, préféra revenir à la méthode d'élimination par les fonctions symétriques des racines, pour l'étendre au cas général. Poisson écrivait :

> Ce théorème important est dû à Bézout ; mais la manière dont il l'a démontré n'est ni directe, ni simple, et même elle n'est pas exempte de toute difficulté. [7]

Après avoir décrit sa propre méthode, Poisson ajoutait cependant :

> La longueur des calculs qu'entraîne cette méthode, la rendra presque toujours impraticable, et nous ne l'avons indiquée ici que parce qu'elle donne la solution la plus directe du problème de l'élimination. Ainsi, dans la pratique, pour obtenir l'équation finale à son véritable degré, il faudra recourir aux méthodes que l'on trouve exposées avec beaucoup de détails dans la *Théorie des équations* de Bézout. [8]

Ayant ainsi présenté le livre de Bézout et évoqué sa postérité immédiate (finalement assez rare, et concentrée chez Lagrange, Waring, Poisson), nous allons maintenant tenter d'en décrire l'impact en théorie de l'élimination dans la première moitié du XIX$^{\text{ème}}$ siècle, en nous concentrant sur un fait dans l'histoire de cette théorie.

5. *Cf.* [29], p. vii.
6. *Cf.* [29], § 52.
7. *Cf.* [252], p. 199.
8. cf. [252], p. 203.

10.1 La théorie de l'élimination dans la première moitié du XIX$^{\text{ème}}$ siècle

Il y a un nouveau commencement, semble-t-il, de la théorie de l'élimination au début des années 1840. Non que la tradition se soit interrompue après Poisson. Deux cas particuliers ont beaucoup progressé : celui des équations linéaires, grâce à la théorie des déterminants, et celui de deux équations à une ou deux inconnues (en particulier le cas du discriminant, résultant de $P(x)$ et $P'(x)$), attisé par l'intérêt pour l'algorithme euclidien à la suite des travaux de Sturm (*cf. infra* chapitre 12). D'autre part, l'élimination algébrique profite aussi à cette période des méthodes de l'analyse. Ainsi, aboutissement de recherches remontant avant 1840, Ossian Bonnet définit en 1847 la multiplicité d'intersection de deux courbes en un point comme nous le verrons au chapitre 13.

Cependant, comme l'ont remarqué Brill et Noether au sujet du livre de Bézout « aussi célèbre que non lu », du temps de Jacobi, « le gros de son ouvrage était tombé dans l'oubli » [9]. Mais à partir des années 40, un ensemble de recherches tournées vers le cas général voient le jour, surtout chez les tenants de la théorie des invariants naissante. On verra apparaître l'opinion chez certains mathématiciens et historiens [10] que Sylvester et Richelot (1808–1875, un élève de Jacobi) auraient alors « découvert » la méthode d'élimination dite « dialytique », avec l'aveu que cette méthode n'est certes pas bien différente de celle d'Euler ou Bézout pour deux équations. Et aussi que Hesse l'aurait redécouverte, en 1843 [11]. Nous avons déjà remarqué que Leibniz connaissait cette méthode ; peut-être était-il encore plus proche des idées de Sylvester sur ce point, qu'Euler et Bézout ne l'étaient. Nous allons essayer de montrer ce qui distingue les travaux de Sylvester, Richelot et Hesse des travaux de Bézout sur l'*équation-somme*.

Sylvester écrit trois articles qui se font suite, en 1839, 1840 et 1841. Ces recherches sont stimulées initialement par le théorème de Sturm, comme le montre l'article de 1839, où Sylvester écrit :

> La théorie me fut suggérée à l'étranger, et dans les intervalles de la maladie... Les réflexions qu'avait d'abord excitées le théorème mémorable de Sturm furent avivées, lorsqu'étant présent à une séance de l'Institut de France, une lettre était lue, du Ministère de l'Instruction Publique, qui demandait une opinion sur l'intérêt de former des tables d'élimination entre deux équations de degrés 5 ou 6 contenant un terme répété. L'offre fut rejetée, sous prétexte du labeur excessif qui eût été requis. Je pense que ceci était très exagéré... [12]

Dans l'article de 1840, il décrit comment obtenir les « *prime derivatives* » c'est-à-dire les restes intermédiaires dans l'algorithme euclidien, ou encore « fonctions de Sturm », au moyen de la méthode dialytique. Aucune démonstration ne figure dans cet article, et Sylvester ajoute finalement :

> Grâce à l'ingéniosité bien connue et l'aide que m'a gentiment accordée un ami distingué, je crois pouvoir faire construire une machine pour exécuter le théorème de Sturm, et en fait tous les problèmes de dérivations, d'après la méthode ici exposée ;... [13]

Il est possible qu'à cette date, Sylvester ne connût pas les travaux d'Euler et Bézout en la matière, puisqu'il ne les cite pas, et il prétendra plus tard (en 1877) avoir découvert la méthode dialytique *en*

9. *Cf.* [49], p. 143 et 147.
10. Par exemple Max Noether, [243], p. 136. Voir Sylvester [297], [298] et [300], ainsi que Richelot [264].
11. *Cf.* [147].
12. *cf.* [297] p. 44. Voici l'extrait des *Comptes rendus hebdomadaire des séances de l'Académie des Sciences*, t. 9 (juillet-décembre 1839), Paris, Bachelier, 1839, p. 119 (section « correspondance » dans la séance du lundi 15 juillet 1839 présidée par Chevreul) :

> M. le Ministre de l'Instruction Publique adresse copie d'une lettre dans laquelle M. A. Vène, Chef de bataillon du Génie, émet le vœu que le Gouvernement, après avoir consulté l'Institut, fasse exécuter par le Bureau des Longitudes des Tables algébriques ayant pour objet l'élimination d'une inconnue entre deux équations littérales de degrés supérieurs en x et y (2^{e}, 3^{e}, 4^{e}, 5^{e}, 6^{e} degré).

13. *cf.* [298] p. 57.

enseignant à un élève [14]. Certes, la méthode dialytique n'est pas l'équivalent exact de la méthode d'Euler et Bézout, mais elle conduit exactement au même résultat. En 1841, Sylvester développe sa méthode dialytique dans le cas de trois équations quadratiques homogènes à trois inconnues x, y, z. Il envisage alors plusieurs procédés distincts, dont l'un [15] consiste à multiplier chaque équation par les trois monômes de degré 1, soit x, y, z. Ainsi pour le système

$$\begin{cases} U = 0 \\ V = 0 \\ W = 0 \end{cases}$$

on obtient les 9 équations de degré 3 :

$$\begin{cases} xU = 0 & yU = 0 & zU = 0 \\ xV = 0 & yV = 0 & zV = 0 \\ xW = 0 & yW = 0 & zW = 0 \end{cases}$$

Et comme il y a 10 monômes de degré 3, il manque une équation pour appliquer la méthode dialytique et former un déterminant. Sylvester propose de former une dixième équation de degré 3 au moyen du jacobien (ou plus exactement, un sous-multiple du jacobien) :

$$\frac{1}{8}.\begin{vmatrix} \frac{\partial U}{\partial x} & \frac{\partial U}{\partial y} & \frac{\partial U}{\partial z} \\ \frac{\partial V}{\partial x} & \frac{\partial V}{\partial y} & \frac{\partial V}{\partial z} \\ \frac{\partial W}{\partial x} & \frac{\partial W}{\partial y} & \frac{\partial W}{\partial z} \end{vmatrix} = 0$$

On obtient ainsi 10 équations en les 10 monômes de degré 3. Si l'on considère ces monômes comme les 10 inconnues indépendantes d'un système de 10 équations linéaires, l'élimination consiste simplement à former le déterminant de ce système. Cette méthode fait l'économie des coefficients indéterminés (auxquels on a recours dans la méthode de l'équation-somme). Enfin, elle se présente comme méthode symbolique, en ceci qu'elle use de l'ambivalence de l'expression symbolique des monômes, tantôt monômes, liés algébriquement, en les inconnues du système initial, tantôt inconnues indépendantes d'un système d'équations linéaires, ce qui permet le transfert des méthodes symboliques de l'algèbre linéaire (déterminants et bientôt matrices) à l'algèbre des formes de degré quelconque.

On ne s'étonnera donc pas que Sylvester, dans un de ses importants articles sur la théorie des invariants *On the principles of the calculus of forms*, décrive le « grand principe de la dialyse, qui, découvert initialement en théorie de l'élimination, s'étend, sous une forme ou sous une autre, à toute la théorie de la concomitance et des invariants » [16].

Mais pour comprendre l'application de la méthode dialytique à la théorie de l'élimination, il faut expliquer comment choisir les équations dont on forme un déterminant. Les premiers lecteurs de l'article de 1841 ont dû être surpris du recours au jacobien. La même année, Sylvester écrit un autre article plus général sur ce sujet [17]. Nous ne pouvons décrire ici en détail sa méthode. Il suffira d'indiquer qu'après avoir obtenu un premier ensemble d'équations dites *augmentatives* en multipliant les équations initiales par des monômes (comme les 9 équations de degré 3 ci-dessus), Sylvester construit d'autres équations dites *secondary derivatives* de la manière suivante. Il écrit chacune des n équations initiales sous la forme $x^\alpha F + y^\beta G + z^\gamma H + ...$ (où x, y, z,... sont les n inconnues à éliminer). Il obtient ainsi un système de n équations :

$$\begin{cases} x^\alpha F + y^\beta G + z^\gamma H + ... = 0 \\ x^\alpha F' + y^\beta G' + z^\gamma H' + ... = 0 \\ \vdots \end{cases}$$

14. Dans [296], Sylvester écrit : « I remember,..., how... when a very young professor, fresh from the University of Cambridge, in the act of teaching a private pupil the simpler parts of Algebra, I discovered the principle now generally adopted into the higher text books, which goes by the name of the *Dialytic Method of Elimination* ».

15. *Cf.* [300], exemple 4, p. 64 ; les autres procédés y sont indiqués en note de bas de page.

16. *Cf.* [305], p. 294. Voir *infra* p. 186.

17. *Cf.* [301].

On peut alors construire l'une des *secondary derivatives* en formant le déterminant :

$$\begin{vmatrix} F & G & H & \dots \\ F' & G' & H' & \dots \\ \vdots & & & \end{vmatrix} = 0$$

La latitude de choix dans l'écriture $x^\alpha F + y^\beta G + z^\gamma H + ...$ permet de former ainsi plusieurs *secondary derivatives*. Dans le cas $n = 3$ et $\deg U = \deg V = \deg W$, en prenant $\alpha = \beta = \gamma = 1$ et

$$\begin{array}{lll} F = \frac{1}{2}.\frac{\partial U}{\partial x}, & G = \frac{1}{2}.\frac{\partial U}{\partial y}, & H = \frac{1}{2}.\frac{\partial U}{\partial z} \\ F' = \frac{1}{2}.\frac{\partial V}{\partial x}, & G' = \frac{1}{2}.\frac{\partial V}{\partial y}, & H' = \frac{1}{2}.\frac{\partial V}{\partial z} \\ F'' = \frac{1}{2}.\frac{\partial W}{\partial x}, & G'' = \frac{1}{2}.\frac{\partial W}{\partial y}, & H'' = \frac{1}{2}.\frac{\partial W}{\partial z} \end{array}$$

on retrouve le jacobien. Observons un instant le cas $n = 2$, $\deg U = \deg V$, que Sylvester ne mentionne pas dans cet article. Si $1 \leq \alpha \leq n$ et $\beta = 0$, en prenant pour G (resp. G') la somme des termes de degré $< \alpha$ de U (resp. V), on obtient pour chaque valeur de α une équation

$$\begin{vmatrix} F & G \\ F' & G' \end{vmatrix} = 0$$

de degré $(n-1)$. Il n'y a pas besoin alors des *augmentatives*. Le déterminant obtenu en éliminant dialytiquement entre les n équations de degré $(n-1)$ n'est autre que le déterminant de la « matrice de Bézout » (*cf. supra* chapitre 8). C'est donc peut-être l'idée de la matrice de Bézout qui a inspiré à Sylvester cette méthode générale. La matrice de Bézout jouera d'ailleurs un rôle important dans son célèbre mémoire *On a theory of syzygetic relations*[18]. Enfin, en général, Sylvester remarque que si le degré des *augmentatives* est au moins $\sum t_i - n + 1$ (où t_i désigne le degré de la i-ème équation initiale), alors il y en a suffisamment pour construire un déterminant et il n'est pas besoin de former de *secondary derivative*. Ce sera le fondement de la méthode développée par Cayley en 1848. Hélas, la méthode d'élimination de Sylvester, sauf dans quelques cas particuliers comme les deux mentionnés ci-dessus, introduit un facteur superflu et ne conduit donc pas immédiatement au résultant. Jouanolou a donné en 1997 une étude complète des *secondary derivatives*, qu'il appelle « formes de Sylvester », dans [170], § 3.10.

Hesse étudie la même question en 1844, en revenant au formalisme de l'équation-somme, dans son article *Über die Elimination der Variabeln aus drei algebraischen Gleichungen vom zweiten Grade mit zwei Variabeln* [148]. Il connaît les articles de Richelot et Sylvester, ainsi que les travaux d'Euler (il étend la méthode d'élimination Euler-Cramer par les fonctions symétriques des racines au cas d'un nombre quelconque d'équations, comme l'avait déjà fait Poisson, mais indépendamment). Surtout, Hesse connaît le livre de Bézout et le cite.

Pour éliminer parmi trois équations quadratiques

$$\begin{cases} U = 0 \\ V = 0 \\ W = 0 \end{cases}$$

Bézout utilisait des « polynômes-multiplicateurs » de degré 2. En homogénéisant, on obtient une équation-somme de degré 4 :

$$AU + BV + CW = z^4 R$$

Et R=0 est alors l'équation finale résultant de l'élimination[19]. Hesse récrit le procédé de Sylvester (décrit ci-dessus) sous la forme d'une équation-somme : si ϕ est le jacobien de U, V, W, on peut obtenir une équation-somme de degré 3 au moyen de polynômes-multiplicateurs de degré 1 :

$$(*) \quad AU + BV + CW + \delta\phi = z^3 R$$

18. *Cf.* [306] et *infra* chapitre 12.

19. Remarquons que ceci coïncide avec l'un des procédé envisagé par Sylvester dans [300], p. 64–65, en note de bas de page ; procédé que nous n'avons pas décrit ci-dessus.

où $\deg A = \deg B = \deg C = 1$ et δ est une constante [20]. Hesse remarque d'ailleurs que le jacobien lui-même s'obtient par une équation-somme avec des polynômes multiplicateurs de degré 2 :

$$AU + BV + CW = z\phi$$

Il montre que les dérivées partielles de ϕ s'obtiennent de même, ainsi

$$AU + BV + CW + \delta\phi = z\frac{\partial\phi}{\partial x}$$

Et alors on parvient finalement à obtenir le résultant R par des polynômes-multiplicateurs de degré 0, en s'aidant des dérivées partielles du jacobien [21] :

$$aU + bV + cW + d\frac{\partial\phi}{\partial x} + e\frac{\partial\phi}{\partial y} + f\frac{\partial\phi}{\partial z} = z^2R$$

Dans le langage adopté plus tard par Hurwitz [22], on pourrait dire que Hesse étudie l'idéal des « *Trägheitsformen* »

$$\mathcal{M}^{-\infty}(U,V,W) = \bigcup_{k\geq 0}\left\{f \quad | \quad \mathcal{M}^k f \subset (U,V,W)\right\}$$

où $\mathcal{M}$ désigne l'idéal (x,y,z).

Hesse déduit aussi de ces travaux une propriété importante du hessien d'une cubique, dont il donne de belles applications géométriques. Nous mentionnerons seulement le résultat algébrique. Si U, V, W sont les dérivées partielles d'une cubique f et que ϕ est le hessien de f, alors le hessien du hessien est à son tour une cubique ψ, et au moyen de l'équation-somme obtenue en $(*)$, il peut se représenter sous la forme [23] :

$$\psi = df + \delta\phi$$

Cette notion d'« équation-somme » apparaît à nouveau, dans un cadre différent, dans un article de Cayley [60] intitulé *On the theory of involution in geometry* (1847). Cayley dit d'une fonction polynomiale homogène Θ de degré r qu'elle est « en involution » avec les fonctions $U, V, ...$ de degrés $m, n, ...$ si elle vérifie :

$$\Theta = AU + BV + ...$$

c'est-à-dire si $\Theta \in (U,V,...)_r$. Cayley indique qu'il y a « une application analytique de cette théorie, d'intérêt considérable, au problème de l'élimination entre un nombre quelconque d'équations [homogènes] contenant le même nombre de variables », élimination qu'il conçoit comme résultant de la méthode dialytique. Mais Cayley situe d'emblée son propos en dressant une liste de travaux dont l'origine est le « paradoxe de Cramer » (Euler, Cramer, Plücker, Jacobi, Hesse). Il s'interroge sur le « nombre de constantes arbitraires » que contient Θ, c'est-à-dire, en langage moderne, sur

$$\dim(U,V,...)_r$$

Suivons Cayley dans son application à la méthode dialytique. On multiplie U par chaque monôme de degré $r-m$ (pour r assez grand), puis V par chaque monôme de degré $r-n$, etc. Notons C l'anneau de polynômes. On obtient ainsi

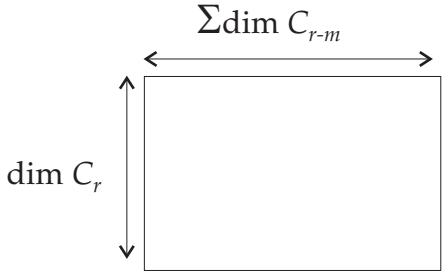
$$\sum \dim C_{r-m} = \dim C_{r-m} + \dim C_{r-n} + ...$$

20. *Cf.* [148], § 8 à 10.
21. *Cf.* [148], § 11 à 14.
22. *Cf.* [159]. Langage adopté aujourd'hui encore par Jouanolou [168]. Voir *infra* p. 159.
23. *Cf.* [148], § 15 et 16.

équations, desquelles on veut éliminer dialytiquement, en considérant les

$$\dim C_r$$

monômes de degré r comme étant les inconnues indépendantes d'un système d'équations linéaires. Représentons ce système par une matrice dont chaque colonne correspond à une de ces équations [24] :



Cette matrice est la matrice d'une application linéaire

$$f_1 : \bigoplus C_{r-m} \longrightarrow C_r$$

Mais si l'on revient à la notion d'« équation-somme » (qu'exprime Cayley par la notion d'« involution »), on voit que cette application s'écrit :

$$f_1(A, B, ...) = AU + BV + ...$$

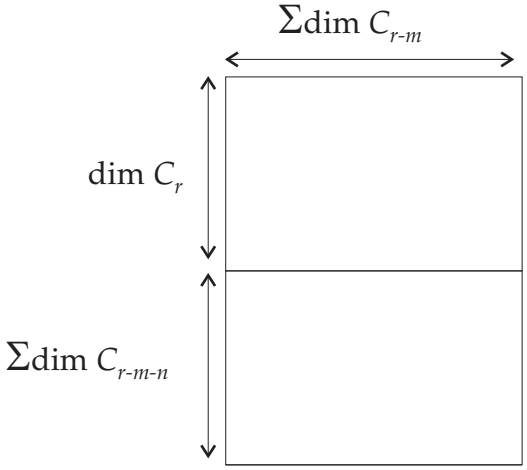
Cayley remarque que les colonnes de la matrice ne sont pas linéairement indépendantes, et il s'interroge sur le nombre de colonnes indépendantes, c'est-à-dire

$$\dim(\mathrm{Im} f_1)$$

Pour que l'élimination dialytique puisse avoir lieu, il faudrait qu'il y ait $\dim C_r$ colonnes indépendantes (ainsi l'on pourrait extraire une sous-matrice carrée, et en construire le déterminant). Pour vérifier que l'élimination est possible, Cayley entreprend donc de calculer

$$N = \dim C_r - \dim(Im f_1)$$

Les relations de dépendance linéaire entre les colonnes sont notées par Cayley au moyen de lignes de coefficients sous la matrice



24. Les figures matricielles suivantes apparaissent dans l'article [62] qui fait suite à [60]. Nous mêlons donc dans l'exposé ci-dessous des concepts qui, originellement, furent présentés de manière séparée par Cayley, son premier article réduisant le problème général de l'élimination à un problème d'algèbre linéaire, et son second article résolvant le problème d'algèbre linéaire (au moyen de ces figures matricielles).

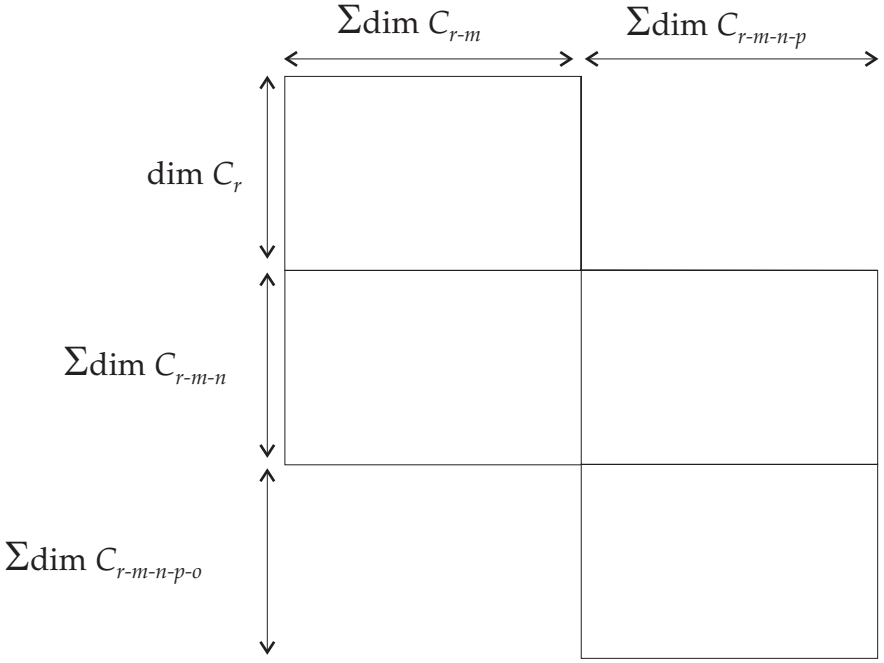
Par rapport à la notion d'équation-somme, ces relations de dépendance linéaire constituent le noyau de l'application f_1. Cayley admet (sans le démontrer) [25] que $\ker f_1$ est engendré par les éléments de $\bigoplus C_{r-m}$ de la forme

$$\left\{ \begin{array}{ll} (MV, -MU, 0, ..., 0) & \text{où} \quad M \in C_{r-m-n} \\ (MW, 0, -MU, 0, ..., 0) & \text{où} \quad M \in C_{r-m-p} \\ \vdots & \end{array} \right\}$$

Il est donc engendré par $\sum \dim C_{r-m-n}$ vecteurs. La hauteur du second bloc matriciel est donc $\sum \dim C_{r-m-n}$, et $\ker f_1$ est engendré par l'image d'une application

$$\begin{array}{rccl} f_2 : & \bigoplus C_{r-m-n} & \longrightarrow & \bigoplus C_{r-m} \\ & (M, 0, ..., 0) & \longmapsto & (MV, -MU, 0, ..., 0) \\ & (0, M, 0, ..., 0) & \longmapsto & (MW, 0, -MU, 0, ..., 0) \\ & ... & & \end{array}$$

De proche en proche, Cayley obtient la figure suivante (mais sans décrire explicitement les étapes suivantes) [26] :



Que l'on écrirait aujourd'hui au moyen de la suite exacte :

$$... \longrightarrow \bigoplus C_{r-m-n-p} \xrightarrow{f_3} \bigoplus C_{r-m-n} \xrightarrow{f_2} \bigoplus C_{r-m} \xrightarrow{f_1} C_r$$

où $\mathrm{Im} f_{i+1} = \ker f_i$, ce qu'hélas Cayley, pas plus que Bézout, n'ont pu démontrer. Cayley annonce qu'alors

$$\begin{array}{rcl} N & = & \dim C_r - \dim(\mathrm{Im} f_1) \\ & = & \dim C_r - \sum \dim C_{r-m} \\ & & + \sum \dim C_{r-m-n} - \sum \dim C_{r-m-n-p} \\ & & + ... \end{array}$$

Pour r assez grand, le calcul de cette quantité, faisant intervenir des coefficients binomiaux, donne $N = 0$. Donc l'élimination dialytique est possible, pour un nombre quelconque d'équations.

25. *Cf.* [60], p. 261, « the number N must be diminished by... ».
26. *Cf.* [62], p. 371.

La lignée géométrique des recherches de Cayley montre un décalage par rapport aux travaux de Bézout, que d'ailleurs Cayley ne cite pas. A-t-il eu connaissance des idées de Bézout par l'intermédiaire de l'esquisse qu'en donne Waring dans la seconde édition de ses *Meditationes algebraicae* (que Cayley connaissait) ? Bien que cette esquisse semble aujourd'hui incompréhensible à qui n'a pas lu Bézout, et qu'ici Cayley n'en fasse pas mention, elle aurait pu au moins lui donner l'idée. Du reste, Cayley met en œuvre les mêmes mécanismes d'algèbre linéaire et rencontre la même pierre d'achoppement (ni Bézout, ni Cayley ne démontrent que $\ker f_1 = \mathrm{Im} f_2$). Sa méthode est exactement la méthode de Bézout, informée par la théorie des déterminants, la méthode dialytique et le symbolisme matriciel. La somme alternée N est présente chez Bézout [27], chez Waring [28], et chez Cayley [29].

Une préoccupation calculatoire chez Cayley constitue un deuxième décalage par rapport aux travaux de Bézout : Cayley cherche une formule du résultant. Dans l'article de 1847, il propose la formule suivante :

$$R = \frac{Q_1 Q_3 \ldots}{Q_2 Q_4 \ldots}$$

où pour chaque i, Q_i est un mineur de la matrice de f_i. Le choix des mineurs obéit à la règle suivante. La matrice de f_i a le même nombre de colonnes que la matrice de f_{i+1} compte de lignes. La règle est que l'ensemble des colonnes sélectionnées pour former le mineur Q_i doit correspondre au complémentaire de l'ensemble des lignes sélectionnées pour former le mineur Q_{i+1}. Quant au dernier, soit Q_j, il doit être choisi parmi les mineurs maximaux de la matrice de f_j de sorte qu'aucun des Q_i ne s'annule. Cayley ne donne pas de règle pour ce choix-ci. Plusieurs auteurs (Salmon, Netto) semblent avoir essayé de démontrer la formule de Cayley, jusqu'à ce que Macaulay crut bon d'y renoncer, ayant trouvé une formule plus simple mais « moins générale » de la forme $\frac{Q_1}{\Delta}$ où Q_1 est encore un mineur de la matrice de f_1, avec moins de liberté de choix cependant [30]. En 1926, E. Fischer est parvenu à démontrer rigoureusement la formule de Cayley [31].

Sylvester, Hesse, Cayley montrent une conception de l'élimination différente de celle de leur prédecesseur Bézout. Conception orientée par une préoccupation calculatoire : ils cherchent tous trois des expressions explicites du résultant, dépendant de déterminants. L'homogénéisation, grand avantage par rapport aux travaux de Bézout surtout dans le cas de plus de deux équations, conduit à des expressions plus symétriques. Une seconde préoccupation, symbolique, les conduit à penser l'élimination comme branche d'une discipline plus vaste, une algèbre des formes, en lien avec la théorie des invariants d'une part, et la géométrie algébrique naissante d'autre part. Les travaux de Leibniz préfiguraient une telle préoccupation, que l'on a observée aussi bien dans la méthode dialytique que dans les « figures matricielles » de Cayley (même s'il ne s'agit pas encore de *matrices* au sens *opératoire* du terme). La nouveauté et le manque de rigueur dans les idées de Bézout mettaient en danger la réception de son traité. Son caractère exclusif, aussi : Bézout écrit un traité sur les équations, sans presque jamais se permettre d'utiliser le concept de racine ou de solution d'une équation. Il se barre la route de la géométrie, alors que les travaux d'Euler avait déjà mis en lumière le rapport entre élimination et projection sur les axes dans un système cartésien. Ces obstacles franchis, les articles de Hesse et Cayley assurent une postérité, au moins indirecte, au traité de Bézout.

10.2 La *Théorie générale des équations algébriques* (1779) de Bézout

Dans la dédicace de son traité intitulé *Théorie générale des équations algébriques*, publié en 1779 à Paris, Bézout indique que le but de son ouvrage est « la perfection d'une partie des Sciences

27. *Cf.* [29] au § 62 que nous avons reproduit en annexe A, où Bézout écrit cette somme en termes de différences finies ; la somme alternée apparaît de manière plus explicite aux § 41–43, 60–61.
28. *Cf.* [324], p. 207.
29. *Cf.* [60], p. 261.
30. *Cf.* [216], et aussi [170] pour une étude récente des formules de Macaulay.
31. *Cf.* [250].

Mathématiques, de laquelle toutes les autres attendent ce qui peut aujourd'hui procurer leur avancement » [32]. Dans son introduction, Bézout oppose deux branches des mathématiques de son temps : l'« analyse algébrique finie » et l'« analyse infinitésimale ». La première consiste en la théorie des équations. Elle jouirait selon lui d'une antériorité historique à l'égard de la seconde. L'analyse infinitésimale aurait récemment captivé tous les efforts des mathématiciens, à cause de son caractère plus attrayant (« plus riant »), des possibilités d'applications plus diverses qu'elle offre, et des obstacles rencontrés en analyse algébrique. Bézout ajoute cependant :

> On a vu que celle-ci même [l'analyse infinitésimale] avait besoin que la première fut perfectionnée.
> (...)
> La nécessité de perfectionner cette partie [l'analyse algébrique] n'a pas échappé à ceux à qui l'analyse infinitésimale est le plus redevable. [33]

L'antériorité historique de l'analyse algébrique semble donc se doubler à ses yeux d'une antériorité logique. Cela se confirme par la composition presque exclusivement algébrique de son traité :

- Bézout ne fait qu'une brève allusion, au § 48, à l'interprétation géométrique des méthodes d'élimination comme recherche des intersections de courbes ;
- il ne fait jamais d'hypothèse sur l'existence ou la forme des racines des équations algébriques [34].

Bézout mentionne Euler et Lagrange, « dont les mémoires sur l'analyse algébrique ne renferment ni moins de profondeur, ni moins de sagacité que les autres productions de ces illustres analystes ». Il fait peut-être allusion ici, plus spécialement, aux *Réflexions sur la résolution algébrique des équations* de Lagrange, et au mémoire d'Euler de 1748, *Démonstration sur le nombre des points où deux lignes des ordres quelconques peuvent se rencontrer*, qu'il connaissait [35]. Bézout oppose analyse algébrique finie et analyse infinitésimale ; qui plus est, son ouvrage se présente sous les traits d'une recherche spécialisée ; sa position est donc singulière [36].

Dans la suite de son introduction, il décrit succinctement « l'état où était l'analyse » lorsqu'il a entrepris son travail, puis il expose le plan de son ouvrage. Le « théorème de Bézout », qui détermine le degré de l'équation finale résultant de l'élimination de $(n-1)$ inconnues parmi n, dans un système de n équations, en forme le cœur. Au § 62, Bézout redémontre « tout ce qu'on a su jusqu'ici sur les équations incomplètes à deux inconnues seulement » : dans le cas de deux équations à deux inconnues x_1, x_2, de la forme

$$\sum_{k_1 \le a_1, k_1+k_2 \le t} A_{k_1 k_2} x_1^{k_1} x_2^{k_2} = 0$$
$$\sum_{k_1 \le a'_1, k_1+k_2 \le t'} A'_{k_1 k_2} x_1^{k_1} x_2^{k_2} = 0$$

le degré de l'équation finale résultant de l'élimination de x_2 est $D = tt' - (t-a_1)(t'-a'_1)$. Cramer en 1750, Euler, puis Bézout lui-même en 1764, avaient déjà connu ce résultat. Mais Bézout obtient cette formule comme un cas particulier de la formule plus générale valable pour un nombre quelconque

32. *Cf.* [29]. La dédicace est adressée à Louis XVI.
33. *Cf.* [29], p.ii.
34. Ses raisonnements ne font même jamais, nous semble-t-il, intervenir la notion de racine. Le terme apparaît aux §§ 48, 117, 280–284, mais n'entre pas de manière cruciale dans les démonstrations. Bézout sait bien sûr que les méthodes connues de résolution algébrique des équations ne s'appliquent pas au delà du quatrième degré.
35. Référence explicite y est faite dans [26].
36. Sur l'usage du terme « analyse » pour désigner une ou des disciplines à la fin du XVIII$^{\text{ème}}$ siècle, H. Sinaceur a pu écrire ([291], p. 51) :

> Le terme « analyse » est un concept générique représentant une sinon la méthode mathématique plutôt qu'une de ses branches particulières. Il est alors normal qu'aucune distinction tranchée n'existe entre algèbre et analyse, et donc aucune spécialisation exclusive. Bien plus, l'analyse des équations, appelée aussi « analyse algébrique » peut être considérée comme une partie d'un tout nommé « Analyse mathématique ».

d'« équations incomplètes de première espèce », de la forme :

$$\sum_{\substack{k_1 \leq a_1, k_2 \leq a_2, k_3 \leq a_3, ... \\ k_1 + k_2 + k_3 + ... \leq t}} A_{k_1 k_2 k_3 ...} x_1^{k_1} x_2^{k_2} x_3^{k_3} ... = 0$$

où

$$\begin{cases} \max\{a_i\} \leq t \\ \min\{a_i + a_j\}_{i<j} \geq t \end{cases}$$

Le degré de l'équation finale résultant de l'élimination de $x_2, x_3, ...$ est alors :

$$\begin{aligned} D = tt't''...&-(t-a_1)(t'-a_1')(t''-a_1'')... \\ &-(t-a_2)(t'-a_2')(t''-a_2'')... \\ &-(t-a_3)(t'-a_3')(t''-a_3'')... \end{aligned}$$

Cette formule sera encore citée telle quelle en 1900 dans les *Vorlesungen über Algebra* de Netto[37]. Ainsi, pour trois équations complètes (trois équations génériques, c'est-à-dire, chacune étant générique de son degré), $D = tt't''$.

10.3 Trois conceptions de l'élimination

Le mémoire de 1764 de Bézout reposait déjà sur l'idée suivante : décomposer le processus d'élimination des inconnues en une suite d'opérations, dont chacune consiste à multiplier certaines des équations précédemment obtenues par des polynômes adéquats, puis à faire la somme des produits. Et certes cette idée était déjà ancienne. Pourtant, elle acquiert perfection dans le livre II du traité de 1779, où Bézout propose de représenter directement l'équation finale résultant de l'élimination, comme une telle somme de produits des équations proposées par des polynômes adéquats. Cette représentation peut sembler assez naturelle au lecteur moderne habitué à la notion d'« idéal », héritée du XIX$^{\text{ème}}$ siècle. Nous allons étudier à présent trois étapes différentes qui jalonnent cette conception de l'« équation-somme », tout en analysant la démonstration du théorème de Bézout telle qu'elle figure dans ce traité.

On se limitera dans un premier temps au cas de trois équations à trois inconnues, équations « complètes » de degrés respectifs t, t', t'', que Bézout écrit sous la forme :

$$\begin{aligned} (u...3)^t &= 0 \\ (u...3)^{t'} &= 0 \\ (u...3)^{t''} &= 0 \end{aligned}$$

En introduisant une inconnue supplémentaire pour homogénéiser, on est ramené à trois équations homogènes, que l'on écrirait aujourd'hui :

$$\begin{aligned} (L_1) \quad & f_1(x_1, x_2, x_3) = 0 \\ (L_2) \quad & f_2(x_1, x_2, x_3) = 0 \\ (L_3) \quad & f_3(x_1, x_2, x_3) = 0 \end{aligned}$$

Posons $n_1 = t$, $n_2 = t'$, $n_3 = t''$.

Bézout suppose dès le départ l'existence d'une « équation finale » unique résultant de l'élimination de deux des inconnues x_2 et x_3, de degré D minimal. On peut distinguer *trois conceptions différentes* de l'élimination dans le traité de 1779. Les deux premières consistent à multiplier f_1 par un « polynôme multiplicateur » ϕ_1 à coefficients indéterminés, de degré $(n_0 - n_\alpha)$, où n_0 est un entier suffisamment grand, puis à faire « disparaître » tous les termes monômes du produit $\phi_1 f_1$ au moyen des équations (L_2) et (L_3), sauf les monômes $1, x_1, x_1^2, ..., x_1^D$, qui composent l'équation finale

37. *Cf.* [236], § 419.

recherchée [38]. Bézout se conforme ensuite à la méthode classique « des coefficients indéterminés ». Après utilisation des équations (L_2) et (L_3), et en vue de l'application de la méthode des coefficients indéterminés, il doit donc rester autant de termes à faire « disparaître » (chacun fournissant une équation), que de coefficients indéterminés dans ϕ_1. Mais la situation se complique encore, du fait que seuls certains des coefficients indéterminés fournis par ϕ_1, selon Bézout, peuvent servir à l'élimination, plusieurs d'entre eux étant « inutiles ». Finalement, Bézout aboutit à une formule :

$$\sharp\{\text{termes restant dans } \phi_1 f_1\} - D = \sharp\{\text{coefficients utiles dans } \phi_1\}$$

Il utilise ensuite cette formule pour calculer D. Comme on le voit, cette méthode (il l'expose en détail dans les paragraphes § 59–67 pour des équations incomplètes de première espèce) est assez ambiguë. On ne sait pas, en effet, dans quel ordre procéder pour appliquer effectivement la méthode des coefficients indéterminés et calculer l'équation finale : les termes que l'on fait disparaître dans $\phi_1 f_1$ au moyen de (2) et (3) risquent de réapparaître lors de l'application de la méthode des coefficients indéterminés. Et Bézout n'en dit pas plus pour le moment, puisqu'il réserve le calcul effectif pour le livre II.

Nous avons dit qu'il y avait dans le traité trois conceptions différentes de l'élimination, et nous venons d'exposer la démarche commune aux deux premières. Ce qui les distingue est la manière de compter le nombre de termes monômes « que l'on peut faire disparaître » dans un polynôme donné, au moyen d'équations données. Dans sa première démonstration (pour les équations complètes, aux paragraphes § 45–48), Bézout affirme que :

$$\sharp\{\text{termes que l'on peut faire disparaître dans } \phi_1 f_1\} = \sharp\{\text{termes divisibles par } x_2^{n_2} \text{ ou } x_3^{n_3}\}$$

Cela ressemble à la méthode newtonienne de l'élimination des plus hautes puissances de l'inconnue, généralisée à un nombre quelconque d'équations et d'inconnues. Il est remarquable que Bézout, en 1764, après avoir indiqué que les méthodes d'élimination d'Euler [39] et de Cramer ne permettent de travailler que sur deux équations à la fois, ajoutait que la méthode de Newton présente le même défaut :

> La méthode de Newton n'exige pas, à la vérité, qu'on élimine par la comparaison des équations deux à deux ; mais elle n'en a pas pour cela aucun avantage sur celle d'Euler et Cramer, dans le cas où l'on a plus de deux équations ; au contraire, elle fait encore acquérir à l'équation finale de nouveaux facteurs inutiles. [40]

Toujours est-il que Bézout ne parle pas explicitement, en 1779, de « la méthode de Newton ». Et la manière dont il utilise l'idée de cette méthode est ambiguë : destinée au calcul effectif de l'équation finale, elle est détournée au profit du calcul du *degré* de l'équation finale. Il ne semble pas s'inquiéter des difficultés concernant le calcul effectif de l'équation finale. [41] Il ajoute, à la suite de sa démonstration :

> L'idée de substitution sur laquelle nos raisonnements ont été appuyés, rapproche le plus qu'il est possible, l'exposition de notre marche, des idées élémentaires de l'élimination dans les équations du premier degré [42]. Quoique nous puissions bien encore appliquer la même idée aux équations incomplètes, nous allons cependant présenter les choses sous un autre point de vue, mais généralement applicable, et toujours de la même manière : au lieu que le principe de substitution, si nous nous y attachions, exigerait des modifications, et des attentions particulières. [43]

38. C'est ainsi que Bézout procède par exemple pour les équations incomplètes de première espèce, *cf.* [29], § 62.
39. Quand Bézout, dans [26], mentionne Euler, il renvoie sûrement à la méthode n°2 d'Euler, *cf. infra* 106.
40. *Cf.* [26], p. 290.
41. La difficulté peut être illustrée par un exemple très simple. Supposons $d^o f_2 = d^o f_3 = 2$ et $d^o \phi_1 = 4$ et cherchons à faire disparaître le plus de termes possibles dans le polynôme ϕ_1 au moyen des équations (L_2) et (L_3). En interprétant de manière naïve la méthode newtonienne, commençons par faire disparaître les termes divisibles par x_3^2 dans f_2 et ϕ_1 au moyen de l'équation (L_3). Si l'on cherche à faire disparaître ensuite les termes divisibles par x_2^2 dans ϕ_1 au moyen de l'équation (L_2), des termes divisibles par x_3^2 vont réapparaître.
42. Pour $n_1 = n_2 = n_3 = 1$, la méthode énoncée ci-dessus ressemble au « pivot de Gauss ».
43. *Cf.* [29], § 54.

Dans sa deuxième démonstration, Bézout explique que de faire disparaître des termes dans un polynôme donné, au moyen des équations (L_2) et (L_3), revient à utiliser de nouveaux polynômes multiplicateurs :

> *On demande combien, à l'aide de ces équations, on peut faire disparaître de termes dans un polynôme, sans en introduire de nouveaux.*
> Ne supposons d'abord qu'une équation ; et concevons que l'ayant multipliée par un polynôme (...), on ajoute le produit (...) au polynôme proposé : il est clair
> 1^o que cette addition ne changera rien à la valeur du polynôme proposé.
> 2^o Que supposant au polynôme multiplicateur les qualités nécessaires pour ne pas introduire de nouveaux termes, on pourra faire disparaître dans le polynôme proposé autant de termes qu'en aura le polynôme multiplicateur, puisque chacun de ceux-ci fournira un coefficient (...) [44]

Ainsi, pour faire disparaître des termes dans $\phi_1 f_1$ au moyen de l'équation (L_2), Bézout utilise donc un polynôme multiplicateur ϕ_2, de degré $n_0 - n_2$, et considère la somme $\phi_1 f_1 + \phi_2 f_2$. Pour faire ensuite disparaître des termes au moyen de l'équation (L_3), il considère de même une somme $\phi_1 f_1 + \phi_3 f_3$. Le recours à la méthode des coefficients indéterminés laisse encore peser l'ambiguïté sur l'ordre dans lequel procéder au calcul effectif.

Enfin, il nous reste à exposer la troisième conception de l'élimination présente dans le traité (c'est cette conception que nous retiendrons pour exposer en détail le calcul du degré de l'équation finale dans la section suivante). Bézout l'expose seulement au livre II. Toutes les étapes deviennent, en quelque sorte, solidaires :

> Dorénavant nous considérerons l'élimination d'une manière qui ne diffère de celle-là qu'en apparence, et qui est la même quant au fond.
> Nous concevrons qu'on multiplie chacune des équations données, par un polynôme particulier, et qu'on ajoute tous ces produits. Le résultat sera ce que nous appellerons l'*Equation-somme*, laquelle deviendra l'équation finale par l'anéantissement de tous les termes affectés des inconnues qu'il s'agit d'éliminer.
> Il s'agit donc actuellement 1^o de fixer la forme que doit avoir chacun de ces polynômes-multiplicateurs. 2^o De déterminer le nombre des coefficients qui, dans chacun, ne peuvent être considérés comme utiles à l'élimination (...) [45]

Le calcul du degré de l'équation finale revient alors à :
- multiplier chaque équation $f_\alpha = 0$ par un polynôme multiplicateur ϕ_α à coefficients indéterminés, de degré $n_0 - n_\alpha$;
- faire l'« équation-somme » ;
- demander que tous les termes s'annulent sauf 1, x_1, x_1^2,..., x_1^D.

Le nombre de coefficients indéterminés est alors $\sum \dim C_{n_0 - n_\alpha}$, et le nombre d'équations fournies par l'annulation des termes de l'équation-somme est $\dim C_{n_0} - D$. La méthode des coefficients indéterminés impose que

$$\sharp\{\text{équations}\} = \sharp\{\text{coefficients indéterminés}\} - \sharp\{\text{coefficients inutiles}\}$$

c'est-à-dire :

$$\dim C_{n_0} - D = \sum \dim C_{n_0 - n_\alpha} - \sharp\{\text{coefficients inutiles}\}$$

10.4 La démonstration de Bézout

Nous allons à présent décrire la démonstration du « théorème de Bézout », conformément à cette « troisième conception de l'élimination » ainsi nommée par nous dans la section précédente. Mais ce faisant, nous adopterons un langage plus moderne que le sien. Nous nous limiterons dans un premier temps au cas de n équations « complètes », c'est-à-dire des équations de degrés donnés (non

44. *Cf.* [29], § 60.
45. *Cf.* [29], § 224.

homogènes), dont tous les coefficients sont algébriquement indépendants. Nous les écrirons sous la forme :

$$\begin{array}{ll}(L_1) & f_1(x_1, x_2, ..., x_n) = 0\\ (L_2) & f_2(x_1, x_2, ..., x_n) = 0\\ \vdots & \\ (L_n) & f_n(x_1, x_2, ..., x_n) = 0\end{array}$$

Posons $t_\alpha = \deg f_\alpha$, et $C = K[x_1, x_2, ..., x_n]$, où $K = \mathbb{Q}\left((a_{\alpha\underline{k}})_{\alpha,\underline{k}}\right)$ est un corps contenant les coefficients (algébriquement indépendants) des équations ci-dessus :

$$(1 \leq \alpha \leq n) \qquad f_\alpha = \sum_{|\underline{k}| \leq t_\alpha} a_{\alpha\underline{k}} x^{\underline{k}}$$

Les notations soulignées désigneront désormais des multi-indices :

$$x^{\underline{k}} = x_1^{k_1} x_2^{k_2} ... x_n^{k_n}, \quad |\underline{k}| = k_1 + k_2 + ... + k_n$$

Nous noterons $C_{\leq t_\alpha}$ le K-espace vectoriel des polynômes de degré inférieur à t_α. Alors $f_\alpha \in C_{\leq t_\alpha}$. Bézout suppose dès le départ l'existence d'une « équation finale » unique, résultant de l'élimination de $(n-1)$ des inconnues, de degré D minimal, et représentable sous la forme d'une « équation-somme »

$$\sum_{\alpha=1}^{n} \phi_\alpha f_\alpha = 0$$

avec des « polynômes-multiplicateurs » ϕ_α bien choisis. Bézout applique la « méthode des coefficients indéterminés ». Il suppose indéterminés les coefficients des polynômes ϕ_α, et leurs degrés $\deg \phi_\alpha = t_0 - t_\alpha$, où l'entier t_0 est lui aussi indéterminé :

$$(1 \leq \alpha \leq n) \qquad \phi_\alpha = \sum_{|\underline{l}| \leq t_0 - t_\alpha} u_{\alpha\underline{l}} x^{\underline{l}}$$

Les coefficients des termes monômes, dans l'« équation-somme » ci-dessus, sont alors de la forme

$$\text{coefficient de } x^{\underline{h}} = \sum_{\substack{\alpha, \underline{k}, \underline{l} \\ \underline{k} + \underline{l} = \underline{h}}} a_{\alpha\underline{k}} u_{\alpha\underline{l}}$$

Ces expressions sont linéaires en les coefficients de ϕ_α. Si l'on impose que tous les termes soient nuls dans $\sum_\alpha f_\alpha \phi_\alpha$ sauf les termes $x_1^D, x_1^{D-1}, ..., x_1$ et le terme constant, et que l'un au moins d'entre ceux-ci soit non nul, par exemple que le coefficient de x_1^D soit égal à 1, on obtient un système de $(\dim C_{\leq t_0} - D)$ équations linéaires de la forme :

$$\sum_{\substack{\alpha, \underline{k}, \underline{l} \\ \underline{k} + \underline{l} = \underline{h}}} a_{\alpha\underline{k}} u_{\alpha\underline{l}} = c_{\underline{h}}$$

où $c_{\underline{h}}$ vaut 0 ou 1. Le problème est alors de trouver une valeur de D minimale telle que ce système ait une solution. Il s'agit d'un problème d'algèbre linéaire.

Les inconnues du système sont les $u_{\alpha\underline{l}}$, au nombre de $\sum_\alpha \dim C_{\leq t_0 - t_\alpha}$. Mais, nous dit Bézout, il y a, parmi ces inconnues, un certain nombre de « coefficients inutiles ». C'est en fait, suivant une autre expression de Bézout, le nombre de coefficients « dont on dispose arbitrairement » dans la solution générale du système linéaire, ou encore, en termes modernes, le nombre de paramètres de la solution, c'est-à-dire la dimension de l'espace vectoriel des solutions. Il suffit de la calculer pour le système linéaire homogène associé. Comme ce système homogène consiste à demander que tous les termes s'annulent dans $\phi_1 f_1 + \phi_2 f_2 + ... + \phi_n f_n$, excepté les termes en x_1^{D-1}, x_1^{D-2},...,x_1 et le

terme constant (cette fois, le terme en x_1^D doit s'annuler aussi), mais que l'« équation finale » est supposée de degré minimal D, alors les solutions de ce système homogène donneront les n-uplets $(\phi_1, \phi_2, ..., \phi_n)$ tels que $\sum \phi_\alpha f_\alpha$ soit identiquement nul. On a donc la formule suivante :

$$\begin{array}{l} \text{« nombre de coefficients inutiles »} \\ = \dim\{(\phi_1, \phi_2, ..., \phi_n) \in \bigoplus C_{\leq t_0 - t_\alpha} \,|\, \sum \phi_\alpha f_\alpha = 0\} \end{array}$$

Ce nombre représente la dimension du noyau d'un morphisme d'espaces vectoriels $(f_1, f_2, ..., f_n)_{\leq t_0}$, que nous noterons pour abréger $f_{\leq t_0}$, et qui est défini de la manière suivante :

$$\begin{array}{ccccccc} 0 \to \ker f_{\leq t_0} \to & \bigoplus C_{\leq t_0 - t_\alpha} & \xrightarrow{f_{\leq t_0}} & C_{\leq t_0} & \to \operatorname{coker} f_{\leq t_0} \to 0 \\ & (\phi_1, \phi_2, ..., \phi_n) & \longmapsto & \sum \phi_\alpha f_\alpha & \end{array}$$

On remarque que t_0 reste inconnu pour le moment.

Bézout utilise la formule suivante :

$$\text{nombre d'équations} = \text{nombre d'inconnues} - \text{nombre de paramètres de la solution}$$

qui n'est vraie que sous la condition que le système linéaire soit libre. Bien sûr, il ne vérifie pas cette condition. On verra que Carl Schmidt (1886) l'a vérifiée pour les équations complètes. Sous cette condition, on a donc :

$$D = \dim C_{\leq t_0} - \sum_\alpha \dim C_{\leq t_0 - t_\alpha} + \dim(\ker f_{\leq t_0})$$

Il reste à espérer que l'indéterminée t_0 disparaisse dans le calcul du membre de droite. Enfin, même si le système n'était pas libre, on pourrait au moins écrire, en vertu de la suite exacte ci-dessus :

$$(1) \quad \dim(\operatorname{coker} f_{\leq t_0}) = \dim C_{\leq t_0} - \sum_\alpha \dim C_{\leq t_0 - t_\alpha} + \dim(\ker f_{\leq t_0})$$

Comme Bézout ne s'occupe jamais de démontrer que le système d'équations linéaires est libre, on peut dire que ce qu'il calcule en fait [46] est $\dim(\operatorname{coker} f_{\leq t_0})$.

Le décompte des « coefficients inutiles », c'est-à-dire le calcul de $\dim(\ker f_{\leq t_0})$, est décrit par Bézout au § 233 dans les termes suivants :

> Si l'on se rappelle ce que nous avons dit dans le premier Livre, on verra facilement, que le nombre des coefficients utiles, dans le premier polynôme-multiplicateur des équations entre lesquelles il s'agit d'éliminer, sera toujours égal au nombre des coefficients de ce polynôme, moins le nombre des termes qu'on peut faire disparaître dans ce polynôme, à l'aide des $n - 1$ autres équations, n étant le nombre total des équations :
> Que le nombre des coefficients utiles du second polynôme-multiplicateur, sera le nombre total des coefficients ou des termes de ce polynôme, moins le nombre de termes qu'on peut faire disparaître dans ce polynôme, à l'aide des $n - 2$ dernières équations :
> Que le nombre des coefficients utiles dans le troisième polynôme-multiplicateur, sera le nombre des termes de ce polynôme, moins le nombre des termes qu'on peut faire disparaître dans ce polynôme, à l'aide des $n - 3$ autres équations ; et ainsi de suite jusqu'au dernier, dont le nombre des coefficients utiles sera précisément égal au nombre de ses termes. [47]

On voit ici se profiler un argument de récurrence : pour calculer le nombre de coefficients que l'on peut faire disparaître dans ϕ_1 avec les équations (L_2) à (L_n), il faut en effet utiliser de nouveaux

46. On doit avoir de toute façon $D \leq \dim C_{\leq t_0}/\mathrm{Im} f_{\leq t_0}$ puisque si D est le degré de l'équation finale, alors les monômes x_1^{D-1},...,x_1 et 1 sont linéairement indépendants dans $C/(f_1, f_2, ..., f_n)$.

47. *Cf.* [29], § 233.

polynômes-multiplicateurs. Intuitivement, on obtiendrait, en suivant Bézout :

$$(2)\quad \left\{\begin{array}{rcl}\dim(\ker f_{\leq t_0}) & = & \text{« nombre de coefficients inutiles »}\\ & = & \text{« nombre de coef. que l'on peut faire disparaître dans } \phi_1 \text{ avec les équ. } (L_2) \text{ à } (L_3) \text{ »}\\ & & +\text{« nombre de coef. que l'on peut faire disparaître dans } \phi_2 \text{ avec les équ. } (L_3) \text{ à } (L_n) \text{ »}\\ & & +...\\ & & +\text{« nombre de coef. que l'on peut faire disparaître dans } \phi_{n-1} \text{ avec l'équ. } (L_n) \text{ »}\\ & = & \dim\left(\mathrm{Im}(f_2,...,f_n)_{\leq t_0-t_1}\right)\\ & & +\dim\left(\mathrm{Im}(f_3,...,f_n)_{\leq t_0-t_2}\right)\\ & & +...\\ & & +\dim\left(\mathrm{Im}(f_n)_{\leq t_0-t_{n-1}}\right)\end{array}\right.$$

Mais comment justifier, de manière rigoureuse, cet argument ? Il y a là une grosse lacune dans la démonstration de Bézout. Nous allons démontrer (2) par récurrence sur le nombre de polynômes f_α. Autrement dit, supposons que pour un certain $r < n$ et $f_{\leq t_0} = (f_1, ..., f_r)_{\leq t_0}$ défini comme ci-dessus, on ait :

$$\begin{array}{rcl}\dim(\ker f_{\leq t_0}) & = & \dim\left(\mathrm{Im}(f_2,...,f_r)_{\leq t_0-t_1}\right)\\ & & +\dim\left(\mathrm{Im}(f_3,...,f_r)_{\leq t_0-t_2}\right)\\ & & +...\\ & & +\dim\left(\mathrm{Im}(f_r)_{\leq t_0-t_{r-1}}\right)\end{array}$$

Soit à présent $f_{\leq t_0} = (f_1, f_2, ..., f_{r+1})_{\leq t_0}$. L'isomorphisme d'espaces vectoriels $E/E\cap F \simeq (E+F)/F$ (pour $E \subset F$), appliqué aux espaces vectoriels :

$$\left(\ker f_{\leq t_0} \cap \bigoplus_{2\leq\alpha\leq r+1} C_{\leq t_0-t_\alpha}\right) \subset \ker f_{\leq t_0}$$

permet d'écrire :

$$\begin{array}{rcl}\dim(\ker f_{\leq t_0}) & = & \dim\left(\frac{\ker f_{\leq t_0}+\bigoplus_{\alpha\geq 2} C_{\leq t_0-t_\alpha}}{\bigoplus_{\alpha\geq 2} C_{\leq t_0-t_\alpha}}\right) + \dim\left(\ker f_{\leq t_0} \cap \bigoplus_{\alpha\geq 2} C_{\leq t_0-t_\alpha}\right)\\ & = & \dim E_{r+1} + \dim\left(\ker(f_2,...,f_{r+1})_{\leq t_0}\right)\end{array}$$

où (pour tout $i \leq n$)

$$E_i = \left\{\phi_1 \in C_{\leq t_0-t_1} | \exists \phi_2 \in C_{\leq t_0-t_2}, ..., \phi_i \in C_{\leq t_0-t_i},\ \sum_{\alpha=1}^{i} \phi_\alpha f_\alpha = 0\right\}$$

Or, par récurrence, on a :

$$\dim\left(\ker(f_2,...,f_{r+1})_{\leq t_0}\right) = \dim(\mathrm{Im}(f_3,...,f_{r+1})_{\leq t_0-t_2}) + ... + \dim(\mathrm{Im}(f_{r+1})_{\leq t_0-t_r})$$

Pour achever la démonstration par récurrence, il reste donc seulement à démontrer que :

$$E_{r+1} = \mathrm{Im}(f_2,...,f_{r+1})_{\leq t_0-t_1}$$

Nous démontrerons cette assertion ultérieurement. Admettons-la pour l'instant, et donc, avec elle, l'assertion (2). Reprenons la démonstration du théorème de Bézout et introduisons à cet effet une notation pour les différences finies[48] d'un polynôme donné $P(T)$:

$$\nabla_t P(T) = P(T) - P(T+t)$$

48. Pour les différences finies, Bézout emploie une autre notation que l'on pourrait définir ainsi, par récurrence :

$$d^r P(t) \cdots \binom{t}{-t_1,...,-t_r} = d^{r-1}P(t) \cdots \binom{t}{-t_2,...,-t_r} - d^{r-1}P(t-t_1) \cdots \binom{t}{-t_2,...,-t_r}$$

Autrement dit,

$$d^r P(t) \cdots \binom{t}{-t_1,...,-t_r} = \nabla_{-t_1} ... \nabla_{-t_r} P(t)$$

De (1) et (2), on déduit en regroupant certains termes :

$$\begin{aligned}
\dim(\operatorname{coker} f_{\leq t_0}) &= \Big(\dim C_{\leq t_0} - \sum_{\alpha=2}^{n} \dim C_{\leq t_0 - t_\alpha} + \dim\left(\operatorname{Im}(f_3, ..., f_n)_{\leq t_0 - t_2}\right) + ... + \dim\left(\operatorname{Im}(f_n)_{\leq t_0 - t_{n-1}}\right)\Big) \\
&\quad - \left(\dim C_{\leq t_0 - t_1} - \dim\left(\operatorname{Im}(f_2, ..., f_n)_{\leq t_0 - t_1}\right)\right) \\
&= \dim\left(\operatorname{coker}(f_2, ..., f_n)_{\leq t_0}\right) - \dim\left(\operatorname{coker}(f_2, ..., f_n)_{\leq t_0 - t_1}\right) \\
&= \nabla_{-t_1} \dim\left(\operatorname{coker}(f_2, ..., f_n)_{\leq t_0}\right) \\
&= \nabla_{-t_1} \nabla_{-t_2} \dim\left(\operatorname{coker}(f_3, ..., f_n)_{\leq t_0}\right) \\
&= ... \\
&= \nabla_{-t_1} \nabla_{-t_2} ... \nabla_{-t_n} \dim C_{\leq t_0}
\end{aligned}$$

Insistons en particulier sur la troisième de ces égalités :

$$(3) \quad \dim\left(\operatorname{coker}(f_1, ..., f_n)_{\leq t_0}\right) = \nabla_{-t_1} \dim\left(\operatorname{coker}(f_2, ..., f_n)_{\leq t_0}\right)$$

Cette formule de récurrence est le cœur du raisonnement de Bézout, traduit en notations modernes. Comme $\dim C_{\leq t_0}$ est un polynôme de degré n en t_0, après n applications successives de l'opérateur ∇, on obtient bien une constante indépendante de t_0. Plus précisément,

$$\dim C_{\leq t_0} = \binom{t_0 + n}{n}$$

et on trouve finalement, pour les équations complètes :

$$\dim\left(\operatorname{coker}(f_1, ..., f_n)_{\leq t_0}\right) = \prod_{\alpha=1}^{n} t_\alpha$$

q. e. d.

Bézout avait d'autre part remarqué que les n applications successives de ∇ peuvent encore s'écrire sous forme d'une somme alternée :

$$\nabla_{-t_1} \nabla_{-t_2} ... \nabla_{-t_n} \dim C_{\leq t_0} = \dim C_{\leq t_0} - \dim \bigoplus_{\alpha} C_{\leq t_0 - t_\alpha} + \dim \bigoplus_{\alpha < \beta} C_{\leq t_0 - t_\alpha - t_\beta} - ...$$

10.5 Des termes d'introduction fictive

Nous avons montré dans la section précédente que la lacune la plus notable dans la théorie de Bézout est l'absence de démonstration de l'assertion ($r \leq n$) :

$$(4) \quad E_r = \operatorname{Im}(f_2, ..., f_r)_{\leq t_0 - t_1}$$

Cette assertion ne se démontre pas aisément. C'est la rigueur du formalisme moderne de l'algèbre linéaire qui nous a permis de percevoir ce manque dans la théorie de Bézout. Avant de montrer comment des techniques postérieures à l'œuvre de Bézout peuvent aussi combler ce manque, remarquons que Bézout avait néanmoins effleuré le problème. Le raisonnement par lequel il dut se convaincre de la légitimité de son intuition transparaît aux § 107 à 118. On démontre facilement les inclusions suivantes (l'inclusion de gauche étant évidente) :

$$(5) \quad \operatorname{Im}(f_2, ..., f_r)_{\leq t_0 - t_1} \subset E_r \subset (f_2, ..., f_r) \cap C_{\leq t_0 - t_1}$$

Autrement dit, si, pour tout $\alpha \leq r$, $\phi_\alpha \in C_{\leq t_0 - t_\alpha}$ et $\sum \phi_\alpha f_\alpha = 0$, alors il existe des polynômes $\phi'_\alpha \in C$ pour $\alpha \geq 2$, tels que $\phi_1 = \sum_{\alpha=2}^{r} \phi'_\alpha f_\alpha$. *Démonstration*[49]. Le corps K est de la forme

49. Démonstration inspirée de Mertens [228], p. 528–529. On peut aussi raccourcir cette démonstration en remarquant que $C/(f_2, ..., f_r)$ est isomorphe à $\mathbb{Q}(a_{1,0}, (a_{\alpha \underline{k}})_{\underline{k} \neq 0})[x_1, ..., x_n]_{(f_2 - a_{2,0}) ... (f_r - a_{r,0})}$. Cet anneau, localisé d'un anneau de polynômes, est intègre.

$\mathbb{Q}((a_{\alpha\underline{k}})_{\alpha,\underline{k}})$ où $(a_{\alpha\underline{k}})_{\underline{k}}$ désignent les coefficients du polynôme f_α. On notera $a_{\alpha,0}$ le terme constant de f_α. On peut toujours, en réduisant au même dénominateur, se ramener au cas où les polynômes ϕ_α sont à coefficients dans $\mathbb{Q}[(a_{\alpha\underline{k}})_{\alpha,\underline{k}}]$. Alors on a un isomorphisme :

$$\begin{array}{ccccc} \mathbb{Q}[(a_{\alpha\underline{k}})_{\alpha,\underline{k}}][x_1,...,x_n]/(f_2,...,f_r) & \simeq & \mathbb{Q}[(a_{\alpha\underline{k}})_{\underline{k}\neq 0}][x_1,...,x_n] & & \\ a_{\alpha 0} & \mapsto & a_{\alpha 0}-f_\alpha & & (2\leq\alpha\leq r) \end{array}$$

L'anneau de droite est un anneau de polynômes sur un corps, et il est donc intègre : $\phi_1 f_1 \in (f_2,...,f_r)$ implique que $\phi_1 \in (f_1,...,f_r)$ *q. e. d.*

Au paragraphe § 107, Bézout expose la notion d'« introduction fictive » sur l'exemple suivant ($n=r=3$, $t_0-t_1=3$) :

$$\left\{\begin{array}{l} f_2 = a'x_1^2+b'x_1x_2+c'x_1x_3+d'x_1+e'x_2+f'x_3+g' \\ f_3 = a''x_1^2+b''x_1x_2+c''x_1x_3+d''x_1+e''x_2+f''x_3+g'' \end{array}\right.$$

(attention, ces deux polynômes ne sont pas génériques, puisqu'il manque les termes en x_2^2, x_2x_3 et x_3^2). Il remarque que $\dim(\mathrm{Im}(f_2,f_3)_{\leq 3})=8$, tandis que $\dim(\mathrm{Im}(f_2,f_3)_{\leq 4}\cap C_{\leq 3})=9$. Dans le langage de Bézout, il s'agit de « faire disparaître des termes » dans un polynôme-multiplicateur $\phi_1\in C_{\leq 3}$, au moyen des équations $f_2=0$ et $f_3=0$. Dans la partie du livre précédent le paragraphe § 107, il supposait que le nombre de termes que l'on peut ainsi « faire disparaître » est $\dim(\mathrm{Im}(f_2,f_3)_{\leq t_0-t_1})$. Mais au § 107, il avoue qu'il faut parfois « introduire fictivement » des termes de degré supérieur à (t_0-t_1) (dans l'exemple ci-dessus, des termes de degré 4) pour « faire disparaître » le plus grand nombre possible de termes de ϕ_1.

Le décompte des « coefficients inutiles » décrit dans notre section précédente, dont nous avions observé le caractère trop peu rigoureux, fait donc à présent l'objet d'une analyse plus précise : Bézout a compris qu'il devait montrer que les inclusions dans (5) sont en fait une égalité. Il entreprend pour cela de montrer que, pour tout $k\geq 0$,

$$(6)\quad \dim(\mathrm{Im}(f_2,...,f_r)_{\leq t_0-t_1})-\dim(\mathrm{Im}(f_2,...,f_r)_{\leq t_0-t_1+k}\cap C_{\leq t_0-t_1})\geq 0$$

Le terme de droite de cette différence est le « nombre de termes que l'on peut faire disparaître dans ϕ_1 » en en introduisant « fictivement » de nouveaux en degré $\leq k$. Mais il ne parvient pas à calculer ce nombre de manière satisfaisante. Son raisonnement peu clair se traduirait peut-être ainsi (§§ 107, 108 et 113) :

$$\begin{array}{ll} (7) & \dim(\mathrm{Im}(f_2,...,f_r)_{\leq t_0-t_1+k}\cap C_{\leq t_0-t_1}) \\ & =\dim(\mathrm{Im}(f_2,...,f_r)_{\leq t_0-t_1+k})-(\dim C_{\leq t_0-t_1+k}-\dim C_{\leq t_0-t_1}) \end{array}$$

Alors (6) s'écrirait ainsi :

$$[\dim C_{\leq t_0-t_1+k}-\dim(\mathrm{Im}(f_2,...,f_r)_{\leq t_0-t_1+k})]-[\dim C_{\leq t_0-t_1}-\dim(\mathrm{Im}(f_2,...,f_r)_{\leq t_0-t_1})]\geq 0$$

ou encore :

$$\dim(C_{\leq t_0-t_1+k}/\mathrm{Im}(f_2,...,f_r)_{\leq t_0-t_1+k})-\dim(C_{\leq t_0-t_1}/\mathrm{Im}(f_2,...,f_r)_{\leq t_0-t_1})\geq 0$$

c'est-à-dire :

$$(8)\quad \nabla_k\mathrm{coker}(f_2,...,f_r)_{\leq t_0-t_1}\leq 0$$

Autrement dit, Bézout pense pouvoir ramener la question à celle de la croissance du polynôme de Hilbert de l'idéal $(f_2,...,f_r)$. Il suffirait de vérifier cette inégalité par le calcul. Hélas, le raisonnement de Bézout, une fois traduit formellement, perd son apparence : la formule (7) est fausse. Quant à la formule (8), bien qu'insuffisante pour la démonstration, elle est néanmoins exacte aussitôt démontré que (5) est une égalité, car on a alors :

$$\frac{C_{\leq t_0-t_1}}{\mathrm{Im}(f_2,...,f_n)_{\leq t_0-t_1}}\quad\xrightarrow{\sim}\quad\frac{C_{\leq t_0-t_1}}{\left(\mathrm{Im}(f_2,...,f_n)_{\leq t_0-t_1+k}\cap C_{\leq t_0-t_1}\right)}\quad\hookrightarrow\quad\frac{C_{\leq t_0-t_1+k}}{\mathrm{Im}(f_2,...,f_n)_{\leq t_0-t_1+k}}$$

Démonstration de la formule (6) (inspirée de Hurwitz [159], prop. 11, p. 122) dans le cas de trois équations génériques ($n = r = 3$). Soit $k > 0$ (pour $k = 0$, il n'y a rien à démontrer), et $\phi_1 = \phi_2 f_2 + \phi_3 f_3$, avec $\phi_2 \in C_{\leq t_0 - t_1 - t_2 + k}$ et $\phi_3 \in C_{\leq t_0 - t_1 - t_3 + k}$, tel que

$$\phi_1 \in \mathrm{Im}(f_2, f_3)_{\leq t_0 - t_1 + k} \cap C_{\leq t_0 - t_1}$$

On va montrer qu'alors

$$\phi_1 \in \mathrm{Im}(f_2, f_3)_{\leq t_0 - t_1 + k - 1} \cap C_{\leq t_0 - t_1}$$

Pour tout polynôme P, notons $[P]_t$ sa composante de degré t. On a donc :

$$0 = [\phi_2]_{t_0 - t_1 - t_2 + k}[f_2]_{t_2} + [\phi_3]_{t_0 - t_1 - t_3 + k}[f_3]_{t_3}$$

D'où :

$$[\phi_2]_{t_0 - t_1 - t_2 + k} = \Lambda[f_3]_{t_3}, \quad [\phi_3]_{t_0 - t_1 - t_3 + k} = -\Lambda[f_2]_{t_2}$$

pour un certain Λ homogène de degré $t_0 - t_1 - t_2 - t_3 + k$. Alors :

$$\phi_1 = \phi_2 f_2 + \phi_3 f_3 = (\Lambda f_3 + r_2) f_2 + (-\Lambda f_2 + r_3) f_3 = r_2 f_2 + r_3 f_3$$

où $r_i \in C_{\leq t_0 - t_1 - t_i + k - 1}$. *q. e. d.*

Avec la démonstration de (6) s'achève pour le moment notre justification des intuitions de Bézout.

10.6 Le complexe de Koszul

Nous avons parcouru toute la démonstration de Bézout, et rendu perceptible une lacune dans son raisonnement. Nous avons vu dans la section précédente comment compléter cette démonstration en comblant cette lacune. Nous allons à présent montrer comment cette lacune a pu être évitée par d'autres mathématiciens, et aussi comment certains outils modernes font écho aux méthodes de Bézout. La propriété (4) est à rapprocher de la notion moderne de « suite régulière » en algèbre commutative [50]. Une suite de polynômes $(f_1, ..., f_n)$ est dite régulière si pour aucun $r \leq n$, f_r n'est diviseur de zéro modulo $(f_1, ..., f_{r-1})$. Démontrer la propriété (4) pour une suite de polynômes génériques (f_1, f_2, f_3) revient en fait à démontrer que la suite des polynômes homogénéisés (f_1^h, f_2^h, f_3^h) est régulière.

Notons $C^h = K[X_0, X_1, X_2, X_3]$ avec $x_1 = \frac{X_1}{X_0}$, $x_2 = \frac{X_2}{X_0}$, $x_3 = \frac{X_3}{X_0}$. Remarquons d'abord que, dans l'hypothèse où la suite (f_1^h, f_2^h, f_3^h) est régulière, on peut construire, pour tout t_0, une résolution de $C^h_{t_0}/(f_1^h, f_2^h, f_3^h)$, intitulée « complexe de Koszul » :

$$(9) \quad 0 \to C^h_{t_0 - t_1 - t_2 - t_3} \to \bigoplus C^h_{t_0 - t_\alpha - t_\beta} \to \bigoplus C^h_{t_0 - t_\alpha} \to C^h_{t_0} \to C^h_{t_0}/(f_1^h, f_2^h, f_3^h) \to 0$$

Alors on trouve immédiatement :

$$(10) \quad \dim C^h_{t_0}/(f_1^h, f_2^h, f_3^h) = \dim C^h_{t_0} - \sum \dim C^h_{t_0 - t_\alpha} + \sum \dim C^h_{t_0 - t_\alpha - t_\beta} - \dim C^h_{t_0 - t_1 - t_2 - t_3}$$

Cette formule coïncide avec la formule aux différences finies de Bézout, obtenue dans la section précédente [51]. C'est la voie qu'empruntera Cayley en 1847–1848.

Dans les deux premières éditions de son *Cours d'algèbre supérieure* (1849, 1854), Serret ne parle pas de la théorie de Bézout, et semble y préférer celle de Poisson. Cette lacune lui sera reprochée, ainsi que plusieurs autres concernant la théorie de l'élimination, dans une critique parue aux *Nouvelles Annales de Mathématiques* :

50. *Cf.* [285], p. 59, où Serre utilise le terme *M-sequence* pour désigner une telle suite régulière. [140], II.8, p. 184 utilise plutôt le terme *regular sequence*.

51. Il est en fait équivalent (*cf.* [285], p. 54, proposition IV.A.3, que l'on peut facilement adapter au cas d'un anneau gradué) de dire que la suite $(f_1^h, ..., f_n^h) \subset C^h$ est régulière et de dire que le complexe de Koszul associé, noté $K_\bullet(f_1^h, ..., f_n^h; C^h)$, est exact, ou même simplement que $H_1 K_\bullet(f_1^h, ..., f_n^h; C^h) = 0$.

> Il est à regretter que M. Serret ne se soit pas enquis de la méthode d'élimination de Bézout, dite du *polynôme-multiplicateur*, qui me semble la plus générale, la plus exacte, la plus conforme à l'état de la question... [52]

Serret se rattrape dans la troisième édition [288] (1866) qui diffère considérablement des deux précédentes. Il y expose la théorie de Bézout mais procède en fait différemment de Bézout dans sa démonstration : il utilise en effet la connaissance préalable du degré de l'équation finale cherchée. La démonstration de Serret fut critiquée et rendue rigoureuse par Carl Schmidt en 1886, et présentée au tournant du siècle sous le nom de Bézout, par exemple par Netto dans ses *Vorlesungen über Algebra* (1900, tome 2). Expliquons-là rapidement dans le cas $n = 3$, en nous basant sur l'exposé de Netto. Il se contente de vérifier que pour $D = t_0 = t_1t_2t_3$, le système de $(\dim C_{\leq t_0} - t_0)$ équations linéaires consistant en l'annulation des termes monômes (*cf.* page 153 ci-dessus) a bien une solution. Il lui suffit pour cela de vérifier que le nombre d'équations est inférieur au nombre d'inconnues, c'est-à-dire

$$\dim C_{\leq t_0} - t_0 \leq \sum_{\alpha} \dim C_{\leq t_0 - t_\alpha}$$

et que le système est libre, ce qu'il démontre par « spécialisation », au moyen de l'exemple trivial [53] :

$$\begin{cases} x_1^{t_1} - a = 0 \\ x_2^{t_2} - x_1 = 0 \\ x_3^{t_3} - x_2 = 0 \end{cases}$$

Il parvient ainsi à montrer que le système linéaire a une solution, et donc qu'il est possible d'obtenir une équation finale de degré $t_1t_2t_3$, mais rien ne dit que cette équation finale est unique, ni donc qu'elle est de degré minimal. On est seulement certain que cette équation est un multiple de la véritable équation finale, appelée par Netto « éliminante ». Il reste donc encore à minorer le degré de l'éliminante, ou bien à montrer que l'équation qu'on a obtenue est irréductible. Netto utilise de nouveau à cet effet la spécialisation, sur le même exemple que précédemment, qui donne effectivement une éliminante $x_3^{t_1t_2t_3} - a = 0$, irréductible [54]. Ainsi, la démonstration de Netto est plus simple que celle de Bézout, mais il s'agit en fait d'une démonstration *a posteriori*. La proximité conceptuelle avec la notion de suite régulière y est absente.

Le problème des « termes d'introduction fictive » a amené Hurwitz en 1913 à introduire la notion de « *Trägheitsform* », en s'inspirant des travaux de Mertens. Dans ce paragraphe et le suivant, nous noterons $f_1, ..., f_r$ des polynômes *homogènes* de degrés respectifs $t_1, ..., t_r$ dans un anneau $C = K[a, b, ...][X_0, ..., X_n]$ où $a, b, ...$ sont des indéterminées. Soit $\mathfrak{M} = (X_0, ..., X_n)$. Un élément f de C est dit « *Trägheitsform* » s'il existe un entier $k \geq 0$ tel que $\mathfrak{M}^k f \subset (f_1, ..., f_r)$. Pour tout k, notons $a_{\alpha k}$ le coefficient du terme en $X_k^{t_\alpha}$ de f_α. Notons pour tout k, comme Hurwitz, $[f]_k$

52. *Cf. Nouvelles Annales de Mathématiques*, t. XIII p. 349–358 et 381–384, t. XIV p. 272–281 et 402–413 ; ici t. XIII p. 356. Il faut rapprocher une telle critique de la citation anonyme suivante (peut-être du même auteur ?) :

> Pour opérer l'élimination, Bézout a le premier indiqué la méthode des polynômes multiplicateurs, (...). Toutefois on n'en parle nulle part. C'est qu'on ne lit pas l'ouvrage où cette méthode est consignée. Auteur d'ouvrages élémentaires, modèles jamais égalés en clarté et en élégance, Bézout a négligé ces deux qualités dans son ouvrage fondamental, dans sa *Théorie générale des équations* (1779), chef d'œuvre d'analyse, dont Jacobi faisait un très grand cas, et qui renferme toutes les prétendues nouveautés en fait d'élimination (*cf.* Nouv. Ann. Math. XVIII, 1859, p. 120–122).

Ces indices suggèrent que le traité de Bézout, reconnu comme un ouvrage important, mais trop difficile à lire, n'a pu avoir qu'une influence rare. Voir aussi notre citation de Fourier *infra* n. 8 p. 203. On est même surpris d'apprendre que Jacobi a connu le traité de Bézout, puisque, commme nous le verrons et comme Henry S. White s'exclame en 1909 :

> What a commentary on the futility of the best efforts is found in the fact that both Jacobi and Minding, only 60 years later, published investigations as new whose methods and results were in effect identical with Bézout's ! (*cf.* Henry S. White (président de l'AMS), « Bézout's theory of resultants and its influence on geometry », *Bulletin of the AMS*, 15 (1909), p. 325–338)

Voir aussi à ce sujet notre citation de Brill et Noether *supra* p. 142.

53. *Cf.* [236], §410.

54. *Cf.* [236], §413. Voir aussi le manuel moderne [191] chap. IX, lemme 3.9, p. 397, où le degré du résultant de n équations homogènes à n inconnues est minoré en spécialisant ces équations à des produits d'équations linéaires.

le résultat de la substitution pour tout α de $\frac{a_{\alpha k}X_k^{t_\alpha}-f_\alpha}{X_k^{t_\alpha}}$ à $a_{\alpha k}$ dans f. Dans le cas de polynômes f_1,...,f_r homogènes génériques, c'est-à-dire dont les coefficients sont les indéterminées $a, b, ...$, les propositions suivantes sont équivalentes [55] :
(i) pour t suffisamment grand $(f_1, ..., f_r)_t = (f, f_1, ..., f_r)_t$
(ii) f est *Trägheitsform*
(iii) il existe un k tel que $[f]_k$ est identiquement nul
(iv) il existe k, m tels que $X_k^m f \in (f_1, ..., f_r)$
(v) pour tout k il existe m tel que $X_k^m f \in (f_1, ..., f_r)$
(vi) il existe m tel que $X_0^m f \in (f_1, ..., f_r)$
On voit que la proposition (vi) ne fait qu'énoncer (après homogénéisation) l'existence de « termes d'introduction fictive » dans le langage de Bézout. Le plus petit entier k tel que $\mathfrak{M}^k f \subset (f_1, ..., f_r)$ est appelé « rang » de f (*Stufe*). Une *Trägheitsform* est dite « propre » si son rang est strictement positif. Hurwitz entreprend une étude générale des *Trägheitsform*. Considérons un instant le cas $n = r$. Hurwitz montre que les *Trägheitsform* propres de rang σ sont de degré $\sum t_\alpha - n - \sigma + 1$. Il montre que toutes les *Trägheitsform* de rang 1 appartiennent $(J, f_1, ..., f_n)$ où J est le jacobien de $f_1, ..., f_n$; en rang > 1, il parvient à donner des formules explicites au moins pour certaines *Trägheitsform*, celles qui sont linéaires en les coefficients de chacun des $f_1, ..., f_n$. Il montre l'existence d'une *Trägheitsform* de degré 0 qui est générateur de l'idéal de toutes les *Trägheitsform* de degré 0 : c'est le résultant. Mertens avait déjà donné une théorie rigoureuse, mais compliquée, du résultant (dans [228]). Pour le cas $r < n$, Hurwitz parvient à démontrer que $(f_1, ..., f_r)$ n'a pas de *Trägheitsform* propre. Considérons les deux assertions suivantes [56] :
(I_n) pour $r < n$, l'idéal $(f_1, ..., f_r)$ n'a pas de *Trägheitsform* propre
(II_n) pour $r \leq n$, si $A_1 f_1 + ... + A_r f_r = 0$, alors il existe des formes L_{ij} telles que $(\forall i)$ $A_i = \sum L_{ij} f_j$ et $(\forall i, j)$ $L_{ij} = -L_{ji}$
Hurwitz montre $I_n \Rightarrow II_n \Rightarrow I_{n+1}$. C'est de cette proposition que nous nous sommes inspirés ci-dessus pour démontrer (6) dans le cas $n = 3$. Pour r et n quelconques, Hurwitz démontre par un argument semblable que $(f_1, ..., f_r)$ n'a pas de *Trägheitsform* propre de degré $> \sum t_\alpha - r$.

Le projet de Hurwitz contient ceux de ces prédecesseurs Bézout, Hesse, Sylvester, Cayley, Mertens : il s'agit de décrire chaque composante homogène de l'idéal des *Trägheitsform* d'un idéal $(f_1, ..., f_r)$, en donnant pour chacune, si possible, une base. Ces composantes homogènes sont des $K[a, ...]$-modules, et il s'agit donc, essentiellement, d'un problème d'algèbre linéaire. Hurwitz avoue cependant que « le problème de déterminer toute les *Trägheitsform* d'un module [=idéal] semble présenter de sérieuses difficultés » [57]. En 1980, Jean-Pierre Jouanolou a entrepris une étude du complexe de Koszul de l'idéal $(f_1, ..., f_r)$ dans le langage de la théorie des schémas de Grothendieck. Le problème des *Trägheitsform* devient alors un problème de « cohomologie locale ». Sans pouvoir décrire ici ces travaux profonds, indiquons simplement comment se traduit l'une des propositions de Hurwitz dans ce cadre conceptuel. Jouanolou démontre [58] que certains groupes de cohomologie à support dans l'idéal $\mathfrak{M} = (X_0, ..., X_n)$ sont nuls, en utilisant des suites spectrales ayant pour aboutissement l'hypercohomologie du complexe de Koszul relativement au foncteur $\Gamma_{\mathfrak{M}}$ des sections à support dans $\mathfrak{M}$. Ainsi par exemple, si $n > r$, on a $H^0_{\mathfrak{M}}(C/(f_1, ..., f_r)) = 0$, ce qui signifie que l'idéal $(f_1, ..., f_n)$ n'a pas de *Trägheitsform* propre.

Insister sur le rôle de la formule de récurrence (3) ou de la somme alternée (10) dans l'ouvrage de Bézout pourrait sembler à certains une erreur épistémologique grossière, et la mention du complexe de Koszul un anachronisme. En 1948, Georges Bouligand soulignait l'« influence profonde », mais

55. *Cf.* [159] ; Hurwitz démontre $(i) \iff (ii)$ dans sa proposition 1, et $(ii) \iff (iii) \iff (iv) \iff (v)$ dans la proposition 9. Il est alors évident que $(iv) \iff (v) \iff (vi)$. L'importance du critère (iii) apparaît dans les travaux de Mertens. Pour $(ii) \iff (vi)$, voir aussi Jouanolou [168], (4.2.3) p. 132.

56. Si l'on traduit, comme l'a fait Jouanolou [168] p. 137–140, ces deux assertions dans un langage plus moderne, on obtient :

$$\begin{array}{lll} (I_n) & H^0_{\mathfrak{M}}(C/(f_1, ..., f_r)) = 0 & \text{lorsque} \quad r < n \\ (II_n) & H_1 K_\bullet(f_1, ..., f_r; C) = 0 & \text{lorsque} \quad r \leq n \end{array}$$

L'on reconnaît en (II_n) un des critères de régularité de la suite $(f_1, ..., f_n)$.

57. *Cf.* [159], p. 614.

58. *Cf.* [167], § 2.1 à 2.8.

« passée inaperçue », de la *Théorie des équations algébrique* dans la constitution de la géométrie algébrique [59]. Bézout lui-même est d'ailleurs fier de son accomplissement :

> Nous nous estimons heureux si considérant le point où nous avons pris les choses, et celui où nous les amenons, on trouve que nous avons acquitté une partie du tribut que tout homme doit à la société dans l'état où il se trouve placé. [60]

Mais force est de constater combien l'influence directe du traité de 1779 est difficile à cerner, à cause du peu de travaux qui s'inscrivent explicitement dans sa continuité (*cf. supra* n. 52 p. 159) ou, encore moins, qui adoptent le langage forgé par Bézout. Nous avons vu au contraire des auteurs construire de toutes pièces un langage différent (celui de la « méthode dialytique » de Sylvester, puis celui de la « théorie de l'involution » de Cayley) pour exprimer une même réalité. Ce traité est donc l'un des recoins obscurs de l'histoire des mathématiques. La raison est évidente : Bézout n'a pu donner une démonstration simple et rigoureuse de son théorème – l'algèbre linéaire à laquelle auraient recours Serret puis C. Schmidt n'existait pas encore, et celle à laquelle a eu recours Bézout n'était pas encore une science exacte. Notre point de vue est que cette incapacité a bien eu une conséquence positive. Refuser de l'admettre serait reléguer la théorie de Bézout au rang d'erreur scientifique, ce qui en soi ne bloque pas toute perspective historique, mais reste tout de même surprenant, car ce n'est pas d'un article ni d'un mémoire ni d'une idée qu'il est question, mais bien des quatre cents pages de l'unique ouvrage de recherche publié par Bézout (si l'on excepte ses quelques articles de jeunesse). Pour comprendre la conséquence positive de l'incapacité où était Bézout de donner une démonstration simple et rigoureuse de son théorème, il serait vain d'étudier ses contemporains, il faut diriger notre regard en avant de son temps. On pourrait préciser le titre que Bézout a donné à son ouvrage, en le qualifiant de « théorie combinatoire des systèmes d'équations algébriques ». Il a entrepris une recherche sur la dimension d'espaces vectoriels de polynômes, dont un cas particulier notable fut repris par Hilbert, et garda son nom : c'est le « polynôme de Hilbert » [61]

$$P(t) = \dim C_{\leq t}/\mathrm{Im}(f_1, ..., f_n)_{\leq t} = \dim\left(\mathrm{coker} f_{\leq t}\right).$$

La section suivante donnera la mesure de l'extension des recherches de Bézout dans ce domaine.

10.7 Les « équations incomplètes »

Les recherches de Bézout sur le résultant d'un système d'équations génériques dans $K[x_1, ..., x_n]$ ne constituent qu'une petite portion de son ouvrage. Il s'intéresse ensuite à d'autres systèmes d'équations, équations qui sont, pour ainsi dire, génériques sauf en certains monômes dont les coefficients sont nuls. Résumons le principe suivant lequel il classe les polynômes. Les polynômes « complets » sont les polynômes génériques de leur degré. Les polynômes « incomplets » sont encore distingués suivant leur « ordre ». Les « polynômes incomplets du premier ordre » sont ainsi décrits par Bézout (il utilise tantôt le terme « polynôme », tantôt le terme « équation ») :

59. *Cf.* Georges Bouligand, « À une étape décisive de l'algèbre. L'œuvre scientifique d'Étienne Bézout », *Revue générale des sciences*, 55 (1948), 121–123.

60. *Cf.* [29], p. xxi.

61. En voici un exemple élémentaire, emprunté au manuel de géométrie algébrique de Hartshorne (*cf.* [140], th. I.7.8), dans la preuve du théorème de Bézout pour l'intersection d'une surface et d'une sous-variété quelconque dans $\mathbb{P}^n$. Plaçons-nous par exemple dans $\mathbb{P}^2$, avec deux polynômes homogènes f_1,f_2 définissant deux courbes.

$$\begin{array}{ccccccccc} 0 & \to & k[X_0,X_1,X_2]/(f_1) & \longrightarrow & k[X_0,X_1,X_2]/(f_1) & \longrightarrow & k[X_0,X_1,X_2]/(f_1,f_2) & \to & 0 \\ & & \phi & \longmapsto & \phi f_2 & & & & \end{array}$$

est exacte. Au moyen du calcul aux différences finies, Hartshorne montre que $P(n) = \dim k[X_0, X_1, X_2]_n/(f_1, f_2)$ est un polynôme à coefficients rationnels, appelé « polynôme de Hilbert » de l'intersection des deux courbes ([140], th. I.7.5), puis il le calcule grâce à la relation de récurrence fournie par la suite exacte (th. I.7.7). Cette relation de récurrence est l'analogue de notre formule (7), laquelle s'expliquerait à son tour par l'exactitude de la suite suivante (en reprenant les notations que nous avons employées dans les sections précédentes) :

$$0 \quad \to \quad C_{\leq t_0-t_1}/(f_2,f_3)_{\leq t_0-t_1} \quad \xrightarrow{f_1} \quad C_{\leq t_0}/(f_2,f_3)_{\leq t_0} \quad \to \quad C_{\leq t_0}/(f_1,f_2,f_3)_{\leq t_0} \quad \to \quad 0$$

elle-même conséquence immédiate de la régularité de la suite (f_1^h, f_2^h, f_3^h).

> 1° que le nombre total des inconnues étant n, leur combinaison n à n serait d'un certain degré quelconque différent pour chaque équation ;
> 2° que leurs combinaisons $n-1$ à $n-1$ seraient de degrés quelconques, différents non seulement pour chaque équation, mais encore pour chacune de ces combinaisons ;
> 3° que leurs combinaisons $n-2$ à $n-2$ seraient de degrés quelconques, différents non seulement pour chaque équation, mais encore pour chacune de ces combinaisons ;
> etc. [62]

Mais Bézout est encore amené à distinguer, au sein de cet « ordre » de polynômes, différentes « espèces » (nous avons déjà mentionné la « première espèce d'équations incomplètes ») :

> Comme il n'est pas possible d'attaquer de front ce problème (celui des polynômes incomplets du premier ordre), je l'ai pris en sens inverse, en ne supposant d'abord que l'absence des plus hautes dimensions des combinaisons une à une, puis l'absence de celles-ci, et des plus hautes dimensions des combinaisons deux à deux, etc. et d'abord, avec quelques restrictions, mais dont l'objet était de faciliter l'intelligence de la méthode (...)

