INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

ProQuest Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

UMI®
Topics in Resultants and Implicitization

by

Ming Zhang

A Thesis Submitted
in Partial Fulfillment of the
Requirements for the Degree

Doctor of Philosophy

APPROVED. THESIS COMMITTEE:

Dr. Ronald Goldman, Chair
Professor of Computer Science

Dr. Joe Warren
Professor of Computer Science