Il introduit les notations suivantes :

$$
\begin{array}{ll}
(u^a ... n)^t & (cf.\ \S 57\text{–}67) \\
[(u^a, x'^a)^b, y''^a ... n]^t & (cf.\ \S 74\text{–}81) \\
([(u^a, x'^a)^b, (u^a, y''^a)'^b, (x'^a, y''^a)''^b]^c, z'''^a ... n)^t & (cf.\ \S 82\text{–}132) \\
\ldots &
\end{array}
$$

La deuxième ligne, par exemple, désigne un polynôme que l'on écrirait en notations modernes :

$$
\begin{array}{ll}
\displaystyle\sum & A_{k\underset{'}{k}\underset{''}{k}} u^k x'^{\underset{'}{k}} y''^{\underset{''}{k}} \ldots \\
k \leq a, \underset{'}{k} \leq \underset{'}{a}, \underset{''}{k} \leq \underset{''}{a}, \ldots, & \\
k + \underset{'}{k} \leq b, & \\
k + \underset{'}{k} + \underset{''}{k} + \ldots \leq t &
\end{array}
$$

où $u, x, y, z, \ldots$ désignent les inconnues. La classification des équations par Bézout ne s'arrête pas là. Dans la section III du livre I, il traite des « polynômes incomplets du second, troisième, quatrième ordres, etc. » au moyen de la notation :

$$(u^{a,\bar{a},\bar{\bar{a}},\ldots} \ldots n)^{t,\bar{t},\bar{\bar{t}},\ldots}$$

où $a \leq \bar{a} \leq \bar{\bar{a}} \leq \ldots$ et $t \geq \bar{t} \geq \bar{\bar{t}} \geq \ldots$. Un tel polynôme s'écrirait de manière moderne sous la forme :

$$
\begin{array}{lll}
 & \displaystyle\sum_{k \leq \bar{\bar{a}}, \ldots, k + \underset{'}{k} + \underset{''}{k} + \ldots \leq \bar{\bar{t}}} & A_{k\underset{'}{k}\underset{''}{k}} u^k x'^{\underset{'}{k}} y''^{\underset{''}{k}} \ldots \\
+ & \displaystyle\sum_{k \leq \bar{a}, \ldots, \bar{\bar{t}} < k + \underset{'}{k} + \underset{''}{k} + \ldots \leq \bar{t}} & A_{k\underset{'}{k}\underset{''}{k}} u^k x'^{\underset{'}{k}} y''^{\underset{''}{k}} \ldots \\
+ & \displaystyle\sum_{k \leq a, \ldots, \bar{t} < k + \underset{'}{k} + \underset{''}{k} + \ldots \leq t} & A_{k\underset{'}{k}\underset{''}{k}} u^k x'^{\underset{'}{k}} y''^{\underset{''}{k}} \ldots \\
+ & \ldots &
\end{array}
$$

Tout le livre I du traité de Bézout est ordonné suivant un degré croissant de généralité. Bézout rappelle plusieurs fois que les équations traitées dans les étapes antérieures sont des cas particuliers des nouvelles formes étudiées [63]. D'autre part, le calcul du degré de l'équation finale pour les équations complètes dès la section I du livre I permet au moins de fixer une limite supérieure dans tous les autres cas : en effet les équations envisagées dans les sections II et III s'obtiennent à partir des équations complètes, en annulant certains de leurs coefficients.

62. *cf.* [29], p. xiii.
63. cf. [29], § 64, 81.

Pour décrire la « première espèce d'équations incomplètes » étudiée par Bézout (aux § 57–67), en gardant le système de notations que nous avons employé dans les sections précédentes, nous allons devoir compliquer ces notations, en introduisant deux systèmes d'indices : l'indice en bas désignera l'inconnue concernée, et l'indice en haut l'équation [64]. Notons alors

$$C_{\leq t,\underline{a}} = C_{\leq t,(a_1,\ldots,a_n)} = \left\{ f \in K[x_1,..,x_n] | \deg f \leq t; \deg_{x_i} f \leq a_i (\forall i \in [1,n]) \right\}$$

La première espèce de polynômes est alors la réunion des $C_{\leq t,\underline{a}}$ vérifiant les conditions :

$$\begin{cases} \max\{a_i\}_{1\leq i\leq n} \leq t \\ \min\{a_i + a_j\}_{1\leq i<j\leq n} \geq t \end{cases}$$

Au § 59, il calcule le nombre de termes d'un tel polynôme, c'est-à-dire :

$$\dim C_{\leq t,\underline{a}} = \binom{n+t}{t} - \sum_{i=1}^{n} \binom{n+t-a_i-1}{n}$$

Introduisons encore l'opérateur aux différences finies $\nabla_{t,\underline{a}}$ défini par :

$$\nabla_{t,\underline{a}} P(T,\underline{A}) = P(T,\underline{A}) - P(T+t,\underline{A}+\underline{a})$$

L'argument que nous avons détaillé ci-dessus pour la détermination du degré du résultant de n polynômes de $K[x_1,...,x_n]$, chacun étant générique de son degré, est appliqué par Bézout au cas de n polynômes génériques :

$$\begin{cases} f \in C_{\leq t,\underline{a}} \\ f' \in C_{\leq t',\underline{a'}} \\ \vdots \\ f^{(n-1)} \in C_{\leq t^{(n-1)},\underline{a^{(n-1)}}} \end{cases}$$

Il en conclut que le degré du résultant est :

$$D = \nabla_{-t,-\underline{a}} \nabla_{-t',-\underline{a'}} \ldots \nabla_{-t^{(n-1)},-\underline{a^{(n-1)}}} \dim C_{\leq T,\underline{A}}$$

où T et $\underline{A}$ sont des variables quelconques, qui jouent le rôle du t_0 ci-avant. Le calcul donne alors :

$$D = \prod_i t^{(i)} - \sum_i \prod_j (t^{(j)} - a_i^{(j)}),$$

résultat que nous avons déjà mentionné. Par exemple, pour $n = 3$ et $f, f', f'' \in C_{\leq 2,(1,1,1)}$, c'est-à-dire :

$$\begin{cases} f = ax_1x_2 + bx_1x_3 + cx_2x_3 + dx_1 + ex_2 + fx_3 + g \\ f' = a'x_1x_2 + b'x_1x_3 + c'x_2x_3 + d'x_1 + e'x_2 + f'x_3 + g' \\ f'' = a''x_1x_2 + b''x_1x_3 + c''x_2x_3 + d''x_1 + e''x_2 + f''x_3 + g'' \end{cases}$$

Bézout trouve que $D = 5$ (au § 65). Hélas, on ne peut appliquer à ce cas, plus général que celui de polynômes génériques de $K[x_1,...,x_n]$, les raisons que nous avons évoquées plus haut pour justifier la propriété (4). Pourtant, nous allons voir que les calculs de Bézout sont exacts.

Mais Bézout ne s'arrête pas là ; il étudie une seconde espèce de polynômes incomplets (§ 74–81), puis une troisième (§ 82–132). Quant à la seconde, il faut remarquer que Waring a reproduit en 1782, dans la préface à la seconde édition de ses *Meditationes algebraicae* (p. *xvii* à *xx*), le résultat donné par Bézout au § 80. Waring écrit :

> Soient h équations de degrés respectifs $n, m, l, k, ...$ contenant autant d'inconnues $x, y, z, v, ...$; soient $p, q, r, s, ..., p', q', r', ..., p'', q'', ...$ les degrés maximaux en les inconnues $x, y, z, v, ...$ des termes présents dans les équations respectives de degrés $n, m, l, k, ...$; alors l'équation dont la racine est x ou y ou z etc. sera de degré au plus $n \times m \times l \times k \times ... - (n-p) \times (m -$

64. Attention, l'indice en haut sera noté à la manière des dérivées : t, t', t'', ..., $t^{(n-1)}$.

$p')\times(l-p'')\times...-(n-q)\times(m-q')\times(l-q'')\times...-(n-r)\times(m-r')\times...-... = P$. Mais si les degrés totaux en (x, y) des équations ci-dessus sont a, a', a'',..., alors l'équation dont la racine est x ou y ou z etc. sera de degré $P+(n-a)\times(m-a')\times...-(p+q-a)\times(m-a')\times...$ [65]

Il oublie cependant une condition nécessaire précisée par Bézout (au § 74). En gardant les notations de Waring, cette condition s'énonce comme suit :

$$\begin{array}{lll} r+a\geq n & r'+a'\geq m & r''+a''\geq l \quad ... \\ s+a\geq n & s'+a'\geq m & s''+a''\geq l \\ \vdots & \vdots & \vdots \end{array}$$

Sans nous attarder sur cette seconde espèce de polynômes incomplets, décrivons plutôt la troisième espèce. Choisissons trois inconnues parmi n, soient x_1, x_2, x_3. On définit le degré d'un polynôme en cette combinaison de trois inconnues, noté $\deg_{123} f$, comme étant le degré de f vu comme polynôme en x_1, x_2, x_3 à coefficients dans $K[x_4, ..., x_n]$. On définit de même le degré de f en une combinaison de deux inconnues ; par exemple $\deg_{13} f$, degré de f vu comme polynôme en x_1, x_3 à coefficients dans $k[x_2, x_4, ..., x_n]$. Considérons à présent

$$C_{\leq t,\underline{a},\underline{b},c} = \left\{ f \in C_{\leq t,\underline{a}} | \deg_{123} f \leq c; \deg_{12} f \leq b_1; \deg_{13} f \leq b_2; \deg_{23} f \leq b_3 \right\}$$

La troisième espèce de polynômes est alors la réunion des $C_{\leq t,\underline{a},\underline{b},c}$ vérifiant les conditions suivantes :

$$\left\{ \begin{array}{ll} (i) & \max\{a_1, a_2\} \leq b_1, \quad \max\{a_1, a_3\} \leq b_2, \quad \max\{a_2, a_3\} \leq b_3, \\ (ii) & a_1 + a_2 \geq b_1, \quad a_1 + a_3 \geq b_2, \quad a_2 + a_3 \geq b_3, \\ (iii) & \max\{c, a_4, ..., a_n\} \leq t, \\ (iv) & a_i + a_j \geq t \quad (\forall i \in [1, n], j \in [4, n], j > i), \\ (v) & \max\{b_1, b_2, b_3\} \leq c, \\ (vi) & a_1 + b_3 \geq c, \quad a_2 + b_2 \geq c, \quad a_3 + b_1 \geq c, \\ (vii) & b_1 + b_2 + b_3 \geq 2c \end{array} \right.$$

Commentons rapidement ces conditions. Les conditions (ii) assurent qu'il existe, dans $C_{\leq t,\underline{a},\underline{b},c}$, des monômes en x_1, x_2 (resp. en x_1, x_3, resp. en x_2, x_3) de degré total $\geq b_1$ (resp. b_2, resp. b_3). Les conditions (iv) assurent, comme le dit Bézout, que les $(n-3)$ dernières inconnues « dans leurs combinaisons deux à deux, trois à trois, quatre à quatre, etc. tant entre elles qu'avec les trois premières, montent à toutes les dimensions possibles, jusqu'à celle du polynôme » [66]. Les conditions (vi) et (vii) assurent qu'il existe bien des monômes en x_1, x_2, x_3 de degré total $\geq c$ (ce dernier point, qui n'est pas totalement évident, est démontré par Bézout au § 83).

Bézout se voit alors confronté à un problème de taille : bien que l'ensemble $C_{\leq t,\underline{a},\underline{b},c}$ soit un sous-espace vectoriel de $C_{\leq t}$, sa dimension n'est pas un polynôme en $t, \underline{a}, \underline{b}, c$. Il s'agit en quelque sorte d'une fonction « polynomiale par morceaux ». Plus précisément, Bézout distingue huit « formes », suivant le signe des trois expressions $E_1 = c-b_1-b_2+a_1$, $E_2 = c-b_1-b_3+a_2$ et $E_3 = c-b_2-b_3+a_3$. Pour chacune de ces « formes », $\dim C_{\leq t,\underline{a},\underline{b},c}$ est un polynôme différent en $t, \underline{a}, \underline{b}, c$. Supposons un instant que l'argument pour la détermination du degré du résultant s'adapte au moins jusqu'à la formule (3) à la troisième espèce, et que l'on trouve une démonstration suffisamment générale de la propriété (4). L'on aurait une formule, analogue aux formules trouvées dans les cas précédents :

$$D = \nabla_{-t,-\underline{a},-\underline{b},-c} \nabla_{-t',-\underline{a'},-\underline{b'},-c'} ... \nabla_{-t^{(n-1)},-\underline{a^{(n-1)}},-\underline{b^{(n-1)}},-c^{(n-1)}} \dim C_{\leq T,\underline{A},\underline{B},C}$$

Mais il serait bien difficile de trouver une expression polynomiale de D à partir de cette formule si $\dim C_{\leq T,\underline{A},\underline{B},C}$ n'est pas lui-même un polynôme. A moins que, après avoir évalué les opérateurs ∇,

65. si sint (h) aequationes $(n, m, l, k, \&c.)$ dimensionum respective totidem incognitas quantitates $(x, y, z, v, \&c.)$ involventes ; et sint p, q, r, s, &c. ; p', q', r', &c. ; p'', q'', &c. ; maximae dimensiones, ad quas ascendunt incognitae quantitates x, y, z, v, &c. , in respectivis aequationibus $(n, m, l, k, \&c.)$ dimensionum ; tum aequationem, cujus radix est x vel y vel z, &c. haud ascendere ad plures quam $n \times m \times l \times k \times \&c. - (n-p) \times (m-p') \times (l-p'') \times \&c. - (n-q) \times (m-q') \times (l-q'') \times \&c. - (n-r) \times (m-r') \times \&c. - \&c. = P$ dimensiones : si vero dimensiones quantitatum (x&c. y) simul sumptarum haud superent dimensiones $a, a', a'', \&c.$ in predictis aequationibus ; tum aequationem, cujus radix est $x, \&c.$ haud plures habere quam $P + (n-a) \times (m-a') \times \&c. - (p+q-a) \times (p'-a') \times \&c.$

66. *Cf.* § 82. Rappelons que Bézout désigne par « dimension » le degré d'un polynôme.

tous les $\dim C$ figurant dans l'expression obtenue appartiennent à une même « forme ». C'est effectivement le cas, on s'en convainc facilement, pour des valeurs suffisamment grandes de $T, \underline{A}, \underline{B}, C$, bien que la « forme » en question dépende justement de la valeur de $T, \underline{A}, \underline{B}, C$. Nous avons fait cette petite digression, afin d'offrir quelque vraisemblance à la méthode de Bézout : il calcule en effet 8 valeurs différentes de D, en supposant dans l'expression ci-dessus que $C_{\leq T,\underline{A},\underline{B},C}$ appartient tour à tour à chacune des 8 formes.

Nous allons détailler le cas $n = 3$, le seul pour lequel Bézout donne un calcul explicite de D. Dans ce cas, $\deg_{123} = \deg$, et l'on peut simplifier les notations ci-dessus en posant $c = t$ et $C_{\leq t,\underline{a},\underline{b}} = C_{\leq t,\underline{a},\underline{b},c}$. Voici les valeurs de $\dim C_{\leq t,\underline{a},\underline{b}}$ obtenues pour chacune des huit « formes »[67]. On numérote les « formes » de 1 à 8, et l'on note $P_i(T, \underline{A}, \underline{B})$ le polynôme exprimant la dimension de $C_{\leq T,\underline{A},\underline{B}}$ pour la i-ème « forme ».

1ère forme ($E_1 \leq 0, \quad E_2 \leq 0, \quad E_3 \leq 0$) :

$$\begin{aligned} P_1(T, \underline{A}, \underline{B}) = &\tbinom{3+T}{3} - \textstyle\sum_{i=1}^{3} \tbinom{3+T-A_i-1}{3} + \sum_{i=1}^{3} \tbinom{3+T-B_i-2}{3} \\ &- \textstyle\sum_{i=1}^{3} \left[(A_1 + A_2 + A_3 - A_{4-i} - B_i).\tbinom{2+T-B_i-1}{2} \right] \end{aligned}$$

2de forme ($E_1 \leq 0, \quad E_2 \leq 0, \quad E_3 \geq 0$) :

$$P_2(T, \underline{A}, \underline{B}) = P_1(T, \underline{A}, \underline{B}) + \binom{3 + T + A_3 - B_2 - B_3 - 2}{3}$$

3ème forme ($E_1 \geq 0, \quad E_2 \leq 0, \quad E_3 \leq 0$) :

$$P_3(T, \underline{A}, \underline{B}) = P_1(T, \underline{A}, \underline{B}) + \binom{3 + T + A_1 - B_1 - B_2 - 2}{3}$$

4ème forme ($E_1 \geq 0, \quad E_2 \leq 0, \quad E_3 \geq 0$) :

$$P_4(T, \underline{A}, \underline{B}) = P_3(T, \underline{A}, \underline{B}) + \binom{3 + T + A_3 - B_2 - B_3 - 2}{3}$$

5ème forme ($E_1 \geq 0, \quad E_2 \geq 0, \quad E_3 \leq 0$) :

$$P_5(T, \underline{A}, \underline{B}) = P_3(T, \underline{A}, \underline{B}) + \binom{3 + T + A_2 - B_1 - B_3 - 2}{3}$$

6ème forme ($E_1 \geq 0, \quad E_2 \geq 0, \quad E_3 \geq 0$) :

$$P_6(T, \underline{A}, \underline{B}) = P_5(T, \underline{A}, \underline{B}) + \binom{3 + T + A_3 - B_2 - B_3 - 2}{3}$$

7ème forme ($E_1 \leq 0, \quad E_2 \geq 0, \quad E_3 \leq 0$) :

$$P_7(T, \underline{A}, \underline{B}) = P_1(T, \underline{A}, \underline{B}) + \binom{3 + T + A_2 - B_1 - B_3 - 2}{3}$$

8ème forme ($E_1 \leq 0, \quad E_2 \geq 0, \quad E_3 \geq 0$) :

$$P_8(T, \underline{A}, \underline{B}) = P_7(T, \underline{A}, \underline{B}) + \binom{3 + T + A_3 - B_2 - B_3 - 2}{3}$$

Notons finalement $D_i = \nabla_{-t,-\underline{a},-\underline{b}} \nabla_{-t',-\underline{a'},-\underline{b'}} \nabla_{-t'',-\underline{a''},-\underline{b''}} P_i(T, \underline{A}, \underline{B})$. Bézout calcule explicitement les D_i (formules qui n'occupent pas moins de huit pages, § 119–127). Il propose ensuite un test (« les symptômes auxquels on reconnaît parmi les différentes expressions de la valeur du degré de l'équation finale, quelle est celle que l'on doit choisir ou rejeter », *cf.* § 117) qui permet de rejeter

67. *Cf.* § 91–99, où l'on pose C=T.

certains des D_i. Ce test est intimimement lié à la conviction de Bézout que la formule (8) est exacte pour tout k. Les autres D_i (ceux qui sont « admissibles ») doivent, affirme Bézout, être égaux entre eux et au degré D du résultant.

Pour mieux cerner le degré de généralité du programme de recherche imaginé par Bézout, nous allons introduire une notation qui contient toutes les précédentes. Soit u une fonction :

$$u : \mathcal{P}(\{1, ..., n\}) \to \mathbb{N}$$

Nous noterons $C_{\leq u}$ le sous-espace vectoriel de $C = k[x_1, ..., x_n]$ défini par :

$$C_{\leq u} = \left\langle x_1^{\alpha_1} ... x_n^{\alpha_n} \quad | \quad \forall I \subset \{1, ..., n\} \quad \sum_{i \in I} \alpha_i \geq u(I) \right\rangle$$

On a aussi une notion de « polynôme générique » dans $C_{\leq u}$ (un polynôme à coefficients indéterminés sauf ceux qui sont nuls par définition de $C_{\leq u}$). Pour n polynômes génériques

$$f^{(i)} \in C_{\leq u^{(i)}} \quad (0 \leq i \leq n-1),$$

la méthode de Bézout revient à chercher une résolution libre de type fini de $C_{\leq U}$, construite comme un complexe de Koszul (pour U « assez grand ») :

$$(11) \quad \ldots \xrightarrow{\wedge f} \bigoplus C_{\leq U - u^{(i)}} \xrightarrow{f = \begin{pmatrix} f \\ f' \\ \vdots \\ f^{(n-1)} \end{pmatrix}} C_{\leq U}$$

Dans le cas des « équations complètes », U et les u_i sont des constantes, et le complexe ci-dessus, après homogénéisation (pour tout $u \in \mathbb{N}$, $C_{\leq u} = \{f \in C \mid \deg f \geq u\}$ devient C_u^h après homogénéisation), apparaît naturellement comme composante graduée d'un complexe de Koszul qui est bien acyclique comme on l'a vu dans la section précédente. Dans le cas général, l'hypothèse implicite dans le raisonnement de Bézout revient aussi à supposer l'acyclicité du complexe (11) pour U « assez grand ». Alors Bézout détermine le « degré de l'équation finale » comme étant $\dim(C_{\leq U}/\mathrm{Im} f)$, égal à une somme alternée, la caractéristique d'Euler-Poincaré du complexe. Une version d'un théorème de D. N. Bernshtein [19] paru en 1975 donne immédiatement les résultats établis par Bézout. Plus encore, ce théorème conduit précisément aux mêmes calculs que ceux faits par Bézout (c'est-à-dire le calcul de la somme alternée). Les résultats sont donc exacts, et les calculs aussi. Mais Bernshtein utilise pour le démontrer des outils étrangers à la pensée de Bézout (développements de Puiseux à plusieurs variables, volume mixte de Minkowski).

Pourtant, la méthode même utilisée par Bézout est exacte. Il suffit pour le voir de lire un article publié l'année suivante par A. G. Kushnirenko (1976). Le théorème de Bernshtein dépasse un peu, en généralité, le projet de Bézout, puisqu'il porte sur les espaces de polynômes à support dans un polyèdre convexe (le support d'un polynôme est l'ensemble des monômes à coefficients non nuls). L'espace de polynômes $C_{\leq u}$ est soumis à un système d'inégalités sur les degrés qui définit bien un polyèdre convexe, mais d'un type particulier. Le raisonnement de Bézout est pourtant général. Kushnirenko parvient à construire un complexe de Koszul dont le complexe d'espaces vectoriels (11) est une composante graduée, afin de donner une démonstration nouvelle du théorème de Bernshtein. L'article de Kushnirenko est intimement lié à la théorie des « variétés toriques » née au début des années 1970. Kushnirenko construit en effet une compactification du « tore » $(k^\star)^n$ et c'est sur cette variété qu'il étudie les zéros du système d'équations initial. Lorsque le système vérifie une certaine « condition de non-singularité », ses zéros sont isolés et contenus dans le tore. Sous cette condition, le complexe de Koszul est acyclique en degré élevé [68]. Appliquons la théorie de Kushnirenko aux

68. Pour démontrer l'acyclicité du complexe de Koszul, Kushnirenko utilise un résultat fondamental sur les anneaux de Cohen-Macaulay (démontré par exemple dans [285] § IV.B.2, théorème 3, p. 65). D'autre part, il n'explicite pas ce

espaces $C_{\leq u}$ que nous avons définis plus haut. On obtient le théorème suivant [69] :

Théorème 1 *Soit $f^{(i)} \in C_{\leq u^{(i)}}$ n polynômes définissant un système d'équations $f^{(i)}(x_1, ..., x_n) = 0$ $(0 \leq i \leq n-1)$. En général (c'est-à-dire quand les coefficients des $f^{(i)}$ appartiennent à un certain ouvert de $\bigoplus_i C_{\leq u^{(i)}}$), les solutions du système sont isolées et de multiplicité 1. Pour un entier q assez grand, la quantité suivante est indépendante de q et elle est le nombre exact de solutions du système contenues dans le tore :*

$$D = \dim \left(C_{\leq q \sum_{i=0}^{n-1} u^{(i)}} \Big/ Imf \right) = \sum_{I \subset \{0,...,n-1\}} (-1)^{|I|} \dim C_{\leq q \sum_{i=0}^{n-1} u^{(i)} - \sum_{i \in I} u^{(i)}}$$

A titre d'exemple, appliquons ce théorème au cas où les $u^{(i)}$ sont chacun définis par un vecteur $v^{(i)} = (v_1^{(i)}, ..., v_n^{(i)}) \in \mathbb{N}^n$, de la manière suivante :

$$\begin{array}{rccc} u^{(i)} : & \mathcal{P}(\{1, ..., n\}) & \longrightarrow & \mathbb{N} \\ & I & \longmapsto & \sum_{j \in I} v_j^{(i)} \end{array}$$

Il est alors facile de voir que :

$$\begin{aligned} D &= \sum_{I \subset \{0,...,n-1\}} (-1)^{|I|} \dim C_{\leq q \sum_{i=0}^{n-1} u^{(i)} - \sum_{i \in I} u^{(i)}} \\ &= \sum_{I \subset \{0,...,n-1\}} (-1)^{|I|} \prod_{j=1}^{n} \left[q \sum_{i=0}^{n-1} v_j^{(i)} - \sum_{i \in I} v_j^{(i)} \right] \\ &= \sum_{\sigma \in \mathfrak{S}_n} \prod_{j=0}^{(n-1)} v_{\sigma(j+1)}^{(j)} \end{aligned}$$

Bézout n'a pas étudié cet exemple mais on trouve cette valeur de D, obtenue par d'autres méthodes, dans certains travaux de la fin du XIX$^{\text{ème}}$ siècle [70].

10.8 Cas de deux coniques : problème des relations entre les coefficients

Dans le sens d'une généralité plus grande, mais coûteuse en calculs et en notations, Bézout adresse en 1779 la critique suivante à la méthode d'Euler–Cramer :

> En prenant le parti de mettre les équations proposées sous la forme d'équations à une inconnue de moins que leur nombre, on abrège immensément les calculs ; mais (...) on est exposé à ne plus reconnaître la possibilité de l'abaissement de l'équation finale lorsque des relations particulières entre les coefficients connus, peuvent y donner lieu (...) [71].

C'est dans le livre II que Bézout étudie les « relations particulières entre les coefficients » susceptibles d'abaisser le degré de l'équation finale. Il traite l'exemple de deux coniques dans les moindres détails :

$$\begin{cases} ax^2 + bxy + cy^2 + dx + ey + f = 0 \\ a'x^2 + b'xy + c'y^2 + d'x + e'y + f' = 0 \end{cases}$$

qu'il appelle « condition de non-singularité », sauf dans le cas où tous les polyèdres coïncident (c'est-à-dire, dans nos notations, le cas où tous les $u^{(i)}$ sont égaux). Il faut revenir à l'article de Bernshtein pour comprendre cette condition de non-singularité. Mais le fait que les systèmes « non-singuliers » forment bien un ouvert de Zariski dans l'espace des coefficients ne nous semble pas suffisamment justifié dans la démonstration de Bernshtein (*cf.* [19], p. 184 : « F_α is a system of n equations in $n-1$ unknowns and does not have any roots in the general case »...). C'est pourtant vrai (voir par exemple [118]).

69. Dans l'article [186] de Kushnirenko, il s'agit du théorème 2.
70. *Cf.* Brill [43] p. 384–386 ; Hilbert [149] p. 440 ; K. Th. Vahlen, journal de Crelle *113*, p. 348 (1894).
71. *Cf.* [29], p. xix.

Il multiplie chacune de ces deux équations par un « polynôme multiplicateur » du second degré à coefficients indéterminés et identifie l'équation finale à :

$$\begin{array}{ll} & (Ax^2+Bxy+Cy^2+Dx+Ey+F)(ax^2+bxy+cy^2+dx+ey+f) \\ + & (A'x^2+B'xy+C'y^2+D'x+E'y+F')(a'x^2+b'xy+c'y^2+d'x+e'y+f')=0 \end{array}$$

En imposant que les coefficients des monômes contenant une puissance de y s'annulent, on obtient un système de 10 équations, à 12 inconnues $A, A', B, B', ..., F, F'$. Il y a une inconnue de trop, ou dans le vocabulaire de Bézout un « coefficient arbitraire », il est donc possible d'ajouter une équation. Bézout ajoute l'équation $Ca+C'a'=0$. On reviendra sur le problème posé par ce choix. Le système linéaire peut donc être représenté, en notation moderne, par la matrice 11×12 suivante :

$$\begin{pmatrix}
 & & & & a & a' & & & & & & \\
b & b' & a & a' & & & & & & & & \\
c & c' & b & b' & a & a' & & & & & & \\
 & & c & c' & b & b' & & & & & & \\
 & & & & c & c' & & & & & & \\
e & e' & d & d' & & & b & b' & a & a' & & \\
 & & e & e' & d & d' & c & c' & b & b' & & \\
 & & & & e & e' & & & c & c' & & \\
 & & f & f' & & & e & e' & d & d' & b & b' \\
 & & & & f & f' & & & e & e' & c & c' \\
 & & & & & & & & f & f' & e & e'
\end{pmatrix}$$

Une fois le système résolu, l'équation finale s'écrira

$$\begin{array}{ll} & (Aa+A'a')x^4 \\ + & (Ad+A'd'+Da+D'a')x^3 \\ + & (Af+A'f'+Fa+F'a'+Dd+D'd')x^2 \\ + & (Df+D'f'+Fd+F'd')x \\ + & (Ff+F'f')=0 \end{array}$$

Bézout utilise la méthode de Cramer pour résoudre le système linéaire mais il améliore cette méthode en donnant une technique de calcul des mineurs par récurrence, et une notation particulièrement adaptée au système ci-dessus : il note $(ab')=ab'-a'b$. Seules les inconnues A, A', D, D', F, F' interviennent dans l'expression de l'équation finale, donc Bézout ne calcule que les mineurs correspondant à ces inconnues. Il remarque que ces mineurs contiennent tous le facteur

$$(*)=[(bc')(cd')(ce')-(ce')^2(ac')+(bc')(cf')]=\begin{vmatrix}
b & b' & a & a' & & \\
c & c' & b & b' & & \\
 & & c & c' & & \\
e & e' & d & d' & b & b' \\
 & & e & e' & c & c' \\
 & & f & f' & e & e'
\end{vmatrix}$$

La mise en évidence de cette factorisation est assez simple pour les mineurs correspondant à A, A', F, F', moins évidente pour ceux correspondant à D et D'. Bézout obtient :

$$\begin{array}{ll} A'= & (*)[(bc')(cd')(ce')-(ce')^2(ac')+(bc')(cf')](ca')[b(bc')-c(ac')] \\ F'= & (*)[(bc')(cd')(ce')-(ce')^2(ac')+(bc')(cf')](ca')[c(cf')-e(ce')] \\ D'= & (*)(ca')\{[(cd')+(be')]c-2(ce')b\} \end{array}$$

L'équation finale peut donc être divisée par le facteur constant $(*)(ca')$, lorsque celui-ci n'est pas nul. Après quelques transformations, Bézout obtient l'équation finale, dégagée de son facteur constant,

sous une forme identique à celle obtenue en 1764 :

$$(**)\quad \begin{array}{lllll} (ac')^2 & x^4+(bc')(bd') & x^3+(bc')(bf') & x^2+2(cd')(cf') & x+(cf')^2 \\ -(ab')(bc') & -2(ac')(cd') & -2(ac')(cf') & -(ce')(bf') & -(ce')(ef') \\ & +(ab')(ce') & +(cd')^2 & +(bc')(cf') & \\ & -(bc')(ae') & -(ce')(bd') & +(ce')(de') & \\ & & -(bc')(de') & & \\ & & +(ae')(ce') & & \end{array} = 0$$

Mais il ajoute que le facteur constant $(*)(ca')$ contient des « connaissances utiles ». Il distingue en effet deux types de « symptômes » indiquant que le degré de l'équation finale est susceptible d'abaissement. Lorsque des relations particulières entre les coefficients des équations initiales font que le coefficient du terme en x^4 s'annule dans l'expression de l'équation finale $(**)$ donnée ci-dessus, il y a bien sûr abaissement du degré [72]. Il existe cependant un deuxième symptôme, qui n'apparaît pas dans $(**)$: lorsque des relations particulières entre les coefficients annulent le facteur $(*)$, on ne peut pas appliquer les formules de Cramer pour résoudre le système linéaire (car tous les mineurs 11×11 sont nuls [73]) et Bézout indique qu'il faut alors chercher une équation finale de degré 3, en utilisant des « polynômes multiplicateurs » de degré 1. Si l'on suppose l'équation finale de degré 3 et les polynômes multiplicateurs de degré 1, le système linéaire à résoudre contient 6 équations à 6 inconnues, et ne possède de solution non nulle qu'à la condition que son déterminant s'annule. Or le déterminant, justement égal au facteur $(*)$, s'annule par hypothèse. L'annulation du facteur $(*)$ est donc un « symptôme » de l'abaissement du degré de l'équation finale. En fait, on peut montrer que, réciproquement, si tous les mineurs 11×11 sont nuls, alors le facteur $(*)$ est identiquement nul [74].

Quant au facteur (ca'), il est lié à l'équation supplémentaire $Ca + C'a' = 0$. Ce sont en effet les deux équations

$$\begin{cases} Ca + C'a' = 0 \\ Cc + C'c' = 0 \end{cases}$$

qui introduisent ce facteur dans le calcul de A, A', D, D', F, F'. Si l'on choisissait l'équation supplémentaire $Cb + C'b' = 0$ au lieu de $Ca + C'a' = 0$, on obtiendrait un facteur (cb') au lieu de (ca'). Par contre, si l'on choisissait $Aa + A'a' = 0$, on serait conduit à une impasse. Cette équation signifie en effet que le coefficient du terme en x^4 s'annule, et donc que l'équation finale est de degré strictement inférieur à 4. Mais cela n'est pas vrai en général. Il suffit pour s'en persuader de trouver deux équations particulières conduisant à une équation finale de degré 4. Par exemple, on montre facilement que

$$\begin{cases} x^2 - y = 0 \\ y^2 - u = 0 \end{cases}$$

conduisent à l'équation finale irréductible $x^4 - u = 0$. Cette équation $x^4 - u = 0$ devrait aussi pouvoir s'obtenir en spécialisant la formule générale de l'équation finale (obtenue lorsque les deux équations proposées sont des équations générales [75]). Mais alors la formule générale de l'équation

72. Bézout n'étudie pas le cas où tous les coefficients s'annulent dans $(**)$. Le fait que le résultant $(**)$ des deux équations initiales, ordonnées suivant les puissances de y, s'annule identiquement, signifie que ces deux équations ont un facteur commun. C'est donc précisément le cas d'une intersection impropre qu'oublie Bézout.

73. On observe aisément que les mineurs correspondant à B, B', C, C', E, E' sont aussi divisibles par $(*)$. On a :

$$\begin{array}{l} B' = (*)(ca')(bc')c \\ C' = 0 \\ E' = -(*)(ce') \end{array}$$

Le facteur $(*)$ s'annule par exemple lorsque $b = c = e = 0$. Dans ce cas la première des deux équations initiales ne contient pas l'inconnue y, et la spécialisée de l'équation finale $(**)$ est égale au carré de cette équation, à une constante multiplicative près (c'^2).

74. Si tous les mineurs 11×11 sont nuls, il existe une relation linéaire entre les 11 lignes de la matrice. Mais la matrice est ainsi faite (elle contient un bloc 5×6 rempli de zéros) qu'une telle relation implique qu'il en existe au moins une entre les 5 premières lignes ou bien une entre les 6 dernières. Dans les deux cas, il est facile de montrer que $(*)$ s'annule identiquement.

75. Quitte à tenir compte des « symptômes » indiquant l'abaissement du degré de l'équation finale.

finale ne peut être de degré strictement inférieur à 4. L'équation supplémentaire $Aa + A'a' = 0$ conduit donc nécessairement à une équation finale identiquement nulle.

Le procédé qui vient d'être décrit (« spécialiser » les équations proposées afin de minorer le degré de l'équation finale) n'a pas été utilisé par Bézout [76]. On pourrait aussi calculer directement la solution du système linéaire avec l'équation $Aa + A'a' = 0$, et vérifier que l'équation finale ainsi obtenue est identiquement nulle. Bézout n'a pas non plus écrit ce calcul. Les raisons invoquées par Bézout pour éviter le choix de l'équation supplémentaire $Aa + A'a' = 0$ sont peu satisfaisantes :

> Si croyant pouvoir abaisser l'équation finale, on employait les douze coefficients qu'offrent les deux polynômes multiplicateurs, tant pour détruire les termes affectés de y, que pour détruire le terme x^4 ; alors on arriverait à une équation qui, si elle n'était pas identique [=identiquement nulle], serait en effet du troisième degré, mais qui n'appartiendrait pas à la question, puisqu'on n'y aurait pas exprimé l'existence des équations particulières. Les valeurs de x qu'on conclurait de cette équation, ne seraient donc nullement propres à satisfaire aux équations proposées : en un mot, cette prétendue équation finale, serait une équation purement arbitraire, et sans aucune liaison avec la question. [77]

Bézout n'a donc pas vu que l'équation finale obtenue au moyen de l'équation supplémentaire $Aa + A'a' = 0$ est nécessairement identiquement nulle. Dans tous les cas semblables où la méthode des coefficients indéterminés le conduit à un système d'équations auquel il faut ajouter des équations supplémentaires pour le rendre déterminé, Bézout laisse donc son lecteur dans le doute : comment en effet s'assurer que l'équation finale obtenue, si elle n'est pas identiquement nulle, n'est pas non plus une telle « équation purement arbitraire et sans aucune liaison avec la question » ? Comment choisir les équations supplémentaires de sorte que cela n'arrive pas ?

Un autre obstacle épistémologique que l'on rencontre dans le traité de Bézout est l'absence d'une conception correcte des points à l'infini. Heureusement, Bézout ne mentionne ce phénomène qu'une seule fois dans tout l'ouvrage, aux § 281–284. Il considère là deux équations de degré 2 :

$$\begin{cases} axy + bx + cy + d = 0 \\ a'xy + b'x + c'y + d' = 0 \end{cases}$$

Le résultant de ces deux équations en y est :

$$\begin{vmatrix} ax + c & bx + d \\ a'x + c' & b'x + d' \end{vmatrix} = 0$$

Cette équation de degré 2, dans les notations de Bézout, est :

$$(ab')x^2 + [(ad') - (bc')]\, x + (cd') = 0$$

Bézout, au moyen de sa méthode des polynômes-multiplicateurs, trouve cette même équation, mais accompagnée d'un facteur constant (ac'). Il s'interroge alors sur la signification de ce facteur, et remarque que lorsqu'il s'annule, il est possible de choisir deux autres polynômes-multiplicateurs pour obtenir une équation finale de degré 1 :

$$a(a'xy + b'x + c'y + d') - a'(axy + bx + cy + d) = (ab')x + (ad')$$

Dans ce cas, le résultant de degré 2 trouvé ci-dessus se factorise de la manière suivante :

$$[(ab')x + (ad')] \left(x + \frac{c}{a}\right) = 0$$

Il contient un « facteur superflu » donnant une racine $x = -\frac{c}{a}$ qui ne correspond à aucune solution finie du système proposé. Bézout en conclut :

76. Nous avons vu Serret l'utiliser dans ce but, *supra* n. 54 p. 159.

77. *Cf.* [29], § 230. Cet extrait est un des rares endroits où Bézout utilise la notion de racine (« les valeurs de x que l'on conclurait de cette équation »). Ici comme ailleurs dans ce traité, cette notion semble suppléer au manque de rigueur, en sortant du cadre conceptuel du traité par un appel à l'intuition de lecteur.

> Nous ne pouvons donc ne pas faire observer ici que la méthode ordinaire d'élimination pour les équations à deux inconnues, la seule que l'on ait eue, jusqu'ici, exempte de donner à l'équation finale un degré plus élevé que ne le comportent généralement les degrés particuliers des deux équations, n'est pas néanmoins à l'abri de donner des racines inutiles et même fausses. En effet, en suivant cette méthode, on est conduit immédiatement à l'équation
>
> $$(ab')x^2 + [(ad') - (bc')]\,x + (cd') = 0,$$
>
> sans aucune indication des cas où il n'y aura qu'une racine de cette équation qui soit admissible. C'est que cette méthode d'élimination est fondée sur une manière trop bornée d'envisager la question, et qui exclut du résultat, les symptômes qu'une Analyse plus générale nous fait ici découvrir. [78]

Interprétons, pour mieux comprendre, la situation ci-dessus en langage géométrique. Les deux équations proposées représentent deux hyperboles équilatères d'asymptotes $\{x = -\frac{c}{a}\}$, $\{y = -\frac{b}{a}\}$, $\left\{x = -\frac{c'}{a'}\right\}$, $\left\{y = -\frac{b'}{a'}\right\}$. Ces hyperboles, ayant deux directions asymptotiques communes, ont deux points d'intersection « à l'infini ». Ce fait, naturel pour qui connaît la structure du plan projectif $\mathbb{P}^2$, l'est moins au XVIII$^{\text{e}}$ siècle, longtemps avant la parution du *Traité des propriétés projectives des figures* de Poncelet. Une seconde manière de compactifier le plan affine $\mathbb{A}^2 \subset \mathbb{P}^1 \times \mathbb{P}^1$ conduit à ignorer ces points d'intersection à l'infini, sauf dans le cas ou deux asymptotes coïncident (et non seulement leurs directions). Dans ce dernier cas, il n'y a plus qu'une seule intersection finie. Le calcul confirme cela. Le fait que deux asymptotes coïncident se traduit par l'équation de condition

$$(ab')(ac') = 0$$

Si $(ab') = 0$, le terme en x^2 du résultant de degré 2 s'évanouit ; il devient donc de degré 1. Si $(ac') = 0$, le résultant de degré 2 se factorise comme nous l'avons indiqué, et son facteur $\left(x + \frac{c}{a}\right)$ indique justement l'abscisse du point d'intersection à l'infini. Ainsi, quelle que soit la manière de concevoir les « points à l'infini », le facteur $\left(x + \frac{c}{a}\right)$ indique bien la présence d'un tel point. Hélas, Bézout nie catégoriquement l'existence d'un point à l'infini dans ce cas. La citation suivante, assez longue, est la seule occurence de l'infini dans le traité de Bézout, et l'infini est ainsi rejeté de sa théorie de l'élimination :

> ...il faut lever une objection qu'on seroit peut-être tenté de faire.
>
> On pourroit peut-être penser que la valeur $x = -\frac{c}{a}$ a pour correspondante en y, une valeur infinie ; car dans la supposition de y infinie, la quantité $axy + bx + cy + d$ se réduit à $axy + cy$; et la quantité $a'xy + b'x + c'y + d'$ se réduit à $a'xy + c'y$; les deux équations dans cette hypothèse semblent donc se réduire à
>
> $$\begin{aligned} & axy + cy = 0, \\ \text{et.....} \quad & a'xy + c'y = 0, \end{aligned}$$
>
> lesquelles dans la supposition de $(ac') = 0$, ont lieu toutes deux en supposant $ax + c = 0$.
>
> Mais il faut bien observer que la quantité $axy + bx + cy + d$ ne se réduit à $axy + cy$ dans l'hypothèse de y infinie, qu'autant que $axy + cy$ peut être censé infini à l'égard de $bx + d$; or le contraire a lieu (...)
>
> Que si l'on insistoit en disant qu'à la vérité, dans le cas de y infinie, on ne doit point négliger $bx + d$ vis-à-vis de $axy + cy$ non plus que $b'x + d'$ vis-à-vis de $a'xy + c'y$; mais que les deux équations
>
> $$\begin{aligned} & axy + cy + bx + d = 0, \\ & a'xy + c'y + b'x + d' = 0 \end{aligned}$$
>
> ne peuvent pas moins avoir lieu dans le cas de $x = -\frac{c}{a}$, en supposant y infinie ; parce que la première devient
>
> $$0y + bx + d = 0 \ ;$$

78. *Cf.* [29] § 284.

et la seconde à cause de $c' = \frac{a'c}{a}$, devient

$$0.\frac{a'}{a}y + b'x + d' = 0, \text{ ou } 0y + \frac{a}{a'}.(b'x + d') = 0,$$

chacune desquelles peut avoir lieu en supposant y infinie.

La réponse seroit, qu'il ne suffit pas que chaque équation soit satisfaite en supposant y infinie ; il faut encore que cet infini soit de même valeur pour chaque équation ; or pour la première il faudroit que $y = -\frac{(bx+d)}{0}$, et pour la seconde $y = -\frac{(b'x+d')a}{0.a'}$ valeurs qui diffèrent, même d'une quantité infinie, lorsque comme on le suppose ici, $x = -\frac{c}{a}$.

Il y a dans la Théorie des Equations beaucoup de cas semblables à celui que nous venons d'examiner, où chaque équation peut être satisfaite en supposant y infinie ; mais pour que cette valeur puisse être regardée comme appartenant à la question, il faut que cette valeur infinie soit la même pour chaque équation, ou du moins que d'une équation à l'autre elle ne diffère que d'une quantité finie. [79]

79. *Cf.* [29] § 283.

Chapitre 11

Naissance d'une théorie

Sous divers aspects, le traité de Bézout dépasse largement les œuvres antérieures qui pourtant l'annoncent. Il faut d'abord prendre acte d'une innovation technique : un judicieux argument de récurrence, impliquant le nombre des équations considérées, mais pas le nombre des inconnues. On élimine toutes les inconnues à la fois, et non plus l'une après l'autre – c'est par l'abandon de l'argument naïf de récurrence sur le nombre d'inconnues que Bézout parvient à franchir l'obstacle des « facteurs superflus » (*cf. supra* p. 139).

Nous avons déjà remarqué l'audace de l'application du calcul aux différences finies pour déterminer la dimension de certains espaces vectoriels lors de l'étude d'anneaux de polynômes et d'idéaux ; nous l'avons comparé à la théorie moderne du « polynôme de Hilbert » (*cf. supra* p. 161).

Nous avons peu insisté sur une évolution non moins importante : l'amélioration des techniques d'algèbre linéaire, due à Bézout et à ses contemporains. Le calcul effectif de l'équation finale dans un système à un nombre quelconque d'équations et d'inconnues se réduit toujours, chez Bézout, à résoudre un système d'équations linéaires. Le fait d'envisager des polynômes dont de nombreux coefficients sont nuls a des conséquences sur le système d'équations linéaires obtenu : en langage moderne, la matrice du système contiendra elle aussi beaucoup de zéros. D'où la spécificité des techniques d'algèbre linéaire développées par Bézout, particulièrement adaptées à ce type de problèmes. Il donne, par exemple, un développement des déterminants en produits de mineurs, apparenté au « développement de Laplace ». Il semble y avoir deux manières de considérer l'histoire de l'algèbre linéaire. Selon un premier point de vue, il s'agit d'un *cas particulier* de la théorie de l'élimination, développé ultérieurement, avec les exposés systématiques qu'ont pu en donner les auteurs du XIX^{e} siècle, au moyen de la théorie des déterminants. Ce point de vue est certainement exact, mais à l'opposé, les démonstrations du théorème de Bézout, pour deux équations chez Euler, pour un nombre quelconque chez Bézout, la solution du paradoxe de Cramer, l'utilisation des déterminants de Vandermonde par Lagrange puis par Gauss, n'auraient pu voir le jour sans que l'algèbre linéaire ait déjà atteint un développement assez poussé, comme *outil* au service de la théorie de l'élimination. Bien sûr, ce deuxième point de vue est absent des études classiques sur l'histoire de l'algèbre linéaire [1].

Enfin, la notion de généricité est une composante essentielle de la théorie de l'élimination naissante. Les équations génériques sont désormais considérées comme un cas particulier, le cas le plus simple de l'élimination. Cela inspire à Bézout une véritable classification des polynômes, quelque peu paradoxale. Elle est construite par généralisations successives, chaque classe englobant comme cas particulier la classe qui la précède ; pourtant la première classe (celle des « équations complètes », les équations génériques), la plus restreinte, est aussi en un certain sens la plus générale, puisque tout polynôme s'obtient par spécialisation à partir d'un polynôme de cette classe. La notion même d'équation générique appelle donc bien vite le procédé de « spécialisation d'indéterminées ». Bézout utilise ce procédé, on l'a vu, dans son mémoire de 1764 (*cf. supra* p. 123 ; la même idée est reprise au paragraphe § 422 du traité de 1779) ; il justifie seulement ce procédé par le fait que deux calculs

1. Comme celle de Thomas Muir, *The theory of determinants in the historical order of its developpement* (1890), qui donne une image inférieure à la réalité des travaux de Bézout en algèbre linéaire.

(celui portant sur les coefficients généraux, et celui portant sur les coefficients spécialisés) « donnent le même résultat littéral ». C'est suffisant dans ce cas. A la fin du XIX^{e} siècle, ce procédé devient incontournable (ainsi chez Eugen Netto qui l'utilise pour montrer que le système linéaire conduisant à l'équation finale est compatible, puis pour minorer le degré de l'équation finale, *cf. supra* p. 159). Mais entre Bézout et Netto, le procédé de spécialisation d'indéterminées n'a pu achever d'éclore sans l'apparition de la notion d'indépendance algébrique. Nous avons aussi vu Gauss utiliser la spécialisation, en 1815 ; il doit à cet effet démontrer que les fonctions symétriques de racines algébriquement indépendantes sont elles-mêmes algébriquement indépendantes. Il règne au contraire une certaine ambiguïté dans la notion de « fonction rationnelle des racines » chez Lagrange : celui-ci ne formule jamais explicitement l'hypothèse, parfois nécessaire dans ses raisonnements, que les seules relations algébriques entre les racines d'une équation générale sont les relations symétriques prévues par le théorème de Newton (*cf. supra* p. 136). Peut-être la prise de conscience de la notion d'indépendance algébrique, fait-elle écho à une autre prise de conscience, celle des relations de dépendance linéaire. On a vu qu'Euler, dès 1747, dans son mémoire sur le paradoxe de Cramer, a l'idée de compter, dans un système d'équations linéaires, le nombre de relations linéaires unissant les équations les unes aux autres. De même Bézout, en 1779, compte le nombre de « coefficients qui restent arbitraires » dans la solution générale d'un système linéaire, c'est-à-dire le nombre de paramètres indépendants, ou le rang de la solution.

Toutes ces caractéristiques et la place centrale qu'elles occupent dans le traité de Bézout font que l'on voudrait parler peut-être pour la première fois, sans arrière-pensée, d'une véritable « théorie de l'élimination ». Lagrange utilisait certes déjà cette expression en 1770 ; mais la spécialisation disciplinaire et la manière exclusive dont Bézout définit l'unique objet de son épais volume (*cf. supra* p. 149), premier traité qui lui soit entièrement consacré, contribuent à ériger en une théorie ce qui auparavant n'était que résultats de recherches isolées.

Pourtant l'on ne peut accorder au traité de 1779 un monopole exclusif sur la théorie de l'élimination puisque Bézout reste étranger, par principe, à la méthode de Cramer–Poisson. Il faut donc se poser la question du partage de la théorie de l'élimination entre différents styles de pensée mathématique au sein de cette vaste « analyse algébrique » du XVIII^{e} siècle. Nous avons déjà distingué deux voies distinctes (*cf. supra* 139), l'une illustrée par Cramer en 1750 et l'autre par Bézout en 1764. La première s'appuyait sur la considération des racines (sans s'interroger sur leur existence) de chacune des deux équations proposées, ce qui permit à Euler de calculer le degré de l'équation finale en 1748, puis à Cramer de calculer effectivement l'équation finale au moyen des fonctions symétriques des racines en 1750, et enfin à Lagrange de donner des critères pour que deux équations aient plusieurs racines communes (il avait été précédé en cela par la méthode n°3 d'Euler en 1764), et de reprendre la méthode de Cramer au moyen du développement en série du logarithme de l'équation finale. La seconde voie, héritée des travaux du XVII^{e}, fut expliquée de manière imparfaite par Euler (*cf.* sa méthode n°1), puis reprise par Bézout ; méthode fondée chez Newton sur l'idée de substitution (substituer aux puissances successives de l'inconnue leurs valeurs données par une autre équation), elle mit à profit l'algèbre linéaire, précocement chez Leibniz, systématiquement chez Bézout. Il serait erroné d'y voir une simple préférence pour des méthodes algébriques effectives. Ces deux voies ont produit deux démonstrations distinctes du théorème de Bézout. Mais ces deux démonstrations ne se valent pas entièrement. L'objet dont il est question dans ce théorème, l'équation finale résultant de l'élimination, est défini tantôt comme une équation de degré minimal de la forme $P_1R_1 + ... + P_nR_n = 0$, tantôt comme une équation telle qu'à chacune de ses racines corresponde une solution du système d'équations proposé. Bézout et Poisson n'ont donc pas, à proprement parler, démontré le même théorème. Ainsi, la démonstration de Bézout ne garantit pas qu'à chaque racine de l'équation finale corresponde au moins une solution commune aux équations proposées ; elle n'enseigne pas non plus le calcul effectif de la solution commune correspondant à une racine donnée de l'équation finale. Inversement, la démonstration de Poisson ne garantit pas que l'équation finale puisse s'obtenir par des « polynômes multiplicateurs ».

Lagrange semble cependant avoir une place particulière dans cette histoire. Il avait adopté, tels Hudde ou Rolle avant lui, l'algorithme euclidien comme méthode d'élimination. D'autre part, le recours aux fonctions symétriques des racines était pour lui un moyen de s'assurer *a priori* de la

réussite de certaines méthodes ; il n'ignorait ni ne rejetait, en pratique, la méthode de Cramer. Lagrange écrivit à Bézout, le 12 juillet 1779, pour le féliciter de son ouvrage (dont Bézout lui avait envoyé un exemplaire) et en particulier du livre I, « chef d'œuvre d'Analyse ». Il lui fit part cependant de l'objection mentionnée ci-dessus, en l'invitant à y trouver remède :

> Un avantage, particulier à votre méthode d'élimination et dont vous n'avez point parlé, consiste en ce qu'on peut avoir avec la même facilité l'expression la plus simple des valeurs des autres inconnues. En effet, si D est le degré de l'équation finale en x, et que, dans l'équation somme, on fasse disparaître la puissance x^D, en conservant à sa place le terme affecté de y ou de z,etc., on aura, pour la détermination de y en x, une équation de la forme
>
> $$ky + (x)^{D-1} = 0$$
>
> il peut arriver que le coefficient k soit nul, ce sera le symptôme de l'abaissement de l'équation en x ; dans ce cas, il faudra conserver dans l'équation somme le terme affecté de y^2, mais cela me paraît susceptible de plusieurs variétés qui mériteraient d'être discutées. Vous êtes, Monsieur, plus en état de remplir cet objet que personne, et j'ose vous y inviter, persuadé qu'il en pourra résulter un nouveau degré de perfection dans votre théorie. [2]

Lagrange donnait ensuite une ébauche de démonstration, qui ressemble beaucoup à celle que l'on trouve dans le traité de Netto [236], § 414. Netto attribuait l'idée d'une telle démonstration à Carl Schmidt (1886) :

> Bézout n'a pas remarqué la nécessité de démontrer cette proposition ; Liouville aussi l'a négligé ; Serret donne une démonstration, qui suppose déjà démontrée cette proposition, comme si de rien n'était, de manière cachée. Carl Schmidt en a donné une preuve, que nous exposerons avec quelques modifications [3].

Nous reproduisons la démonstration de Lagrange ci-dessous, marquant ainsi la rencontre encore inachevée entre ces deux versants de la théorie de l'élimination, dès 1779 :

> Ayant, par votre méthode, ces équations identiques
>
> $$\begin{aligned}
> &(x)^D = (x...n)^T(x...n)^t + (x...n)^{T'}(x...n)^{t'} + ...,\\
> &ky + (x)^{D-1} = ((x...n))^T(x...n)^t + ((x...n))^{T'}(x...n)^{t'} + ...,\\
> &...,\\
> &k'z + ((x))^{D-1} = ... ;
> \end{aligned}$$
>
> et ainsi de suite (les doubles crochets indiquant des polynômes de la même forme, mais avec des coefficients différents) ; il est clair que les équations proposées
>
> $$(x...n)^t = 0, \qquad (x...n)^{t'} = 0, \qquad ...$$
>
> donneront nécessairement celles-ci
>
> $$(x)^D = 0, \quad ky + (x)^{D-1} = 0, \quad k'z + ((x))^{D-1} = 0,$$
>
> qui sont le résultat de l'élimination ; mais, pour que ces dernières puissent être censées épuiser entièrement les premières, ou leur être tout à fait équivalentes, ne faudrait-il pas aussi que, les dernières étant posées, les premières s'ensuivissent nécessairement ? et cette condition ne demanderait-elle pas que, en éliminant des équations
>
> $$\begin{aligned}
> &0 = (x...n)^T(x...n)^t + (x...n)^{T'}(x...n)^{t'} + ...,\\
> &0 = ((x...n))^T(x...n)^t + ((x...n))^{T'}(x...n)^{t'} + ...,\\
> &...,
> \end{aligned}$$

2. *Cf.* Lettre de Lagrange à Bézout, Berlin, le 12 juillet 1779, *in* Œuvres de Lagrange, vol. 14 p. 276–278.

3. *Bézout hat die Nothwendigkeit dieses Nachweises nicht bemerkt ; auch Liouville hat sie übersehen ; Serret giebt einen Beweis, der freilich den Satz versteckter Weise schon als richtig voraussetzt. Herr C. Schmidt hat einen Beweis geliefert, den wir mit einigen nothwendig erscheinenden Änderungen folgen lassen wollen* (*cf.* [236] § 414).

> les quantités $(x...n)^t$, $(x...n)^{t'}$,..., regardées comme des inconnues particulières, l'équation de condition qui en résulterait entre les polynômes multiplicateurs $(x...n)^T$, $(x...n)^{T'}$,... ; $((x...n))^T$, $((x...n))^{T'}$,... ne pût jamais être satisfaite au moyen des mêmes équations

$$(x)^D = 0, \quad ky + (x)^{D-1} = 0, \quad k'z + ((x))^{D-1} = 0, \quad ...?$$

Traduisons en notations modernes ce que nous propose Lagrange dans le cas ($n = 3$) de trois équations

$$\begin{cases} f_1(u, y, z) = 0 \\ f_2(u, y, z) = 0 \\ f_3(u, y, z) = 0 \end{cases}$$

Lagrange cherche trois « équations-sommes » de la forme :

$$\begin{aligned} E(u) &= \phi_1 f_1 + \phi_2 f_2 + \phi_3 f_3 \\ y + F(u) &= \psi_1 f_1 + \psi_2 f_2 + \psi_3 f_3 \\ z + G(u) &= \chi_1 f_1 + \chi_2 f_2 + \chi_3 f_3 \end{aligned}$$

A chaque racine u_0 de E, on peut donc associer $y_0 = -F(u_0)$ et $z_0 = -G(u_0)$. Pour montrer que (u_0, y_0, z_0) est bien solution des trois équations initiales, il suffit de montrer que pour ces valeurs des inconnues le déterminant suivant ne s'annule pas :

$$\Delta(u, y, z) = \begin{vmatrix} \phi_1 & \phi_2 & \phi_3 \\ \psi_1 & \psi_2 & \psi_3 \\ \chi_1 & \chi_2 & \chi_3 \end{vmatrix}$$

Il faut donc montrer que $\Delta(u, -F(u), -G(u))$ ne s'annule pour aucune racine de E. Lagrange s'arrête là. Carl Schmidt démontrera en effet que $\Delta(u, -F(u), -G(u))$ n'est pas divisible par E qui est lui-même irréductible, *q. e. d.*

A ce stade, malgré l'ampleur du travail accompli par Bézout, Lagrange et Poisson, les problèmes conceptuels soulevés lors de l'application de l'élimination à la géométrie restent intacts : absence d'une définition exacte de la multiplicité d'intersection (nous l'avons vu chez Maclaurin et Euler), décompte problématique des « points à l'infini » (chez Bézout), notion de « point imaginaire » dépendante d'un concept non encore géométrisé de nombre complexe [4].

La théorie de l'élimination, aux XVII^e^ et XVIII^e^ siècles, s'est heurtée, conjointement, à des difficultés techniques et à des problèmes conceptuels. On aurait pu croire, à ses débuts, avec Faulhaber ou Leibniz, que la difficulté était purement technique, ou au mieux, *combinatoire* (*cf. supra* p. 71). La complexité des calculs d'élimination a pu ralentir l'histoire, et le problème n'était, en tout cas, pas à la portée de tous, ni même les résultats déjà atteints, comme le montrera encore au XIX^e^ siècle le jugement trop pessimiste d'Auguste Comte :

> Quelque pénibles que soient les méthodes à l'aide desquelles on surmonte cette difficulté, elles ne sont pas même applicables d'une manière entièrement générale, à l'élimination d'une inconnue entre deux équations de forme quelconque. [5]

Pourtant la difficulté technique n'a pu finalement ni masquer les problèmes conceptuels, ni empêcher le développement de la théorie ; elle a peut-être même, à certains moments, fait sentir d'autant plus l'urgence d'un nouvel éclairage conceptuel : pensons aux réflexions de Leibniz sur la nécessité de chercher une formule plutôt qu'un algorithme d'élimination, à celles de Cramer sur le « paradoxe », ou à celles de Bézout sur le problème des facteurs superflus [6].

4. Nous y reviendrons au moment d'étudier des travaux postérieurs à la naissance de l'analyse complexe *infra* section 13.2.

5. *Cf. Cours de philosophie positive*, tome premier (Rouen Frères, Paris, 1830), cinquième leçon, p. 208.

6. Sur les facteurs superflus, *cf. supra* n. 37 p. 98, p. 120 et p. 140.

11.1 Raisonnement *a priori* et calcul effectif

A la lecture du traité de Bézout, un fait nous frappe : tout procédé algorithmique est absent du livre I (il n'est même pas besoin de résoudre de manière effective un quelconque système linéaire). Bien qu'on y trouve des traces de la méthode newtonienne d'élimination successive des plus hautes puissances, ou de la méthode de substitution en algèbre linéaire (le « pivot de Gauss »), ces outils ne sont pas utilisés par Bézout comme méthodes effectives de calcul [7]. Le livre II est au contraire entièrement dédié au calcul effectif de l'« équation finale ». Cet ordre d'exposition était déjà présent dans son mémoire de 1764. La raison en était que la connaissance préalable du degré de l'équation finale servait à Bézout de « pierre de touche à l'aide de laquelle on peut juger du mérite des méthodes qu'on se propose d'employer pour éliminer ». Il s'agit d'un trait commun avec la méthode de Lagrange, dans un autre contexte : Lagrange a forgé des outils qui lui permettent de juger *a priori* du succès d'une méthode de résolution algébrique des équations, et même de formuler *a priori* toutes les méthodes possibles de résolution algébrique pour un degré donné [8]. De même Poisson écrira bientôt : « Je me proposais seulement de déterminer à priori, le degré de l'équation finale résultante d'un nombre quelconque d'équations complètes ; (...) dans la pratique, pour obtenir l'équation finale à son véritable degré, il faudra recourir aux méthodes que l'on trouve exposées avec beaucoup de détails dans la *Théorie des équations* de Bézout » [9]. Voici donc un trait distinctif des travaux de ces auteurs : ils tournent le dos au désir aveugle d'un algorithme efficace.

Le problème de l'élimination effective reste alors en suspens. Les chercheurs de la période suivante s'y intéresseront de nouveau (nous avons déjà mentionné les travaux d'Ossian Bonnet ; ceux de Cayley et Sylvester feront l'objet du chapitre suivant). La méthode enseignée par Bézout dans son livre II présente un inconvénient, dont nous avons vu un exemple (*supra* p. 170). Le système d'équations linéaires déterminant les coefficients des « polynômes-multiplicateurs » intervenant dans la représentation de l'équation finale sous forme d'« équation-somme » possède une infinité de solutions. Ces solutions sont paramétrées par les quantités que Bézout désignait sous le nom de « coefficients inutiles » ou « coefficients dont on dispose arbitrairement ». Mais Bézout ne détermine pas de manière effective un tel paramétrage. Il nous enjoint seulement à ajouter d'autres équations linéaires au système pour le rendre déterminé (*cf.* l'équation $Ca + C'a = 0$ dans l'exemple que nous avons mentionné), et il n'explique pas de manière satisfaisante comment choisir ces équations supplémentaires [10]. Cette manière de rendre le système déterminé, qui consiste à imposer des conditions supplémentaires aux polynômes-multiplicateurs, sera aussi adoptée par Macaulay en 1903, mais de manière plus circonspecte.

Soient $f_1, ..., f_l$ des polynômes génériques homogènes en n indéterminées $X_1, ..., X_n$, de degrés respectifs $m_1, ..., m_l$, avec $l \leq n$. Un monôme est dit *réduit* en X_i quand il n'est pas divisible par $X_i^{m_i}$ (donc l'ordre des indéterminées et celui des polynômes importent). Notons :

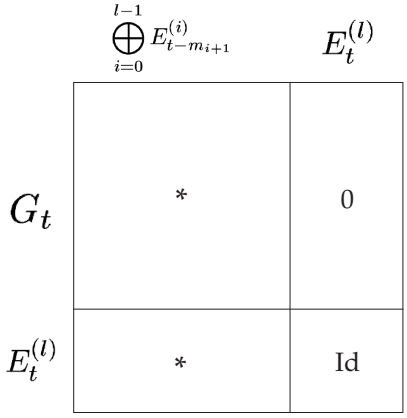
$$E_t^{(i)} = \langle \text{monômes de degré } t \text{ réduits en } X_1, ..., X_i \rangle \subset K[X_1, ..., X_n]_t$$

On pose de plus $E_t^{(0)} = K[X_1, ..., X_n]_t$. Quand $l = n$, soit $t_n = m_1 + ... + m_n - n + 1$. Alors $E_{t_n}^{(l)} = 0$. Macaulay impose aux polynômes-multiplicateurs d'appartenir à certains de ces espaces $E_t^{(i)}$. Plus précisément, il considère l'application vectorielle suivante :

$$\begin{array}{ccc} E_{t-m_1}^{(0)} \oplus E_{t-m_2}^{(1)} \oplus \ldots \oplus E_{t-m_l}^{(l-1)} \oplus E_t^{(l)} & \xrightarrow{\begin{pmatrix} f_1 \\ \vdots \\ f_l \\ 1 \end{pmatrix}} & k[X_1, ..., X_n]_t \\ (g_1, g_2, ..., g_l, h) & \longmapsto & h + \sum f_i g_i \end{array}$$

7. *Cf. supra* p. 151.
8. *Cf. supra* p. 135.
9. *Cf.* [252], reproduit en annexe B.
10. Nous ne prétendons pas cependant avoir entièrement épuisé les richesses du livre II du traité de Bézout dans notre commentaire.

Macaulay étudie la matrice de cette application (dans les bases constituées des monômes). Cette matrice est de la forme :



où $G_t = \langle$monômes de degré t non réduits en au moins l'une des variables$\rangle \subset K[X_1, ..., X_n]_t$. Notons avec Macaulay $D(n,t)$ le déterminant du bloc supérieur gauche. Quand $l = n$, un raisonnement analogue à celui utilisé pour justifier la méthode dialytique de Sylvester montre que le résultant des polynômes $f_1, ..., f_n$ divise $D(n, t_n)$; ce déterminant est en fait un mineur de la matrice à laquelle conduit la méthode dialytique, et c'est ainsi que Macaulay démontre la formule que nous avons mentionnée *supra* p. 148.

Il vérifie de plus que l'application vectorielle ci-dessus est un isomorphisme, et que cet isomorphisme identifie $(f_1, ..., f_l)_t$ au sous-espace vectoriel défini par $h \equiv 0$ (*cf.* [216] p. 7). Peut-être Macaulay n'est-il pas le premier à énoncer ce théorème, mais, étrangement, il croit même qu'il s'agit d'un théorème connu depuis très longtemps :

> Ce théorème est fondamental dans la méthode de Bézout, et il est probablement contenu dans sa *Théorie générale des Equations Algébriques* (Paris, 1779, 4to, 471 pp.), que je n'ai pas pu consulter. [11]

Dans le cas $l = 1$, $n = 2$, on a :

$$E^{(0)}_{t-m_1} = \langle\text{monômes de degré } t - m_1\rangle$$
$$E^{(1)}_t = \langle\text{monômes de degré } t \text{ non divisibles par } X_1^{m_1}\rangle$$

L'isomorphisme inverse $\begin{pmatrix} f_1 \\ 1 \end{pmatrix}^{-1} : f \in k[X_1, X_2]_t \longmapsto (g, h)$, tel que $f = f_1 g + h$, est simplement la division euclidienne de f par f_1, de quotient g et de reste h. Le reste h est de degré strictement inférieur à m_1 en X_1.

On trouve une idée semblable à celle de Macaulay dans la théorie moderne des « bases standards » ou « bases de Groebner ». Cette théorie, outre de nombreuses applications, offre justement une méthode effective d'élimination. Nous ne voulons pas ici décrire l'histoire, très récente, de cette théorie ; d'ailleurs, D. Eisenbud a déjà écrit quelques pages instructives sur ce sujet [12]. Il fait remonter le concept d'ordre monomial, fondamental dans cette théorie, à un article de Macaulay paru en 1927 ; il indique toutefois un exemple antérieur de base de Groebner dans un article de Gordan (1900), écrit en réaction aux méthodes nouvelles – non effectives – introduites par Hilbert en théorie des invariants. Mais la théorie ne commence vraiment que dans la deuxième moitié du siècle, avec les travaux de Groebner et de son élève Buchberger, et ceux, indépendants, de Hironaka. L'« idée » à laquelle nous avons fait allusion consiste en la généralisation du concept de division euclidienne. On l'appelle parfois aujourd'hui « division de Hironaka ». Soit l'idéal $I \subset K[X_1, ..., X_n]$ engendré par

11. *Cf.* [216] p. 6.
12. *Cf.* [100] p. 337–338. Eisenbud donne aussi des références sur la complexité algorithmique du calcul d'une base standard, p. 333–334.

des polynômes homogènes (mais *pas nécessairement* génériques) $f_1, ..., f_l$. Choisissons un « ordre monomial », c'est-à-dire un ordre total sur les monômes, compatible avec la multiplication. Pour tout polynôme $f \in K[X_1, ..., X_n]$ homogène, notons $\mathrm{in}f$ son « terme initial », c'est-à-dire le monôme dominant de f pour l'ordre monomial, affecté de son coefficient ; soit $\exp f$ le multi-indice tel que $\mathrm{in}f = cX^{\exp f}$ avec $c \in K$. Notons de plus pour tout $1 \leq i \leq l-1$:

$$E_t^{(i)} = \langle \text{monômes } M \text{ de degré } t \text{ tels que } M\mathrm{in}f_{i+1} \text{ ne soit divisible par aucun des monômes } \mathrm{in}f_1, ..., \mathrm{in}f_i \rangle$$

$E_t^{(0)} = K[X_1, ..., X_n]_t$
$E_t^{(l)} = \langle \text{monômes de degré } t \text{ divisibles par aucun des monômes } \mathrm{in}f_1, ..., \mathrm{in}f_l \rangle$

Alors on vérifie facilement que l'application vectorielle suivante est un isomorphisme :

$$E_{t-m_1}^{(0)} \oplus E_{t-m_2}^{(1)} \oplus \ldots \oplus E_{t-m_l}^{(l-1)} \oplus E_t^{(l)} \xrightarrow{\begin{pmatrix} \mathrm{in}f_1 \\ \vdots \\ \mathrm{in}f_l \\ 1 \end{pmatrix}} k[X_1, ..., X_n]_t$$

Il est moins évident que l'application vectorielle suivante en est un aussi :

$$\begin{array}{ccc} E_{t-m_1}^{(0)} \oplus E_{t-m_2}^{(1)} \oplus \ldots \oplus E_{t-m_l}^{(l-1)} \oplus E_t^{(l)} & \xrightarrow{\begin{pmatrix} f_1 \\ \vdots \\ f_l \\ 1 \end{pmatrix}} & k[X_1, ..., X_n]_t \\ (g_1, g_2, ..., g_l, h) & \longmapsto & h + \sum f_i g_i \end{array}$$

Démonstration. Notons

$$\phi = \begin{pmatrix} f_1 \\ \vdots \\ f_l \\ 1 \end{pmatrix}, \quad \mathrm{in}\phi = \begin{pmatrix} \mathrm{in}f_1 \\ \vdots \\ \mathrm{in}f_l \\ 1 \end{pmatrix}$$

Soit $y \in K[X_1, ..., X_n]_t$. Soit $M_0 \in \left(\bigoplus_i E_{t-m_{i+1}}^{(i)} \right) \oplus E_t^{(l)}$ le monôme tel que $\mathrm{in}\phi(M_0) = \mathrm{in}y$. Alors on a (pour l'ordre monomial) :

$$\mathrm{in}(y - \phi(M_0)) < \mathrm{in}y$$

Soit de même le monôme M_1 tel que $\mathrm{in}\phi(M_1) = \mathrm{in}(y - \phi(M_0))$. Alors

$$\mathrm{in}(y - \phi(M_0 + M_1)) < \mathrm{in}(y - \phi(M_0))$$

De proche en proche, on construit ainsi un antécédent $M_0 + M_1 + ...$ de y par ϕ. Cette application est donc surjective. Comme c'est une application entre deux espaces vectoriels de même dimension finie, l'injectivité en résulte aussitôt. *q. e. d.*

Cette application est donc inversible, mais ici (contrairement à la situation étudiée par Macaulay), l'image inverse de I_t n'est pas nécessairement incluse dans le sous-espace défini par $h \equiv 0$. Il faut, pour garantir cela, imposer certaines conditions sur les générateurs $f_1, ..., f_l$ de I. Notons $\mathrm{in}I$ l'idéal engendré par tous les termes initiaux des éléments de I (c'est l'« idéal initial » de I), et supposons en effet que $\mathrm{in}f_1, ..., \mathrm{in}f_l$ engendrent $\mathrm{in}I_t$. Il est facile de vérifier que l'application suivante est surjective :

$$\bigoplus_{i=0}^{l-1} E_{t-m_{i+1}}^{(i)} \xrightarrow{\begin{pmatrix} \mathrm{in}f_1 \\ \vdots \\ \mathrm{in}f_l \end{pmatrix}} (\mathrm{in}f_1, ..., \mathrm{in}f_l)_t = \mathrm{in}I_t$$

Une démonstration semblable à celle qui précède permet d'en déduire aussi la surjectivité de l'application :

$$\bigoplus_{i=0}^{l-1} E^{(i)}_{t-m_{i+1}} \xrightarrow{\begin{pmatrix} f_1 \\ \vdots \\ f_l \end{pmatrix}} I_t$$

Lorsque $\mathrm{in}f_1, ..., \mathrm{in}f_l$ suffisent ainsi à engendrer toutes les composantes homogènes de $\mathrm{in}I$, on dit que $f_1, ..., f_l$ constituent une « base de Groebner » ou « base standard » de I. A cause du caractère noethérien de l'anneau de polynômes, il existe toujours une telle base finie. Dans ce cas, l'isomorphisme que l'on vient de décrire généralise la division euclidienne (cas $l = 1$) ; la notion de base de Groebner généralise aussi celle de système « triangulaire » d'équations linéaires (cas ou $\deg f_1 = ... = \deg f_l = 1$). Enfin, la division de Hironaka est facile à calculer de manière effective.

Mais il faut encore décrire le moyen effectif de déterminer une telle base standard d'un idéal donné. Buchburger a donné un critère qui permet de rendre algorithmique la recherche d'une base standard. Pour tout polynôme f, notons

$$f\mathcal{R}f_1, ..., f_l$$

le reste h de la division de Hironaka de f par $f_1, ..., f_l$. Pour tous polynômes f, g, notons de plus

$$\exp f \vee \exp g = (\max\{(\exp f)_1, (\exp g)_1\}, ..., \max\{(\exp f)_n, (\exp g)_n\})$$

et

$$f\mathcal{S}g = X^{\exp f \vee \exp g - \exp f - \exp g}\,(f\mathrm{in}g - g\mathrm{in}f)$$

Alors $f_1, ..., f_l$ est une base standard si et seulement si [13]

$$(\forall i, j) \quad (f_i\mathcal{S}f_j)\mathcal{R}f_1, ..., f_l = 0$$

En partant d'un système $f_1, ..., f_l$ de générateurs de I, on peut donc concevoir l'opération consistant à ajouter à ce système tous les polynômes de la forme $(f_i\mathcal{S}f_j)\mathcal{R}f_1, ..., f_l$. La clôture, pour cette opération, d'un système quelconque de générateurs de I est alors une base standard de I.

L'application de cette théorie à l'élimination demande de choisir un ordre monomial adéquat. On peut ainsi faire en sorte que tous les éléments f_i d'une base standard de I qui ne dépendent pas des $(n - e)$ dernières variables $X_{e+1}, ..., X_n$ forment une base standard de l'idéal $I \cap K[X_1, ..., X_e]$. Rappelons d'ailleurs que $I \cap K[X_1, ..., X_e]$ est l'idéal de la projection de la variété des zéros de I sur le sous-espace linéaire $\{X_{e+1} = ... = X_n = 0\}$ *à condition que* la variété des zéros de I ne coupe pas le sous-espace linéaire [14] $\{X_1 = ... = X_e = 0\}$. Or, un autre ordre monomial permet aussi de calculer une base standard de l'idéal $I + (X_1, ..., X_e)$.

13. *Cf.* démonstration dans la thèse de Bayer [14] p. 14. Comme nous l'a indiqué M. Martin-Deschamps, il y a cependant une lacune dans cette démonstration ; le résultat est exact.

14. On le démontre en identifiant le complémentaire de $\{X_1 = ... = X_e = 0\}$ à $\mathbb{P}^{e-1} \times \mathbb{A}^{n-e}$ et en remarquant que la projection $\mathbb{P}^{e-1} \times \mathbb{A}^{n-e} \to \mathbb{A}^{n-e}$ est fermée. On a déjà utilisé un raisonnement semblable *supra* section 6.2.

Troisième partie

De Jacobi à Kronecker

Chapitre 12

Jacobi et Sylvester

12.1 Le bézoutiant de Sylvester

La matrice de Bézout connaît, vers le milieu de XIX$^{\text{ème}}$ siècle, un nouveau succès, auprès des tenants de la théorie des invariants. Cayley, Sylvester, Salmon, Hermite, Jordan, Brioschi, et en Allemagne Hesse puis bientôt Aronhold, Clebsch, Gordan, de 1841 (date de publication des premiers articles de Boole, qui inspirent bientôt les travaux de Cayley) à 1888 (quand Hilbert introduit des méthodes non constructives en théorie des invariants, pour démontrer le « théorème des syzygies de Hilbert », commençant en quelque sorte l'« algèbre commutative » moderne), vont rivaliser à la recherche d'une méthode efficace de calcul d'invariants, efforts culminant au moment où Gordan fera triompher la méthode symbolique allemande en donnant la première étude complète des invariants des formes binaires. Jusqu'à cette date, la « théorie des invariants » est-elle d'ailleurs vraiment théorie ? N'est-elle pas plus simplement un « calcul des invariants », et la recherche d'algorithmes, guidée par l'induction à partir de méthodes éprouvées au calcul des invariants de formes binaires de bas degré ? Pas seulement pourtant, dans la mesure justement où les premiers « invariantistes » Cayley et Sylvester avaient des vues suffisamment larges [1] pour formuler les questions qu'ils se posaient en termes généraux, et ainsi forger des concepts (tels ceux de covariant et contravariant, ou encore de « combinant » – il en sera question plus loin), un vocabulaire (mais combien abscons pour le lecteur moderne, il n'est besoin pour s'en convaincre que de parcourir le glossaire dressé par Sylvester à la fin de [306]), et sinon une théorie, du moins le plan d'une théorie future.

Le corpus témoignant de la genèse de cette théorie n'est pas d'abord facile [2]. Rappelons simplement, à titre d'exemple, les étapes qui menèrent à la connaissance complète des invariants d'une forme binaire quartique (c'est-à-dire un polynôme homogène de degré 4 en deux indéterminées) [3].

1. M. Noether a bien décrit ce caractère, dans sa biographie de Cayley :

   > Ce que l'on dit des anglais : « qu'ils sont des esprits endurants, dont l'être s'élève à partir de la succession et de l'embrassement d'un vaste ensemble de faits étroits (*deren Wesen auf Aneinanderreihen und Umspannen einer Menge kleiner Thatsachen aufgeht*), pour y trouver une loi – les meilleurs d'entre eux sont des cerveaux aristotéliciens... », s'applique pleinement à Cayley : ce n'est pas dans son style d'analyser les concepts jusque dans leurs derniers éléments, ni de construire un système à partir de peu d'hypothèses, ni de donner à ses développements une forme valable en général et recélant en soi tous les détails. Mais il est maître dans l'usage empirique du matériel. Il le rassemble en une pensée abstraite, qu'il généralise et soumet à l'épreuve du calcul. Puis, des données obtenues, l'idée générale apparaît d'un coup, et des années de travail seront consacrées à sa vérification par le calcul. Cayley est donc le physicien parmi les mathématiciens (*der Naturforscher unter den Mathematikern*). ([242], p. 479).

   Sylvester vénérait son collègue et ami Cayley comme son aîné en mathématiques bien qu'il fût son cadet en âge. En 1885, il écrivait :

   > ...though younger than myself, Cayley is my spiritual progenitor – who first opened my eyes and purged them of the dross so that they could see and accept the higher mysteries of our common mathematical faith... (*cf.* [308], t. 4, p. 300).

2. Dieudonné a donné un exposé moderne de la théorie classique des invariants dans [95], chapitre 1 et 2.
3. L'histoire de l'utilisation de ces invariants dans l'étude des intégrales elliptiques est exposée dans [157] p. 95–96.

Boole, après avoir remarqué l'invariance du discriminant dans ses premiers travaux, calcule le discriminant de la quartique binaire et en donne l'expression explicite en 1844 (*cf.* [34], p. 171). A la même époque, dans un court article en allemand [101] sur la résolution algébrique des équations de degré inférieur ou égal à 4, Eisenstein utilise les invariants de degrés 2 et 3 d'une quartique binaire, sans toutefois remarquer leur caractère invariant. Sylvester estimait grandement cette découverte, et la décrivit lors d'une allocution publique en ces termes évocateurs de l'ambiance de ses recherches :

> Now this gigantic outcome [la « Nouvelle algèbre », c'est-à-dire la théorie des invariants] of modern analytical thought, itself, too, only the precursor and progenitor of a future still more heaven-reaching theory, which will comprise a complete study of the inter-operation, the actions and reactions, of algebraic forms (Analytical Morphology in its absolute sense), how did this originate ? In the accidental observation by Eisenstein, some score or more years ago, of a single invariant (the Quadrinvariant of a Binary Quartic) which he met with in the course of certain researches just as accidentally and unexpectedly as M. Du Chaillu might meet a Gorilla in the country of the Fantees, or any one of us in London a White Polar Bear escaped from the Zoological Gardens. Fortunately he pounced down upon his prey and preserved it for the contemplation and study of future mathematicians. It occupies only part of a page in his collected posthumous works. This single result of observation (...), which remained unproductive in the hands of its distinguished author, has served to set in motion a train of thought and to propagate an impulse which have led to a complete revolution in the whole aspect of modern analysis, and whose consequences will continue to be felt until Mathematics are forgotten and British Associations meet no more. [4]

En 1845, dans un article célèbre [56], Cayley expose une méthode symbolique pour calculer les invariants de formes multilinéaires, au moyen d'une vaste généralisation du concept de déterminant, les « hyperdéterminants ». Il tente d'en tirer les fruits pour le calcul des invariants des formes binaires (en cherchant les invariants d'une forme binaire de degré n au moyen des invariants d'une forme n-linéaire sur un espace de dimension 2). Ainsi, il parvient en particulier à retrouver l'invariant Π de degré 2 d'une quartique binaire. A la suite de la rédaction de cet article, Boole lui apprend [5] l'existence d'un invariant T de degré 3 de la quartique binaire, et que le discriminant s'écrit alors sous la forme $\Pi^3 - 27T^2$. Dans une courte note [55] en français au Journal de Crelle, Cayley retrouve les invariants Π et T dans un autre contexte. Il cherchait alors à généraliser la formule $\psi = df + \delta\phi$ due à Hesse, où f est une forme ternaire cubique, ϕ son hessien, et ψ le hessien de ϕ, et il trouvait alors une formule analogue pour les formes binaires de degrés quelconques. Nous mentionnons cet épisode parce qu'il nous montre que Cayley connaissait bien, à cette époque, le mémoire [148] de Hesse si important dans l'histoire de la théorie de l'élimination. Enfin, en 1846, dans le « Mémoire sur les hyperdéterminants » paru dans le Journal de Crelle [57] et qui se veut version française de ses deux premiers articles sur la théorie des invariants (« On the theory of linear transformations » [56] et « On linear transformations »), Cayley confesse que la méthode mentionnée ci-dessus ne permet pas de calculer T, car la forme quadrilinéaire associée – dans le symbolisme de Cayley – à la forme quartique binaire n'a pas d'invariant de degré 2 (mais bien un invariant de degré 4, qui permet au moins d'obtenir T^2). La situation se complique encore à cause d'une légère erreur, qui donnera prise aux critiques rigoureuses de Schläfli en 1851 : l'invariant de degré 6 d'une forme quadrilinéaire, duquel Cayley déduit l'expression du discriminant d'une quartique, n'est pas, contrairement à son affirmation (*cf.* [57], p. 18), le discriminant de la forme quadrilinéaire (Schläfli montre que le discriminant de la forme quadrilinéaire est en fait de degré 24). Mais cette erreur est sans conséquence. Et cette méthode n'est certes pas la seule proposée par Cayley. Dans la deuxième partie de son mémoire, il montre comment calculer T au moyen d'un autre outil symbolique [6], qui a gardé le nom de « procédé Ω ».

Si nous sommes entrés dans cette longue digression, c'était pour mieux comprendre l'intérêt

4. *Cf.* [295], p. 656.
5. Cayley raconte cela à la fin de [56].
6. *Cf.* [57] p. 22–37. Le procédé Ω n'est qu'un cas particulier de ce que Dieudonné appelle « Aronhold's Rule », *cf.* [95] p. 33.

des mathématiciens de cette époque vis-à-vis des formes binaires, en replaçant leurs travaux dans le cadre de la théorie des invariants naissante. Certes, pour raconter la genèse de cette théorie, il faudrait, sur le modèle adopté ci-dessus pour les invariants de la quartique, suivre l'usage des méthodes dans le calcul des invariants de la cubique, puis de la quintique, puis leurs covariants respectifs, avant d'aborder le cas des formes ternaires cubiques, d'abord chez Hesse et Cayley, cas pour lequel Aronhold introduisit les méthodes qui feraient finalement le succès de l'école allemande entre les mains de Gordan. La théorie de l'élimination semble s'être développé sur un patron moins morcelé. Le résultant de deux formes binaires, et son expression par le bézoutiant, formule si efficace dans les calculs, se voient réserver une place de choix dans les travaux des invariantistes, et tout particulièrement de Sylvester, et l'on dispose ainsi très tôt d'une théorie complète du résultant de deux formes binaires de degré quelconque. Boole s'était déjà intéressé à l'élimination. Dans [33], il étudie l'invariance par transformations linéaires du discriminant d'une équation. Il formule à cet effet une méthode d'élimination pour un nombre d'équations supérieur à 2, qui semble cependant naïve et est erronée : elle introduit de gros facteurs superflus. Mais cela est sans conséquence sur le reste de son mémoire (Boole renvoie certes à Sylvester les lecteurs interessés par cette question) [7]. L'importance que Boole accordait à la théorie de l'élimination est encore visible dans son « algèbre de la logique », où elle permet de formaliser les syllogismes, préfigurant la reprise du concept d'élimination dans les recherches en logique et fondements des mathématiques au vingtième siècle. Cayley lui-même cherche un lien entre la théorie des invariants et l'élimination. Dans son « Mémoire sur les hyperdéterminants », il écrit en effet :

> On serait probablement conduit à des résultats importants en cherchant d'établir une connexion entre la théorie des hyperdéterminants et celle de l'élimination, je n'ai rien encore trouvé de satisfaisant sur ce sujet. [8]

Il mentionne le problème du calcul du résultant de deux formes binaires U, V, et il indique qu'il a communiqué avec Boole sur ce sujet. En posant $W = \lambda U + \mu V$, il donne une expression du résultant de deux formes binaires quadratiques dans son symbolisme :

$$\overline{12}^2.[(\overline{12}^2.WW)(\overline{12}^2.WW)]$$

où le premier symbole $\overline{12}^2$ signifie $\frac{\partial^2}{\partial\lambda_1\partial\mu_2} - \frac{\partial^2}{\partial\lambda_2\partial\mu_1}$ et les suivants $\frac{\partial^2}{\partial x_1\partial y_2} - \frac{\partial^2}{\partial x_2\partial y_1}$. Poursuivons les calculs de Cayley quelque peu :

$$\overline{12}^2.WW = (2ac - 2b^2)\lambda^2 + (2a'c + 2ac' - 4b'b)\mu\lambda + (2a'c' - 2b'^2)\mu^2$$

D'où :

$$\overline{12}^2.[(\overline{12}^2.WW)(\overline{12}^2.WW)] = 8.\begin{vmatrix} 2ac - 2b^2 & a'c + ac' - 2b'b \\ a'c + ac' - 2b'b & 2a'c' - 2b'^2 \end{vmatrix}$$

Autrement dit, Cayley représente le résultant comme discriminant d'une certaine forme quadratique. Et il ajoute : « cela doit se rattacher à quelque théorie générale que j'ignore » [9]. Sylvester est parvenu par la suite, suivant « une idée de Cayley » nous dit-il dans son mémoire « On a theory of the syzygetic relations of two rational integral functions » (1853), à exprimer le résultant de deux formes binaires f, ϕ de degré m quelconque comme discriminant d'une certaine forme quadratique [10]. Soit $\Omega(D_1, D_2) = \sum_{i=0}^{m-1} (-1)^i \binom{m}{i} u_{i+1} D_1^i D_2^{m-1-i}$ où D_1 et D_2 sont des opérateurs différentiels quelconques, et $u_1, ..., u_m$ des indéterminées ; soit encore

$$\theta = \frac{f(x,y)\phi(x',y') - f(x',y')\phi(x,y)}{xy' - x'y}$$

7. La méthode d'élimination de Boole est exposée [33] p. 5–7. Il élimine les inconnues les unes après les autres ; à chaque étape, si l'on a m équations (1), (2),..., (m), on élimine l'une des inconnues entre (1) et (2), entre (2) et (3), entre (3) et (4), etc. pour obtenir ainsi $(m-1)$ équations, et ainsi de suite. Boole fait référence à Sylvester en [33] p. 119.

8. *Cf.* [57], p. 37.

9. *Cf.* [57], p. 25.

10. Cayley publie son « idée » au Journal de Crelle en 1857, dans une note [65] datée de 1855. Voir aussi la note [67] de 1861 qui généralise cette méthode.

Alors l'expression :

$$\Omega(\frac{\partial}{\partial y'}, -\frac{\partial}{\partial x'})\Omega(\frac{\partial}{\partial y}, -\frac{\partial}{\partial x})\theta$$

est une forme quadratique en les indéterminées $u_1, ..., u_m$, le « bézoutiant », dont la matrice est la matrice de Bézout de f, ϕ, et son déterminant est le résultant de f, ϕ. D'autre part, il est facile de voir que, par construction, cette forme quadratique est un covariant de f, ϕ. En résulte immédiatement l'invariance du résultant [11]. Mais la propriété d'invariance du résultant résulte [12], plus simplement, de la méthode dialytique que nous avons décrite section 10.1.

Le bézoutiant constitue l'objet principal du grand mémoire de Sylvester « On a theory of the syzygetic relations... » que nous allons commenter à présent. Ce mémoire en cinq sections s'ouvre (section I) sur une étude de l'algorithme euclidien. Sylvester y démontre que le résultant et tous les restes intermédiaires obtenus au moyen de l'algorithme euclidien peuvent aussi être calculés par la méthode des coefficients indéterminés, ou bien encore en échelonnant la matrice de Bézout (par le pivot de Gauss). Pour ce faire, il étudie les propriétés d'homogénéité du résultant et des restes intermédiaires. Soient les deux polynômes génériques P et Q de degrés respectifs m et $n < m$ en l'inconnue x. La méthode des coefficients indéterminés consiste ici à chercher des polynômes R_ι, Q_ι, P_ι de degrés en x respectivement $(n-\iota)$, $(m-n+\iota-1)$ et $(\iota-1)$ vérifiant l'« équation syzygétique »

$$(1) \quad R_\iota = Q_\iota Q + P_\iota P$$

On obtient ainsi (comme le faisait Bézout) un système d'équations linéaires, que Sylvester suppose résolu au moyen des formules de Cramer (cette méthode n'est en fait qu'un usage particulier de la méthode que Sylvester qualifiait de « dialytique »). Il en déduit les propriétés d'homogénéité suivantes [13] : R_ι est de degré ι en les coefficients de P et $(m-n+\iota)$ en ceux de Q. Sylvester appelle les R_ι « résidus simplifiés » et il est amené dans ses démonstrations, de manière plus ou moins implicite, à les supposer irréductibles. Quand $m = n$, il remarque qu'en échelonnant la matrice de Bézout de P et Q (par le pivot de Gauss), la ι-ème ligne de la matrice obtenue fournit à son tour les coefficients d'un polynôme R_ι de degré $(n-\iota)$ en x, ι en les coefficients de P et ι aussi en ceux de Q, et vérifiant une relation (1). Alors, les R_ι ainsi obtenus sont égaux aux « résidus simplifiés » à un coefficient numérique près [14]. Par contre, les restes intermédiaires obtenus par l'algorithme euclidien du plus grand commun diviseur ne sont égaux aux « résidus simplifiés » qu'à un facteur près *dépendant des coefficients de P et Q*. On retrouve le problème des « facteurs superflus », que Sylvester dénomme *allotrious factors*. Il est le premier, à notre connaissance, à donner une étude complète des facteurs superflus introduits par l'algorithme euclidien [15]. Nous allons donc décrire sa méthode, et nous l'appliquerons ensuite à une autre situation : l'étude des facteurs superflus dans l'algorithme d'élimination des plus hautes puissances de l'inconnue qu'utilisait Newton.

12.2 Les facteurs superflus dans l'algorithme euclidien

Dans ce paragraphe, nous gardons des notations proches de celles de Sylvester, mais nous détaillons son raisonnement afin de le rendre plus aisément compréhensible. Soient P et Q des polynômes en x :

$$\begin{cases} P = ax^m + bx^{m-1} + ... \\ Q = \alpha x^n + \beta x^{n-1} + ... \end{cases}$$

Notons R_ι le ι-ème « résidu simplifié » :

$$R_\iota = \alpha_\iota x^{n-\iota} + \beta_\iota x^{n-\iota-1} + ...$$

11. *Cf.* [306], § 61*, 62, 63.
12. *Cf.* [305] dernier paragraphe p. 291 ; p. 346.
13. *Cf.* [306], § 2.
14. *Cf.* [306], § 5.
15. *Cf.* [306], § 3, 4.

Par la méthode des coefficients indéterminés, on trouve l'expression suivante du premier résidu simplifié :

$$R_1 = \begin{vmatrix} \alpha & & a \\ \beta & \alpha & b \\ \gamma & \beta & c \end{vmatrix} x^{n-1} + \begin{vmatrix} \alpha & & a \\ \beta & \alpha & b \\ \delta & \gamma & d \end{vmatrix} x^{n-2} + \dots$$

On va d'abord montrer que les autres résidus simplifiés s'obtiennent au moyen d'une formule de récurrence :

Lemme 1 *Pour $\iota \geq 0$, on a (à un facteur numérique près)*

$$R_{\iota+2} = \frac{1}{\alpha_\iota^2} \left(\begin{vmatrix} \alpha_{\iota+1} & & \alpha_\iota \\ \beta_{\iota+1} & \alpha_{\iota+1} & \beta_\iota \\ \gamma_{\iota+1} & \beta_{\iota+1} & \gamma_\iota \end{vmatrix} x^{n-\iota-2} + \begin{vmatrix} \alpha_{\iota+1} & & \alpha_\iota \\ \beta_{\iota+1} & \alpha_{\iota+1} & \beta_\iota \\ \delta_{\iota+1} & \gamma_{\iota+1} & \delta_\iota \end{vmatrix} x^{n-\iota-3} + \dots \right)$$

Démonstration. Définissons les R_ι par la formule donnée dans l'énoncé du lemme, et démontrons que ce sont bien des fonctions entières en les coefficients de P, Q et vérifiant les propriétés d'homogénéité et de degré des « résidus simplifiés ». Supposons que ce soit le cas pour R_ι et $R_{\iota+1}$. On vérifie aisément que $R_{\iota+2}$ est alors de degré $(\iota+2)$ en les coefficients de P et $(m-n+\iota+2)$ en ceux de Q. Pour vérifier qu'il s'agit d'une fonction entière, il faut démontrer que α_ι^2 divise chaque coefficient de

$$\begin{vmatrix} \alpha_{\iota+1} & & \alpha_\iota \\ \beta_{\iota+1} & \alpha_{\iota+1} & \beta_\iota \\ \gamma_{\iota+1} & \beta_{\iota+1} & \gamma_\iota \end{vmatrix} x^{n-\iota-2} + \begin{vmatrix} \alpha_{\iota+1} & & \alpha_\iota \\ \beta_{\iota+1} & \alpha_{\iota+1} & \beta_\iota \\ \delta_{\iota+1} & \gamma_{\iota+1} & \delta_\iota \end{vmatrix} x^{n-\iota-3} + \dots$$

Considérons par exemple le premier coefficient, dans lequel nous développons les quantités $\alpha_{\iota+1}$, $\beta_{\iota+1}$,... :

$$\begin{vmatrix} \frac{1}{\alpha_{\iota-1}^2} \begin{vmatrix} \alpha_\iota & & \alpha_{\iota-1} \\ \beta_\iota & \alpha_\iota & \beta_{\iota-1} \\ \gamma_\iota & \beta_\iota & \gamma_{\iota-1} \end{vmatrix} & & & \alpha_{\iota+1} \\ \frac{1}{\alpha_{\iota-1}^2} \begin{vmatrix} \alpha_\iota & & \alpha_{\iota-1} \\ \beta_\iota & \alpha_\iota & \beta_{\iota-1} \\ \delta_\iota & \gamma_\iota & \delta_{\iota-1} \end{vmatrix} & \frac{1}{\alpha_{\iota-1}^2} \begin{vmatrix} \alpha_\iota & & \alpha_{\iota-1} \\ \beta_\iota & \alpha_\iota & \beta_{\iota-1} \\ \gamma_\iota & \beta_\iota & \gamma_{\iota-1} \end{vmatrix} & & \beta_{\iota+1} \\ \frac{1}{\alpha_{\iota-1}^2} \begin{vmatrix} \alpha_\iota & & \alpha_{\iota-1} \\ \beta_\iota & \alpha_\iota & \beta_{\iota-1} \\ \epsilon_\iota & \delta_\iota & \epsilon_{\iota-1} \end{vmatrix} & \frac{1}{\alpha_{\iota-1}^2} \begin{vmatrix} \alpha_\iota & & \alpha_{\iota-1} \\ \beta_\iota & \alpha_\iota & \beta_{\iota-1} \\ \delta_\iota & \gamma_\iota & \delta_{\iota-1} \end{vmatrix} & & \gamma_{\iota+1} \end{vmatrix}$$

Sylvester a montré de deux manières (par un calcul explicite[16], ou bien par un raisonnement *a priori*[17]) que α_ι^2 divise bien le déterminant :

$$\begin{vmatrix} \begin{vmatrix} \alpha_\iota & & \alpha_{\iota-1} \\ \beta_\iota & \alpha_\iota & \beta_{\iota-1} \\ \gamma_\iota & \beta_\iota & \gamma_{\iota-1} \end{vmatrix} & & & \alpha_{\iota+1} \\ \begin{vmatrix} \alpha_\iota & & \alpha_{\iota-1} \\ \beta_\iota & \alpha_\iota & \beta_{\iota-1} \\ \delta_\iota & \gamma_\iota & \delta_{\iota-1} \end{vmatrix} & \begin{vmatrix} \alpha_\iota & & \alpha_{\iota-1} \\ \beta_\iota & \alpha_\iota & \beta_{\iota-1} \\ \gamma_\iota & \beta_\iota & \gamma_{\iota-1} \end{vmatrix} & & \beta_{\iota+1} \\ \begin{vmatrix} \alpha_\iota & & \alpha_{\iota-1} \\ \beta_\iota & \alpha_\iota & \beta_{\iota-1} \\ \epsilon_\iota & \delta_\iota & \epsilon_{\iota-1} \end{vmatrix} & \begin{vmatrix} \alpha_\iota & & \alpha_{\iota-1} \\ \beta_\iota & \alpha_\iota & \beta_{\iota-1} \\ \delta_\iota & \gamma_\iota & \delta_{\iota-1} \end{vmatrix} & & \gamma_{\iota+1} \end{vmatrix}$$

Pour conclure, il suffit donc de remarquer que α_ι et $\alpha_{\iota-1}$ sont premiers entre eux. *q. e. d.*

16. *Cf.* [306], p. 441–442.
17. *Cf.* [306], p. 440.

Notons à présent λ_ι le ι-ème facteur superflu. Le ι-ème « résidu complet » est $\lambda_\iota R_\iota$. L'algorithme euclidien détermine $\lambda_{\iota+2}R_{\iota+2}$ comme reste de la division de $\lambda_\iota R_\iota$ par $\lambda_{\iota+1}R_{\iota+1}$. Notons $(px+q)$ le quotient de cette division. Le lemme va permettre d'établir une relation de récurrence entre les facteurs superflus (à un facteur numérique près) :

$$\begin{aligned}
&\lambda_{\iota+2}R_{\iota+2} = \lambda_\iota R_\iota - (px+q)\lambda_{\iota+1}R_{\iota+1}\\
&= \tfrac{1}{\lambda_{\iota+1}^2\alpha_{\iota+1}^2}\left(\begin{vmatrix}\lambda_{\iota+1}\alpha_{\iota+1} & & \lambda_\iota\alpha_\iota\\ \lambda_{\iota+1}\beta_{\iota+1} & \lambda_{\iota+1}\alpha_{\iota+1} & \lambda_\iota\beta_\iota\\ \lambda_{\iota+1}\gamma_{\iota+1} & \lambda_{\iota+1}\beta_{\iota+1} & \lambda_\iota\gamma_\iota\end{vmatrix} x^{n-\iota-2} + \begin{vmatrix}\lambda_{\iota+1}\alpha_{\iota+1} & & \lambda_\iota\alpha_\iota\\ \lambda_{\iota+1}\beta_{\iota+1} & \lambda_{\iota+1}\alpha_{\iota+1} & \lambda_\iota\beta_\iota\\ \lambda_{\iota+1}\delta_{\iota+1} & \lambda_{\iota+1}\gamma_{\iota+1} & \lambda_\iota\delta_\iota\end{vmatrix} x^{n-\iota-3} + ...\right)\\
&= \tfrac{\lambda_\iota}{\alpha_{\iota+1}^2}\left(\begin{vmatrix}\alpha_{\iota+1} & & \alpha_\iota\\ \beta_{\iota+1} & \alpha_{\iota+1} & \beta_\iota\\ \gamma_{\iota+1} & \beta_{\iota+1} & \gamma_\iota\end{vmatrix} x^{n-\iota-2} + \begin{vmatrix}\alpha_{\iota+1} & & \alpha_\iota\\ \beta_{\iota+1} & \alpha_{\iota+1} & \beta_\iota\\ \delta_{\iota+1} & \gamma_{\iota+1} & \delta_\iota\end{vmatrix} x^{n-\iota-3} + ...\right)\\
&= \tfrac{\lambda_\iota\alpha_\iota^2}{\alpha_{\iota+1}^2}R_{\iota+2}
\end{aligned}$$

On obtient donc l'expression suivante des facteurs superflus introduits par l'algorithme euclidien :

$$\lambda_{\iota+2} = \frac{\lambda_\iota\alpha_\iota^2}{\alpha_{\iota+1}^2}$$

D'où finalement :

$$\lambda_{\iota+2} = \frac{\alpha_\iota^2\alpha_{\iota-2}^2\alpha_{\iota-4}^2\cdots}{\alpha_{\iota+1}^2\alpha_{\iota-1}^2\alpha_{\iota-3}^2\cdots}$$

Sylvester mentionne, au § 10, l'application possible de son raisonnement sur les facteurs superflus à d'autres situations que l'algorithme euclidien. Nous allons à présent l'utiliser afin d'étudier l'algorithme d'élimination successive des plus hautes puissances de l'inconnue. Nous nous éloignons à cet effet des notations de Sylvester. A nouveau, les raisonnements ci-dessous ne sont valables qu'à un facteur numérique près. Soient f_0, ϕ_0 deux polynômes initiaux de même degré n en x. Pour tout polynôme f, on notera $\mathrm{lc}f$ (resp. $\mathrm{sc}f$) le coefficient dominant de f (resp. son avant-dernier coefficient). L'élimination des plus hautes puissances consiste à construire une suite de polynômes f_1, ϕ_1, ..., f_i, ϕ_i,..., où f_i et ϕ_i sont de degré $(n-i)$, en éliminant à chaque étape la plus haute puissance de l'inconnue dans les deux derniers polynômes obtenus, par combinaison linéaire :

$$\begin{cases} f_{i+1} = f_i - \frac{\mathrm{lc}f_i}{\mathrm{lc}\phi_i}\phi_i\\ \phi_{i+1} = \phi_i - x\frac{\mathrm{lc}\phi_i}{\mathrm{lc}f_{i+1}}f_{i+1}\end{cases}$$

Cet algorithme n'introduit pas de facteur superflu au numérateur (on l'a déjà mentionné à propos de Newton, *cf. supra* section 3.5). Pour le montrer, écrivons :

$$\begin{cases} f_i = l_iF_i\\ \phi_i = \lambda_i\Phi_i\end{cases}$$

où l_i, λ_i désignent les facteurs superflus. La première étape de l'algorithme conduit à :

$$\begin{aligned} F_1 &= \mathrm{lc}\phi_0.f_0 - \mathrm{lc}f_0.\phi_0\\ \Phi_1 &= \phi_0\mathrm{lc}F_1 - x\mathrm{lc}\phi_0.F_1\end{aligned}$$

Pour tout i, il existe des expressions de F_i et Φ_i en termes de déterminants (sans détailler ce point, remarquons que F_i n'est autre que le résidu simplifié R_i de la section précédente). Il en résulte que leurs coefficients dominants $\mathrm{lc}F_i$, $\mathrm{lc}\Phi_i$ sont tous premiers entre eux, et que F_i (resp. Φ_i) est de degré i (resp. i) en les coefficients de f_0, et de degré i (resp. $i+1$) en ceux de ϕ_0. Commençons par établir des formules de récurrence entre les F_i, Φ_i :

Lemme 2 *Pour tout $i \geq 0$, on a :*

$$\begin{aligned} F_{i+2} &= \tfrac{lc\Phi_{i+1}.F_{i+1} - lcF_{i+1}.\Phi_{i+1}}{lc\Phi_i}\\ \Phi_{i+2} &= \tfrac{\Phi_{i+1}lcF_{i+2} - xlc\Phi_{i+1}.F_{i+2}}{lcF_{i+1}}\end{aligned}$$

Démonstration. Définissons F_i et Φ_i par les formules de l'énoncé du lemme. Il est facile de voir que ces fonctions vérifient alors les propriétés d'homogénéité et de degré que l'on vient de mentionner. Il nous reste à démontrer que ce sont bien des fonctions entières en les coefficients de f_0, ϕ_0. Supposons le résultat acquis jusqu'au rang $(i+1)$. Notons de plus :

$$\begin{array}{l} F'_{i+1} = \mathrm{lc}\Phi_i.F_i - \mathrm{lc}F_i.\Phi_i \quad (= F_{i+1} lc\Phi_{i-1}) \\ \Phi'_{i+1} = \Phi_i(\mathrm{lc}\Phi_i.\mathrm{sc}F_i - \mathrm{lc}F_i.\mathrm{sc}\Phi_i) - x\mathrm{lc}\Phi_i.(\mathrm{lc}\Phi_i.F_i - \mathrm{lc}F_i.\Phi_i) \quad (= \Phi_{i+1}\mathrm{lc}F_i.\mathrm{lc}\Phi_{i-1}) \end{array}$$

On peut démontrer, par un calcul explicite (voir annexe C), que $\mathrm{lc}\Phi_i$ divise bien l'expression :

$$\mathrm{lc}\Phi'_{i+1}.F'_{i+1} - \mathrm{lc}F'_{i+1}.\Phi'_{i+1}$$

De plus $\mathrm{lc}\Phi_i$ est premier avec $\mathrm{lc}F_i$ et avec $\mathrm{lc}\Phi_{i-1}$. Or

$$F_{i+2} = \frac{\mathrm{lc}\Phi'_{i+1}.F'_{i+1} - \mathrm{lc}F'_{i+1}.\Phi'_{i+1}}{\mathrm{lc}F_i.(\mathrm{lc}\Phi_{i-1})^2.\mathrm{lc}\Phi_i}$$

F_{i+2} est donc bien une fonction entière. De même, on peut démontrer, par un calcul explicite, que $\mathrm{lc}F_{i+1}.\mathrm{lc}\Phi_i$ divise bien l'expression :

$$\Phi'_{i+1}(\mathrm{lc}\Phi'_{i+1}.\mathrm{sc}F'_{i+1} - \mathrm{lc}F'_{i+1}.\mathrm{sc}\Phi'_{i+1}) - x\mathrm{lc}\Phi'_{i+1}.(\mathrm{lc}\Phi'_{i+1}.F'_{i+1} - \mathrm{lc}F'_{i+1}.\Phi'_{i+1})$$

De plus $\mathrm{lc}F_{i+1}.\mathrm{lc}\Phi_i$ est premier avec $\mathrm{lc}F_i$ et $\mathrm{lc}\Phi_{i-1}$. Or

$$\Phi_{i+2} = \frac{\Phi'_{i+1}(\mathrm{lc}\Phi'_{i+1}.\mathrm{sc}F'_{i+1} - \mathrm{lc}F'_{i+1}.\mathrm{sc}\Phi'_{i+1}) - x\mathrm{lc}\Phi'_{i+1}.(\mathrm{lc}\Phi'_{i+1}.F'_{i+1} - \mathrm{lc}F'_{i+1}.\Phi'_{i+1})}{(\mathrm{lc}F_i)^2.(\mathrm{lc}\Phi_{i-1})^3\mathrm{lc}F_{i+1}\mathrm{lc}\Phi_i}$$

Et Φ_{i+2} est donc bien une fonction entière en les coefficients de f_0, ϕ_0. *q. e. d.*

Cherchons à présent une formule de récurrence entre les facteurs superflus :

$$\left\{ \begin{array}{l} f_{i+1} = (\mathrm{lc}\Phi_i.F_i - \mathrm{lc}F_i.\Phi_i)\frac{l_i}{\mathrm{lc}\Phi_i} = F_{i+1}.\frac{l_i\mathrm{lc}\Phi_{i-1}}{\mathrm{lc}\Phi_i} \\ \phi_{i+1} = (\Phi_i\mathrm{lc}F_{i+1} - x\mathrm{lc}\Phi_i.F_{i+1})\frac{\lambda_i}{\mathrm{lc}F_{i+1}} = \Phi_{i+1}.\frac{\lambda_i\mathrm{lc}F_i}{\mathrm{lc}F_{i+1}} \end{array} \right.$$

D'où :

$$\left\{ \begin{array}{l} l_{i+1} = \frac{l_i\mathrm{lc}\Phi_{i-1}}{\mathrm{lc}\Phi_i} = \frac{\mathrm{lc}\Phi_{i-1}.\mathrm{lc}\Phi_{i-2}\ldots}{\mathrm{lc}\Phi_i.\mathrm{lc}\Phi_{i-1}\ldots} = \frac{1}{\mathrm{lc}\Phi_i} \\ \lambda_{i+1} = \frac{\lambda_i\mathrm{lc}F_i}{\mathrm{lc}F_{i+1}} = \frac{\mathrm{lc}F_i.\mathrm{lc}F_{i-1}\ldots}{\mathrm{lc}F_{i+1}.\mathrm{lc}F_i\ldots} = \frac{1}{\mathrm{lc}F_{i+1}} \end{array} \right.$$

En particulier, on voit que l'algorithme d'élimination des plus hautes puissances de l'inconnue *n*'introduit *pas* de facteur superflu au numérateur !

12.3 Le « cercle d'idées Sturmiennes »

Comme l'une des principales applications qu'envisage Sylvester dans son mémoire est le théorème de Sturm sur le nombre de racines réelles d'une équation, il est amené, comme l'avait été Sturm, à modifier l'algorithme euclidien en changeant le signe des restes successifs (nous n'en avons pas tenu compte dans les paragraphes précédents, bien que Sylvester l'explique dès le commencement). La $(\iota+2)$-ème division conduira donc au résultat suivant :

$$\lambda_{\iota+2}R_{\iota+2} + \lambda_\iota R_\iota - (px+q)\lambda_{\iota+1}R_{\iota+1} = 0$$

Les $\lambda_\iota R_\iota$ ainsi obtenus sont appelés par Sylvester « résidus complets ». Ces résidus complets vérifient une propriété importante, qui en font une suite de fonctions utilisable dans la théorie de Sturm : l'annulation du $(\iota+1)$-ème implique que les ι-ème et $(\iota+2)$-ème sont de signes contraires. Sylvester montre [18] qu'il en est de même des « résidus simplifiés » (bien sûr on a vu ci-dessus que les facteurs superflus sont des carrés, et sont donc essentiellement positifs, mais nos raisonnements n'étaient

18. *Cf.* [306], § 11 et 12.

justifiés qu'à un facteur numérique près, et donc au signe près). Pour ce faire, il démontre que leurs coefficients dominants vérifient la propriété de Sturm. En effet ces coefficients dominants sont les mineurs principaux de la matrice de Bézout (puisque les résidus simplifiés s'obtiennent en échelonnant la matrice de Bézout) et la propriété de Sturm est conséquence d'une formule générale sur les déterminants. Soit A une matrice $m \times m$, M le $(m-2)$-ème mineur principal, et $M_{i,j}$ le mineur obtenu en rayant la i-ème ligne et la j-ème colonne. Alors

$$\det M . \det A = \det M_{m+1,m+1} . \det M_{m,m} - \det M_{m+1,m} . \det M_{m,m+1}$$

Si A est une matrice symétrique (c'est le cas de la matrice de Bézout), on a $M_{m+1,m} = M_{m,m+1}$, d'où, lorsque $\det M_{m+1,m+1} = 0$:

$$\det M . \det A = -(\det M_{m+1,m})^2$$

Ainsi, lorsque l'un des mineurs principaux d'une matrice symétrique à coefficients réels s'annule, les mineurs principaux précédent et suivant sont de signes contraires.

Sylvester est donc parvenu à réduire tout l'algorithme euclidien à une méthode matricielle, fait important pour la théorie de l'élimination et pour la théorie de Sturm. L'identification de la matrice de Bézout comme matrice d'une forme quadratique permet d'ailleurs de traduire la théorie de Sturm dans le langage de la théorie des formes quadratiques. Le nombre de continuations de signes dans la suite des coefficients dominants des résidus simplifiés (en comptant f' lui-même comme premier résidu simplifié), qui mesure le nombre de racines réelles de f, *moins un*, est en effet égal au nombre de termes positifs dans la suite des quotients des mineurs successifs de la matrice de Bézout [19]. Or il est possible de diagonaliser la forme quadratique, au moyen de transformations linéaires à coefficients réels, de manière à rendre les termes sur la diagonale respectivement égaux aux quotients des mineurs successifs [20]. Sylvester montre justement à cette occasion le théorème que le nombre de termes positifs situés sur la diagonale est un invariant de la forme quadratique (c'est-à-dire qu'il ne dépend pas des transformations linéaires à coefficients réels choisies pour diagonaliser) [21]. Il nomme ce nombre « inertie » de la forme quadratique. Finalement, le nombre de racines réelles de f est donc égal à l'inertie du bézoutiant, plus un. Sylvester avait démontré que le polynôme caractéristique d'une matrice symétrique a toutes ses racines réelles, et que la matrice d'une forme quadratique est donc diagonalisable dans une base orthogonale [22]. Dans une telle base, les éléments de la diagonale (valeurs propres) sont les racines du polynôme caractéristique. Il peut donc finalement énoncer le théorème suivant :

> Etant donné une équation algébrique de degré quelconque, on peut former une équation de degré un de moins, dont toutes les racines sont réelles, et dont le nombre de racines positives est un de moins que le nombre total de racines réelles de l'équation donnée. [23]

Sylvester utilise, dans l'introduction de son mémoire, deux termes qui décrivent bien deux versants jusqu'alors inconciliables de la théorie de l'élimination :

> *Endoscopique, Exoscopique :* Lorsque, dans une investigation, les coefficients des fonctions concernées sont vus comme des monades entières indécomposables, la méthode est dite exoscopique. Elle est dite endoscopique lorsque les coefficients sont traités par rapport à leur constitution interne, en tant qu'ils sont composés de racines ou d'autres éléments. [24]

Ailleurs, il ajoute :

> Ces mots admettent une application étendue et importante en analyse. Ainsi les méthodes de résolution des équations (...) fournies par Tschirnhausen et Jerrard (...) sont essentiellement exoscopiques ; d'un autre côté, les méthodes de Lagrange et d'Abel pour résoudre les mêmes problèmes sont endoscopiques. [25]

19. *Cf.* [306], § 50, 56, 58.
20. *Cf.* [306], § 45.
21. *Cf.* [306], § 44.
22. *Cf.* [302].
23. *Cf.* [306], § 59.
24. *cf.* [306], p. 582.
25. *cf.* [306], p. 431.

Sylvester ne semble pas vouloir restreindre le sens de ces mots à la théorie de l'élimination, mais nous pouvons y voir deux catégories utiles dans notre analyse historique, décrivant bien les deux « voies » que nous avons distinguées dans notre chapitre précédent. Ainsi Hudde, Newton, Leibniz, Rolle et Bézout avaient utilisé des méthodes d'élimination « exoscopiques » conduisant à l'expression de l'équation finale en fonction des coefficients des équations initiales, alors que celle développée par Euler (méthode n°2 d'Euler), Cramer, Lagrange puis Poisson était « endoscopique », conduisant au moins dans un premier temps à une expression en fonction des racines des équations initiales qui pouvait ensuite être exprimée en fonction des coefficients au moyen des relations entre coefficients et racines. Dans son mémoire de 1853, Sylvester prétend « mettre en relation les deux manières de penser provenant de ces différents points de vue » [26]. Et en effet, dans la deuxième section, il étudie à nouveau l'algorithme euclidien et il cherche cette fois des formules des restes et quotients successifs en fonction des racines des deux équations proposées. Notant h_i $(1 \leq i \leq m)$ les racines de la première équation et η_i $(1 \leq i \leq n)$ celles de la seconde, il montre [27] que l'expression suivante est égale au « résidu simplifié » de degré i :

$$(-1)^{i\frac{i-1}{2}} . \sum \left[\prod_{j=1}^{m} (x - h_{q_{i+j}}) \frac{\prod_{1\leq j\leq i,\ 1\leq k\leq n} (h_{q_j} - \eta_k)}{\prod_{1\leq j\leq i,\ i+1\leq k\leq m} (h_{q_j} - h_{q_k})} \right]$$

où la sommation s'étend à toutes les permutations $(q_1, q_2, ..., q_m)$ de m éléments. Il trouve de même des expressions pour les Q_ι et P_ι dans la relation (1). Posons $\nu = \iota - 1$ et $\mu = m - n + \iota - 1$, alors :

$$P_\iota = (-1)^{n-\iota} \sum \left[\prod_{j=1}^{\mu} (x - h_{q_j}) \frac{\prod_{1\leq j\leq \mu,\ 1\leq k\leq n} (h_{q_j} - \eta_k)}{\prod_{1\leq j\leq \mu,\ \mu+1\leq k\leq m} (h_{q_j} - h_{q_k})} \right]$$

et

$$Q_\iota = (-1)^{m\iota+m-1} \sum \left[\prod_{j=1}^{\nu} (x - \eta_{\xi_j}) \frac{\prod_{1\leq j\leq m,\ 1\leq k\leq \nu} (h_j - \eta_{\xi_k})}{\prod_{1\leq j\leq \nu,\ \nu+1\leq k\leq n} (\eta_{\xi_j} - \eta_{\xi_k})} \right]$$

où la sommation s'étend à toutes les permutations $(\xi_1, \xi_2, ..., \xi_n)$ de n éléments [28]. Dans le cas du théorème de Sturm où l'algorithme euclidien est appliqué au couple (f, f'), si l'on pose

$$\zeta(h_{q_1}, h_{q_2}, ..., h_{q_i}) = \prod_{1\leq j<k\leq i} (h_{q_j} - h_{q_k})^2,$$

où les h_i sont les racines de f, on obtient la formule suivante [29] pour le résidu simplifié de degré i :

$$\sum \left[\zeta(h_{q_1}, h_{q_2}, ..., h_{q_i}) \prod_{j=i+1}^{m} (x - h_{q_j}) \right]$$

Dans ce cas, Sylvester trouve aussi des formules pour les Q_ι et P_ι, mais la formule pour Q_ι est nettement plus compliquée [30]. Il est remarquable que Sylvester ait travaillé sur les deux fronts (endoscopique et exoscopique) dès le début de ses recherches. Ces formules endoscopiques remontent à l'article [297] et la notation des produits ζ à l'article [299], parus en 1839-1840. Dans ces travaux, le rôle principal et paradigmatique est joué par les produits des différences d'un certain ensemble de racines, c'est-à-dire des déterminants de Vandermonde : c'est en effet par le biais du déterminant de Vandermonde que Sylvester semble aborder la théorie des déterminants (non encore pleinement constituée), en définissant ses « produits ζ-iques ». Le produit ζ-ique des différences consiste à

26. *Cf.* [306], p. 431, « the two trains of thought arising out of these distinct views are brought into mutual relation ».
27. *Cf.* [306], § 14–28.
28. *Cf.* [306], § 29.
29. *Cf.* [306], § 35.
30. *Cf.* [306], § 36–43. Nous n'avons pas analysé le « supplément à la section III » [306] p. 495–510.

développer le produit des différences des n quantités a, b, ..., k, puis à changer tous les exposants en indices. On obtient naturellement, en passant par le déterminant de Vandermonde, un déterminant noté par Sylvester $\zeta PD(ab...k)$ que l'on écrirait aujourd'hui :

$$\begin{vmatrix} a_0 & b_0 & \dots & k_0 \\ a_1 & b_1 & \dots & k_1 \\ \vdots & \vdots & & \vdots \\ a_{n-1} & b_{n-1} & \dots & k_{n-1} \end{vmatrix}$$

Bien que Sylvester n'ait pas entièrement conservé cette notation dans le mémoire de 1853, on perçoit aisément l'importance du déterminant de Vandermonde dans ces recherches.

Sturm donne en 1842 une démonstration rigoureuse des formules de Sylvester pour les résidus simplifiés en fonction des racines [31]. En 1846, Cayley parvient à développer ces formules en fonction des sommes de puissances des racines, ce qui fournit une écriture plus synthétique [32]. Soit f de degré n, et f_m le résidu simplifié de degré $(n-m)$ quand on applique l'algorithme euclidien au couple (f, f'). Notons avec Cayley S_i la somme des i-èmes puissances des racines de f. On a alors :

$$\frac{f_m(x)}{f(x)} = \begin{vmatrix} S_0 & S_1 & \dots & S_{m-2} & \sum\limits_{\forall a | f(a)=0} \frac{1}{x-a} \\ S_1 & S_2 & \dots & S_{m-1} & \sum\limits_{\forall a | f(a)=0} \frac{a}{x-a} \\ \vdots & & & & \vdots \\ S_{m-1} & S_m & \dots & S_{2m-3} & \sum\limits_{\forall a | f(a)=0} \frac{a^{m-1}}{x-a} \end{vmatrix}$$

Cayley en déduit une formule pour $f_m(x)$. Retenons simplement la forme du coefficient dominant de $f_m(x)$:

$$\begin{vmatrix} S_0 & S_1 & \dots & S_{m-1} \\ S_1 & S_2 & \dots & S_m \\ \vdots & & & \vdots \\ S_{m-1} & S_m & \dots & S_{2m-2} \end{vmatrix}$$

Ainsi, les coefficients dominants des résidus simplifiés sont représentés de nouveau comme suite des mineurs successifs d'une matrice, la matrice symétrique $(S_{i+j})_{i,j\geq 0}$ des sommes des puissances des racines, et l'on obtient une autre démonstration du fait qu'ils vérifient la propriété de Sturm mentionnée ci-dessus [33]. A partir de 1847 et jusqu'à la publication du mémoire de Sylvester en 1853, ce « cercle d'idées sturmiennes » continue de s'agrandir à travers les travaux de Borchardt [35], Jacobi puis Hermite. Chez Jacobi, sous l'influence des travaux de Borchardt, apparaît dès 1847 (mais seulement publiée de manière posthume en 1856) une démonstration du théorème de Sturm faisant entièrement l'économie de la notion de continuité, et utilisant la loi d'inertie qu'il connaissait donc dès cette époque. Max Noether, dans sa biographie de Sylvester [243], a évoqué cette contribution et son importance dans les travaux ultérieurs d'Hermite qui à leur tour étaient connus de Sylvester. Pour mieux comprendre l'origine de ce qui, aux yeux de Sylvester, forme « l'achèvement du cercle d'idées sturmiennes » [34], il nous faut donc étudier les travaux de Jacobi.

12.4 C. G. Jacobi

Jacobi avait, en 1835, rédigé un article en latin [163] sur l'élimination d'une inconnue entre deux équations, paru en 1836 au Journal de Crelle. C'est peut-être au sujet de cet article, entre autres,

31. *Cf.* [294].
32. *Cf.* [58].
33. Cayley ne pousse ce raisonnement qu'en 1848 ?
34. *Cf.* [306], § 47.

que certains ont pu déplorer l'oubli du traité majeur de Bézout au début du XIX^{e} siècle [35]. Il est vrai que Jacobi ne mentionne pas les recherches de Bézout sur la théorie de l'élimination générale (pour un nombre d'équations supérieur à deux), mais ses propres recherches, aussi spéciales qu'elles soient (elles ne concernent que l'élimination d'une seule inconnue), ne sont pas désuètes pour autant, et elles montrent même une avance certaine par rapport aux premiers travaux de Richelot, Hesse et Sylvester sur l'élimination après 1839. Elles anticipent en fait le mémoire de Sylvester sur les syzygies. Sylvester l'avoue d'ailleurs à la fin de la première section de son mémoire :

> Depuis la rédaction de la section ci-dessus m'est apparu l'estimable mémoire de Jacobi, « De Eliminatione Variabilis è duabus Equationibus Algebraicis », *Crelle*, Vol. XVI. Ce mémoire se restreint à la considération de deux équations de même degré, et les principaux résultats dans cette section concernant le carré Bézoutique [=matrice de Bézout] et les facteurs superflus [*allotrious factors*], appliqués à ce cas, s'y trouvent. Le mode de traitement est cependant suffisamment dissemblable pour justifier que cette section soit préservée sous sa forme originale. [36]

L'on s'étonnera par contre à bon droit de ce que Jacobi n'ait pas remarqué, dans cet article, l'application de ses méthodes à la théorie de Sturm.

Dans l'étude de la matrice de Bézout, Jacobi attache une importance singulière à sa matrice adjointe, matrice qui possède des propriétés analogues à celles de la matrice $(S_{i+j})_{i,j\geq 0}$ des sommes des puissances des racines vue ci-dessus. Kronecker généralisera l'étude de telles matrices dans la théorie de Sturm et l'algorithme euclidien [37].

Soient $f(x) = \sum a_k x^k$ et $\phi(x) = \sum b_k x^k$ les deux polynômes initiaux, de degré n. Notons $A = (\alpha_{r,s})_{r,s}$ la matrice de Bézout de f et ϕ, et $m_i = \sum_j \alpha_{j,i} x^j$ pour $0 \leq i \leq (n-1)$. Jacobi remarque que :

$$m_0 + m_1 x + ... + m_{n-1} x^{n-1} = \phi . \frac{df}{dx} - f . \frac{d\phi}{dx}$$

En homogénéisant f et ϕ, on obtient ainsi une expression du déterminant fonctionnel $\frac{\partial \phi}{\partial y} . \frac{\partial f}{\partial x} - \frac{\partial f}{\partial y} . \frac{\partial \phi}{\partial x}$ (déterminant qui prendra plus tard le nom de *jacobien*) en fonction des coefficients de la matrice de Bézout.

Notons avec Jacobi $(A_{r,s})_{r,s} = \mathrm{ad}^{\mathrm{t}} A$. Il est clair que $(A_{r,s})$ est symétrique. D'autre part, les formules de Cramer (appliquées au système d'équations dont la matrice est la matrice de Bézout) montrent que, si x est une racine commune de f et ϕ, on a pour tout r :

$$1 : x : x^2 : ... : x^{n-1} = A_{0,r} : A_{1,r} : A_{2,r} : ... : A_{n-1,r}$$

On en déduit des relations entre les $A_{r,s}$: pour tout r, s, r', s' vérifiant $r + s = r' + s'$, on a en particulier $A_{r,s} = A_{r',s'}$, toujours sous l'hypothèse qu'il existe une racine commune, c'est-à-dire que le résultant est nul. Pour montrer que ces relations ont même lieu en général, Jacobi démontre en fait que le résultant $L = \det A$ est un générateur de l'idéal

$$\left\{ P \mid \exists k, P x^k \in (f, \phi) \right\}$$

dans l'anneau des coefficients (à la suite de Hurwitz, nous dirions aujourd'hui que le résultant engendre l'idéal des *Trägheitsformen*). Il suffit alors de remarquer que L ne divise pas $(A_{r,s} - A_{r',s'})$ qui est donc identiquement nul. La matrice adjointe de la matrice de Bézout est donc de la forme $(A_{r+s})_{r,s}$.

35. *Cf. supra* n. 9 p. 142 et n. 52 p. 159.
36. *Cf.* [306], § 13.
37. *Cf.* [180] où Kronecker écrit l'histoire de ces matrices « persymétriques » et annonce des résultats qu'il publiera dans [182].

Les éléments de la matrice adjointe vérifient deux systèmes d'équations linéaires récurrentes :

$$\begin{cases} 0 = a_0A_0 + a_1A_1 + ... + a_nA_n \\ 0 = a_0A_1 + a_1A_2 + ... + a_nA_{n+1} \\ \vdots \\ 0 = a_0A_{n-2} + a_1A_{n-1} + ... + a_nA_{2n-2} \end{cases}$$
$$\begin{cases} 0 = b_0A_0 + b_1A_1 + ... + b_nA_n \\ 0 = b_0A_1 + b_1A_2 + ... + b_nA_{n+1} \\ \vdots \\ 0 = b_0A_{n-2} + b_1A_{n-1} + ... + b_nA_{2n-2} \end{cases}$$

Notons que ces équations sont analogues aux « formules de Newton » vérifiées par les éléments de la matrice des sommes de puissances des racines $0 = \sum a_i S_{i,j}$.

En remarquant que

$$Lx^r = \sum_{j=0}^{n-1} A_{j+r} m_j$$

où $0 \leq r \leq (n-1)$, Jacobi donne une solution explicite de l'équation suivante en fonction des A_j et des coefficients de f et ϕ (mais pas *la* solution générale) [38] :

$$Mf + N\phi = LQ$$

où Q est un polynôme quelconque donné de degré $(2n-1)$. Pour $0 \leq r \leq (n-1)$, posons :

$$\begin{cases} -M_r = \sum_{j=0}^{n-1} A_{r+j}[b_n x^{n-j-1} + b_{n-1}x^{n-j-2} + ... + b_{j+1}] \\ N_r = \sum_{j=0}^{n-1} A_{r+j}[a_n x^{n-j-1} + a_{n-1}x^{n-j-2} + ... + a_{j+1}] \\ M_{2n-1-r} = \sum_{j=0}^{n-1} A_{2n-2-r-j}[b_0 + b_1 x + ... + b_{n-j-1}x^{n-j-1}] \\ -N_{2n-1-r} = \sum_{j=0}^{n-1} A_{2n-2-r-j}[a_0 + a_1 x + ... + a_{n-j-1}x^{n-j-1}] \end{cases}$$

On a alors pour $0 \leq r \leq (2n-1)$:

$$M_r f + N_r \phi = L.x^r.$$

Jacobi montre comment calculer de proche en proche les m_r, et donc la matrice de Bézout, en fonction des déterminants de la forme [39]

$$(a_r b_s) = \begin{vmatrix} a_r & a_s \\ b_r & b_s \end{vmatrix}$$

Chaque élément $\alpha_{r,s}$ de la matrice de Bézout est une somme de tels déterminants, et on a la relation :

$$\alpha_{r-1,s} - \alpha_{r,s-1} = (a_r b_s)$$

En 1853, Sylvester donne une description plus imagée de la matrice de Bézout, en fait équivalente à cette dernière relation :

> Ce carré [=la matrice de Bézout] peut être conçu comme une sorte de pyramide étagée, formée par la superposition successive de couches carrées, lequelles possèdent non seulement une symétrie par rapport à une diagonale (ce qui est le propre d'une table de multiplication), mais aussi une symétrie plus élevée (comme il en existe dans une table d'addition), tous les termes dans n'importe quelle ligne de termes parallèle à la diagonale transverse à l'axe de symétrie étant égaux [40].

38. *Cf.* [163], § 7 à 9.
39. *Cf.* [163] § 11.
40. *Cf.* [306] § 7.

Sylvester introduit le terme « persymétrique » (*persymmetric*) pour désigner cette « symétrie plus élevée » observée dans chacune des couches. C'est justement le cas de l'adjointe de la matrice de Bézout que nous avons observé ci-dessus : $A_{r,s} = A_{r',s'}$ si $(r+s) = (r'+s')$.

La structure de la matrice de Bézout est donc compliquée, mais celle de son adjointe l'est beaucoup moins. C'est là l'idée motrice que poursuit Jacobi. A la recherche des relations existant entre les $\alpha_{r,s}$, il part des relations élémentaires entre les mineurs de l'adjointe $(A_{r+s})_{r,s}$ et les transforme en relations entre les $\alpha_{r,s}$ au moyen de la dualité entre mineurs d'une matrice et mineurs de l'adjointe. Enfin, il montre que les A_{r+s} sont algébriquement indépendants [41]. Autrement dit, étant donné une matrice persymétrique quelconque d'ordre n, Jacobi détermine des polynômes f et ϕ tels que cette matrice soit l'adjointe de la matrice de Bézout de f et ϕ.

A la fin de son mémoire, Jacobi note u_r ce que Sylvester appellera « résidu simplifié » de degré $(n-1-r)$. Il s'interroge alors sur la relation de divisibilité (de la forme $c_r u_{r-2} - v_r u_{r-1} = u_r$ où u_{r-1} est le diviseur et v_r le quotient) existant entre trois résidus simplifiés successifs, sans toutefois faire l'étude complète des facteurs superflus introduits par l'algorithme euclidien. Les u_r sont des polynômes en x dont les coefficients sont des mineurs de la matrice de Bézout. C'est encore au moyen de la dualité entre mineurs de la matrice de Bézout et mineurs de l'adjointe qu'il parvient à établir la relation de divisibilité suivante :

$$\left|(\alpha_{r,s})_{n-r\leq r,s\leq n-1}\right|^2 u_r + \left|(\alpha_{r,s})_{n-r-1\leq r,s\leq n-1}\right|^2 u_{r-2}$$

est divisible par u_{r-1}.

On a retrouvé la propriété de Sturm : si u_{r-1} s'annule, alors u_r et u_{r-2} sont de signes contraires. Pourtant Jacobi n'en dit rien dans ce mémoire, ni dans ses travaux ultérieurs jusqu'à 1847 ; Max Noether s'étonne à ce sujet de ce que « le travail de Jacobi de 1835 *De eliminatione* était perdu pour la théorie des restes de Sturm, non seulement par rapport à Sylvester qui le découvrit pendant l'impression de son mémoire de 1853, mais aussi par rapport à Jacobi lui-même ». En 1847, Jacobi aborde la question du théorème de Sturm par un autre biais. Mais c'est peut-être par une réflexion critique sur les méthodes déjà existantes qu'il fut amené à chercher une version purement algébrique du théorème. C'est en tout cas ce que laissent croire les paroles suivantes rapportées par Borchardt en 1856. Borchardt avait rédigé un article [35] dans lequel il utilisait le théorème de Sturm pour démontrer que les racines du polynôme caractéristique d'une matrice symétrique réelle sont toutes réelles ; avant publication de l'article, Borchardt informe Jacobi sur les résultats atteints, et celui-ci lui répond au sujet du théorème de Sturm :

> Je ne sais pas comment vous avez démontré votre proposition ; peut-être au moyen d'un mémoire de Sturm, où il présente sa proposition par des formules combinatoires. J'ai cherché une démonstration simple, qui ne suppose pas de proposition sur les équations, ni sur les déterminants. [42]

En tout cas, malgré l'apparente continuité entre les travaux de Jacobi, Sylvester et Sturm des années 1835–1842, et les réflexions de Jacobi en 1847, le problème n'est plus, pour Jacobi, de *préciser* l'énoncé du théorème de Sturm en donnant des formules explicites pour les « résidus » successifs ou bien un critère d'autre nature (le nombre de carrés négatifs d'une certaine forme quadratique). Il s'agit à présent de donner une démonstration nouvelle (justement Borchardt, dans son article de 1847, ne donne pas de démonstration du théorème de Sturm).

Borchardt rapporte également la démonstration que Jacobi lui a communiquée oralement à la suite de cet échange (« peu après », *bald darauf*, nous dit-il). Il n'est plus question de la matrice de Bézout. Jacobi prend comme point de départ la forme quadratique associée à la matrice $(S_{i,j})$:

$$S = \sum_{0\leq i,k\leq n-1} S_{i+k} x_i x_k$$

Cette forme quadratique peut se transformer de deux manières différentes en une somme de carrés. La première méthode consiste à appliquer des transformations linéaires à coefficients réels pour

41. *Cf.* [163] § 13, 14.
42. *Cf.* [36], p. 281.

obtenir une somme de carrés dont les coefficients dépendent des mineurs successifs de la matrice $(S_{i,j})$ (à peu près comme le fait Sylvester en 1853, § 45) de sorte que le nombre de changements de signes dans la suite des mineurs successifs est égal au nombre de carrés ayant leurs coefficients négatifs. La seconde méthode part du fait que :

$$S = \sum_{i=1}^{n} (x_0 + \alpha_i x_1 + ... + \alpha_i^{n-1} x_{n-1})^2$$

où les α_i sont les racines de la proposée. Cette égalité repose évidemment sur la formule matricielle

$$V\ {}^{\mathrm{t}}V = (S_{i,j})$$

où V est le déterminant de Vandermonde

$$V = (\alpha_k^i)_{\substack{0 \leq i \leq n-1, \\ 1 \leq k \leq n}}$$

Dans cette représentation de S, les paires de termes correspondant à des racines imaginaires conjuguées donnent exactement un carré à coefficient positif et un carré à coefficient négatif, à cause de l'égalité

$$(P + iQ)^2 + (P - iQ)^2 = 2P^2 - 2Q^2$$

Par conséquent [43]

> S est une somme de carrés réels, parmi lesquels autant ont le signe négatif qu'il y a de paires de racines imaginaires parmi $\alpha_1, \alpha_2, ..., \alpha_n$.

Finalement, Jacobi applique la loi d'inertie, selon laquelle le nombre de carrés négatifs est invariant par transformations linéaires à coefficients réels. Il en conclut que le nombre de changements de signe dans la suite des mineurs successifs est égal au nombre de paires de racines imaginaires.

En 1853, Hermite, qui semble avoir eu vent des réflexions de Jacobi [44] sans toutefois en connaître le détail [45], parvient à les généraliser en représentant les résidus successifs eux-mêmes (et non plus seulement leurs coefficients dominants avec Jacobi) comme mineurs successifs de la matrice

$$\left(\sum_{i=1}^{m} \frac{h_i^{r+s}}{x - h_i} \right)_{0 \leq r,s \leq m-1}$$

Sylvester poursuit les recherches d'Hermite, comme l'avait fait Jacobi pour la matrice $(S_{i,j})$, en concevant une démonstration purement algébrique du théorème de Sturm pour la détermination du nombre de racines dans un intervalle quelconque de la droite réelle. Il prétend ainsi « achever le cercle d'idées sturmiennes » [46].

L'« algébrisation » du théorème de Sturm a été analysée par H. Sinaceur [291]. Elle y voit d'abord une conséquence du « style abstrait » de Sturm, qui appelle sur ce théorème le règne de l'algèbre, même si par lui-même « Sturm n'a certes pas fait grand chose pour annoncer ou hâter ce règne » [47], et qui confère à Sturm une place singulière parmi les auteurs ayant étudié la question de la « séparation des racines réelles » (Lagrange, Cauchy, Fourier, Budan). Il pourrait être qualifié, selon H. Sinaceur, de « structuraliste avant la lettre », et son œuvre, comparable à celle de Galois pour la résolution algébrique des équations. Elle lit aussi dans cette algébrisation du théorème le contraste entre une école anglaise et l'école de Paris :

> A une époque où l'école de Paris était, en dépit de quelques exceptions, dominée par l'esprit analytique, Sylvester veut promouvoir l'esprit algébrique jusqu'à y soumettre l'analyse toute entière ! [48]

43. *Cf.* [36], p. 282.
44. *Cf.* [36] p. 281.
45. *Cf.* la remarque de Sylvester sur Hermite [306] § 47.
46. *Cf.* [143], et [306] § 44–47.
47. *Cf.* [291] p. 144.
48. *Cf.* [291] p. 126.

Enfin, elle éclaircit pour nous le lien entre les travaux de Sylvester et ceux de Hermite, en remarquant chez celui-ci une « rupture de style » entre son mémoire [142] de 1852 (dont Max Noether n'avait pu tenir compte dans sa biographie [243] de Sylvester) et ses « Remarques » [143] de 1853, puis la convergence dans ses travaux ultérieurs avec la conception de l'algèbre de Sylvester. H. Sinaceur retrace aussi le sort du théorème de Sturm en logique au XX^{e} siècle dans la « méthode d'élimination des quantificateurs » de Tarski. Mais par rapport à la théorie de l'élimination *telle que la conçoit Sylvester*, théorie algébrique, il ne s'agit que d'une application (comme l'indique le titre du mémoire de 1853) qui certes a donné occasion à des découvertes importantes pour l'algorithme euclidien (par exemple des formules explicites des facteurs superflus), et qui, enfin, a trouvé une certaine autonomie, grâce à la loi d'inertie des formes quadratiques. On aurait d'ailleurs déploré cette autonomie si elle s'était manifestée plus tôt, comme le confirment les paroles de Sylvester :

> Happily this cycle [=le cercle d'idées sturmiennes] was commenced from the other end, for it would have been difficult to have suspected that the root-expressions for the terms in the rhizoristic series [=suite des fonctions de Sturm] could be identified with the residues, had the former been first to be discovered, and much of the theory of algebraical common measure laid open by means of this identification would probably have remained unknown [49].

En tout cas, l'achèvement du cercle d'idées sturmiennes n'empêche pas, parallèlement, la généralisation du théorème de Sturm lui-même : par Hermite en 1852 aux systèmes de plusieurs équations à plusieurs inconnues, par Sylvester dans le mémoire de 1853, avec sa « théorie des intercalations » qui étudie la position *relative* des racines réelles de deux équations [50], et avec sa méthode d'encadrement des racines réelles d'une équation au moyen d'un système d'inéquations linéaires formées à partir des quotients successifs de l'algorithme euclidien [51]. Mais nous ne pouvons insister ici davantage sur l'histoire du théorème de Sturm.

12.5 Le bézoutiant est un combinant

Il nous reste à évoquer la section V, dernière section du grand mémoire de Sylvester. Après avoir exposé la formule du résultant attribuée à Cayley (*cf. supra* p. 185), il en déduit que le bézoutiant $B(u_1, u_2, ..., u_m)$ de deux polynômes f et ϕ homogènes de degré m en x et y est un *covariant* de f et ϕ : lorsque les variables x, y de f et ϕ subissent une transformation linéaire, les variables $u_1, u_2, ..., u_m$ de B subissent la même transformation linéaire que $x^{m-1}, x^{m-2}y, ..., y^{m-1}$. Ou encore [52], on peut dire que $B(u_1, u_2, ..., u_m)$ est un *invariant* des trois formes en x, y :

$$\begin{array}{l} f(x,y) \\ \phi(x,y) \\ \Omega(x,y) = \sum_{i=1}^{m} (-1)^{i-1} \binom{m-1}{i-1} u_i x^{i-1} y^{m-i} \end{array}$$

Plus encore, Sylvester remarque [53] que B est un *combinant* par rapport à f et ϕ, ce qui signifie que, outre son invariance, il vérifie la propriété suivante :

$$\text{bézoutiant}(kf + i\phi, k'f + i'\phi) = \begin{vmatrix} k & k' \\ i & i' \end{vmatrix}^{m} .\text{bézoutiant}(f, \phi)$$

Il remarque d'autre part [54] qu'en substituant dans B aux variables $u_1, u_2, ..., u_m$ les monômes $x^{m-1}, x^{m-2}y, ..., y^{m-1}$, on obtient le jacobien, unique covariant [55] de f et ϕ de degré $2(m-1)$

49. *Cf.* [306] § 47.
50. *Cf.* [306] § 48–51.
51. *Cf.* [306] § 52–54 et supplément à la section IV.
52. *Cf.* [306] § 61*, où une erreur figure cependant dans $\Omega(x, y)$.
53. *Cf.* [306] § 64. Jacobi avait déjà fait cette observation [163] § 12 p. 310.
54. *Cf.* [306] § 65.
55. C'est-à-dire que le jacobien est un invariant de f, ϕ et d'un vecteur (x, y). L'unicité du jacobien résulte aisément de la méthode symbolique en théorie des invariants – *cf.* [95] p. 34.

en x et y, et linéaire en les coefficients de f ainsi qu'en ceux de ϕ. Jacobi avait déjà remarqué que l'on pouvait exprimer le jacobien au moyen des coefficients de la matrice de Bézout (voir plus haut, p. 193). Plus exactement, on obtient ainsi le jacobien délivré d'un facteur entier numérique. Sylvester remarque d'ailleurs que cette manière de former le jacobien est plus simple que son expression comme déterminant de la matrice des dérivées partielles, car elle ne nécessite d'effectuer aucune soustraction. La suite de la section V fait état de recherches alors inachevées : Sylvester, à la recherche d'un procédé inverse qui permette de calculer le bézoutiant à partir du jacobien, s'est aperçu [56] que le bézoutiant n'est pas le seul invariant de f, ϕ, Ω linéaire en les coefficients de f ainsi qu'en ceux de ϕ, quadratique en $u_1, u_2, ..., u_m$, et combinant par rapport à f, ϕ. Il se propose alors [57] de former une « échelle » de tels invariants :

> Pour le moment je me contenterai de donner une méthode pour former l'échelle constitutive (sans cependant chercher la preuve que toutes les formes hors d'une telle échelle sont fonctions linéaires de celles comprises dans l'échelle), et de déterminer les relations numériques entre les formes de cette échelle et le bézoutiant pour un nombre limité de valeurs de m.

Il le fait pour $1 \leq m \leq 6$.

56. *Cf.* [306] § 66.
57. *Cf.* [306] § 67–71, et note au § 66. La citation suivante est à la p. 559.

Chapitre 13

Courbes algébriques planes et polygone de Newton

Nous avons vu Maclaurin, dès 1720 dans sa *Geometria organica*, compter les points d'intersection de deux courbes planes avec leurs multiplicités qu'il déterminait *en général* comme produit des ordres de singularité du point d'intersection sur chaque courbe (les exceptions étant les cas de tangence entre deux branches des courbes). L'absence de démonstration du théorème de Bézout le dispensait d'avoir à justifier ce nouveau concept. Il fallait d'autre part une classification des singularités des courbes algébriques planes, associant à chaque point de la courbe un certain ordre de singularité (ainsi, points doubles et points de rebroussement étaient comptés, en général, comme singularités d'ordre 2).

Nous avons rencontré pour la première fois chez Ossian Bonnet une véritable *définition* de la multiplicité d'intersection, d'apparence anodine, dans un article sur le « théorème de Bret » dont nous avons déjà parlé [1]. Sa définition découle naturellement de l'expression du résultant par la formule de Poisson. Avant de poursuivre l'histoire de ce concept, nous exposerons quelques faits élémentaires, dans le cas de deux courbes algébriques planes.

13.1 Exposé élémentaire : intersection de deux courbes algébriques planes

Supposons que l'abscisse du point d'intersection considéré est $x = x_0 < \infty$. Le résultant peut s'écrire, en développant en série les racines y de chaque équation suivant les puissances croissantes de $(x - x_0)$ (au moyen du polygone de Newton), sous la forme (formule de Poisson) :

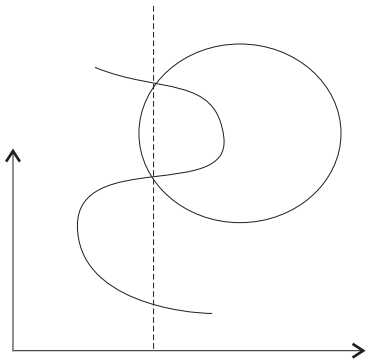
$$u_0^n v_0^m \prod_{1\leq i\leq m,\ 1\leq j\leq n} (P_i(x) - Q_j(x))$$

où m et n sont les degrés respectifs en y des équations des deux courbes, u_0 et v_0 leurs termes dominants respectifs (polynômes en x), et pour tout i, j, $y = P_i(x)$ sont les racines de la première équation et $y = Q_j(x)$ celles de la seconde. Définissons la *multiplicité d'intersection* en le point d'intersection fini (x_0, y_0) comme étant le nombre rationnel k :

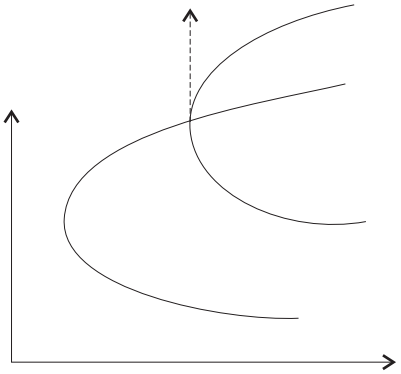
$$(1)\quad k = \operatorname{ord}_{(x-x_0)} \prod_{\substack{1\leq i\leq m,\ 1\leq j\leq n \\ P_i(x_0) = Q_j(x_0) = y_0}} (P_i(x) - Q_j(x))$$

Notre définition coïncide alors avec celle d'Ossian Bonnet.

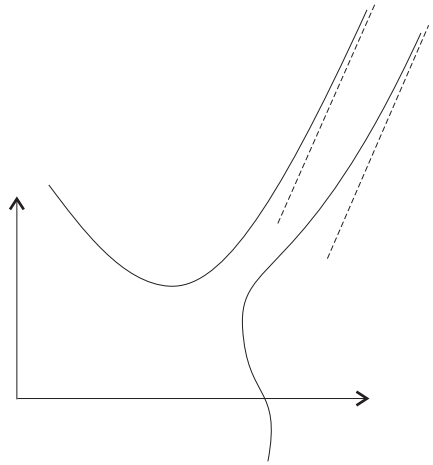
1. *Cf.* [32], « degré de multiplicité » ; *cf. supra* p. 82.



Deux points d'intersection distincts
ont même abscisse



Tangente verticale à l'une des courbes
en l'un des points d'intersection



Une direction asymptotique commune

Sous l'hypothèse qu'il n'y a aucune asymptote verticale aux courbes (c'est-à-dire que u_0 et v_0 sont indépendants de x), et qu'il n'y a qu'un unique point d'intersection d'abscisse x_0, k est égal à l'ordre du produit de tous les $(P_i(x) - Q_j(x))$, c'est-à-dire à l'ordre de multiplicité de la racine x_0 du résultant, et est donc en particulier un entier [2].

Faisons à présent l'hypothèse qu'il n'y a aucune asymptote verticale, ni de tangente verticale en aucun des points d'intersection. Alors les branches des deux courbes en $x = x_0$ sont de la forme :

$$P_i(x) = a_i + b_i(x - x_0) + ...$$
$$Q_j(x) = \alpha_j + \beta_j(x - x_0) + ...$$

(le premier terme est d'exposant ≥ 0 à cause de l'absence d'asymptote verticale, et il n'y a pas de terme d'exposant ≤ 1 fractionnaire non entier à cause de l'absence de tangente verticale [3]). Alors

$$k = \mathrm{ord}_{(x-x_0)} \prod_{\substack{1 \leq i \leq m,\ 1 \leq j \leq n \\ a_i = \alpha_j = y_0}} [(b_i - \beta_j)(x - x_0) + ...]$$

S'il n'y a pas de *contact* entre les deux courbes au point (x_0, y_0), chaque terme du produit est donc d'ordre 1 en $(x - x_0)$, et l'on a

$$k = \mathrm{Card}\{i \mid a_i = y_0\}.\mathrm{Card}\{j \mid \alpha_j = y_0\}$$

2. Ces hypothèses ne sont pas essentielles : on s'y ramène par un changement de coordonnées affine.

3. Bien que la série ne soit pas nécessairement entière (par exemple, en un point de rebroussement à tangente non verticale).

Plus généralement même s'il y a contact, on aura toujours (et c'est ainsi que Maclaurin concevait la multiplicité d'intersection) :

$$k \geq \text{Card}\{i \mid a_i = y_0\}.\text{Card}\{j \mid \alpha_j = y_0\}$$

Supposons que les courbes n'ont aucune asymptote verticale, ni de point de ramification à l'infini, ni de « branche parabolique » [4]. L'absence d'asymptote verticale permet d'écrire les équations des courbes sous la forme $P(x,y) = 0$, $Q(x,y) = 0$, avec :

$$\begin{cases} P(x,y) = \prod(y - P_{x_0,i}(x)) \\ Q(x,y) = \prod(y - Q_{x_0,j}(x)) \end{cases}$$

(on suppose que $u_0 = v_0 = 1$) où $P_{x_0,i}$ et $Q_{x_0,j}$ sont les développements en séries au voisinage d'une abscisse x_0 quelconque. Alors le degré total du résultant est la somme des ordres de multiplicité de ses racines, c'est-à-dire la somme des multiplicités d'intersection de tous les points des deux courbes. D'autre part, l'absence de point de ramification à l'infini implique que les développements en séries au voisinage de $x_0 = \infty$ sont des séries de Laurent (uniformisante $\frac{1}{x}$, développements suivant les puissances *entières* décroissantes de x, avec un nombre fini d'exposants positifs). On dit que les courbes n'ont pas de « branche parabolique » lorsque ces séries sont de la forme :

$$\begin{aligned} P_{\infty,i}(x) &= a_i x + b_i + \tfrac{c_i}{x} + ... \\ Q_{\infty,j}(x) &= \alpha_j x + \beta_j + \tfrac{\gamma_j}{x} + ... \end{aligned}$$

(il suffit pour cela de s'assurer que les coefficients des termes en $y^{\deg P}$ et $y^{\deg Q}$ dans $P(x,y)$ et $Q(x,y)$ sont non nuls). Alors le degré du résultant en x est :

$$\deg_x \prod_{1 \leq j \leq n} P(x, Q_{\infty,j}(x)) = \deg_x \prod_{i \leq m,\ j \leq n} [(a_i - \alpha_j)x + (b_i - \beta_j) + ...]$$

Du moment qu'il n'y a pas de direction asymptotique commune aux deux courbes (c'est-à-dire de point d'intersection à l'infini), on a toujours $(a_i - \alpha_j) \neq 0$, et chaque facteur du produit compte pour 1 dans le degré du résultant. On aura donc, dans ce cas : [5]

$$\deg \text{Res}_x(P,Q) = \deg_y P. \deg_y Q$$

On sait déjà que le degré du résultant (somme des ordres de multiplicité de toutes ses racines) est égal à la somme des multiplicités d'intersection de tous les points. En notant $I(P,Q;(x_0,y_0))$ la multiplicité d'intersection du point (x_0,y_0) sur les deux courbes, on a donc finalement le théorème de Bézout (sous la forme qu'employait Maclaurin) :

$$\deg P. \deg Q = \sum_{(x_0,y_0)} I(P,Q;(x_0,y_0))$$

Malgré le caractère intuitif élémentaire des considérations précédentes, conjuguant développements en séries, algèbre et géométrie, certes au prix de plusieurs hypothèses simplificatrices que nous

4. A nouveau, on verra que ces hypothèses ne sont pas essentielles : on s'y ramène en choisissant la droite à l'infini de sorte que les points à l'infini des deux courbes soient tous distincts.

5. On trouve dans un mémoire de Cauchy une esquisse de démonstration qui ressemble aux considérations précédentes. Soient deux équations

$$\begin{cases} x^n + Ax^{n-1} + Bx^{n-2} + ... = 0 \\ x^m + Px^{m-1} + Qx^{m-2} + ... = 0 \end{cases}$$

où $\deg_y A = \deg_y P = 1$, $\deg_y B = \deg_y Q = 2, ...$ Le résultant $U = 0$ est le produit des mn différences des racines. Cauchy fait le changement de variable $z = \frac{x}{y}$, et il remarque que les valeurs de z « resteront finies pour des valeurs infinies de y » (car il en est ainsi des coefficients des deux équations en z). Donc le résultant des équations transformées, qui est $\frac{U}{y^{mn}}$, restera fini pour des valeurs infinies de y. Donc $\deg_y U = mn$. *cf.* [53] p. 160–161.

discuterons ensuite (et qui ne sont pas essentielles puisque l'on peut s'y ramener par transformations affines ou projectives), il semble que personne n'ait été prêt à décrire la théorie de l'élimination sous cet angle jusque dans les années 1870. Alors plusieurs auteurs abordent cette question : Max Noether, Weierstrass, Halphen, et d'autres. On voit se dessiner, à la même époque, un concept plus général de multiplicité d'intersection, et des formules pour l'exprimer, aussi bien dans les cas d'intersections simples que dans les phénomènes de « contact », et bientôt valables en dimension supérieure. L'on trouve chez ces auteurs pourtant, recomposées, les traces d'idées présentes dans les travaux du passé. C'est d'abord ce passé que nous allons étudier.

13.2 Polygone de Newton et ramification

A la fin du XVII^{e} siècle, Newton avait décrit son algorithme de développement en séries des solutions d'une équation polynomiale à deux inconnues (le « polygone de Newton ») dans un ouvrage célèbre, mais il ne semble pas l'avoir utilisé pour étudier le comportement local des courbes. Bien qu'Euler s'intéressât à ces questions, eût-il été possible alors de concevoir le polynôme $P(x, y)$ comme produit de facteurs linéaires en y de la forme $\prod(y - P_i(x))$ où les $P_i(x)$ sont développables en séries suivant les puissances de x ? Bien des obstacles s'y opposaient. Il fallait d'abord concevoir les développements en séries donnés par Newton non plus seulement comme exprimant par une formule la dépendance fonctionnelle d'une racine y par rapport x, mais comme une véritable expression algébrique des racines de l'équation, c'est-à-dire entrant dans la composition de ses coefficients (polynômes en x) suivant les relations mêmes qui unissent coefficients et racines dans une équation à une seule inconnue. Passons outre. L'algorithme de Newton fait intervenir, dans le développement en série, des quantités algébriques solutions d'équations à une inconnue dont le degré peut être élevé et qui n'ont pas toujours de racines réelles. En l'absence d'une théorie géométrique achevée du nombre complexe, que dire de ces développements en séries ? Bien qu'Euler fût familier de la notion de « nombre imaginaire » et qu'il sût l'utiliser à bon escient, nous avons déjà montré qu'il ne maîtrisait pas suffisamment l'usage des développements en série pour l'étude locale d'une courbe en un point. D'Alembert, qui avait peut-être plus conscience de l'importance du polygone de Newton dans l'étude locale d'une courbe, l'utilisait pour exprimer une racine x d'une équation de la forme $xQ(x) + y = 0$, comme fonction de y, afin de démontrer le théorème fondamental de l'algèbre[6]. Mais il se restreignait alors à un développement en série suivant les puissances de y *à coefficients réels*. Eût-il été prêt à travailler avec une fonction dont le développement en série a des coefficients imaginaires ?

Les raisonnements d'Euler ou de d'Alembert, en général exacts, reposaient sur une conception souple mais étroite du « nombre imaginaire » : assez souple pour s'adapter en bien des cas à l'intuition infaillible d'un grand mathématicien, mais trop étroite pour embrasser l'expression analytique de l'ensemble des racines d'une équation en fonction d'un paramètre. L'œuvre fondatrice de Cauchy en analyse complexe, dans la première moitié du XIX^{e} siècle, va renverser cet obstacle. Mais bien que Cauchy s'intéresse à l'expression analytique de ces « fonctions irrationnelles », solutions d'une équation algébrique dont les coefficients dépendent d'un paramètre, il n'a pas recours au polygone de Newton ni aux séries à exposants fractionnaires. Brill décrit en détail ces travaux[7]. Rappelons les réflexions de Brill, sur les « lacunes dans la théorie de Cauchy » qui démarquent précisément les travaux de Puiseux de cette dernière. Brill identifie trois lacunes :
(i) Cauchy « ne traite pas les points de ramification séparément des pôles de la fonction »,
(ii) Cauchy « ne translate pas le centre du disque de convergence en un point variable du plan complexe » (autrement dit le développement en série n'est pas pour lui l'outil d'une étude locale),
(iii) Cauchy n'étudie pas les « singularités de la fonction » plus élevées que les points de ramification simples ou les pôles, alors que le polygone de Newton aurait pu servir ce but.
Quant à la lacune (i), autant dire que Cauchy ne décrit pas le phénomène de ramification en un point. S'intéresse-t-il jamais aux « singularités de la fonction », son but n'était pas d'étudier localement une telle fonction au voisinage d'une quelconque valeur du paramètre, mais au contraire de fournir

6. *Cf.* [157] p. 29–33.
7. *Cf.* [49], p. 176–197. Les citations suivantes se rapportent à la p. 197.

une expression de la fonction par une série à exposants entiers valable dans un domaine dont tout point de ramification ou pôle sont exclus et d'en étudier le rayon de convergence. Les « singularités de la fonction » ne s'insèrent dans l'étude qu'en tant qu'elles limitent le rayon de convergence de la série entière représentant la fonction, et le polygone de Newton n'a pas sa place dans une telle théorie.

Pourtant, Cauchy, de même que Poisson et O. Bonnet, n'hésite pas à écrire le résultant sous la forme $\prod(P_i(x) - Q_j(x))$, où les P_i et les Q_j, à défaut d'un développement en série, ont au moins le caractère de « fonctions irrationnelles » de x (*cf. supra* note 5). La théorie de l'élimination, en cette première moitié du XIX$^{\text{e}}$ siècle, semble se transformer au contact de l'analyse. Nous en avons un autre témoignage en la personne de Joseph Fourier. Son ouvrage posthume *L'analyse des équations déterminées* [115] édité par Navier en 1831 devait contenir sept livres, mais seule la première partie (les deux premiers livres) vit le jour. Dans sa préface, Fourier indique quel devait être le sujet du quatrième livre : la « résolution des équations littérales », par la généralisation du « polygone de Newton » à des systèmes de plusieurs équations. Pour une seule équation à deux inconnues, la méthode du polygone de Newton ramène le calcul du développement en série d'une des deux variables suivant les puissances de l'autre variable à la recherche de l'enveloppe convexe d'un ensemble de points dans le plan. Dans le cas d'un système de plusieurs équations à autant d'inconnues, la recherche de l'enveloppe convexe se traduit par la résolution d'un système d'inéquations linéaires, sujet que Fourier connaissait bien (le septième livre de *L'analyse des équations déterminées* devait être consacré à l'« analyse des inégalités linéaires »). Fourier veut substituer cette méthode de résolution des systèmes d'équations à toutes les méthodes par élimination, qui ne donnent, selon lui, que des résultats « théoriques » [8].

Minding, en 1841 [230], propose d'utiliser le même outil, le polygone de Newton, pour résoudre justement un problème théorique : la détermination du degré de l'équation finale résultant de l'élimination d'une inconnue parmi deux dans un système de deux équations (pas des équations génériques, sinon le théorème de Bézout suffit...) :

$$\begin{cases} f(x,y) = A_0 y^m + A_1 y^{m-1} + ... \\ \phi(x,y) = B_0 y^n + B_1 y^{n-1} + ... \end{cases}$$

Le résultant est (formule de Poisson) :

$$\psi(x) = B_0^m \prod_{i=1}^{n} f(x, y_i)$$

où les y_i sont les racines de $\phi(x,y) = 0$ développées en séries suivant les puissances décroissantes de x :

$$y_i = c_i x^{h_i} + ...$$

(ce développement en série est obtenu au moyen du polygone de Newton, les exposants peuvent être fractionnaires). Notons k_i l'exposant du terme dominant de $f(x, y_i)$, et $b_0 = \deg B_0$. Alors :

$$\deg \psi = mb_0 + \sum_{i=1}^{n} k_i$$

Lorsque $B_0 = 1$, dans un langage qui sera celui de Weierstrass, Minding détermine donc le degré de ψ en comptant le nombre de pôles (avec multiplicité) de f sur la courbe $\phi(x,y) = 0$ (tous ces pôles sont situés en $x = \infty$ puisque f est une fonction entière), et ce nombre – en tant que degré

8. Fourier semble s'être intéressé à l'élimination assez tôt ; il écrivait, à l'âge de 21 ans, en 1789, alors que la Révolution rendait difficile l'accès aux livres (lettre du 22 mars à Claude-Louis Bonard, citée dans *Fourier, créateur de la physique-mathématique*, J. Dhombres et J.-B. Robert, Paris, Belin, 2000, p. 690) :

> J'ai encore travaillé ces méthodes d'élimination, il n'est pas difficile de reconnaître combien celles dont on fait usage sont défectueuses, mais il l'est beaucoup de leur en substituer de meilleures. Vous voyez bien qu'il faudrait que j'eusse sous les yeux l'ouvrage de M. Bézout sur le même sujet.

du résultant – est aussi égal au nombre de points d'intersection finis (comptés avec multiplicité) des deux courbes $f(x,y)=0$ et $\phi(x,y)=0$. Si l'on échange les rôles de f et ψ en écrivant

$$\psi(x) = A_0^n \prod_{i=1}^{m} \phi(x, \eta_i),$$

on obtient une expression différente de $\deg \psi$. Dans un article ultérieur, Minding cherche donc une expression plus « symétrique » du degré de ψ, et il démontre ce théorème [9] :

> On ordonne les racines y et η des équations ci-dessus en une suite, dans l'ordre croissant de leurs degrés ; alors on obtient le degré de l'équation finale $\psi(x)=0$ en multipliant le degré de chaque racine, aussi bien de la suite des y que de la suite des η, par le nombre de racines dans l'autre suite qui la suivent dans l'ordre mentionné puis en additionnant tous ces produits, et en ajoutant $na_0 + mb_0$ à la somme ; où a_0 et b_0 sont les degrés de A_0 et B_0.

Mais c'est seulement à la suite des travaux de Puiseux en 1850 que le polygone de Newton donnera tous ses fruits. Décrivons à présent rapidement la méthode de Puiseux. Soit l'équation

$$f(u,z)=0$$

définissant u comme « fonction algébrique », donc multivaluée, de z. Puiseux étudie le comportement de cette fonction au voisinage du point $(z,u)=(a,b)$; il ne s'intéresse pas tant à un développement en série de u qu'à la manière dont les différentes valeurs de la fonction en un point a du plan complexe « se changent les unes dans les autres » [10] lorsque l'on parcourt un chemin fermé autour de a. Nous décrirons seulement ici l'algorithme qui en résulte pour le développement en série de u. Posons

$$\alpha = z-a, \quad \beta = u-b$$

On cherche le premier coefficient dans le développement en série de β en posant :

$$\beta = v(\alpha)\alpha^{\mu}$$

où $\mu = \frac{r}{s}$ est une fraction irréductible inconnue pour le moment, et $v(0)=v_0$ est différent de zéro. Alors $v(0)\alpha^{\mu}$ sera le premier terme du développement en série de β. Notons :

$$f(\beta) := f(\beta+b, \alpha+a) = \sum A_{p,q}\beta^p\alpha^q$$

L'équation initiale devient alors $f(v\alpha^{\mu})=0$, ou encore :

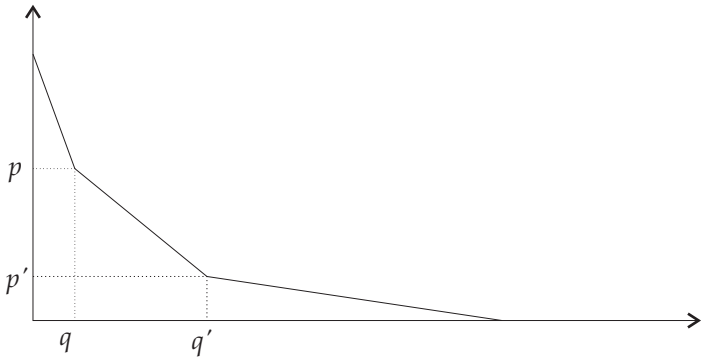
$$\sum_i \left(\alpha^i \sum_{q+\mu p=i} A_{p,q} v^p \right) = 0$$

Pour μ donné, lorsque α est un « infiniment petit », cette équation devient

$$\sum_{q+\mu p=i_0} A_{p,q} v_0^p = 0$$

où $i_0 = \min\limits_{(q,p)|A_{p,q}\neq 0} (q+\mu p)$. Pour trouver les valeurs de μ telles que cette somme ne soit pas réduite à un monôme (auquel cas $v_0=0$ contrairement à l'hypothèse), on cherche l'enveloppe convexe des points de coordonnées (q,p) dans le quadrant supérieur droit du plan (remarquons qu'il y a au moins un point sur chaque axe, quitte à diviser l'équation par un monôme) :

9. *Cf.* [232] p. 2.
10. *Cf.* [256], § 18, p. 384, et pour l'algorithme de développement en série § 21–23.



C'est ce polygone que l'on appelait traditionnellement « polygone de Newton », bien que Puiseux n'emploie pas cette expression. On obtient ω valeurs distinctes de μ, donnant chacune une équation en v_0 dont chaque racine donne à son tour le premier terme $v_0\alpha^\mu$ du développement en série d'une ou plusieurs des valeurs de la fonction β de α. Considérons l'une de ces ω équations :

$$\sum_{q+\mu p=i_0} A_{p,q} v_0^p = 0$$

Dans le cas où les racines de cette équation sont toutes distinctes, l'équation initiale en v et α, développée suivant les puissances de $(v-v_0)$ et $\alpha^{1/s}$, contient un terme linéaire, et il était bien connu qu'il existe alors un développement en série entière de $(v-v_0)$ suivant les puissances croissantes de $\alpha^{1/s}$. Dans le cas où les racines de l'équation ci-dessus ne sont pas toutes distinctes, soit v_0 une racine multiple. On pose

$$\beta = v_0\alpha^{r/s} + \beta'$$

et on réitère la méthode qui vient d'être exposée en remplaçant α, β respectivement par $\alpha^{1/s}, \beta'$ afin d'obtenir les termes successifs des développements en séries des valeurs de β commençant par $v_0\alpha^{r/s}$. *Remarque* : pour tous (p,q) et (p',q') tels que $q+p\mu = i$, $q'+p'\mu = i$, on a

$$\mu = \frac{q'-q}{p-p'} = \frac{r}{s}$$

donc s divise tous les $(p-p')$. En divisant l'équation en v_0 par sa puissance $v_0^{p'}$ d'exposant minimal, elle devient :

$$P(v_0) = 0$$

avec

$$P(X) = \sum_{q+\mu p=i_0} A_{p,q} X^{p-p'}$$

On remarque alors que le degré ϕs de cette équation, divisible par s d'après ce qui précède, est la largeur d'un des côtés du polygone de Newton. Puiseux montre que « les valeurs infiniments petites de β » données par cette équation se partagent alors « en ϕ groupes (...) et que les s fonctions qui composent un même groupe peuvent être rangées circulairement dans un ordre tel, que chacune d'elles devienne égale, après une révolution (...), à la valeur initiale de la suivante » [11].

Sans entrer davantage dans les détails de la méthode de Puiseux, indiquons seulement qu'elle rend difficile à prédire la largeur des polygones de Newton successifs dans la réitération de l'algorithme. En particulier, au cas où f est irréductible, Puiseux n'a pas démontré que l'on aboutit à des racines toutes distinctes après un nombre fini d'étapes. On le voit plus facilement en choisissant un changement de variables légèrement différent du précédent :

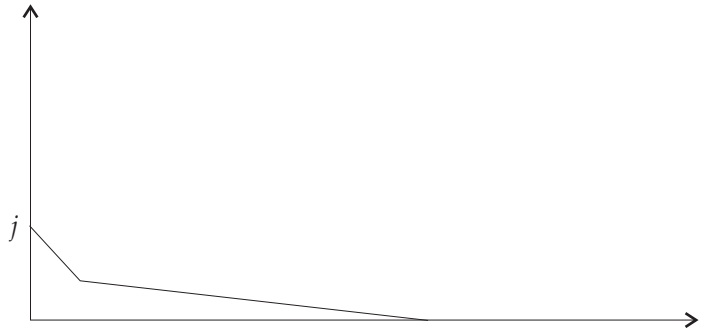
$$\beta = \alpha^\mu(v_0 + \beta')$$

11. *Cf.* [256], p. 395.

On aura :

$$f(\alpha^{\mu}(v_0+\beta')) = \alpha^{i_0}(v_0+\beta')^{p'}P(v_0+\beta') + \sum_{\substack{i>i_0 \\ q+\mu p=i}} \alpha^i A_{p,q}(v_0+\beta')^p,$$

qui est divisible par α^{i_0}. Pour connaître la hauteur du nouveau polygone de Newton, on regarde le terme en α^{i_0} de degré minimal en β' (correspondant au point le plus bas situé sur l'axe des ordonnées). Si v_0 est une racine de P de multiplicité j, alors $P(v_0+\beta')$ est divisible par $(\beta')^j$, et la hauteur du nouveau polygone est j :



En particulier cette hauteur est inférieure ou égale à ϕs, hauteur du segment du polygone précédent qui nous a permis de trouver v_0. Ainsi la hauteur du polygone décroît constamment. Il faut encore démontrer qu'elle ne peut rester constante indéfiniment. En effet, supposons au contraire qu'elle soit constante à partir du premier polygone de Newton. Alors $j = \phi s$ et ϕs est aussi la hauteur du premier polygone. On a (c est une constante) :

$$P(X) = \sum A_{p,q}X^{p-p'} = c(X-v_0)^j$$

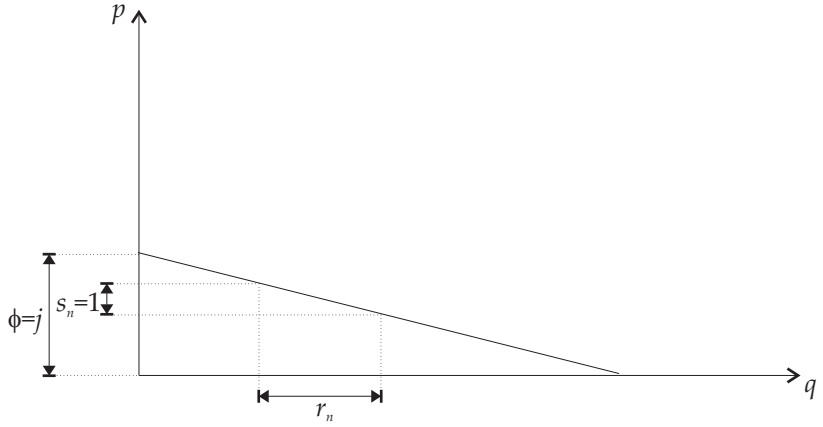
Les $(p-p')$ qui entrent dans cette expression sont donc $0, 1, ..., j$. Donc s, qui les divise tous, est égal à 1. Donc le premier polygone et tous les suivants sont de hauteur $\phi = j$. Notons

$$\nu_n = \mu + \mu_1 + \mu_2 + ... + \mu_n$$

les exposants des termes successifs du développement en série que l'on obtient au moyen de cette suite de polygones de Newton, et

$$\mu_n = \frac{r_n}{s_n}.$$

Alors le $(n+1)$-ème polygone ($n \geq 0$) sera de la forme :



Notons $i_n = q + \mu_n p$ pour (q, p) appartenant au $(n+1)$-ème polygone. On voit donc que

$$i_n = \mu_n j$$

Les $(n+1)$ premiers polygones permettent d'obtenir la somme partielle β_n suivante :

$$\beta_n = v_0\alpha^{\nu_0} + v_1\alpha^{\nu_1} + ... + v_n\alpha^{\nu_n}$$

Il est alors facile de voir que la valuation en α de $f(\beta_n)$, notée $V(f(\beta_n))$, tend vers l'infini avec n. La construction du $(n+2)$-ème polygone de Newton conduit à calculer :

$$f(\beta_n + \alpha^{\nu_n}\beta^{(n+1)}) = \alpha^{i_0+i_1+...+i_n}.(...)$$

Dérivons cette expression :

$$\frac{\partial\left(f(\beta_n + \alpha^{\nu_n}\beta^{(n+1)})\right)}{\partial\beta^{(n+1)}} = \alpha^{\nu_n} f'(\beta_n + \alpha^{\nu_n}\beta^{(n+1)}) = \alpha^{i_0+i_1+...+i_n}.(...)$$

En posant $\beta^{(n+1)} = 0$ dans cette expression et en prenant la valuation, on a alors :

$$V(f'(\beta_n)) \geq (i_0 - \mu) + (i_1 - \mu_1) + ... + (i_n - \mu_n) = (j-1)\nu_n$$

De plus $V(\mathrm{disc}(f, f')) \geq \min\{V(f(\beta_n)), V(f'(\beta_n))\}$. Mais, comme f est irréductible, son discriminant a une valuation finie. Il est donc impossible que $V(f'(\beta_n))$ tende vers l'infini avec n. Donc $j = 1$, *q. e. d.*

Hamburger [139] n'était pas entièrement satisfait de la méthode de Puiseux, et en particulier de l'usage systématique de la notion d'« ordre d'un infiniment petit ». En effet pour déterminer l'ordre μ du premier terme du développement en série de β, le raisonnement de Puiseux revenait à admettre *a priori* l'existence d'un développement en série, puis à donner les moyens d'en calculer le premier terme. Par exemple, avant de donner l'interprétation graphique du polygone de Newton, il décrivait ainsi le choix des termes de l'équation $\Lambda = f(u, z)$ formant un des côtés du polygone :

> Dans le polynôme Λ, choisir de toutes les manières possibles un groupe de deux ou de plusieurs termes tels, qu'en y regardant α comme un infiniment petit du premier ordre et β comme un infiniment petit d'un ordre convenable, ces termes soient d'un même ordre inférieur à celui de tous les autres termes du polynôme. Quand on aura trouvé tous ces groupes, on formera, en les égalant à zéro, des équations dont chacune déterminera approximativement une ou plusieurs des p valeurs infiniment petites de β. [12]

Hamburger écrit :

> Si expéditive cette méthode soit elle pour le calcul effectif des séries, et si peu puisse-t-on élever d'objection fondée contre la solidité de sa démonstration, il semble pourtant intéressant de chercher si l'on peut parvenir de manière purement analytique aux développements en séries, sans détermination de dimension (*Dimensionsbestimmungen*) préalable, détermination dont seules les séries une fois trouvées donne la certitude. [13]

La méthode de Hamburger, moins directe que celle de Puiseux, est plus simple car elle n'a recours qu'à un seul type de changement de variable, une transformation quadratique du plan $(x, y) \mapsto (x, t)$ définie par :

$$t = \frac{y - \eta}{x - \xi}$$

Soit donc l'équation $f(x, y) = 0$ de degré μ. On cherche le développement en série de $(y - \eta)$ au point (ξ, η). Supposons η racine n-uple de $f(\xi, y)$, $n \leq \mu$. On procède alors au changement de variable $(x, y) \mapsto (x, t)$ ci-dessus. Notons la nouvelle équation $g(x, t) = 0$. Dans l'équation

$$g(\xi, t) = 0$$

12. *Cf.* [256], p. 389.
13. *Cf.* [139], p. 462. Le terme *Dimensionsbestimmungen* renvoie ici à la détermination de l'« ordre d'un infiniment petit ».

Hamburger reconnaît l'équation en t de degré $\nu \leq n$ dont les racines sont les différentes valeurs de $\frac{\partial y}{\partial x}$ (attention, ν dénote ici pour nous, comme pour Hamburger, le nombre des racines de l'équation y *compris* les « racines infinies » correspondant aux racines nulles de l'équation réciproque en $\frac{\partial x}{\partial y}$). Elle a des racines infinies si et seulement si $\nu < n$. On considère chaque racine de cette équation, suivant trois cas.
Cas (i) : racine simple. Les racines simples donnent immédiatement des développements en séries de $y(x)$ par des méthodes bien connues.
Cas (ii) : cas d'une racine finie multiple τ. On réitère avec $g(x,t)$ de degré $\nu \leq \mu$, en (ξ, τ), ce qu'on a fait ci-dessus avec $f(x,y)$ de degré μ, en (ξ, η). On obtient ainsi les différents développements en séries de $t(x)$.
Cas (iii) : cas d'une racine infinie. On reprend l'équation initiale $f(x,y) = 0$ et on fait plutôt un changement de variable de même forme mais en échangeant le rôle de x et y, en posant :

$$u = \frac{x - \xi}{y - \eta}$$

Soit $g(y,u) = 0$ l'équation transformée. L'équation $g(\eta, u) = 0$ est alors l'équation réciproque de celle précédemment obtenue, elle est de même degré ν, et ces racines sont les valeurs de $\frac{\partial x}{\partial y}$, dont certaines sont nulles. On recommence alors avec $g(y,u) = 0$ de degré $\nu < n \leq \mu$, en $(\eta, 0)$, ce qu'on a fait ci-dessus pour $f(x,y) = 0$ en (ξ, η). On obtient ainsi les développements en séries de $u(y)$, qui, inversés par une méthode connue, fournissent des développements de $y(u)$.

On voit que le degré des équations en $\frac{\partial y}{\partial x}$, $\frac{\partial t}{\partial x}$, etc. décroît strictement à chaque étape, à condition que l'équation considérée possède une racine infinie. Pour montrer que l'on doit nécessairement rencontrer une équation possédant une racine infinie, ou bien n'ayant que des racines distinctes, Hamburger remarque que la composée de k transformations dans le cas (ii) est une transformation $(x,y) \mapsto (x, t_k)$ de la forme :

$$y - \eta = (x - \xi)\frac{\partial y}{\partial x}(\xi, \eta) + ... + \frac{(x-\xi)^{k-1}}{(k-1)!}\frac{\partial^{k-1} y}{\partial x^{k-1}}(\xi, \eta) + (x - \xi)^k t_k$$

En $x = \xi$, les racines de l'équation transformée ne sont autres que les différentes valeurs de

$$t_k = \frac{\frac{\partial^k y}{\partial x^k}(\xi, \eta)}{k!}$$

Hamburger donne alors une borne sur k grâce à un résultat sur le discriminant, à peu près comme nous l'avons fait ci-dessus pour démontrer la finitude de la méthode de Newton-Puiseux.

L'usage des transformations quadratiques avait déjà été suggéré, peu auparavant, par Noether [244] comme méthode de résolution des singularités d'une courbe algébrique. En 1892, Noether parlera donc du résultat de Hamburger en ces termes :

> Le résultat principal exprimant que l'on peut transformer birationnellement (et certes par une transformation birationnelle du plan) une courbe avec des singularités quelconques en une courbe avec des singularités ordinaires, c'est-à-dire seulement avec des points multiples à tangentes distinctes, a été énoncé explicitement par Noether ; la démonstration du fait que la transformation doit s'arrêter après un nombre fini d'opérations se trouve pour la première fois publié par Hamburger. [14]

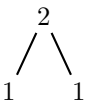
Aussi, bien que le travail de Hamburger ne soit pas entièrement original, il a le mérite d'offrir pour la première fois une démonstration complète et rigoureuse (y compris quant à la convergence des séries obtenues), et se prêtant aisément à une interprétation géométrique en théorie des courbes (alors que l'intuition sous-jacente aux travaux de Puiseux est plutôt celle de fonction sur une surface de Riemann).

Pour décrire l'effet d'une suite de transformations quadratiques sur une courbe, Noether invente en 1871 le concept de « point infiniment voisin ». Il cherche à exprimer, à la suite de Cayley, de

14. *Cf.* [49] p. 371.

combien un point singulier abaisse le genre d'une courbe, puis il généralise sa méthode au cas des surfaces dans $\mathbb{P}^3$. Il semble partir de la constatation que l'ordre d'un point sur une courbe [15] ne suffit pas à décrire le caractère singulier de ce point, même pour les singularités les plus simples (point double et point de rebroussement ordinaires sont tous deux d'ordre deux). Cette constatation avait aussi été le point de départ d'une tentative de Sylvester pour classer les singularités [16]. Il faut, pour compléter cette information, y ajouter l'information relative à tous les points correspondant à ce point sur la transformée propre de la courbe par une transformation quadratique centrée en ce point. Cette information supplémentaire s'obtient à son tour au moyen de transformations quadratiques successives. Le fait qu'une telle suite de transformations quadratiques conduise en un nombre fini d'étapes à des points tous d'ordre un implique que l'on peut représenter le point singulier par un arbre fini [17], dont voici quelques exemples :

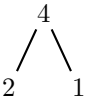
1) point double ordinaire. En $(0,0)$, l'équation $x^2 - y^2 = 0$ a un point double ordinaire, transformé en deux points simples $(0, \pm 1)$ par la transformation $t = \frac{y}{x}$. On représentera cette situation ainsi :



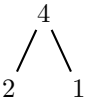
2) point de rebroussement ordinaire. En $(0,0)$, l'équation $x^3 - y^2 = 0$ a un point de rebroussement ordinaire, transformé en un point simple $(0,0)$ par la transformation $t = \frac{y}{x}$:

3) point de contact. En $(0,0)$, l'équation $y^2 - x^2y + x^5 = 0$ a un point de contact, transformé en un point double ordinaire $(0,0)$ par la transformation $t = \frac{y}{x}$:

4) Point de rebroussement et point de contact, les deux tangentes étant distinctes.



5) Point de rebroussement et point double, la tangente du point de rebroussement étant confondue avec l'une des deux tangentes du point double. On obtient le même arbre que dans l'exemple précédent :



6) Exemple étudié par Puiseux ([256], § 27 p. 408) puis repris par Hamburger ([139] p. 485) :

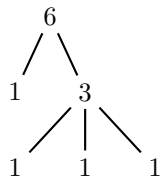
$$A(y-y_0)^7 + B(y-y_0)^5(x-x_0) + C(y-y_0)^4(x-x_0)^4 + D(y-y_0)^2(x-x_0)^5$$
$$+E(y-y_0)(x-x_0)^7 + F(x-x_0)^9 + G(y-y_0)^8 + H(y-y_0)^4(x-x_0)^5 + I(x-x_0)^{10} = 0$$

Si $E^2 - 4DF \neq 0$, on obtient l'arbre :

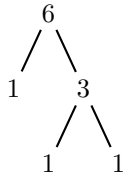
15. Soit la courbe d'équation $f(x,y) = 0$. L'ordre du point $(0,0)$ sur cette courbe est défini comme étant le degré total minimal des monômes présents dans le polynôme $f(x,y)$.

16. *Cf. infra* p. 264.

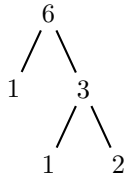
17. On s'inspire ici des explications d'Abhyankar [4], § 39 p. 425.



Si $E^2 - 4DF = 0$ et $I \neq 0$, on a :



Et si $E^2 - 4DF = 0$ et $I = 0$, on a :



Mais dans la note de 1871, Noether se contente de décrire la méthode générale en des termes encore vagues, avant de l'appliquer à quelques exemples :

> Par une telle transformation, dans la courbe transformée C', la singularité provenant du point P en tant qu'il est point multiple d'ordre ν se résout, et il reste donc sur C', correspondant au point P, des points de singularité moindre qui, réunis et alliés à un point d'ordre ν général, sont équivalentes au point P de C. Par une application répétée de transformations en les points singuliers des courbes ainsi obtenues C', C'',..., la singularité des points correspondant à P s'abaisse constamment, jusqu'à ce que finalement une suite de points simples de la courbe transformée corresponde au point P. [18]

Dans un article ultérieur, Noether qualifiera de « points infiment voisins de P » les points ainsi obtenus par la suite des transformations quadratiques :

> D'après cette définition, un point singulier est constitué d'un nombre fini de points *multiples* qui, seulement dans certaines directions, se serrent, infiniment proches, les uns contre les autres, mais de telle façon qu'il ne résulte pas de cela un point d'ordre plus élevé [19].

Noether conçoit donc une relation d'équivalence pour laquelle tout point singulier est équivalent à un certain nombre de points multiples « généraux » et de points simples. Une telle notion d'équivalence suffit certes à donner une formule universelle du genre de la courbe :

$$g = \frac{(n-1)(n-2)}{2} - \sum \frac{\nu(\nu-1)}{2}$$

où la somme s'étend à tous les points d'ordre ν « généraux » obtenus par la suite des transformations quadratiques à partir de chaque point singulier [20]. Elle suffira aussi à définir la multiplicité d'intersection de deux courbes en un point, au moyen de la formule suivante [21] :

$$\sum \mu\nu$$

18. *Cf.* [244] p. 268.
19. *Cf.* [241] p. 166–182, ici p. 167 : *Nach dieser Definition besteht ein singulärer Punkt aus einer endlichen Anzahl vielfacher Punkte, die nur in bestimmten Richtungen unendlich nahe an einander gerückt sind, aber so, dass hierdurch kein Punkt von höherer Ordnung der Vielfachheit entsteht.*
20. Cette formule du genre figure aussi dans les *Vorlesungen über die Theorie der Abelschen Transcendenten* de Weierstrass, *cf.* [326] t. 4 p. 166–171, mais sans recours au concept géométrique de point infiniment voisin.
21. *Cf.* [49], p. 383.

où la somme s'étend à tous les points communs aux deux courbes obtenus par la suite des transformations quadratiques à partir du point P, chacun de ces points étant compté comme point d'ordre μ « général » sur la première courbe et comme point d'ordre ν « général » sur la seconde [22]. Mais cette notion d'équivalence ne peut suffire à classer les singularités de manière exhaustive, comme le montre les exemples 4 et 5 ci-dessus qui conduisent au même arbre bien que les singularités soient de nature très différente. Zariski [334] a montré comment définir un critère d'équivalence plus fin et l'a utilisé dans ses recherches sur l'« équisingularité ». Sa méthode consiste à considérer séparément chaque composante irréductible locale de la courbe [23]. Deux singularités sont dites équivalentes quand les composantes irréductibles locales des deux courbes en ce point se correspondent deux à deux de sorte 1) que l'ordre du point sur chacune soit le même pour les deux courbes, 2) que deux composantes d'une courbe aient une tangente commune si et seulement si les deux composantes correspondantes de l'autre courbe ont aussi une tangente commune, et 3) que ces conditions soient encore vérifiées après chaque transformation quadratique.

Ainsi, non seulement les recherches sur le polygone de Newton, à partir de Puiseux, donnent une base solide à l'étude locale des fonctions algébriques et donnent sens (que les auteurs aient recours explicite aux développements en séries, ou bien qu'ils continuent – tel Halphen – d'utiliser l'ancien langage d'« ordre d'un infiniment petit » pour désigner l'ordre d'une fonction en un point sans faire mention explicite d'un développement en série) à la définition de la multiplicité d'intersection comme ordre du résultant calculé par la formule de Poisson – c'est ce qui manquait par exemple à Euler [24] – mais l'étude des singularités par Noether et ses successeurs permet une description précise des phénomènes de contact entre deux branches de courbe, description qui manquait à Maclaurin ou Poncelet pour donner une définition *géométrique* de la multiplicité d'intersection. Il n'est donc guère étonnant qu'après ces efforts, une théorie élémentaire des intersections de deux courbes planes, telle que nous l'avons exposée au début de ce chapitre (section 13.1), apparaisse chez plusieurs auteurs dans les années 1870 ; nos hypothèses simplificatrices vont aussi pouvoir être abandonnées, et l'on passe d'énoncés généraux à des énoncés universels, justement grâce à l'étude des singularités, et aussi, ne l'oublions pas, grâce à l'usage de la géométrie projective dans le décompte des « points à l'infini » (souvenons-nous en effet des obstacles au concept de point à l'infini rencontrés au XVIII$^{\text{e}}$ siècle, *cf. supra* p. 170).

13.3 Application : le genre (Riemann, Weierstrass, M. Noether)

En 1857, Riemann définit le genre d'une surface de Riemann de manière topologique [25], comme étant l'entier p tel qu'il faille au moins $2p$ « rétrosections » pour rendre la surface simplement connexe. Le genre est donc un invariant de la surface construite par Riemann pour étudier la fonction s de z définie par une équation algébrique $F(s,z) = 0$. La surface de Riemann est un modèle non-singulier de la courbe d'équation $F(s,z) = 0$, qui n'est elle-même jamais mentionnée explicitement par Riemann. Mais les successeurs de Riemann appliqueront la théorie du genre aux courbes algébriques, et plus généralement à des variétés de dimension quelconque [26].

22. *Cf.* [241] p. 175–176, où Noether ne détermine pas explicitement la multiplicité d'intersection mais montre bien que chaque point d'intersection contribue d'exactement $\sum \mu\nu$ facteurs linéaires au résultant, afin de montrer la finitude du nombre de transformations rationnelles dans la résolution des singularités.

23. Soit $f(x,y) = 0$ l'équation de la courbe sur un corps k. Au point $(0,0)$, les composantes irréductibles locales sont données par la factorisation du polynôme $f(x,y)$ dans l'anneau des séries entières $k[[x,y]]$.

24. *Cf. supra* p. 108.

25. *Cf.* [157] p. 169–178 pour un commentaire du mémoire de Riemann.

26. Attention à ne pas confondre ce genre géométrique et le genre arithmétique, qui coïncident pour les courbes mais pas pour les surfaces (*cf.* [140] p. 57 et 247, et [157] p. 205). Dans [140], le genre géométrique n'est défini que pour des variétés non-singulières. Pour les courbes, le genre arithmétique l'est en général (III Ex. 5.3, IV Ex. 1.8). Pour une courbe plane X de degré d (qu'elle soit ou non singulière), le genre arithmétique est $p_a(X) = \frac{1}{2}(d-1)(d-2)$ (*cf.* p. 54). Soit $\tilde{X}$ sa normalisation,

$$p_a(\tilde{X}) = \frac{1}{2}(d-1)(d-2) - \sum_{P \in X} \delta_P$$

Après avoir démontré que l'espace vectoriel des « intégrales de première espèce » sur la surface de Riemann est exactement de dimension p (*cf.* [265], section I, § IV), Riemann s'intéresse (§ V) aux fonctions méromorphes sur la surface de Riemann, ayant exactement m pôles simples donnés. Si $m \geq p+1$ l'espace vectoriel de ces fonctions est en général de dimension $(m-p+1)$. En 1864, Roch précise ce théorème qui a gardé depuis le nom de « théorème de Riemann-Roch » : l'espace vectoriel de ces fonctions est de dimension $(m-p+1+q)$ où q est la dimension de l'espace des différentielles de première espèce qui s'annulent en les m points donnés (si les m points sont dans une position suffisamment générale, $q=0$). Dans son mémoire, Riemann utilise la théorie de l'élimination au moins à trois reprises. Tout d'abord aux § VI–VII, où il calcule le genre $p=(n-1)(m-1)-r$ de la surface de Riemann associée à l'équation $F(s,z)=0$ de degré n en la variable s et m en la variable z, la courbe $F(s,z)=0$ ayant pour points singuliers r points doubles sans compter les deux points singuliers à l'infini d'ordres respectifs m et n. Riemann ne parle pas de « points doubles », puisqu'il n'est jamais question dans son mémoire de la courbe d'équation $F(s,z)=0$, mais il conçoit chacun de ces points comme lieu où deux points de ramification coïncident en « se détruisant ». Il étudie le discriminant de F, résultant de F et $\frac{\partial F}{\partial s}$. Notons $F(s,z)=a_0 s^n+...+a_n$, son discriminant est

$$a_0^{n-1}\prod_i \frac{\partial F}{\partial s}(s_i,z)$$

Riemann calcule plutôt

$$a_0^{n-2}\prod_i \frac{\partial F}{\partial s}(s_i,z)$$

ce qui a pour effet de ne pas tenir compte du point m-uple à l'infini. Le degré en z de cette expression est alors [27] $2m(n-1)$. En distinguant w points de ramifications simples et r points doubles (que Riemann conçoit comme couples de points de ramification coïncidant en se détruisant), il trouve donc

$$w+2r=2m(n-1)$$

En calculant w d'une autre manière (topologique), il trouve

$$w=2n+2(p-1)$$

D'où [28] :

$$p=(n-1)(m-1)-r$$

Or pour les courbes non-singulières $p_a=p_g$ (*cf.* p. 256). Et $\tilde{X}$ est en fait l'unique modèle non-singulier de X ([290] t. 1 p. 129). On peut donc poser par définition, pour une courbe X ayant des points singuliers,

$$p_g(X)=\frac{1}{2}(d-1)(d-2)-\sum_{P\in X}\delta_P$$

Mais pour une variété X de dimension $n>1$ non-singulière, la définition moderne du genre géométrique s'écrirait

$$p_g(X)=\dim\Gamma(X,\omega_X)$$

où ω_X est le faisceau inversible $\Lambda^n\Omega_X$, et Ω_X le faisceau des formes différentielles sur X. Mais si X est singulière, il n'y a pas, en général, un modèle non-singulier unique dans sa classe d'équivalence rationnelle. Il faut donc définir le genre autrement. Nous verrons plus loin l'approche suivie par Noether en 1870, exposée de manière plus moderne par Zariski dans le cas $n=2$ ([335], § III.8).

27. Riemann redémontre en fait, sans le dire, un cas particulier du théorème de Bézout, en calculant le degré du résultant de deux polynômes à deux indéterminées, étant donné pour chaque polynôme une borne sur son degré en chacune des indéterminées (cas particulier bien connu de Bézout, *cf.* [29] § 62). Soient deux polynômes généraux :

$$F(\overset{n}{s},\overset{m}{z})=0 \quad \chi(\overset{\nu}{s},\overset{\mu}{z})=0$$

Cette notation, qui est celle de Riemann, signifie que F est de degré n en s et m en z, et que χ est de degré ν en s et μ en z. Le résultant est alors de degré $(m+n)(\mu+\nu)-m\mu-n\nu=m\nu+n\mu$.

28. Weierstrass ([326] t. 4 p. 172–174) démontrera cette formule à partir de celle obtenue pour une équation $F(s,z)=0$ de degré total $d=m+n$, en observant que les deux points singuliers à l'infini comptent pour $\frac{1}{2}m(m-1)+\frac{1}{2}n(n-1)$ points doubles :

$$p=\frac{1}{2}(d-1)(d-2)-\sum\delta_P=\frac{1}{2}(m+n-1)(m+n-2)-\frac{1}{2}m(m-1)-\frac{1}{2}n(n-1)-r$$

Riemann a recours de nouveau à la formule de Poisson et au théorème de Bézout au § VIII, pour déterminer la dimension de l'espace des fonctions rationnelles $\frac{\psi}{\chi}$ ayant m pôles donnés sur la surface de Riemann. Il trouve que ce nombre est aussi $(m-p+1)$, donc toutes les fonctions méromorphes sur la surface de Riemann sont rationnelles. Cela lui permet (§ IX) d'établir la forme prise par les intégrales de première espèce :

$$\int \frac{\phi(s,z)dz}{\frac{\partial F}{\partial s}} = \int \frac{\phi(s,z)ds}{\frac{\partial F}{\partial z}}$$

où ϕ est de degré $(n-2)$ en s et $(m-2)$ en z, et s'annule en les points doubles.

Enfin, au § X, du théorème de Bézout appliqué à F et ϕ, Riemann déduit que le nombre de zéros d'une différentielle de première espèce de la forme

$$\frac{\phi(s,z)dz}{\frac{\partial F}{\partial s}}$$

est $(2p-2)$.

Dans ses *Vorlesungen über die Theorie der Abelschen Transcendenten*[29], Weierstrass développe une théorie des intersections de deux courbes planes au chapitre VI. La notion de base des *Vorlesungen* est celle d'*algebraisches Gebilde* (*cf.* chapitre I), « figure algébrique » dont les « points » sont les solutions d'une équation $f(x,y)=0$ de degré r, mais qui est avant tout l'ensemble des « éléments » centrés en chacun de ces points (a,b), un élément étant défini comme l'ensemble des représentations paramétriques locales équivalentes en séries entières

$$x=\phi(t), \quad y=\psi(t)$$

où $(\phi(0),\psi(0))=(a,b)$. Il ne faut pas confondre la notion d'« élément » et celle de « branche » d'une courbe que nous avons employée au début de ce chapitre. Ainsi pour l'*algebraisches Gebilde* d'équation $y^2=x^3$, il n'y a qu'un seul élément centré en $(0,0)$, soit $(x,y)=(t^2,t^3)$, mais la courbe en question a deux branches $y=\pm x^{\frac{3}{2}}$. L'usage des transformations quadratiques dans la résolution des singularités est exposé par Weierstrass de manière élégante grâce à son « théorème de préparation »[30]. Soit un *algebraisches Gebilde* d'équation

$$f_0(\xi,\eta)=0$$

ayant un point singulier en $(0,0)$. Ecrivons la forme de degré μ minimal de f_0 :

$$(\xi,\eta)_\mu = C(g\eta-h\xi)^{\mu_0}(g'\eta-h'\xi)^{\mu_1}...(g^{(\nu)}\eta-h^{(\nu)}\xi)^{\mu_\nu}$$

Weierstrass applique une transformation quadratique de la forme :

$$\begin{cases} \xi=(g+\alpha\eta_1)\xi_1 \\ \eta=(h+\beta\eta_1)\xi_1 \end{cases}$$

Alors $f_0(\xi,\eta)=\xi_1^\mu f_1(\xi_1,\eta_1)$, avec $f_1(0,\eta_1)$ divisible par $\eta_1^{\mu_0}$, et le théorème de préparation affirme qu'il existe donc une factorisation :

$$f_1(\xi_1,\eta_1)=(\eta_1^{\mu_0}+G_1(\xi_1)\eta_1^{\mu_0-1}+...+G_{\mu_0}(\xi_1))\overline{G}(\xi_1,\eta_1)$$

où $\overline{G}(0,0)\neq 0$, $G_1(0)=...=G_{\mu_0}(0)=0$, et les G_i sont des séries entières. Cela « détermine μ_0 valeurs infiniment petites de η_1 pour chaque valeur infiniment petite de ξ_1 »[31]. Quand $\mu_0=1$,

29. *Cf.* [326], t. 4. Pour un commentaire de ce texte, *cf.* [157] p. 178–184.

30. *Cf.* [326], t. 4 p. 19–24. Le *Vorbereitungssatz* est un théorème concernant les fonctions de plusieurs variables (au moins deux), que semble avoir enseigné Weierstrass dès 1860 (d'après [49] p. 403) ; il est exposé dans un texte autographe de 1879 dans [326] t. 2 p. 135–188, puis dans les *Abhandlungen aus der Functionenlehre*. Quant au devenir de ce théorème, *cf.* [51] p. 155–168.

31. *Cf.* [326] t. 4 p. 22 : *... werden zu jedem unendlich kleinen Werthe von* ξ_1 μ_0 *unendlich kleine Werthe von* η_1 *bestimmt.*

ce même théorème de préparation donne un développement en série entière [32] de $\eta_1 = -G_1(\xi_1)$. Weierstrass montre ensuite qu'un nombre fini de transformations quadratiques suffit toujours à résoudre complètement un point singulier [33].

Clebsch avait déjà traduit la théorie des fonctions algébriques de Puiseux et Riemann dans le langage de la géométrie des courbes [34]. Le concept de courbe n'est alors pas introduit autrement par Weierstrass que par ces quelques mots : « En géométrie analytique, où ces *Gebilde* sont des courbes, on en est d'abord venu naturellement à les classer par degré... » [35]. A plusieurs endroits des *Vorlesungen*, les deux notions semblent coïncider, et l'on peut donc concevoir l'*algebraisches Gebilde* comme courbe algébrique à laquelle on adjoint toutes ses représentations paramétriques locales. L'on peut douter, dans les premiers chapitres, du sens à donner à la notion de « point à l'infini » (*unendlich ferne Stelle*). Weierstrass distingue trois formes de points à l'infini [36] :

$$(a, \infty), \quad (\infty, b), \quad (\infty, \infty)$$

Mais faute d'utiliser des coordonnées homogènes pour appréhender le plan projectif, il n'indique pas clairement la structure de la droite à l'infini. Si $a \neq a'$, pour Weierstrass, (a, ∞) et (a', ∞) sont-ils distincts ou confondus ? Et comment décider si deux points de la forme (∞, ∞) sont distincts ou confondus ? Ainsi, en particulier, comment décider si un point à l'infini est le centre de plusieurs « éléments » distincts, et est donc un point singulier ? Supposer (a, ∞) distinct de (a', ∞) revient – en langage moderne – à travailler dans $\mathbb{P}^1 \times \mathbb{P}^1$, ce qui est assez naturel si l'on conçoit l'*algebraisches Gebilde* comme graphe d'une fonction multivaluée $\mathbb{P}^1 \longrightarrow \mathbb{P}^1$. Mais quand Weierstrass se tourne davantage vers la géométrie des courbes, il semble au contraire travailler dans $\mathbb{P}^2$, bien qu'il n'introduise pas de nouvelle notation. C'est particulièrement clair au chapitre VI où il applique une transformation birationnelle linéaire (une « colinéation ») [37] :

$$\begin{cases} u = a_0 + a_1 x + a_2 y \\ v = b_0 + b_1 x + b_2 y \\ w = c_0 + c_1 x + c_2 y \end{cases}$$

$$\xi = \frac{u}{w}, \quad \eta = \frac{v}{w}, \quad f(x, y) = w^r \phi(\xi, \eta)$$

Il définit seulement alors un « point singulier à l'infini » de $f(x, y) = 0$ comme étant un point à l'infini (x, y) tel que le point fini correspondant (ξ, η) de la courbe transformée $\phi(\xi, \eta) = 0$ soit un point singulier. Weierstrass définit toutefois dès le premier chapitre un *algebraisches Gebilde* « régulier à l'infini » (*im Unendlichen regulär*), c'est-à-dire tel que la courbe de degré r associée ait ses r points à l'infini tous distincts, en imposant que la forme dominante de $f(x, y)$ soit factorisable en r facteurs linéaires distincts.

En choisissant la droite à l'infini au moyen de la colinéation, Weierstrass peut même imposer que les $(r + r')$ points à l'infini de deux courbes $f(x, y) = 0$ et $F(x, y) = 0$ de degrés respectifs r et r' soient tous distincts. Les deux courbes vérifient alors les hypothèses faites au début de notre chapitre : pas de ramification à l'infini, pas de branche parabolique, pas de direction asymptotique commune.

La multiplicité d'intersection d'un point fini (x_0, y_0) est conçue par Weierstrass, sans être nommée, comme étant l'« ordre » (*Ordnungszahl*) de la fonction $F(x, y)$ en le point (x_0, y_0) de l'*algebraisches Gebilde* $f(x, y) = 0$, c'est-à-dire

$$\sum \mathrm{ord}_t F(\phi(t), \psi(t))$$

où la somme s'étend à tous les éléments $(\phi(t), \psi(t))$ de l'*algebraisches Gebilde* centrés en (x_0, y_0). Cette définition porte trois conséquences. La multiplicité d'intersection est indépendante des paramétrages locaux choisis [38] et *a fortiori* du système de coordonnées x, y ; il n'est plus besoin

32. *Cf.* [326] t. 4 p. 15.
33. *Cf.* [326] t. 4 p. 25–28.
34. *Cf.* [157] p. 184–186.
35. *Cf.* [326] t. 4, p. 59. *cf.* aussi p. 62 : *wenn man das durch die Gleichung* $f(x, y) = 0$ *definirte algebraische Gebilde als Curve auffasst.*
36. *Cf.* p. 33
37. *Cf.* [326] t. 4 p. 149 et 163.
38. *Cf.* [326] t. 4, p. 55

d'hypothèse sur l'absence de tangente verticale. Ce qui fait dire à Weierstrass :

> Ce qui précède montre clairement que l'on ne peut faire l'économie du concept d'élément. Car c'est seulement avec son aide que l'on obtient la détermination juste des ordres. [39]

D'autre part, les paramétrages étant des séries entières, la multiplicité d'intersection est elle-même un entier. Par contre, cette définition de la multiplicité comme ordre d'une fonction n'est pas invariante par projection, et ne peut donc s'étendre sans précaution aux points d'intersection à l'infini que Weierstrass doit traiter séparément en les transformant en des points finis au moyen d'une colinéation [40].

Après quoi le raisonnement de Weierstrass suit les deux moments du nôtre, mais dans le langage, légèrement différent, de la théorie des fonctions algébriques. 1°) Sur l'*algebraisches Gebilde* $f(x,y)=0$, « le nombre de points en lesquels une fonction rationnelle [ici $F(x,y)$] de (x,y) prend une valeur finie donnée ou bien est infinie, est égal au nombre de points où elle est nulle, chaque point étant compté avec son ordre » [41]. Au moins dans le cas où $f(x,y)=0$ n'a pas d'asymptote verticale, ce nombre n'est autre que le degré en x du résultant de f et F ; Weierstrass l'appelle « degré (*Grad*) de la fonction rationnelle F » (p. 47–50). 2°) Dans le cas où les $(r+r')$ points à l'infini sont tous distincts, le nombre de points où $F(x,y)$ est infini, c'est-à-dire le nombre de pôles, chacun étant compté « avec son ordre », est rr' (p. 152).

Le théorème que l'on vient d'énoncer se généralise d'ailleurs aux formes différentielles sur la courbe [42] :

$$\text{nombre de zéros} = \text{nombre de pôles} + 2p - 2$$

En particulier pour les formes de première espèce

$$\text{nombre de zéros} = 2p - 2$$

comme l'avaient remarqué Weierstrass (p. 125 et 589) et Riemann avant lui.

Après avoir ainsi traduit le théorème de Bézout dans la théorie des fonctions algébriques, Weierstrass l'applique au calcul du genre [43].

En 1869, Noether cherche à démontrer qu'il est possible de généraliser la notion de genre introduite par Riemann à une hypersurface dans un espace projectif de dimension quelconque, et que le genre ainsi défini est alors un invariant birationnel. Il définit le genre p d'une hypersurface de dimension r comme étant la dimension de l'espace vectoriel des r-formes différentielles partout finies sur la variété. Soit $f(x_1,...,x_{r+2})=0$ l'équation homogène irréductible de l'hypersurface, de degré n. On choisit des paramètres λ_1,...,λ_r définissant un paramétrage local de l'hypersurface. Au moyen de ce paramétrage, Noether exprime les r-formes différentielles finies sous la forme suivante,

39. *Cf.* [326] t. 4, p. 56 : *Aus dem Bisherigen ist ersichtlich, daß der Begriff eines Elementes nicht entbehrt werden kann. Denn nur mit dessen Hülfe ergiebt sich die richtige Bestimmung der Ordnungszahlen.*

40. *Cf.* [326] t. 4, p. 153–155. Nous comprendrons mieux ce phénomène, que Weierstrass n'explique pas très clairement, avec l'exemple suivant. Soit $f(x,y)=y+1$, et $F(x,y)=y$. Dans $\mathbb{P}^1\times\mathbb{P}^1$ en $x=\infty$, l'*algebraisches Gebilde* $f(x,y)=0$ est paramétré par l'uniformisante $t=\frac{1}{x}$, avec $(x=\frac{1}{t}, y=-1)$. Alors la fonction $F\left(\frac{1}{t},-1\right)=-1$ est d'ordre 0 en ce point. Il semblerait donc qu'il faille définir la multiplicité d'intersection de ces deux courbes en leur point à l'infini comme étant zéro, ce qui est toutefois contraire à l'intuition du plan projectif. Mais, avec Weierstrass, déplaçons la droite à l'infini en la droite d'équation $x=0$ au moyen de la transformation birationnelle linéaire

$$\begin{cases} u=y \\ v=1 \\ w=x \end{cases}$$

$$\xi=\frac{y}{x},\quad \eta=\frac{1}{x},\quad f(x,y)=w\phi(\xi,\eta),\quad F(x,y)=w\Phi(\xi,\eta)$$

On a $\phi(\xi,\eta)=\xi+\eta$ et $\Phi(\xi,\eta)=\xi$. Le point à l'infini $(x,y)=(\infty,-1)$ de l'*algebraisches Gebilde* $f(x,y)=0$ devient par cette transformation $(\xi,\eta)=(0,0)$. L'élément centré en $(0,0)$ de $\phi(\xi,\eta)=0$ est paramétré par l'uniformisante ξ, et l'ordre de $\Phi(\xi,\eta)$ en ce point est donc 1. Weierstrass définit donc la multipicité d'intersection de $y=0$ et $y+1=0$ en leur point à l'infini comme étant égale à 1 et non à 0.

41. *Cf.* [326] t. 4 p. 56.

42. *Cf.* [140] IV.1.3.3.

43. *Cf.* [326] t. 4 p. 156–157, où Weierstrass redémontre de manière purement algébrique la formule du genre d'une courbe plane de degré et de singularités donnés que Riemann avait déjà déterminée.

symétrique vis-à-vis des coordonnées, reprenant semble-t-il [44] une idée due à Aronhold puis utilisée systématiquement par Clebsch et Gordan [77] dans le cas $r = 1$:

$$\frac{\Theta \sum \pm c_1 x_2 \frac{\partial x_3}{\partial \lambda_1} \frac{\partial x_4}{\partial \lambda_2} ... \frac{\partial x_{r+2}}{\partial \lambda_r}}{\sum_i c_i \frac{\partial f}{\partial x_i}} d\lambda_1 d\lambda_2 ... d\lambda_r$$

où c est un vecteur arbitraire et Θ un polynôme homogène de degré $(n - r - 2)$, vérifiant certaines conditions locales en le lieu singulier de l'hypersurface. Pour démontrer l'invariance birationnelle du genre, Noether étudie l'effet d'une transformation birationnelle

$$x_1 : ... : x_{r+2} = \phi_1(y_1, ..., y_{r+2}) : ... : \phi_{r+2}(y_1, ..., y_{r+2})$$

sur cette expression. L'équation $f(x) = 0$ se transforme en une équation de la forme $M.F(y) = 0$, avec $F(y)$ irréductible. L'expression ci-dessus devient alors :

$$\frac{\Theta.D. \sum \pm k_1 y_2 \frac{\partial y_3}{\partial \lambda_1} \frac{\partial y_4}{\partial \lambda_2} ... \frac{\partial y_{r+2}}{\partial \lambda_r}}{s.M. \sum_i k_i \frac{\partial F}{\partial y_i}} d\lambda_1 d\lambda_2 ... d\lambda_r$$

avec :

$$\begin{array}{l} D = \det d\phi \\ (\forall j) \quad k_j = \sum c_i \frac{\partial y_j}{\partial x_i} \\ s = \deg \phi \end{array}$$

Le problème revient alors à montrer que M divise ΘD modulo F, et que $\frac{\Theta D}{M}$ peut donc être remplacé par un polynôme dans cette expression, polynôme devant de plus vérifier les conditions locales requises en le lieu singulier de l'hypersurface transformée.

Nous ne sommes pas capable de juger *en général* de la validité des conclusions de Noether, qui de toute façon, comme nous le verrons lors de l'énoncé des conditions locales vérifiées par Θ, se limite explicitement à des hypersurfaces pas trop singulières. Au moins au sein de la classe des hypersurfaces non-singulières, le genre défini par Noether est bien un invariant birationnel (*cf.* [140] II.8.19). Dans le cas des surfaces ($r = 2$), Zariski donne un exposé moderne de la théorie de Noether (*cf.* [335] III.8). Nous voudrions seulement, avant de reprendre le fil des recherches de Noether, donner une justification moderne (en termes de faisceaux) de l'expression symétrique des r-formes différentielles ci-dessus, dans le cas non-singulier, r quelconque.

Soit donc une hypersurface de degré d, $Y \subset X = \mathbb{P}^{r+1}$. Pour la topologie de Zariski, il n'existe pas, en général, de « paramétrage local » de Y. On considère en fait un morphisme $\lambda : U \longrightarrow \mathbb{A}^r$ qui fait d'un ouvert de Zariski U de Y un revêtement ramifié de $\mathbb{A}^r$ (c'est le lemme de normalisation de Noether qui affirme l'existence d'un tel morphisme ; c'est aussi une conséquence de la méthode d'élimination de Kronecker, *cf. infra* p. 238). Localement (c'est-à-dire sur un ouvert de Zariski de $\mathbb{A}^r$), on peut relever les sections $\delta\lambda_1$,...,$\delta\lambda_r$ du faisceau tangent $\mathcal{T}_{\mathbb{A}^r}$ en des sections de $\mathcal{T}_Y$. Le faisceau normal $\mathcal{N}_Y = (\mathcal{T}_X \otimes \mathcal{O}_Y) / \mathcal{T}_Y$ se relève localement en un sous-faisceau de $\mathcal{T}_X \otimes \mathcal{O}_Y$ engendré par une section δc que l'on peut choisir de sorte que $\mathcal{T}_X \otimes \mathcal{O}_Y = \delta c \oplus \mathcal{T}_Y$ (dans ce qui suit, on sous-entendra désormais la tensorisation par $\mathcal{O}_Y$, pour alléger les notations). Ce choix permet de relever les sections $d\lambda_1$,...,$d\lambda_r$ du faisceau Ω_Y des formes différentielles sur Y à Ω_X (plus exactement, à $\Omega_X \otimes \mathcal{O}_Y$), de sorte que $(\forall i) \quad d\lambda_i.\delta c = 0$. On considère la suite exacte (*cf.* [140] II.8.13)

$$\begin{array}{ccccccc} 0 \longrightarrow \Omega_X & \longrightarrow & \mathcal{O}_X(-1)^{r+2} & \longrightarrow & \mathcal{O}_X & \longrightarrow 0 \\ & & e_i & \longmapsto & x_i & \end{array}$$

où $\mathcal{O}_X(-1) = \mathcal{O}_X(1)^\vee$ est le dual du *twisting sheaf*, et où $< e_1, ..., e_{r+2} >$ est la base canonique de $\mathcal{O}_X(-1)^{r+2}$. On choisit une forme linéaire $w = \sum w_i x_i$ (on pourra alors donner un sens aux rapports

44. *Cf.* [49] p. 301.

différentiels $\frac{\partial x_i}{\partial \lambda_j}$, en se plaçant sur l'ouvert affine $w \neq 0$ de $\mathbb{P}^{r+1}$ et en imposant $(\forall j)\ \frac{\partial w}{\partial \lambda_j} = 0$). Alors localement

$$\begin{array}{l} d\lambda_1, ..., d\lambda_r, \\ dw = \sum w_i e_i, \\ df = \sum \frac{\partial f}{\partial x_i} e_i \end{array}$$

forment une base de $\mathcal{O}_X(-1)^{r+2}$. Le choix de dw permet de relever $\delta\lambda_1$,...,$\delta\lambda_r$,δc de $\mathcal{T}_X$ à $\mathcal{O}_X(1)^{r+2}$ au moyen de la suite exacte, duale de la précédente (*cf.* [140] II.8.20.1) :

$$\begin{array}{ccccccc} 0 \longrightarrow & \mathcal{O}_X & \longrightarrow & \mathcal{O}_X(1)^{r+2} & \longrightarrow \mathcal{T}_X \longrightarrow 0 \\ & 1 & \longmapsto & \sum f_i x_i & & \end{array}$$

où $< f_1, ..., f_{r+2} >$ est la base duale de $< e_1, ..., e_{r+2} >$, de sorte que $(\forall i)\ dw.\delta\lambda_i = 0$ et $dw.\delta c = 0$. Notons $c_i = e_i.\delta c$. Alors

$$\left\langle \delta\lambda_1, ..., \delta\lambda_r, \frac{\delta c}{\sum \frac{\partial f}{\partial x_i} c_i}, \frac{\sum f_i x_i}{w} \right\rangle$$

est la base duale de $< d\lambda_1, ..., d\lambda_r, df, dw >$. Alors, en remarquant que

$$e_i - (e_i.\frac{\sum f_i x_i}{w})dw = \frac{we_i - x_i dw}{w} = wd\left(\frac{x_i}{w}\right)$$

et donc que

$$e_i.\delta\lambda_j = \frac{\partial x_i}{\partial \lambda_j}$$

on a finalement

$$e_1 \wedge ... \wedge e_{r+2} = \begin{vmatrix} c_1 & \cdots & c_{r+2} \\ x_1 & \cdots & x_{r+2} \\ \frac{\partial x_1}{\partial \lambda_1} & \cdots & \frac{\partial x_{r+2}}{\partial \lambda_1} \\ \vdots & & \vdots \\ \frac{\partial x_1}{\partial \lambda_r} & \cdots & \frac{\partial x_{r+2}}{\partial \lambda_r} \end{vmatrix} . \frac{d\lambda_1 \wedge ... \wedge d\lambda_r \wedge df \wedge dw}{w \sum \frac{\partial f}{\partial x_i} c_i}$$

Nous noterons $\Delta_x(c)$ le déterminant ci-dessus. Notons ω_X, ω_Y les puissances extérieures maximales (« faisceaux canoniques ») de Ω_X, Ω_Y. On a un isomorphisme (appliquer [140] exercice II.5.16d à la suite exacte [140] II.8.13 reproduite ci-dessus)

$$\bigwedge^{r+2} \mathcal{O}_X(-1)^{r+2} \simeq \omega_X$$

qui identifie aux sections globales $\Theta e_1 \wedge ... \wedge e_{r+2}$ de

$$\bigwedge^{r+2} \mathcal{O}_X(-1)^{r+2} \otimes \mathcal{O}_X(d) \otimes \mathcal{O}_Y$$

les sections de $\omega_X \otimes \mathcal{O}_X(d) \otimes \mathcal{O}_Y$ de la forme :

$$\Theta \frac{\Delta_x(c)}{\sum \frac{\partial f}{\partial x_i} c_i} d\lambda_1 \wedge ... \wedge d\lambda_r \wedge df$$

(où Θ est un polynôme de degré $d - r - 2$). Notons de plus $\mathcal{I}$ l'idéal de Y. Alors l'isomorphisme $\omega_X \otimes \mathcal{O}_Y \simeq \omega_Y \otimes \mathcal{I}/\mathcal{I}^2$ (*cf.* [140] II.8.20 ; $\mathcal{I}/\mathcal{I}^2$ est le faisceau conormal, dual de $\mathcal{N}_Y$) permet de conclure :

$$\begin{array}{ccc} \omega_X \otimes \mathcal{O}_Y \otimes \mathcal{O}_X(d) \to \omega_Y \otimes \mathcal{I}/\mathcal{I}^2 \otimes \mathcal{O}_X(d) \simeq & & \omega_Y \\ \Theta e_1 \wedge ... \wedge e_{r+2} & \longmapsto & \Theta \frac{\Delta_x(c)}{\sum \frac{\partial f}{\partial x_i} c_i} d\lambda_1 \wedge ... \wedge d\lambda_r \end{array}$$

Autrement dit, la formule écrite par Noether donne bien l'expression *locale* (sur l'ouvert « paramétré » par λ) des sections *globales* de ω_Y. On a d'ailleurs ainsi redémontré l'isomorphisme ([140] II.8.20.3)

$$\omega_Y \simeq \mathcal{O}_Y(d-r-2).$$

Revenons à l'article de Noether. Il prétend que sa théorie s'applique aux surfaces contenant des courbes multiples, des « nœuds » (*höhern conischen Knotenpunkten*), des courbes de rebroussements, et de même en dimension supérieure [45], mais il exclut tout de même dans le cours de l'article certaines « singularités spéciales », telles que « celles qui interviennent dans une surface lorsque coïncident des plans tangents le long d'une courbe multiple ou lorsqu'un nœud conique (*conische Knotenpunkt*) devient un nœud plan » [46]. En chaque composante de dimension complexe h du lieu singulier en laquelle f s'annule à l'ordre μ, la condition locale requise pour Θ est alors qu'il s'annule à l'ordre $(\mu-r+h)$. La démonstration d'invariance du genre est alors détaillée par Noether dans le cas $r=3$: il s'agit de montrer qu'en chaque sous-variété de dimension complexe h de l'hypersurface $F(y)=0$, en laquelle F s'annule à l'ordre μ et M à l'ordre q, ΘD s'annule à l'ordre $(q+\mu-r+h)$, et que d'autre part, sous ces conditions, M divise ΘD modulo F. Nous ne pouvons ici nous attarder sur la première de ces deux assertions [47]. Quant à la seconde, elle consiste en le théorème qui serait plus tard appelé « théorème de Noether ». Noether écrit [48] :

> ...il faudrait encore démontrer cette proposition constamment utilisée en théorie des courbes et des surfaces : ΘD se laisse alors nécessairement mettre sous la forme $\Theta D = AM + BF$.

Mais la démonstration qu'en propose ensuite Noether est plus une esquisse, qui ne peut convaincre le lecteur que dans le cas le plus simple où F, M s'annulent seulement à l'ordre 1, et ΘD au moins à l'ordre 1, en chaque composante (de dimension complexe $r-1$) de l'intersection de $F=0$ et $M=0$.

Noether revient sur cette question en 1872, dans un article intitulé *Über einen Satz aus der Theorie der algebraischen Functionen* [240]. Il s'intéresse alors surtout au cas $r=1$ (théorie des courbes), et précise son théorème : si, en chaque composante de l'intersection de $\phi=0$ et $\psi=0$, f s'annule à l'ordre $(q+r-1)$ où q et r désignent les ordres auxquels s'annulent respectivement ϕ et ψ en cette composante, alors il existe des polynômes A et B tels que $f = A\phi + B\psi$. Ce théorème a parfois gardé le nom de « théorème $A\phi+B\psi$ ». Ici encore, Noether ne donne qu'une esquisse de démonstration en dimension quelconque. Par contre, il traite le cas de deux courbes planes $\phi=0$, $\psi=0$ de manière détaillée. Il montre que l'existence de polynômes A et B tels que $f = A\phi + B\psi$ dépend en fait uniquement du comportement local de f au voisinage des points d'intersection des deux courbes : deux tels polynômes existent si et seulement si en chaque point $P=(x_0,y_0)$ d'intersection il existe deux séries entières A', B' en $(x-x_0)$, $(y-y_0)$ telles que $f = A'\phi + B'\psi$. Pour ce faire, il est amené à exprimer les conditions d'existence de deux telles séries entières, par la méthode des coefficients indéterminés, sous forme d'un système d'équations linéaires en les coefficients de A' et B'. L'élimination des inconnues dans ce système conduit alors aux conditions que doivent vérifier les coefficients de f pour permettre la représentation $f = A'\phi + B'\psi$. Noether démontre qu'en général, c'est-à-dire mis à part « les cas particuliers dans lesquels plus de rq points d'intersection des deux courbes ϕ, ψ coïncident au point P », le nombre de ces équations de conditions est justement rq, où q et r désignent les ordres auxquels s'annulent respectivement

45. *Cf.* [239] p. 294.

46. *Cf.* [239] p. 304. Pour des variétés plus générales, le problème est laissé en suspens par Noether (*cf.* aussi p. 315).

47. Voici la liste des cas de figures envisagés par Noether ([239] § 5, 6) pour $r=3$. Il distingue d'abord les cas où une sous-variété de F correspond à une sous-variété de f de dimension inférieure. En une surface de F correspondant à un point d'ordre μ de f, Θ s'annule à l'ordre $(\mu-1)$, D à l'ordre 3 et M à l'ordre μ ; en une courbe de F correspondant à un point de f, D s'annule à l'ordre 2 ; en une surface de F correspondant à une courbe de f d'ordre μ, Θ s'annule à l'ordre $(\mu-1)$, de même que M. Il considère ensuite les cas où une sous-variété de F correspond à une sous-variété de f de dimension supérieure. Quand il s'agit d'un point d'ordre $\mu>2$ sur F, D s'y annule 2 fois, M $(n-\mu)$ fois, et Θ $(n-5)$ fois (mais $(n-3)$ fois si $\mu=1$, et $(n-4)$ fois si $\mu=2$) ; quand il s'agit d'une courbe d'ordre $\mu>1$ sur F, D s'y annule 3 fois, M $(n-\mu)$ fois, et Θ $(n-5)$ fois (mais $(n-4)$ fois si $\mu=1$) ; quand il s'agit d'une surface d'ordre μ sur F, D s'y annule 4 fois, M $(n-\mu)$ fois, et Θ $(n-5)$ fois.

48. *es müsste noch der in Curve- und Flächentheorie fortwährend gebrauchte Satz dargethan werden, dass ΘD sich deshalb nothwendigen Weise in die Form setzen lässt : $\Theta D = AM + BF$.*

ϕ et ψ en P. On verra que le nombre des équations de condition n'est autre que la multiplicité d'intersection telle que la définit Macaulay.

Chapitre 14

La multiplicité d'intersection

14.1 Le concept de multiplicité d'intersection

La multiplicité d'intersection est le « nombre avec lequel on compte » les répétitions d'un phénomène fini. Ces répétitions sont primitivement indiscernables : l'intuition géométrique ne donne qu'un seul « point » en chaque point d'intersection. L'objet du décompte n'est pas l'occurence, toujours solitaire, de ce phénomène primitif, mais celles d'autres phénomènes en nombre fini qui sont reliés au premier dans une expérience. C'est cette expérience qui varie d'une définition à l'autre de la multiplicité d'intersection. Prenons l'exemple de la parabole d'équation $y-(x-a)^2=0$ et de sa tangente $y=0$ au point $P=(a,0)$. Si l'on adopte la définition basée sur la formule de Poisson, le point P compte pour autant qu'il y a de facteurs linéaires $(x-a)$ dans le résultant en x, c'est-à-dire pour deux ; on a ainsi greffé à l'intuition géométrique du point une notion algébrique, la multiplicité d'une racine d'une équation. Si l'on adopte la définition basée sur la notion de « point infiniment voisin », on conçoit à l'intersection, outre le point P, un « point P' infiniment voisin dans la direction de la tangente ». L'intuition sous-jacente, éloignée de l'intuition géométrique primitive, semble très leibnizienne : deux points P, P' indiscernables dans l'étendue et pourtant distincts, que l'on représenterait peut-être aux extrémités d'un segment infiniment petit de la tangente. Si l'on adopte, à la manière de Poncelet, une définition de la multiplicité d'intersection basée sur le principe de continuité, on comptera non pas le nombre de points d'intersection avec la tangente elle-même, mais le nombre de points d'intersection avec une droite quelconque voisine de la tangente. Enfin, nous verrons une autre définition de la multiplicité d'intersection, qui a son origine dans les travaux de Noether puis Macaulay. La théorie des *modular systems* de Macaulay s'interprète aisément en termes d'anneaux et d'idéaux, et la multiplicité d'intersection est alors conçue, dans l'exemple ci-dessus, comme dimension du produit tensoriel des anneaux des variétés intersectées : on a ainsi traduit l'opération – géométrique – « intersection » en une opération – algébrique – de « tensorisation ». Dans notre cas

$$\dim\left(\mathbb{C}[x,y]/(y-(x-a)^2)\otimes_{\mathbb{C}[x,y]}\mathbb{C}[x,y]/(y)\right)=\dim\mathbb{C}[x]/((x-a)^2)=2.$$

Dans cette section, pour expliquer de manière plus précise le lien historique et mathématique entre les différentes définitions de la multiplicité d'intersection, nous allons choisir une notation unique. Soient $X=\mathbb{A}^2$ le plan affine, Y et Z les deux courbes, et $P=(a,b)$ le point d'intersection considéré (un point fini). Soient $f(x,y)=0$ et $g(x,y)=0$ les équations de Y et Z. La première définition possible suppose un système de coordonnées (x,y) suffisamment général :

$$(I_1)\quad I(Y,Z;P)=\mathrm{ord}_{(x-a)}\mathrm{Res}(f,g)$$

Cette définition est précisée par le recours à la formule de Poisson pour le résultant, mais reste inadaptée jusqu'aux années 1870 où, comme nous l'avons vu, l'étude locale des courbes, en théorie des fonctions algébriques, a permis de récrire (I_1) sous la forme

$$(I_2)\quad I(Y,Z;P)=\sum\mathrm{ord}_t g(\phi(t),\psi(t))$$

où la somme s'étend à toutes les paramétrisations locales $(\phi(t), \psi(t))$ essentiellement distinctes de la courbe Y en P. C'est ainsi que procède Weierstrass en 1873. D'autre part, la multiplicité d'intersection se traduit alors dans le langage de la théorie des fonctions algébriques comme *Ordnungszahl* de la fonction $g(x, y)$ au point P de Y ; dans le langage plus récent de la théorie des diviseurs, elle se traduit comme coefficient de P dans le diviseur (g) sur la courbe $f = 0$. Il est certes possible de justifier (I_2) par le recours à une autre définition. Ainsi Weierstrass remarque que ce nombre est égal au nombre de zéros de $(g(x, y) - s)$ infiniment proches de P pour s infiniment petit, ce qui montre en fait l'équivalence de cette définition avec une définition de la multiplicité d'intersection basée sur le principe de continuité. Mais en géométrie algébrique où l'objet géométrique est toujours conçu avec son équation, ces définitions (I_1) et (I_2) s'imposent aussi d'elles-mêmes à notre intuition. Ainsi Euler pouvait-il repérer les lignes de contacts entre deux surfaces par la présence d'un facteur carré dans le résultant, et justifiait-il sa démarche *sans* avoir recours au principe de continuité :

> Ces sortes de contacts se connaîtraient par l'équation même, si on en trouve une pour la projection, qui contienne deux racines réelles, parce que ce contact n'est rien autre chose que le concours de deux intersections. [1]

La notion de « point infiniment voisin » apparue chez Noether dans une note de 1871 a permis, comme nous l'avons vu, une autre définition de la multiplicité d'intersection :

$$(I_3) \quad I(Y, Z; P) = \sum \mu\nu$$

L'intuition sous-jacente à cette définition, que nous avons qualifié de « leibnizienne » ci-dessus, et qui décrit bien le phénomène de multiplicité en un point de contact entre deux courbes, n'éclaire cependant pas le fait qu'un point « général » d'ordre μ sur l'une des deux courbes et « général » d'ordre ν sur l'autre doive être compté avec multiplicité $\mu\nu$.

Sans passer par la notion de point infiniment voisin, en supposant au contraire déjà connus les développements en séries de deux branches de courbes en un point (calculés par la méthode du polygone de Newton ou bien par celle de Hamburger), il est possible de donner une formule (I_4) équivalente à (I_3) et exprimant la multiplicité d'intersection de deux composantes irréductibles locales en ce point en fonction des exposants successifs intervenant dans les développements en séries. C'est la théorie des « paires caractéristiques » développées par Noether, Smith, Halphen.

La théorie des « systèmes inverses » (*inverse systems*) de Macaulay (1916) permet une définition de la multiplicité d'intersection que l'on pourrait écrire ainsi aujourd'hui :

$$(I_5) \quad I(Y, Z; P) = \dim_{\mathbb{C}} \left(\mathbb{C}[[x, y]]/(f, g)\right).$$

(I_1) et (I_5) se généralisent d'ailleurs en dimension supérieure, au moins dans les cas où la théorie classique du résultant s'applique (c'est-à-dire n équations à n inconnues ayant un ensemble fini de solutions). (I_5) s'écrit alors :

$$I(H_1, H_2, ..., H_n; P) = \dim_{\mathbb{C}} \left(\mathbb{C}[[x_1, x_2, ..., x_n]]/(f_1, f_2, ..., f_n)\right).$$

Macaulay montre d'ailleurs l'équivalence entre (I_1) et (I_5) dans ce cas. Ajoutons que, lorsque les deux courbes n'ont qu'un seul point d'intersection fini, (I_5) peut aussi s'écrire :

$$I(Y, Z; P) = \dim_{\mathbb{C}} \left(\mathbb{C}[x, y]/(f, g)\right).$$

Nous allons décrire comment s'exprimait Macaulay dans son formalisme des « systèmes inverses ». En fait, ce que Macaulay nommait *multiplicity* n'est pas une multiplicité d'intersection proprement dite. Il s'agit d'un nombre attaché à chaque « module simple », c'est-à-dire à chaque idéal primaire de dimension 0. Un idéal quelconque dont l'ensemble des zéros est un ensemble fini de points se décompose en idéaux primaires de dimension 0 correspondant à chacun des points (par le théorème de décomposition en idéaux primaires démontré par Lasker et simplifié par Macaulay lui-même). Macaulay se retient d'ailleurs de généraliser trop hâtivement :

1. *Cf.* [104] § 139, que nous avons déjà cité *supra* p. 112.

> This number has a geometrical interpretation when the theory of the resultant is applicable ; but in general it has only an algebraical interpretation. [2]

Soit donc un idéal $I \subset k[x_1, ..., x_n]$ de dimension 0. Macaulay donne une méthode effective pour calculer une base de l'espace vectoriel $I_{\leq t}$ des polynômes de degré $\leq t$ appartenant à l'idéal. Il considère d'autre part le dual

$$(k[x_1, ..., x_n]_{\leq t})^\vee$$

qu'il identifie à $k[x_1^{-1}, ..., x_n^{-1}]_{\geq -t}$, la base constituée des monômes de degré $\leq t$ ayant pour base duale la base des monômes inverses. Les sous-espaces orthogonaux $(I_{\leq t})^\perp \subset k[x_1^{-1}, ..., x_n^{-1}]_{\geq -t}$ vérifient alors le diagramme commutatif d'espaces vectoriels suivant :

$$\begin{array}{ccccc} (I_{\leq t+1})^\perp & \longrightarrow & k[x_1^{-1}, ..., x_n^{-1}]_{\geq -t-1} & & \text{monôme } M \\ {\scriptstyle p_t}\downarrow & & \downarrow & & \downarrow \\ (I_{\leq t})^\perp & \longrightarrow & k[x_1^{-1}, ..., x_n^{-1}]_{\geq -t} & \left\{ \begin{array}{l} 0 \text{ si } \deg M = -t-1 \\ M \text{ si } \deg M \geq -t \end{array} \right. & \end{array}$$

En particulier, les $(I_{\leq t})^\perp$ forment un système projectif, dont la limite projective est le sous-espace vectoriel orthogonal à I :

$$I^\perp = \varprojlim (I_{\leq t})^\perp \subset \varprojlim \left(k[x_1^{-1}, ..., x_n^{-1}]_{\geq -t}\right) = k[[x_1^{-1}, ..., x_n^{-1}]] \simeq (k[x_1, ..., x_n])^\vee$$

Macaulay montre que p_t est un isomorphisme pour t assez grand, donc $I^\perp$ est de dimension finie. Il appelle *inverse system* la base $< e_1, ..., e_N >$ de ce sous-espace fournie par sa méthode. Ses éléments fournissent un système d'équations linéaires exprimant l'appartenance d'un polynôme à l'idéal I :

$$F \in I \iff \left\{ \begin{array}{l} \langle F, e_1 \rangle = 0 \\ \vdots \\ \langle F, e_N \rangle = 0 \end{array} \right.$$

Macaulay les appelle « équations modulaires » de I. On a

$$I^\perp \simeq k[x_1, ..., x_n]/I$$

donc $N = \dim k[x_1, ..., x_n]/I$ et les résultats donnés par Macaulay peuvent s'interpréter comme une étude de l'anneau artinien $k[x_1, ..., x_n]/I$ (voir par exemple [217] p. 79 où il parle d'un « ensemble complet de restes » pour I, *a complete set of remainders*, c'est-à-dire une base de $k[x_1, ..., x_n]/I$). D'autre part, la décomposition en composantes primaires donne un isomorphisme :

$$k[x_1, ..., x_n]/I \simeq \bigoplus (k[x_1, ..., x_n]/\mathfrak{q}_i)$$

où les $\mathfrak{q}_i$ sont les idéaux primaires correspondant à chaque point de la variété. Chaque $\mathfrak{q}_i$ a à son tour un « système inverse ». Considérons l'un des $\mathfrak{q}_i$, soit $\mathfrak{q}$. La « multiplicité » μ de $\mathfrak{q}$ est définie comme étant la dimension de son système inverse. Macaulay montre par ailleurs que la décomposition primaire n'est pas essentielle : après « déplacement de l'origine » en le point support de $\mathfrak{q}$ par un changement de coordonnées affines, on a $\mathfrak{q}^\perp = I^\perp \cap k[x_1^{-1}, ..., x_n^{-1}]$, c'est-à-dire que les équations modulaires de $\mathfrak{q}$ sont les équations modulaires *finies* de I. Macaulay appelle ces dernières « équations noethériennes » de I, peut-être pour rappeler que Noether avait déjà décrit ces équations dans le cas $I = (f, g)$ et avait remarqué qu'elles forment en général un espace vectoriel de dimension rq lorsque l'origine est un point d'ordre r sur f et q sur g (*cf. supra* p. 218 et [240] p. 353). Lorsque la théorie du résultant s'applique ($I = (f_1, ..., f_n)$), Macaulay montre que cette « multiplicité » est égale à celle donnée par la formule de Poisson ([217] p. 77). Comme les équations modulaires de $\mathfrak{q}$ sont finies, on peut à présent utiliser la dualité entre $k[[x_1, ..., x_n]]$ et

$$(k[[x_1, ..., x_n]])^\perp \simeq k[x_1^{-1}, ..., x_n^{-1}].$$

2. *Cf.* [217], p. 78.

On montre que $\mathfrak{q}^\perp \subset (k[x_1,...,x_n])^\perp$ n'est autre que l'orthogonal de $\mathfrak{q}k[[x_1,...,x_n]] = Ik[[x_1,...,x_n]]$ dans $(k[[x_1,...,x_n]])^\perp$. Autrement dit, le système inverse de $\mathfrak{q}$ fournit un système d'équations linéaires exprimant l'appartenance d'une série entière P à l'idéal $Ik[[x_1,...,x_n]]$. Finalement [3]

$$\mathfrak{q}^\perp \simeq k[[x_1,...,x_n]]/Ik[[x_1,...,x_n]]$$

et la multiplicité de $\mathfrak{q}$ est

$$\mu = \dim\left(k[[x,...,x_n]]/Ik[[x_1,...,x_n]]\right)$$

Contrairement aux formules (I_1) à (I_4), il semble aisé de généraliser (I_5) à des variétés Y, Z de dimensions quelconques de l'espace affine X. Si $V \subset Y \cap Z$ est une composante irréductible de l'intersection, on lui attribuerait la multiplicité :

$$\begin{aligned}(I_5') \quad I(Y,Z;V) &= \ell\left(\mathcal{O}_{X,V}/\left(\mathcal{I}(Y)+\mathcal{I}(Z)\right)\right)\\ &= \ell\left(\left(\mathcal{O}_{X,V}/\mathcal{I}(Y)\right)\otimes_{\mathcal{O}_{X,V}}\left(\mathcal{O}_{X,V}/\mathcal{I}(Z)\right)\right)\end{aligned}$$

où $\mathcal{O}_{X,V}$ est l'anneau local de X en V. Hélas, une telle définition offre de forts inconvénients, car elle ne vérifie pas une propriété pourtant raisonnable du point de vue du géomètre. Dans le cas où l'intersection est un ensemble fini de points, on s'attend à ce que, si l'on fait varier légèrement les deux variétés au voisinage d'un de ces points, le nombre de points d'intersection dans ce voisinage, comptés avec multiplicité, reste constant. Macaulay avait lui-même donné l'exemple d'une surface de degré 4 dans $\mathbb{A}^4$ d'idéal $(x_1^2x_3 - x_2^3, x_1x_4 - x_2x_3, x_1x_3^2 - x_2^2x_4, x_2x_4^2 - x_3^3)$ dont l'intersection avec le plan $\{x_1 = x_4 = 0\}$ aurait pour multiplicité 5 à l'origine suivant cette définition, tandis que son intersection avec un plan général consiste en seulement 4 points [4]. C'est pourtant la définition que Gröbner choisit d'adopter, en toute connaissance de cause. Dans sa « géométrie algébrique », Gröbner conçoit en effet une correspondance biunivoque entre idéaux et « variétés algébriques », et il définit explicitement l'intersection de deux « variétés algébriques » comme correspondant à la somme de leurs idéaux. Gröbner était conscient qu'il devait alors renoncer au moins provisoirement à de nombreux énoncés élégants de la géométrie énumérative tels que le théorème de Bézout, et il étudie de manière détaillée l'exemple donné par Macaulay. Il exprime ainsi son point de vue :

> Anstelle des Bestrebens, den Bézoutschen Satz mit umfassender Allgemeinheit aussprechen zu können, setzen wir das Bestreben, die tieferliegenden Gründe aufzudecken, von denen es abhängt, daß er in gewissen Fällen nicht gilt. [5]

Pourtant, une définition géométrique de la multiplicité d'intersection restait souhaitable. Il fallait pour cela établir l'usage du « principe de continuité » de Poncelet et Schubert, et aussi expliciter ce qu'on était en droit d'attendre d'une telle définition de la multiplicité d'intersection. Le concept de « spécialisation » a joué un rôle dans ces recherches sur le « principe de continuité » ; l'histoire en a déjà été écrite par Van der Waerden [317] qui y a lui-même participé. Dans le cas de deux courbes Y et Z de degrés e et d, on pourrait définir la multiplicité d'intersection ainsi :

$$(I_6) \quad I(Y,Z;P) = \begin{array}{l}\text{nombre de points généraux de l'intersection de deux}\\ \text{courbes générales de degrés } e \text{ et } d \text{ qui se spécialisent en le}\\ \text{point } P \text{ quand on spécialise les deux courbes en } Y \text{ et } Z.\end{array}$$

Plus généralement, dans ses *Foundations of algebraic geometry*, Weil définit la « multiplicité d'une spécialisation » au moyen du théorème suivant, valable pour un corps k de caractéristique quelconque [6] :

3. L'égalité $Ik[[x_1,...,x_n]] = \mathfrak{q}k[[x_1,...,x_n]]$ est une conséquence immédiate du « théorème de Lasker-Noether » [217] p. 61. A vrai dire, Macaulay montre seulement, *cf.* [217] p. 74–75, que les équations noethériennes expriment l'appartenance d'un *polynôme* P à l'idéal $Ik[[x_1,...,x_n]]$. Mais $\mathfrak{q} = I + O^\gamma$ où $O = (x_1,...,x_n)$ et γ est un certain entier, appelé par Macaulay *characteristic number* de l'idéal primaire $\mathfrak{q}$. Pour une série entière P, il suffit donc de vérifier que sa partie de degré $< \gamma$ appartient à $\mathfrak{q}k[[x_1,...,x_n]]$, et la démonstration de Macaulay est donc aussi valable dans ce cas.

4. *Cf.* [217] p. 98.

5. *Cf.* [135] p. 179.

6. *Cf.* [327] ch. III § 3 proposition 7, p. 60–61. En caractéristique 0, $[k(\alpha) : k]_i = 1$.

> Let k be a field, (x) a set of independant variables over k, (y) a set of algebraic quantities over the field $k(x)$, and $(y^{(1)}, ..., y^{(n)})$ a complet set of conjugates of (y) over $k(x)$. Let (a) be a set of elements of k, and (α) a set of algebraic quantities over k, such that (α) is an isolated specialization of (y) over $(x) \to (a)$, with reference to k. Then there is an integer μ such that, in every specialization $(\overline{y}^{(1)}, ..., \overline{y}^{(n)})$ of $(y^{(1)}, ..., y^{(n)})$ over $(x) \to (a)$ with reference to k, each conjugate of (α) over k appears exactly μ times among the $(\overline{y}^{(\nu)})$; and μ is a multiple of $[k(\alpha) : k]_i$.

C'est cet entier μ qu'il appelle *multiplicity of a specialization*. Sous certaines hypothèses, cet entier coïncide avec la « multiplicité d'un idéal primaire » de Macaulay [7].

On peut déceler dans cet ouvrage de Weil une certaine affinité avec le style adopté par Kronecker dans ses *Grundzüge* (*cf. infra* chapitre 15), par exemple dans l'usage fondamental qui y est fait de la notion de *algebraic quantity* (*algebraische Grösse*) dans un contexte géométrique. Mais Weil réduit au minimum tout recours à la théorie de l'élimination aussi bien qu'à la théorie des idéaux. Il semble même faire preuve d'ironie à l'égard d'une certaine algèbre lorsqu'il écrit, dans une note de bas de page, qu'un procédé démonstratif emprunté à Chevalley « élimine enfin de la géométrie algébrique les dernières traces de la théorie de l'élimination ». Van der Waerden avait en effet démontré le même énoncé en 1927 au moyen de la méthode d'élimination de Kronecker (*cf. infra* p. 269). Dans les années 1970, les articles historiques de van der Waerden [317] et Abhyankar [4] ont rendu célèbre cette citation de Weil. Van der Waerden ajoutait que « visiblement Weil et Chevalley n'aiment pas la théorie de l'élimination », et Abhyankar allait plus loin en écrivant : « Ce qui est surprenant est que, sous l'influence de Bourbaki, il est devenu à la mode de jeter le discrédit sur la théorie de l'élimination. (...) Il me semble que ce que Bourbaki a atteint en cela, est d'échanger des démonstrations constructives contre de simples démonstrations d'existence ». Enfin, lors de l'*International Congress of Mathematicians* à Nice en 1970, à la suite de l'exposé de van der Waerden (qui est à l'origine de son article [317]), Abhyankar rédigea un poème *Polynomials and power series* dans lequel il répondit à Weil :« eliminate the eliminators of elimination theory » [8]. Mais le point de vue de Weil était controversé dès la parution de son livre. Son but était, en réaction aux défauts des méthodes algébriques de son temps, de fournir un fondement solide pour la géométrie algébrique sur lequel s'appuieraient les recherches présentes et futures. [9] Dans un compte-rendu du

7. *Cf.* [327], chap. III § 2 lemme 6.

8. Le poème d'Abhyankar aurait été publié dans *The Mathematical Intelligencer* en septembre 1972, mais nous n'avons pu consulter ce numéro de la revue. Ce poème se trouve cependant à la Bibliothèque Hadamard de l'Université Paris XI (Orsay), dans un fascicule dactylographié, joint à une version de l'article [4] légèrement différente de la version parue dans l'American Mathematical Monthly. Voici les premiers vers du poème :

> Polynomials and power series
> May they forever rule the world.
> Eliminate, eliminate, eliminate
> Eliminate the eliminators of elimination theory.
> As you must resist the superbourbaki coup
> So must you fight the little bourbakis too.
> Kronecker, Kronecker, Kronecker above all
> Kronecker, Mertens, Macauley (*sic*), and Sylvester.

Dans son article [4], Abhyankar oppose les preuves constructives (explicites, algorithmiques) aux preuves d'existence, en géométrie algébrique. Il reproche aux traités écrits après 1940 d'être illisibles, et juge les travaux des années 1960 (en algèbre homologique et en théorie des schémas) « hors de proportion ». Il insiste sur le fait que l'influence de Bourbaki a discrédité la théorie de l'élimination, tandis que cette théorie était auparavant considérée comme un outil inévitable et qu'elle permettait des preuves constructives. Son article commence ainsi :

> La méthode de l'algèbre des lycées est puissante, belle et accessible. Alors ne nous laissons pas submerger par les groupes-anneaux-corps ou les flèches fonctorielles (...) et ne perdons pas de vue la puissance des procédés algorithmiques explicites que nous ont donnés Newton, Tschirnhaus, Kronecker et Sylvester.

C'est donc une sorte de retour aux anciens que nous propose Abhyankar. Il entreprend de « traduire » les théories modernes dans le langage classique et expose en effet dans cet article de nombreux résultats avancés en théorie des courbes algébriques planes, dans le cadre de l'algèbre élémentaire des polynômes ou des séries formelles.

9. Dans l'introduction à la première édition de son livre en 1946, Weil écrit (la citation grecque est une phrase célèbre attribuée à Archimède (*cf.* [83] t. 2 p. 836–837) : « donne-moi où je puisse me tenir ferme [et j'ébranlerai la Terre] ») :

> Δὸς ποῦ στῶ
> Algebraic geometry, in spite of its beauty and importance, has long been held in disrepute by many

livre de Weil, Zariski a écrit :

> C'est un trait caractéristique de ce livre – à une exception près (le chapitre III) – qu'aucun usage n'y est fait des méthodes supérieures de l'algèbre moderne [the higher methods of modern algebra]. L'auteur a décidé de ne pas supposer ni utiliser l'algèbre moderne « au delà des faits les plus élémentaires sur les corps abstraits et leurs extensions et des premiers rudiments de la théorie des idéaux ». (...) L'auteur justifie cette procédure par un argument de continuité historique, prônant un retour aux « palais qui sont héréditairement nôtres ». Mais il est très improbable que nos prédécesseurs reconnaîtront dans le livre de Weil leur propre édifice familier bien qu'amélioré et complété. Si le géomètre traditionnel était invité à choisir entre « des installations de fortune pleines d'anneaux, d'idéaux et de valuations » d'un côté, et des installations pleines de corps, de corps linéairement disjoints, d'extensions régulières, d'extensions indépendantes, de spécialisations génériques, de spécialisations finies et de spécialisations de spécialisations de l'autre côté, il déclinerait très probablement l'invitation au choix, et il dirait : « Maudites soient vos maisons ! ». Alors autant se servir d'algèbre moderne le plus possible. [10]

Pourtant Zariski conclut son compte-rendu en écrivant que « le livre de Weil est le premier exposé purement arithmétique d'un secteur important de la géométrie algébrique » – la théorie des intersections – « et est donc une référence dans la littérature de ce domaine ». Et malgré le renouveau de la géométrie algébrique avec la théorie des schémas de Grothendieck dans les années 1960-1970, le livre de Weil est en effet resté pendant longtemps une référence pour la théorie des intersections – il était encore cité comme tel par Hartshorne en 1977 [11].

Revenons à notre sujet : la multiplicité d'intersection. En offrant la rigueur au principe de continuité, la « multiplicité d'intersection » de Weil devait donc perdre l'élégance et la concision des formules algébriques (I_1) à (I_5) ci-dessus. Weil donne en appendice une définition axiomatique de la multiplicité d'intersection. C'est peut-être surtout en cela qu'il se montre « bourbakiste ». La multiplicité d'intersection est alors décrite par les propriétés qu'elles doit vérifier, et non plus par une formule algébrique. Il démontre aussi l'unicité du concept dans cet appendice ; il le construit dans le corps de l'ouvrage. Il définit d'abord la « multiplicité d'intersection » de deux variétés dont l'une est une variété linéaire, en spécialisant une variété linéaire générale. Dans le cas de deux variétés quelconques, il se ramène au cas où l'une est linéaire par « réduction à la diagonale » : la multiplicité du point P d'intersection de Y, $Z \subset X$ est celle du point (P, P) d'intersection de $Y \times Z$ et de la diagonale $\Delta \subset X \times X$.

Pourtant Samuel a donné quelques années plus tard une expression de la multiplicité d'intersection, équivalente à la définition de Weil , par une formule algébrique. Soit à nouveau P un point d'intersection isolé de Y et Z, variétés quelconques dans $X = \mathbb{A}^n$, et $\Delta \subset X \times X$. Soit la fonction :

$$F(t) = \ell\left(\mathcal{O}_{Y\times Z,(P,P)}/\mathcal{I}(\Delta)^t\right)$$

où $\mathcal{O}_{Y\times Z,(P,P)} = \mathcal{O}_{X\times X,(P,P)}/\left(\mathcal{I}(Y)\otimes 1 + 1\otimes\mathcal{I}(Z)\right)$ est l'anneau local de $Y \times Z$ en (P, P). Alors

---

mathematicians as lacking proper foundations. (...) There is no doubt that, in this field, the work of consolidation has so long been overdue that the delay is now seriously hampering progress in this and other branches of mathematics. To take one instance, a personal one, this book has arisen from the necessity of giving a firm bases to Serveri's theory of correspondences on algebraic curves, especially in the case of characteristic $p \neq 0$ (...), this being required for the solution of a long outstanding problem, the proof of the Riemann hypothesis in function-fields.

10. *Cf.* [333]. Plus tard, Pierre Samuel pouvait écrire ([275], p. 124–125) :

    Le théorème des zéros est dû à Hilbert (...). Le théorème des zéros a récemment été l'objet de nombreuses recherches, tendant à le démontrer sans utiliser la théorie de l'élimination (...). Les notions de spécialisation et de point générique apparaissent dans les premiers travaux de van der Waerden et celui-ci démontre le théorème d'extension des spécialisations par la théorie de l'élimination. Les travaux (souvent collectifs) de l'Ecole Américaine [=Chevalley, Weil, Zariski] ont porté sur l'élimination de cette théorie (...). Ce n'est pas ici le lieu de dénouer l'écheveau complexe des influences que Chevalley, Weil et Zariski eurent les uns sur les autres.

11. *Cf.* [140] appendice A « Intersection Theory ».

F est un polynôme pour t assez grand, notons-le $S(t)$. On a [12]

$$(I_7) \quad I(Y,Z;P) = \text{coefficient de } \frac{t^n}{n!} \text{ dans } S(t)$$

Pour montrer l'avantage de cette définition sur (I_5) et (I'_5), considérons l'exemple déjà mentionné :

$$\begin{aligned} &X = \mathbb{A}^4 \\ &\mathcal{I}(Y) = \mathfrak{p} = (x_1^2x_3 - x_2^3, x_1x_4 - x_2x_3, x_1x_3^2 - x_2^2x_4, x_2x_4^2 - x_3^3) \\ &\mathcal{I}(Z) = \mathfrak{l} = (x_1, x_4) \end{aligned}$$

On a :

$$k[[x_1, ..., x_4]]/(\mathfrak{p} + \mathfrak{l}) \simeq k[[x_2, x_3]]/(x_2^3, x_2x_3, x_3^3)$$

Donc (I_5) donne une multiplicité égale à 5, que l'on ait ou non recours à la réduction à la diagonale.

$$F(t) = \ell\left(\mathcal{O}_{X\times X,(P,P)}/\left(\mathfrak{p} + \mathfrak{l} + (x_1 - x'_1, ..., x_4 - x'_4)^t\right)\right)$$

et la multiplicité d'intersection donnée par (I_7) est 4, comme le principe de continuité le laissait prévoir (Y est une surface de degré 4, paramétrée par $(\lambda^4, \lambda^3\mu, \lambda\mu^3, \mu^4)$, qu'un plan générique coupe en 4 points). Ici, la réduction à la diagonale est essentielle. Si l'on posait en effet :

$$F(t) = \ell\left(\mathcal{O}_{X,P}/(\mathfrak{p} + \mathfrak{l})^t\right) = \ell(k[[x_1, ..., x_4]]/(x_2^3, x_2x_3, x_3^3, x_1, x_4)^t),$$

le polynôme de Samuel serait :

$$P(t) = 6\frac{t^4}{4!} + \frac{4}{3}t^3 + \frac{9}{4}t^2 + \frac{7}{6}t$$

d'où 6 pour multiplicité.

Finalement, J.-P. Serre, appliquant les méthodes cohomologiques à la théorie de Samuel, a montré que le coefficient de $\frac{t^n}{n!}$ dans un polynôme de Samuel peut s'exprimer sous forme d'une caractéristique d'Euler-Poincaré (somme alternée des dimensions de certains espaces de cohomologie) [13]. Utilisant la théorie de la profondeur ou « dimension homologique » qu'il avait développée dans [284], il parvient alors à la formule suivante :

$$(I_8) \quad I(Y,Z;P) = \sum_{i=0}^{\dim X} (-1)^i \ell\left(\mathrm{Tor}_i^{\mathcal{O}_{X,P}}\left(\mathcal{O}_{X,P}/\mathcal{I}(Y), \mathcal{O}_{X,P}/\mathcal{I}(Z)\right)\right)$$

Le premier terme de cette somme n'est autre que la multiplicité d'intersection définie selon (I'_5). Serre montre que cette définition (I_8) vérifie les axiomes d'une multiplicité d'intersection [14], et est donc équivalente à (I_6) et (I_7) dans les cas où celles-ci sont valides, bien que la réduction à la diagonale ne soit pas requise dans sa formulation. D'autre part, (I_8) est plus générale que (I_5) et (I_6), qui ne s'appliquent qu'au cas où la variété ambiante X est non-singulière.

Cette pluralité même de définitions possibles de la multiplicité d'intersection, et surtout les démonstrations d'équivalence entre ces définitions, tendent à convaincre de l'existence d'un *objet* « point d'intersection multiple ». Ce concept dépasse certes l'intuition géométrique primitive du « point » aussi bien que l'idée élémentaire de multiplicité d'une racine d'une équation à une inconnue en algèbre. Mais sa vérité s'impose encore à notre esprit quand on pense aux avantages qui résultent de son admission. Ce concept permet de transformer des énoncés *généraux* de théorèmes en des énoncés *universels*.

12. *Cf.* [274]. Pour l'équivalence, sous certaines hypothèses, de (I_7) et (I'_5), voir par exemple [285] p. 65. Aujourd'hui, on appelle $S(t)$ « polynôme de Samuel ». Sa différence première $(S(t+1) - S(t))$ est polynôme de Hilbert d'un certain anneau gradué.

13. *Cf.* [285], théorème 1 p. 57.

14. *Cf.* [285], p. 114–116.

On a vu dès l'apparition du théorème de Bézout (*supra* p. 50) Newton l'utiliser comme un théorème général valable pour deux courbes suffisamment quelconques mais faux dans certains cas particuliers, à cause des points à l'infini et des points d'intersection multiples. Peu après, Maclaurin franchissait un pas de plus en comptant les points d'intersection avec leur multiplicité. Il fallait certes encore pour l'ériger en théorème universel, tenir compte aussi des points imaginaires (Euler en était conscient), développer la géométrie projective (première moitié du XIX$^{\text{ème}}$) pour donner un sens plus précis à la notion de point à l'infini en géométrie algébrique, et enfin *définir* la multiplicité d'intersection. L'ironie du sort veut que Maclaurin et Euler fussent peut-être plus proches de concevoir ce théorème de Bézout comme théorème universel que ne l'était Bézout lui-même : on a vu qu'il s'attacha, sans avoir une compréhension suffisante de l'espace projectif, à montrer que l'abaissement du degré de l'équation finale par rapport au produit des degrés des équations initiales n'était pas toujours dû à la présence d'un point d'intersection à l'infini (*supra* p. 170) !

Plus généralement, tous les énoncés de la géométrie énumérative requièrent de compter les objets géométriques *avec multiplicité* : alors seulement ont-ils un sens universel. Nous verrons plus loin l'exemple de la formule de Chasles déterminant le nombre de sections coniques vérifiant des conditions données. Parmi les énoncés qui de généraux deviennent universels *via* la multiplicité d'intersection, on pourrait encore citer le théorème de Bret, la théorie des courbes adjointes pour définir le genre à la manière de Clebsch et Gordan [77], le théorème $A\phi + B\psi$ de Noether (*cf. supra* p. 218),...

Par ailleurs, le caractère universel de ces énoncés est la condition de possibilité de leur application à certains problèmes particulièrement intéressants. Autrement dit, les problèmes les plus intéressants sont parfois justement la proie des exceptions au théorème général. Seul l'énoncé universel permettra alors de régler la difficulté. Illustrons notre propos par un exemple. Nous avons déjà dit qu'il fallait probablement chercher l'origine du théorème de Bézout dans la critique de la classification cartésienne des courbes par « genre » (*supra* section 2.3). Mais il est remarquable que le premier auteur à l'énoncer clairement, Newton, ait compté certainement parmi les auteurs les plus conscients de deux *exceptions* à ce théorème, exceptions particulièrement fécondes : la méthode des tangentes (en analyse infinitésimale) et la méthode de la corde (méthode diophantienne).

Mentionnons encore une autre motivation importante pour admettre les notions de « point d'intersection multiple » et de « multiplicité d'intersection ». Il s'agit de l'idée d'une « théorie des cycles ». Nous avons déjà mentionné que Gröbner s'attachait à définir algébriquement une opération « intersection » entre deux variétés algébriques, correspondant à la somme des idéaux des variétés. En fait, dès les années 1930, Severi et van der Waerden avaient défini la notion de « cycle » [15], somme formelle de variétés de même dimension affectées de coefficients numériques. Autrement dit les cycles forment un groupe additif, et il est assez naturel de définir un « produit d'intersection » commutatif et distributif par rapport à l'addition, tel que, dans un cycle intersection de deux variétés et de dimension zéro, le coefficient de chaque point soit la multiplicité d'intersection des deux variétés en ce point. L'idée d'un anneau de cycles, muni d'une opération « produit d'intersection », apparaît aussi explicitement dans le mémoire de Halphen que nous commentons dans la prochaine section, et comme telle, elle est à l'origine de la géométrie énumérative d'H. Schubert. Nous verrons aussi Kronecker s'efforcer de comprendre la nature arithmétique de l'opération « intersection » et forger à cet effet le concept de « système de diviseurs » (*infra* p. 251). Nous pourrions ajouter encore d'autres occurences de cette idée au XIX$^{\text{ème}}$ siècle [16]. Dans ces exemples, autant sinon plus que le désir de rendre un énoncé universel ou celui de rendre une démonstration rigoureuse par le truchement d'un concept de multiplicité mieux défini, compte celui d'établir une structure opératoire. Malgré le fait que la définition (I_8) de Serre soit applicable à des situations singulières, en 1984, il n'existait pas encore de théorie des cycles sur une variété non-singulière [17].

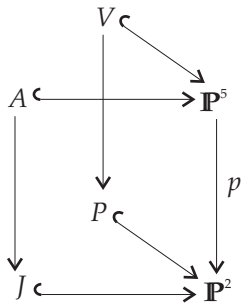
15. *Cf.* [317] p. 177 où van der Waerden lui-même décrit l'histoire de cette notion dans les travaux de Severi et dans ses propres travaux dont certains en commun avec Chow (d'où la dénomination ultérieure d'« anneau de Chow » par exemple dans les rédactions du Séminaire C. Chevalley *Anneaux de Chow et Applications* [73] p. 3-14 à 3-18).

16. Ainsi l'opération $[A, B]$ définie par O. Bonnet dans sa démonstration du théorème de Bret, ou encore celle semblable utilisée par Brill en 1872 (*cf. infra* p. 267).

17. *cf.* [117] p. 41. Fulton cite à cet égard un contre-exemple intéressant que nous mentionnerons à nouveau *infra* note 32 p. 269.

14.2 Le mémoire de Halphen sur la détermination des coniques (1873)

En mêlant les notations de Halphen à des notations modernes, on résume la construction employée dans son mémoire par le diagramme commutatif suivant :



$\mathbb{P}^5$ est ici l'espace des coefficients de l'équation d'une conique plane, V (qui fut appelée plus tard « Variété de Véronèse ») est la sous-variété de $\mathbb{P}^5$ correspondant aux coniques dégénérées en un produit de deux droites confondues (Halphen les appelle les « droites-coniques »), et A est un « système » quelconque de coniques, c'est-à-dire une sous-variété de $\mathbb{P}^5$ de dimension 1. Halphen construit une projection $p : \mathbb{P}^5 \to \mathbb{P}^2$ en associant à chaque conique (chaque point de $\mathbb{P}^5$) son intersection avec une transversale Δ donnée : « un couple quelconque de points m et m' de la droite Δ peut être représenté sur le plan par un point dont les coordonnées sont : $x = \frac{1}{2}(am + am')$, $y = \frac{am.am'}{K}$ » [18]. L'image de V est une conique $P \subset \mathbb{P}^2$ (une parabole, dans l'ouvert affine considéré par Halphen), et l'image de A est une certaine variété J de degré μ. Alors $P \cap J$ comprend l'image des droites-coniques de A et l'image des coniques de A tangentes à Δ :

$$\text{Card } P \cap J = \omega + \nu$$

où ν est le nombre des coniques tangentes à Δ dans A (chacune étant comptée avec multiplicité 1), et ω le nombre des droites-coniques de A comptées avec une multiplicité à déterminer, multiplicité que Halphen appelle la « valeur d'une droite-conique ». Question de style, Halphen ne laisse transparaître aucune notion algébrique dans la solution d'un problème de géométrie, si ce n'est dans la première section de son mémoire qui contient justement des « théorèmes généraux sur les intersections des courbes planes algébriques » et dans les sections où il décrit un anneau de cycles, comme nous le verrons. Dans le théorème II de cette première section, il traduit ainsi notre définition (I_1) de la multiplicité d'intersection dans un langage qui lui permet par la suite d'éviter tout recours à l'algèbre :

> Le nombre des intersections de deux courbes, réunies en un point O, est égal à la somme des ordres des segments infiniment petits interceptés par les deux courbes sur une sécante dont la distance au point O est infiniment petite du premier ordre, et qui ne coïncide avec aucun tangente à l'une des courbes en ce point [19].

Halphen montre en particulier que pour une droite Δ suffisamment générale et une droite-conique de valeur 1, « le déplacement angulaire du grand axe, dans les coniques voisines [appartenant au système A], est d'ordre égal ou supérieur au second, relativement à l'angle des asymptotes » [20] ; sinon la droite-conique est de valeur ≥ 2.

Afin de déterminer le nombre des coniques du système A vérifiant une condition donnée, c'est-à-dire appartenant à une certaine hypersurface $A' \subset \mathbb{P}^5$, il considère pour chaque conique $S \in A$ les coniques « en involution » avec S et une conique fixe Σ et vérifiant la condition donnée, c'est-à-dire

18. *Cf.* [137] p. 104.
19. *Cf.* [137] p. 102.
20. *Cf.* [137] p. 108.

les coniques passant par les quatre points d'intersection de S et Σ et vérifiant la condition donnée. Il en existe en général un nombre a, si a est le degré de l'hypersurface définissant la condition donnée ; et quand S vérifie elle-même la condition donnée, elle est confondue avec l'une de ces a coniques. Après projection par p, cela se traduit dans $\mathbb{P}^2$ par une correspondance entre les points de J et ceux de $p(A')$, telle qu'à chaque point $x \in J$ corresponde a points $y \in p(A')$ alignés avec $p(\Sigma)$ et x. Halphen montre que le « nombre de couples de points correspondants confondus » est alors $a\mu$ (comptés avec multiplicité) [21]. Mais parmi ces $a\mu$ points, il faut rejeter les points $O \in p(A') \cap J \cap P$ correspondant à des droites-coniques, comptés avec leurs multiplicités [22] : Halphen montre que pour tout point $O \in J \cap P$ correspondant à une droite-conique de valeur donnée, il faut compter ce point, dans $p(A') \cap J \cap P$, « b fois la valeur de la droite-conique répondant au point O », où b est défini par le fait que, parmi les a coniques correspondant à une « conique infiniment aplatie » du système A, il y en a $(a-b)$ ordinaires et b infiniment aplaties. Alors le nombre de coniques de A vérifiant la condition donnée est

$$N = a\mu - b\omega = a\mu - b(2\mu - \nu) = (a-2b)\mu + b\nu$$

Le fait que N soit une forme linéaire en les « caractéristiques » μ et ν du système A avait été énoncé par Chasles.

Si l'on remplace à présent A et A' par des sous-variétés de $\mathbb{P}^5$ de dimensions 2 et 3, on obtient un théorème dû à Cremona, que Halphen démontre par une méthode semblable :

> Le nombre de coniques satisfaisant à une condition triple et à une condition double indépendantes est de la forme $\gamma\rho + \delta\sigma + \epsilon\tau$, les nombres ρ, σ, τ ne dépendant que de la condition triple (ce sont les caractéristiques du complexe défini par cette condition), et les nombres γ, δ, ϵ ne dépendant que de la condition double. [23]

Plus généralement, Halphen considère des complexes de coniques, sous-variétés de $\mathbb{P}^5$ de codimension $1 \leq r \leq 5$. A tout complexe de coniques il associe un « module », expression symbolique de la forme

$$\sum_{i=0}^{r} \alpha_i p^i d^{r-i}$$

où p et d sont des lettres « initiales des mots *point*, *droite* » [24] et pour tout i, $\alpha_i \in \mathbb{Q}$. Par exemple, une condition simple définit un complexe de coniques de codimension 1, et on lui associe son module $\alpha p + \beta d$, tel que le nombre de coniques qui vérifient cette condition dans un complexe de coniques de codimension 4 et de caractéristiques μ et ν soit $\alpha\mu + \beta\nu$. De plus, pour chaque r, les symboles $p^r, p^{r-1}d, ..., d^r$ sont liés par certaines relations linéaires que Halphen décrit explicitement. Halphen a ainsi défini un anneau, gradué par le degré total en p, d. Il vérifie d'autre part que le module de l'intersection de deux complexes de codimension i et j, pour $i + j < 5$, est égal au produit des modules. Lorsque $i + j = 5$, si l'on remplace, dans le module produit, $p^i d^{5-i}$ par le nombre de coniques qui passent en i points donnés et touchent $(5-i)$ droites données (pour i de 0 à 5, ce nombre est respectivement égal à 1, 2, 4, 4, 2, 1), on obtient le nombre de coniques dans l'intersection des deux complexes. La composante de degré 5 de l'anneau de Halphen est ainsi identifiée a $\mathbb{Q}$. Notons A^i la composante de degré i. Le produit d'intersection $A^i \times A^{5-i} \to A^5 \simeq \mathbb{Q}$ définit une dualité entre A^i et A^{5-i}. Halphen montre par exemple [25] que

$$< \frac{2d^4 - p^4}{3}, \frac{2p^4 - d^4}{3} >$$

est une base duale de $< p, d >$.

La démonstration par Halphen du théorème de Chasles n'est ni la première ni la dernière ; l'histoire de la justification de cette théorie a déjà été écrite, par S. L. Kleiman [26]. Halphen a le mérite d'avoir senti l'importance d'un tel calcul symbolique, et cette construction est à l'origine de la

21. Il utilise à cet effet le théorème III de la première section de son mémoire.
22. Pour déterminer leurs multiplicités, il utilise le théorème IV de la première section de son mémoire.
23. *Cf.* [137] p. 121. Halphen utilise alors la locution « complexe de coniques » pour désigner les sous-variétés de $\mathbb{P}^5$, et non plus « système de coniques ».
24. *Cf.* [137] p. 124.
25. *Cf.* [137] p. 125.
26. *Cf.* [172], surtout p. 123–127.

géométrie énumérative de H. Schubert [27]. Elle peut aussi, comme le suggère S. L. Kleiman [28], être justifiée par les arguments modernes de la théorie des intersections, en remarquant que $A^i = A^i(B) \otimes \mathbb{Q}$, où B est l'éclaté de $\mathbb{P}^5$ le long de V, et $A^\bullet(B)$ est l'« anneau de Chow » ou anneau d'intersection de B, dont les éléments sont appelés « cycles ». On considère le diagramme cartésien suivant, où $\dim V = 2$ et $\dim V' = 4$:

$$\begin{array}{ccc} V' & \longrightarrow & B \\ \downarrow & & \downarrow \\ V & \longrightarrow & \mathbb{P}^5 \end{array}$$

Le théorème de structure de l'anneau de Chow d'un éclaté fournit dans notre cas les isomorphismes suivant [29] :

$$\begin{array}{l} A^1(B) \simeq A^1(\mathbb{P}^5) \oplus A^0(V) =< p, l > \\ A^2(B) \simeq A^2(\mathbb{P}^5) \oplus A^0(V) \oplus A^1(V) \\ A^3(B) \simeq A^3(\mathbb{P}^5) \oplus A^1(V) \oplus A^2(V) \\ A^4(B) \simeq A^4(\mathbb{P}^5) \oplus A^2(V) \end{array}$$

où le cycle l est un générateur de $A^0(V)$ (et on a $l = 2p - d$). La deuxième ligne s'obtient au moyen des isomorphismes suivants (où l'on adopte les notations de S.G.A. 5) :

$$\begin{array}{l} A^2(B) \simeq A^2(\mathbb{P}^5) \oplus A^1(V') \\ A^1(V') \simeq g^* A^1(V) \oplus g^* A^0(V)\xi \end{array}$$

Or :

$$\begin{array}{l} A^2(\mathbb{P}^5) =< p^2 > \\ g^* A^1(V) \otimes \mathbb{Q} =< \frac{1}{2} pl > \text{ (à vérifier)} \\ g^* A^0(V)\xi \otimes \mathbb{Q} =< \frac{1}{2} kl > \text{ (à vérifier)} \end{array}$$

où k est le cycle des coniques dégénérées en un produit de droites distinctes ou en une droite-conique à foyers confondus (l'éclatement de $\mathbb{P}^5$ le long de V revient à considérer qu'une même droite du plan est support de plusieurs droites-coniques distinctes selon la position des foyers des coniques qui lui sont infiniment voisines). On peut de plus montrer que $pl = 2p^2 - pd$ et $lk = 5pd - 2p^2 - 2d^2$, d'où finalement

$$A^2(B) \otimes \mathbb{Q} =< p^2, pd, d^2 >,$$

justifiant ainsi le théorème de Cremona.

27. *ibid.* p. 121.
28. *ibid.* p. 130.
29. *Cf.* S.G.A. 5, exposé VII (J.-P. Jouanolou), (9.10.2) p. 350.

Chapitre 15

Théorie arithmétique de la grandeur algébrique : Kronecker (1882)

La préface annonce en quelques phrases le projet de Kronecker, qui, selon Dieudonné ([93], p. 27), « was the first to dream of one vast algebraico-geometric construction comprising the theory of numbers and algebraic geometry ». La première partie des *Grundzüge* est une théorie des extensions de corps et des entiers algébriques. Comme l'indique la préface [1] et comme l'a démontré H. M. Edwards, une bonne définition du concept même d'entier est à cet égard cruciale. Dans la deuxième partie, il s'agit alors de faire une théorie de la divisibilité dans les anneaux d'entiers. Mais, comme le remarque J. Boniface, à la différence de Dedekind, Kronecker, lors du passage des entiers naturels aux entiers algébriques, ne s'intéresse pas tant à préserver les lois (théorème d'existence et d'unicité de la décomposition en facteurs premiers, qui joue un rôle majeur chez Dedekind, mais est secondaire, comme l'a montré Edwards, chez Kronecker) qu'à préserver les objets, ici le plus grand commun diviseur de plusieurs entiers, que Kronecker définit pour un anneau d'entiers quelconque. Enfin, la théorie des corps construite par Kronecker concerne aussi bien les extensions transcendantes, qu'il manipule géométriquement au moyen d'un concept de « variété » (*Mannigfaltigkeit*), et il laisse entendre à la fin de sa préface que le concept de « Stufe » (ce qui, traduit dans un langage algébrico-géométrique plus moderne, désigne la codimension d'une sous-variété de l'espace affine) aura un rôle important dans l'édifice [2]. Nous verrons que ce concept est au cœur de la théorie de l'élimination élaborée par Kronecker.

15.1 Théorie des corps

Kronecker choisit délibéremment (il s'en justifie dans le § 1) d'éviter le mot « corps » utilisé par Dedekind, et il préfère parler de « domaine de rationalité » (*Rationalitäts-Bereich*) pour désigner un corps $\mathbb{Q}(\mathfrak{R}', \mathfrak{R}'', \mathfrak{R}''', ...)$ qu'il note lui-même de la manière suivante :

$$(\mathfrak{R}', \mathfrak{R}'', \mathfrak{R}''', ...)$$

où $\mathfrak{R}'$, $\mathfrak{R}''$, $\mathfrak{R}'''$,... sont des « grandeurs » données. Kronecker insiste d'emblée sur le fait que le concept de « grandeur » recouvre ici les notions usuelles de nombres rationnels et nombres irrationnels algébriques (tel $\sqrt{2}$), aussi bien que celle de « grandeur indéterminée » ou de « fonction rationnelle

1. *Cf.* p. 2 où Kronecker parle du problème de « la représentation de toutes les grandeurs entières algébriques d'un genre... par des fonctions linéaires de plusieurs grandeurs algébriques... afin de rendre réelles les grandeurs fractionnaires *idéales* »

2. Selon Kronecker, à l'issue du dernier paragraphe des *Grundzüge*, il apparaîtra que « le concept de *Stufe* convient pour remplacer le concept d'irrationalité algébrique ». Nous y reviendrons *infra* section 15.6.

de plusieurs grandeurs indéterminées ». Pour motiver sa démarche, il indique dès le § 1 que la notion d'irréductibilité d'un polynôme (à coefficients dans un domaine de rationalité) n'a de sens que relativement au domaine de rationalité choisi, et il remarque que Galois et Abel étaient déjà conscients du besoin de préciser le sens du mot « rationnel ». Il renvoie au manuscrit d'Abel *Sur la résolution algébrique des équations* publié après sa mort en 1839. Abel y définit en effet les différentes modalités du concept de « grandeur » presque comme Kronecker, et il écrit :

> D'après la nature des quantités connues nous ferons les distinctions suivantes :
>
> 1. Une quantité qui peut s'exprimer algébriquement par l'unité s'appelle un nombre algébrique ; si elle peut s'exprimer rationnellement par l'unité, elle s'appelle un nombre rationnel, et si elle peut être formée de l'unité par addition, soustraction et multiplication, elle s'appelle un nombre entier.
> 2. Si les quantités connues contiennent une ou plusieurs quantités variables, la quantité y est dite fonction algébrique, rationnelle ou entière de ces quantités selon la nature des opérations nécessaires pour la former. Dans ce cas on regarde comme quantité connue toute quantité constante. [3]

Le lexique introduit par Kronecker au § 2 de son mémoire permet de décrire le contexte relatif, que l'on exprime aujourd'hui en précisant de quel corps de base un corps donné est l'extension. Kronecker nomme « domaine de base » (*Stammbereich*) un corps de base $\mathbb{Q}(\mathfrak{R}', \mathfrak{R}'', \mathfrak{R}''', ...)$, et « domaine de genre $\mathfrak{G}'$ » (*der Bereich der Gattung* $\mathfrak{G}'$) une extension engendrée par la grandeur $\mathfrak{G}'$, c'est-à-dire de la forme $\mathbb{Q}(\mathfrak{G}', \mathfrak{R}', \mathfrak{R}'', \mathfrak{R}''', ...)$, ce qu'il note lui-même :

$$(\mathfrak{G}', \mathfrak{R}', \mathfrak{R}'', \mathfrak{R}''', ...)$$

Un genre $\mathfrak{G}'$ est dit « d'ordre n » quand $\mathfrak{G}'$ est de degré n sur le corps de base ; il est dit « genre galoisien » quand l'extension est galoisienne (c'est-à-dire que tous les conjugués de $\mathfrak{G}'$ appartiennent à $\mathbb{Q}(\mathfrak{G}', \mathfrak{R}', \mathfrak{R}'', \mathfrak{R}''', ...)$). Dans la terminologie de Kronecker, l'expression « genre $\mathfrak{G}'$ » ne s'applique en fait qu'à l'ensemble des éléments primitifs (c'est-à-dire les éléments de même degré que $\mathfrak{G}'$ sur le corps de base) de l'extension considérée, tandis que l'expression « domaine de genre $\mathfrak{G}'$ » désigne toute l'extension. Kronecker démontre que le degré d'une sous-extension divise le degré de l'extension.

Au § 3, Kronecker remarque que l'on peut construire toute extension de $\mathbb{Q}$ par l'adjonction « d'un certain nombre de grandeurs variables ou indépendantes, et de fonctions algébriques de celles-ci ». Il démontre ensuite le théorème de l'élément primitif, qui permet d'écrire toute extension de type fini de $\mathbb{Q}$ sous la forme $\mathbb{Q}(\mathfrak{R}', \mathfrak{R}'', \mathfrak{R}''', ...)$ où tous les $\mathfrak{R}$ sauf un sont des indéterminées. Il mentionne aussi l'existence d'extensions qui ne sont pas de type fini, telle la clôture algébrique de $\mathbb{Q}$, mais il choisit de restreindre son propos aux extensions de type fini [4].

Au § 4 apparaît une modalité importante de la pensée de Kronecker : l'exigence d'effectivité. La théorie des corps qu'il vient d'exposer permet certes une définition rigoureuse de la notion d'irréductibilité, relative à un corps de base, mais cette définition reste incomplète, selon Kronecker, tant qu'elle n'est pas accompagnée d'un algorithme pour tester l'irréductibilité d'un polynôme donné. Plus précisément, il exige un algorithme pour tester l'irréductibilité d'un polynôme $F(x, x', x'', ...)$ à coefficients dans une extension de type fini de $\mathbb{Q}$. Il remarque d'emblée que d'habitude on avait recours à la théorie de l'élimination pour résoudre ce problème. Et que l'on pense en effet aux travaux des algébristes du XVII$^{\text{ème}}$ pour la recherche de facteurs rationnels (Jan Hudde) par la méthode des coefficients indéterminés, puis aux travaux de Lagrange que nous avons déjà commentés. Les raisonnements sur lesquels Bézout (et plus tard Cayley) fonde sa théorie de l'élimination, et que les progrès de l'algèbre linéaire jusqu'à Kronecker ne peuvent éclaircir d'avantage, ont un défaut que la théorie de l'élimination conçue par Poisson (puis Hesse et Schläfli) ne peut relever, puisque celle-ci se place d'emblée dans le cadre restreint du résultant d'un système de n équations ayant un nombre

3. *cf.* [2] t. 2 p. 224. Sur le caractère relatif de la notion d'irréductibilité chez Abel, voir aussi *supra* n. 5 p. 41.

4. En fait, Kronecker n'impose pas de limite au nombre d'interminées $\mathfrak{R}', \mathfrak{R}'', \mathfrak{R}'', ...$ Il pourrait y en avoir une infinité, et il serait donc plus exact de parler, de manière générale, d'« extensions finies d'extensions purement transcendantes de $\mathbb{Q}$ ».

fini de solutions. Une théorie plus générale était encore à naître. Revenons au § 4 des *Grundzüge*. Kronecker juge que :

> ...lors du développement naturel et complet de la théorie de l'élimination on utilise la décomposition d'une fonction entière en ses facteurs. C'est pourquoi on aura recours ici à une nouvelle méthode [de décomposition] qui ne nécessite que des moyens simples, déjà applicable ici.

Ce qui semblait d'abord n'être qu'une exigence gratuite d'effectivité revêt donc une autre dimension. Kronecker entend fonder la théorie de l'élimination. Certes la théorie kroneckerienne de l'élimination est une théorie effective : il ne s'agit ni de donner une définition *a priori* du résultant, ni de donner seulement une borne au nombre de solutions d'un système d'équations polynomiales, mais vraiment de résoudre un tel système absolument quelconque. La méthode d'élimination proposée par Kronecker consiste en fait, interprétée en termes modernes, en la décomposition d'une variété algébrique affine (ensemble des solutions d'un système d'équations polynomiales) en ses composantes irréductibles. Et cette méthode requiert en effet un algorithme pour décomposer un polynôme en ses facteurs irréductibles. Décrivons rapidement celui que donne Kronecker. En faisant le produit des conjugués du polynôme $F(x, x', x'', ...)$ et en posant

$$x' = c_1 x^g, \quad x'' = c_2 x^{g^2}, ... \quad x^{(n)} = c_n x^{g^n}$$

pour g assez grand, il se ramène au cas d'un polynôme $F(x)$ à une indéterminée, à coefficients dans $\mathbb{Z}$. Dans ce cas, il représente le facteur $f(x)$ cherché par une formule d'interpolation

$$f(x) = f(r_0)g_0(x) + f(r_1)g_1(x) + ... + f(r_n)g_n(x)$$

où les r_i sont des entiers distincts quelconques, et les $g_n(x)$ certains polynômes. Si $f(x)$ est un facteur de $F(x)$, alors pour tout i, $f(r_i)$ divise $F(r_i)$, et il n'y a donc qu'un nombre fini de facteurs $f(x)$ possibles. A ce stade plane un mystère : quelle place peut bien avoir la théorie de l'élimination au sein d'une « théorie arithmétique des grandeurs algébriques » ? Une première réponse nous sera donnée par le rôle qu'occupe le discriminant dans la théorie arithmétique de Kronecker.

15.2 Grandeur entière et discriminant

Au §§ 5–7, Kronecker définit la « grandeur entière ». Il utilise à cet effet une nouvelle notation

$$[\mathfrak{R}', \mathfrak{R}'', \mathfrak{R}''', ...]$$

pour désigner l'anneau que l'on noterait aujourd'hui $\mathbb{Z}[\mathfrak{R}', \mathfrak{R}'', \mathfrak{R}''', ...]$. Kronecker classe les grandeurs entières en « espèces » (*Art*) au sein d'un « genre ». Si $\mathfrak{R}'$, $\mathfrak{R}''$, $\mathfrak{R}'''$,... sont des indéterminées, les entiers de $\mathbb{Q}(\mathfrak{R}', \mathfrak{R}'', \mathfrak{R}''', ...)$ sont simplement les éléments de $\mathbb{Z}[\mathfrak{R}', \mathfrak{R}'', \mathfrak{R}''', ...]$ [5]. Dans une extension de corps, une grandeur est dite entière si elle est solution d'une équation à coefficients entiers dans le corps de base et coefficient dominant égal à 1. Si $\mathfrak{S}'$, $\mathfrak{S}''$, $\mathfrak{S}'''$,... sont des entiers algébriques dans un « domaine de genre » $K = \mathbb{Q}(\mathfrak{G}, \mathfrak{R}', \mathfrak{R}'', \mathfrak{R}''', ...)$ sur un corps de base $k = \mathbb{Q}(\mathfrak{R}', \mathfrak{R}'', \mathfrak{R}''', ...)$, Kronecker nomme l'anneau $\mathbb{Z}[\mathfrak{R}', \mathfrak{R}'', \mathfrak{R}''', ..., \mathfrak{S}', \mathfrak{S}'', \mathfrak{S}''', ...]$ un « domaine d'espèce ». La réunion de tous les domaines d'espèces d'un même domaine de genre constitue la clôture intégrale de l'anneau d'entiers du corps de base dans son extension. L'objectif des §§ 6 et 7 est de démontrer que cet ensemble est bien lui-même un « domaine d'espèce », que Kronecker qualifiera d'« espèce principale » (*Haupt-Art*). Kronecker démontre un peu plus : notons O_k l'anneau d'entiers du corps de

5. Une définition de la notion d'entier algébrique commence nécessairement par la donnée, assez arbitraire, d'un anneau d'entiers en quelque sorte « absolu », intégralement clos, dont chaque autre anneau d'entiers sera une certaine clôture intégrale dans une extension de son corps de fractions. En caractéristique zéro, le choix naturel est celui de $\mathbb{Z}$. Mais parfois, Kronecker énonce des résultats « en faisant abstraction des coefficients entiers » (*sofern von den Zahlcoefficienten abgesehen wird*, *cf.* par exemple début du § 7). Cela revient alors à définir l'anneau d'entiers absolu comme étant $\mathbb{Q}[\mathfrak{R}', \mathfrak{R}'', \mathfrak{R}''', ...]$, où $\mathfrak{R}'$, $\mathfrak{R}''$, $\mathfrak{R}'''$,... sont des indéterminées (remarquez que dans ces conditions, aussi bien $\mathbb{Z}[\mathfrak{R}', \mathfrak{R}'', \mathfrak{R}''', ...]$ que $\mathbb{Q}[\mathfrak{R}', \mathfrak{R}'', \mathfrak{R}''', ...]$ sont intégralement clos dans leur corps de fraction commun $\mathbb{Q}(\mathfrak{R}', \mathfrak{R}'', \mathfrak{R}''', ...)$).

base k et O_K sa clôture intégrale dans K, alors O_K est un O_k-module de type fini. Au § 6, il montre que O_K est un O_k-module de type fini, en se ramenant au cas où k est une extension purement transcendante de $\mathbb{Q}$ (au § 7, il remarque que dans certains cas, un nombre n de générateurs suffisent, où n est le degré de K sur k)[6]. A cet effet, il introduit le concept de *discriminant de n grandeurs* $x^{(1)}, x^{(2)}, ..., x^{(n)}$ d'un genre d'ordre n : c'est le carré du déterminant $\left|x_i^{(h)}\right|$ $(h, i = 1, 2, ..., n)$ où pour tout h, $x_1^{(h)}$, $x_2^{(h)}$,..., $x_n^{(h)}$ sont les conjugués[7] de $x^{(h)}$. Ce concept était déjà présent, sans être nommé, dans son article de 1862 (publié en 1881) *Über die Discriminante algebraischer Functionen einer Variabeln*. Comme son titre l'indique, ce texte se restreignait au cas des grandeurs algébriques sur un corps de base de la forme $k = \mathbb{Q}(v)$ (où v est une indéterminée). Kronecker y étudiait le *discriminant d'une grandeur x* solution d'une équation irréductible de degré n à coefficients dans ce corps, discriminant défini comme étant le carré de $\left|x_i^h\right|$ où les x_i^h (pour $1 \leq i \leq n$ et $0 \leq h \leq n-1$) sont les conjugués des n premières puissances de x. Le *discriminant d'une grandeur x* est donc le *discriminant des n grandeurs* $1, x, ..., x^{n-1}$, et il coïncide d'ailleurs avec le *discriminant de l'équation*, défini comme produit des différences des racines. Soit x une grandeur entière algébrique sur k, et $K = k(x)$. Dans ce cas particulier, n générateurs suffisent à engendrer le O_k-module O_K. Kronecker montre que, pour tout autre grandeur algébrique entière y sur k telle que $K = k(y)$, le discriminant Δ de ces n générateurs divise le discriminant de y. Il l'appelle alors « facteur essentiel » du discriminant de y (*wesentliche Theiler*). Il montre même que Δ est le plus grand commun diviseur de tous ces discriminants[8], ce qui lui donne un sens intrinsèque (indépendant du choix des générateurs). Il montre de plus que le « facteur non essentiel » du discriminant de y est toujours un carré. Dans la préface de 1881 à cet article, Kronecker écrivait :

> Je pourrais décrire comme étant l'un des principaux fruits de ce travail le fait que le traitement du cas pour ainsi dire général ou régulier des équations algébriques entre deux variables, auquel Riemann s'est restreint en remarquant que « l'on peut aisément étendre les résultats aux autres cas, regardés comme cas limites », suffit aussi parce que les autres équations peuvent être transformées en de telles.[9]

Riemann, dans sa *Théorie des fonctions abéliennes*, s'était en effet restreint au cas des fonctions dont les points de ramification « coïncident seulement par paires, et en se détruisant » (*cf.* § VI). Kronecker démontre qu'il existe un $\xi \in K$ dans le discriminant duquel le facteur essentiel et le facteur non essentiel n'ont aucune racine commune et les racines du facteur non essentiel sont toutes des racines de multiplicité exactement égale à 2. Géométriquement, les racines du facteur essentiel sont le lieu de ramification de ξ, vu comme fonction de la variable v, et les racines du facteur non essentiel correspondent à ce que Riemann appelait des « points de ramification qui coïncident », c'est-à-dire, dans un langage plus moderne, à des points singuliers de la courbe d'équation $F(\xi, v) = 0$ (remarquons que la « surface de Riemann » de ξ n'est autre qu'un modèle non singulier de cette courbe, au sein de sa classe d'équivalence birationnelle). Kronecker est donc parvenu à transformer l'équation en une équation dont les points singuliers sont tous des points doubles d'abscisses distinctes et situés hors de son lieu de ramification. Noether (*cf.* [49], p. 370–375) y a vu la première méthode de résolution des singularités des courbes algébriques planes, préfigurant les résultats de peu postérieurs (début des années 1870) de Noether, Hamburger et Weierstrass (*cf. supra* section 13.2).

Mais revenons aux *Grundzüge*. Le discriminant (et le lieu de ramification) était un invariant important en théorie des nombres (Dirichlet avait par exemple donné une formule dépendant du discriminant pour le nombre de classes d'un corps quadratique) ainsi qu'en théorie des fonctions d'une variable (par exemple pour déterminer le genre d'une surface de Riemann). Kronecker cherche

6. A partir du § 6, et au moins jusqu'au § 18, Kronecker utilise souvent l'hypothèse que k est une extension purement transcendante de $\mathbb{Q}$ (*natürliche Rationalitäts-Bereich*). C'est le fait que k ait un anneau d'entiers factoriel qui importe dans les démonstrations. On n'obtient donc une théorie relative de la divisibilité dans les anneaux d'entiers que si l'anneau d'entiers du corps de base a lui-même « naturellement » une telle théorie. De toute façon, comme on l'a déjà remarqué, tous les corps considérés par Kronecker sont des extensions finies d'extensions purement transcendantes de $\mathbb{Q}$.

7. Lorsque les n grandeurs sont des entiers algébriques, le fait que les conjugués d'un entier soient eux-mêmes des entiers implique que le discriminant est un entier algébrique.

8. *Cf.* p. 221.

9. *Cf.* p. 199.

à présent (§ 8) à généraliser cette notion au cas de plusieurs variables. Mais, dans le cas général, le système de générateurs du O_k-module O_K a plus de n éléments, et Kronecker considère le « système fondamental de discriminants », ensemble des discriminants des générateurs pris n à n, ainsi que (dans le langage d'aujourd'hui) l'idéal engendré par tous ces discriminants [10]. Géométriquement, c'est-à-dire quand les indéterminées du corps de base sont conçues comme des variables, il explique que la variété des zéros de cet idéal est un « invariant du genre ». Dans son langage, il représente en fait cet idéal par une forme linéaire dont les coefficients sont les discriminants des générateurs pris n à n (on verra plus loin le rapport avec le concept d'idéal). Dans le cas où l'anneau d'entiers du corps de base est factoriel, il propose de définir le *discriminant du genre* comme étant le plus grand commun diviseur Δ des discriminants des générateurs pris n à n (sa variété des zéros est, nous dit Kronecker, une partie de la variété des zéros de l'idéal ci-dessus). Kronecker montre qu'alors, si x', x'',..., $x^{(n+m)}$ sont les générateurs de O_K, et $u_1x' + u_2x'' + ... + u_{n+m}x^{(n+m)}$ un élément générique de O_K, le facteur de son discriminant [11] indépendant des u divise une puissance de Δ.

Jusqu'ici, Kronecker a donc posé les fondements d'une théorie des extensions de corps et des anneaux d'entiers, dans laquelle le polynôme minimal d'une grandeur (qui détermine ses conjugués, l'ordre du genre auquel elle appartient, son discriminant, et est unitaire à coefficients entiers si et seulement si la grandeur est entière [12]) joue un rôle important. Mais une théorie complète de la « grandeur algébrique » ne peut faire l'économie de l'étude des « systèmes absolument quelconques d'équations » (*ganz allgemeiner Gleichungssysteme*). Il faudrait par exemple généraliser le concept de discriminant d'une grandeur algébrique (solution d'une équation à coefficients dans le corps de base) au cas d'un système de grandeurs, solution d'un système d'équations à plusieurs inconnues. Au § 10, Kronecker remarque que les rapports des $(n+1)$ grandeurs $x^0, x', ..., x^{(n)}$ solutions d'un système de n équations homogènes $F_1 = 0, F_2 = 0, ..., F_n = 0$ sont algébriques et définissent un « genre » (en toute rigueur, il faut supposer ici qu'il n'y a qu'un nombre fini de solutions) : cet ensemble de grandeurs engendre une extension algébrique du corps des coefficients des équations. Dans ce cas, il propose de déterminer directement le discriminant du genre, au moyen d'un jacobien (généralisant ainsi le discriminant d'une équation à une inconnue, résultant de l'équation et de sa dérivée). Il se donne une $(n+1)$-ième forme de degré quelconque F_0 à coefficients indéterminés et considère le résultant des $(n+1)$ équations

$$F_1 = 0, F_2 = 0, ..., F_n = 0, |F_{gh}| = 0$$

où $|F_{gh}|$ est le jacobien de $F_0, F_1, ..., F_n$. Il affirme alors, sans le démontrer, que le discriminant du genre [13] est un facteur irréductible (quand les F sont génériques) de ce résultant, indépendant des coefficients de F_0, et dont l'annulation est une condition nécessaire et suffisante pour que deux des solutions du système d'équations initial soient égales.

15.3 Théorie de l'élimination

Kronecker se propose de publier ultérieurement une étude générale sur les « systèmes absolument quelconques d'équations » et d'en indiquer seulement les principaux résultats dans les *Grundzüge*. Il décrit alors, en un paragraphe que nous allons traduire intégralement, un véritable programme de recherches pour la théorie de l'élimination, au sein de laquelle le contexte relatif des « domaine de base » et « domaine de genre » aura désormais un rôle à jouer :

10. Dans la terminologie moderne, cet idéal est un « idéal déterminantiel », *cf.* [123].

11. On remarquera de plus que la non-annulation du discriminant de $u_1x' + u_2x'' + ... + u_{n+m}x^{(n+m)}$, vu comme polynôme en $u_1, ..., u_n$, est une condition pour que cet élément soit un élément primitif (puisqu'alors tous ses conjugués sont distincts). Kronecker utilise ainsi le discriminant dans un cas particulier, fin § 12.

12. On est ici en caractéristique 0, donc si α est un entier algébrique, son polynôme minimal est $\prod(X - \alpha_i)$, où les α_i sont les conjugués distincts de α, entiers algébriques eux aussi. L'équation $\prod(X - \alpha_i) = 0$ est alors une équation de dépendance intégrale de α.

13. Pour donner un sens à cette affirmation, il faut préciser quel est l'anneau d'entiers du corps de base, soit par exemple $\mathbb{Z}[\mathfrak{j}', \mathfrak{j}'', ...]$ où $\mathfrak{j}', \mathfrak{j}'', ...$ sont les coefficients (génériques) des équations. Pour une expression plus classique de ce « discriminant », *cf.* [234], tome I, volume 2, § 66, p. 147–148.

> Un système d'un nombre quelconque d'équations algébriques pour $z^0, z', z'', ..., z^{(n-1)}$, dont les coefficients appartiennent au domaine de rationalité $(\mathfrak{R}', \mathfrak{R}'', \mathfrak{R}''', ...)$, définit des relations algébriques entre les grandeurs z et $\mathfrak{R}$, dont la connaissance et la représentation constituent le but de la théorie de l'élimination. Les m fonctions $G_1, G_2, ..., G_m$ qui, posées égales à zéro, constituent les équations, doivent être supposées fonctions rationnelles entières des n grandeurs z, mais dont les coefficients sont seulement fonctions rationnelles (entières ou fractionnaires) des grandeurs $\mathfrak{R}$ à coefficients entiers. Quant au nombre des fonctions, que l'on note m, il ne faut aucunement le restreindre ; il peut, comme dans le cas particulier – dit général – ci-dessus [cas du discriminant de n équations homogènes à $(n+1)$ inconnues] être égal à n, c'est-à-dire au nombre des grandeurs z à déterminer, mais il peut aussi être plus grand ou plus petit que ce nombre. Si l'on conçoit les grandeurs z comme des variables libres (*als unbeschränkt veränderlich*), alors les équations $G = 0$ constituent une certaine restriction de cette variabilité (*eine gewisse Beschränkung dieser Variabilität*), dont une caractérisation plus précise peut être décrite comme étant le problème de l'élimination. Il faut alors maintenir l'indétermination des grandeurs indéterminées parmi les grandeurs $\mathfrak{R}', \mathfrak{R}'', \mathfrak{R}''', ...$, ou leur libre variabilité si elles sont conçues comme des variables ; c'est-à-dire que seule la variabilité des grandeurs z est à caractériser, ce qui ne demande aucune restriction de la variabilité des grandeurs $\mathfrak{R}$. La distinction ainsi faite entre les variables z et celles qui surviennent parmi les grandeurs $\mathfrak{R}$ est de la plus grande importance ; elle n'implique aucune restriction de la généralité, elle sépare plutôt seulement les problèmes distincts qui peuvent être posés par l'élimination, selon leur contenu conceptuel.

On dirait aujourd'hui que Kronecker se propose d'étudier un idéal définissant une variété plongée dans un espace affine $\mathbb{A}^n_{\mathbb{Q}(\mathfrak{R}', \mathfrak{R}'', \mathfrak{R}''', ...)}$. Un des « principaux résultats » de la théorie de Kronecker réside en la possibilité de ramener le cas d'un système quelconque d'équations au cas d'une unique équation. Kronecker l'explique géométriquement à la fin du § 10 : toute composante Y de dimension k de la variété $X \subset \mathbb{A}^n$ peut être projetée birationnellement [14] sur une hypersurface d'une certaine sous-variété linéaire de dimension $(k+1)$. Soient $x^{(0)}, x^{(1)}, ..., x^{(n-1)}$ des coordonnées suffisamment générales de l'espace affine $\mathbb{A}^n$; il suffira de projeter sur la sous-variété linéaire

$$\{x^{(k+1)} = x^{(k+2)} = ... = x^{(n-1)} = 0\} \simeq \mathbb{A}^{k+1}.$$

D'ailleurs, la projection de la composante de dimension k sur la sous-variété

$$\{x^{(0)} = x^{(k+1)} = x^{(k+2)} = ... = x^{(n-1)} = 0\} \simeq \mathbb{A}^k$$

en fait un revêtement ramifié. On a en effet le diagramme commutatif suivant :

$$\begin{array}{ccccc}
(x^{(0)}, x^{(1)}, ..., x^{(n-1)}) & \longmapsto & (x^{(0)}, x^{(1)}, ..., x^{(k)}, 0, ..., 0) & \longmapsto & (0, x^{(1)}, x^{(2)}, ..., x^{(k)}, 0, ..., 0) \\
\mathbb{A}^n & \longrightarrow & \mathbb{A}^n & \longrightarrow & \mathbb{A}^n \\
\uparrow & & \uparrow & & \uparrow \\
Y & \xrightarrow{p_k} & \mathbb{A}^{k+1} & \xrightarrow{q_k} & \mathbb{A}^k \\
 & & (x^{(0)}, x^{(1)}, ..., x^{(k)}) & \longmapsto & (x^{(1)}, x^{(2)}, ..., x^{(k)})
\end{array}$$

Notons $\pi_k = q_k \circ p_k$. Pour établir ces résultats, Kronecker donne une méthode effective de décomposition d'une variété en composantes irréductibles, qui consiste à éliminer les inconnues du système d'équations successivement l'une après l'autre. Chaque composante peut enfin être représentée comme lieu des zéros d'un système d'au plus $(n+1)$ équations.

14. Pour retrouver ces résultats dans un contexte moderne, *cf.* [140], exercice I.4.8 p. 31, et l'article récent de Marc Giusti et Joos Heintz [125], qui étudie le coût (théorie de la complexité) de la méthode de Kronecker. Enfin [128] et [208] suggèrent que la méthode de Kronecker est une alternative efficace aux algorithmes d'élimination ayant recours à des « techniques de réécriture » (c'est-à-dire aux bases de Gröbner).

Kronecker commence par décrire sa méthode d'élimination. Supposons $x^{(n-1)}$, $x^{(n-2)}$,...,$x^{(k+1)}$ déjà éliminées, et l'image par p_k des composantes de X de dimension $\leq k$ définie par un idéal $(H_1, H_2, ..., H_m)$ qui, outre l'image de ces composantes, peut encore s'annuler sur certaines variétés « immergées » [15], images de variétés incluses dans des composantes de X de dimension $> k$.

Le choix d'un système de coordonnées suffisamment général permet d'assurer que l'image de toute composante irréductible de dimension $\leq k$ de X par p_k est un fermé, que π_k fait de toute composante irréductible de X de dimension k un revêtement ramifié de $\mathbb{A}^k$, et que les images par p_k de deux composantes de dimension $\leq k$ distinctes sont distinctes. Démontrons la première de ces trois assertions. La projection p_k vérifie le diagramme commutatif suivant :

$$\begin{array}{ccc}
(Z : X^{(0)} : X^{(1)} : ... : X^{(n-1)}) & \longmapsto & \{X^{(k+1)} = X^{(k+2)} = ... = X^{(n-1)} = 0\} \cap L_{(Z:X^{(0)}:X^{(1)}:...:X^{(n-1)})} \\
\mathbb{P}^n - \{Z = X^{(0)} = X^{(1)} = ... = X^{(k)} = 0\} & \longrightarrow & \mathbb{P}^n \\
\downarrow & & \downarrow \\
\mathbb{A}^n & \xrightarrow{p_k} & \mathbb{A}^n \\
(\frac{X^{(0)}}{Z}, \frac{X^{(1)}}{Z}, ..., \frac{X^{(n-1)}}{Z}) & \longmapsto & (\frac{X^{(0)}}{Z}, \frac{X^{(1)}}{Z}, ..., \frac{X^{(k)}}{Z}, 0, ..., 0)
\end{array}$$

où $L_{(Z:X^{(0)}:X^{(1)}:...:X^{(n-1)})}$ est la sous-variété linéaire projective de codimension $(k+1)$ contenant $(Z : X^{(0)} : X^{(1)} : ... : X^{(n-1)})$ et $\{Z = X^{(0)} = X^{(1)} = ... = X^{(k)} = 0\}$. Notons $\overline{Y}$ la clôture projective d'une composante Y de X de dimension $\leq k$. Alors

$$\dim \overline{Y} \cap \{Z = 0\} \leq k - 1.$$

Or

$$\dim\{Z = X^{(0)} = X^{(1)} = ... = X^{(k)} = 0\} = n - k - 2$$

Donc, d'après le théorème de Bertini, pour un choix de coordonnées $X^{(0)}$,...,$X^{(k)}$ suffisamment générales, on aura :

$$\dim \overline{Y} \cap \{Z = X^{(0)} = X^{(1)} = ... = X^{(k)} = 0\} < 0$$

et $\overline{Y}$ évite $\{Z = X^{(0)} = X^{(1)} = ... = X^{(k)} = 0\}$. L'image de Y par p_k est donc fermée.

On sépare de $(H_1, H_2, ..., H_m)$ ses composantes de dimension k en cherchant le plus grand commun diviseur des polynômes H_i, soit F_{n-k} (ce qui suppose un algorithme de factorisation dans les anneaux de polynômes, dont on a déjà parlé). Pour tout i, notons aussi $K_i = \frac{H_i}{F_{n-k}}$. On obtient ainsi un idéal $(K_1, K_2, ..., K_m)$ contenant l'image par p_k des composantes de dimension $\leq (k-1)$. On élimine ensuite $x^{(k)}$. Pour ce faire, Kronecker calcule simplement le résultant des deux polynômes en $x^{(k)}$

$$\left\{ \begin{array}{l} U_1K_1 + U_2K_2 + ... + U_mK_m \\ V_1K_1 + V_2K_2 + ... + V_mK_m \end{array} \right.$$

où $U_1, ..., U_m$ et $V_1, ..., V_m$ sont des indéterminées. Ce résultant est une forme en ces indéterminées, dont les coefficients engendrent à leur tour un idéal dont le lieu des zéros contient l'image par p_{k-1} des composantes de X de dimension $\leq k-1$ (ainsi que, peut-être, certaines variétés « immergées »). Conformément aux idées de Kronecker, pour étudier les composantes de dimension k, il faut « maintenir la variabilité » des grandeurs $x^{(1)}$, $x^{(2)}$,...,$x^{(k)}$. Le polynôme $F_{n-k} \in k(x^{(1)}, ..., x^{(k)})[x^{(0)}]$ fournit alors une équation

$$F_{n-k} = 0$$

décrivant la variabilité de la grandeur $x^{(0)}$ dans les composantes de dimension k, et cette équation a un nombre fini de racines dans $\overline{k(x^{(1)}, ..., x^{(k)})}$. Autrement dit, l'hypersurface d'idéal (F_{n-k}) est conçue comme un revêtement ramifié de $\{x^{(0)} = 0\} \subset \mathbb{A}^{k+1}$ (le choix de coordonnées suffisamment générales permet d'assurer que $x^{(0)}$ ne prend pas de valeur infinie).

15. Macaulay les appellera ainsi. Kronecker semble ne pas être conscient de l'existence de ces variétés immergées, dont Macaulay donnera un exemple (*cf.* [217], exemple (*ii*) p. 22).

Pour étudier de manière effective la relation entre chaque composante irréductible Y de X de dimension k et son image $p_k(Y)$, et en particulier la variabilité des autres variables $x^{(k+1)}$,...,$x^{(n-1)}$ en fonction de $x^{(1)}$,...,$x^{(k)}$, Kronecker utilise le même changement de variables que Poisson en 1802 ; il choisit comme nouveau système de coordonnées

$$x = u_0x^{(0)} + u_1x^{(1)} + ... + u_{n-1}x^{(n-1)},\ x^{(1)},\ x^{(2)},\ ...,\ x^{(n-1)}$$

où $u_0, u_1, ..., u_{n-1}$ sont des indéterminées [16]. On applique alors seulement la méthode d'élimination que nous venons de décrire, et on obtient

$$F_{n-k} \in k(u_0, u_1, ..., u_{n-1}, x^{(1)}, x^{(2)}, ..., x^{(k)})[x]$$

F_{n-k} est appelée « résolvente partielle » [17]. Kronecker appelle « résolvente totale » (*Gesammtresolvente*) le produit $\prod F_i = 0$ des résolventes partielles. A chaque racine ξ de F_{n_k} correspond un feuillet de l'une des composantes de dimension k, paramétré par les variables $x^{(1)}$,...,$x^{(k)}$, et dont les points ont pour coordonnées $(\xi, x^{(1)}, ..., x^{(k)}, \xi^{(k+1)}, \xi^{(k+2)}, ..., \xi^{(n-1)})$. Ces points ont pour coordonnées, dans le système de coordonnées initial :

$$(\xi^{(0)}, x^{(1)}, ..., x^{(k)}, \xi^{(k+1)}, ..., \xi^{(n-1)})$$

où l'on pose $\xi^{(0)} = \frac{1}{u_0}(\xi - u_1x^{(1)} - ... - u_kx^{(k)} - u_{k+1}\xi^{(k+1)} - ... - u_{n-1}\xi^{(n-1)})$. On peut donc écrire :

$$F_{n-k} = \prod_i (x - \xi_i) = \prod_i \left(x - (u_0\xi_i^{(0)} + u_1x^{(1)} + ... + u_kx^{(k)} + u_{k+1}\xi_i^{(k+1)} + ... + u_{n-1}\xi_i^{(n-1)}\right)$$

où, pour tout i,

$$(\xi_i^{(0)}, \xi_i^{(k+1)}, ..., \xi_i^{(n-1)}) \in \left(\overline{k(u_0, ..., u_{n-1}, x^{(1)}, ..., x^{(k)})}\right)^{n-k}$$

On a ainsi justifié l'affirmation de Kronecker, que les F_{n-k} sont « décomposables en facteurs linéaires » [18] en u_0,...,u_{n-1}. Le choix initial de coordonnées suffisamment générales permet encore d'assurer que les $\xi_i^{(0)}$,$\xi_i^{(k+1)}$,...,$\xi_i^{(n-1)}$ sont toutes finies, et l'on a ainsi décrit la réunion des composantes de dimension k de X comme revêtement ramifié de $\mathbb{A}^k$ par π_k. Mais il est possible que certains des feuillets

$$(\xi^{(0)}, x^{(1)}, ..., x^{(k)}, \xi^{(k+1)}, ..., \xi^{(n-1)})$$

ainsi décrits dépendent de $u_0, ..., u_{n-1}$. Macaulay a montré que si un tel feuillet existe, il appartient à une variété « immergée ». Donc toutes les composantes Y de dimension k de X sont décrites par des paramétrages indépendants de $u_0, ..., u_{n-1}$:

$$(\xi_i^{(0)}, \xi_i^{(k+1)}, ..., \xi_i^{(n-1)}) \in \left(\overline{k(x^{(1)}, ..., x^{(k)})}\right)^{n-k}$$

Un produit de d facteurs $(x - \xi)$ décrivant tous les feuillets d'une même composante irréductible Y est un facteur irréductible de $F_{n-k} \in k(u_0, ..., u_{n-1})[x, x^{(1)}, ..., x^{(k)}]$. Substituons à nouveau à x, dans ce facteur, sa valeur

$$x = u_0x^{(0)} + u_1x^{(1)} + ... + u_{n-1}x^{(n-1)},$$

16. Ici, nous changeons légèrement les notations de Kronecker. De plus, dans les *Grundzüge*, Kronecker parle en fait de *deux* changements de variables en même temps (nous faisions allusion au premier en disant qu'il fallait choisir un système de coordonnées « suffisamment général »). On pourrait croire qu'un deuxième changement de variables est superflu, puisque on a déjà supposé le premier suffisamment général pour éviter les cas d'exception. Mais le rôle du second changement de variables, comme l'a bien vu Macaulay, n'est pas d'éviter des cas d'exception. Il s'agit plutôt d'utiliser l'auxiliaire des coefficients indéterminés pour donner une expression algébrique de la relation entre Y et $p_k(Y)$. *Cf.* [217] p. 18, et aussi [234], II.2, p. 154 et 158.

17. Nous avons choisi, comme certains auteurs, l'orthographe française « résolvente » plutôt que « résolvante », car le mot « résolvante » est utilisé dans un autre contexte pour désigner les « résolvantes de Lagrange » et les « résolvantes de Galois ».

18. *cf.* [183] p. 30.

on obtient alors un polynôme $\Phi \in (k[u_0, ..., u_{n-1}])\,[x^{(0)}, x^{(1)}, ..., x^{(n-1)}]$. Kronecker écrit : [19]

> En donnant aux grandeurs u un certain nombre de système de valeurs, on voit facilement que $(n+1)$ tels systèmes de valeurs suffisent toujours à faire en sorte que les équations ainsi obtenues $\Phi_1 = 0$, $\Phi_2 = 0$,...,$\Phi_{n+1} = 0$ aient pour résolvente totale $\Phi = 0$.

En effet [20], pour $1 \leq i \leq n+1$, remplaçons les indéterminées $u_0, ..., u_{n-1}$ respectivement par d'autres indéterminées $u_0^{(i)}, ..., u_{n-1}^{(i)}$, on obtient un polynôme Φ_i qui s'annule identiquement sur Y. On peut voir la variété d'idéal (Φ_1) comme un fibré sur l'espace affine de coordonnées $u_0^{(1)}, ..., u_{n-1}^{(1)}$. Ecrivons $V(\Phi_1)$ comme réunion de ses composantes irréductibles

$$V(\Phi_1) = \bigcup_j Y_j^{(1)} \cup \bigcup_k Z_k^{(1)}$$

de sorte que les $Y_j^{(1)}$ soient les composantes irréductibles dont la fibre en $u_0^{(1)}, ..., u_{n-1}^{(1)}$ est constante, et que les $Z_k^{(1)}$ soient les composantes irréductibles dont la fibre varie avec $u_0^{(1)}, ..., u_{n-1}^{(1)}$. Ecrivons aussi $V(\Phi_1, \Phi_2)$ comme réunion de ses composantes irréductibles, en adoptant la même convention :

$$V(\Phi_1, \Phi_2) = \bigcup_j Y_j^{(2)} \cup \bigcup_k Z_k^{(2)}$$

Si l'on spécialise à présent $u_0^{(1)}, ..., u_{n-1}^{(1)}$ en des valeurs quelconques dans k, il est possible de spécialiser $u_0^{(2)}, ..., u_{n-1}^{(2)}$ dans k de sorte que :

$$\dim \bigcup_k Z_k^{(2)} = \dim \bigcup_k Z_k^{(1)} - 1$$

Ecrivons de même, pour tout $1 \leq i \leq n+1$,

$$V(\Phi_1, ..., \Phi_i) = \bigcup_j Y_j^{(i)} \cup \bigcup_k Z_k^{(i)}$$

On peut alors spécialiser les autres indéterminées $u_0^{(3)}, ..., u_{n-1}^{(3)}$,...,$u_0^{(n+1)}, ..., u_{n-1}^{(n+1)}$ de proche en proche, de sorte que :

$$n > \dim \bigcup_k Z_k^{(1)} > \dim \bigcup_k Z_k^{(2)} > ... > \dim \bigcup_k Z_k^{(n+1)}$$

Alors nécessairement $\bigcup_k Z_k^{(n+1)} = \varnothing$, et $V(\Phi_1, ..., \Phi_{n+1}) = Y$.

Avec Macaulay, on peut aussi utiliser Φ pour décrire l'équivalence birationnelle entre Y et $p_k(Y)$. Macaulay remarque [21] que le coefficient de u_0^d dans Φ donne une équation de $p_k(Y)$:

$$\phi(x^{(0)}, x^{(1)}, ..., x^{(k)}) = 0$$

Pour tout $k+1 \leq i \leq n-1$, le coefficient de $u_i u_0^{d-1}$ donne une équation de la forme :

$$x^{(i)}\phi' - \phi_i = 0,$$

où $\phi', \phi_i \in k[x^{(0)}, x^{(1)}, ..., x^{(k)}]$. Cela donne un sens précis à l'assertion faite par Kronecker dans un langage encore vague à la fin de cette section : [22]

19. *cf.* [183] p. 30. Kronecker ne démontre pas cette assertion. Selon Macaulay, König l'a démontrée dans [176] p. 234. Selon Netto et Le Vavasseur, [234] I 9 § 69, K. Th. Vahlen aurait montré que $(n+1)$ équations sont parfois *nécessaires*. Mais l'exemple qu'il donne est en fait erroné, *cf.* les cours de S. S. Abhyankar à Montréal (1970) sur le « nombre minimal d'équations définissant une courbe algébrique gauche », [3] p. 12.

20. La méthode démonstrative exposée ici nous a été enseignée par Marc Giusti ; nous l'utiliserons encore *infra* dans la démonstration du théorème 5 p. 253. Le choix des valeurs en lesquelles nous spécialisons les indéterminées peut être rendu effectif (grâce aux techniques exposées dans [125]).

21. *cf.* [217] p. 27–28.

22. *es lässt sich daher jedes m-fach ausgedehnte Gebilde einer beliebig grossen Mannigfaltigkeit auf ein solches eindeutig beziehen, das aus einer nur $(m+1)$-fachen Mannigfaltigkeit entnommen ist*, [183] p. 31.

Pour toute composante de dimension m d'une variété de dimension quelconque, il existe donc une correspondance biunivoque entre ses points et ceux d'une sous-variété d'une variété de dimension $(m+1)$.

15.4 Application : la théorie de Galois

Aux § 11 et 12, Kronecker nous offre un premier exemple d'application de sa théorie. « La théorie de l'élimination, nous dit-il, montre la source propre de la nouvelle lumière qui fut apportée par Galois dans la théorie des équations algébriques, il y a un demi-siècle ». Le corps de décomposition d'une équation irréductible

$$x^n - c_1 x^{n-1} + ... \pm c_n = 0$$

à coefficients dans le domaine de base $\mathbb{Q}(\mathfrak{R}', \mathfrak{R}'', \mathfrak{R}''', ...)$ est une extension finie, engendrée par ses racines $\xi_1, \xi_2, ..., \xi_n$. L'on peut aussi interpréter ces racines comme la variété solution d'un système de $(n+i)$ équations, que Kronecker définit explicitement :

$$(\bar{C}) \quad \mathfrak{f}_k(\xi_1, \xi_2, ..., \xi_n) = c_k, \quad \mathfrak{g}_i(\xi_1, \xi_2, ..., \xi_n) = c_i'$$

où, pour $k = 1, ..., n$, les $\mathfrak{f}_k$ sont les fonctions symétriques élémentaires des racines, et $\mathfrak{g}_i$ sont des fonctions rationnelles bien choisies [23] . Conformément aux résultats du paragraphe précédent, Kronecker montrera plus loin (à la fin du § 12) que $(n+1)$ équations suffisent en général. La résolvente du système d'équations est de la forme :

$$g(x, u_1, u_2, ..., u_n) = \prod (x - u_1 \xi_{r_1} - u_2 \xi_{r_2} - ... - u_n \xi_{r_n})$$

où le produit est étendu à toutes les permutations $(r_1, ..., r_n)$ appartenant au groupe de l'équation. En changeant les indices, on trouve :

$$g(x, u_1, u_2, ..., u_n) = \prod (x - u_{r_1} \xi_1 - u_{r_2} \xi_2 - ... - u_{r_n} \xi_n)$$

et l'on voit ainsi que g est une fonction des indéterminées $u_1, u_2, ..., u_n$ *invariante par les permutations du groupe*. Kronecker se flatte d'avoir ainsi donné à la théorie de Galois un « perfectionnement formel », en ayant remplacé la considération du groupe par celle d'une « fonction concrète, invariable par le groupe ». Le lecteur moderne s'étonnera de cette préférence pour une « fonction concrète », justement de la part de Kronecker, qui fit œuvre de pionnier dans l'axiomatisation de la notion de groupe abstrait [24] ! Mais le calcul de la résolvente donne ici encore une méthode *effective* pour déterminer le groupe de l'équation, étant donné la représentation de son corps de décomposition comme quotient d'un anneau de polynômes :

$$\mathbb{Q}(\mathfrak{f}_1, ...\mathfrak{f}_n)[x_1, ..., x_n] / (\mathfrak{f}_1(x_1, ..., x_n) - c_1, ..., \mathfrak{f}_n(x_1, ..., x_n) - c_n, \mathfrak{g}_1(x_1, ..., x_n) - c_0', ..., \mathfrak{g}_i(x_1, ..., x_n) - c_i')$$

Bien sûr, tout le problème est alors de déterminer les $\mathfrak{g}_i$, mais Kronecker contourne d'abord ce problème en remarquant que la résolvente g est un facteur irréductible de la résolvente du système d'équations

$$(C) \quad \mathfrak{f}_k(x_1, ..., x_n) = c_k \quad (k = 1, ..., n)$$

qui s'écrit :

$$\prod (x - u_1 \xi_{r_1} - u_2 \xi_{r_2} - ... - u_n \xi_{r_n})$$

23. *Cf.* § 12. On remarquera que la variété des solutions est un ensemble fini de points, obtenus en faisant agir le groupe de l'équation sur les coordonnées de l'un quelconque d'entre eux. Mais vue comme variété sur $\mathbb{Q}(\mathfrak{R}', \mathfrak{R}'', \mathfrak{R}''', ...)$, elle est irréductible. Par contre, la variété solution des équations $\mathfrak{f}_k(\xi_1, \xi_2, ..., \xi_n) = c_k$, qui consiste en l'ensemble des points dont les coordonnées sont une permutation quelconque des racines, n'est en général pas irréductible sur $\mathbb{Q}(\mathfrak{R}', \mathfrak{R}'', \mathfrak{R}''', ...)$. Les équations supplémentaires $\mathfrak{g}_i(\xi_1, \xi_2, ..., \xi_n) = c_i'$ sont donc nécessaires pour séparer les composantes irréductibles. Certainement Kronecker était-il conscient, en quelque manière, d'une telle interprétation géométrique (voir nos remarques plus loin sur ce sujet). Si l'on conçoit les coefficients de l'équation de départ comme étant des variables, on obtient un « revêtement galoisien ».

24. *Cf.* [332], pp. 44–48.

où le produit est cette fois étendu à toutes les permutations $(r_1, ..., r_n)$ du groupe symétrique. La théorie de Galois est donc ramenée à un problème de factorisation. Enfin les coefficients de g sont des fonctions des racines invariantes par les permutations du groupe, et ils fournissent d'ailleurs le système d'équations $\mathfrak{g}_i(\xi_1, ..., \xi_n) = c_i'$ cherché. Mais comme dans la fonction g, les rôles des indéterminées u et des racines ξ sont interchangeables, Kronecker préfère se ramener à l'étude des fonctions rationnelles des n indéterminées u dans le § 12 : après adjonction des indéterminées $u_1, ..., u_n$, le corps de décomposition de l'équation de départ est devenu corps de rupture de l'équation $g(x) = 0$ sur $\mathbb{Q}(\mathfrak{R}', \mathfrak{R}'', \mathfrak{R}''', ..., u_1, u_2, ..., u_n)$. Ce dernier corps $\mathbb{Q}(\mathfrak{R}', \mathfrak{R}'', \mathfrak{R}''', ..., u_1, u_2, ..., u_n)$ est lui-même corps de décomposition d'une équation en u, sur un sous-corps contenant les fonctions symétriques élémentaires de $u_1, ..., u_n$ et les coefficients de g. Le groupe de Galois de ce nouveau corps de décomposition est le même que le groupe de Galois de l'équation de départ. On a donc ramené la question à l'étude du corps des fonctions rationnelles de n indéterminées $u_1, ..., u_n$. Le paragraphe § 11 se présente donc comme une incursion rapide, et difficile à lire, en théorie de Galois : Kronecker n'écrit jamais le système d'équations $(\bar{C})$ (bien qu'il utilise cette notation pour le désigner), et une erreur typographique change le $(\bar{C})$ en (C) (fin de la page 33).

Après avoir ainsi ramené la question à l'étude du corps des fonctions rationnelles de n indéterminées, Kronocker propose (au § 12) de prendre à présent « comme point de départ », en théorie de Galois, l'étude de ce corps. Cela revient justement à faire la théorie de Galois d'une équation *générique* de degré n. Le corps des fonctions rationnelles à n indéterminées sur $\mathbb{Q}$ peut en effet se représenter comme extension finie du corps des fonctions symétriques, comme le montre le diagramme suivant, où l'on a aussi indiqué les anneaux d'entiers étudiés par Kronecker :

$$\begin{array}{lll} K = \mathbb{Q}(\mathfrak{f}_1, ..., \mathfrak{f}_n)[x_1, ..., x_n]/(\mathfrak{f}_1 - \sum x_i, ..., \mathfrak{f}_n - \prod x_i) & \longleftarrow & O_K = \mathbb{Q}[x_1, ..., x_n] \\ \quad = \mathbb{Q}(x_1, ..., x_n) & & \\ \uparrow & & \uparrow \\ k = \mathbb{Q}(\mathfrak{f}_1, ..., \mathfrak{f}_n) & \longleftarrow & O_k = \mathbb{Q}[\mathfrak{f}_1, ..., \mathfrak{f}_n] \end{array}$$

K est de degré $n!$ sur k. Kronecker montre que le O_k-module O_K est libre de rang $n!$, en exhibant une base [25]

$$\left\{ x_1^{h_1} ... x_{n-1}^{h_{n-1}} \right\} \begin{array}{l} h_k = 0, ..., n-k \\ k = 1, ..., n-1 \end{array}$$

Le *discriminant du genre* est alors le carré du déterminant

$$\left| \left(x_{\sigma(1)}^{h_1} ... x_{\sigma(n-1)}^{h_{n-1}} \right)_{h_1, ..., h_{n-1}, \sigma} \right|$$

où $h_k = 0, ..., n-k$ et σ parcourt le groupe symétrique Σ_n. En ordonnant convenablement les éléments de la base, on peut réécrire cette matrice comme produit AB des deux matrices suivantes. La matrice A est une matrice diagonale par blocs, le i-ème bloc sur la diagonale étant de la forme

$$\left(x_{\sigma\tau^i(2)}^{h_2} ... x_{\sigma\tau^i(n-1)}^{h_{n-1}} \right)_{h_2, ..., h_{n-1}, \sigma}$$

où $h_k = 0, ..., n-k$, où τ est la permutation cyclique $(12...n)$, et où σ parcourt le sous-groupe de Σ_n laissant i invariant. La matrice B est constituée elle aussi de $n \times n$ blocs de taille $(n-1)! \times (n-1)!$, le bloc (i, j) étant simplement $x_j^{i-1} \mathrm{Id}_{(n-1)!}$. Notons $\mathcal{D}$ le carré du déterminant de Vandermonde (c'est-à-dire le discriminant de l'équation). On vérifie alors aisément que $\det B = \pm \mathcal{D}^{\frac{(n-1)!}{2}}$. Par récurrence sur n, on montre alors que

$$\left| \left(x_{\sigma(1)}^{h_1} ... x_{\sigma(n-1)}^{h_{n-1}} \right)_{h_1, ..., h_{n-1}, \sigma} \right|^2 = \mathcal{D}^{\frac{n!}{2}}$$

25. On a donc un nouveau cas particulier dans lequel le nombre de générateurs d'une base de O_K est égal au degré de K sur k (Kronecker l'avait déjà annoncé au § 7).

comme l'annonce Kronecker sans le démontrer à la fin du § 12.

De même, si $L = \mathbb{Q}(\mathfrak{f}_1, ..., \mathfrak{f}_n, \mathfrak{g})$ est une sous-extension de degré ρ sur k, Kronecker démontre que ρ générateurs suffisent à engendrer le O_k-module $O_L = L \cap O_K$. Si l'équation

$$x^n - \mathfrak{f}_1 x^{n-1} + ... \pm \mathfrak{f}_n = 0$$

reste irréductible sur L, Kronecker dit que l'extension L est « propre » (*eigentlich*). Dans ce cas, le groupe de l'équation sur L est d'ordre $r = \frac{n!}{\rho}$, et Kronecker propose, plutôt que de parler du « groupe de l'équation » (notion qui renvoie à la nature de l'indétermination des racines), d'introduire le terme « genre d'affection » (*Affect-Gattung*), pour désigner le corps des fonctions rationnelles de n indéterminées invariantes par les permutations du groupe, notion qui renvoie donc, en quelque sorte, à la nature arithmétique des racines et des fonctions rationnelles des racines. L'équation est alors dite avoir une certaine « affection » d'ordre r, et ses racines appartenir à une certaine « classe » d'ordre r (*cf.* [183], p. 37). Ce faisant, Kronecker se réclame plus d'Abel que de Galois, et mentionne à cet égard un problème de Galois inverse :

> Il échappe à Galois un problème des plus intéressants, dans la théorie des équations algébriques, qu'Abel a trouvé et aussi traité. C'est le problème de déterminer toutes les équations d'une certaine classe pour un domaine de rationalité donné, et je vais ici l'exposer plus en détail, aussi parce qu'il montre clairement la nature arithmétique des questions algébriques. [26]

Si $\mathfrak{g}$ est un élément primitif de L sur k, déterminer les équations de genre d'affection L (c'est-à-dire les équations dont le groupe de Galois est le groupe de K sur L), à coefficients dans un corps donné, revient à trouver certaines des solutions dans ce corps d'une équation $\Phi(\mathfrak{g}, \mathfrak{f}_1, ..., \mathfrak{f}_n) = 0$ (qui n'est autre que l'équation de $\mathfrak{g}$ sur k). Il s'agit donc, remarque Kronecker, d'un « problème diophantien ». Pour mieux comprendre cette attitude, il suffira de rappeler ses travaux sur le théorème dit « de Kronecker-Weber ». Ces travaux s'inscrivent dans la continuité des recherches menées par Abel dans son manuscrit *Sur la résolution algébrique des équations*. Abel y proposait un certain problème de Galois inverse : « Trouver toutes les équations d'un degré déterminé quelconque qui soient résolubles algébriquement » [27]. Le théorème de Kronecker-Weber résout le problème de Galois inverse suivant : quelle est la forme générale d'une extension galoisienne de $\mathbb{Q}$ dont le groupe est abélien ? En 1853, Kronecker affirmait déjà que toute extension de $\mathbb{Q}$ dont le groupe est cyclique se plonge dans une extension engendrée par des racines de l'unité [28], et il en donnait l'esquisse d'une démonstration (Weber en donnerait plus tard une démonstration complète). En 1877, il généralise et énonce le théorème qui a gardé son nom, et souligne son intérêt pour la théorie des nombres :

> Toutes les racines des équations abéliennes à coefficients entiers sont fonctions rationnelles de racines de l'unité, et toutes les fonctions rationnelles de racines de l'unité sont racines d'équations abéliennes à coefficients entiers.
> Cette proposition donne, me semble-t-il, un précieux aperçu en théorie des nombres algébriques ; car elle constitue un premier progrès quant à leur classification naturelle, qui dépasse leur rassemblement en genres, jusque aujourd'hui le seul que l'on ait considéré. [29]

Dans ce même article, il affirme qu'il avait été conduit, en 1857, dans ses recherches sur les fonctions elliptiques et la multiplication complexe, à formuler une conjecture analogue sur les extensions abéliennes d'un corps quadratique imaginaire. Nous reviendrons sur ce sujet, étroitement lié à la théorie du corps de classes, quand nous commenterons la section § 19 des *Grundzüge*.

---

26. *Cf.* [183], p. 38.
27. *cf.* [2] t. 2 p. 219. Voir aussi [156] p. 62–65.
28. Kronecker écrit dans [177] p. 10 que « la racine de toute équation abélienne à coefficients entiers peut être représentée comme fonction rationnelle de racines de l'unité ». Mais ce sont seulement les équations de groupe cyclique qu'il désigne alors par l'expression « équation abélienne ».
29. *Cf.* [179], p. 69. Dans cet article, Kronecker utilise l'expression *Abelsche Gleichung* pour désigner une équation abélienne, et l'expression *einfache Abelsche Gleichung* pour désigner une équation dont le groupe est cyclique.

15.5 Théorie de la divisibilité dans les anneaux d'entiers algébriques

On a vu ci-dessus, dans la construction de la *résolvente* d'un système d'équations par Kronecker, une première utilisation de l'« auxiliaire des coefficients indéterminés » (*methodische Hülfsmittel der unbestimmten Coefficienten*). Il utilise à nouveau cette méthode de l'adjonction d'indéterminées au domaine de base dans les sections § 14 à 18, pour construire une théorie de la divisibilité dans les anneaux d'entiers. Pour qualifier cette méthode, il parlera plus loin d'un accroissement de la « dimension », en utilisant un langage géométrique :

> De même que la ligne des nombres réels s'étend par une « unité latérale » en le plan des nombres complexes, un domaine de grandeurs $[\mathfrak{R}', \mathfrak{R}'', \mathfrak{R}''', ...]$ est agrandi par les indéterminées [...], d'une certaine manière, en rapport à sa « dimension ». [30]

Il ne s'agit donc pas seulement d'un auxiliaire algébrique formel. Cette méthode est à la fois censée être la plus simple, mais aussi produire la description la plus simple de la réalité mathématique. Kronecker la compare à l'un des grands progrès des mathématiques du XIX$^{\text{ème}}$ siècle : bien des phénomènes de l'analyse gagnent en « simplicité » quand ils sont étudiés du point de vue de l'analyse complexe [31]. Quelle que soit l'interprétation suggérée ici par Kronecker, ses successeurs ont d'ailleurs naturellement adopté une interprétation géométrique des *Grundzüge*, en particulier de sa méthode d'élimination, comme on l'a déjà remarqué. Donc, bien que le langage de Kronecker soit rarement géométrique, le plus souvent arithmético-algébrique, il semble légitime d'en pousser l'interprétation au delà du formel, bien qu'il faille avoir recours pour cela au substrat intuitif auquel nous sommes aujourd'hui habitués, celui de la théorie des ensembles et des théories de structures (telle la notion d'idéal). C'est ce que l'on a fait ci-dessus en traduisant son langage dans les notations modernes de l'algèbre commutative. Pour faire une théorie de la divisibilité dans les anneaux d'entiers, qui ne sont pas toujours factoriels, l'on a recours aujourd'hui à la théorie des idéaux de Dedekind. Dans les § 14 à 18, Kronecker propose une autre solution, commentée de manière détaillée par Harold M. Edwards dans [97]. Nous allons seulement résumer son analyse, avant de proposer notre propre interprétation [32]. Kummer, dès 1844, avait construit une théorie de la divisibilité pour les entiers cyclotomiques (éléments de $\mathbb{Z}[\zeta_p]$, où ζ_p est une racine p-ième de l'unité, p premier), en introduisant des « nombres idéaux », afin de rétablir l'existence et l'unicité de la décomposition d'un entier en facteurs premiers. Un tel « nombre idéal » n'était pas défini autrement par Kummer qu'en énonçant la relation d'équivalence entre deux entiers cyclotomiques modulo ce nombre. Dedekind avait au contraire choisi comme fondement, à partir de 1871, l'« idéal », ensemble des entiers divisibles par un « nombre idéal » donné. Il introduisit plus tard (années 1890) des idéaux fractionnaires, et donna ainsi à la théorie la forme qu'elle a gardée jusqu'à aujourd'hui. Kronecker choisit de définir, plutôt que des « nombres idéaux » ou des « idéaux », des « diviseurs », définis (sur le modèle de Kummer) par une relation d'équivalence. Seulement (au contraire de Kummer) il exprime cette relation d'équivalence comme une véritable relation de divisibilité, à condition d'adjoindre à l'anneau d'entiers des indéterminées. Un diviseur est noté

$$\mathrm{mod}\ [\phi x + \phi' x' + \phi'' x'' + ...]$$

où $\phi x + \phi' x' + ...$ est une forme à coefficients $x, x', ...$ dans l'anneau d'entiers. Kronecker définit la norme $\mathrm{Nm}(\phi x + \phi' x' + ...)$ d'une forme (produit des conjugués, la conjugaison agissant sur les coefficients de la forme). La forme est dite « primitive » si le plus grand commun diviseur des coefficients de sa norme est 1 (il faut que l'anneau d'entiers du corps de base soit factoriel pour que cela ait un sens). Kronecker définit aussi la forme primitive $\mathrm{Fm}(\phi x + \phi' x' + ...)$, telle que :

$$P.\mathrm{Fm}(\phi x + \phi' x' + ...) = \mathrm{Nm}(\phi x + \phi' x' + ...)$$

30. *Cf.* § 22.
31. *Cf.* [183], p. 94.
32. H. Weyl a proposé encore une autre lecture, axiomatique, de la théorie des diviseurs de Kronecker, dans [328], chapitre 2. Il dresse une liste d'axiomes que doit vérifier le concept de « diviseur ». Puis il définit le diviseur $\mathfrak{A}$ comme une suite finie d'entiers algébriques $(\alpha_1, ..., \alpha_r)$, en donnant un critère de divisibilité d'un entier algébrique quelconque α par $\mathfrak{A}$, et il montre que cette définition vérifie les axiomes (p. 49).

où P est le plus grand commun diviseur des coefficients de la norme (appelé aujourd'hui son « contenu »). Un entier quelconque z est alors dit « divisible » (*theilbar*) par ce diviseur si et seulement si le quotient

$$\frac{z.\mathrm{Fm}(\phi x + \phi' x + ...)}{\phi x + \phi' x + ...}$$

est une forme Q à coefficients entiers algébriques. H. M. Edwards a remarqué que les lecteurs de Kronecker (Dedekind, Hurwitz) ont buté sur ce point, Kronecker ne laissant pas clairement entendre si le quotient Q doit être une forme à coefficients entiers algébriques, ou si Q doit simplement vérifier une équation $Nm(X-Q)=0$ à coefficients entiers dans le corps de base (les deux étant équivalents, bien que Kronecker ne le démontre pas)[33]. Deux diviseurs sont dits « absolument équivalents » lorsqu'ils définissent la même relation de divisibilité. Kronecker construit ainsi toute une théorie de la divisibilité. A la section § 14, il montre que le diviseur $\mod [x + u'x' + u''x'' + ...]$, défini par une forme *linéaire*, est en un certain sens le plus grand commun diviseur des entiers $x, x', x'', ...$ A la section § 15, il définit le vocabulaire et les notations, puis il énonce et il démontre son « premier théorème fondamental » :

> Quand le produit de deux formes algébriques entières [=formes dont les coefficients sont des grandeurs algébriques entières], dont l'une est primitive, est congruent à zéro modulo un diviseur algébrique, alors l'autre forme doit elle-même être divisible par le diviseur.[34]

Les sections § 16 et 17 sont consacrées à la démonstration d'un « second théorème fondamental » :

> Des diviseurs algébriques qui ont les mêmes éléments [=les mêmes coefficients] sont absolument équivalents.[35]

En particulier tout diviseur est absolument équivalent à un diviseur défini par une forme *linéaire*. Mais pour démontrer le « second théorème fondamental », Kronecker commence justement par démontrer (§ 16) plusieurs propriétés des diviseurs définis par des formes linéaires, dont un analogue du lemme de Gauss :

> Si le produit de deux diviseurs algébriques est divisible par un troisième, et que le premier diviseur n'a aucun diviseur commun avec le troisième, alors le deuxième est divisible par le troisième.[36]

H. M. Edwards remarque que Kronecker utilise, dans la démonstration du second théorème fondamental, l'existence, pour tout entier du corps de base, d'un diviseur qui le divise et soit défini par une forme linéaire, fait essentiel qu'il ne démontre proprement qu'à la section § 18. Enfin, à la section § 18, Kronecker décrit une méthode effective pour obtenir tous les diviseurs d'un entier du corps de base, et il en déduit l'existence et l'unicité de la décomposition en facteurs premiers d'un quelconque diviseur.

De ce point de vue, la théorie de Kronecker est donc une modification et une généralisation de celle de Kummer, en donnant le rôle central à la notion de plus grand commun diviseur (incarnée par les diviseurs définis par des formes linéaires, et qui apparaît dès le § 14) plutôt qu'à celle de facteur premier (l'existence et l'unicité de la décomposition en facteurs premiers n'étant décrite qu'au § 18, sans même qu'en soit proprement énoncé un théorème). Kronecker se flatte cependant d'avoir donné une existence « véritable » (*wirklich*) aux diviseurs en les représentant par des formes, grâce à l'« auxiliaire des coefficients indéterminés ». L'on pourrait en donner l'interprétation moderne suivante.

Soit une extension galoisienne K sur k, avec des anneaux d'entiers O_k et O_K :

$$\begin{array}{ccc} K & \longleftarrow & O_K \\ \uparrow & & \uparrow \\ k & \longleftarrow & O_k \end{array}$$

33. *Cf.* [97], p. 356 et 364–368. Ce problème amena Dedekind à concevoir son « théorème de Prague », qui permet, comme le montre Edwards, de simplifier considérablement la théorie de Kronecker. *Cf. infra* p. 255.
34. *Cf.* § 14, IX.
35. *Cf.* § 16, II.
36. *Cf.* § 16, V.

On suppose O_k factoriel. Soit $S \subset O_k[u_1, u_2, ...]$ l'ensemble multiplicatif des « formes primitives » en les indéterminées $u_1, u_2, ...$, à coefficients dans O_k (les indéterminées forment un ensemble dénombrable $\{u_1, u_2, ...\}$; on emploiera dans tout ce qui suit le mot « forme » pour désigner un polynôme en $u_1, u_2, ...$ *non nécessairement homogène*, suivant ainsi l'habitude de Kronecker). Alors $S^{-1}O_k[u_1, u_2, ...]$ est intégralement clos dans son corps de fractions $k(u_1, u_2, ...)$, et $O_K \otimes S^{-1}O_k[u_1, u_2, ...]$ est entier sur $S^{-1}O_k[u_1, u_2, ...]$. On a donc plongé O_k et O_K dans deux nouveaux anneaux d'entiers, et l'on a à présent la situation suivante :

$$\begin{array}{ccc} K(u_1, u_2, ...) & \longleftarrow & O_K \otimes S^{-1}O_k[u_1, u_2, ...] \\ \uparrow & & \uparrow \\ k(u_1, u_2, ...) & \longleftarrow & S^{-1}O_k[u_1, u_2, ...] \end{array}$$

Théorème 2 *L'anneau $O_K \otimes S^{-1}O_k[u_1, u_2, ...]$ est factoriel.*

Corollaire 1 *Cet anneau est noethérien et principal*[37]*.*

Démonstration. On démontrera d'abord que tout idéal $(x_1, ..., x_k)$ engendré par un nombre fini de générateurs est principal, engendré par $\text{pgcd}(x_1, ..., x_k) = x_1u_1 + ... + x_ku_k$, où $u_1, ..., u_k$ sont des indéterminées qui ne figurent pas dans $x_1, ..., x_k$.

On prolonge les morphismes de conjugaison de K à $K(u_1, u_2, ...)$, ce qui permet de définir une norme dans $K(u_1, u_2, ...)$ (il n'y a même pas besoin que l'extension K sur k soit galoisienne pour cela). On définit de même $\text{Fm}\alpha$, pour tout $\alpha \in O_K \otimes S^{-1}O_k[u_1, u_2, ...]$, tel que :

$$\text{Nm}\alpha = P_\alpha.\text{Fm}\alpha$$

où P_α est le plus grand commun diviseur des coefficients du numérateur de $\text{Nm}\alpha$. Alors $\text{Fm}\alpha$ est un quotient de formes primitives, donc une unité.

Soit $\alpha = x_1u_1 + ... + x_ku_k$. Montrons que pour tout i, $\frac{x_i\text{Fm}\alpha}{\alpha}$ est entier sur $S^{-1}O_k[u_1, u_2, ...]$. Il suffit pour cela de voir que l'équation

$$\text{Nm}\left(X - \frac{x_i\text{Fm}\alpha}{\alpha}\right) = 0$$

est une équation à coefficients entiers. Or,

$$\text{Nm}\left(X - \frac{x_i\text{Fm}\alpha}{\alpha}\right) = \frac{\text{Nm}(X\alpha - x_i\text{Fm}\alpha)}{P_\alpha.\text{Fm}\alpha}$$

et l'on voit facilement que $\text{Nm}(X\alpha - x_i\text{Fm}\alpha)$ est bien divisible par P_α (*cf.* la section § 14 de Kronecker). Donc pour tout i, x_i est divisible par α. Donc en termes d'idéaux, $(x_1, ..., x_k) = (\alpha)$.

Pour la suite de la démonstration, on aura besoin de deux lemmes. On dira qu'un diviseur de la forme $x_1u_1 + ... + x_ku_k$, avec $x_i \in O_K$ pour tout i, est un « diviseur linéaire ».

Lemme 3 *Soit p premier dans O_k. Alors p possède un diviseur linéaire premier qui n'est pas une unité.*

Démonstration (*cf.* section § 18). Il existe au moins un diviseur linéaire de p qui n'est pas une unité, à savoir pu_0. S'il en existe d'autres, on peut les écrire sous la forme $pu_0 + x_1u_1 + ... + x_ku_k$ (à une unité près), car si $(x_1, ..., x_k)$ est un diviseur de p alors $(x_1, ..., x_k) = (p, x_1, ..., x_k)$. Soit Σ l'ensemble de ces diviseurs. Tout diviseur linéaire de p est alors, à une unité près, un élément de Σ. Si $O_k = \mathbb{Z}$, quitte à réduire modulo p les entiers $x_1, .., x_k$, on voit que l'on peut d'ailleurs se limiter à un ensemble Σ fini[38]. On ordonne Σ par la relation de divisibilité. Montrons qu'il a un élément minimal. Soient

37. En fait, suivant Weyl ([328], p. 60), on peut même montrer que tout idéal est de la forme $(x_1u_1 + x_2u_2)$ où $x_1, x_2 \in O_K$.

38. Remarquons en effet que pour $O_k = \mathbb{Z}$, l'anneau des restes $O_K/(p)$ est de cardinal fini $p^{[K:k]}$.

$\alpha, \beta, \gamma \in \Sigma$, avec $\beta = \alpha\gamma$. Alors $P_\beta = P_\alpha P_\gamma$. Donc si α divise strictement β, $P_\gamma \neq 1$, et P_α divisera strictement P_β. Si $P_\alpha = 1$, α est minimal, sinon on recommence avec α ce que l'on vient de faire avec β. Comme O_k est factoriel, on obtient ainsi un élément minimal α après un nombre fini d'étapes. [39]. Montrons à présent que α est premier. Soient $\beta, \gamma \in O_K \otimes S^{-1}O_k[u_1, u_2, ...]$, tels que $\alpha \nmid \beta$, $\alpha \nmid \gamma$. Montrons que $\alpha \nmid \beta\gamma$. On fait une récurrence sur le nombre m d'indéterminées présentes dans β ou γ. On peut supposer qu'aucune de ces indéterminées n'apparaît dans α, quitte à multiplier par une unité [40]. Si $m = 0$, $\text{pgcd}(\beta, \alpha) = \beta u_1 + \alpha u_2$ est un diviseur linéaire de α, et, à cause du caractère minimal de α, est donc trivial. Donc $\beta\gamma u_1 + \alpha\gamma u_2 = \epsilon\gamma$ où ϵ est une unité, et comme α ne divise pas γ, il ne divise pas non plus $\beta\gamma$. Si $m > 0$, soit u l'une des indéterminées présentes dans β ou γ. On écrit :

$$\beta = \sum \beta_i u^i, \quad \gamma = \sum \gamma_j u^j$$

Alors si s et t sont les plus grands indices tels que $\alpha \nmid \beta_s$ et $\alpha \nmid \gamma_t$, le coefficient de u^{s+t} dans $\beta\gamma$ ne sera pas divisible par α (par récurrence sur m). Donc $\beta\gamma$ non plus.

Lemme 4 (Lemme de Gauss) *Soient $p, a, b \in O_K \otimes S^{-1}O_k[u_1, u_2, ...]$ avec p premier, $p \nmid a$ et $p \mid ab$. Alors $p \mid b$.*

Démonstration. Par hypothèse, $\text{pgcd}(p, a) = pu_1 + au_2 = \epsilon$ est une unité. Donc $bpu_1 + bau_2 = b\epsilon$, et $p \mid ab \Longrightarrow p \mid b$.

Corollaire 2 *Aux unités près, il y a unicité de la décomposition en facteurs premiers dans $O_K \otimes S^{-1}O_k[u_1, u_2, ...]$.*

Fin de la démonstration du théorème 2 (existence de la décomposition en facteurs premiers). Soit $\alpha \in O_K \otimes S^{-1}O_k[u_1, u_2, ...]$. Si P_α est une unité de O_k, α est aussi une unité. Sinon, soit $p \in O_k$ premier tel que $p \mid P_\alpha$. Il existe (lemme 3) un diviseur linéaire premier π de p, donc de $\text{Nm}\alpha$. Alors (lemme 4) π divise l'un des conjugués $\sigma(\alpha)$ de α. Donc $\sigma^{-1}(\pi)$ est un diviseur linéaire premier de α (il faut ici utiliser l'hypothèse que K est galoisien sur k ; *cf. remarque* ci-dessous). Soit

$$\beta = \frac{\alpha}{\sigma^{-1}(\pi)}$$

On a alors :

$$P_\beta . P_{\sigma^{-1}(\pi)} = P_\alpha$$

Si $P_\beta = 1$, c'est fini ; sinon l'on recommence avec β ce que l'on a fait avec α et l'on conclut par récurrence sur le nombre de facteurs premiers de P_α (car O_k est factoriel). *q. e. d.*
Remarque : on peut sûrement se passer de l'hypothèse que l'extension K est galoisienne dans la démonstration du théorème 2, en plongeant K dans une extension galoisienne.

Dans la section § 18, Kronecker annonce un théorème sur la ramification (il renvoie au § 25 pour plus de détails). Un entier premier $p \in O_k$ est dit *ramifié* si ses facteurs premiers dans $O_K \otimes S^{-1}O_k[u_1, u_2, ...]$ ne sont pas tous distincts.

Théorème 3 *Un entier premier $p \in O_k$ est ramifié si et seulement s'il divise le discriminant du genre.*

A la suite de ce théorème, Kronecker énonce quelques assertions qui sont en contradiction avec sa définition de la norme, comme produit des conjugués *distincts*. Désignons-donc par « norme absolue » le produit des n conjugués, où $n = [K : k]$. Soit $p \in O_k$ premier. Sa norme absolue est p^n. Alors la norme absolue de chaque diviseur premier de p divise p^n, et est donc elle-même de la forme

39. Remarquons toutefois que, dans le cas où Σ est infini, cette méthode n'est pas effective car elle ne précise pas comment tester si α est minimal, ni comment trouver, sinon, un nouveau diviseur β. Kronecker annonce qu'il donnera une méthode meilleure, dans la section § 25, pour la décomposition en facteurs premiers ; cette autre méthode marche dans le cas $O_k = \mathbb{Z}$, mais il semble douteux qu'elle soit applicable en général.

40. Deux diviseurs linéaires ayant les mêmes coefficients sont égaux à une unité près. Il s'agit là d'un théorème énoncé par Kronecker au § 16. En effet, pour tout i, $x_1 v_1 + ... + x_k v_k$ divise x_i. Alors $x_1 v_1 + ... + x_k v_k$ divise $x_1 u_1 + ... + x_k u_k$.

p^f. L'exposant f est appelé « ordre » du diviseur. Si p n'est pas ramifié, la somme des ordres des diviseurs premiers de p est donc égale à n. Lorsque $O_k = \mathbb{Z}$, le discriminant du genre est différent de 1. On a donc le théorème [41] :

Théorème 4 *Il n'existe pas d'extension finie non ramifiée de* $\mathbb{Q}$.

Sous les mêmes hypothèses, Kronecker montre que le nombres des restes dans O_K modulo un diviseur est égal à la norme du diviseur. A la section 19, il continue son incursion en théorie des nombres et propose d'appliquer le « principe d'équivalence de Kummer », c'est-à-dire de considérer, en termes modernes, le groupe des classes d'idéaux. Un diviseur, élément de $O_K \otimes S^{-1}O_k[u_1, u_2, ...]$, est dit appartenir à la « classe principale » (*Hauptclasse*) s'il appartient (à une unité près) à O_K. Le groupe des classes est alors défini comme quotient du monoïde multiplicatif de l'anneau $O_K \otimes S^{-1}O_k[u_1, u_2, ...]$ par la classe principale. La relation d'équivalence sous-jacente entre diviseurs est appelée par Kronecker « équivalence relative », pour la distinguer de la relation d'« équivalence absolue » entre deux diviseurs qui ne diffèrent qu'à une unité près. Kronecker donne deux critères équivalents d'équivalence relative. Notons H la classe principale, et supposons l'extension galoisienne (hypothèse que Kronecker omet mais qui semble ici nécessaire pour garantir que le monoïde quotient soit un groupe : l'inverse d'un diviseur, dans le groupe des classes, est en effet représenté par le produit de ses conjugués).
1^er^ critère d'équivalence relative : les diviseurs ϕ et ψ sont équivalents si et seulement s'il existe un diviseur χ tel que $\phi\chi$ et $\psi\chi$ appartiennent à H.
2^d^ critère d'équivalence relative : les diviseurs ϕ et ψ sont équivalents si et seulement s'il existe deux diviseurs η, $\theta \in H$ tels que $\phi = \frac{\eta}{\theta}\psi$.
Kronecker esquisse une démonstration de la finitude du groupe des classes, en construisant un ensemble fini de diviseurs, tel que pour tout diviseur ϕ, son inverse (dans le groupe des classes) soit l'un d'eux. *Démonstration.* Soit k un entier tel que :

$$k^n \leq \mathrm{Nm}\phi < (k+1)^n$$

On choisit une base de O_K, et on considère l'ensemble de tous les entiers algébriques à coefficients positifs $\leq k$. Il y en a $(k+1)^n > \mathrm{Nm}\phi$, donc deux d'entre eux sont égaux modulo ϕ, et leur différence est un diviseur principal $C\phi$, dont les coefficients sont inférieurs à k en valeur absolue. Soit $C\phi = k\eta$ où η est un nombre algébrique à coefficients inférieurs à 1. Alors $\mathrm{Nm}\eta$ est borné par une grandeur M indépendante de ϕ, et

$$\mathrm{Nm}C.\mathrm{Nm}\phi = k^n \mathrm{Nm}\eta \leq \mathrm{Nm}\phi.M$$

Donc $\mathrm{Nm}C \leq M$, et l'inverse C de ϕ peut être choisi dans l'ensemble fini des diviseurs de norme inférieure à M, *q. e. d.*

Les recherches décrites par Kronecker dans le reste de la section 19 constituent la préhistoire de la théorie du corps de classe. Helmut Hasse (*cf.* [52], ch. XI, « History of Class Field Theory ») a comparé les contributions respectives de Kronecker, Weber et Hilbert à la genèse de cette théorie. Dedekind avait déjà remarqué [42], dans le cas d'un corps de nombres, qu'à cause de la finitude du groupe des classes, pour tout diviseur α, il existe $m \in \mathbb{N}$ tel que α^m est principal, c'est-à-dire (à une unité près), égal à un élément $a \in O_K$:

$$\alpha^m = \epsilon a$$

Dans l'anneau d'entiers de l'extension $K(\sqrt[m]{a})$, le diviseur α^m est donc égal, *à une unité près*, à la puissance m-ième du diviseur principal $\sqrt[m]{a}$:

$$\alpha^m = \epsilon(\sqrt[m]{a})^m$$

En décomposant en diviseurs premiers les deux membres de cette égalité, on en déduit que le diviseur α est égal à $\sqrt[m]{a}$ à une unité près, et est donc principal. Comme il n'y a qu'un nombre

41. Kronecker énonce ce théorème à la section 8, p. 21. C'est une conséquence immédiate de la méthode de factorisation qu'il expose dans la dernière section des *Grundzüge*. *Cf.* [98], § 2.8, qui attribue ce théorème à Dedekind. Abhyankar [4] p. 444 l'attribue à Kronecker mais fait une remarque intéressante. Cassels, Fröhlich [52] l'attribuent à Minkowski qui en a aussi donné une démonstration.
42. *Cf.* [99], p. 61, remarque n°33 de Dedekind.

fini de classes, il suffit d'adjoindre un nombre fini de radicaux au corps K pour rendre principal chaque diviseur. Mais Kronecker ne se satisfait pas d'un tel procédé, qui n'offre pas une description suffisamment explicite des extensions de corps ainsi obtenues. Il explique qu'il avait, depuis l'hiver 1856, trouvé comment associer à certains corps de nombres une extension (un « genre associé », *associerte Gattung*) dont le degré est égal au nombre de classes, et telle que « toutes les propriétés profondes en rapport avec la composition et la répartition en classes du genre [...] se reflètent pour ainsi dire dans les propriétés élémentaires du genre associé » [43]. Pour mieux saisir la différence entre ces deux méthodes, soit $K = \mathbb{Q}(\sqrt{-23})$. On montre facilement que le groupe des classes est $\mathbb{Z}/3\mathbb{Z}$, et qu'il est engendré par la classe de $2u + \frac{1-\sqrt{-23}}{2}v$, qui est l'un des deux facteurs premiers de 2. La théorie hilbertienne du corps de classes consiste à déterminer une extension abélienne maximale non-ramifiée [44] L de K, qui peut d'ailleurs s'obtenir, dans ce cas particulier, au moyen de la théorie des fonctions elliptiques. Il s'agit du corps de décomposition L de l'équation $X^3 - X - 1 = 0$ sur K. Dans L, tous les diviseurs de K deviennent principaux (mais L peut très bien avoir de nouveaux diviseurs non principaux). Et $[L : K]$ est bien égal au nombre de classes de K, comme Kronecker l'affirme ($[L : K] = 3$). Cette extension, non-ramifiée, fournit d'ailleurs l'exemple d'un « genre qui n'a pas de discriminant », situation qui semblait susciter l'intérêt de Kronecker [45]. Démontrons-le directement. Soient R, T et S les anneaux d'entiers des corps $\mathbb{Q}$, K et L. On a alors ([52], p. 17, *tower formula*) :

$$\delta(S/R) = \delta(T/R)^{[L:K]} N_{K/\mathbb{Q}}\delta(S/T)$$

Or $T = \mathbb{Z}\left[1, \frac{1+i\sqrt{23}}{2}\right]$ et $\delta(T/R) = -23$. L'algorithme et la démonstration de Kronecker (§ 6 des *Grundzüge*) pour construire une base du Z-module O_L, ont pour conséquence immédiate que $\delta(S/R)$ est inférieur en valeur absolue au discriminant des grandeurs $1, x_2, x_1, x_1x_2, x_1^2, x_1^2x_2$, où x_1 et x_2 désignent deux des racines de $X^3 - X - 1$. Or ce discriminant de 6 grandeurs n'est autre que $\mathcal{D}^3$, où $\mathcal{D}$ est le discriminant de $X^3 - X - 1$, comme Kronecker l'affirme au § 12, et comme nous l'avons démontré plus haut. Le calcul donne donc :

$$|\mathcal{D}^3| = 23^3 \geq |\delta(S/R)| = 23^3 |N_{K/\mathbb{Q}}\delta(S/T)|$$

donc $|\delta(S/T)| = 1$, comme annoncé. Kronecker faisait certainement allusion à de tels exemples, tirés de la théorie des fonctions elliptiques. Sur le même exemple, l'autre méthode, critiquée par Kronecker, nous aurait conduit à calculer le diviseur

$$\left(2u + \frac{1-\sqrt{-23}}{2}v\right)^3 = \epsilon\frac{3+i\sqrt{23}}{2}$$

où ϵ est une unité, puis à adjoindre au corps K le radical $\sqrt[3]{\frac{3+i\sqrt{23}}{2}}$, rendant ainsi principal $2u + \frac{1-\sqrt{-23}}{2}$. Mais en procédant ainsi, on risque d'obtenir une extension ramifiée, et l'on est pas sûr d'obtenir une extension unique, indépendante du choix des radicaux adjoints au corps de base. L'histoire de la théorie du corps de classes a donc justifié la préférence de Kronecker pour l'autre méthode, bien que celui-ci n'ait pas pu, semble-t-il [46], s'élever des exemples tirés de la théorie des fonctions elliptiques à une théorie plus générale.

15.6 Systèmes de diviseurs et codimension

Kronecker est donc parvenu à faire valoir de nouveau la théorie classique de la divisibilité (plus grand commun diviseur, décomposition en facteurs premiers) en étendant l'anneau d'entiers d'une

43. *Cf.* [183], p. 67.

44. *Cf.* [52], exercice 3. Plus précisément, l'extension abélienne en question doit être non-ramifiée et « totalement décomposée en les valuations archimédiennes de K ». Dans le cas particulier ci-dessus, cela signifie qu'il y a exactement trois K-plongements de L dans $\mathbb{C}$ essentiellement distincts vis-à-vis de la topologie induite par la norme dans $\mathbb{C}$.

45. *cf.* [183], p. 22.

46. *Cf.* les remarques de H. Hasse dans [52], ainsi que les remarques 32 à 36 de Dedekind et le commentaire de Edwards, Neumann et Purkert dans [99].

extension algébrique par l'adjonction d'indéterminées. Au § 20, il remarque que cette « conservation des déterminations conceptuelles lors du passage du rationnel à l'algébrique » ne vaut pas, par contre, lors du passage d'un corps ne contenant pas de variable à un corps contenant des variables (c'est-à-dire une extension transcendante d'un corps de nombres). Bien que la théorie développée dans les sections § 14 à 17 s'applique aussi à de tels corps, Kronecker la juge à présent insuffisante. Le concept de diviseur devant exprimer « quelque chose de commun » (*Gemeinsames*) à deux ou plusieurs grandeurs du domaine, il avait en effet été défini sur le modèle du « plus grand commun diviseur », comme combinaison linéaire $x_1u_1 + ... + x_nu_n$ des grandeurs en question. Mais dans le cas où certaines des indéterminées de l'anneau d'entiers considéré sont conçues comme des *variables*, le plus grand commun diviseur ne suffit plus à exprimer tout le « commun » de n grandeurs. En effet, l'intuition géométrique fait correspondre au plus grand commun diviseur de n polynômes une hypersurface, sur laquelle les n polynômes s'annulent ensemble. Mais, en général, le lieu commun des zéros des n polynômes peut contenir d'autres variétés, de dimensions inférieures. Le plus grand commun diviseur ne saurait en rendre compte. C'est donc l'intuition géométrique qui présente d'abord à notre conscience un « commun » autre que le « plus grand commun diviseur », sous la forme d'une intersection, lieu des zéros communs à un ensemble de polynômes [47]. Le concept de diviseur étant donc insuffisant pour rendre compte de ce « commun », Kronecker définit à présent (§ 20–22) le concept de « système modulaire » (*Modulsystem*) [48] ou « système de diviseurs » (*Divisoren-System*). Un élément G est dit « contenir le système de diviseurs $(F_1, ..., F_n)$ » lorsque $G \equiv 0$ modulo $(F_1, ..., F_n)$. Mais comme pour échapper de nouveau à la notion ensembliste d'« idéal », Kronecker représente tout système de diviseurs par une forme à coefficients indéterminés (par exemple $F_1u_1 + ... + F_nu_n$). Contrairement au cas des « diviseurs », la relation « être contenu dans », pour les systèmes de diviseurs, *ne se traduit pas* par une relation de divisibilité.

Kronecker introduit au § 21 un concept de codimension (*Stufe*) sans le définir de manière rigoureuse, en ayant recours encore une fois à l'intuition géométrique. On a vu que la méthode d'élimination du § 10 permet de décomposer une variété affine, lieu des zéros d'un idéal de $k[x_1, ..., x_n]$ sur $\overline{k}$, en composantes irréductibles. On peut alors définir la dimension comme dimension maximale de ses composantes. Au § 20, il traite en détail le cas des variétés affines de dimension 0 (ensemble de points isolés) sous l'hypothèse que le discriminant est non nul ; il utilise à cet effet la théorie du résultant de n polynômes à $(n-1)$ indéterminées, sans d'ailleurs se soucier de savoir si la « résolvente » définie au § 10 coïncide ou non avec le « résultant » classique [49]. Dans ce cas particulier, la relation « être contenu dans » se traduit bien par une relation de divisibilité. Quant à l'origine de ces recherches, Kronecker renvoie à ses travaux de 1865 sur la formule d'interpolation de Lagrange. Nous avons déjà rencontré, au § 4, l'usage qu'il fait de cette formule pour factoriser les polynômes. Dans un article de 1865, *Über einige Interpolationsformeln für ganze Functionen mehrer Variabeln*, il généralise cette formule aux polynômes à plusieurs indéterminées. Soit un ensemble de m points $\xi_k = (\xi_{1k}, ..., \xi_{nk}), 1 \leq k \leq m$, solution d'un système d'équations non homogènes à n inconnues $F_1(x_1, ..., x_n) = 0, ..., F_n(x_1, ..., x_n) = 0$. On peut alors écrire (on le justifierait aujourd'hui au moyen du théorème des zéros de Hilbert) :

$$F_i = (x_1 - \xi_{1k})F_{1i}^{(k)} + ... + (x_n - \xi_{nk})F_{ni}^{(k)}$$

Les déterminants $D_k(x_1, ..., x_n) = \left|F_{hi}^{(k)}\right|_{1 \leq h,i \leq n}$ permettent alors d'écrire la formule d'interpolation suivante qui définit une fonction $\mathfrak{F}(x_1, ..., x_n)$ prenant la valeur $\mathfrak{F}_k$ au point ξ_k :

$$\mathfrak{F}(x_1, ..., x_n) = \sum_k \mathfrak{F}_k \frac{D_k(x_1, ..., x_n)}{D_k(\xi_{1k}, ..., \xi_{nk})}$$

47. L'opération intellectuelle par laquelle ce « quelque chose de commun », concept d'abord vide de tout substrat mathématique, pourra légitimement subsumer le concept de plus grand commun diviseur et le concept d'intersection, n'est-il pas un bel exemple de synthèse (au sens kantien), renversant l'image d'un Kronecker conservateur obsédé par une vision arithmétisante de la réalité mathématique ?

48. Ce terme est devenu *modular system* chez Macaulay [217].

49. Macaulay a donné un contre-exemple, *cf.* [217], p. 21.

De plus, en remarquant qu'au point ξ_k

$$F_{hi}^{(k)} = \frac{\partial F_i}{\partial x_h},$$

on a :

$$D_k(\xi_{1k}, ..., \xi_{nk}) = J(\xi_{1k}, ..., \xi_{nk})$$

où J est le déterminant fonctionnel $\left|\frac{\partial F_i}{\partial x_h}\right|_{1\leq h,i\leq n}$. En termes modernes, la formule d'interpolation de Kronecker décrit l'anneau de fonctions d'une variété affine de dimension zéro, ensemble de m points, comme étant de la forme k^m (où k est le corps des constantes), en associant à chaque fonction $\mathfrak{F}$ ses valeurs en ces points $(\mathfrak{F}_1, ..., \mathfrak{F}_m) \in k^m$. Dans le même article, Kronecker déduit de sa formule une démonstration d'un théorème de Jacobi : si $\mathfrak{F}$ est de degré $< \deg J$, on a

$$\sum \frac{\mathfrak{F}(\xi_{1k}, ..., \xi_{nk})}{J(\xi_{1k}, ..., \xi_{nk})} = 0.$$

La démonstration semble néanmoins incomplète : Kronecker remarque que pour un tel $\mathfrak{F}$, la forme dominante de sa formule d'interpolation est (à un facteur constant près égal au produit des degrés des F_i)

$$R(x_1, .., x_n) \sum \frac{\mathfrak{F}(\xi_{1k}, ..., \xi_{nk})}{J(\xi_{1k}, ..., \xi_{nk})}$$

où R est le jacobien des formes dominantes $f_1, ..., f_n$ de $F_1, ..., F_n$, soit $\left|\frac{\partial f_i}{\partial x_h}\right|_{1\leq h,i\leq n}$. Il en conclut que

$$R(x_1, .., x_n) \sum \frac{\mathfrak{F}(\xi_{1k}, ..., \xi_{nk})}{J(\xi_{1k}, ..., \xi_{nk})} \in (f_1, ..., f_n),$$

puis finalement que

$$R(x_1, .., x_n) \sum \frac{\mathfrak{F}(\xi_{1k}, ..., \xi_{nk})}{J(\xi_{1k}, ..., \xi_{nk})} = 0.$$

Il ne justifie pas les deux dernières étapes de son raisonnement, bien qu'elles soient valables en général. En utilisant le formalisme de J.-P. Jouanolou, il suffit en effet de remarquer que l'idéal $(F_1, ..., F_n)$ n'a pas de forme d'inertie de degré $\deg J$, et que l'idéal $(f_1, ..., f_n)$ a exactement une forme d'inertie de degré $\deg J$ non triviale, le jacobien R.

Suivons à nouveau le fil des *Grundzüge*. Au § 22, Kronecker expose une théorie des formes représentant les systèmes de diviseurs. Soit $A = k[x_1, ..., x_{n-1}]$. Soit $S \subset A[u_1, u_2, ...]$ l'ensemble (multiplicatif) des formes « proprement primitives », c'est-à-dire dont les coefficients engendrent l'idéal (1) de l'anneau A.

Lemme 5 *Si les formes E et F sont proprement primitives alors la forme EF l'est aussi, et réciproquement.*

Démonstration. Supposons EF proprement primitive. Son idéal n'a donc pas de zéro dans $\overline{k}$. Or le produit des idéaux engendrés par les coefficients de E et de F contient l'idéal engendré par les coefficients de EF. Donc ce produit d'idéaux n'a pas de zéro non plus, donc E et F sont proprement primitives (théorème des zéros de Hilbert). Réciproquement, supposons E et F proprement primitives. Soit z un zéro de l'idéal engendré par les coefficients de EF. Alors $EF(z) \equiv 0$, donc soit $E(z) \equiv 0$, soit $F(z) \equiv 0$. Absurde. Donc EF est proprement primitive. *q. e. d.*

Kronecker travaille dans l'anneau $S^{-1}A[u_1, u_2, ...]$. Une forme (en $u_1, u_2, ...$) est dite de codimension m quand la variété des zéros (dans $\overline{k}$) de ses coefficients (éléments de A) est de codimension m. Alors $m \leq n$, et $m = n$ si et seulement si la variété des zéros est vide, c'est-à-dire si la forme appartient à S (Kronecker a ici recours à un cas du théorème des zéros de Hilbert). La forme F_0 « contient » la forme F si et seulement si l'idéal des coefficients de F_0 est inclus dans l'idéal des coefficients de F. Mais cela *ne se traduit pas* par une relation de divisibilité. Et cette relation « être contenu dans » donne lieu à une relation d'équivalence entre formes, qui ne se traduit pas non

plus par l'égalité des formes dans $S^{-1}A[u_1, u_2, ...]$. Le « plus grand commun contenu » (*grössten gemeinsamen Inhalt*) de F et F_0 est la forme $uF + F_0$ (où u est une indéterminée n'entrant pas dans l'expression de F et F_0). En termes de variétés, il s'agit de l'intersection des deux variétés.

De même que l'adjonction d'indéterminées avait permis à Kronecker, aux § 14–18, de transformer un anneau d'entiers en un anneau principal, elle lui permet, aux § 20–22, de transformer A en un anneau $S^{-1}A[u_1, u_2, ...]$, faisant de toute variété, de dimension donnée, une intersection complète. Afin de donner une formulation algébrique de la relation entre formes « être contenu dans » (qui n'est pas une relation de divisibilité), Kronecker énonce en effet le résultat suivant, pour deux « formes pures » de codimension m (*reine Formen m^{ter} Stufe*) F_0, F_1 :

> Soient $F_1, F_2, ..., F_m$ des formes algébriques entières, qui ont chacune les mêmes coefficients et diffèrent donc entre elles seulement quant au système des indéterminées, alors l'une de ces formes est contenue dans une forme F_0 si F_0 est absolument équivalente [c'est-à-dire égale, à un élément de S près] à une fonction linéaire homogène entière des m formes $F_1, F_2, ..., F_m$.[50]

Il nous est difficile de juger cet énoncé, mais nous pouvons au moins démontrer l'énoncé suivant, un peu moins fort :

Théorème 5 *Soient F_1,F_2,...,F_m des formes à coefficients dans A, qui ont chacune les mêmes coefficients et diffèrent donc entre elles seulement quant au système des indéterminées. Soit $I \subset A$ l'idéal engendré par les coefficients. Alors prenant un nombre fini suffisant d'indéterminées $u_1, u_2, ...$ (comprenant les indéterminées des formes $F_1, ..., F_m$), et notant S l'ensemble multiplicatif des formes de codimension strictement supérieure à m, et B l'anneau $S^{-1}A[u_1, u_2, ...]$, on aura dans B :*

$$(F_1, F_2, ..., F_m) \subset I \subset \sqrt{(F_1, F_2, ..., F_m)}$$

Autrement dit, « géométriquement », I est un idéal d'intersection complète dans B.

Démonstration. On utilise la méthode employée précédemment (*cf. supra* p. 241) et on écrit, pour $1 \leq i \leq m$,

$$V((F_1, F_2, ..., F_i)) = \bigcup_j Y_j^{(i)} \cup \bigcup_k Z_k^{(i)}$$

où les $Y_j^{(i)}$ sont les composantes irréductibles de $V((F_1, F_2, ..., F_i))$ dont la fibre en $u_1, u_2, ...$ est constante, et où les $Z_k^{(i)}$ sont celles dont la fibre est variable. Si on spécialise $u_1, u_2, ...$ en des valeurs suffisamment générales de k, on voit facilement que :

$$1 = \operatorname{codim} \bigcup_k Z_k^{(1)} = \operatorname{codim} \bigcup_k Z_k^{(2)} - 1 = ... = \operatorname{codim} \bigcup_k Z_k^{(m)} - m + 1$$

Autrement dit, en général, la fibre des $Z_k^{(m)}$ en $u_1, u_2, ...$ est de codimension m. Il est, certes, possible que $\bigcup_k Z_k^{(m)}$ contienne une sous-variété dont la fibre en $u_1, u_2, ...$ soit constante, mais une telle sous-variété serait alors de codimension strictement supérieure à m. Exprimons ce fait algébriquement. Dans $A[u_1, u_2, ...]$, notons J l'idéal de $\bigcup_k Z_k^{(m)}$. Soient $N_1, N_2, ...$ des générateurs de J (ce sont des formes à coefficients dans A), et $v_1, v_2, ...$ des indéterminées, choisies parmi $u_1, u_2, ...$, mais qui n'apparaissent pas dans $N_1, N_2, ...$ Soit $e = v_1N_1 + v_2N_2 + ...$; alors $e \in S$. Or, $IJ \subset \sqrt{(F_1, F_2, ..., F_m)}$ (théorème des zéros de Hilbert). En particulier, $eI \subset \sqrt{(F_1, F_2, ..., F_m)}$, et, dans $S^{-1}A[u_1, u_2, ...]$:

$$I \subset \sqrt{(F_1, F_2, ..., F_m)}$$

q. e. d.

Kronecker ajoute :

50. *cf.* [183], § 22, p. 91.

> Ceci peut servir de définition de la codimension (*Stufenzahl*) m d'une forme pure F, si l'on ajoute seulement qu'aucun nombre plus petit que m ne peut servir à une telle représentation. [51]

Une définition algébrique de la codimension est bien ce qui manquait à Kronecker, pour pouvoir étendre sa théorie des « systèmes de diviseurs » à des anneaux A plus généraux. La théorie de l'élimination sur laquelle il fonde, au § 10, la notion de *Stufe* se rapporte exclusivement à des anneaux de polynômes *sur un corps*, et à l'intuition géométrique sous-jacente des variétés affines. Il faudrait faire une théorie de l'élimination sur $\mathbb{Z}$ pour dépasser ce cadre. Kronecker semblait convaincu qu'il était possible de franchir le pas. Au § 21, il décrit à quoi ressemblerait la théorie de la dimension dans le cas $A = \mathbb{Z}[x_1, x_2, x_3]$:

> Prenons par exemple $n = 4$ et pensons les trois variables comme des coordonnées quelconques de l'espace, alors les diviseurs de codimension 1 sont soit des nombres soit des fonctions entières des coordonnées, dont l'annulation représente des surfaces. Parmi les systèmes de diviseurs de codimension 2, il s'en trouve dont les éléments ne peuvent tous s'annuler simultanément, et dont l'un des éléments peut être choisi comme nombre, mais il s'en trouve aussi, dont l'annulation simultanée des éléments représente une courbe ; parmi les systèmes modulaires de codimension 3, l'on en trouve qui représente de même des systèmes de points... [52]

Il indique même que la théorie des systèmes de diviseurs de codimension 2 dans $\mathbb{Z}[x]$ produit une théorie des nombres algébriques, la décomposition du système de diviseurs $(F(x), p)$ correspondant à la décomposition de p en facteurs premiers dans l'anneau d'entiers algébriques sur $\mathbb{Q}$ de $\mathbb{Q}[X]/F(X)$ (sauf pour un ensemble fini de valeurs de p). En fait, il faut attendre la théorie des schémas de Grothendieck pour avoir, comme le souhaitait Kronecker, une théorie des variétés (les « schémas ») sur $\mathbb{Z}$, ou un anneau quelconque.

Quant à la « grandeur entière », il y avait en tout cas une lacune dans les *Grundzüge* : bien que le concept de grandeur entière et celui de plus grand commun diviseur aient subi avec succès le passage à des extensions algébriques quelconques de $\mathbb{Q}$, voire de $\mathbb{Q}(T)$, dans les sections § 14 à 18, et bien qu'au concept de plus grand commun diviseur ait été substitué celui d'intersection dans les sections § 20 à 22, effectuant ainsi le passage d'un corps de nombres k à des extensions purement transcendantes de dimension quelconque $k(x_1, ..., x_{n-1})$ au moyen d'une théorie de l'élimination dans $k[x_1, ..., x_{n-1}]$ qui rend compte des variétés de codimension supérieure à 1, il restait à faire une théorie de la « grandeur entière » dans les extensions algébriques de $k(x_1, ..., x_{n-1})$, qui tienne compte elle aussi des variétés de codimension supérieure à 1. Qu'il nous soit permis d'employer un vocabulaire moderne : il fallait en quelque sorte définir et étudier des « morphismes finis » entre variétés affines.

Un court article [184] publié en 1883 témoigne encore du caractère non achevé des *Grundzüge* et montre que Kronecker était à la recherche d'une autre formulation algébrique de la relation entre systèmes de diviseurs « être contenu dans ». Il y démontre le théorème suivant : [53]

> Soient
>
> $$M_0, M_1, M_2, ..., M_{n+1}$$
>
> des grandeurs entières d'un domaine de rationalité $[\mathfrak{R}', \mathfrak{R}'', \mathfrak{R}''', ...]$, alors l'équation :
>
> $$\sum_h M_h U^h . \sum_i M_{m+i} U^{i-1} = \sum_k M'_k U^k \quad \left(\begin{array}{l} h = 0, 1, ...m \\ i = 1, 2, ...n - m + 1 \\ k = 0, 1, ...n \end{array} \right),$$
>
> où U est une indéterminée, définit $n + 1$ grandeurs $M'_0, M'_1, ...M'_n$ du même domaine $[\mathfrak{R}', \mathfrak{R}'', \mathfrak{R}''', ...]$.

51. *Cf.* [183], p. 92.
52. *Cf.* [183], § 21.
53. *cf.* [184], p. 419–420. Nous avons corrigé une faute dans la dernière formule de cette citation (le texte des *Werke* contient r au lieu de $\rho - r$).

(...)
Le produit de formes :

$$\sum_h M_h U_h . \sum_i M_{m+i} U_{m+i} \quad \left(\begin{array}{l} h = 0, 1, ...m \\ i = 1, 2, ...n - m + 1 \end{array} \right)$$

vérifie une équation algébrique de degré ρ, dans laquelle le coefficient de la puissance $\rho^{\text{ième}}$ est égal à 1, mais celui de la puissance $r^{\text{ième}}$ *est une forme contenant le produit de formes*

$$\prod_h \sum_k M'_k V_{hk} \quad \left(\begin{array}{l} h = 1, 2, ...\rho - r \\ k = 0, 1, ...n \end{array} \right)$$

La relation « être contenu dans » qu'utilise Kronecker dans l'énoncé de ce théorème est simplement la relation d'inclusion entre les idéaux engendrés par les coefficients des formes. Autrement dit, le coefficient de la puissance $r^{\text{ième}}$ $(0 \leq r \leq \rho)$ appartient à l'idéal $(M'_0, M'_1, ..., M'_n)^{\rho - r}$. A la suite de ce théorème, Kronecker définit une nouvelle relation « être contenu dans » (plus faible que l'inclusion entre idéaux, mais qui implique au moins l'inclusion entre les radicaux des idéaux). Il explique ensuite, dans un paragraphe quelque peu obscur, que cette nouvelle relation permet de définir un nouveau concept de « grandeur entière » [54]. Cet article, son titre l'indique, se rapportait principalement aux variétés de codimension supérieure à 1. Sans recours à l'historiographie, il nous serait impossible de comprendre ce que Kronecker visait.

J. Molk, dans un long commentaire en français des *Grundzüge*, nous offre une explication possible quand il se propose de décomposer les systèmes de diviseurs en produits de systèmes irréductibles [55], pour $A = k[x, y]$. A cet effet, dans le cas d'un système de diviseurs $(\Phi(x, y), \Psi(x, y))$, il montre que ce système est équivalent au système des coefficients de la résolvente de Φ et Ψ (*cf.* [233] p. 79-97). Il en est de même en général pour un système de diviseurs $(\Lambda_1(x, y), \Lambda_2(x, y), ..., \Lambda_\mu(x, y))$, *à condition* d'affaiblir la relation d'équivalence entre systèmes de diviseurs comme le faisait Kronecker en 1883 (*cf.* [233] p. 104-105). Malgré ses efforts, Molk semble abandonner tout espoir d'une théorie générale des systèmes de diviseurs qui rende compte des facteurs multiples dans la résolvente. Or la relation d'équivalence entre systèmes de diviseurs consistant en l'égalité des radicaux des idéaux donne une théorie générale sans notion de multiplicité ; mais Kronecker et Molk cherchaient une relation d'équivalence plus forte. La recherche d'une telle relation d'équivalence était probablement motivée par la recherche d'une notion de multiplicité, et renoncer à toute notion de multiplicité était donc un constat d'échec.

D'autres commentateurs des *Grundzüge* (Hurwitz, H. M. Edwards) ont vu dans l'article de 1883 le chaînon manquant qui permet au moins de démontrer rigoureusement les considérations des sections 14 à 18. Il y avait en effet, comme nous l'avons mentionné, un point obscur dans la section 14, qui a fait dire à Dedekind :

> Ici se cache le cœur de la théorie des idéaux de Kronecker ; la simplicité qui nous étonnait au début apparaît bientôt sous une autre lumière, si l'on exige que la démonstration soit complètement achevée. [56].

Kronecker semble en effet déduire, du fait qu'une forme Q vérifie une équation $\mathrm{Nm}(X - Q) = 0$ à coefficients entiers sur le corps de base, le fait que cette forme soit elle-même à coefficients entiers. Dans le théorème de l'article de 1883, si l'on spécialise les indéterminées U_i $(0 \leq i \leq n)$, on obtient facilement le corollaire suivant :

Corollaire 3 *Les grandeurs $M_k M_{m+i}$ (pour tout k, i tels que $0 \leq k \leq m$ et $1 \leq i \leq n - m + 1$) sont entières sur $\mathbb{Z}[M'_0, M'_1, ..., M'_n]$.*

54. Comme nous l'avons mentionné, Kronecker souhaitait réduire la théorie des nombres algébriques à celle des systèmes de diviseurs de rang 2 dans $\mathbb{Z}[x]$, et plus généralement toute la théorie des « formes entières algébriques de rang $\mathfrak{m}$ » à une théorie des « formes entières rationnelles de rang $\mathfrak{m} + 1$ » (*cf.* [183] p. 113).

55. *Cf.* [233] § IV.1. Il faut ici entendre la notion de produit, avec Molk, au sens du produit des idéaux correspondant aux systèmes de diviseurs. Molk affirme (cf. [233] p. 107) que sa méthode, décrite pour $A = k[x, y]$, est aussi valable pour $A = \mathbb{Z}[x]$.

56. *Cf.* [99], remarque n°20 de Dedekind, et le commentaire des éditeurs que nous résumons ci-dessous.

Ce corollaire se généralise facilement au produit de formes à un nombre quelconque d'indéterminées :

$$\sum_h M_h \phi_h . \sum_i M_{m+i} \psi_i = \sum_k M'_k \chi_k$$

où ϕ_h, ψ_i, χ_k désignent des produits quelconques des indéterminées [57]. En prenant $\sum_h M_h \phi_h = X - Q$ et $\sum_i M_{m+i} \psi_i = \frac{\mathrm{Nm}(X-Q)}{X-Q}$, on en déduit la proposition annoncée.

Reprenons le fil des *Grundzüge*. Kronecker entendait réduire la théorie des « formes entières algébriques de rang $\mathfrak{m}$ » à une théorie des « formes entières rationnelles de rang $\mathfrak{m} + \mathfrak{1}$ ». Dans la vingt-cinquième et dernière section des *Grundzüge*, il tente d'appliquer ce principe pour $\mathfrak{m} = 1$; il donne ainsi un nouvel algorithme (p. 114–117), déjà annoncé ailleurs dans les *Grundzüge*, pour la décomposition d'une grandeur algébrique entière en ses diviseurs irréductibles. Dans le cas $O_k = \mathbb{Z}$, l'idée de cette méthode n'est pas nouvelle. Plus généralement, soit

$$\begin{aligned} O_k &= \mathbb{Q}[\mathfrak{R}', \mathfrak{R}'', ...] \\ k &= \mathbb{Q}(\mathfrak{R}', \mathfrak{R}'', ...) \\ K &= \mathbb{Q}(\mathfrak{G}, \mathfrak{R}', \mathfrak{R}'', ...) \end{aligned}$$

où $\mathfrak{G}$ vérifie une équation [58]

$$\mathfrak{F}(\mathfrak{G}) = 0$$

avec $\mathfrak{F}(\mathfrak{R}) \in \mathbb{Q}[\mathfrak{R}', \mathfrak{R}'', ...][\mathfrak{R}]$.

Ici Kronecker ne parvient pas à exprimer une base agréable de O_K avec $\mathfrak{G}$ et ses puissances (p. 111-112). En tout cas, O_K contient l'anneau engendré par $\mathfrak{G}, \mathfrak{R}', \mathfrak{R}'', ...$:

$$\mathbb{Q}[\mathfrak{G}, \mathfrak{R}', \mathfrak{R}'', ...] \subset O_K$$

Nous allons préciser cette affirmation. Remarquons que K, anneau engendré par $\mathfrak{G}$ sur k, est isomorphe à un anneau quotient :

$$K \simeq k[\mathfrak{R}]/\mathfrak{F}(\mathfrak{R})$$

Soit Δ le déterminant de $\mathfrak{F}(\mathfrak{R})$. Alors $\Delta \in O_k$. On va montrer que

$$O_K \subset O_{k,(P)}[\mathfrak{G}]$$

où l'anneau de droite est engendré par $\mathfrak{G}$ sur $O_{k,(P)}$, localisé de O_k par rapport à n'importe quel idéal premier (P) avec P ne divisant pas Δ. Soit en effet $x \in O_K$, alors il existe des $a_i \in k$ tels que :

$$x = \sum_{i=0}^{\deg \mathfrak{F} - 1} a_i \mathfrak{G}^i$$

Les $(\deg \mathfrak{F})$ équations conjuguées forment un système d'équations linéaires en les a_i, dont le déterminant est un déterminant de Vandermonde ayant pour carré Δ. Les formules de Cramer permettent d'exprimer chaque a_i comme quotient d'un élément de $O_k[\mathfrak{G}]$ par ce déterminant. Si $P \in O_k$ ne divise pas Δ, il est donc clair que chaque a_i appartient à l'anneau $O_{k,(P)}[\mathfrak{G}]$. *q. e. d.*

Soit une telle grandeur entière P irréductible dans O_k et ne divisant pas Δ. Essayons de la décomposer en produit de diviseurs (rappelons, conformément à la théorie des diviseurs exposée plus haut, qu'il s'agit de décomposer P dans $O_K \otimes S^{-1} O_k[u_1, u_2, ...]$). Remarquons tout d'abord

57. *Cf.* par exemple [98], partie 0, corollaire 3 p. 4.

58. Nous ne décrivons ici que le cas où K est une extension algébrique de k ayant un élément primitif $\mathfrak{G}$. Mais on peut toujours se ramener à cette situation en adjoignant des indéterminées $u', u'', ...$; on aurait alors $K = \mathbb{Q}(u', u'', ...)(\mathfrak{G}, \mathfrak{R}', \mathfrak{R}'', ...)$ et cela ne nuit pas à la suite du raisonnement. Selon Edwards, c'est Hensel qui a démontré, dans le cas particulier $k = \mathbb{Q}$, la validité de l'algorithme décrit ici par Kronecker. Edwards avoue « *I do not know of a proof of Kronecker's more general case. (If one could be given, it would be a large step toward the solution of the problem of factoring divisors in the general case.).* ». Nous essayons ici de reconstituer le raisonnement, imparfait, de Kronecker dans le cas où $k = \mathbb{Q}(\mathfrak{R}', \mathfrak{R}'', ...)$ et $O_k = \mathbb{Q}[\mathfrak{R}', \mathfrak{R}'', ...]$.

que l'on a bien une théorie de la codimension pour les systèmes de diviseurs dans $\mathbb{Q}[\mathfrak{R}', \mathfrak{R}'', ...][\mathfrak{R}]$, théorie garantie par la méthode d'élimination de Kronecker et son interprétation géométrique [59]. Or pour décomposer (P) en ses composantes de codimension 1 (c'est-à-dire en éléments de O_K), on va décomposer $(P, \mathfrak{F}(\mathfrak{R}))$ en ses composantes de codimension 2 (c'est-à-dire que l'on cherche un produit d'idéaux qui, en codimension 2, présente le même lieu des zéros que $(P, \mathfrak{F}(\mathfrak{R}))$). A cet effet, on a naturellement recours à une décomposition de $\mathfrak{F}$:

$$\mathfrak{F}(\mathfrak{R}) = f_1(\mathfrak{R})^{n_1} f_2(\mathfrak{R})^{n_2} ... + P\overline{\mathfrak{F}}(\mathfrak{R})$$

avec $f_1(\mathfrak{R}), f_2(\mathfrak{R}), ...$ irréductibles modulo P (cela a bien un sens puisque l'anneau $O_{k,(P)}/(P)$ est un corps et que l'anneau de polynômes $\left(O_{k,(P)}/(P)\right)[\mathfrak{R}]$ est donc factoriel). D'ailleurs, comme le discriminant Δ de $\mathfrak{F}$ est non nul modulo P, une telle décomposition ne compte pas de facteur double et $n_1 = n_2 = ... = 1$.

Remarque : une telle décomposition de $\mathfrak{F}$ donne bien une décomposition « géométrique » de $(P, \mathfrak{F}(\mathfrak{R}))$, au sens suivant. Mettons tous les fractions rationnelles au même dénominateur :

$$f_1(\mathfrak{R}) = \frac{\tilde{f}_1(\mathfrak{R})}{c(\mathfrak{R}', \mathfrak{R}'', ...)}, \quad f_2(\mathfrak{R}) = \frac{\tilde{f}_2(\mathfrak{R})}{c(\mathfrak{R}', \mathfrak{R}'', ...)}, ... \quad \overline{\mathfrak{F}}(\mathfrak{R}) = \frac{\tilde{\mathfrak{F}}(\mathfrak{R})}{c(\mathfrak{R}', \mathfrak{R}'', ...)}$$

avec $c(\mathfrak{R}', \mathfrak{R}'', ...)$ non divisible par P, de sorte que

$$c(\mathfrak{R}', \mathfrak{R}'', ...)\mathfrak{F}(\mathfrak{R}) = \tilde{f}_1(\mathfrak{R})\tilde{f}_2(\mathfrak{R}) ... + P\tilde{\mathfrak{F}}(\mathfrak{R})$$

Les lieux géométriques suivants sont alors confondus :

$$\begin{aligned} Z\left(\prod_i \left(P, \tilde{f}_i(\mathfrak{R})\right)\right) &= Z\left(P, \prod_i \tilde{f}_i(\mathfrak{R})\right) \\ &= Z(P, c(\mathfrak{R}', \mathfrak{R}'', ...)\mathfrak{F}(\mathfrak{R})) \\ &= Z\left((P, c(\mathfrak{R}', \mathfrak{R}'', ...))(P, \mathfrak{F})\right) \end{aligned}$$

D'où

$$\bigcup_i Z\left(P, \tilde{f}_i(\mathfrak{R})\right) = Z\left(P, c(\mathfrak{R}', \mathfrak{R}'', ...)\right) \cup Z\left(P, \mathfrak{F}\right)$$

Revenons à la théorie des diviseurs. La décomposition de $\mathfrak{F}(\mathfrak{R})$ donne, en posant $\mathfrak{R} = \mathfrak{G}$:

$$\tilde{f}_1(\mathfrak{G})\tilde{f}_2(\mathfrak{G}) ... + P\tilde{\mathfrak{F}}(\mathfrak{G}) = 0$$

Donc dans O_K, P divise $\prod \tilde{f}_i(\mathfrak{G})$. Donc dans $O_K \otimes S^{-1}O_k[u_1, u_2, ...]$, il divise

$$\prod \left(P + u_i \tilde{f}_i(\mathfrak{G})\right)$$

Il est clair d'autre part que $P + u_i\tilde{f}_i(\mathfrak{G})$ divise P pour tout i.

Lemme 6 *Pour tous $i \neq j$, le diviseur $P + u_i\tilde{f}_i(\mathfrak{G}) + u_j\tilde{f}_j(\mathfrak{G})$ est une unité.*

Démonstration : on écrit l'égalité de Bézout suivante dans l'anneau $\left(O_{k,(P)}/(P)\right)[\mathfrak{R}]$:

$$A(\mathfrak{R})f_i(\mathfrak{R}) + B(\mathfrak{R})f_j(\mathfrak{R}) = C$$

Et $C \neq 0$ car f_i et f_j sont premiers entre eux. On relève cette égalité dans l'anneau $O_{k,(P)}[\mathfrak{R}]$:

$$\tilde{A}(\mathfrak{R})\tilde{f}_i(\mathfrak{R}) + \tilde{B}(\mathfrak{R})\tilde{f}_j(\mathfrak{R}) = \tilde{C}$$

avec $\tilde{C}$ non divisible par P. On pose dans cette égalité $\mathfrak{R} = \mathfrak{G}$; comme le diviseur $u_i\tilde{f}_i(\mathfrak{G}) + u_j\tilde{f}_j(\mathfrak{G})$ divise $\tilde{f}_i(\mathfrak{G})$ et $\tilde{f}_j(\mathfrak{G})$, il divise $\tilde{C}$ aussi. Donc le diviseur $u_i\tilde{f}_i(\mathfrak{G}) + u_j\tilde{f}_j(\mathfrak{G}) + P$ divise le diviseur $P + u\tilde{C}$ qui lui-même est une unité, donc celui-là en est une aussi, *q. e. d.*

Ce lemme montre que les $P + u_i\tilde{f}_i(\mathfrak{G})$ sont tous premiers entre eux. Mais donc comme chacun divise P, leur produit aussi. Finalement les diviseurs P et $\prod \left(P + u_i\tilde{f}_i(\mathfrak{G})\right)$ se divisent mutuellement et sont égaux, à une unité près. On a donc trouvé une décomposition de P. Il reste à vérifier que les diviseurs $P + u_i\tilde{f}_i(\mathfrak{G})$ sont irréductibles.

59. On serait gêné par l'absence d'une théorie de la codimension dans le cas $O_k = \mathbb{Z}[\mathfrak{R}', \mathfrak{R}'', ...]$.

Lemme 7 *Pour tout i, le diviseur* $P + u_i\tilde{f}_i(\mathfrak{G})$ *est irréductible.*

Démonstration : supposons qu'il existe $f(\mathfrak{G}) \in O_K$ tel que $P + u_i\tilde{f}_i(\mathfrak{G}) + uf(\mathfrak{G})$ soit différent d'une unité, et montrons qu'alors $P + u_i\tilde{f}_i(\mathfrak{G})$ divise $f(\mathfrak{G})$. Ecrivons une égalité de Bézout dans $\big(O_{k,(P)}/(P)\big)[\mathfrak{R}]$ que l'on relève immédiatement à $O_{k,(P)}[\mathfrak{R}]$:

$$A(\mathfrak{R})\tilde{f}_i(\mathfrak{R}) + B(\mathfrak{R})f(\mathfrak{R}) = C + P.D(\mathfrak{R})$$

Quitte à réduire au même dénominateur $A(\mathfrak{R})$, $B(\mathfrak{R})$, C et $D(\mathfrak{R})$, on peut même supposer qu'ils sont tous éléments de $O_k[\mathfrak{R}]$. Posons $\mathfrak{R} = \mathfrak{G}$, on obtient une égalité de la forme :

$$A(\mathfrak{G})\tilde{f}_i(\mathfrak{G}) + B(\mathfrak{G})f(\mathfrak{G}) = C + P.D(\mathfrak{G})$$

avec $A(\mathfrak{G})$, $B(\mathfrak{G})$, C et $D(\mathfrak{G})$ éléments de O_K. On en déduit que le diviseur $u_i\tilde{f}_i(\mathfrak{G}) + uf(\mathfrak{G}) + P$ divise le diviseur $C + vP$. Ce dernier ne peut donc pas être une unité, et if faut que P divise C. Alors $f_i(\mathfrak{R})$ et $f(\mathfrak{R})$ ont un facteur commun modulo P (c'est-à-dire si l'on considère ces deux polynômes comme éléments de l'anneau $\big(O_{k,(P)}/(P)\big)[\mathfrak{R}]$). Mais f_i est irréductible modulo P, donc il divise f modulo P. Cette relation de divisibilité, relevée à $O_{k,(P)}[\mathfrak{R}]$ et quitte à réduire toutes les fractions au même dénominateur, implique que dans $O_k[\mathfrak{R}]$, il existe un $C \in O_k$ non divisible par P tel que

$$C.f(\mathfrak{R}) \in (\tilde{f}_i(\mathfrak{R}), P)$$

D'où, en posant $\mathfrak{R} = \mathfrak{G}$,

$$P + u_i\tilde{f}_i(\mathfrak{G}) \quad | \quad C.f(\mathfrak{G})$$

et comme C n'est pas divisible par P, on a finalement

$$P + u_i\tilde{f}_i(\mathfrak{G}) \quad | \quad f(\mathfrak{G})$$

Q. e. d.

Retenons seulement que le texte des *Grundzüge* est de nature programmatique. Kronecker n'hésite pas à sauter les points obscurs et les détails à l'attention des commentateurs et de la recherche future, dans un ouvrage qui est le fruit d'une longue recherche menée sur des cas particuliers, champs d'essai des algorithmes, un ouvrage qui n'est pas une démonstration *a priori* de résultats pressentis. Il n'a pourtant pas négligé de chercher les fondements simples de sa théorie (*einfachste Grundlagen*, comme il le dit dans l'article de 1883) ; il définit certes des concepts (ceux de « genre », de « grandeur entière », de discriminant, de « système de diviseurs », de codimension,...). Mais il ne semble pas vouloir en donner une construction logique suffisante. L'ouvrage s'organise alors autour d'algorithmes qui mettent en œuvre ces concepts et en donnent une certaine légitimation (l'algorithme de factorisation au § 4, l'algorithme de calcul d'une base du O_k-module O_K au § 6, l'algorithme pour le calcul de la *Gesammtresolvente* au § 10, l'algorithme de décomposition en diviseurs premiers aux § 18 et 25...), et autour de cas particuliers (celui du corps de décomposition d'une équation aux § 11 et 12, celui des corps de nombres aux § 14 à 19, celui des variétés de dimension 0 au § 20...) et de l'évocation de travaux antérieurs (par exemple ses recherches sur le corps de classes et la multiplication complexe au § 19, ou celles sur les extensions non ramifiées – les « genres sans discriminant » – au § 8), cas particuliers qui entraînent peu à peu l'intuition vers son nouveau vêtement.

Chapitre 16

Elimination et géométrie dans la seconde moitié du XIX^{e} siècle

16.1 Cayley, Brill : le résultant réduit

Quand l'anneau des coefficients des polynômes U, V, W,... est factoriel, il est possible de factoriser le résultant $\mathrm{Res}(U, V, W, ...)$. Supposons que l'anneau des coefficients soit engendré sur un corps par un certain nombre d'indéterminées. Cayley introduit dans [61] en 1847 la notion de résultant réduit (*reduced resultant*), notion relative à une certaine interprétation géométrique de l'anneau des coefficients. Cayley distingue, parmi les indéterminées dans l'anneau des coefficients, des indéterminées constantes et des indéterminées variables. Les facteurs du résultant ne contenant que des constantes sont alors rejetés en tant que superflus, et les autres facteurs constituent le résultant réduit. Mais cette dénomination ne sera pas conservée par Brill [42], qui utilisera le terme de résultant réduit dans un sens différent, se rapportant à un problème plus spécifique : le résultant réduit de Brill est la condition pour que n équations à $(n-1)$ inconnues (équations dont les coefficients doivent contenir certains paramètres variables) ayant déjà un certain nombre de solutions en aient au moins une de plus. Cayley fait d'ailleurs un usage fâcheux du terme en 1847, ce qui explique peut-être qu'il n'ait pas gardé son sens originel : un des exemples qu'il propose pour illustrer cette notion est en effet erroné (bien que cela soit sans conséquence sur le reste de l'article). Cayley écrit que l'équation de la polaire réciproque $\mathrm{F}U = 0$ d'une courbe projective plane $U = 0$ est le résultant réduit du système :

$$\begin{cases} U = 0 \\ \alpha\frac{\partial U}{\partial x} + \beta\frac{\partial U}{\partial y} + \gamma\frac{\partial U}{\partial z} = 0 \\ x\mathrm{x} + y\mathrm{y} + z\mathrm{z} = 0 \end{cases}$$

où x, y, z sont les inconnues, α, β, γ sont vues comme constantes et x,y,z comme variables. Cette assertion est évidemment fausse, quelle que soit la manière dont on interprète la notion de résultant réduit. Qui plus est, Cayley avait publié au *Cambridge and Dublin Mathematical Journal* (en 1846 [59]) le calcul explicite de la polaire réciproque d'une cubique plane ; la méthode est correcte, mais l'expression finale est erronée [1]. Schläfli le remarque [2] en 1852 et donne l'expression exacte. Néanmoins, les types de problèmes étudiés par Brill sont très liés aux recherches exposés par Cayley en 1847.

1. Et il ne s'agit pas seulement d'une faute d'impression ; comparer par exemple le coefficient de x^5y avec celui qui figure dans l'expression donnée par Schläfli en 1852.

2. *cf.* [277] p. 74. Dans l'édition des *Collected mathematical papers* de Cayley, dont une partie fut supervisée par Cayley lui-même, Cayley ne fait pas mention du calcul effectué par Schläfli ; il édite pourtant l'article de 1846 en supprimant l'expression finale de la polaire réciproque. Il se justifie en écrivant que cette expression est aussi présente dans le mémoire [64]. Mais l'expression parue en 1856 dans [64] (*cf.* p. 328–329 du tome 2 des *Collected mathematical papers*) est différente de celle de 1846, et semble exacte (à en juger par le coefficient de x^5y). Enfin dans l'article [61] de 1847, Cayley affirme que les méthodes qu'il décrit pourront servir à « vérifier » l'expression parue en 1846, et il calcule, à titre d'exemple, le coefficient de z^6, qui était hélas déjà exact en 1846.

Cayley s'intéresse dans cet article à la théorie des courbes. Il expose plusieurs méthodes pour déterminer l'équation de la polaire réciproque d'une courbe projective plane, ses points d'inflexion (à la suite des travaux de Hesse) et ses tangentes doubles, au moyen de l'élimination. Il commence par déterminer l'équation du système des droites passant par un point (α, β, γ) et tangentes à la courbe $U = 0$ de degré n (α, β, γ sont alors conçues comme des indéterminées constantes). Cette équation peut s'écrire comme résultant réduit d'un quelconque des trois systèmes suivants :

$$\begin{cases} U = 0 \\ \alpha\frac{\partial U}{\partial x} + \beta\frac{\partial U}{\partial y} + \gamma\frac{\partial U}{\partial z} = 0 \\ \begin{vmatrix} x & y & z \\ \alpha & \beta & \gamma \\ \xi & \eta & \zeta \end{vmatrix} = 0 \end{cases} \qquad \text{(facteur superflu : } U(\alpha, \beta, \gamma)\text{)}$$

$$\begin{cases} U = 0 \\ \alpha\frac{\partial U}{\partial x} + \beta\frac{\partial U}{\partial y} + \gamma\frac{\partial U}{\partial z} = 0 \\ \xi\frac{\partial U}{\partial x} + \eta\frac{\partial U}{\partial y} + \zeta\frac{\partial U}{\partial z} = 0 \end{cases} \qquad \text{(facteur superflu : discriminant de } U\text{)}$$

$$\begin{cases} \alpha\frac{\partial U}{\partial x} + \beta\frac{\partial U}{\partial y} + \gamma\frac{\partial U}{\partial z} = 0 \\ \xi\frac{\partial U}{\partial x} + \eta\frac{\partial U}{\partial y} + \zeta\frac{\partial U}{\partial z} = 0 \\ \begin{vmatrix} x & y & z \\ \alpha & \beta & \gamma \\ \xi & \eta & \zeta \end{vmatrix} = 0 \end{cases} \qquad \text{(pas de facteur superflu)}$$

Puis il remplace, dans l'équation du système des tangentes, respectivement $\begin{vmatrix} \beta & \gamma \\ \eta & \zeta \end{vmatrix}, \begin{vmatrix} \gamma & \alpha \\ \zeta & \xi \end{vmatrix}, \begin{vmatrix} \alpha & \beta \\ \xi & \eta \end{vmatrix}$ par x,y,z pour obtenir l'équation de la polaire réciproque. Ce procédé est correct et a été démontré rigoureusement, et généralisé pour le calcul de l'hypersurface duale d'une hypersurface de $\mathbb{P}^n$ en dimension n quelconque, par Schläfli [3] (1852). Le résultant du troisième système peut encore s'écrire comme discriminant de $[U]$, où $[U]$ est une forme binaire en deux indéterminées l, m, résultat de la substitution de x, y, z par $\alpha l + \xi m$, $\beta l + \eta m$, $\gamma l + \zeta m$ dans U. Notons

$$\mathrm{d}^k U = \sum_{|\iota|=k} \binom{\iota}{k} \frac{\partial^k U}{\partial x^{\iota_1} \partial x^{\iota_2} \partial x^{\iota_3}} (\alpha, \beta, \gamma) \xi^{\iota_1} \eta^{\iota_2} \zeta^{\iota_3}$$

où $\iota = (\iota_1, \iota_2, \iota_3)$ désigne un multi-indice, et $\binom{\iota}{k}$ est le coefficient multinomial $\binom{\iota}{k} = \frac{k!}{\iota_1! \iota_2! \iota_3!}$. Alors (formule de Taylor) :

$$[U] = l^n U(\alpha, \beta, \gamma) + l^{n-1} m \mathrm{d}U + \frac{1}{2} l^{n-2} m^2 \mathrm{d}^2 U + \ldots + \frac{1}{n!} m^n \mathrm{d}^n U.$$

Si $U(\alpha, \beta, \gamma) = 0$, Cayley remarque que l'on a :

$$\frac{\mathrm{Disc}[U]}{(\mathrm{d}U)^2} = X \mathrm{d}U + Y (\mathrm{d}^2 U)^2$$

Si (α, β, γ) est un point d'inflexion, $\mathrm{d}U \mid \mathrm{d}^2 U$. Si (α, β, γ) est point de contact d'une tangente double, $\mathrm{d}U \mid Y$. Cayley retrouve ainsi, comme condition pour que (α, β, γ) soit point d'inflexion, l'annulation du hessien. Quant à la condition pour que (α, β, γ) soit point de contact d'une tangente double, Cayley trouve une certaine équation $\Pi U = 0$ de degré $(n-2)(n^2-9)$. La détermination de ΠU est légèrement simplifiée par Salmon dans son traité *Higher Plane Curves* [4] : il la fait reposer sur un lemme déterminant la condition pour que deux courbes de degré p, q, dont les pq points d'intersection sont confondus en un point unique, aient d'autres points d'intersection (leurs équations auront alors un facteur commun). On obtient à nouveau cette condition comme facteur d'un certain résultant.

3. *Cf.* [277] § 28.
4. *Cf.* [272] ; dans l'édition de 1879, *cf.* § 381.

Dans notre cas, tous les points d'intersection de dU et Y sont confondus en (α, β, γ), et comme dU est linéaire en ξ, η, ζ, la condition en question signifie que d$U \mid Y$. Dans son article de 1847, Cayley calcule aussi les degrés des équations du système des tangentes doubles (notée P$U = 0$) et du système des tangentes aux points d'inflexion (notée Q$U = 0$). Soit FF$U = 0$ l'équation de la polaire réciproque de la polaire réciproque de U ; comme $(\mathrm{P}U)^2(\mathrm{Q}U)^3U \mid \mathrm{FF}U$ (à cause des relations de Plücker) la connaissance des degrés de chaque membre en les coefficients de U permet à Cayley de conclure :

$$\mathrm{FF}U = \mathrm{K}U(\mathrm{P}U)^2(\mathrm{Q}U)^3U$$

où KU est le discriminant de la courbe.

Plusieurs articles de Cayley font suite à celui-ci, en 1864, 1865, 1871. L'article de 1864 nous intéresse [5]. On y retrouve en effet une situation semblable à celle que nous venons d'observer dans le lemme de Salmon : Cayley cherche à présent la condition pour qu'une courbe ayant déjà α points doubles et β points de rebroussement ait un point double supplémentaire. Il nomme « discriminant spécial » l'expression dont l'annulation manifeste cette condition. Il montre que lorsque les coefficients de l'équation de la courbe dépendent linéairement de paramètres, le degré du discriminant spécial en les paramètres est $3(n-1)^2 - 7\alpha - 11\beta$. Ce résultat et d'autres semblables lui permettent en 1865 de généraliser la formule pour FFU (qui n'était valable que pour une courbe suffisamment générale) au cas d'une courbe ayant α points doubles et β points de rebroussement à tangentes données, mais il avoue que cette investigation « a cependant par rapport à quelques points besoin de confirmation » [6]. En 1871, il vérifie [7] par un calcul explicite que sa formule est exacte dans le cas $n = 4$, $\alpha = 0$, $\beta = 2$.

On remarque chez Cayley, dans l'article de 1864, une conception de la théorie de l'élimination nettement influencée par ses travaux sur les invariants. L'élimination consiste, nous dit-il, à chercher « la relation qui doit exister entre les coefficients d'une fonction ou d'un système de fonctions pour que quelque circonstance particulière (ou singularité) puisse avoir lieu » [8]. Par conséquent, toutes les méthodes sont bonnes pour former des expressions algébriques ne contenant plus les variables à éliminer – en particulier les méthodes issues de la théorie des invariants, et non plus seulement les méthodes d'élimination algébrique proprement dites. Si cette idée était déjà présente dans les travaux antérieurs de Sylvester et Cayley, elle prend ici tout son sens, puisque l'objet (le résultant réduit) n'est pas le simple résultat d'opérations algébriques (multiplication, somme, méthode des coefficients indéterminés, etc.) visant à éliminer les inconnues, et que sa construction même renvoie à sa nature géométrique. C'est dans cette ordre d'idées que Cayley propose une formule symbolique pour la polaire réciproque d'une cubique. Notons avec lui :

$$\begin{aligned}
&(\mathrm{P}12) = \mathrm{x}\begin{vmatrix} \partial_{y_1} & \partial_{z_1} \\ \partial_{y_2} & \partial_{z_2} \end{vmatrix} + \mathrm{y}\begin{vmatrix} \partial_{z_1} & \partial_{x_1} \\ \partial_{z_2} & \partial_{x_2} \end{vmatrix} + \mathrm{z}\begin{vmatrix} \partial_{x_1} & \partial_{y_1} \\ \partial_{x_2} & \partial_{y_2} \end{vmatrix} \\
&(\mathrm{P}34) = \ldots \\
&\ldots \\
&UUUU = \prod_{1 \le i \le 4} U(x_i, y_i, z_i)
\end{aligned}$$

Alors on a :

$$-2\mathrm{F}U = (\mathrm{P}12)^2(\mathrm{P}34)^2(\mathrm{P}13)(\mathrm{P}42)UUUU$$

Enfin Cayley a l'espoir que ces recherches « puissent en quelque manière faciliter la solution du problème des réciproques des *surfaces* : objet, qui est resté encore dans une complète obscurité » [9]. C'est précisément ce que va faire Schläfli dans [277], § 28. Schläfli remarque de plus que l'élimination

5. *Cf.* [68].
6. *Cf.* [69] p. 416.
7. *Cf.* [71].
8. *Cf.* [68] p. 162.
9. *Cf.* [61] p. 345.

de x, y, z, θ dans les équations

$$\begin{cases} \mathrm{x}x + \mathrm{y}y + \mathrm{z}z = 0 \\ \frac{\partial U}{\partial x} = \mathrm{x}\theta^{n-1} \\ \frac{\partial U}{\partial y} = \mathrm{y}\theta^{n-1} \\ \frac{\partial U}{\partial z} = \mathrm{z}\theta^{n-1} \end{cases}$$

conduit seulement à une puissance $(\mathrm{F}U)^{n-1}$ de la polaire réciproque (et de même en dimension supérieure) [10]. Il expose d'emblée sa méthode en dimension quelconque, ce qui est certainement, en 1852, faire preuve d'audace. Il doit en particulier définir les concepts géométriques suivants :

- espace projectif $\mathbb{P}^n$ (*vollständige Totalität* ∞^n)
- hypersurface (*höhere einfach bedingte Totalität* ∞^{n-1})
- hyperplan (*höhere einfach bedingte lineare Totalität* ∞^{n-1})
- point (*direkte Lösung*)
- espace projectif dual $\check{\mathbb{P}}^n$ (*vollständige reziproke Totalität*)
- hypersurface de $\check{\mathbb{P}}^n$ (*einfach bedingte reziproke Totalität*)
- point de $\check{\mathbb{P}}^n$ (*reziproke Lösung*)

Il faut à présent tourner notre attention vers l'article de Brill [42], dans lequel celui-ci présente une théorie nouvelle du résultant réduit, dont nous avons cependant déjà vu deux antécédents – le lemme de Salmon et le « discriminant spécial » de Cayley. Brill désigne d'abord par la locution « résultant réduit » la condition pour que trois équations à deux inconnues (trois courbes planes) qui ont une solution commune en aient une autre en plus (il traite ensuite de cas plus compliqués, où sont données plusieurs solutions communes, ou bien une solution en laquelle les courbes, deux à deux, ont certaines multiplicités d'intersection, etc. – nous n'envisagerons ici que le cas le plus simple). Cette définition du résultant réduit ne peut concerner que des systèmes d'équations particuliers (puisqu'ils doivent déjà posséder une solution commune, alors qu'en général un système de trois équations à deux inconnues n'a pas de solution). Brill considère deux cas de figure :

(i) la solution commune est un point (α, β) entièrement indéterminé (les coefficients des trois équations $f = 0$, $\phi = 0$, $\psi = 0$ contiennent les indéterminées α et β de sorte que ce point appartienne aux trois courbes) ;

(ii) le point commun (α, β) est un point quelconque d'une courbe $f = 0$ donnée (les indéterminées α, β sont donc soumises à l'équation $f(\alpha, \beta) = 0$, et seuls les coefficients des équations des deux autres courbes $\phi = 0$, $\psi = 0$ contiennent les indéterminés α, β).

Le cas (ii) peut aussi s'interpréter comme la recherche de « paires coïncidentes », c'est-à-dire des paires de points $(x, y, z), (\alpha, \beta, \gamma)$ de la courbe $f = 0$ vérifiant simultanément les deux équations $\phi = 0$, $\psi = 0$. Résumons la méthode de Brill dans le cas (i). Il s'agit, pour l'essentiel, de dériver la formule de Poisson. Notons

$$R = \mathrm{Res}(f, \phi, \psi) = \psi(\alpha_1, \beta_1)\psi(\alpha_2, \beta_2)...\psi(\alpha_{mn}, \beta_{mn})$$

où $(\alpha_1, \beta_1) = (\alpha, \beta)$, et

$$\psi(x, y) = \sum_\iota w_\iota x^{\iota_1} y^{\iota_2}$$

où ι est un multi-indice. Alors :

$$\frac{\partial R}{\partial w_\iota} = \alpha^{\iota_1}\beta^{\iota_2}\psi(\alpha_2, \beta_2)...\psi(\alpha_{mn}, \beta_{mn}).$$

Quand le résultant réduit s'annule, l'un des $\psi(\alpha_k, \beta_k)$ $(k \geq 2)$ doit s'annuler, et donc $\frac{\partial R}{\partial w_i}$ aussi. Par conséquent le résultant réduit divise tous les $\frac{\partial R}{\partial w_\iota}$; mais leur facteur commun contient encore un

10. Dans [168] § 6.3.11, J.-P. Jouanolou, après avoir qualifié cette situation de « paradoxe apparent », montre que l'on peut obtenir l'équation de la polaire réciproque (l'« hypersurface duale ») plus directement, sous forme d'un « résultant anisotrope ».

facteur superflu, exprimant la condition pour que la multiplicité d'intersection de $f = 0$ et $\phi = 0$ en (α, β) soit > 1 (auquel cas il existe un $k \geq 2$ tel que $(\alpha_k, \beta_k) = (\alpha, \beta)$).

On pourrait souhaiter une démonstration intrinsèque de la propriété fondamentale du résultant réduit (c'est-à-dire indépendante de la clôture algébrique du corps contenant les coefficients). Il faut pour cela trouver une alternative à la formule de Poisson. Rappelons d'abord comment l'on démontre de manière intrinsèque la propriété fondamentale du résultant. Soit $R = \mathrm{Res}(f_1, f_2, f_2)$ où f_1, f_2, f_3 sont homogènes en les indéterminées X_1, X_2, X_3, et génériques dans un premier temps. Les lettres grecques $\iota, \kappa, \lambda, \mu, N$ désigneront ci-dessous des multi-indices.

$$\begin{aligned} f_1 &= \sum u_\kappa X^\kappa \\ f_2 &= \sum v_\kappa X^\kappa \\ f_3 &= \sum w_\kappa X^\kappa \end{aligned}$$

Dérivons l'identité $X^\iota R = \sum_{1 \leq t \leq 3} \phi_t f_t$ selon les coefficients w_κ de f_3 :

$$(*) \quad X^\iota \frac{\partial R}{\partial w_\kappa} = \phi_3 X^\kappa + \sum \frac{\partial \phi_t}{\partial w_\kappa} f_t$$

Multiplions cette égalité par $X^\iota \frac{\partial R}{\partial v_{N-\kappa}}$, on obtient :

$$X^{2\iota} \frac{\partial R}{\partial w_\kappa} \frac{\partial R}{\partial v_{N-\kappa}} = (\phi_3 X^\lambda).(X^\iota \frac{\partial R}{v_{N-\kappa}}).X^{\kappa-\lambda} \qquad \mod (f_1, f_2, f_3)$$

En appliquant au membre de droite, à trois reprises, l'égalité $(*)$ ou une égalité semblable obtenue en dérivant selon un autre coefficient, on obtient finalement :

$$X^{2\iota} \frac{\partial R}{\partial w_\kappa} \frac{\partial R}{\partial v_{N-\kappa}} = X^{2\iota} \frac{\partial R}{\partial w_\lambda} \frac{\partial R}{\partial v_{N-\lambda}} \qquad \mod (f_1, f_2, f_3)$$

Donc la quantité

$$\frac{\partial R}{\partial w_\kappa} \frac{\partial R}{\partial v_{N-\kappa}} - \frac{\partial R}{\partial w_\lambda} \frac{\partial R}{\partial v_{N-\lambda}}$$

appartient à l'idéal (R). Spécialisons f_1, f_2, f_3 de sorte que R s'annule. Alors pour tous κ, λ, N on a

$$\frac{\partial R}{\partial w_\kappa} \frac{\partial R}{\partial v_{N-\kappa}} - \frac{\partial R}{\partial w_\lambda} \frac{\partial R}{\partial v_{N-\lambda}} = 0$$

On démontre de même que pour tous κ, λ, N on a :

$$\frac{\partial R}{\partial w_\kappa} \frac{\partial R}{\partial w_{N-\kappa}} - \frac{\partial R}{\partial w_\lambda} \frac{\partial R}{\partial w_{N-\lambda}} = 0$$

Ces relations impliquent qu'il existe x_1, x_2, x_3 tels que les monômes x^κ en x_1, x_2, x_3 soient proportionnels aux $\frac{\partial R}{\partial w_\kappa}$ pour tout multi-indice κ tel que $|\kappa| = \deg f_3$. Or pour un tel x,

$$R \sim \sum w_\kappa \frac{\partial R}{\partial w_\kappa} \sim \sum w_\kappa x^\kappa = f_3(x)$$

donc $f_3(x) = 0$. En remplaçant w par v (resp. u) dans $(*)$, on montre de même que $f_2(x) = 0$ (resp. $f_1(x) = 0$) *pour la même valeur de* x. Si les dérivées partielles du résultant ne sont pas toutes nulles, l'existence d'une solution au système d'équations proposé est ainsi démontrée.

Pour démontrer la propriété du résultant réduit, revenons au cas générique et étudions les dérivées secondes du résultant, en dérivant $(*)$:

$$(**) \quad X^\iota \frac{\partial^2 R}{\partial w_\kappa \partial w_\lambda} = \frac{\partial \phi_3}{\partial w_\lambda} X^\kappa + \frac{\partial \phi_3}{\partial w_\kappa} X^\lambda \qquad \mod (f_1, f_2, f_3)$$

En écrivant des égalités semblables pour $\frac{\partial^2 R}{\partial w_\lambda \partial w_\mu}$ et $\frac{\partial^2 R}{\partial w_\mu \partial w_\kappa}$, on obtient un système de trois équations linéaires en $\frac{\partial \phi_3}{\partial w_\kappa}$, $\frac{\partial \phi_3}{\partial w_\lambda}$ et $\frac{\partial \phi_3}{\partial w_\mu}$, dont la solution modulo (f_1, f_2, f_3) est de la forme suivante (formule de Cramer) :

$$\frac{\partial \phi_3}{\partial w_\kappa} = \frac{\begin{vmatrix} \frac{\partial^2 R}{\partial w_\lambda \partial w_\mu} & X^\mu & X^\lambda \\ \frac{\partial^2 R}{\partial w_\kappa \partial w_\mu} & & X^\kappa \\ \frac{\partial^2 R}{\partial w_\kappa \partial w_\lambda} & X^\kappa & \end{vmatrix}}{\begin{vmatrix} & X^\mu & X^\lambda \\ X^\mu & & X^\kappa \\ X^\lambda & X^\kappa & \end{vmatrix}} = \frac{-X^{2\kappa} \frac{\partial^2 R}{\partial w_\lambda \partial w_\mu} + X^{\kappa+\lambda} \frac{\partial^2 R}{\partial w_\kappa \partial w_\mu} + X^{\kappa+\mu} \frac{\partial^2 R}{\partial w_\kappa \partial w_\lambda}}{2X^{\kappa+\lambda+\mu}}$$

Il est alors facile de montrer que, modulo f_1, f_2, f_3 et les dérivées premières du résultant, pour tous κ, λ, N :

$$\frac{\partial \phi_3}{\partial w_\kappa} \frac{\partial \phi_2}{\partial v_{N-\kappa}} - \frac{\partial \phi_3}{\partial w_\lambda} \frac{\partial \phi_2}{\partial v_{N-\lambda}} = 0$$

On a de même, modulo f_1, f_2, f_3 et les dérivées premières du résultant :

$$\frac{\partial \phi_3}{\partial w_\kappa} \frac{\partial \phi_3}{\partial w_{N-\kappa}} - \frac{\partial \phi_3}{\partial w_\lambda} \frac{\partial \phi_3}{\partial w_{N-\lambda}} = 0$$

Spécialisons à présent f_1, f_2, f_3 de sorte qu'elles aient un zéro commun en $(x_1 : x_2 : x_3)$. Alors $R = 0$, mais supposons de plus que les dérivées premières de R s'annulent toutes. Les relations que l'on vient d'obtenir montrent qu'il existe y_1, y_2, y_3 tels que les monômes y^κ en y_1, y_2, y_3 soient proportionnels aux $\frac{\partial \phi_3}{\partial w_\kappa}(x)$ pour tout multi-indice κ tel que $|\kappa| = \deg f_3$. Or pour un tel y, $(**)$ implique que :

$$\frac{\partial R}{\partial w_\kappa} \sim \sum w_\lambda \frac{\partial^2 R}{\partial w_\kappa \partial w_\lambda} \sim x^\kappa f_3(y) + y^\kappa f_3(x)$$

On en conclut que $f_3(y) = 0$, et l'on montrerait de même que, *pour la même valeur de* y, f_1 et f_2 s'annulent. La propriété fondamentale du résultant réduit est ainsi démontrée.

Il faut aussi rattacher les travaux sur le résultant réduit aux recherches de Sylvester ([304] et [303]) qui tentait, en 1852, de classer les singularités des courbes planes de sorte que la somme totale des ordres de multiplicité des singularités d'une courbe $U = 0$ soit égale au rang k minimal tel que la différentielle $\mathrm{d}^k D$ du discriminant D de la courbe ne s'annule pas. C'est en fait la multiplicité, en le point singulier $(x : y : z)$, de l'idéal $\left(\frac{\partial U}{\partial x}, \frac{\partial U}{\partial y}, \frac{\partial U}{\partial z}\right)$ des premières polaires de U que Sylvester cherche à définir et à déterminer par ce critère. Il écrit par exemple :

> (...) de même qu'en un point double les premières polaires se coupent simplement, mais qu'en un point de rebroussement elles ont toutes l'une avec l'autre un contact du premier degré, de même nous devrions nous attendre à ce qu'il existe une espèce de point singulier telle que toutes les premières polaires aient l'une avec l'autre un contact du second degré en ce point (...) [11]

Toutes les recherches sur le résultant réduit ou le discriminant spécial renvoient d'ailleurs à la même question : comment définir la multiplicité d'un point d'intersection de trois courbes planes ? C'est l'étude du phénomène analogue en dimension 0 qui semble avoir servi de modèle à Sylvester. En dimension 0, l'approche de Sylvester est correcte et a été exposée plus rigoureusement par Hilbert [149] en 1887. Hilbert énonce comme suit un critère déterminant la multiplicité et la distribution des racines d'un polynôme $X(t) = a_1 t^n + a_2 t^{n-1} + ... + a_{n+1}$. Supposons que :

$$X(t) = (t - t_1)^{\mu_1} (t - t_2)^{\mu_2} ... (t - t_k)^{\mu_k}$$

11. *Cf.* [303] p. 372.

où $\mu_1, ..., \mu_k \geq 1$ et $t_1, ..., t_k$ sont des racines distinctes du polynôme X. Alors $\mathrm{d}D$,...,$\mathrm{d}^{n-k-1}D$ s'annulent identiquement et

$$\mathrm{d}^{n-k}D = c(a_1t_1^n+a_2t_1^{n-1}+...+a_{n+1})^{\mu_1-1}(a_1t_2^n+a_2t_2^{n-1}+...+a_{n+1})^{\mu_2-1}...(a_1t_k^n+a_2t_k^{n-1}+...+a_{n+1})^{\mu_k-1}$$

où c est une constante. Réciproquement, si les différentielles de D vérifient ces conditions, alors $X(t)$ se factorise comme ci-dessus.

16.2 L'histoire de la théorie de l'élimination selon Brill

L'article encyclopédique [235] de E. Netto et plus encore sa version française, cossignée par Netto et Le Vavasseur [12], dans l'*Encyclopédie des sciences mathématiques pures et appliquées* de Jules Molk, sont une mine de références bibliographiques dont certaines renvoient même à des ouvrages du XVII$^{\text{ème}}$ siècle. Mais la réflexion historique en est absente et, fait inévitable dans une entreprise si vaste, l'historiographie elle-même se révèle laborieuse : Netto et Le Vavasseur avouent que « le grand nombre et la variété des mémoires relatifs à l'élimination rendent difficile une classification rationnelle de ces mémoires » [13]. On pouvait s'attendre à un travail d'une autre nature de la part de A. Brill, collaborateur de M. Noether dans la rédaction du célèbre rapport sur l'histoire de la géométrie algébrique *Die Entwicklung der Theorie der algebraischen Functionen in älterer und neuerer Zeit* [49] paru en 1892. Et en effet, lors du congrès international à Heidelberg en 1904, Brill fait une communication sur « l'élimination et la géométrie dans les dernières décennies » [40], soi-disant « une sorte d'annexe géométrique au rapport détaillé sur l'élimination fait d'un point de vue algébrique dans l'encyclopédie ». Si Brill s'y restreint à la théorie de l'élimination « dans son rapport à la géométrie », c'est peut-être aussi ce point de vue qui confère à sa réflexion l'unité qui manquait à celle de Netto. Brill commence son discours en évoquant les travaux de Hesse : son usage de la théorie des déterminants était au cœur du renouveau de la géométrie algébrique dans les années 1840. S'ouvrait alors à cette discipline un champ d'investigations, entre géométrie et algèbre, qui contient la théorie de l'élimination. L'article de Brill paru en 1915 dans les actes du congrès est hélas court et peu documenté. Sans vouloir ici combler cette lacune en détaillant chacun de ses arguments, nous nous contenterons de commenter quelques-unes des idées qui y sont abordées. Il est remarquable que Brill semble d'abord attacher une certaine importance à la recherche d'une formule du résultant comme quotient de déterminants ; il cite ainsi les articles [60] et [62] de Cayley (*cf. supra* section 10.1) qui n'occupent peut-être pas tout l'espace qu'ils méritent dans l'article de Netto et Le Vavasseur. Les problèmes proprement géométriques qu'évoque ensuite Brill se répartissent en trois thèmes : des problèmes concernant ce que nous appelons aujourd'hui des « variétés déterminantielles » [14], la théorie du résultant réduit, et le vaste domaine de la géométrie énumérative.

Nous mentionnerons seulement un des problèmes se rattachant au premier thème et que Brill connaissait bien : l'étude des *ausgezeichnete Punktgruppen* [15]. Il y a en effet consacré une section du rapport *Die Entwicklung der Theorie der algebraischen Functionen* [16], et ce sujet lui avait aussi inspiré un article [43] en 1872. Il faut peut-être chercher l'origine de ce problème dans le mémoire de Riemann sur les fonctions abéliennes [17], mais c'est à la *Theorie der Abelschen Functionen* de Clebsch et Gordan [77] que Brill renvoie plutôt. La situation qui nous intéresse y est décrite comme suit.

12. *Cf.* [234] t. 1, vol. 2, § 33–72, p. 73–169 ; ce fascicule est paru le 19 novembre 1907. Les trois derniers articles § 70–72 ont été entièrement rédigés par J. Kürschák et E. Delassus. G. Enestrom, A. Brill, G. Loria, G. Vivanti, Gy. (J.) König ont aussi participé à la rédaction de certains passages, repérés par des astérisques.
13. *Cf.* [234] t. 1, vol. 2, p. 104.
14. Pour des références récentes sur les variétés déterminantielles, *cf.* par exemple [123]. Nous avons déjà croisé un exemple de variété déterminantielle *supra* p. 237.
15. Nous ne traduirons pas ce terme, qui désigne des points bases de ce que l'on appelle aujourd'hui un « diviseur spécial » (*cf.* [140] p. 296). Le diviseur spécial lui-même était alors désigné par le terme *Specialgruppe*.
16. [49] p. 542–552.
17. *Cf.* les remarques de Noether [49] p. 348.

Etant donné une courbe algébrique plane X de degré n, le problème de l'inversion des intégrales abéliennes proposé par Jacobi consistait à chercher un inverse à l'application

$$\begin{array}{rlll} v=(v_1,...,v_p): & X^p/\mathfrak{S}_p & \longrightarrow & \mathbb{C}^p/\Omega \\ & (x^{(1)},...,x^{(p)}) & \longmapsto & \left(\sum\limits_{i=1}^{p}\int\limits_{\alpha^{(i)}}^{x^{(i)}} \mathrm{d}u_1,...,\sum\limits_{i=1}^{p}\int\limits_{\alpha^{(i)}}^{x^{(i)}} \mathrm{d}u_p\right) \end{array}$$

où $\langle \mathrm{d}u_1,...,\mathrm{d}u_p\rangle$ est une base de l'espace des formes différentielles de première espèce sur X, les $\alpha^{(i)}$ sont des constantes arbitraires, le groupe symétrique $\mathfrak{S}_p$ agit par permutation sur les p coordonnées dans X^p, et Ω est le réseau de $\mathbb{C}^p$ engendré par les systèmes de périodes des intégrales $\int \mathrm{d}u_j$. Le problème de l'inversion avait été résolu peu auparavant par Riemann et Weierstrass [18]. La méthode trouvée par Riemann (et suivie par Clebsch et Gordan) consistait à définir certaines « fonctions Θ » de p variables complexes et à considérer l'expression :

$$\Theta\left(v_1(x^{(1)},...,x^{(p)})-\int_\mu^\xi \mathrm{d}u_1,\quad ...,\quad v_p(x^{(1)},...,x^{(p)})-\int_\mu^\xi \mathrm{d}u_p\right)$$

où μ et ξ sont des indéterminées, mais où les limites d'intégration $\alpha^{(i)}$ font l'objet d'un choix approprié (et dépendent de μ). On montre alors que cette expression ne s'annule que dans deux cas de figure :

- si ξ coïncide avec l'un des $x^{(i)}$;
- s'il existe une courbe adjointe de X de degré $(n-3)$ qui passe par tous les $x^{(i)}$.

Donc, sauf dans le second cas, les racines ξ de l'équation

$$\Theta\left(v_1-\int_\mu^\xi \mathrm{d}u_1,\quad ...,\quad v_p-\int_\mu^\xi \mathrm{d}u_p\right)=0$$

constituent la solution du problème de l'inversion [19]. Au contraire, dans le second cas, Clebsch et Gordan montrent que $v(x^{(1)},...,x^{(p)})$ n'a pas un antécédent unique. Soit en effet une courbe adjointe passant par les points $x^{(1)},...,x^{(p)},x^{(p+1)},...,x^{(2p-2)}$. Ils montrent qu'il existe alors une infinité d'autres courbes adjointes passant par les $(p-2)$ points $x^{(p+1)},...,x^{(2p-2)}$. Si l'une d'elles découpe sur la courbe le groupe de points $x_0^{(1)},...,x_0^{(p)},x^{(p+1)},...,x^{(2p-2)}$, on aura pour tout $1\leq j\leq p$ (théorème d'Abel) :

$$\sum_{i=1}^{p}\int_{x_0^{(i)}}^{x^{(i)}} \mathrm{d}u_j\equiv 0$$

modulo les périodes. D'où $v(x_0^{(1)},...,x_0^{(p)})=v(x^{(1)},...,x^{(p)})$.

Pour mieux comprendre les résultats atteints par Clebsch et Gordan, qu'il nous soit permis de traduire cette situation dans un langage plus moderne. Le quotient $\mathbb{C}^p/\Omega$ est isomorphe à ce que l'on appelle aujourd'hui la « variété jacobienne » de X, dont les points sont les classes de diviseurs de degré 0. Comme tout diviseur de degré p est équivalent à un diviseur effectif $(x^{(1)})+...+(x^{(p)})$ auquel on peut à son tour associer le diviseur de degré zéro $(x^{(1)})+...+(x^{(p)})-O^p$ (où O est un diviseur effectif de degré 1 arbitraire), on peut identifier les points de la jacobienne aux classes de diviseurs de degré p. Dire qu'il existe une courbe adjointe passant par $x^{(1)},...,x^{(p)}$ revient à dire qu'il existe un diviseur $(x^{(1)})+...+(x^{(p)})+(x^{(p+1)})+...+(x^{(2p-2)})$ dans la classe canonique K, c'est-à-dire que $l(K-(x^{(1)})-...-(x^{(p)}))>0$. D'après le théorème de Riemann-Roch ceci signifie que $l((x^{(1)})+...+(x^{(p)}))>1$, c'est-à-dire qu'il existe un diviseur effectif $(y^{(1)})+...+(y^{(p)})$ équivalent

18. L'histoire de ce problème a été écrite par Christian Houzel, *cf.* [157] p. 169–184.

19. Clebsch et Gordan définissent la fonction Θ à partir d'une fonction $V(x^{(1)},...,x^{(p)};\xi)$ en [77] § 55 p. 195. La fonction V elle-même est étudiée – en particulier ses zéros – en § 50 p. 174. Les cas indéterminés du problème de l'inversion sont mis en rapport avec l'étude des zéros de la fonction Θ au § 60 p. 206.

à $(x^{(1)}) + ... + (x^{(p)})$ mais de support différent : autrement dit $v(x^{(1)}, ..., x^{(p)})$ n'a pas un antécédent unique [20]. Clebsch et Gordan vont donc s'attacher à étudier la sous-variété de $X^p/\mathfrak{S}_p$ définie par la condition $l(K - (x^{(1)}) - ... - (x^{(p)})) > 0$. Plus précisément, ils montrent que le système linéaire des diviseurs découpés sur X par les courbes adjointes passant par $(p-1)$ points donnés $x^{(1)}$,...,$x^{(p-1)}$ de X est *en général* de dimension zéro (alors $l(K - (x^{(1)}) - ... - (x^{(p-1)})) - 1 = 0$, c'est-à-dire qu'il existe une seule courbe adjointe passant par ces $(p-1)$ points), mais que, pour tout $k < p-1$ et pour tout groupe de $(2k-p+1)$ points, il existe un nombre fini de manières de le compléter en un groupe de $(k+1)$ points qui soient points bases d'un système linéaire de dimension $(p-k-1)$ de diviseurs découpés sur X par des courbes adjointes ; le terme *ausgezeichnete Punktgruppe* désigne un tel groupe de $(k+1)$ points.

Dans l'article de 1872 déjà mentionné, Brill s'est posé un problème semblable, mais où la courbe X est remplacée par une droite (d'équation $\mu = 0$) et le système des courbes adjointes par un système de courbes de degré m donné et dépendant de $(k+i)$ paramètres (avec $i \leq (k-1)$ et $2k \leq m$). Sa méthode est semblable à celle suivie par Clebsch et Gordan, mais il détermine exactement, au moyen de la théorie de l'élimination, le nombre de manières de compléter le groupe de points donné. L'équation générale d'une courbe du système est de la forme :

$$\alpha_1\phi_1(\lambda,\mu) + \alpha_2\phi_2(\lambda,\mu) + ... + \alpha_{k+i}\phi_{k+i}(\lambda,\mu) = 0$$

où les $\alpha_1, ..., \alpha_{k+i}$ sont les paramètres. Brill démontre le théorème suivant :

> Il y a
> $$(k+i)_k = \frac{(m-k+1)(m-k+2)...(m-k-i+1)}{1.2....(i+1)}$$
> manières différentes de trouver sur une droite un groupe de i points qui constitue, réuni à $(k-i-1)$ points donnés de la droite $(i \leq k-1)$, un groupe de points tel que toutes les courbes d'un système donné de courbes de degré m $(m \geq 2k)$ à $(k+i-1)$ degrés de liberté [21] qui passent par ces points se coupent encore en un autre $k^{\text{ème}}$ point de la droite [22].

Il étudie à cet effet la variété déterminantielle définie par l'annulation des mineurs $k \times k$ de la matrice suivante (où $q = k+i$) :

$$\begin{pmatrix} \phi_1(\lambda_1,\mu_1) & \phi_2(\lambda_1,\mu_1) & \dots & \phi_q(\lambda_1,\mu_1) \\ \phi_1(\lambda_2,\mu_2) & \phi_2(\lambda_2,\mu_2) & \dots & \phi_q(\lambda_2,\mu_2) \\ \vdots & & & \vdots \\ \phi_1(\lambda_k,\mu_k) & \phi_2(\lambda_k,\mu_k) & \dots & \phi_q(\lambda_k,\mu_k) \end{pmatrix}$$

Brill écrit cette variété déterminantielle comme *somme alternée* d'intersections de variétés déterminantielles définies par des matrices plus petites. La justification d'une telle démarche requiert une théorie des cycles ; l'exposé de Brill manque donc de rigueur, mais il se ramène ainsi au cas d'une variété définie comme ensemble solution d'un système de ρ équations à ρ inconnues. Ces équations appartiennent à une certaine classe, qu'avait déjà étudiée Bézout, et dont Brill décrit ainsi la forme générale :

> Nous dirons qu'une équation à ρ variables $x_1, x_2, ..., x_\rho$, qui est de degré p_1 en x_1, p_2 en x_2, ..., p_ρ en x_ρ, est générale, quand chacun de ses termes de la forme :
> $$x_1^{i_1} x_2^{i_2} ... x_\rho^{i_\rho} \quad (i_1 \leq p_1,\ i_2 \leq p_2, ...,\ i_\rho \leq p_\rho)$$
> survient affecté d'un coefficient quelconque [23].

20. Nous nous sommes inspirés ici de [290] p. 242–243. Dans ce qui suit, nous traduisons les résultats démontrés par Clebsch et Gordan dans [77] § 61 vers le langage des « systèmes linéaires de diviseurs ». Brill et Noether adopteront bientôt (dans les années 1870) un langage proche de celui-ci, celui des « séries linéaires de groupes de points ».

21. *Bestimmungsstücken* dans le texte ; c'est-à-dire que l'équation générale d'une courbe du système dépend linéairement de $(k+i)$ paramètres comme nous l'avons décrit ci-dessus.

22. *Cf.* [43] p. 395.

23. *Cf.* [43] p. 383.

En étudiant le degré de certaines des fonctions symétriques intervenant dans la formule de Poisson du résultant, Brill montre que l'élimination de $(\rho - 1)$ inconnues conduit à un résultant de degré

$$\sum_i \prod_j m_i^{(j)}$$

en la dernière inconnue, où $m_i^{(j)}$ désigne le degré en x_i de la $j^{\text{ème}}$ équation du système [24]. Dans le cas $\rho = 3$, ce résultat est inclus dans les formules qu'avaient données Bézout [25].

Revenons à l'article historique de 1904. Après avoir évoqué divers problèmes de géométrie impliquant la théorie de l'élimination, Brill conclut :

> Une analyse plus exacte et une comparaison des méthodes mentionnées avec d'autres méthodes servant au même but – ce dont je dois ici m'abstenir – montreraient leur force et leur souplesse. Mais s'il ne s'agit que de déterminer un degré, c'est-à-dire le nombre de solutions, et non pas du calcul effectif (*die wirkliche Darstellung*) du résultat de l'élimination, alors l'algèbre est encore aujourd'hui dépassée par les méthodes énumératives géométriques connues qui depuis le milieu des années soixante ont franchement inondé la géométrie avec leurs résultats énumératifs ; je veux parler de la méthode des caractéristiques, du principe de correspondance et récemment du principe de conservation du nombre.
>
> Ces principes expriment comment le degré du résultant de certains groupes d'équations de conditions dépend du degré de celles-ci ; ce en quoi l'on renonce à construire le résultant lui-même, ou même les termes intermédiaires (*der Zwischenglieder*). Avec cela on renonce surtout aussi à la pleine rigueur démonstrative. Le rôle de ces principes est d'applanir et de préparer le terrain, dans les domaines moins connus de la géométrie, par un travail de pionnier souvent difficile, en vue du travail particulier ultérieur de l'algèbre.
>
> Certes l'épreuve des limites de validité de ces principes restera toujours l'affaire de l'algèbre. Dans ce but, Clebsch a élaboré, à partir de la théorie des formes quadratiques ternaires, une démonstration de la formule des caractéristiques, démonstration étendue ensuite par Halphen et d'autres [26]. La formule de correspondance a aussi fait l'objet de plusieurs recherches algébriques [27]. Seul le principe de conservation du nombre [28] attend encore son dû de la part de l'algèbre. Juger s'il est valable, et dans quelles limites, ne pourra être fait qu'à partir des fondements généraux de la théorie de l'élimination. [29]

Dans cette longue citation, Brill décrit bien un certain rapport disciplinaire entre théorie de l'élimination et géométrie ; rapport selon lequel la rigueur démonstrative reste du ressort exclusif de l'algèbre, aussi bien que le calcul effectif d'objets dont les principes de la géométrie énumérative ne donnent qu'une connaissance *a priori*. La démonstration des énoncés adoptés comme « principes » par la géométrie énumérative appartient dès lors aux fondements de la théorie de l'élimination. Mais la théorie de l'élimination ne pouvait servir de fondement à la géométrie algébrique, à moins d'être elle-même suffisamment fondée : beaucoup de travail restait à accomplir. Brill place son espoir dans la méthode d'élimination de Kronecker :

> On doit [les progrès les plus récents de la théorie de l'élimination], on ne peut le cacher, à l'influence réciproque (*Eingreifen*) entre l'arithmétique et la théorie des formes, après

24. *Cf.* [43] p. 386.
25. On le vérifie en calculant, dans le langage de Bézout, le degré de l'équation finale pour la « première forme d'équations incomplètes de seconde espèce ».
26. On a vu cependant que la rigueur, dans la « méthode des caractéristiques », laissait encore à désirer dans le mémoire de Halphen, *cf. supra* section 14.2.
27. Nous en avons aussi vu un exemple dans le mémoire de Halphen, *cf. supra* le « nombre de couples de points correspondants confondus » p. 230.
28. Le « principe de conservation du nombre » n'est autre que le principe de continuité que nous avons rencontré chez Poncelet (*cf. supra* p. 100).
29. *Cf.* [43] p. 278.

> que l'incitation au développement des méthodes d'élimination a été pendant longtemps presque exclusivement issue de la géométrie [30].

Etant donné une variété intersection, rappelons que la méthode de Kronecker détermine de manière effective des systèmes d'équations dont chacun a pour ensemble solution l'une des composantes équidimensionnelles de cette variété. Ces systèmes d'équations sont chacun représentés par une forme, facteur de la « résolvente totale » (*Gesamtresolvente*). Brill remarque :

> On peut tirer profit de ces circonstances dans les énumérations géométriques, tant qu'elles ne concernent que des déterminations de degré. En effet entre les limites définies par l'annulation identique de la résolvente totale, une spécialisation des variétés géométriques que l'on intersecte ne peut avoir pour effet aucune variation du degré des variétés à l'intersection [c'est-à-dire des degrés des composantes équidimensionnelles de la variété intersection] [31].

Essayons de préciser l'affirmation de Brill. Supposons que les équations des variétés que l'on intersecte dépendent d'un paramètre t. La résolvente totale, et ses facteurs les « résolventes partielles », sont d'abord calculés pour t indéterminé au moyen de l'algorithme de Kronecker, et ils dépendent donc de t. On obtient ainsi la résolvente totale pour une valeur générale de t, et chaque composante équidimensionnelle est déterminée par la résolvente partielle correspondante, sauf dans le cas des valeurs spéciales de t qui font s'annuler identiquement l'une des résolventes partielles (et donc aussi la résolvente totale) [32]. Van der Waerden suivit d'ailleurs une idée analogue dans son article de 1927 pour définir la multiplicité d'une spécialisation, en dimension 0, en étudiant l'effet de la spécialisation sur la résolvente totale [33]. Nous avons déjà mentionné cet article qui compte peut-être involontairement, dans les années 1940–1950, parmi les causes de l'image déchue de la théorie de l'élimination (*cf. supra* p. 225).

16.3 Conclusion : géométrie algébrique et élimination de l'élimination

> Il n'est pas de brouillards, comme il n'est point d'algèbres,
> Qui résistent, au fond des nombres ou des cieux,
> A la fixité calme et profonde des yeux ;
>
> Victor Hugo, *La légende des siècles*

L'histoire de la théorie de l'élimination se résume à trois périodes. Au cours de la première, dès le début même de l'algèbre, les mathématiciens utilisent certaines techniques d'élimination dans des cas particuliers, sans toutefois posséder de méthode générale. Avec la seconde période commence ce que nous avons appelé l'« histoire courte » de l'élimination. A partir de Hudde et Fermat au XVII$^{\text{ème}}$ siècle, on dispose au moins d'une méthode générale d'élimination d'une inconnue dans un système de deux équations polynomiales quelconques. Nous avons vu cependant qu'un mathématicien chinois utilisait une méthode semblable dès le XIII$^{\text{ème}}$ siècle, sans toutefois posséder la notation et les concepts nécessaires pour l'appliquer à deux équations ayant un nombre quelconque d'inconnues. Dans le dernier tiers du XVII$^{\text{ème}}$ siècle, commence à naître la *théorie de l'élimination*, à la suite des travaux de Leibniz, Newton et Rolle. Newton avait énoncé un théorème : le théorème de Bézout. Il avait calculé les équations finales, délivrées de facteur superflu, pour deux équations générales de bas

30. *Cf.* [43] p. 282.
31. *Cf.* [43] p. 281.
32. Méfions-nous des apparences, *cf.* l'exemple décrit dans [117] p. 41 : sur une certaine hypersurface singulière de $\mathbb{P}^4$ il existe deux droites dont l'une peut être continuement déplacée vers la position de l'autre (elles sont « rationnellement équivalentes »), sans pour autant qu'il existe une droite générale sur cette hypersurface, dont elles soient toutes deux des spécialisées. Nous allons donner un exemple plus élémentaire, ayant cependant l'inconvénient de faire intervenir des variétés non irréductibles. Plaçons-nous dans une variété ambiante singulière, la réunion X de deux plans P_1 et P_2 dans $\mathbb{A}^3$. Soit alors Y la variété réunion d'une conique dans P_1 et d'une cubique dans P_2. Il existe alors deux familles de droites dans X (celles du plan P_1 et celles du plan P_2). Les unes coupent Y en deux points, les autres en trois points, et il n'existe pas de droite générale de X dont ces droites seraient toutes deux des spécialisées.
33. Dans l'article de van der Waerden, Kronecker n'est pas cité et la résolvente totale est appelée *die u-Resultante*, *cf.* [316] p. 759.

degrés. Leibniz avait élaboré un véritable programme de recherche, qu'il ne put hélas poursuivre, en rapport avec ses travaux d'algèbre linéaire. Rolle, enfin, étudia l'objet « système d'équations » et décrivit une suite d'opérations pour transformer un tel système, au terme de la construction d'un arbre, en un ensemble de systèmes triangulaires. Ces diverses recherches méritent-elles le nom de *théorie* ? Les démonstrations sont encore absentes, l'équation finale cherchée n'est pas caractérisée autrement que par le fait de ne plus contenir les inconnues éliminées. La théorie n'achève donc de naître qu'à la fin du XVIII$^{\text{ème}}$ s., avec Lagrange, Waring, Bézout et Poisson. Enfin une troisième période s'étend au moins jusqu'au premier quart du XX$^{\text{ème}}$ siècle, période au cours de laquelle la théorie de l'élimination joue un rôle fondamental en géométrie comme en algèbre. Bien que les recherches de certains auteurs (tels Hesse et Cayley) continuent celles des auteurs de la période précédente, ceux-ci appartiennent au passé et sont souvent oubliés.

Etrangement, ce progrès de presque trois siècles s'arrête sur une image déchue de la théorie de l'élimination à partir des années 1940 : l'« élimination de l'élimination » dont témoignent de nombreux auteurs. A quelle réalité historique correspond-elle ? A défaut d'une réponse définitive à cette question, nous souhaitons seulement apporter ici quelques éléments de refléxion. Cette image déchue est surtout attestée par ses opposants. Certains d'entre eux semblent y réagir par l'idée et le désir d'un certain retour aux anciens et à leurs méthodes plus algorithmiques [34] (polygone de Newton, formule de Poisson, méthode d'élimination de Kronecker).

D'autres semblent y voir l'effet d'un manque de rigueur dans les travaux du passé, auquel on peut trouver remède en traduisant ces travaux dans un langage plus solide. Nous avons vu combien il avait fallu attendre une théorie rigoureuse du cas pourtant élémentaire de l'intersection de deux courbes algébriques planes (*supra* p. 213–215) ; et il est vrai qu'en 1916, Macaulay avouait encore craindre de tomber dans le travers de ses prédecesseurs en offrant au public un ouvrage semé d'erreurs. Il écrivait :

> *Le sujet est plein de pièges. J'ai indiqué certaines erreurs commises par d'autres, mais je ne doute pas en avoir moi-même commis de nouvelles.* On peut s'attendre à ce que toutes les erreurs soient découvertes et éliminées en temps utile, puisque des démonstrations et des références sont données pour tous les énoncés importants et pour la plupart des énoncés plus petits [35].

Macaulay avait ainsi décelé et corrigé plusieurs erreurs concernant la méthode d'élimination de Kronecker [36]. Mais ces erreurs ne sont-elles pas dues, autant qu'à un manque de rigueur, à la difficulté intrinsèque de la théorie ? Les méthodes et le langage même adoptés par Macaulay ne sont pas fondamentalement différents de ceux de ses prédecesseurs, et son manuel de 1916 a pourtant fait date. C'est peut-être l'ancêtre commun des traités d'algèbre commutative : tous le citent [37].

D'ailleurs, à la difficulté de la théorie s'ajoute la complexité des calculs d'élimination. Nous avons remarqué qu'à une autre époque, cette difficulté technique avait peut-être joué le rôle de ferment plus que d'obstacle [38]. Mais selon certains, il pourrait s'agir d'un facteur responsable de l'abandon provisoire des méthodes d'élimination au XX$^{\text{ème}}$ siècle : les méthodes connues, même effectives, étaient trop complexes pour faire réellement les calculs, même avec un ordinateur. Des recherches plus récentes visent justement à l'obtention de méthodes plus efficaces et à l'analyse des méthodes déjà connues pour en déterminer la complexité algorithmique [39].

D'autre part, certains auteurs ont remarqué les limites trop étroites d'une théorie de l'élimination dans les anneaux de polynômes sur un corps (de caractéristique nulle), théorie qui a d'ailleurs souvent recours à une clôture algébrique du corps de base, compliquant ainsi de manière artificielle la situation étudiée. Une telle théorie pouvait seulement suffire dans les débuts de la géométrie algébrique. Dans les années 1930–1940, celle-ci s'est progressivement étendue à des corps quelconques (plus nécessairement algébriquement clos, ni de caractéristique nulle), dans l'œuvre de Zariski, puis

34. *Cf. supra* note 8 p. 225.
35. Nous soulignons, *cf.* [217] p. vi.
36. *Cf. supra* n. 15 p. 239, p. 240 et n. 49 p. 251.
37. D'ailleurs la dénomination d'« anneau de Cohen-Macaulay » garde la trace des recherches de Macaulay sur la notion d'intersection complète, qui, comme nous l'avons vu, remonte à Kronecker.
38. *Cf. supra* p. 176.
39. *Cf.* la première page de [206], et les références que nous avons données *supra* n. 12 p. 178 et n. 14 p. 238.

celle de Weil. Survient enfin un renouveau global de cette discipline dans les années 1960–1970 avec l'œuvre de Grothendieck qui choisit, dans sa théorie des schémas, de travailler sur un anneau de base quelconque (et non plus seulement sur des corps). La formule de Poisson, par exemple, semblait fortement dépendre du cadre étroit de la théorie classique. J.-P. Jouanolou montre cependant, dans le langage de la théorie des schémas, comment interpréter cette formule sans faire intervenir les « solutions » des équations considérées. Soit donc un système de n équations non homogènes à $(n-1)$ inconnues

$$\begin{cases} P_1(x,y,...) = 0 \\ \vdots \\ P_n(x,y,...) = 0 \end{cases}$$

génériques sur un anneau de base k, et soit A l'anneau engendré sur k par les coefficients de ces équations. Soit ρ le résultant des formes dominantes de $P_1, ..., P_{n-1}$. Dans le cas où k est un corps, Poisson aurait écrit le produit [40]

$$\prod_i P_n(x_i, y_i, ...)$$

où les $(x_i, y_i, ...)$ sont les « solutions », dans une clôture algébrique du corps des fractions de A, des $(n-1)$ premières équations. Mais ce produit n'est autre que la norme $N(P_n)$ de la multiplication par P_n dans l'anneau

$$A_\rho[x,y,...]/(P_1, ..., P_{n-1})$$

et ce concept ne fait plus intervenir les « solutions » [41]. De même J.-P. Jouanolou construit toute sa théorie de l'élimination sur un anneau de base k qui n'est pas nécessairement un corps.

Plus généralement, il faut avouer que toute l'algèbre s'est transformée avec l'avénement des *structures* en mathématiques. Ce mouvement remonte peut-être au début du XIX$^{\text{ème}}$ siècle [42]. Nous en avons vu un exemple intéressant dans le mémoire de Halphen sur la détermination des coniques, où celui-ci découvre la structure de l'anneau d'intersection de la variété des coniques, en concevant l'opération « intersection » comme un *produit* dans cet anneau. La théorie de l'élimination semble absente d'une telle construction achevée. On peut d'ailleurs se demander quelle place aurait eu la méthode d'élimination de Kronecker dans une « théorie arithmétique de la grandeur algébrique » achevée, si Kronecker était parvenu à accomplir le programme de recherches qu'il a décrit dans ses *Grundzüge*. Nous l'avons vu, il cherchait en effet, dans les dernières sections de son mémoire, une autre définition de la codimension, indépendante de sa méthode d'élimination. Il cherchait lui aussi à définir une opération « quelque chose de commun » (*Gemeinsames*) entre deux ou plusieurs variétés algébriques, généralisant à la fois l'opération « intersection » et celle de « plus grand commun diviseur » [43]. Plus tôt encore, l'idée de l'élimination comme recherche de « la relation qui doit exister entre les coefficients d'une fonction ou d'un système de fonctions pour que quelque circonstance particulière (ou singularité) puisse avoir lieu », idée que nous avons suivie chez Cayley plus haut dans ce chapitre, montrait déjà un décalage par rapport à la conception traditionnelle de l'élimination. Brill aussi, bien qu'il semble toujours percevoir la théorie de l'élimination comme pouvant être le fondement suffisant de la géométrie algébrique, participe à cette évolution : ses travaux sont motivés par le problème de l'inversion des intégrales abéliennes, problème qui s'inscrit dans l'étude de structures (telle la variété jacobienne) liées aux courbes algébriques.

Le concept de l'« inconnue » n'était donc plus l'unique et seul concept de l'algèbre. L'« élimination de l'élimination » était alors inévitable : l'élimination, opération corrélative de l'inconnue, ne pouvait plus être l'opération par excellence de l'algèbre. Il était naturel que certains tentent d'accélérer cette évolution, et d'autres, plus conservateurs ou plus prudents, d'y réagir, ou mieux, de chercher à traduire les travaux du passé dans les langages des théories contemporaines. Bien sûr, les théories mathématiques, comme tout objet de pensée, ne meurent pas. La théorie de l'élimination existe en tant que telle aujourd'hui encore, bien qu'elle n'ait plus le rôle fondamental qu'elle avait

40. *Cf. supra* section 7.1.
41. *Cf.* [168] prop. 2.7 p. 124.
42. *Cf.* [321].
43. *Cf. supra* section 15.6 p. 250.

dans les siècles passés. On pourrait parler d'une certaine « école contemporaine » en théorie de l'élimination. Singularité de cette école : certains de ses membres ont déjà une conscience aiguë de l'histoire (au moins récente) de leur discipline [44]. D'ailleurs, si notre ouvrage était à refaire en connaissance de ce fait, nous procéderions peut-être autrement. Un choix s'imposerait : il faudrait soit assimiler d'abord l'essence des travaux contemporains sur la question, afin de rédiger une histoire informée et globale, soit au contraire nous restreindre à la période antérieure à 1839, en laissant à des auteurs plus compétents le soin d'écrire eux-mêmes l'histoire récente de leur propre discipline.

Le fait que nous n'ayons pas fait ce choix a pour conséquence le caractère inachevé de notre travail, surtout concernant l'histoire récente à partir de la seconde moitié du XIX$^{\text{ème}}$ siècle. Néanmoins, ceci a eu une conséquence positive. Le temps consacré à une meilleure assimilation des travaux contemporains aurait bien sûr retardé nos recherches sur les périodes anciennes, recherches qui nous ont au contraire occupé dès le début. Nous aurions alors été confrontés à une aporie face à l'œuvre de Bézout, que seule une bonne connaissance des travaux anciens nous a permis de situer précisément à la charnière entre deux époques. Mais, d'autre part, une connaissance minimale des travaux contemporains restait aussi nécessaire. A les ignorer complètement, nous n'aurions pu, non plus, comprendre ni juger l'intérêt scientifique de l'ouvrage de Bézout. Cette connaissance nous a aussi permis de mieux cerner, rétrospectivement, certains concepts qui apparaissent d'abord de manière floue dans le développement historique de la théorie. Enfin et surtout, elle nous a permis de décrire et de commenter les principaux travaux, dans ce domaine, de quelques auteurs importants du XIX$^{\text{ème}}$ siècle : Jacobi, Hesse, Bonnet, Cayley, Sylvester, Weierstrass, Halphen, Noether, Kronecker.

44. Aux auteurs que nous avons déjà cités ailleurs, il convient d'ajouter W. Vogel. Son livre [331] contient un excellent chapitre historique.

Quatrième partie

Annexes

Annexe A

Le paragraphe § 62 de la *Théorie générale des équations algébriques*

[1] *On demande quel est le degré de l'équation finale résultante d'un nombre quelconque n d'équations de la forme* $(u^a...n)^t$*, et comprenant un pareil nombre d'inconnues.*

Concevons que les équations soient $(u^a...n)^t = 0$, $(u^{a'}...n)^{t'} = 0$, $(u^{a''}...n)^{t''} = 0$, $(u^{a'''}...n)^{t'''} = 0$, etc. et qu'ayant multiplié la première, par le polynôme $(u^A...n)^T$ qui ait les conditions mentionnées dans l'énoncé du Problème (V) [2], on fasse ensuite, à l'aide des $n-1$ autres équations, disparoître de l'équation-produit, tous les termes qu'il est possible d'en faire disparoître sans en introduire de nouveaux. Alors (61) il ne restera plus dans l'équation-produit $(u^{A+a}...n)^{T+t} = 0$, qu'un nombre de termes exprimé par

$$d^{n-1}[N(u^{A+a}...n)^{T+t}]...\left(\begin{array}{ccccc} T+t & & A+a & & \underset{,}{A}+\underset{,}{a} \\ -t', -t'', -t''', \text{etc.}: & & -a', -a'', -a''', \text{etc.}: & & -\underset{,}{a}', -\underset{,}{a}'', -\underset{,}{a}''', \text{etc.}: \\ & & & & \underset{,,}{A}+\underset{,,}{a} \\ & & & & -\underset{,,}{a}', -\underset{,,}{a}'', -\underset{,,}{a}''', \text{etc.}: \text{etc.} \end{array}\right)$$

Soit D le degré de l'équation finale, et par conséquent $D+1$ le nombre des termes dont elle est composée ; alors le nombre des termes qu'il reste à faire disparoître, est

$$d^{n-1}[N(u^{A+a}...n)^{T+t}]...\left(\begin{array}{ccccc} T+t & & A+a & & \underset{,}{A}+\underset{,}{a} \\ -t', -t'', -t''', \text{etc.} & : & -a', -a'', -a''', \text{etc.} & : & -\underset{,}{a}', -\underset{,}{a}'', -\underset{,}{a}''', \text{etc.} \quad : \quad \text{etc.} \end{array}\right) - D - 1.$$

Or, pour qu'on puisse les faire disparoître, il faut que le polynôme multiplicateur fournisse autant de coëfficiens plus un.

Mais eu égard au nombre de termes qu'on peut faire disparoître dans le polynôme multiplicateur, à l'aide des $n-1$ dernières équations, ce polynôme ne peut fournir qu'un nombre de coëfficiens

$$= d^{n-1}[N(u^A...n)^T]...\left(\begin{array}{ccccc} T & & A & & \underset{,}{A} \\ -t', -t'', -t''', \text{etc.} & : & -a', -a'', -a''', \text{etc.} & : & -\underset{,}{a}', -\underset{,}{a}'', -\underset{,}{a}''', \text{etc.} \quad : \quad \text{etc.} \end{array}\right),$$

on aura donc l'équation suivante

$$\begin{array}{l} d^{n-1}[N(u^{A+a}...n)^{T+t}]...\left(\begin{array}{ccccc} T+t & & A+a & & \underset{,}{A}+\underset{,}{a} \\ -t', -t'', -t''', \text{etc.} & : & -a', -a'', -a''', \text{etc.} & : & -\underset{,}{a}', -\underset{,}{a}'', -\underset{,}{a}''', \text{etc.} \quad : \quad \text{etc.} \end{array}\right) - D - 1 \\ = d^{n-1}[N(u^A...n)^T]...\left(\begin{array}{ccccc} T & & A & & \underset{,}{A} \\ -t', -t'', -t''', \text{etc.} & : & -a', -a'', -a''', \text{etc.} & : & -\underset{,}{a}', -\underset{,}{a}'', -\underset{,}{a}''', \text{etc.} \quad : \quad \text{etc.} \end{array}\right) - 1, \end{array}$$

1. Nous reproduisons dans cette annexe le paragraphe § 62 du traité de Bézout [29].
2. Ces conditions sont : $A - a' - a'' - ... + \underset{,}{A} - \underset{,}{a}' - \underset{,}{a}'' - ... + \text{etc.} > T - t' - t'' - ...$

d'où l'on tire

$$D = d^{n-1}[N(u^{A+a}...n)^{T+t}]... \left(\begin{array}{ccccc} T+t & & A+a & & \underset{,}{A}+\underset{,}{a} \\ -t', -t'', -t''', \text{etc.} & : & -a', -a'', -a''', \text{etc.} & : & -\underset{,}{a}', -\underset{,}{a}'', -\underset{,}{a}''', \text{etc.} \quad : \quad \text{etc.} \end{array} \right)$$
$$-d^{n-1}[N(u^{A}...n)^{T}]... \left(\begin{array}{ccccc} T & & A & & \underset{,}{A} \\ -t', -t'', -t''', \text{etc.} & : & -a', -a'', -a''', \text{etc.} & : & -\underset{,}{a}', -\underset{,}{a}'', -\underset{,}{a}''', \text{etc.} \quad : \quad \text{etc.} \end{array} \right),$$

c'est-à-dire,

$$D = d^{n}[N(u^{A+a}...n)^{T+t}]... \left(\begin{array}{ccccc} T+t & & A+a & & \underset{,}{A}+\underset{,}{a} \\ -t, -t', -t'', -t''', \text{etc.} & : & -a, -a', -a'', -a''', \text{etc.} & : & -\underset{,}{a}, -\underset{,}{a}', -\underset{,}{a}'', -\underset{,}{a}''', \text{etc.} \quad : \quad \text{etc.} \end{array} \right),$$

Donc si on différencie la quantité $N(u...n)^{T+t} - N(u...n)^{T+t-A-a-1} - N(u...n)^{T+t-\underset{,}{A}-\underset{,}{a}-1} - N(u...n)^{T+t-\underset{''}{A}-\underset{''}{a}-1}$, etc. qui (59) est la valeur de $N(u^{A+a}...n)^{T+t}$; si on la différencie n fois de suite selon les règles données (13 et 14) on aura enfin

$$\begin{aligned} D = tt't''t''', \text{etc.} & -(t-a).(t'-a').(t''-a'').(t'''-a'''), \text{etc.} \\ & -(t-\underset{,}{a}).(t'-\underset{,}{a}').(t''-\underset{,}{a}'').(t'''-\underset{,}{a}'''), \text{etc.} \\ & -(t-\underset{''}{a}).(t'-\underset{''}{a}').(t''-\underset{''}{a}'').(t'''-\underset{''}{a}'''), \text{etc.} \\ & -(t-\underset{'''}{a}).(t'-\underset{'''}{a}').(t''-\underset{'''}{a}'').(t'''-\underset{'''}{a}'''), \text{etc.} \\ & -\text{etc.} \end{aligned}$$

Chaque produit ayant autant de facteurs qu'il y a d'inconnues.

Donc si on n'a que deux inconnues, le degré de l'équation finale sera $D = tt' - (t-a).(t'-a') - (t-\underset{,}{a}).(t'-\underset{,}{a}')$.

Annexe B

Le *Mémoire sur l'élimination dans les équations algébriques* de Poisson

[1]Le degré de l'équation finale résultant de l'élimination de toutes les inconnues moins une, entre un nombre d'équations égal à celui des inconnues, ne peut être plus grand que le produit des exposants de ces équations ; et il est justement égal à ce produit, lorsque les équations données sont les plus générales de leurs degrés. Ce théorème important est dû à Bézout ; mais la manière dont il l'a démontré n'est ni directe, ni simple, et même elle n'est pas exempte de toute difficultés. Avant Bézout, ce théorème était connu pour le cas de deux équations seulement : Cramer en avait donné une démonstration que l'on trouve maintenant dans les livres élémentaires, et qui ne laisse rien à désirer. Je me propose ici de l'étendre au cas d'un nombre quelconque d'équations ; mais, pour simplifier, je n'en considérerai que quatre, attendu qu'il sera facile d'appliquer au cas le plus général, les raisonnements que je ferai sur ce cas particulier. Représentons donc par (a),(b),(c) et (m) quatre équations ; la première du degré a, la seconde du degré b, la troisième du degré c, et la quatrième du degré m, entre les quatres inconnues x,y,z et u ; et supposons que ces équations sont complètes, c'est-à-dire que (a), par exemple, désigne l'équation la plus générale du degré a, à quatre inconnues. Si l'on élimine successivement, entre les trois premières, z et y, z et x, y et x, on obtiendra trois équations du même degré, la première entre x et u, la seconde entre y et u, et la troisième entre z et u ; et si l'on désigne par n le degré de ces équations, elles donneront pour chacune des inconnues x,y et z, un nombre n de valeurs en fonction de u. Soient a_1,a_2,a_3...a_n les valeurs de x ; b_1,b_2,b_3...b_n celles de y ; c_1,c_2,c_3...c_n celles de z ; si l'on prenait indistinctement une valeur de x, une de y et une de z, ces trois valeurs ne satisferaient pas ensemble aux équations (a),(b) et (c) : supposons donc que a_1,b_1 et c_1, a_2, b_2 et c_2, et, en général, a_p,b_p et c_p sont les valeurs de x,y et z, qui satisfont ensemble à ces équations ; et concevons que l'on substitue successivement chacun de ces système de valeurs à la place de x,y et z dans le premier membre de l'équation (m), le second étant supposé zéro, on obtiendra par là un nombre n de fonctions irrationnelles de u ; et il est facile de voir, 1° que toute quantité qui, mise à la place de u dans l'une quelconque de ces fonctions, la rendra nulle, sera une des valeurs de l'inconnue u ; 2° que réciproquement chacune des valeurs que peut avoir cette inconnue, doit rendre nulle une de ces fonctions. Il suit de là que si l'on égale à zéro le produit de toutes ces fonctions, l'équation qu'on obtiendra, sera, sans aucun facteur étranger à la question, l'équation finale résultant de l'élimination de x,y et z entre les équations données. C'est donc de cette équation qu'il s'agit de déterminer le degré.

Or, le premier membre de l'équation (m) est de la forme $u^m+Au^{m-1}+Bu^{m-2}+Cu^{m-3}+...+K$, A,B,C,...,K représentant des polynomes en x,y et z, savoir, A le polynome le plus général du premier degré, B le polynome le plus général du second degré, et ainsi de suite ; et il en résulte d'abord que le

1. « Mémoire sur l'élimination dans les équations algébriques », Poisson, *Journal de l'école polytechnique*, tome IV, 11e cahier (juin 1802), p. 199.

premier terme du produit des fonctions en question sera u^{mn}, et que, dans tous les autres termes, la somme des exposants de u et des valeurs ci-dessus de x,y et z, ne sera jamais plus grande que mn. De plus, sans effectuer le produit de ces fonctions, on voit qu'il doit être tel qu'il reste le même lorsqu'on y change a_p en a_q, b_p en b_q, c_p en c_q et réciproquement, quelque soient les indices p et q ; car si l'on opérait ces changements dans les fonctions mêmes, on ne ferait rien autre chose qu'échanger entre elles deux de ces fonctions. Si donc ce produit doit renfermer un terme comme $u^k a_p^\alpha b_p^\beta c_p^\gamma a_{p'}^{\alpha'} b_{p'}^{\beta'} c_{p'}^{\gamma'}$etc. (dans lequel k,α,β,γ,α',etc. représentent des exposants entiers positifs, dont la somme est moindre que mn, ou égale à ce nombre, et où p,p',p'',etc. sont des indices différents entre eux), il devra aussi contenir tous les termes que l'on peut déduire de celui-là, en y mettant successivement pour p,p',p'',etc., tous les nombres possibles depuis 1 jusqu'à n inclusivement, pourvu que l'on ne mette jamais dans un même terme, le même nombre pour deux de ces indices. La somme de tous ces termes est une certaine fonction de u, que je désignerai désormais par $u^k \int a_p^\alpha b_p^\beta c_p^\gamma a_{p'}^{\alpha'} b_{p'}^{\beta'} c_{p'}^{\gamma'}$etc., et dont il faut déterminer la forme.

Pour cela, soit l'équation $t = x + gy + hz$, dans laquelle t est une nouvelle inconnue, et g et h sont des coefficients arbitraires. En éliminant x,y et z entre cette équation et les équations (a), (b) et (c), on obtiendra une équation du degré n en u et t, que je représenterai par $t^n + Mt^{n-1} + Nt^{n-2} + ...S = 0$, en désignant par M,N,...,S, des fonctions rationnelles et entières de u, et le plus haut exposant de cette inconnue étant un dans M, deux dans N, et ainsi de suite. Les n racines de cette équation, seront $a_1 + gb_1 + hc_1$, $a_2 + gb_2 + hc_2$,...$a_n + gb_n + hc_n$; en sorte que $\int (a_p + gb_p + hc_p)^l$, qui désigne la somme des puissances l de ces quantités, sera, l étant un nombre entier et positif, une fonction rationnelle et entière des coefficients M,N,etc. : ce sera donc une fonction rationnelle et entière de u, et il est facile de s'assurer que le plus haut exposant de u dans cette fonction, sera égal à l. Mais on conçoit que g et h étant des quantités arbitraires et indépendantes de u, ce résultat ne peut avoir lieu pour la fonction $\int (a_p + gb_p + hc_p)^l$, sans qu'il ait aussi lieu pour chaque terme de cette fonction développée suivant les puissances et les produits des quantités g et h ; or le terme général de ce développement peut être représenté par $Lg^\beta h^\gamma \int a_p^\alpha b_p^\beta c_p^\gamma$, α, β et γ étant des exposants entiers et positifs, dont la somme est égale à l, et L étant une certaines fonction de ces nombres ; on peut donc dire que $\int a_p^\alpha b_p^\beta c_p^\gamma$ est une fonction rationnelle et entière de u, dans laquelle le plus haut exposant de cette inconnue est égal à $\alpha + \beta + \gamma$. Il suit de là que la fonction ci- dessus $u^k \int a_p^\alpha b_p^\beta c_p^\gamma a_{p'}^{\alpha'} b_{p'}^{\beta'} c_{p'}^{\gamma'}$etc. est une fonction rationnelle et entière de u, dans laquelle le plus haut exposant de cette inconnue est égal à la somme des exposants k,α,β,etc. ; car on a l'équation

$$\int a_p^\alpha b_p^\beta c_p^\gamma . \int a_{p'}^{\alpha'} b_{p'}^{\beta'} c_{p'}^{\gamma'} = \int a_p^{\alpha+\alpha'} b_p^{\beta+\beta'} c_p^{\gamma+\gamma'} + \int a_p^\alpha b_p^\beta c_p^\gamma a_{p'}^{\alpha'} b_{p'}^{\beta'} c_{p'}^{\gamma'}$$

qui fait voir que si cette proposition est démontrée pour la fonction $\int a_p^\alpha b_p^\beta c_p^\gamma$, elle le sera aussi pour la fonction $\int a_p^\alpha b_p^\beta c_p^\gamma a_{p'}^{\alpha'} b_{p'}^{\beta'} c_{p'}^{\gamma'}$ et ainsi de suite.

Cela posé, on a vu plus haut que le premier membre de l'équation finale en u (le second étant zéro), a pour premier terme u^{mn}, et que le reste de ce premier membre ne peut être autre chose que la somme d'un nombre fini de fonctions pareilles à la précédente, multipliées chacune par un coefficient connu, et dans lesquelles la somme des exposants k,α,β,etc., n'est jamais plus grande que mn ; ce premier membre est donc une fonction rationnelle de u, dans laquelle mn est le plus haut exposant de cette inconnue, et par conséquent ce nombre mn marque le degré de l'équation finale. On prouvera de même, en généralisant les raisonnements ci- dessus, et en les appliquant à un nombre quelconque d'équations complètes, que si n désigne le degré de l'équation finale résultante de toutes les équations moins une, et que m soit le degré de celle-ci, le degré de l'équation finale résultante de toutes les équations, sera égal à mn ; d'où il est facile de conclure que ce degré sera égal au produit des exposants des équations données, ainsi qu'on l'a annoncé au commencement. Dans tout ce qui précède, on a supposé que les équations entre lesquelles il s'agissait d'éliminer, étaient les plus générales de leur degré ; mais lorsque cela n'aura pas lieu, on conçoit que le degré de l'équation finale pourra être moindre que celui qu'on vient de déterminer, c'est-à-dire que les coefficients des plus hautes puissances de l'inconnue dans l'équation finale, pourront se trouver nuls en vertu des valeurs particulières des coefficients des inconnues dans les équations données. On doit donc dire en général que le degré de l'équation finale résultante d'un nombre quelconque d'équations, ne peut jamais être

plus grand que le produit des exposants de ces équations. Je me proposais seulement de déterminer à priori, le degré de l'équation finale résultante d'un nombre quelconque d'équations complètes ; mais on a pu voir aussi qu'en suivant la marche indiquée par l'analyse précédente, on parviendrait à former cette équation finale. Cette analyse nous fournit donc une méthode d'élimination, qui n'introduit dans l'équation finale aucun facteur étranger à la question, et qui n'exige la résolution d'aucune équation ; en sorte que toutes les opérations sont exécutables dans l'état actuel de l'algèbre ; mais la longueur des calculs qu'entraîne cette méthode, la rendra presque toujours impraticable, et nous ne l'avons indiquée ici que parce qu'elle donne la solution la plus directe du problème de l'élimination. Ainsi, dans la pratique, pour obtenir l'équation finale à son véritable degré, il faudra recourir aux méthodes que l'on trouve exposées avec beaucoup de détails dans la Théorie des équations de Bézout.

Annexe C

Etude des facteurs superflus dans les méthodes du XVII^{e} siècle avec Maple

On note :

$$P := A_0 + A_1\,x + A_2\,x^2 + A_3\,x^3$$
$$p := a_0 + a_1\,x + a_2\,x^2 + a_3\,x^3$$
$$Q := A_0 + A_1\,x + A_2\,x^2 + A_3\,x^3 + A_4\,x^4$$
$$q := a_0 + a_1\,x + a_2\,x^2 + a_3\,x^3 + a_4\,x^4$$

On calcule d'abord le résultant (3,3) de P et p au moyen de la méthode dialytique de Sylvester :

$$\begin{aligned} R := {} & 3\,{a_3}^2\,{A_0}^2\,A_3\,a_0 - A_3\,{a_2}^2\,a_1\,A_1\,A_0 + A_3\,{a_2}^2\,a_0\,{A_1}^2 - 3\,a_3\,{A_0}^2\,a_1\,A_3\,a_2 - 3\,{a_3}^2\,A_0\,a_0\,A_2\,A_1 \\ & + a_3\,A_0\,a_0\,A_3\,a_2\,A_1 - a_3\,A_3\,a_0\,a_1\,A_2\,A_0 + 3\,{A_3}^2\,a_0\,a_1\,a_2\,A_0 + 3\,a_3\,A_3\,{a_0}^2\,A_2\,A_1 \\ & - 2\,{A_3}^2\,{a_0}^2\,a_2\,A_1 - a_1\,A_2\,{A_3}^2\,{a_0}^2 + {a_1}^2\,A_2\,A_3\,a_2\,A_0 + a_1\,a_3\,{A_2}^2\,a_0\,A_1 \\ & - a_1\,A_2\,a_0\,A_3\,a_2\,A_1 - 2\,a_3\,{A_1}^2\,a_1\,A_3\,a_0 + 2\,a_3\,A_1\,A_3\,{a_1}^2\,A_0 + {A_3}^2\,{a_1}^2\,A_1\,a_0 \\ & - 2\,{a_2}^2\,A_2\,A_0\,A_3\,a_0 + 2\,a_2\,a_3\,{A_2}^2\,a_0\,A_0 + a_2\,{A_2}^2\,{a_0}^2\,A_3 + a_2\,a_3\,A_2\,a_1\,A_1\,A_0 \\ & - a_2\,a_3\,A_2\,a_0\,{A_1}^2 + {A_3}^3\,{a_0}^3 + {a_3}^2\,{A_1}^3\,a_0 - a_3\,{A_2}^3\,{a_0}^2 - 3\,a_3\,A_0\,{A_3}^2\,{a_0}^2 \\ & + 2\,{a_3}^2\,{A_0}^2\,a_1\,A_2 - {a_1}^2\,a_3\,{A_2}^2\,A_0 + {a_3}^2\,A_1\,a_2\,{A_0}^2 - {a_3}^2\,{A_1}^2\,a_1\,A_0 - {A_3}^2\,{a_1}^3\,A_0 \\ & - {a_2}^2\,a_3\,A_2\,{A_0}^2 + A_3\,{a_2}^3\,{A_0}^2 - {a_3}^3\,{A_0}^3 \end{aligned}$$

On donne ensuite les procédures $F(n,i)$ et $f(n,i)$ pour calculer les coefficients des restes intermédiaires délivrés de leurs facteurs superflus dans la méthode d'élimination successive des plus hautes puissances de l'inconnue d'un système de deux équations de degré d (on utilise la formule de récurrence de mon lemme p. 188 ; ici $F(n,i)$ resp. $f(n,i)$ désigne le coefficient du terme de degré i du polynôme en x que je note là-bas F_{n-1} resp. Φ_{n-1}) :

```
F := proc(n, i)
    if n = 1 then A_i
    else
        if n = 2 then f(1, d) × F(1, i) − F(1, d) × f(1, i)
        else
            (f(n − 1, d − n + 2) × F(n − 1, i) − F(n − 1, d − n + 2) × f(n − 1, i))/f(n − 2, d − n + 3)
        fi
    fi
end
```

$f := \mathbf{proc}(n,\, i)$
  if $n = 1$ **then** a_i
  else
    if $n = 2$ **then** $\mathrm{f}(1,\, d-1) \times (\mathrm{f}(1,\, d) \times \mathrm{F}(1,\, i) - \mathrm{F}(1,\, d) \times \mathrm{f}(1,\, i)) - \mathrm{f}(1,\, d) \times$
      $(\mathrm{f}(1,\, d) \times \mathrm{F}(1,\, i-1) + \mathrm{f}(1,\, d-1) \times \mathrm{F}(1,\, i) - \mathrm{F}(1,\, d) \times \mathrm{f}(1,\, i-1) - \mathrm{F}(1,\, d-1) \times \mathrm{f}(1,\, i))$
    else(
      $\mathrm{f}(n-1,\, d-n+1) \times (\mathrm{f}(n-1,\, d-n+2) \times \mathrm{F}(n-1,\, i) - \mathrm{F}(n-1,\, d-n+2) \times \mathrm{f}(n-1,\, i))$
      $- \mathrm{f}(n-1,\, d-n+2) \times (\mathrm{f}(n-1,\, d-n+2) \times \mathrm{F}(n-1,\, i-1)$
      $+ \mathrm{f}(n-1,\, d-n+1) \times \mathrm{F}(n-1,\, i) - \mathrm{F}(n-1,\, d-n+2) \times \mathrm{f}(n-1,\, i-1)$
      $- \mathrm{F}(n-1,\, d-n+1) \times \mathrm{f}(n-1,\, i)))/(\mathrm{F}(n-1,\, d-n+2) \times \mathrm{f}(n-2,\, d-n+3))$
    fi
  fi
end

On vérifie par exemple que $f(3, i)$ est entier :

$$
\begin{aligned}
& a_{d-1}\, {A_d}^2\, a_{d-2}\, a_{i-1} + {A_d}^2\, a_{d-3}\, a_{d-1}\, a_i - {a_{d-1}}^2\, A_{d-2}\, A_d\, a_i - {a_d}^2\, A_{d-2}\, A_d\, a_{i-2} \\
& \quad + {a_d}^2\, A_{d-3}\, A_d\, a_{i-1} + {a_d}^2\, A_{d-2}\, a_{d-1}\, A_{i-1} + {a_d}^2\, A_d\, a_{d-3}\, A_{i-1} \\
& \quad - a_d\, {A_d}^2\, a_{d-3}\, a_{i-1} + a_d\, {A_d}^2\, a_{d-2}\, a_{i-2} + a_d\, {a_{d-1}}^2\, A_d\, A_{i-2} \\
& \quad + {A_{d-1}}^2\, a_d\, a_{d-1}\, a_{i-1} - {A_{d-1}}^2\, a_d\, a_{d-2}\, a_i - A_{d-1}\, {a_d}^2\, A_{d-2}\, a_{i-1} \\
& \quad - A_{d-1}\, {a_d}^2\, a_{d-1}\, A_{i-2} + A_{d-1}\, {a_d}^2\, A_{d-3}\, a_i + A_{d-1}\, {a_d}^2\, a_{d-2}\, A_{i-1} \\
& \quad - A_{d-1}\, a_d\, {a_{d-1}}^2\, A_{i-1} - A_{d-1}\, {a_{d-1}}^2\, A_d\, a_{i-1} - {a_d}^2\, A_d\, a_{d-2}\, A_{i-2} \\
& \quad + 2\, a_d\, a_{d-2}\, A_{d-2}\, A_d\, a_i - 2\, a_d\, a_{d-1}\, A_d\, a_{d-2}\, A_{i-1} - a_d\, A_{d-3}\, a_{d-1}\, A_d\, a_i \\
& \quad - A_{d-1}\, a_d\, A_d\, a_{d-3}\, a_i + A_{d-1}\, a_d\, a_{d-1}\, A_{d-2}\, a_i + A_{d-1}\, a_d\, a_{d-1}\, A_d\, a_{i-2} \\
& \quad + A_{d-1}\, a_{d-1}\, A_d\, a_{d-2}\, a_i + {a_{d-1}}^3\, A_d\, A_{i-1} - {a_{d-2}}^2\, {A_d}^2\, a_i - {a_{d-1}}^2\, {A_d}^2\, a_{i-2} \\
& \quad - {a_d}^3\, A_{d-3}\, A_{i-1} + {a_d}^3\, A_{d-2}\, A_{i-2} - {a_d}^2\, {A_{d-2}}^2\, a_i
\end{aligned}
$$

On applique ces procédures au calcul du résultant (3,3). Bien-sûr, il ne faut pas tenir compte des termes dans lesquels il y a des coefficients d'indices négatifs :

```
> simplify(subs(d=3,a[-1]=0,A[-1]=0,a[-2]=0,A[-2]=0,F(4,0)));is(%=R);
```

$$
\begin{aligned}
& 3\, {a_3}^2\, {A_0}^2\, A_3\, a_0 - A_3\, {a_2}^2\, a_1\, A_1\, A_0 + A_3\, {a_2}^2\, a_0\, {A_1}^2 - 3\, a_3\, {A_0}^2\, a_1\, A_3\, a_2 - 3\, {a_3}^2\, A_0\, a_0\, A_2\, A_1 \\
& \quad + a_3\, A_0\, a_0\, A_3\, a_2\, A_1 - a_3\, A_3\, a_0\, a_1\, A_2\, A_0 + 3\, {A_3}^2\, a_0\, a_1\, a_2\, A_0 + 3\, a_3\, A_3\, {a_0}^2\, A_2\, A_1 \\
& \quad - 2\, {A_3}^2\, {a_0}^2\, a_2\, A_1 - a_1\, A_2\, {A_3}^2\, {a_0}^2 + {a_1}^2\, A_2\, A_3\, a_2\, A_0 + a_1\, a_3\, {A_2}^2\, a_0\, A_1 \\
& \quad - a_1\, A_2\, a_0\, A_3\, a_2\, A_1 - 2\, a_3\, {A_1}^2\, a_1\, A_3\, a_0 + 2\, a_3\, A_1\, A_3\, {a_1}^2\, A_0 + {A_3}^2\, {a_1}^2\, A_1\, a_0 \\
& \quad - 2\, {a_2}^2\, A_2\, A_0\, A_3\, a_0 + 2\, a_2\, a_3\, {A_2}^2\, a_0\, A_0 + a_2\, {A_2}^2\, {a_0}^2\, A_3 + a_2\, a_3\, A_2\, a_1\, A_1\, A_0 \\
& \quad - a_2\, a_3\, A_2\, a_0\, {A_1}^2 + {A_3}^3\, {a_0}^3 + {a_3}^2\, {A_1}^3\, a_0 - a_3\, {A_2}^3\, {a_0}^2 - 3\, a_3\, A_0\, {A_3}^2\, {a_0}^2 \\
& \quad + 2\, {a_3}^2\, {A_0}^2\, a_1\, A_2 - {a_1}^2\, a_3\, {A_2}^2\, A_0 + {a_3}^2\, A_1\, a_2\, {A_0}^2 - {a_3}^2\, {A_1}^2\, a_1\, A_0 - {A_3}^2\, {a_1}^3\, A_0 \\
& \quad - {a_2}^2\, a_3\, A_2\, {A_0}^2 + A_3\, {a_2}^3\, {A_0}^2 - {a_3}^3\, {A_0}^3
\end{aligned}
$$

true

On vérifie que le numérateur de l'expression obtenue par élimination successive des plus hautes puissances de l'inconnue est égal au résultant (3,3). A cet effet, on définit au préalable la procédure *elim* :

$elim := \mathbf{proc}(P,\, Q)$
  $\mathrm{collect}(P - \mathrm{lcoeff}(P,\, x) \times x^{(\mathrm{degree}(P,\, x) - \mathrm{degree}(Q,\, x))} \times Q/\mathrm{lcoeff}(Q,\, x),\, x)$
end

```
> F2:=elim(P,p): f2:=elim(p,F2): F3:=elim(F2,f2): f3:=elim(f2,F3):
final:=normal(elim(F3,f3)): simplify(numer(final));is(-%=R);
```

$$
\begin{aligned}
&-3\,{a_3}^2\,{A_0}^2\,A_3\,a_0 + A_3\,{a_2}^2\,a_1\,A_1\,A_0 - A_3\,{a_2}^2\,a_0\,{A_1}^2 + 3\,a_3\,{A_0}^2\,a_1\,A_3\,a_2 + 3\,{a_3}^2\,A_0\,a_0\,A_2\,A_1\\
&\quad - a_3\,A_0\,a_0\,A_3\,a_2\,A_1 + a_3\,A_3\,a_0\,a_1\,A_2\,A_0 - 3\,{A_3}^2\,a_0\,a_1\,a_2\,A_0 - 3\,a_3\,A_3\,{a_0}^2\,A_2\,A_1\\
&\quad + 2\,{A_3}^2\,{a_0}^2\,a_2\,A_1 + a_1\,A_2\,{A_3}^2\,{a_0}^2 - {a_1}^2\,A_2\,A_3\,a_2\,A_0 - a_1\,a_3\,{A_2}^2\,a_0\,A_1\\
&\quad + a_1\,A_2\,a_0\,A_3\,a_2\,A_1 + 2\,a_3\,{A_1}^2\,a_1\,A_3\,a_0 - 2\,a_3\,A_1\,A_3\,{a_1}^2\,A_0 - {A_3}^2\,{a_1}^2\,A_1\,a_0\\
&\quad + 2\,{a_2}^2\,A_2\,A_0\,A_3\,a_0 - 2\,a_2\,a_3\,{A_2}^2\,a_0\,A_0 - a_2\,{A_2}^2\,{a_0}^2\,A_3 - a_2\,a_3\,A_2\,a_1\,A_1\,A_0\\
&\quad + a_2\,a_3\,A_2\,a_0\,{A_1}^2 - {A_3}^3\,{a_0}^3 - {a_3}^2\,{A_1}^3\,a_0 + a_3\,{A_2}^3\,{a_0}^2 + 3\,a_3\,A_0\,{A_3}^2\,{a_0}^2\\
&\quad - 2\,{a_3}^2\,{A_0}^2\,a_1\,A_2 + {a_1}^2\,a_3\,{A_2}^2\,A_0 - {a_3}^2\,A_1\,a_2\,{A_0}^2 + {a_3}^2\,{A_1}^2\,a_1\,A_0 + {A_3}^2\,{a_1}^3\,A_0\\
&\quad + {a_2}^2\,a_3\,A_2\,{A_0}^2 - A_3\,{a_2}^3\,{A_0}^2 + {a_3}^3\,{A_0}^3
\end{aligned}
$$

true

On détermine le facteur superflu introduit par la méthode d'élimination de Fermat dans le calcul du résultant (3,3). A cet effet, on définit au préalable la procédure *elim1* :

$$\mathit{elim1} := \mathbf{proc}(P,\,Q)\,\mathrm{collect}((\mathrm{tcoeff}(Q,\,x)\times P - \mathrm{tcoeff}(P,\,x)\times Q)/x,\,x)\,\mathbf{end}$$

```
> P3:=elim1(P,p): P4:=elim1(p,P3): P5:=elim1(P3,P4): P6:=elim1(P4,P5):
final:=elim1(P5,P6): factor(final/R);
```

$$(-A_1\,a_0 + A_0\,a_1)\,(a_1\,A_1\,a_0 - {a_1}^2\,A_0 - {a_0}^2\,A_2 + a_2\,A_0\,a_0)\,{a_0}^2$$

On détermine le facteur superflu introduit par la méthode d'élimination de Leibniz-Euler dans le calcul du résultant (3,3). A cet effet, on définit au préalable la procédure *elim2* :

$$\mathit{elim2} := \mathbf{proc}(P,\,Q)\,\mathrm{collect}(\mathrm{lcoeff}(Q,\,x)\times P - \mathrm{lcoeff}(P,\,x)\times Q,\,x)\,\mathbf{end}$$

```
> P3:=elim2(P,p): P4:=elim1(P,p): P5:=elim2(P3,P4): P6:=elim1(P3,P4):
final:=elim2(P5,P6): factor(final/R);
```

$$a_3\,A_0 - A_3\,a_0$$

On détermine le facteur superflu introduit par l'algorithme euclidien dans le calcul du résultant (3,3). On va observer un facteur superflu (numérateur et dénominateur) qui contient deux carrés parfaits, et qui est de la forme prédite par Sylvester. Il faut avant tout définir la procédure *eucl* :

$$\begin{aligned}
\mathit{eucl} := \mathbf{proc}&(P,\,Q)\\
&\mathrm{collect}(P - \mathrm{lcoeff}(P,\,x)\times x^{(\mathrm{degree}(P,\,x)-\mathrm{degree}(Q,\,x))}\times Q/\mathrm{lcoeff}(Q,\,x),\,x)\\
&\mathbf{end}
\end{aligned}$$

```
> P3:=eucl(P,p): P4:=eucl(p,P3): P5:=eucl(P4,P3): P6:=eucl(P3,P5):
final:=eucl(P6,P5): factor(final/R);
```

$$
\begin{aligned}
&-(a_3\,A_2 - A_3\,a_2)^2 \Big/ ((a_2\,a_3\,A_2\,A_1 + a_2\,A_2\,A_3\,a_1 - A_3\,{a_2}^2\,A_1 - {a_3}^2\,{A_1}^2 + 2\,a_3\,A_1\,A_3\,a_1\\
&\quad - {A_3}^2\,{a_1}^2 + {a_3}^2\,A_2\,A_0 - a_3\,A_2\,A_3\,a_0 - a_3\,{A_2}^2\,a_1 - A_3\,a_2\,a_3\,A_0 + {A_3}^2\,a_2\,a_0)^2 a_3)
\end{aligned}
$$

A présent, on applique l'élimination des plus hautes puissances de l'inconnue au calcul du résultant (4,3), et on vérifie qu'à nouveau, le numérateur ne contient pas de facteur superflu :

```
> P3:=elim(Q,p): P4:=elim(p,P3): P5:=elim(P3,P4): P6:=elim(P4,P5): P7:=elim(P5,P6):
final:=normal(elim(P6,P7)): simplify(numer(final));is(-%=resultant(Q,p,x));
```

$$
\begin{aligned}
&2\,A_2\,{a_3}^3\,{A_0}^2\,a_1 - {a_2}^4\,{A_0}^2\,A_4 + a_3\,{A_3}^3\,{a_0}^3 + {a_3}^3\,a_0\,{A_1}^3 + 3\,{a_3}^3\,{A_0}^2\,a_0\,A_3 - 3\,{a_3}^2\,A_0\,{A_3}^2\,{a_0}^2\\
&\quad - 2\,{a_3}^2\,{A_1}^2\,a_1\,A_3\,a_0 + {a_3}^2\,A_0\,a_0\,A_3\,a_2\,A_1 + 3\,a_3\,{A_3}^2\,a_0\,a_1\,a_2\,A_0 + a_3\,{A_3}^2\,{a_1}^2\,A_1\,a_0\\
&\quad - A_2\,a_0\,A_3\,a_1\,{a_3}^2\,A_0 - A_2\,{A_3}^2\,a_1\,a_3\,{a_0}^2 + 3\,A_2\,{a_3}^2\,A_1\,A_3\,{a_0}^2 - 2\,A_2\,a_0\,{a_2}^2\,a_3\,A_0\,A_3\\
&\quad - A_2\,a_3\,a_2\,a_0\,a_1\,A_3\,A_1 + 4\,A_0\,{A_4}^2\,{a_1}^2\,a_0\,a_2 + 4\,a_3\,{a_2}^2\,{A_0}^2\,a_1\,A_4 - a_3\,A_0\,A_4\,a_0\,{a_2}^2\,A_1\\
&\quad + 3\,a_3\,a_0\,{A_1}^2\,a_1\,A_4\,a_2 - 3\,a_3\,A_0\,A_1\,a_2\,A_4\,{a_1}^2 - 4\,a_3\,A_0\,{A_4}^2\,{a_0}^2\,a_1\\
&\quad + 5\,{a_3}^2\,A_0\,A_1\,A_4\,a_0\,a_1 - 4\,{a_3}^2\,A_4\,a_0\,a_2\,{A_0}^2 + A_0\,{a_2}^3\,A_1\,a_1\,A_4 - A_0\,{A_4}^2\,{a_1}^4\\
&\quad - 3\,{a_0}^2\,A_1\,{A_4}^2\,a_1\,a_2 - {A_0}^3\,{a_3}^4 - {A_4}^3\,{a_0}^4 - 2\,{A_3}^2\,a_3\,a_2\,{a_0}^2\,A_1 - {A_3}^2\,a_3\,A_0\,{a_1}^3\\
&\quad - 3\,A_3\,A_0\,{a_2}^2\,A_4\,a_0\,a_1 - A_3\,a_0\,A_1\,a_2\,A_4\,{a_1}^2 + A_3\,A_0\,a_2\,A_4\,{a_1}^3 - A_3\,a_3\,A_0\,{a_2}^2\,A_1\,a_1\\
&\quad - A_3\,a_3\,A_0\,a_0\,{a_1}^2\,A_4 + A_3\,a_3\,a_0\,{A_1}^2\,{a_2}^2 + A_3\,{A_4}^2\,a_1\,{a_0}^3 + 5\,A_3\,a_3\,A_0\,A_4\,{a_0}^2\,a_2\\
&\quad + A_3\,a_3\,A_4\,a_1\,{a_0}^2\,A_1 + 2\,A_3\,{a_3}^2\,A_0\,{a_1}^2\,A_1 - {A_3}^2\,a_2\,A_4\,{a_0}^3 + A_3\,a_3\,{a_2}^3\,{A_0}^2\\
&\quad - 3\,A_3\,{a_3}^2\,a_2\,{A_0}^2\,a_1 + {A_2}^2\,a_3\,{a_0}^2\,A_3\,a_2 - 2\,A_0\,{A_4}^2\,{a_0}^2\,{a_2}^2 + 2\,A_3\,{a_2}^2\,{a_0}^2\,A_4\,A_1\\
&\quad + a_0\,A_1\,{A_4}^2\,{a_1}^3 - {a_0}^2\,{a_3}^2\,{A_2}^3 - a_0\,{a_2}^3\,{A_1}^2\,A_4 - {a_3}^3\,A_0\,{A_1}^2\,a_1 + {a_3}^3\,{A_0}^2\,A_1\,a_2\\
&\quad - {A_2}^2\,{a_2}^2\,{a_0}^2\,A_4 + 2\,{A_2}^2\,{a_3}^2\,A_0\,a_2\,a_0 + 2\,{A_2}^2\,{a_0}^2\,A_4\,a_1\,a_3 + {A_2}^2\,{a_3}^2\,a_0\,A_1\,a_1\\
&\quad + A_2\,a_0\,{a_2}^2\,A_1\,a_1\,A_4 + 2\,A_2\,{A_4}^2\,a_2\,{a_0}^3 - A_2\,a_0\,{a_0}^2\,{A_4}^2\,{a_1}^2 - {A_2}^2\,{a_3}^2\,A_0\,{a_1}^2\\
&\quad - 4\,A_2\,a_3\,A_0\,A_4\,a_0\,a_2\,a_1 + 2\,A_2\,a_3\,A_0\,A_4\,{a_1}^3 - 2\,A_2\,a_3\,a_0\,A_4\,{a_1}^2\,A_1\\
&\quad - A_2\,a_3\,{a_0}^2\,A_4\,a_2\,A_1 - A_2\,{a_3}^2\,a_0\,{A_1}^2\,a_2 - A_2\,A_0\,{a_2}^2\,{a_1}^2\,A_4 + A_2\,{a_3}^2\,A_0\,A_1\,a_2\,a_1\\
&\quad + 3\,A_2\,{a_3}^2\,A_0\,A_4\,{a_0}^2 + A_2\,A_3\,a_2\,{a_0}^2\,A_4\,a_1 - 3\,A_2\,{a_3}^3\,A_0\,a_0\,A_1 + A_2\,A_3\,a_3\,A_0\,a_2\,{a_1}^2\\
&\quad - 3\,A_2\,A_3\,a_3\,A_4\,{a_0}^3 - A_2\,{a_3}^2\,{a_2}^2\,{A_0}^2 - 3\,{a_3}^2\,{A_1}^2\,A_4\,{a_0}^2 + 2\,A_2\,A_0\,{a_2}^3\,a_0\,A_4\\
&\quad - 2\,{a_3}^2\,{A_0}^2\,A_4\,{a_1}^2 + 3\,a_3\,A_1\,{A_4}^2\,{a_0}^3
\end{aligned}
$$

true

Annexe D

Vérification expérimentale des formules de Bézout avec le logiciel Macaulay2

```
comb=(liste,i)->if i==0 then {{}} else
 if i>#liste then {} else
 comb(drop(liste,1),i)|apply(comb(drop(liste,1),i-1),c->c|{liste#0})
listcomb=liste->toList apply(0..#liste, i->comb(liste,i))
expKoszul=(about,functs)->apply(listcomb(functs),
 d->apply(d,c->if #c==0 then about else about-(sum c)))
```

Syntaxe : `expKoszul({6,5,5,5},{{3,1,1,1},{3,1,1,1},{3,1,1,1}})`

```
Koszul=(functs)->(bases=listcomb(toList(0..(#functs-1)));
 toList apply(0..#bases-2,i->(b1=bases#(i+1);b2=bases#i;
 transpose matrix toList apply(0..(#b1-1),
 j->toList apply(0..(#b2-1),k->if isSubset(b2#k,b1#j) then
 (pos=position(b1#j,v->(not member(v,b2#k)));
 (-1)^pos*(functs#(b1#j#pos))) else 0)))))
```

Pour ce qui suit, il faut avoir défini `K` (par exemple `K=ZZ/32749`) et `C` tel que la variable homogénéisante soit la dernière des variables (par exemple `C=K[x_0..x_2,z]`).
`randpolys(expfuncts,espece2)` génère les polynômes de secondes espèces à trois variables (plus la variable homogénéisante) dont les exposants sont donnés par la liste de listes `expfuncts`.

```
espece1=n->if (n#0<0) then {} else apply(select(exponents (sum
 gens C)^(n#0),m->((m#0 <=n#1) and (m#1 <= n#2) and (m#2 <=n#3))),m->C_m)
espece2=n->if (n#0<0) then {} else apply(select(exponents (sum
 gens C)^(n#0),m->((m#0 <= n#1) and (m#1 <= n#2) and (m#2 <= n#3)
 and (m#1+m#2 <= n#4) and (m#0+m#2 <= n#5) and (m#0+m#1 <= n#6))),m->C_m)
dimC=(n,espece)->#(espece(n))
randpolys=(expfuncts,espece)->apply(expfuncts,n->(monomes=espece(n);
 coefs=flatten entries random(K^1,K^#monomes);
 sum(#monomes,i->(monomes#i)*(coefs#i))))
```

`basesKoszul(abouts,functs,espece)` génère la liste des listes de bases des espaces vectoriels sur `K` du complexe de Koszul des fonctions d'exposants `functs` et d'espèce `espece`, et d'aboutissement `abouts`.

```
basesKoszul=(abouts,functs,espece)->apply(expKoszul(abouts,functs),
 degre->apply(degre, comp->(espece comp)))
```

`matr2matr(matrice, liste des bases d'arrivée, liste des bases de départ)` transforme une matrice de polynômes en une matrice à coefficients dans `K`. Cette procédure a recours à la fonction `transBase`, qui transforme les bases d'arrivée en `HashTable` pour les rendre plus facile à manipuler (ce n'est pas utile pour les bases de départ).

```
transBase=base->apply(base,M->(keys standardForm M)#0)
poly2matr=(P,B)->(S=standardForm P; apply(B, mon->
 if S#?mon then S#mon else 0))
matr2matr=(M,bases1,bases2)->(M2=matrix apply(entries M, ligne->flatten
 apply(ligne, bases2, (P,B)->apply(B, monome->P*monome)));
 bases1trans=apply(bases1, B->transBase B);transpose matrix
 apply(entries transpose M2, col->flatten apply(col, bases1trans,
 (poly,B)->poly2matr(poly,B))))
```

`dimKoszul(bases)` renvoie la liste des dimensions (linéaires) des différentes degrés d'un complexe dont on a obtenu des bases par la fonction `basesKoszul`. `altersum(liste)` renvoie alors la somme alternée de ces dimensions.

```
dimKoszul=b->apply(b,deg->#(flatten deg))
altersum=l->sum apply(l,0..#l-1,(e,n)->e*(-1)^n)
```

`dimCesp1(n)` calcule la dimension de l'espace de polynômes de première espèce C_n, au moyen de la formule donnée par Bézout au paragraphe 59. Il faut donc que les sommes deux à deux des trois derniers exposants dépassent le premier.

```
dimCesp1=n->binomial(3+n#0,3)-sum(1..3,i->binomial(3+n#0-n#i-1,3))
```

`tablesp1(D,M)` construit un tableau (`HashTable`) d'exposants d'équations de première espèce de degré $\leq$ `D` à moins de `M` termes. Les clefs sont des entiers à partir de 0.

```
perm=(d,i,j,k)->toList set({d,i,j,k},{d,j,k,i},{d,k,i,j},
 {d,k,j,i},{d,j,i,k},{d,i,k,j})
tablesp1=(D,M)->(liste={}; c=0; for i from 1 to D do (for j from
 i to D do (for k from j to D do
 for d from max(i,j,k) to min(i+j,i+k,j+k,D) do
 (m=dimCesp1 {d,i,j,k}; if m<=M then apply(perm(d,i,j,k),
 eq->(liste=liste|{c=>eq};c=c+1))))); new HashTable from liste)
```

`tablpseudoesp1(D,M)` construit un tableau (`HashTable`) d'exposants d'équations de "pseudo-première-espèce" (c'est-à-dire des exposants du type `{d,i,j,k}` avec `min(i+j,i+k,j+k)<d`) de degré $\leq$ `D` à moins de `M` termes. Les clefs sont des entiers à partir de 0.

```
tablpseudoesp1=(D,M)->(liste={}; c=0; for i from 1 to D do (for j
 from i to D do (for k from j to D do
 for d from max(min(i+j,i+k,j+k,D)+1,i,j,k) to min(D,i+j+k) do
 (m=dimC({d,i,j,k},espece1); if m<=M then apply(perm(d,i,j,k),
 eq->(liste=liste|{c=>eq};c=c+1))))); new HashTable from liste)
```

`tabl3(T)` construit un tableau à 3 entrées `(f1,f2,f3)` où `f1`, `f2`, `f3` sont des éléments du tableau `T` d'exposants, sans répéter les triplets obtenus par permutation des variables. Pour tester si `(f1,f2,f3)` s'obtiennent par permutation des variables à partir d'un des triplets précédemment ajoutés au tableau, on a recours à `perm3plet(f1,f2,f3)` qui construit un `Set` des triplets obtenus à partir de `(f1,f2,f3)` par permutation des variables.

```
perm3plet=(f1,f2,f3)->set apply({{1,2,3},{2,3,1},{3,1,2},
 {3,2,1},{2,1,3},{1,3,2}},p->apply((f1,f2,f3),
 f->{f#0,f#(p#0),f#(p#1),f#(p#2)}))
tabl3=T->(N=#T;liste={};setPerm=new Set from {};for i
```

```
 from 0 to N-1 do (for j from i to N-1 do (for k from j to N-1 do
 (el=(T#i,T#j,T#k); if ((i,j,k)==(0,0,0)) or (not member(el,setPerm))
 then (el=(T#i,T#j,T#k);liste=liste|{(i,j,k)=>el};
 setPerm=setPerm+perm3plet el))));
 new HashTable from liste)
```

Pour ce qui suit, il faut choisir `K=ZZ/32749;C=K[x_0..x_2,z,MonomialOrder=>Eliminate 2]`. `testBezoutEsp1(T,T3,D)` calcule un triplet de polynômes aléatoires (de première espèce) pour chaque entrée de `T3=tabl3 T`, puis calcule, pour ce triplet, la somme alternée du complexe construit en utilisant le "polynôme-générateur" d'exposants nuls (pour Bézout, il s'agit donc du degré de l'équation finale), ainsi que l'idéal résultant, en arrêtant le calcul au degré `D`. L'inconnue restante est la dernière des trois (ainsi que la variable homogénéisante `z`).

```
testBezoutEsp1=(T,T3,D)->(
 print "";
 N=#T;
 polys={};
 for i from 0 to N-1 do polys=polys|{i=>apply({0,1,2},
  j->(randpolys({T#i},espece1))#0)};
 polys=new HashTable from polys;
 for i from 0 to N-1 do (for j from i to N-1 do (for k from j to N-1 do
  (if T3#?(i,j,k) then
  (print concatenate("exposants: ",toString T3#(i,j,k));
  kos=expKoszul({0,0,0,0},T3#(i,j,k));
  polgen=-((kos#(#kos-1))#0);
  print concatenate("somme alternee: ",toString altersum apply(kos,
  degr->sum apply(degr,pol->dimCesp1(pol+polgen))));
  g=gb(ideal ((polys#i)#0,(polys#j)#1,(polys#k)#2),DegreeLimit => D);
  print concatenate("ideal resultant (en degre <=",toString D,"): ");
  print selectInSubring(1,gens g);
  print "";
  )))))
```

Pour l'essayer, utiliser par exemple le code suivant :

```
T=tablesp1(2,8);T3=tabl3 T;
testBezoutEsp1(T,T3,30);
```

`testBezoutPseudoesp1(T,T3,D)` fait la même chose que `tesBezoutEsp1`, mais peut travailler avec une table de polynômes de "pseudo-première-espèce". Elle a en effet recours a `dimC` plutôt qu'à `dimCesp1`, ce qui allonge d'ailleurs le temps de calcul (pour remédier à cela, il faudrait implémenter les formules données par Bézout pour le calcul du nombre de termes d'un polynôme de seconde espèce).

```
testBezoutPseudoesp1=(T,T3,D)->(
 print "";
 N=#T;
 polys={};
 for i from 0 to N-1 do polys=polys|{i=>apply({0,1,2},
  j->(randpolys({T#i},espece1))#0)};
 polys=new HashTable from polys;
 for i from 0 to N-1 do (for j from i to N-1 do (for k from j to N-1 do
  (if T3#?(i,j,k) then
  (print concatenate("exposants: ",toString T3#(i,j,k));
  kos=expKoszul({0,0,0,0},T3#(i,j,k));
  polgen=-((kos#(#kos-1))#0);
```

```
  print concatenate("somme alternee: ",toString altersum
   apply(kos,degr->sum apply(degr,pol->dimC(pol+polgen,espece1))));
  g=gb(ideal ((polys#i)#0,(polys#j)#1,(polys#k)#2),DegreeLimit => D);
  print concatenate("ideal resultant (en degre <=",toString D,"): ");
  print selectInSubring(1,gens g);
  print "";
  )))))
```

Pour l'essayer, utiliser par exemple le code suivant :

```
T=tablpseudoesp1(4,15);T3=tabl3 T;
testBezoutPseudoesp1(T,T3,50);
```

On implémente ensuite les formules de Bézout pour le nombre de termes d'un polynôme de seconde espèce de forme donnée.
Syntaxe : `Nesp2(forme,3,{t#0,t#1,t#2,t#3,t#4,t#5,t#6})`.

```
Ncomp=(n,t)->product(1..n,i->t+i)/product(1..n,i->i)
Nesp2=(forme,n,t)->
 if forme==1 then (Ncomp(n,t#0)-sum(1..3,i->Ncomp(n,t#0-t#i-1))
 +sum(4..6,i->Ncomp(n,t#0-t#i-2))-sum(1..3,i->(t#1+t#2+t#3-t#i-t#(3+i))
 *Ncomp(n-1,t#0-t#(3+i)-1))) else
 (if forme==2 then Nesp2(1,n,t)+Ncomp(n,t#0+t#3-t#5-t#4-2)
 else (if forme==3 then Nesp2(1,n,t)+Ncomp(n,t#0+t#1-t#6-t#5-2)
 else (if forme==4 then Nesp2(3,n,t)+Ncomp(n,t#0+t#3-t#5-t#4-2)
 else (if forme==5 then Nesp2(3,n,t)+Ncomp(n,t#0+t#2-t#6-t#4-2)
 else (if forme==6 then Nesp2(5,n,t)+Ncomp(n,t#0+t#3-t#5-t#4-2)
 else (if forme==7 then Nesp2(1,n,t)+Ncomp(n,t#0+t#2-t#6-t#4-2)
 else (if forme==8 then Nesp2(7,n,t)+Ncomp(n,t#0+t#3-t#5-t#4-2)))))))
```

`forme(t)` renvoie la liste des formes (entre 1 et 8) auquelles appartient un polynôme donné de seconde espèce.

```
forme=t->(l={};
 if (t#0-t#6)<=min(t#5-t#1,t#4-t#2) then
 (if (t#4-t#3)>=(t#0-t#5) then l={1};if (t#4-t#3)<=(t#0-t#5) then l=l|{2});
 if (t#0-t#6)>=(t#5-t#1) and (t#0-t#6)<=(t#4-t#2) then
 (if (t#0-t#5)<=(t#4-t#3) then l=l|{3};if (t#0-t#5)>=(t#4-t#3) then l=l|{4});
 if (t#0-t#6)>=max(t#5-t#1,t#4-t#2) then
 (if (t#0-t#5)<=(t#4-t#3) then l=l|{5}; if (t#0-t#5)>=(t#4-t#3) then l=l|{6});
 if (t#0-t#6)<=(t#5-t#1) and (t#0-t#6)>=(t#4-t#2) then
 (if (t#0-t#5)<=(t#4-t#3) then l=l|{7}; if (t#0-t#5)>=(t#4-t#3) then l=l|{8});
 l)
```

`pseudoesp1ToEsp2(t)` transforme un multi-indice `t` de pseudo-première-espèce en un multi-indice de seconde espèce :

```
pseudoesp1ToEsp2=t->t|{min(t#0,t#2+t#3),min(t#0,t#1+t#3),min(t#0,t#1+t#2)}
```

Voici comment tester `Nesp2` en cherchant le degré de l'équation finale $D = 17$ pour trois polynômes `C_{4,2,1,1},C_{3,1,3,1},C_{3,1,3,1}` :

```
T=tablpseudoesp1(4,20);T3=tabl3 T;
triplet=toList apply(T3#(4,8,8),t->pseudoesp1ToEsp2 t);
K=Frac(ZZ/32749[T0,A1,A2,A3,B1,B2,B3]);
kos=expKoszul({T0,A1,A2,A3,B1,B2,B3},triplet);
altersum apply(kos,degr->sum apply(degr,pol->Nesp2(1,3,pol)))
```

`test8formesPseudoesp1(T,T3,D)` fait la même chose que `testBezoutPseudoesp1`, mais a recours aux formules données par Bézout plutôt qu'à `dimCesp1` pour calculer la somme alternée. Elle affiche pour chacune des 8 formes la valeur de D trouvée, suivie d'un astérisque si cette forme est "admissible", au sens de Bézout.

```
test8formesPseudoesp1=(T,T3,D)->(
 print "";
 N=#T;
 polys={};
 for i from 0 to N-1 do polys=polys|{i=>apply({0,1,2},
  j->(randpolys({T#i},espece1))#0)};
 polys=new HashTable from polys;
 C'=QQ[S_0..S_6];
 for i from 0 to N-1 do (for j from i to N-1 do (for k from j to N-1 do
  (if T3#?(i,j,k) then
  (print concatenate("exposants: ",toString T3#(i,j,k));
  triplet=toList apply(T3#(i,j,k),t->pseudoesp1ToEsp2 t);
  for formes from 1 to 8 do
   (differentielles={};
   for multnd from 0 to 2 do
    (tri=apply(0..2,num->if num==multnd then
     toList(S_0..S_6) else triplet#num);
    kos=expKoszul({0,0,0,0,0,0,0},tri);
    differentielles=differentielles|values standardForm lift(
     altersum apply(kos,degr->sum apply(degr,pol->Nesp2(formes,3,pol))),C');
    );
   admissible=fold((x,y)->(x and y),apply(differentielles,dif->(dif>=0)));
   kos=expKoszul({0,0,0,0,0,0,0},triplet);
   s=lift(altersum apply(kos,degr->sum
    apply(degr,pol->Nesp2(formes,3,pol))),ZZ);
   if formes==1 then admissibles="" else
    admissibles=concatenate(admissibles,",");
   admissibles=concatenate(admissibles,toString s);
   if admissible then admissibles=concatenate(admissibles,"*");
   );
  print concatenate("somme alternee pour chacune des 8 formes: ",admissibles);
  use C;
  g=gb(ideal ((polys#i)#0,(polys#j)#1,(polys#k)#2),DegreeLimit => D);
  resultant=selectInSubring(1,gens g);
  if #(entries resultant)#0==1 then print concatenate(
  "terme dominant du resultant (en degre <=",toString D,"): ",
  toString leadMonomial((entries resultant)#0#0)) else
   print "Resultant non trouve";
  print "";
 )))))
```

Pour l'essayer, utiliser par exemple le code suivant :

```
T=tablpseudoesp1(4,15);T3=tabl3 T;
test8formesPseudoesp1(T,T3,50);
```

Ci-dessous figurent les résultats affichés par cette dernière commande. On remarque que les formes admissibles renvoient toujours la valeur de D maximale parmi les 8 valeurs trouvées, et que la première forme est toujours admissible :

```
exposants: ({3, 1, 1, 1},{3, 1, 1, 1},{3, 1, 1, 1})
```

```
somme alternee pour chacune des 8 formes: 6*,6*,6*,6*,6*,6*,6*,6*
terme dominant du resultant (en degre <=50): x_2^6*z^3

exposants: ({3, 1, 1, 1},{3, 1, 1, 1},{3, 2, 1, 1})
somme alternee pour chacune des 8 formes: 8*,8*,8*,8*,8*,8*,8*,8*
terme dominant du resultant (en degre <=50): x_2^8*z^3

exposants: ({3, 1, 1, 1},{3, 1, 1, 1},{4, 2, 1, 1})
somme alternee pour chacune des 8 formes: 8*,8*,8*,8*,8*,8*,8*,8*
terme dominant du resultant (en degre <=50): x_2^8*z^4

exposants: ({3, 1, 1, 1},{3, 1, 1, 1},{3, 3, 1, 1})
somme alternee pour chacune des 8 formes: 10*,10*,10*,10*,10*,10*,10*,10*
terme dominant du resultant (en degre <=50): x_2^10*z^3

exposants: ({3, 1, 1, 1},{3, 1, 1, 1},{4, 3, 1, 1})
somme alternee pour chacune des 8 formes: 10*,10*,10*,10*,10*,10*,10*,10*
terme dominant du resultant (en degre <=50): x_2^10*z^4

exposants: ({3, 1, 1, 1},{3, 2, 1, 1},{3, 2, 1, 1})
somme alternee pour chacune des 8 formes: 10*,10,10,10,10,10,10,10
terme dominant du resultant (en degre <=50): x_2^10*z^3

exposants: ({3, 1, 1, 1},{3, 2, 1, 1},{3, 1, 2, 1})
somme alternee pour chacune des 8 formes: 11*,11,11,11,11,11,11,11
terme dominant du resultant (en degre <=50): x_2^11*z^4

exposants: ({3, 1, 1, 1},{3, 2, 1, 1},{4, 2, 1, 1})
somme alternee pour chacune des 8 formes: 10*,10*,10*,10*,10*,10*,10*,10*
terme dominant du resultant (en degre <=50): x_2^10*z^4

exposants: ({3, 1, 1, 1},{3, 2, 1, 1},{4, 1, 2, 1})
somme alternee pour chacune des 8 formes: 11*,11*,11*,11*,11*,11*,11*,11*
terme dominant du resultant (en degre <=50): x_2^11*z^4

exposants: ({3, 1, 1, 1},{3, 2, 1, 1},{3, 3, 1, 1})
somme alternee pour chacune des 8 formes: 12*,12,12*,12,12,12,12,12
terme dominant du resultant (en degre <=50): x_2^12*z^3

exposants: ({3, 1, 1, 1},{3, 2, 1, 1},{3, 1, 3, 1})
somme alternee pour chacune des 8 formes: 14*,14,14,14,14,14,14*,14
terme dominant du resultant (en degre <=50): x_2^14*z^4

exposants: ({3, 1, 1, 1},{3, 2, 1, 1},{4, 3, 1, 1})
somme alternee pour chacune des 8 formes: 12*,12,12,12,12,12,12,12
terme dominant du resultant (en degre <=50): x_2^12*z^4

exposants: ({3, 1, 1, 1},{3, 2, 1, 1},{4, 1, 3, 1})
somme alternee pour chacune des 8 formes: 14*,14,14,14,14,14,14,14
terme dominant du resultant (en degre <=50): x_2^14*z^5

exposants: ({3, 1, 1, 1},{4, 2, 1, 1},{4, 2, 1, 1})
somme alternee pour chacune des 8 formes: 10*,10*,10*,10*,10*,10*,10*,10*
terme dominant du resultant (en degre <=50): x_2^10*z^4
```

```
exposants: ({3, 1, 1, 1},{4, 2, 1, 1},{4, 1, 2, 1})
somme alternee pour chacune des 8 formes: 11*,11*,11*,11*,11*,11*,11*,11*
terme dominant du resultant (en degre <=50): x_2^11*z^5

exposants: ({3, 1, 1, 1},{4, 2, 1, 1},{3, 3, 1, 1})
somme alternee pour chacune des 8 formes: 12*,12*,12*,12*,12*,12*,12*,12*
terme dominant du resultant (en degre <=50): x_2^12*z^4

exposants: ({3, 1, 1, 1},{4, 2, 1, 1},{3, 1, 3, 1})
somme alternee pour chacune des 8 formes: 14*,14*,14*,14*,14*,14*,14*,14*
terme dominant du resultant (en degre <=50): x_2^14*z^4

exposants: ({3, 1, 1, 1},{4, 2, 1, 1},{4, 3, 1, 1})
somme alternee pour chacune des 8 formes: 12*,12*,12*,12*,12*,12*,12*,12*
terme dominant du resultant (en degre <=50): x_2^12*z^5

exposants: ({3, 1, 1, 1},{4, 2, 1, 1},{4, 1, 3, 1})
somme alternee pour chacune des 8 formes: 14*,14*,14*,14*,14*,14*,14*,14*
terme dominant du resultant (en degre <=50): x_2^14*z^5

exposants: ({3, 1, 1, 1},{3, 3, 1, 1},{3, 3, 1, 1})
somme alternee pour chacune des 8 formes: 12*,12,12*,12,12,12,12,12
terme dominant du resultant (en degre <=50): x_2^12*z^3

exposants: ({3, 1, 1, 1},{3, 3, 1, 1},{3, 1, 3, 1})
somme alternee pour chacune des 8 formes: 17*,17*,17*,17*,17*,17*,17*,17*
terme dominant du resultant (en degre <=50): x_2^17*z^4

exposants: ({3, 1, 1, 1},{3, 3, 1, 1},{4, 3, 1, 1})
somme alternee pour chacune des 8 formes: 14*,14,14*,14,14,14,14,14
terme dominant du resultant (en degre <=50): x_2^14*z^4

exposants: ({3, 1, 1, 1},{3, 3, 1, 1},{4, 1, 3, 1})
somme alternee pour chacune des 8 formes: 18*,18,18*,18,18,18,18,18
terme dominant du resultant (en degre <=50): x_2^18*z^5

exposants: ({3, 1, 1, 1},{4, 3, 1, 1},{4, 3, 1, 1})
somme alternee pour chacune des 8 formes: 14*,14,14,14,14,14,14,14
terme dominant du resultant (en degre <=50): x_2^14*z^5

exposants: ({3, 1, 1, 1},{4, 3, 1, 1},{4, 1, 3, 1})
somme alternee pour chacune des 8 formes: 18*,18,18,18,18,18,18,18
terme dominant du resultant (en degre <=50): x_2^18*z^6

exposants: ({3, 2, 1, 1},{3, 2, 1, 1},{3, 2, 1, 1})
somme alternee pour chacune des 8 formes: 11*,10,10,9,9,8,10,9
terme dominant du resultant (en degre <=50): x_2^11*z^3

exposants: ({3, 2, 1, 1},{3, 2, 1, 1},{3, 1, 2, 1})
somme alternee pour chacune des 8 formes: 13*,12,12,11,11,10,12,11
terme dominant du resultant (en degre <=50): x_2^13*z^4

exposants: ({3, 2, 1, 1},{3, 2, 1, 1},{4, 2, 1, 1})
```

```
somme alternee pour chacune des 8 formes: 12*,12,12,12,12,12,12,12
terme dominant du resultant (en degre <=50): x_2^12*z^4

exposants: ({3, 2, 1, 1},{3, 2, 1, 1},{4, 1, 2, 1})
somme alternee pour chacune des 8 formes: 14*,14,14,14,14,14,14,14
terme dominant du resultant (en degre <=50): x_2^14*z^4

exposants: ({3, 2, 1, 1},{3, 2, 1, 1},{3, 3, 1, 1})
somme alternee pour chacune des 8 formes: 12*,11,12,11,11,10,11,10
terme dominant du resultant (en degre <=50): x_2^12*z^3

exposants: ({3, 2, 1, 1},{3, 2, 1, 1},{3, 1, 3, 1})
somme alternee pour chacune des 8 formes: 16*,15,15,14,15,14,16,15
terme dominant du resultant (en degre <=50): x_2^16*z^4

exposants: ({3, 2, 1, 1},{3, 2, 1, 1},{4, 3, 1, 1})
somme alternee pour chacune des 8 formes: 13*,12,12,11,11,10,12,11
terme dominant du resultant (en degre <=50): x_2^13*z^4

exposants: ({3, 2, 1, 1},{3, 2, 1, 1},{4, 1, 3, 1})
somme alternee pour chacune des 8 formes: 17*,16,16,15,15,14,16,15
terme dominant du resultant (en degre <=50): x_2^17*z^5

exposants: ({3, 2, 1, 1},{3, 1, 2, 1},{3, 1, 2, 1})
somme alternee pour chacune des 8 formes: 13*,12,12,11,11,10,12,11
terme dominant du resultant (en degre <=50): x_2^13*z^4

exposants: ({3, 2, 1, 1},{3, 1, 2, 1},{3, 1, 1, 2})
somme alternee pour chacune des 8 formes: 15*,14,14,13,13,12,14,13
terme dominant du resultant (en degre <=50): x_2^15*z^3

exposants: ({3, 2, 1, 1},{3, 1, 2, 1},{4, 2, 1, 1})
somme alternee pour chacune des 8 formes: 14*,14,14,14,14,14,14,14
terme dominant du resultant (en degre <=50): x_2^14*z^5

exposants: ({3, 2, 1, 1},{3, 1, 2, 1},{4, 1, 2, 1})
somme alternee pour chacune des 8 formes: 14*,14,14,14,14,14,14,14
terme dominant du resultant (en degre <=50): x_2^14*z^5

exposants: ({3, 2, 1, 1},{3, 1, 2, 1},{4, 1, 1, 2})
somme alternee pour chacune des 8 formes: 16*,16,16,16,16,16,16,16
terme dominant du resultant (en degre <=50): x_2^16*z^4

exposants: ({3, 2, 1, 1},{3, 1, 2, 1},{3, 3, 1, 1})
somme alternee pour chacune des 8 formes: 15*,14,15,14,14,13,14,13
terme dominant du resultant (en degre <=50): x_2^15*z^4

exposants: ({3, 2, 1, 1},{3, 1, 2, 1},{3, 1, 3, 1})
somme alternee pour chacune des 8 formes: 15*,14,14,13,14,13,15,14
terme dominant du resultant (en degre <=50): x_2^15*z^4

exposants: ({3, 2, 1, 1},{3, 1, 2, 1},{3, 1, 1, 3})
somme alternee pour chacune des 8 formes: 19*,19,18,18,17,17,18,18
terme dominant du resultant (en degre <=50): x_2^19
```

```
exposants: ({3, 2, 1, 1},{3, 1, 2, 1},{4, 3, 1, 1})
somme alternee pour chacune des 8 formes: 16*,15,15,14,14,13,15,14
terme dominant du resultant (en degre <=50): x_2^16*z^5

exposants: ({3, 2, 1, 1},{3, 1, 2, 1},{4, 1, 3, 1})
somme alternee pour chacune des 8 formes: 16*,15,15,14,14,13,15,14
terme dominant du resultant (en degre <=50): x_2^16*z^5

exposants: ({3, 2, 1, 1},{3, 1, 2, 1},{4, 1, 1, 3})
somme alternee pour chacune des 8 formes: 20*,19,19,18,18,17,19,18
terme dominant du resultant (en degre <=50): x_2^20*z^3

exposants: ({3, 2, 1, 1},{4, 2, 1, 1},{4, 2, 1, 1})
somme alternee pour chacune des 8 formes: 12*,12*,12*,12*,12*,12*,12*,12*
terme dominant du resultant (en degre <=50): x_2^12*z^5

exposants: ({3, 2, 1, 1},{4, 2, 1, 1},{4, 1, 2, 1})
somme alternee pour chacune des 8 formes: 14*,14*,14*,14*,14*,14*,14*,14*
terme dominant du resultant (en degre <=50): x_2^14*z^5

exposants: ({3, 2, 1, 1},{4, 2, 1, 1},{3, 3, 1, 1})
somme alternee pour chacune des 8 formes: 14*,14,14*,14,14,14,14,14
terme dominant du resultant (en degre <=50): x_2^14*z^4

exposants: ({3, 2, 1, 1},{4, 2, 1, 1},{3, 1, 3, 1})
somme alternee pour chacune des 8 formes: 18*,18,18,18,18,18,18*,18
terme dominant du resultant (en degre <=50): x_2^18*z^5

exposants: ({3, 2, 1, 1},{4, 2, 1, 1},{4, 3, 1, 1})
somme alternee pour chacune des 8 formes: 14*,14,14,14,14,14,14,14
terme dominant du resultant (en degre <=50): x_2^14*z^5

exposants: ({3, 2, 1, 1},{4, 2, 1, 1},{4, 1, 3, 1})
somme alternee pour chacune des 8 formes: 18*,18,18,18,18,18,18,18
terme dominant du resultant (en degre <=50): x_2^18*z^6

exposants: ({3, 2, 1, 1},{4, 1, 2, 1},{4, 1, 2, 1})
somme alternee pour chacune des 8 formes: 14*,14*,14*,14*,14*,14*,14*,14*
terme dominant du resultant (en degre <=50): x_2^14*z^5

exposants: ({3, 2, 1, 1},{4, 1, 2, 1},{4, 1, 1, 2})
somme alternee pour chacune des 8 formes: 16*,16*,16*,16*,16*,16*,16*,16*
terme dominant du resultant (en degre <=50): x_2^16*z^4

exposants: ({3, 2, 1, 1},{4, 1, 2, 1},{3, 3, 1, 1})
somme alternee pour chacune des 8 formes: 17*,17,17*,17,17,17,17,17
terme dominant du resultant (en degre <=50): x_2^17*z^4

exposants: ({3, 2, 1, 1},{4, 1, 2, 1},{3, 1, 3, 1})
somme alternee pour chacune des 8 formes: 17*,17,17,17,17,17,17*,17
terme dominant du resultant (en degre <=50): x_2^17*z^5

exposants: ({3, 2, 1, 1},{4, 1, 2, 1},{3, 1, 1, 3})
```

```
somme alternee pour chacune des 8 formes: 21*,21*,21,21,21,21,21,21
terme dominant du resultant (en degre <=50): x_2^21*z

exposants: ({3, 2, 1, 1},{4, 1, 2, 1},{4, 3, 1, 1})
somme alternee pour chacune des 8 formes: 17*,17,17,17,17,17,17,17
terme dominant du resultant (en degre <=50): x_2^17*z^5

exposants: ({3, 2, 1, 1},{4, 1, 2, 1},{4, 1, 3, 1})
somme alternee pour chacune des 8 formes: 17*,17,17,17,17,17,17,17
terme dominant du resultant (en degre <=50): x_2^17*z^6

exposants: ({3, 2, 1, 1},{4, 1, 2, 1},{4, 1, 1, 3})
somme alternee pour chacune des 8 formes: 21*,21,21,21,21,21,21,21
terme dominant du resultant (en degre <=50): x_2^21*z^4

exposants: ({3, 2, 1, 1},{3, 3, 1, 1},{3, 3, 1, 1})
somme alternee pour chacune des 8 formes: 12*,11,12*,11,11,10,11,10
terme dominant du resultant (en degre <=50): x_2^12*z^3

exposants: ({3, 2, 1, 1},{3, 3, 1, 1},{3, 1, 3, 1})
somme alternee pour chacune des 8 formes: 18*,17,18,17,18,17,18,17
terme dominant du resultant (en degre <=50): x_2^18*z^4

exposants: ({3, 2, 1, 1},{3, 3, 1, 1},{4, 3, 1, 1})
somme alternee pour chacune des 8 formes: 14*,13,14,13,13,12,13,12
terme dominant du resultant (en degre <=50): x_2^14*z^4

exposants: ({3, 2, 1, 1},{3, 3, 1, 1},{4, 1, 3, 1})
somme alternee pour chacune des 8 formes: 20*,19,20,19,19,18,19,18
terme dominant du resultant (en degre <=50): x_2^20*z^5

exposants: ({3, 2, 1, 1},{3, 1, 3, 1},{3, 1, 3, 1})
somme alternee pour chacune des 8 formes: 15*,14,14,13,14,13,15*,14
terme dominant du resultant (en degre <=50): x_2^15*z^4

exposants: ({3, 2, 1, 1},{3, 1, 3, 1},{3, 1, 1, 3})
somme alternee pour chacune des 8 formes: 23*,23,22,22,22,22,23,23
terme dominant du resultant (en degre <=50): x_2^23

exposants: ({3, 2, 1, 1},{3, 1, 3, 1},{4, 3, 1, 1})
somme alternee pour chacune des 8 formes: 20*,19,19,18,19,18,20,19
terme dominant du resultant (en degre <=50): x_2^20*z^5

exposants: ({3, 2, 1, 1},{3, 1, 3, 1},{4, 1, 3, 1})
somme alternee pour chacune des 8 formes: 18*,17,17,16,17,16,18,17
terme dominant du resultant (en degre <=50): x_2^18*z^5

exposants: ({3, 2, 1, 1},{3, 1, 3, 1},{4, 1, 1, 3})
somme alternee pour chacune des 8 formes: 26*,25,25,24,25,24,26,25
terme dominant du resultant (en degre <=50): x_2^26*z^3

exposants: ({3, 2, 1, 1},{4, 3, 1, 1},{4, 3, 1, 1})
somme alternee pour chacune des 8 formes: 15*,14,14,13,13,12,14,13
terme dominant du resultant (en degre <=50): x_2^15*z^5
```

```
exposants: ({3, 2, 1, 1},{4, 3, 1, 1},{4, 1, 3, 1})
somme alternee pour chacune des 8 formes: 21*,20,20,19,19,18,20,19
terme dominant du resultant (en degre <=50): x_2^21*z^6

exposants: ({3, 2, 1, 1},{4, 1, 3, 1},{4, 1, 3, 1})
somme alternee pour chacune des 8 formes: 19*,18,18,17,17,16,18,17
terme dominant du resultant (en degre <=50): x_2^19*z^6

exposants: ({3, 2, 1, 1},{4, 1, 3, 1},{4, 1, 1, 3})
somme alternee pour chacune des 8 formes: 27*,26,26,25,25,24,26,25
terme dominant du resultant (en degre <=50): x_2^27*z^4

exposants: ({3, 1, 2, 1},{4, 2, 1, 1},{4, 1, 2, 1})
somme alternee pour chacune des 8 formes: 14*,14*,14*,14*,14*,14*,14*,14*
terme dominant du resultant (en degre <=50): x_2^14*z^5

exposants: ({3, 1, 2, 1},{3, 3, 1, 1},{3, 1, 3, 1})
somme alternee pour chacune des 8 formes: 18*,17,18,17,18,17,18,17
terme dominant du resultant (en degre <=50): x_2^18*z^4

exposants: ({3, 1, 2, 1},{4, 3, 1, 1},{4, 1, 3, 1})
somme alternee pour chacune des 8 formes: 21*,20,20,19,19,18,20,19
terme dominant du resultant (en degre <=50): x_2^21*z^6

exposants: ({4, 2, 1, 1},{4, 2, 1, 1},{4, 2, 1, 1})
somme alternee pour chacune des 8 formes: 12*,12*,12*,12*,12*,12*,12*,12*
terme dominant du resultant (en degre <=50): x_2^12*z^5

exposants: ({4, 2, 1, 1},{4, 2, 1, 1},{4, 1, 2, 1})
somme alternee pour chacune des 8 formes: 14*,14*,14*,14*,14*,14*,14*,14*
terme dominant du resultant (en degre <=50): x_2^14*z^5

exposants: ({4, 2, 1, 1},{4, 2, 1, 1},{3, 3, 1, 1})
somme alternee pour chacune des 8 formes: 14*,14*,14*,14*,14*,14*,14*,14*
terme dominant du resultant (en degre <=50): x_2^14*z^5

exposants: ({4, 2, 1, 1},{4, 2, 1, 1},{3, 1, 3, 1})
somme alternee pour chacune des 8 formes: 18*,18*,18*,18*,18*,18*,18*,18*
terme dominant du resultant (en degre <=50): x_2^18*z^5

exposants: ({4, 2, 1, 1},{4, 2, 1, 1},{4, 3, 1, 1})
somme alternee pour chacune des 8 formes: 14*,14*,14*,14*,14*,14*,14*,14*
terme dominant du resultant (en degre <=50): x_2^14*z^5

exposants: ({4, 2, 1, 1},{4, 2, 1, 1},{4, 1, 3, 1})
somme alternee pour chacune des 8 formes: 18*,18*,18*,18*,18*,18*,18*,18*
terme dominant du resultant (en degre <=50): x_2^18*z^6

exposants: ({4, 2, 1, 1},{4, 1, 2, 1},{4, 1, 2, 1})
somme alternee pour chacune des 8 formes: 14*,14*,14*,14*,14*,14*,14*,14*
terme dominant du resultant (en degre <=50): x_2^14*z^5

exposants: ({4, 2, 1, 1},{4, 1, 2, 1},{4, 1, 1, 2})
```

```
somme alternee pour chacune des 8 formes: 16*,16*,16*,16*,16*,16*,16*,16*
terme dominant du resultant (en degre <=50): x_2^16*z^5

exposants: ({4, 2, 1, 1},{4, 1, 2, 1},{3, 3, 1, 1})
somme alternee pour chacune des 8 formes: 17*,17*,17*,17*,17*,17*,17*,17*
terme dominant du resultant (en degre <=50): x_2^17*z^5

exposants: ({4, 2, 1, 1},{4, 1, 2, 1},{3, 1, 3, 1})
somme alternee pour chacune des 8 formes: 17*,17*,17*,17*,17*,17*,17*,17*
terme dominant du resultant (en degre <=50): x_2^17*z^5

exposants: ({4, 2, 1, 1},{4, 1, 2, 1},{3, 1, 1, 3})
somme alternee pour chacune des 8 formes: 21*,21*,21*,21*,21*,21*,21*,21*
terme dominant du resultant (en degre <=50): x_2^21*z^2

exposants: ({4, 2, 1, 1},{4, 1, 2, 1},{4, 3, 1, 1})
somme alternee pour chacune des 8 formes: 17*,17*,17*,17*,17*,17*,17*,17*
terme dominant du resultant (en degre <=50): x_2^17*z^6

exposants: ({4, 2, 1, 1},{4, 1, 2, 1},{4, 1, 3, 1})
somme alternee pour chacune des 8 formes: 17*,17*,17*,17*,17*,17*,17*,17*
terme dominant du resultant (en degre <=50): x_2^17*z^6

exposants: ({4, 2, 1, 1},{4, 1, 2, 1},{4, 1, 1, 3})
somme alternee pour chacune des 8 formes: 21*,21*,21*,21*,21*,21*,21*,21*
terme dominant du resultant (en degre <=50): x_2^21*z^5

exposants: ({4, 2, 1, 1},{3, 3, 1, 1},{3, 3, 1, 1})
somme alternee pour chacune des 8 formes: 14*,14,14*,14,14,14,14,14
terme dominant du resultant (en degre <=50): x_2^14*z^4

exposants: ({4, 2, 1, 1},{3, 3, 1, 1},{3, 1, 3, 1})
somme alternee pour chacune des 8 formes: 21*,21*,21*,21*,21*,21*,21*,21*
terme dominant du resultant (en degre <=50): x_2^21*z^5

exposants: ({4, 2, 1, 1},{3, 3, 1, 1},{4, 3, 1, 1})
somme alternee pour chacune des 8 formes: 16*,16,16*,16,16,16,16,16
terme dominant du resultant (en degre <=50): x_2^16*z^5

exposants: ({4, 2, 1, 1},{3, 3, 1, 1},{4, 1, 3, 1})
somme alternee pour chacune des 8 formes: 22*,22,22*,22,22,22,22,22
terme dominant du resultant (en degre <=50): x_2^22*z^6

exposants: ({4, 2, 1, 1},{3, 1, 3, 1},{3, 1, 3, 1})
somme alternee pour chacune des 8 formes: 17*,17,17,17,17,17,17*,17
terme dominant du resultant (en degre <=50): x_2^17*z^4

exposants: ({4, 2, 1, 1},{3, 1, 3, 1},{3, 1, 1, 3})
somme alternee pour chacune des 8 formes: 26*,26*,26*,26*,26*,26*,26*,26*
terme dominant du resultant (en degre <=50): x_2^26*z

exposants: ({4, 2, 1, 1},{3, 1, 3, 1},{4, 3, 1, 1})
somme alternee pour chacune des 8 formes: 22*,22,22,22,22,22,22*,22
terme dominant du resultant (en degre <=50): x_2^22*z^6
```

```
exposants: ({4, 2, 1, 1},{3, 1, 3, 1},{4, 1, 3, 1})
somme alternee pour chacune des 8 formes: 20*,20,20,20,20,20,20*,20
terme dominant du resultant (en degre <=50): x_2^20*z^5

exposants: ({4, 2, 1, 1},{3, 1, 3, 1},{4, 1, 1, 3})
somme alternee pour chacune des 8 formes: 28*,28,28,28,28,28,28*,28
terme dominant du resultant (en degre <=50): x_2^28*z^4

exposants: ({4, 2, 1, 1},{4, 3, 1, 1},{4, 3, 1, 1})
somme alternee pour chacune des 8 formes: 16*,16,16,16,16,16,16,16
terme dominant du resultant (en degre <=50): x_2^16*z^5

exposants: ({4, 2, 1, 1},{4, 3, 1, 1},{4, 1, 3, 1})
somme alternee pour chacune des 8 formes: 22*,22,22,22,22,22,22,22
terme dominant du resultant (en degre <=50): x_2^22*z^7

exposants: ({4, 2, 1, 1},{4, 1, 3, 1},{4, 1, 3, 1})
somme alternee pour chacune des 8 formes: 20*,20,20,20,20,20,20,20
terme dominant du resultant (en degre <=50): x_2^20*z^6

exposants: ({4, 2, 1, 1},{4, 1, 3, 1},{4, 1, 1, 3})
somme alternee pour chacune des 8 formes: 28*,28,28,28,28,28,28,28
terme dominant du resultant (en degre <=50): x_2^28*z^5

exposants: ({4, 1, 2, 1},{3, 3, 1, 1},{3, 1, 3, 1})
somme alternee pour chacune des 8 formes: 21*,21*,21*,21*,21*,21*,21*,21*
terme dominant du resultant (en degre <=50): x_2^21*z^5

exposants: ({4, 1, 2, 1},{4, 3, 1, 1},{4, 1, 3, 1})
somme alternee pour chacune des 8 formes: 22*,22,22,22,22,22,22,22
terme dominant du resultant (en degre <=50): x_2^22*z^7

exposants: ({3, 3, 1, 1},{3, 3, 1, 1},{3, 3, 1, 1})
somme alternee pour chacune des 8 formes: 12*,11,12*,11,11,10,11,10
terme dominant du resultant (en degre <=50): x_2^12*z^3

exposants: ({3, 3, 1, 1},{3, 3, 1, 1},{3, 1, 3, 1})
somme alternee pour chacune des 8 formes: 18*,17,18*,17,18,17,18,17
terme dominant du resultant (en degre <=50): x_2^18*z^4

exposants: ({3, 3, 1, 1},{3, 3, 1, 1},{4, 3, 1, 1})
somme alternee pour chacune des 8 formes: 14*,13,14*,13,13,12,13,12
terme dominant du resultant (en degre <=50): x_2^14*z^4

exposants: ({3, 3, 1, 1},{3, 3, 1, 1},{4, 1, 3, 1})
somme alternee pour chacune des 8 formes: 20*,19,20*,19,19,18,19,18
terme dominant du resultant (en degre <=50): x_2^20*z^5

exposants: ({3, 3, 1, 1},{3, 1, 3, 1},{3, 1, 3, 1})
somme alternee pour chacune des 8 formes: 18*,17,18,17,18,17,18*,17
terme dominant du resultant (en degre <=50): x_2^18*z^4

exposants: ({3, 3, 1, 1},{3, 1, 3, 1},{3, 1, 1, 3})
```

```
somme alternee pour chacune des 8 formes: 27*,27*,27*,27*,27*,27*,27*,27*
terme dominant du resultant (en degre <=50): x_2^27

exposants: ({3, 3, 1, 1},{3, 1, 3, 1},{4, 3, 1, 1})
somme alternee pour chacune des 8 formes: 22*,21,22,21,22,21,22,21
terme dominant du resultant (en degre <=50): x_2^22*z^5

exposants: ({3, 3, 1, 1},{3, 1, 3, 1},{4, 1, 3, 1})
somme alternee pour chacune des 8 formes: 22*,21,22,21,22,21,22,21
terme dominant du resultant (en degre <=50): x_2^22*z^5

exposants: ({3, 3, 1, 1},{3, 1, 3, 1},{4, 1, 1, 3})
somme alternee pour chacune des 8 formes: 32*,31,32,31,32,31,32,31
terme dominant du resultant (en degre <=50): x_2^32*z^3

exposants: ({3, 3, 1, 1},{4, 3, 1, 1},{4, 3, 1, 1})
somme alternee pour chacune des 8 formes: 16*,15,16,15,15,14,15,14
terme dominant du resultant (en degre <=50): x_2^16*z^5

exposants: ({3, 3, 1, 1},{4, 3, 1, 1},{4, 1, 3, 1})
somme alternee pour chacune des 8 formes: 24*,23,24,23,23,22,23,22
terme dominant du resultant (en degre <=50): x_2^24*z^6

exposants: ({3, 3, 1, 1},{4, 1, 3, 1},{4, 1, 3, 1})
somme alternee pour chacune des 8 formes: 24*,23,24,23,23,22,23,22
terme dominant du resultant (en degre <=50): x_2^24*z^6

exposants: ({3, 3, 1, 1},{4, 1, 3, 1},{4, 1, 1, 3})
somme alternee pour chacune des 8 formes: 36*,35,36,35,35,34,35,34
terme dominant du resultant (en degre <=50): x_2^36*z^4

exposants: ({3, 1, 3, 1},{4, 3, 1, 1},{4, 1, 3, 1})
somme alternee pour chacune des 8 formes: 24*,23,23,22,23,22,24,23
terme dominant du resultant (en degre <=50): x_2^24*z^6

exposants: ({4, 3, 1, 1},{4, 3, 1, 1},{4, 3, 1, 1})
somme alternee pour chacune des 8 formes: 17*,16,16,15,15,14,16,15
terme dominant du resultant (en degre <=50): x_2^17*z^5

exposants: ({4, 3, 1, 1},{4, 3, 1, 1},{4, 1, 3, 1})
somme alternee pour chacune des 8 formes: 25*,24,24,23,23,22,24,23
terme dominant du resultant (en degre <=50): x_2^25*z^7

exposants: ({4, 3, 1, 1},{4, 1, 3, 1},{4, 1, 3, 1})
somme alternee pour chacune des 8 formes: 25*,24,24,23,23,22,24,23
terme dominant du resultant (en degre <=50): x_2^25*z^7

exposants: ({4, 3, 1, 1},{4, 1, 3, 1},{4, 1, 1, 3})
somme alternee pour chacune des 8 formes: 37*,36,36,35,35,34,36,35
terme dominant du resultant (en degre <=50): x_2^37*z^5
```

Annexe E

Extrait de la préface des *Meditationes Algebraicae* de Waring (1782)

In eâdem editione asseritur, si modo duae aequationes duas incognitas quantitates $(x\&y)$ involvant, correspondentes valores quantitatum $(x\&y)$ semper eandem habere irrationalitatem, ni duo vel plures (m) valores unius (e. g. x) sint inter se aequales, tum autem eandem irrationalitatem habere quadraticam, &c. aequationem, cujus radix est x, quam habet ejus correspondens valor quantitatis (y).

Methodum reducendi duas aequationes $y^n + ay^{n-1} + by^{n-2} +$ &c. $= 0$ et $y^m + Ay^{m-1} + By^{m-2} +$ &c. $= 0$ in unam, ita ut exterminetur incognita quantitas y, docuit *Eulerus* ducendo praedictas duas aequationes in duas quantitates $Py^{m-1} + a'y^{m-2} + b'y^{m-3} +$ &c. , et $P'y^{n-1} + A'y^{n-2} + B'y^{n-3} +$ &c. , et rectangula addendo et ita assumendo coefficientes $P, a', b',$ &c. ; $P', A', B',$ &c. ; ut evanescant singuli termini resultantis aequationis, in quibus invenitur quantitas (y) : consimilia etiam applicari possunt ad tres vel plures aequationes tres vel plures incognitas quantitates habentes.

Bezout reduxit duas vel plures aequationes in unam, ita ut incognitae quantitates exterminentur, ducendo datas aequationes $Ay^n + (a + bx + cz +$ &c. $)y^{n-1} + (d + ex + fz + gx^2 + hxz +$ &c. $)y^{n-2} +$ &c. $= 0$, $A'y^m + (a' + b'x + c'z +$ &c. $)y^{m-1} +$ &c. $= 0$, $A''y^r + (a'' + b''x + c''z +$ &c. $)y^{r-1} +$ &c. $= 0$, &c. in quantitates subsequentium formularum $Py^{nmr\&c.\ -n} + (p + qx + rz +$ &c. $)y^{nmr\&c.\ -n-1} +$ &c. , $P'y^{nmr\&c.\ -m} + (p' + q'x + r'z +$ &c. $)y^{nmr\&c.\ -m-1} +$ &c. , $P''y^{nmr\&c.\ -r} +$ &c. ; &c. respective ; et rec-

[1]Dans cette même édition [la seconde (1770)] on affirme que, étant données deux équations à deux inconnues x et y, les valeurs correspondantes des quantités x et y ont toujours la même irrationalité[2], sauf si deux ou plusieurs (m) valeurs de l'une (par exemple y)[3] sont égales entre elles, auquel cas l'équation – par exemple quadratique – dont la racine est x a la même irrationalité que la valeur de y lui correspondant.

Euler enseigne une méthode réduisant deux équations $y^n + ay^{n-1} + by^{n-2} + ... = 0$ et $y^m + Ay^{m-1} + By^{m-2} + ... = 0$ à une seule de sorte que soit éliminée l'inconnue y. Cette méthode consiste à multiplier les deux équations ci-dessus par deux quantités $Py^{m-1} + a'y^{m-2} + b'y^{m-3} + ...$ et $P'y^{n-1} + A'y^{n-2} + B'y^{n-3} + ...$, à faire la somme des deux produits et à déterminer les coefficients $P, a', b',$ &c. , $P', A', B',$ &c. pour que chaque terme de l'équation résultante dans lequel se trouve y s'évanouisse ; on peut aussi appliquer des méthodes semblables à trois équations ou plus, à trois inconnues ou plus.

Bézout a réduit deux équations ou plus à une seule équation, de sorte que soient éliminées les inconnues, en multipliant les équations données $Ay^n + (a + bx + cz + ...)y^{n-1} + (d + ex + fz + gx^2 + hxz + ...)y^{n-2} + ... = 0$, $A'y^m + (a' + b'x + c'z + ...)y^{m-1} + ... = 0$, $A''y^r + (a'' + b''x + c''z + ...)y^{r-1} + ... = 0$,... par des quantités respectivement de la forme $Py^{nmr...-n} + (p + qx + rz + ...)y^{nmr...-n-1} + ...$, $P'y^{nmr...-m} + (p' + q'x + r'z + ...)y^{nmr...-m-1} + ...$, $P''y^{nmr...-r} + ...$,... Il fait la somme des pro-

tangula exinde resultantia addendo ; deinde ita assumendo coefficientes harum quantitatum, ut nihilo evadant aequales omnes termini, in quibus inveniuntur quantitates exterminandae : in multis casibus dedit etiam magis simplices formulas multiplicatorum, quae exterminationem saepe magis facilem reddent : et subsequentes elegantes propositiones detexit ; ex principiis Taylori incrementorum ; etiamque propositione facile deducendâ, attamen à me in primâ editione hujusce operis datâ, viz. quod numerus coefficientium in aequatione (n) dimensionum $(y^n + (a + bx + cz + \&c.\ )y^{n-1} + \&c.\ ) = 0$, (m) incognitas quantitates habente, erit $(n+1) \times \frac{n+2}{2} \times \frac{n+3}{3} .. \frac{n+m}{m} = S$; et ex propositione eleganti a se inventâ, viz. quod $N(h)^n$ denotante numerum terminorum in aequatione (n) dimensionum, quae continet (h) incognitas quantitates $(x, y, z, v, \&c.\ )$; numerus terminorum in aequatione (n) dimensionum divisibilium vel per x^p vel y^q vel z^r vel v^s &c. erit $= (N(h)^{n-p} + N(h)^{n-q} + N(h)^{n-r} + N(h)^{n-s} + \&c.\ ) - (N(h)^{n-p-q} + N(h)^{n-p-r} + N(h)^{n-p-s} + N(h)^{n-q-s} + \&c.\ ) + (N(h)^{n-p-q-r} + N(h)^{n-p-q-s} + \&c.\ ) - (N(h)^{n-p-q-r-s} + \&c.\ ) + \&c.$: et si sint (h) aequationes $(n, m, l, k, \&c.\ )$ dimensionum respective totidem incognitas quantitates $(x, y, z, v, \&c.\ )$ involventes ; et sint p, q, r, s, &c. ; p', q', r', &c. ; p'', q'', &c. ; maximae dimensiones, ad quas ascendunt incognitae quantitates x, y, z, v, &c. , in respectivis aequationibus $(n, m, l, k, \&c.\ )$ dimensionum ; tum aequationem, cujus radix est x vel y vel z, &c. haud ascendere ad plures quam $n \times m \times l \times k \times \&c. - (n-p) \times (m-p') \times (l-p'') \times \&c. - (n-q) \times (m-q') \times (l-q'') \times \&c. - (n-r) \times (m-r') \times \&c. - \&c. = P$ dimensiones : si vero dimensiones quantitatum $(x \&c.\ y)$ simul sumptarum haud superent dimensiones $a, a', a'', \&c.$ in predictis aequationibus ; tum aequationem, cujus radix est $x, \&c.$ haud plures habere quam $P + (n-a) \times (m-a') \times \&c. - (p+q-a) \times (p'-a') \times \&c.$; hae propositiones forsan facilius e principiis in prob. 41, a me traditis, quae prius edita fuere in secundâ edtione hujusce operis, erui possunt.

In primâ editione hujusce operis reducuntur duae aequationes in unam ita ut incognitae quantitates exterminentur infinitarum serierum ope, quae methodus in hâc editione ad plures aequationes plures incognitas quantitates habentes applicatur ; etiamque datur modus hoc problema resolvendi, quae conjungit methodos

duits ainsi obtenus, puis il détermine les coefficients de ces quantités pour que s'annulent identiquement tous les termes dans lesquels se trouvent les quantités à éliminer ; dans bien des cas il donne aussi des multiplicateurs de formes plus simples, qui facilitent souvent l'élimination. A partir du principe des incréments de Taylor[4], d'une proposition facile à démontrer et que j'ai donnée dans la première édition de cet ouvrage (à savoir, que le nombre de coefficients dans l'équation de degré n à m inconnues $y^n + (a + bx + cz + ...)y^{n-1} + ... = 0$ sera $(n+1) \times \frac{n+2}{2} \times \frac{n+3}{3} ... \frac{n+m}{m} = S$) et d'une proposition élégante qu'il a trouvée (à savoir que, si $N(h)^n$ désigne le nombre de termes dans l'équation de degré n à h inconnues $x, y, z, v, ...$, alors le nombre de termes dans l'équation de degré n divisibles par x^p ou par x^q ou par z^r ou par v^s etc. sera $= (N(h)^{n-p} + N(h)^{n-q} + N(h)^{n-r} + N(h)^{n-s} + ...) - (N(h)^{n-p-q} + N(h)^{n-p-r} + N(h)^{n-p-s} + N(h)^{n-q-s} + ...) + (N(h)^{n-p-q-r} + N(h)^{n-p-q-s} + ...) - (N(h)^{n-p-q-r-s} + ...) + ...$), il a découvert les élégantes propositions suivantes. Soient h équations de degrés respectifs $n, m, l, k, ...$ contenant autant d'inconnues $x, y, z, v, ...$; soient $p, q, r, s, ..., p', q', r', ..., p'', q'', ...$ les degrés maximaux en les inconnues $x, y, z, v, ...$ des termes présents dans les équations respectives de degrés $n, m, l, k, ...$; alors l'équation dont la racine est x ou y ou z etc. sera de degré au plus $n \times m \times l \times k \times ... - (n-p) \times (m-p') \times (l-p'') \times ... - (n-q) \times (m-q') \times (l-q'') \times ... - (n-r) \times (m-r') \times ... - ... = P$. Mais si les degrés totaux en $(x,\ y)$ des équations ci-dessus sont a, a', a'',..., alors l'équation dont la racine est x ou y ou z etc. sera de degré $P + (n-a) \times (m-a') \times ... - (p+q-a) \times (m-a') \times ...$[5] Ces propositions seront peut-être plus faciles à établir en suivant le principe que j'énonce dans le problème 41 (que j'ai déjà publié dans la seconde édition de cet ouvrage).

Dans la première édition de cet ouvrage, on réduit deux équations à une seule en éliminant les inconnues par l'usage des séries infinies. Cette méthode est appliquée dans la présente édition à un plus grand nombre d'équations ayant un plus grand nombre d'inconnues ; on donne ainsi une manière de résoudre ce pro-

Crameri et *Bezout*. *Le Grange*[6] ope fluxionum et logarithmorum idem problema resolvit.

In primâ editione hujusce operis primum, quod scio, animadvertitur, si termini duarum datarum aequationum (n&m) dimensionum respective duas incognitas quantitates (x&y) involventium, in quibus maximae inveniuntur dimensiones praedictarum quantitatum, communem habeant divisorem, aequationem haud assurgere ad $n \times m$ dimensiones ; et in eâdem editione etiam adjicitur in proprietatibus curvarum, si modo tres sint aequationes (n, m, r) dimensionum tres incognitas quantitates (x, y, z) habentes, aequationem, cujus radix est x vel y vel z ascendere posse ad $n \times m \times r$ dimensiones. In secundâ editione hujusce operis, si modo dentur (h) aequationes, quarum dimensiones sunt respective $n, m, r, s,$ &c. ; totidem (h) incognitas quantitates $(x, y, z,$ &c.) habentes, docetur methodus e terminis maximarum datarum aequationum dimensionum inveniendi, annon aequatio, cujus radix est x vel y vel z, &c. , habet $n \times m \times r \times s \times$ &c. dimensiones ; quae methodus in hâc editione fusius explicatur : eâdem editione animadvertitur, si modo plures sint aequationes quam incognitae quantitates, et inveniatur communis divisor unam incognitam quantitatem (x) solummodo involvens, tot esse datarum aequationum radices, quot diversos valores habet incognita quantitas (x) ; ni duas vel plures radices aequales habeat praedicta incognita quantitas, &c. ; etiamque si numerus aequationum major sit quam numerus incognitarum quantitatum, et exterminentur omnes incognitae quantitates, resultare aequationes cognitas quantitates solummodo involventes, quae detegunt quibus casibus possunt esse aequationes hujusmodi ; &c. : dantur etiam exempla, in quibus continentur tres aequationes totidem incognitas quantitates habentes ; attamen una ex iis ad libitum assumi potest ; in his exemplis datur aequatio relationem inter cognitas quantitates exprimens : et in hâc editione asseritur, si modo dentur n aequationes (m) incognitas quantitates habentes, et in praedictis contineantur (r) independentes aequationes relationem solummodo inter cognitas quantitates exhibentes, tum assumi posse $(n - m + r)$ incognitas quantitates ad libitum ; haud possum dicere, annon quaedam de his in aequationibus dimensionum unam superantium ab aliis prius tradita fuere : *Bezout* dedit novam methodum inveniendi valorem e singulis incognitis quantitatibus in simplicibus aequatio-

blème qui conjugue les méthodes de Cramer et de Bézout. Lagrange a résolu le même problème au moyen des fluxions et des logarithmes.

Dans la première édition de cet ouvrage, on remarque, pour la première fois à ma connaissance, que l'équation ne sera pas de degré mn si, dans les deux équations données de degrés respectifs n et m et contenant les deux inconnues x et y, les termes de degrés maximaux en ces inconnues ont un diviseur commun. Et dans la même édition, j'ai ajouté, dans les *Proprietatibus curvarum*[7], que si on a trois équations de degrés n, m, r ayant trois inconnues x, y, z, alors l'équation dont la racine est x ou y ou z peut atteindre le degré $n \times m \times r$. Soient h équations, de degrés respectifs $n, m, r, s, ...,$ contenant le même nombre h d'inconnues $x, y, z, ...$ On enseigne dans la seconde édition de cet ouvrage une méthode pour trouver, à partir des termes de degrés maximaux des équations données, si oui ou non l'équation dont la racine est x ou y ou z etc. sera de degré $n \times m \times r \times s....$ Cette méthode est expliquée davantage dans la présente édition. Dans la seconde édition, on remarque le fait suivant. Soient des équations, en nombre supérieur à celui des inconnues. Si l'on trouve un diviseur commun ne contenant qu'une seule inconnue x, alors il y a autant de racines aux équations données, que de valeurs distinctes de l'inconnue x, sauf si cette inconnue a deux ou plusieurs racines égales, etc. Et ainsi si le nombre des équations est supérieur à celui des inconnues et que toutes les inconnues peuvent être éliminées, on obtient des équations qui ne contiennent que les quantités connues et qui révèlent dans quels cas de telles équations peuvent avoir lieu. On y donne aussi des exemples, dans lesquels il y a trois équations à trois inconnues, mais où la valeur de l'une d'elles peut être prise à volonté. Dans ces exemples il y a une équation qui exprime une relation entre les quantités connues. Et dans la présente édition on affirme que, si l'on a n équations à m inconnues et que parmi ces équations il y en a r indépendantes qui expriment une relation seulement entre des quantités connues, alors on peut prendre à volonté la valeur de $(n-m+r)$ de ces inconnues. Je ne peux pas dire si d'autres avant moi ont traité ce sujet pour les équations de degré supérieur à un. Bézout a donné une nouvelle méthode pour trouver la valeur de chaque inconnue dans les équations linéaires. Dans le présent ouvrage on expose aussi une nouvelle méthode pour trouver ces valeurs.

nibus ; nova methodus inveniendi praedictos valores etiam in hoc opere traditur.

1 Nous reproduisons et traduisons dans cette annexe un extrait de la préface à la seconde édition (1782) des *Meditationes Algebraicae* de Waring, p. *xvii* à *xx*. Quand, dans ce texte, Waring parle de la « seconde édition », il s'agit en fait de la première édition des *Meditationes Algebraicae* (1770 [323] ; un exemplaire de cette édition plus rare se trouve à la Sorbonne, sous la cote SXt 267). Waring désigne en effet comme « première édition » un ouvrage autre : les *Miscellanea analytica de aequationibus algebraicis et curvarum proprietatibus*, ouvrage paru en 1762. La traduction de cet extrait en anglais par D. Weeks (dans sa traduction des *Meditationes Algebraicae* [324]) est en partie erronée, peut-être à cause du fait qu'il n'avait pas à sa disposition le traité de Bézout.

2 C'est-à-dire que chacune s'exprime rationnellement en fonction de l'autre.

3 Ici nous corrigeons « par exemple x » en « par exemple y ».

4 C'est-à-dire le calcul au différences finies.

5 Pour traduire cette seconde proposition, nous nous aidons de Bézout [29] qui l'énonce plus clairement, § 80 p. 57.

6 *Sic*.

7 Il s'agit d'une partie des *Miscellanea analytica de aequationibus algebraicis et curvarum proprietatibus*, Cambridge 1762 (le passage auquel Waring fait ici allusion est dans le chapitre 3, p. 130 de cet ouvrage).

Bibliographie

[1] Niels Henrik Abel. « Recherche de la quantité qui satisfait à la fois à deux équations algébriques données ». *Annales de Mathématiques pures et appliquées rédigées par Gergonne, Paris*, *17*, 1827. ([2], t. I, p. 212–218).

[2] ..... *Œuvres complètes de Niels Henrik Abel.* Grondahl & Son, Christiania, 1881.

[3] Shreeram S. Abhyankar. *Algebraic space curves.* Presses de l'Université de Montréal, Montréal, 1971.

[4] ..... « Historical ramblings in algebraic geometry and related algebra ». *American Mathematical Monthly*, *83* : 409–448, 1976.

[5] Abū Kāmil. *The Algebra of Abū Kāmil*, éd. M. Levey. University of Wisconsin, Madison, Milwaukee, London, 1966.

[6] 'Umar al-Khayyām. « Traité du savant unique Abū al-Fatḥ 'Umar ibn Ibrāhīm al-Khayyāmī d'algèbre et d'*al-muqābala* ». *in* [315] p. 116-237.

[7] Al-Samaw'al. *Al-Bāhir en Algèbre d'As-Samaw'al*, éd. S. Ahmad et R. Rashed. Université de Damas, Damas, 1972.

[8] Kenza Asselah. « Arithmétique et algèbre dans la seconde moitié du XVII^{e} siècle français : les « Elemens & Nouveaux Elemens de Mathématiques » de Jean Prestet ». Thèse de doctorat de l'Université Paris 7, soutenue le 6 décembre 2005.

[9] Richard Baltzer. *Theorie und Anwendung der Determinanten mit Beziehung Auf Die Originalquellen.* Hirzel, Leipzig, 1857.

[10] ..... *Ueber die Auflösungen eines Systems von Gleichungen.* Dr. v. E. Blachmann u. Sohn, Dresde, 1868.

[11] ..... « Mathematische Bemerkungen ». *Berichte über die Verhandlungen der Kgl. Sächsischen Gesellschaft der Wissenschaften zu Leipzig*, *25* : 523–537, 1873.

[12] Izabella G. Bashmakova. « Sur l'histoire de l'algèbre commutative ». *Revue de Synthèse*, *89* : 185–202, 1968.

[13] ..... « Arithmetic of algebraic curves from Diophantus to Poincaré ». *Historia Mathematica*, *8* : 393–416, 1981.

[14] Dave Bayer. *The Division Algorithm and the Hilbert Scheme,*thèse présentée par D. Bayer. Harvard University, Cambridge, Massachusetts, 1982.

[15] Dave Bayer et Mike Stillman. « Computation of Hilbert functions ». *Journal of Symbolic Computation*, *14* : 31–50, 1992.

[16] Marouane Ben Miled. « Les quantités irrationnelles dans l'œuvre d'al-Karajī ». *in De Zénon d'Élée à Poincaré. Recueil d'études en hommage à Roshdi Rashed.* éd. Régis Morelon et Ahmad Hasnawi, Editions Peeters, Louvain – Paris, 2004, p. 27–54.

[17] Jacques Bernoulli. « Animadversio in geometriam cartesianam, et constructio quorundam problematum hypersolidorum ». *Acta Eruditorum*, pages 323–330, juin 1688.

[18] Jakob Bernoulli. *Ars Conjectandi.* Thurnisiorum fratrum, Basel, 1713.

[19] D. N. Bernshtein. « The number of roots of a system of equations ». *Functional Analysis and its Applications*, *9* : 183–185, 1975. traduit de *Funktsional'nyi Analiz i Ego Prilozheniya*, *9*(3) : 1–4, 1975.

[20] K. Bes. « La dépendance ou l'indépendance d'un système d'équations algébriques ». *Verhandelingen der Koninklijke Akademie van Wetenschappen te Amsterdam. Eerste sectie*, *8*(6) : 1–29, janvier 1904 (1903).

[21] ..... « L'équation finale ». *Verhandelingen der Koninklijke Akademie van Wetenschappen te Amsterdam. Eerste sectie*, *8*(1) : 1–61, juin 1901 (1903).

[22] ..... « Les systèmes de racines d'un système de n équations homogènes à $(n+1)$ variables ». *Verhandelingen der Koninklijke Akademie van Wetenschappen te Amsterdam. Eerste sectie*, *8*(2) : 1–52, mai 1902 (1903).

[23] ..... « Théorie générale de l'élimination d'après la méthode de Bézout, suivant un nouveau procédé ». *Verhandelingen der Koninklijke Akademie van Wetenschappen te Amsterdam. Eerste sectie*, *6*(7) : 1–121, septembre 1899.

[24] Etienne Bézout. « Mémoire sur les courbes dont la rectification dépend d'une quantité donnée ». *Histoire de l'Académie Royale des Sciences*, page 65, 1758.

[25] ..... « Mémoire sur plusieurs classes d'équations de tous les degrés qui admettent une solution algébrique ». *Histoire de l'Académie Royale des Sciences*, pages 17–52, 1762.

[26] ..... « Sur le degré des équations résultantes de l'évanouissement des inconnues ». *Histoire de l'Académie Royale des Sciences*, pages 288–338, 1764.

[27] ..... « Mémoire sur la résolution générale des équations de tous les degrés ». *Histoire de l'Académie Royale des Sciences*, pages 533–552, 1765.

[28] ..... « Expériences faites par ordre de l'Académie, sur le froid de l'année 1776 ». *Histoire de l'Académie Royale des Sciences*, page 505, 1777.

[29] ..... *Théorie générale des équations algébriques*. Ph.-D. Pierres, Paris, 1779.

[30] Rudolf Bkouche. « La naissance du projectif : de la perspective à la géométrie projective ». *in Mathématiques et philosophie de l'antiquité à l'âge classique*, ed. R. Rashed, Editions du CNRS, 1991, pp. 239–285.

[31] Martin Bloch. « Beiträge zur Theorie der Resultantensysteme, welche bei der Bestimmung des größten gemeinschaftlichen Teilers zweier ganzen Funktionen einer Variablen auftreten ». Thèse Université Justus Liebig, Gießen, 1892.

[32] Ossian Bonnet. « Mémoire sur la résolution de deux équations à deux inconnues ». *Nouvelles Annales de Mathématiques*, *6* : 54–63, 135–150 et 243–252, 1847.

[33] G. Boole. « Exposition of a general theory of linear transformations ». *Math. Journ. Cambridge Dublin*, *3* : 1–20, 106–119, 1843.

[34] ..... « Notes on linear transformations ». *Cambridge Mathematical Journal*, *4* : 167–171, 1843–1845.

[35] C. W. Borchardt. « Développements sur l'équation à l'aide de laquelle on détermine les inégalités séculaires du mouvement des planètes ». *Journal de Mathématiques pures et appliquées, fondé par Liouville, Paris*, *12* : 50–67, 1847.

[36] ..... « Bemerkung über die beiden vorstehenden Aufsätze ». *Journal für die reine und angewandte Mathematik, fondé par Crelle, Leipzig*, *53* : 281–283, 1857. (fait suite à un article de Hermite, et à un texte posthume de Jacobi, concernant tous deux la loi d'inertie des formes quadratiques).

[37] H. J. M. Bos. « On the representation of curves in Descartes' Géométrie ». *Archives for History of Exact Sciences*, *24*(4) : 295–338, 1981.

[38] W. Braikenridge. *Exercitatio geometrica de descriptione linearum curvarum*. Ric. Hett & Joh. Nourse, Londres, 1733.

[39] Bret. « Théorie de l'élimination, entre deux équations de degrés quelconques, fondée sur la méthode du plus grand commun diviseur ». *Annales de Mathématiques pures et appliquées rédigées par Gergonne, Paris*, *3* : 13–18, 1812-1813.

[40] A. Brill. « Elimination und Geometrie in den letzten Jahrzehnten ». *in Verhandlungen des dritten internationalen Mathematiker-Kongresses in Heidelberg 1904*, Leipzig 1915, p. 275–283.

[41] ..... « Ueber zwei Berührungsprobleme ». *Mathematische Annalen*, *4* : 527–549, 1871.

[42] ..... « Zur Theorie der Elimination und der algebraischen Curven ». *Mathematische Annalen*, *4* : 510–526, 1871.

[43] ..... « Über Elimination aus einem gewissen System von Gleichungen ». *Mathematische Annalen*, *5* : 378–396, 1872.

[44] ..... « Ueber die Discriminante ». *Mathematische Annalen*, *12* : 87–89, 1877.

[45] ..... « Ueber eine Eigenschaft der Resultante ». *Mathematische Annalen*, *16* : 345–347, 1880.

[46] ..... « Ueber Singularitäten ebener algebraischer Curven und eine neue Curvenspecies ». *Mathematische Annalen*, *16* : 348–408, 1880.

[47] ..... « Ueber binäre Formen und die Gleichung sechsten Grades ». *Mathematische Annalen*, *20* : 330–356, 1882.

[48] ..... « Ueber die reducirte Resultante ». *Abhandlungen der Königl. Kaiserl. Akademie der Wissenschaften zu München*, *17* : 91–101, 1888/1889 (1892).

[49] A. Brill et M. Noether. « Die Entwicklung der Theorie der algebraischen Functionen in älterer und neuerer Zeit ». *Jahresbericht der Deutschen Mathematiker Vereinigung*, *3* : 111–566, 1892–1893.

[50] Girolamo Cardano. *Artis Magnae, sive de regulis algebraicis liber unus.* Ioh. Petreium, Nürnberg, 1545.

[51] Henri Cartan. « Sur le théorème de préparation de Weierstrass ». *in Festschrift zur Gedächtnisfeier für Karl Weierstrass 1815–1965*, ed. Heinrich Behnke et Klaus Kopfermann, Westdeutscher Verlag, Köln Opladen, 1966, p. 155–168.

[52] J. W. S. Cassels et A. Fröhlich. *Algebraic Number Theory.* Academic Press, London, New York, 1967.

[53] Augustin Cauchy. « Sur la résolution des équations numériques et sur la théorie de l'élimination (1829) ». *in* [54], série II, vol. IX, p. 87–161.

[54] ..... *Œuvres complètes d'Augustin Cauchy.* Série 1, 12 vol. ; série 2, 15 vol. Gauthiers-Villars, Paris, 1882–1974.

[55] Arthur Cayley. « Notes sur deux formules données par M. M. Eisenstein et Hesse ». *Journal für die reine und angewandte Mathematik, fondé par Crelle, Leipzig*, *29* : 54–57, 1845. (*in* [72] t. 1, p. 113–116).

[56] ..... « On the theory of linear transformations ». *Cambridge Mathematical Journal*, *3* : 193–209, 1845. (*in* [72] t. 1, p. 80–94).

[57] ..... « Mémoire sur les hyperdéterminants ». *Journal für die reine und angewandte Mathematik, fondé par Crelle, Leipzig*, *30* : 1–37, 1846. (version française améliorée de « On the Theory of Linear Transformations », *Camb. Math. Journ.*, *4* : 193–209, 1845, *in* [72] t. 1, p. 80–94).

[58] ..... « Note sur les fonctions de M. Sturm ». *Journal de Mathématiques pures et appliquées, fondé par Liouville, Paris*, *11* : 297–299, 1846. (*in* [72] t. 1, p. 306–308).

[59] ..... « On homogeneous functions of the third order with three variables ». *Cambridge and Dublin Mathematical Journal*, *1* : 97–104, 1846. (*in* [72] t. 1 p. 230–233).

[60] ..... « On the theory of involution in geometry ». *Cambridge and Dublin Mathematical Journal*, *2* : 52–61, 1847. (*in* [72], t. 1, p. 259–266).

[61] ..... « Recherches sur l'élimination, et sur la théorie des courbes ». *Journal für die reine und angewandte Mathematik, fondé par Crelle, Leipzig*, *34* : 30–45, 1847. (*in* [72], t. 1, p. 337–351).

[62] ..... « On the theory of elimination ». *Cambridge and Dublin Mathematical Journal*, *3* : 116–120, 1848. (*in* [72], t. 1, p. 370–374).

[63] ..... « Nouvelles recherches sur les covariants ». *Journal für die reine und angewandte Mathematik, fondé par Crelle, Leipzig*, *47* : 109–124, 1854. (*in* [72], t. 2 p. 164).

[64] ..... « A Third Memoir upon Quantics ». *Philosophical Transactions*, *146* : 627–647, 1856. (*in* [72] t. 2 p. 310–335).

[65] ..... « Note sur la méthode d'élimination de Bézout ». *Journal für die reine und angewandte Mathematik, fondé par Crelle, Leipzig*, *53* : 366–367, 1857. (*in* [72], t. 4, p. 38).

[66] ..... « A Fifth Memoir upon Quantics ». *Philosophical Transactions*, *148* : 429–460, 1858. (*in* [72], t. 2, p. 527–557).

[67] ..... « Note sur l'élimination ». *Journal für die reine und angewandte Mathematik, fondé par Crelle, Leipzig*, *60* : 373–374, 1861. (*in* [72], t. 5 p. 157–159).

[68] ..... « Nouvelles recherches sur l'élimination et la théorie des courbes ». *Journal für die reine und angewandte Mathematik, fondé par Crelle, Leipzig*, *63* : 34–39, 1864. (*in* [72] t. 5 p. 162–167).

[69] ..... « Suite des recherches sur l'élimination et la théorie des courbes ». *Journal für die reine und angewandte Mathematik, fondé par Crelle, Leipzig*, *64* : 167–171, 1865. (*in* [72] t. 5 p. 416–420).

[70] ..... « Note on the Discriminant of a Binary Quantic ». *Quart. Math. Journ.*, *10* : 23, 1870. (*in* [72] t. 7 p. 303).

[71] ..... « Example of a special discriminant ». *Quarterly Journal of Pure and Applied Mathematics*, *11* : 211–213, 1871. (*in* [72] t. 8 p. 46–47).

[72] ..... *The collected mathematical papers.* 13 volumes et un volume supplémentaire. Cambridge University Press, Cambridge, 1889–1898.

[73] C. Chevalley. « Les classes d'équivalence rationnelle ». *in Séminaire C. Chevalley*, E. N. S. Paris, 1958, p. 2-01 à 2-14 et 3-01 à 3-18.

[74] A. Clebsch. « Über eine Klasse von Eliminations-Problemen und über einige Punkte der Theorie der Polaren ». *Journal für die reine und angewandte Mathematik, fondé par Crelle, Leipzig*, *58* : 273–291, 1861.

[75] ..... « Über die Anwendung der Abelschen Functionen in der Geometrie ». *Journal für die reine und angewandte Mathematik, fondé par Crelle, Leipzig*, *63* : 189–243, 1863.

[76] ..... « Über die Elimination aus zwei Gleichungen dritten Grades ». *Journal für die reine und angewandte Mathematik, fondé par Crelle, Leipzig*, *64* : 95–97, 1865.

[77] A. Clebsch et P. Gordan. *Theorie der Abelschen Functionen.* Teubner, Leipzig, 1866.

[78] ..... « Sulla rappresentazione tipica delle forme binarie ». *Annali di Matematica Pura e Applicata*, *2*(1) : 23–79, 1867/1868.

[79] Condorcet. « Éloge de Mr. Bézout ». *Histoire de l'Académie Royale des Sciences*, (1783) 1786.

[80] P. Costabel. « L'initiation mathématique de Descartes ». *Archives de Philosophie*, *46* : 637–646, 1983.

[81] Gabriel Cramer. *Introduction à l'analyse des lignes courbes algébriques.* Frères Cramer et Cl. Philibert, Genève, 1750.

[82] M. J. Crowe. *A history of vector analysis. The evolution of the idea of a vectorial system.* Notre Dame, Londres, 1967.

[83] Pappus d'Alexandrie. *La collection mathématique*, trad. Paul Ver Eecke, 2 tomes. A. Blanchard, Paris, 1982.

[84] Florimond de Beaune. « De Aequationum natura, constitutione, & limitibus opuscula duo ». éd. Erasme Bartholin, dans [89].

[85] Jean Paul de Gua de Malves. *Usages de l'analyse de Descartes pour découvrir, sans le secours du calcul différentiel, les propriétés, ou affections principales des lignes géométriques de tous les ordres.* Briasson, Paris, 1740.

[86] de Jonquières. « Note sur la Géométrie Organique de Maclaurin ». *Journal de Mathématiques pures et appliquées, fondé par Liouville, Paris*, 1857.

[87] de L'Hospital. *Traité analytique des sections coniques et de leur usage pour la résolution des équations dans les problèmes tant déterminez qu'indéterminez.* Veuve Jean Bordot et fils, Paris, 1707.

[88] René Descartes. *La géométrie.* Jean Maire, Leyde, 1637. (cf. [90], t. VI, p. 368–485).

[89] ..... *Geometria Renati Cartesii*, éd. Frans van Schooten. Elzevir, Amsterdam, 1659–1661.

[90] ..... *Œuvres de Descartes publiées par Charles Adam et Paul Tannery.* Vrin, Paris, 1986.

[91] ..... *Exercices pour les éléments des solides*, éd. P. Costabel. P.U.F., Paris, 1987.

[92] A. Dickenstein, N. Fitchas, M. Giusti, et C. Sessa. « The membership problem for unmixed polynomial ideals is solvable in subexponential time ». to appear *in* Discrete Appl. Math. Proc. AAECC-7, Toulouse, 1989.

[93] Jean Dieudonné. *Cours de géométrie algébrique*, vol. I, *Aperçu historique sur le développement de la géométrie algébrique.* Presses Universitaires de France, Paris, 1974.

[94] ..... *Abrégé d'histoire des mathématiques 1700–1900*, 2 vol. Hermann, Paris, 1978.

[95] Jean A. Dieudonné et James B. Carrell. *Invariant Theory, Old and New.* Academic Press, New York, London, 1971.

[96] Diophante. *Les arithmétiques*, éd. R. Rashed. Les Belles Lettres, Paris, 1984.

[97] Harold M. Edwards. « The Genesis of Ideal Theory ». *Archive for History of Exact Sciences*, *23* : 321–378, 1980.

[98] ..... *Divisor Theory.* Birkhäuser, Boston, 1990.

[99] Harold M. Edwards, O. Neumann, et W. Purkert. « Dedekinds « Bunte Bemerkungen » zu Kroneckers « Grundzüge » ». *Archive for History of Exact Science*, *27* : 49–85, 1982.

[100] David Eisenbud. *Commutative Algebra with a View Toward Algebraic Geometry.* Springer-Verlag, 1995.

[101] Eisenstein. « Allgemeine Auflösung der Gleichungen von den ersten vier Graden ». *Journal für die reine und angewandte Mathematik, fondé par Crelle, Leipzig*, *27* : 81–83, 1844.

[102] Leonhard Euler. « De formis radicum aequationum cuiusque ordinis coniectatio ». *Commentarii Academiae Scientiarum Petropolitanae*, *6* : 216–231, (1732/1733) 1738. (cf. [108], E30).

[103] ..... « Démonstration sur le nombre des points où deux lignes des ordres quelconques peuvent se rencontrer ». *Mémoires de l'académie des sciences de Berlin*, *4* : 234–248, 1748. (cf. [108], série 1, vol. XXVI, p. 46–59).

[104] ..... *Introductio in analysin infinitorum.* Bousquet, Lausanne, 1748. (trad. franç. J. B. Labey, *Introduction à l'analyse infinitésimale*, Barrois, Paris, 1797).

[105] ..... « Sur une contradiction apparente dans la doctrine des lignes courbes ». *Mémoires de l'académie des sciences de Berlin*, *4* : 219–233, 1748. (cf. [108], série 1, vol. XXVI, p. 33–45).

[106] ..... « De resolutione aequationum cuiusvis gradus ». *Novi Commentarii Academiae Scientiarum Petropolitanae*, *9* : 70–98, 1764. (cf. [108], E282).

[107] ..... « Nouvelle méthode d'éliminer les quantités inconnues des équations ». *Mémoires de l'académie des sciences de Berlin*, *20* : 91–104, 1764. (cf. [108], serie 1, vol. VI, p. 197–211).

[108] ..... *Leonhardi Euleri Opera Omnia*, éd., F. Rudio, A. Krazer, P. Stackel. Série 1, 29 vol. ; série 2, 31 vol. ; série 3, 12 vol. ; série 4A, correspondance, 7 vol. B. G. Teubner, Leipzig Berlin Zurich Basel, 1911.

[109] Johannes Faulhaber. *Newer Arithmetischer Wegweyser. Zu der Hochnutzlichen freyen Rechenkunst, mit Newen Inventionibus geziert. Daraus ein fleissiger Praeceptor, mit Göttlicher Hülff / auch die harten Ingenia der Jugent (vermittelst der didactica) von einer Staffel zu der andern / fruchtbarlich laiten und füren kan / biß sie die Species und Exempla, in ganzen vnd gebrochenen Zahlen gründlich erlernen mögen. Auß den Aller erfahrnesten / bewehrtesten vnnd Kunstreichsten Authorn dieser Kunst / mit fleiß extrahirt, vnd zusamen getragen.* Johann Meder, Ulm, 1617.

[110] ..... *Miracula Arithmetica. Zu der Continuation seiner Arithmetischen Wegweisers gehörig.* David Franck, Augsburg, 1622.

[111] ..... *Academia Algebrae. Darinnen die miraculosische Inventiones / zu den höchsten Cossen weiters continuiert vnd proficiert werden. Dergleichen zwar vor 15. Jahren den Gelehrten auff allen Vniversiteten in gantzem Europa proponiert, darauff continuirt, auch allen Mathematicis inn der gantzen weiten Welt dedicirt, aber biß hero / noch nie so hoch / bißauff die regulierte / Zensicubiccubic Coß/ durch offnen Truck publiciert worden. Welcher vorgesetzet ein kurtz Bedencken / Was einer für Authores nach ordnung gebrauchen solle / welcher die Coß fruchtbarlich / bald / auch fundamentaliter lehrnen und ergreiffen will.* Johann Ulrich Schönigk, Augsburg, 1631.

[112] Pierre Fermat. *Œuvres de Fermat*, éd. P. Tannery et Ch. Henry. Gauthier-Villars, Paris, 1891.

[113] ..... *Œuvres de Pierre de Fermat. I. La théorie des nombres.* éd. R. Rashed, C. Houzel, G. Christol. Albert Blanchard, Paris, 1999.

[114] Menso Folkerts. « Conrad Landvogt, ein bisher unbekannter Algebraiker um 1500 ». *in Amphora, Festschrift für Hans Wussing zu seinem 65. Geburtstag*, éd. S. S. Demidov, M. Folkerts, D. E. Rowl et O. J. Scriba, Birkhäuser, Basel 1992.

[115] Joseph Fourier. *Analyse des équations déterminées. Première partie* (édité par C. L. M. H. Navier). Firmin-Didot et fr., Paris, 1831.

[116] William Fulton. *Intersection Theory.* Springer-Verlag, 1983.

[117] ..... *Introduction to intersection theory in algebraic geometry.* Regional conference series in mathematics, no. 54. American Mathematical Society, 1984.

[118] ..... *Introduction to Toric Varieties.* Princeton University Press, 1993.

[119] Evariste Galois. *Ecrits et mémoires mathématiques d'Evariste Galois*, éd., Robert Bourgne et J.-P. Azra. Gauthier-Villars, Paris, 1962.

[120] C. F. Gauss. *Disquisitiones Arithmeticae.* Gerh. Fleischer jun., Lipsiae, 1801.

[121] Carl Friedrich Gauss. *Werke*, 12 vol. Göttingen, 1863–1929.

[122] Albert Girard. *Invention nouvelle en l'algèbre.* Guillaume Ianssen Blaeuw, Amsterdam, 1629.

[123] Marc Giusti. « Sections des variétés déterminantielles par les plans de coordonnées ». *in Proc. Int. Conf. on Alg. Geometry*, La Rabida 1981, Springer Verlag, 1982, p. 103–118.

[124] ..... « Combinatorial dimension theory of algebraic varieties ». *Journal of Symbolic Computation*, *6* : 249–265, 1988.

[125] Marc Giusti et Joos Heintz. « Algorithmes –disons rapides– pour la décomposition d'une variété algébrique en composantes irréductibles et équidimensionnelles ». *in* Effective Methods in Algebraic Geometry (Proceedings of MEGA'90), T. Mora et C. Traverso (eds.), Progress in Math. 94, Birkhäuser, 1991, p. 169–193.

[126] ..... « La détermination des points isolés et de la dimension d'une variété algébrique peut se faire en temps polynomial ». *in* Computational Algebraic Geometry and Commutative

Algebra, ed. D. Eisenbud et L. Robbiano, Cambridge University Press, Cambridge, 1993, p. 216–256.

[127] Marc Giusti, Joos Heintz, et Juan Sabia. « On the efficiency of effective Nullstellensätze ». *Computational Complexity*, *3* : 56–95, 1993.

[128] Marc Giusti, Grégoire Lecerf, et Bruno Salvy. « A Gröbner free alternative for polynomial system solving ». *Journal of Complexity*, *17* : 154–211, 2001.

[129] Judith V. Grabiner. article « Bézout, Étienne » *in* C. C. Gillispie, éd., *Dictionary of Scientific Biography*, 16 vol., vol. II, p. 111–114. Charles Scribner's Sons, New York, 1970.

[130] Gilles-Gaston Granger. *Formes, opérations, objets*. Vrin, Paris, 1994.

[131] Hermann Günther Grassmann. *Die lineale Ausdehnungslehre, eine neuer Zweig der Mathematik, dargestellt und durch Anwendungen auf die übrigen Zweiger der Mathematik, wie auch auf die Statik, Mechanik, die Lehre vom Magnetismus und die Krystallonomie erlaütert*. Otto Wigand, Leipzig, 1844. (trad. D. Flament, *La science de la grandeur extensive, la lineale Ausdehnungslehre*, Albert Blanchard, Paris, 1994).

[132] Jeremy J. Gray. « Algebra in der Geometrie von Newton bis Plücker ». *Math. Semesterber.*, *36* : 175–204, 1989.

[133] ..... « Algebraic Geometry between Noether and Noether – a forgotten chapter in the history of algebraic geometry ». *Revue d'Histoire des Mathématiques*, *3* : 1–48, 1997.

[134] L. Gruson. « Un aperçu des travaux mathématiques de G. H. Halphen (1844–1889) ». *in Complex Projective Geometry*, eds. Ellingsrud, Peskine, Sacchiero, Stromme, Cambridge University Press 1992, p. 189–198.

[135] W. Gröbner. *Moderne algebraische Geometrie*. Springer-Verlag, 1949.

[136] Karlheinz Haas. « Die mathematischen Arbeiten von Johann Hudde (1628–1704) Bürgmeister von Amsterdam ». *Centaurus*, *4*(3) : 235–284, 1956.

[137] Georges-Henri Halphen. « Mémoire sur la détermination des coniques et des surfaces du second ordre ». *Bulletin de la Société Mathématique*, pages t. I p. 130 et 226, t. II p. 11, 1873. (*in* [138] t. 1 p. 98–157.

[138] ..... *Œuvres*. Gauthier-Villars, Paris, 1916.

[139] M. Hamburger. « Über die Entwickelung algebraischer Functionen in Reihen ». *Zeitschrift für Math. und Phys.*, *16* : 461–491, 1871.

[140] Robin Hartshorne. *Algebraic Geometry*. Springer, New York, 1977.

[141] Kurt Hawlitschek. *Johann Faulhaber 1580–1635. Eine Blütezeit der mathematischen Wissenschaften in Ulm*. Stadtbibliothek Ulm, Ulm, 1995.

[142] Charles Hermite. « Sur l'extension du théorème de M. Sturm à un système d'équations simultanées ». présenté à l'Académie en 1852, fait l'objet d'un extrait dans les Comptes rendus des séances de l'Académie des Sciences vol. 35, puis d'une édition posthume en 1912 *in* [144] vol. 3 p. 1–34.

[143] ..... « Remarques sur le théorème de M. Sturm ». *Comptes rendus des séances de l'Académie des Sciences*, *36*, 1853. (*in* [144] vol. I p. 284–287).

[144] ..... *Œuvres*, ed. E. Picard. Gauthier-Villars, Paris, 1905–1917. (4 volumes).

[145] Otto Hesse. « De curvis et superficiebus secundi ordinis ». *Journal für die reine und angewandte Mathematik, fondé par Crelle, Leipzig*, *20* : 285–308, 1840.

[146] ..... « Ueber die lineare Construction des achten Schnittpunktes dreier Oberflächen zweiter Ordnung, wenn sieben Schnittpunkte derselben gegeben sind ». *Journal für die reine und angewandte Mathematik, fondé par Crelle, Leipzig*, *26* : 147–154, 1843.

[147] ..... « Über die Bildung der Endgleichung, welche durch Elimination einer Variabeln aus zwei algebraischen Gleichungen hervorgeht, und die Bestimmung ihres Grades ». *Journal für die reine und angewandte Mathematik, fondé par Crelle, Leipzig*, *27*(1) : 1–5, 1844.

[148] ..... « Über die Elimination der Variabeln aus drei algebraischen Gleichungen vom zweiten Grade mit zwei Variabeln ». *Journal für die reine und angewandte Mathematik, fondé par Crelle, Leipzig*, *28*(1) : 68–96, 1844.

[149] David Hilbert. « Über die Singularitäten der Discriminantenfläche ». *Mathematische Annalen*, *30* : 437–441, 1887.

[150] John Hoe. *Les systèmes d'équations polynômes dans le Siyuan Yujian (1303)*. Mémoires de l'Institut des Hautes Études Chinoises, vol. VI. Collège de France, 1977.

[151] Joseph E. Hofmann. *Michael Stifel (1487?–1567) : Leben, Wirken und Bedeutung für die Mathematik seiner Zeit*. Steiner, Wiesbaden, 1968.

[152] ..... *Leibniz in Paris*. Cambridge, 1974.

[153] Annick Horiuchi. *Les mathématiques japonaises à l'époque d'Edo*. Vrin, Paris, 1994.

[154] Christian Houzel. « Euler et l'apparition du formalisme ». *in Philosophie et calcul de l'infini*, C. Houzel, J.-L. Ovaert, P. Raymond, J.-J. Sansuc, 1976, François Maspero, Paris.

[155] ..... « La résolution algébrique des équations ». *in Sciences à l'époque de la Révolution française*, Albert Blanchard, 1988, pp. 17-37.

[156] ..... « The Work of Niels Henrik Abel ». *in Legacy of Niels Henrik Abel : The Abel Bicentennial, Oslo, 2002*, ed. Olav Arnfinn Laudal, Springer Verlag, 2004, p. 21–178.

[157] ..... *La géométrie algébrique. Recherches historiques*. Albert Blanchard, Paris, 2002.

[158] Johann Hudde. « Epistola prima de reductione aequationum ». *in* [89], p. 401–506.

[159] A. Hurwitz. « Über dit Trägheitsformen eines algebraischen Moduls ». *Annali die Matematica pura ed applica*, *3*(20) : 113–151, 1913.

[160] Ibn Sinān. *Ibrāhīm Ibn Sinān, logique et géométrie au X^e siècle*, éd. R. Rashed et H. Bellosta. Brill, Leiden, Boston, Köln, 2000.

[161] Jean Itard. *Essais d'histoire des mathématiques*. Albert Blanchard, Paris, 1984.

[162] C. G. Jacobi. « Theoremata nova algebraica circa systema duarum aequationum, inter duas variabiles propositarum ». *Journal für die reine und angewandte Mathematik, fondé par Crelle, Leipzig*, *14*(4) : 281–288, 1835.

[163] ..... « De eliminatione variabilis e duabus aequationibus algebraicis ». *Journal für die reine und angewandte Mathematik, fondé par Crelle, Leipzig*, *15* : 101–124, 1836.

[164] George B. Jerrard. *Mathematical Researches*. William Strong, Bristol, 1832–1835.

[165] F. Joachimsthal. « Remarques sur la condition de l'égalité de deux racines d'une équation algébrique ; et sur quelques théorèmes de géométrie, qui en suivent ». *Journal für die reine und angewandte Mathematik, fondé par Crelle, Leipzig*, *33*(4) : 371–376, 1846.

[166] Phillip S. Jones. article « Cramer, Gabriel » *in* C. C. Gillispie, éd., *Dictionary of Scientific Biography*, 16 vol., vol. III, p. 459–462. Charles Scribner's Sons, New York, 1970.

[167] Jean-Pierre Jouanolou. *Idéaux résultants*. Université Louis Pasteur, Département de Mathématiques, Strasbourg, 1980. (aussi paru dans *Adv. in Math.*, *37*(3) : 212–238, 1980).

[168] ..... « Le formalisme du résultant ». *Advances in Mathematics*, *90*(2) : 117–263, 1991.

[169] ..... « Aspects invariants de l'élimination ». *Advances in Mathematics*, *114* : 1–174, 1995.

[170] ..... « Formes d'inertie et résultant : un formulaire ». *Advances in Mathematics*, *126* : 93–118, 1997.

[171] Wolfgang Kaunzner. « Über Charakteristika in der mittelalterlichen abendländischen Mathematik ». *Mathematische Semesterberichte*, *34*(2) : 143–186, 1987.

[172] Steven L. Kleiman. « Chasles's enumerative theory of conics : a historical introduction ». *in Studies in Algebraic Geometry*, ed. A. Seidenberg, *The Mathematical Association of America, Studies in Mathematics* vol. 20, 1980, p. 117–143.

[173] Eberhard KNOBLOCH. *Die mathematischen Studien von G. W. Leibniz zur Kombinatorik. Auf Grund fast ausschliesslich handschriftlicher Aufzeichnungen dargelegt und kommentiert.* Franz Steiner Verlag GmbH, Wiesbaden, 1973.

[174] ..... « Unbekannte Studien von Leibniz zur Eliminations- und Explicationstheorie ». *Archive for history of exact sciences*, *12* : 142–173, 1974.

[175] ..... « Déterminants et élimination chez Leibniz ». *Revue d'histoire des sciences*, *54*(2) : 143–164, 2001.

[176] Julius KÖNIG. *Einleitung in die allgemeine Theorie der algebraischen Gröszen.* 1903.

[177] L. KRONECKER. « Über die algebraisch auflösbaren Gleichungen I ». *Monatsberichte der Königlich Preussischen Akademie der Wissenschaft zu Berlin*, pages 365–374, 1853. (cf. *Werke*, t. IV, p. 3–11).

[178] ..... « Über einige Interpolationsformeln für ganze Functionen mehrer Variabeln ». *Monatsberichte der Königlich Preussischen Akademie der Wissenschaft zu Berlin*, pages 95–121, 1865. (cf. *Werke*, t. I, p. 133-142).

[179] ..... « Über Abelsche Gleichungen ». *Monatsberichte der Königlich Preussischen Akademie der Wissenschaft zu Berlin*, pages 845–851, 1877. (cf. *Werke*, t. IV, p. 63–71).

[180] ..... « Auszug aus einem Briefe des Herrn Kronecker an E. Schering ». *Nachrichten von der Königl. Gesellschaft der Wissenschaften und der Georg-Augusts-Universität zu Göttingen*, pages 271–279, 1881.

[181] ..... « Über die Discriminante algebraischer Functionen einer Variabeln ». *Journal für die reine und angewandte Mathematik*, *91* : 301–334, 1881.

[182] ..... « Zur Theorie der Elimination einer Variabeln aus zwei algebraischen Gleichungen ». *Monatsberichte der Königlich Preussischen Akademie der Wissenschaft zu Berlin*, pages 535–600, 1881. (cf. *Werke*, t. II, p. 113–192).

[183] ..... « Grundzüge einer arithmetischen Theorie der algebraischen Grössen ». *Journal für die reine und angewandte Mathematik*, *92* : 1–122, 1882.

[184] ..... « Zur Theorie der Formen höherer Stufen ». *Monatsber. Akad. Wiss. Berlin*, pages 957–960, 1883. (cf. *Werke*, t. II, p. 419–424).

[185] Abraham Gotthelf KÄSTNER. *Geschichte der Mathematik.* Johann Georg Rosenbusch, Göttingen, 1796–1800.

[186] A. G. KUSHNIRENKO. « Newton polytopes and the Bezout theorem ». *Functional Analysis and its Applications*, *10* : 233–235, 1976. traduit de *Funktsional'nyi Analiz i Ego Prilozheniya*, *10*(3) : 82–83, 1976.

[187] LABBATIE. *Méthode d'élimination par le plus grand commun diviseur.* Bachelier, Paris, 1832, 1833 ou 1835. (8+58 pages).

[188] Joseph LAGRANGE. « Sur l'élimination des inconnues dans les équations ». *Mémoires de l'Académie de Berlin*, *25* : 303–318, (1769) 1771. (cf. Œuvres, t. III, p. 141–154).

[189] ..... « Réflexions sur la résolution algébrique des équations ». *Nouveaux Mémoires de l'Académie royale des Sciences et Belles-Lettres de Berlin*, pages 134–215 (1770), 138–253 (1771), 1770–1771. (cf. Œuvres, t. III, p. 205–421).

[190] ..... « Sur la forme des racines imaginaires des équations ». *Nouveaux Mémoires de l'Académie royale des Sciences et Belles-Lettres de Berlin*, pages ?– ?, 1772. (cf. Œuvres, t. III, p. 479–518).

[191] Serge LANG. *Algebra, Third Edition.* Addison-Wesley, Reading, 1993.

[192] LAPLACE, LAGRANGE, et MONGE. *L'Ecole Normale de l'an III. Leçons de mathématiques*(ed. J. Dhombres). Dunod, Paris, 1992.

[193] E. LASKER. « Zur Theorie der Moduln und Ideale ». *Mathematische Annalen*, *60* : 20–116 et 607–608, 1905.

[194] H. Laurent. « Mémoire sur la théorie de l'élimination ». *Nouvelles Annales de Mathématiques*, *3*(2) : 145–160, 1883.

[195] ..... *Traité d'analyse*. Gauthier-Villars, Paris, 1885. 6 volumes.

[196] ..... « Mémoire sur les équivalences algébriques et l'élimination ». *Nouvelles Annales de Mathématiques*, *3*(5) : 432–447 et 456–460, 1886.

[197] ..... « Sur le calcul d'une fonction symétrique ». *Nouvelles Annales de Mathématiques*, *3*(6) : 416, 1887.

[198] ..... « Sur la théorie de l'élimination ». *Nouvelles Annales de Mathématiques*, *3*(7) : 60–65 et 116–119, 1888.

[199] ..... « Sur l'élimination ». *Nouvelles Annales de Mathématiques*, *3*(11) : 5–7, 1892.

[200] ..... « Démonstration d'une formule qui donne, sous forme explicite, la résultante de plusieurs équations algébriques ». *Nouvelles Annales de Mathématiques*, *3*(12) : 305–315, 1893.

[201] ..... « Sur la détermination des facteurs étrangers introduits dans l'élimination ». *Journal de Mathématiques Spéciales*, *4*(2) : 97–98 et 121–122, 1893.

[202] ..... « Sur les fonctions symétriques ». *Journal de Mathématiques Spéciales*, *4*(2) : 49–51 et 217–220, 1893.

[203] ..... « Sur l'élimination ». *Nouvelles Annales de Mathématiques*, *3*(12) : 355–359, 1893.

[204] ..... « Sur les fonctions entières ». *Nouvelles Annales de Mathématiques*, *3*(15) : 23–28, 1896.

[205] ..... *L'élimination*. C. Naud, Paris, 1900.

[206] D. Lazard. « Algèbre linéaire sur $K[X_1, ..., X_n]$ et élimination ». *Bulletin de la Société Mathématique de France*, *105* : 165–190, 1977.

[207] Luis Laïta. « The Influence of Boole's Search for a Universal Method in Analysis on the Creation of his Logic ». *Annals of Science*, *34* : 163–176, 1977.

[208] Grégoire Lecerf. « Une alternative aux méthodes de réécriture pour la résolution de systèmes algébriques ». thèse de doctorat de l'Ecole Polytechnique, 2001.

[209] Gottfried Wilhelm Leibniz. « De aequationibus pluribus ad unam reducendo ». (7. Februar 1676) *in* [213], $7^{\text{ème}}$ série, t. 1, $2^{\text{ème}}$ partie.

[210] ..... « De resolutionibus aequationum cubicarum triradicalium. De radicibus realibus, quae interventu imaginariarum exprimuntur. Deque sexta quadam operatione arithmetica ». (Oktober 1675) *in* [213], $7^{\text{ème}}$ série, t. 2, p. 678–700.

[211] ..... *Der Briefwechsel mit Mathematikern*, éd. C. I. Gerhardt, seulement un volume paru. Berlin, 1899.

[212] ..... *Der Beginn der Determinantentheorie, Leibnizens nachgelassene Studien zum Determinantenkalkül*, éd. Eberhard Knobloch. Gerstenberg, Hildesheim, 1980.

[213] ..... *Sämtliche Schriften und Briefe*, éd. Preuß. Akad. d. Wiss. Darmstadt, Leipzig, Berlin, depuis 1923.

[214] Joseph Liouville. « Mémoire sur quelques propositions générales de géométrie et sur la théorie de l'élimination dans les équations algébriques ». *Journal de Mathématiques pures et appliquées, fondé par Liouville, Paris*, *6* : 345–411, 1841.

[215] ..... « Sur les équations algébriques à plusieurs inconnues ». *Journal de Mathématiques pures et appliquées, fondé par Liouville, Paris*, *12* : 68–72, 1847.

[216] F. S. Macaulay. « Some formulae in elimination ». *Proceedings of the London Mathematical Society*, *1*(35) : 3–27, 1903.

[217] ..... *The Algebraic Theory of Modular Systems*. Cambridge University Press, Cambridge, 1916.

[218] Colin MACLAURIN. « De linearum geometricarum proprietatibus generalibus tractatus ». 65 pages + 12 tables de figures, publié de manière posthume en 1748 en appendice à [222].

[219] ..... « Nova methodus universalis curvas omnes cujuscunque ordinis mechanicae describendi sola datorum angulorum et rectarum ope ». *Philosophical transactions*, *30* : 939–945, 1717–1719.

[220] ..... *Geometria organica : sive descriptio linearum curvarum universalis.* Gul. et Joh. Innys, London, 1720.

[221] ..... « The description of curve lines ». *Philosophical transactions*, *39* : 143–165, 1735–1736.

[222] ..... *Treatise of Algebra.* 1748.

[223] ..... *The collected letters*, ed. Stella Mills. Shiva, Nantwich, Cheshire, 1982.

[224] L. J. MAGNUS. « Über eine Methode den Grad einer durch Elimination hervorgehenden Gleichung zu finden ». *Journal für die reine und angewandte Mathematik, fondé par Crelle, Leipzig*, *26*(4) : 365–367, 1843.

[225] Kenneth L. MANDERS. « Algebra in Roth, Faulhaber, and Descartes ». manuscrit communiqué par l'auteur en 2005, issu d'une intervention au colloque *Transmission and Transformartion of Mathematical Thought*, Wuhan, Chine, octobre 1998.

[226] ..... « Descartes et Faulhaber ». *Archives de Philosophie, Bulletin Cartésien*, *58*(3) : 1–12, 1995.

[227] Marta MENGHINI. « Notes on the correspondance between Luigi Cremona and Max Noether ». *Historia Mathematica*, *13* : 341–351, 1986.

[228] F. MERTENS. « Über die bestimmenden Eigenschaften der Resultante von n Formen mit n Veränderlichen ». *Sitzungsberichte der Mathematisch-Naturwissenschaftlichen Classe der Kaiserlichen Akademie der Wissenschaften, II. Abtheilung*, *93* : 527–566, 1886.

[229] Stella MILLS. « Note on the Braikenridge-Maclaurin theorem ». *Notes Rec. Roy. Soc. Lond.*, *38* : 235–240, 1984.

[230] Ferd. MINDING. « Über die Bestimmung des Grades einer durch Elimination hervorgehende Gleichung ». *Journal für die reine und angewandte Mathematik, fondé par Crelle, Leipzig*, *22*(2) : 178–183, 1841.

[231] ..... « Erwiderung auf den Artikel 23. im 26^{sten} Bande dieses Journals ». *Journal für die reine und angewandte Mathematik, fondé par Crelle, Leipzig*, *27*(4) : 379–380, 1844.

[232] ..... « Entwicklung eines symmetrischen Ausdrucks für den Grad einer durch Elimination hervorgehenden Gleichung ». *Journal für die reine und angewandte Mathematik, fondé par Crelle, Leipzig*, *31*(1) : 1–11, 1846.

[233] Jules MOLK. « Sur une notion qui comprend celle de la divisibilité et sur la théorie générale de l'élimination ». *Acta Mathematica*, *6* : 1–166, 1885.

[234] ..... *Encyclopédie des sciences mathématiques pures et appliquées.* Jacques Gabay, Sceaux, 1992. Réimpression de l'édition originale Gauthiers-Villars et B. G. Teubner (1904–1916), complétée ; 7 tomes.

[235] Eugen NETTO. « Rationale Funktionen mehrerer Veranderlichen ». *Enzykl. der Math. Wiss.*, vol. I B 1 *b* : 255–282, 1894.

[236] ..... *Vorlesungen über algebra.* B. G. Teubner, Leipzig, 1900.

[237] Isaac NEWTON. « The Deposited Lucasian Lectures on Algebra ». *in* [238], vol. V (1683-1684), p. 3–534.

[238] ..... *The Mathematical Paper of Isaac Newton*, éd., D. T. Whiteside. Cambridge University Press, Cambridge, 1972.

[239] Max NOETHER. « Zur Theorie des eindeutigen Entsprechens algebraischer Gebilde von beliebig vielen Dimensionen ». *Mathematische Annalen*, *2* : 293–316, 1870.

[240] ..... « Über einen Satz aus der Theorie der algebraischen Functionen ». *Mathematische Annalen*, *6* : 351–359, 1873.

[241] ..... « Über die singulären Wertsysteme einer algebraischen Function und die singulären Punkte einer algebraischen Curve ». *Mathematische Annalen*, *9* : 166–182, 1875.

[242] ..... « Arthur Cayley ». *Mathematische Annalen*, *46* : 462–480, 1895.

[243] ..... « James Joseph Sylvester ». *Mathematische Annalen*, *50* : 133–156, 1898.

[244] ..... « Über die algebraischen Functionen einer und zweier Variabeln ». *Nachrichten von der Königl. Gesellschaft der Wissenschaften und der Georg-Augusts Universität zu Göttingen*, *9* : 267–278, 7 juin 1871.

[245] Karen Hunger PARSHALL. *James Joseph Sylvester ; life and work in letters*. Clarendon Press, Oxford, 1998.

[246] Erwan PENCHÈVRE. « *Geometry and Dioptrics in Classical Islam*, Roshdi Rashed ». compte-rendu à paraître dans la *Gazette des Mathématiciens*.

[247] ..... « *Le développement de la géométrie aux IX^e–XI^e siècles : Abū Sahl al-Qūhī*, Philippe Abgrall ». compte-rendu à paraître dans *Arabic Sciences and Philosophy*.

[248] ..... « L'œuvre algébrique de Johannes Faulhaber ». *Oriens – Occidens, Cahiers du Centre d'Histoire des Sciences et des Philosophies Arabes et Médiévales*, *5* : 187–222, 2004.

[249] ..... « L'élimination en algèbre aux XVII^e et XVIII^e siècles ». *Historia Scientiarum*, *14*(2) : 101–117, 2004.

[250] M. PIEL. « Eine Begründung der Resultantentheorie durch einen elementaren Beweis der Cayleyschen Konstruktion ». *Mathematische Zeitschrift*, *38* : 650–668, 1934.

[251] Julius PLÜCKER. « Solution d'une question fondamentale concernant la théorie générale des courbes ». *Journal für die reine und angewandte Mathematik, fondé par Crelle, Leipzig*, *12* : 105–108, 1834. (*in* Ges. Math. Abh., Leipzig 1895, vol. 1, p. 298–301).

[252] Siméon-Denis POISSON. « Mémoire sur l'élimination dans les équations algébriques ». *Journal de l'école polytechnique*, *4*(11) : 199, juin 1802.

[253] Jean-Victor PONCELET. *Applications d'analyse et de géométrie, qui ont servi, en 1822, de principal fondement au « Traité des propriétés projectives des figures »*, 2 vol. Mallet-Bachelier puis Gauthier-Villars, Paris, 1862–1864.

[254] ..... *Traité des propriétés projectives des figures*, 2^e éd. Gauthier-Villars, Paris, 1865.

[255] Jean PRESTET. *Nouveaux élémens de mathématiques, ou Principes généraux de toutes les sciences qui ont les grandeurs pour objet. 2^e édition plus ample et mieux digérée*. A. Pralard, Paris, 1689.

[256] V. PUISEUX. « Recherches sur les fonctions algébriques ». *Journal de Mathématiques pures et appliquées, fondé par Liouville, Paris*, *15* : 365–480, 1850.

[257] R. RASHED. *Entre arithmétique et algèbre : Recherches sur l'histoire des mathématiques arabes*. Les Belles Lettres, Paris, 1984.

[258] Roshdi RASHED. article « Al-Karajī » *in* C. C. GILLISPIE, éd., *Dictionary of Scientific Biography*, 16 vol., vol. VII, p. 240–246. Charles Scribner's Sons, New York, 1970.

[259] ..... Edition critique, traduction et commentaire des traités d'algèbre d'Al-Khwārizmī et d'Abū Kāmil, à paraître.

[260] ..... « Science en Islam et modernité classique ». *Al-Furqān Publications* : No. 71. Al-Furqān Islamic Heritage Foundation, Londres, 2002.

[261] ..... « Fermat and Algebraic Geometry ». *Historia Scientiarum*, *11*(1) : 24–47, 2001.

[262] ..... « Al-Qūhī et Al-Sijzī : sur le compas parfait et le tracé continu des sections coniques ». *Arabic Sciences and Philosophy*, *13* : 9–43, 2003.

[263] ..... *Geometry and Dioptrics in Classical Islam*. Al-Furqān Islamic Heritage Foundation, Londres, 2005.

[264] Richelot. « Nota ad theoriam eliminationis pertinens ». *Journal für die reine und angewandte Mathematik, fondé par Crelle, Leipzig*, *21*(3) : 226–234, 1840.

[265] B. Riemann. « Theorie der Abel'schen Functionen ». *Journal für die reine und angewandte Mathematik, fondé par Crelle, Leipzig*, *54* : ?–?, 1857.

[266] Adam Ries. *Coß*, ed. Wolfgang Kaunzner et Hans Wussing. Teubner, Stuttgart, Leipzig, 1992.

[267] Michel Rolle. *Traité d'algèbre ; ou principes généraux pour résoudre les questions de mathématique*. Estienne Michallet, Paris, 1690.

[268] ..... « De l'évanouissement des quantitez inconnues dans la géométrie analytique ». *Mémoires de l'Académie des Sciences*, pages 419–450, 1709.

[269] Georgio Rosenhain. « Exercitationes analiticae in theorema Abelianum de integralibus functionum algebraicarum ». *Journal für die reine und angewandte Mathematik, fondé par Crelle, Leipzig*, *28*(3) : 249–278, 1844.

[270] ..... « Neue Darstellung der Resultante der Elimination von z aus zwei algebraischen Gleichungen $f(z) = 0$ und $\phi(z) = 0$ vermittelst der Werthe welche die Functionen $f(z)$ und $\phi(z)$ für gegebene Werthe von z annehmen ». *Journal für die reine und angewandte Mathematik, fondé par Crelle, Leipzig*, *30*(2) : 157–165, 1846.

[271] Peter Roth. *Arithmetica Philosophica, Oder schöne newe wolgegründte Vberauß Künstliche Rechnung der Coß oder Algebrae / In drey vnterschiedliche Theil getheilt. [...]*. Johann Lanßenberger, Nürnberg, 1608.

[272] George Salmon. *A treatise on the higher plane curves*. Hodges and Smith, Dublin, 1852.

[273] ..... *Lessons Introductory to the Modern Higher Algebra*. Hodges and Smith, Dublin, 1859.

[274] Pierre Samuel. « La notion de multiplicité en algèbre et en géométrie algébrique ». *Journal de Mathématiques pures et appliquées, fondé par Liouville, Paris*, *30* : 159–274, 1951.

[275] ..... *Méthodes d'algèbre abstraite en géométrie algébrique*, seconde édition, corrigée. Springer-Verlag, Berlin-Heidelberg-New York, 1967. (édition originale 1955).

[276] P.-Frédéric Sarrus. *Méthode d'élimination par le plus grand commun diviseur*. Bachelier, Paris, 1834. (15 pages).

[277] Ludwig Schläfli. « Über die Resultante eines Systemes mehrerer algebraischer Gleichungen. Ein Beitrag zur Theorie der Elimination ». *Denkschrift der Kaiserlichen Akademie der Wissenschaften, mathematisch-naturwissenschaftliche Klasse*, 4. Band (1852), Wien. Verzeichnis Graf, Nr. 24. (*in* [278], vol. 2, p. 9–106, suivi d'une postface de Burckardt et de van der Waerden p. 107–112).

[278] ..... *Gesammelte Mathematische Abhandlungen*, 3 vol. Birkhaüser, Basel, 1950–1956.

[279] Carl Schmidt. « Zur Theorie der Elimination ». *Zeitschrift für Mathematik und Physik*, *31* : 214–222, 1886.

[280] Ivo Schneider. *Johannes Faulhaber 1580–1635, Rechenmeister in einer Welt des Umbruchs*. Birkhäuser, Basel, 1993.

[281] Frans Van Schooten. *Exercitationum Mathematicarum libri quinque*. Elzevir, Leyden, 1657.

[282] Charlotte Angas Scott. « On the intersections of plane curves ». *Bulletin of the AMS*, *4* : 260–273, 1897–1898.

[283] ..... « A proof of Noether's fundamental theorem ». *Math. Ann.*, *52* : 593–597, 1899.

[284] Jean-Pierre Serre. « Sur la dimension homologique des anneaux et des modules noethériens ». *in* Symp. Tokyo-Nikko, 1955, p. 175–189.

[285] ..... *Local Algebra*. Springer, 2000.

[286] Michel Serres. « Les anamnèses mathématiques ». *Archives Internationales d'Histoire des Sciences*, *20* : 3–38, 1967.

[287] Joseph-Alfred Serret. *Cours d'algèbre supérieure*. Bachelier, Paris, 1849.

[288] ..... *Cours d'algèbre supérieure*, 3e édition. Bachelier, Paris, 1866.

[289] Francesco Severi. « La géométrie algébrique ». *in* Proceedings of the International Mathematical Congress, University of Toronto, 1924, vol. I, p. 149–154.

[290] Igor R. Shafarevich. *Basic Algebraic Geometry*, 2e édition. Springer-Verlag, 1994–1996. (2 tomes).

[291] Hourya Sinaceur. *Corps et modèles*. Vrin, Paris, 1991.

[292] Henry John Stephen Smith. *The Collected Mathematical Papers*. Chelsea, New York, 1894. (2 tomes).

[293] Simon Stévin. *L'Arithmétique de Simon Stévin de Bruges*. Christophle Plantin, Leyde, 1585.

[294] Ch. F. Sturm. « Démonstration d'un théorème d'algèbre de M. Sylvester ». *Journal de Mathématiques pures et appliquées, fondé par Liouville, Paris*, *7* : 356–368, 1842.

[295] James Joseph Sylvester. « Presidential address to section 'A' of the British Association », Exeter British Association Report (1869) p. 1–9. *in* [308], vol. II, p. 650–661.

[296] ..... « Address on Commemoration Day at Johns Hopkins University », 22 february, 1877 (Baltimore). *in* [308], vol. III, p. 72–87.

[297] ..... « On rational derivation from equations of coexistence, that is to say, a new and extended theory of elimination, Part I ». *The London and Edinburgh philosophical magazine and journal of science*, *15* : 428–435, 1839. (*in* [308], vol. I, p. 40–46).

[298] ..... « A method of determining by mere inspection the derivations from two equations of any degree ». *The London and Edinburgh philosophical magazine and journal of science*, *16* : 132–135, 1840. (*in* [308], vol. I, p. 54–57).

[299] ..... « On derivation of coexistence. Part II. Being the theory of simultaneous simple homogeneous equations ». *Philosophical Magazine*, *16* : 37–40, 1840. (*in* [308], vol. I p. 47–53).

[300] ..... « Exemples of the dialytic method of elimination as applied to ternary systems of equations ». *Cambridge Mathematical Journal*, *2* : 232–236, 1841. (*in* [308], vol. I, p. 61–65).

[301] ..... « On a linear method of eliminating between double, *treble*, and other systems of algebraic equations ». *Philosophical Magazine*, *18* : 425–435, 1841. (*in* [308], vol. I p. 75–85).

[302] ..... « A demonstration of the theorem that every homogeneous quadratic polynomial is reducible by real orthogonal substitutions to the form of positive and negative squares ». *Philosophical Magazine*, *4* : 138–142, 1852. (*in* [308], vol. I p. 378–381).

[303] ..... « Observations on a new theory of multiplicity ». *The London and Edinburgh philosophical magazine and journal of science*, *3* : 460–467, 1852. (*in* [308], vol. I, p. 370–377).

[304] ..... « On a remarkable theorem in the theory of equal roots and multiple points ». *The London and Edinburgh philosophical magazine and journal of science*, *3* : 375–378, 1852. (*in* [308], vol. I, p. 367–369).

[305] ..... « On the principles of the calculus of forms ». *Cambridge and Dublin Mathematical Journal*, *7* : 52–97 et 179–217, 1852. (*in* [308], vol. I, p. 285–327 et 328–363).

[306] ..... « On a theory of the syzygetic relations of two rational integral functions, comprising an application to the theory of Sturm's functions, and that of the greatest algebraical common measure ». *Philosophical Transactions*, *143*(3) : 407–548, 1853. (cf. [308], vol. I (1837-1853), p. 429).

[307] ..... « On the calculus of forms, otherwise the theory of invariants ». *Cambridge and Dublin Mathematical Journal*, *8* : 256–269, 1853. (*in* [308], vol. I, p. 411–422).

[308] ..... *The Collected Mathematical Papers of James Joseph Sylvester*. Chelsea Publishing Co., New York, 1973.

[309] L. Tits. « Sur la sommation des puissances numériques ». *Mathesis*, *37* : 353–355, 1923.

[310] Tschirnhaus. « Methodus generalis omnium aequationum radices exhibendi ». (10 avril 1678). *in* [213], 3ème série, t. 2, p. 369–381.

[311] ..... « Methodus auferendi omnes terminos intermedios ex data aequatione ». *Acta Eruditorum*, pages 204–207, 1683.

[312] H. W. Turnbull. *Bi-centenary of the death of Colin Maclaurin (1698–1746).* University Press, Aberdeen, 1951. lecture by Turnbull, february 4th 1947.

[313] Charles Tweedie. « A study of the life and writings of Colin Maclaurin ». *Mathematical Gazette*, *8* : 133–151, 1915. supplemented by notes *9* : 303–306 (1917) et *10* : 209 (1918).

[314] ..... « The Geometria organica of Colin Maclaurin ». *Proceedings of the Royal Society of Edinburgh*, *36* : 87–150, 1916.

[315] B. Vahabzadeh et R. Rashed. *Al-Khayyām mathématicien.* Albert Blanchard, Paris, 1999.

[316] Bartel Leendert van der Waerden. « Der Multiplizitätsbegriff der algebraischen Geometrie ». *Mathematische Annalen*, *97* : 756–774, 1927.

[317] ..... « The foundation of algebraic geometry from Severi to André Weil ». *Archive for History of Exact Sciences*, *7* : 171–180, 1970/1971.

[318] ..... « On the sources of my book *Moderne Algebra* ». *Historia Mathematica*, *2* : 31–40, 1975.

[319] Vandermonde. « Mémoire sur la résolution des équations ». *Histoire de l'Académie Royale des Sciences*, pages 365–416, 1771.

[320] François Viète. *Francisci Vietae Opera*, éd. Franz Van Schooten. Elzevir, Leyden, 1646.

[321] Jules Vuillemin. *La philosophie de l'algèbre.* Presses universitaires de France, Paris, 1952.

[322] ..... *Mathématiques et métaphysique chez Descartes.* Presses Universitaires de France, Paris, 1960.

[323] Edward Waring. *Meditationes algebraicae.* University of Cambridge, Cambridge, 1770.

[324] ..... *Meditationes algebraicae : an English translation of the work of Edward Waring*, trad. Dennis Weeks. American Mathematical Society, Providence, Rhode Island, 1991. (traduction de la 2de édition, parue en 1782).

[325] Dennis Weeks. « The life and mathematics of George Campbell, F. R. S. ». *Historia Mathematica*, *18* : 328–343, 1991.

[326] Karl Weierstrass. *Mathematische Werke.* Mayer und Müller, Berlin, 1894–1927. (7 volumes).

[327] André Weil. *Foundations of algebraic geometry*, revised edition. American Mathematical Society, 1962 (1ère édition 1946).

[328] H. Weyl. *Algebraic Theory of Numbers.* Princeton U. Press, Princeton, 1940.

[329] Henry S. White. « Bézout's theory of resultants and its influence on geometry ». *Bulletin of the American Mathematical Society*, *15* : 325–338, 1909.

[330] F. Winkler, B. Buchberger, F. Lichtenberger, et H. Rolletschek. « Algorithm 628 : an algorithm for constructing canonical bases of polynomial ideals ». *ACM Transactions on Mathematical Software*, *11*(1) : 66–78, 1985.

[331] Wolfgang Wogel. *Lectures on results on Bezout's Theorem.* Tata Institute of Fundamental Research, Bombay, 1984.

[332] Hans Wussing. *Die Genesis des abstrakten Gruppen Begriffes.* Deutscher Verlag der Wissenschaften, Berlin, 1969.

[333] Oscar Zariski. Compte-rendu du livre [327] de Weil, *Bulletin of the American Mathematical Society*, 54 (1948), p. 671–675.

[334] ..... « Studies in equisingularity I. Equivalent singularities of plane algebroid curves ». *American Journal of Mathematics*, *90*(3) : 507–536, 1968.

[335] ..... *Algebraic surfaces.* Springer-Verlag, Berlin Heidelberg New York, 1971.

Les techniques d'élimination sont aussi vieilles que l'algèbre elle-même, mais dans la seconde moitié du XVII$^{\text{ème}}$ s. apparaissent des méthodes générales, algorithmes d'élimination des inconnues dans des systèmes d'équations polynomiales, chez Fermat, Hudde, Newton, Leibniz, Rolle. Nous comparons ces méthodes, aussi bien au plan historique qu'au plan mathématique, et nous retraçons l'histoire de la théorie de l'élimination, qui ne naît vraiment qu'au XVIII$^{\text{ème}}$ s., lorsque Euler, Cramer, Lagrange, Waring, Bézout et bientôt Poisson s'occupent de démontrer le « théorème de Bézout », de trouver des formules exprimant l'équation finale résultant de l'élimination, de déceler les « facteurs superflus » introduits par certaines méthodes, et d'appliquer cette théorie à la géométrie et aux grands problèmes contemporains en algèbre. Le traité de Bézout (1779) est en avance sur son temps et prépare une troisième période, au cours de laquelle l'élimination acquiert un rôle fondamental, aussi bien en lien avec la géométrie algébrique que dans ses rapports avec la théorie des invariants, par exemple chez Jacobi, Hesse, Cayley, Sylvester, Weierstrass, Halphen, Noether, Kronecker. Au profit de l'étude des structures en algèbre au XX$^{\text{ème}}$ s. l'élimination passe au second plan sans toutefois disparaître. Nous utilisons dans nos analyses les travaux les plus récents sur l'élimination, comme ceux de J.-P. Jouanolou. Enfin, de même que pour l'histoire du concept de multiplicité d'intersection, nous avons cru pouvoir consacrer des chapitres importants à l'œuvre de Faulhaber, à celle de Maclaurin, et au grand mémoire de 1882 de Kronecker.

History of Elimination Theory

Elimination techniques are as old as algebra itself, however in second half of the XVIIth century general methods appear, which are algorithms to eliminate the unknowns in systems of polynomial equations, found by Fermat, Hudde, Newton, Leibniz, Rolle. We compare these methods, from a historical as well as mathematical point of view, and we retrace the history of elimination theory, which truly starts in the XVIIIth century, when Euler, Cramer, Lagrange, Waring, Bézout and soon Poisson deal with demonstrating « Bezout's theorem », finding formulas for the final equation resulting from elimination, discovering « superfluous factors » introduced by some methods, and applying this theory to geometry and major contemporary issues in algebra. Bezout's treatise (1779) is ahead of its time and prepares a third period, along which elimation acquires a fundamental role, in relation to both algebraic geometry and invariant theory, as in Jacobi, Hesse, Cayley, Sylvester, Weierstrass, Halphen, Noether, Kronecker. In XXth century, in favor of the study of algebraic structures, elimination theory is receding although not disappearing. In our analysis, we use the most recent works on elimination, as those of J.-P. Jouanolou. Finally, as we have done for the history of the concept of intersection multiplicity, we have also devoted important chapters to the works of Faulhaber, to the works of Maclaurin, and to the great memoir (1882) of Kronecker.

Mots-clés :

élimination, équation, inconnue, polynôme, Fermat, Hudde, Newton, Leibniz, Rolle, Euler, Cramer, Lagrange, Waring, Bézout, Poisson, Jacobi, Hesse, Cayley, Sylvester, Weierstrass, Halphen, Noether, Kronecker, Jouanolou, Faulhaber, Maclaurin, multiplicité d'intersection, coefficients indéterminés, polygone de Newton, système d'équations, codimension, dimension, facteur superflu, théorie des intersections, théorie des invariants, théorème de Sturm, géométrie projective, algèbre effective, géométrie énumérative, algorithme, formule, résultant, structure, plus grand commun diviseur, Descartes, arithmétisation, algèbre, géométrie, géométrie algébrique, Bernshtein, Kushnirenko, Abu Kamil, Cardan, Zhu Shijie, matrice de Bézout, complexe de Koszul, Macaulay, Abhyankar, Brill, Seki Takakazu, méthode, règle, théorie.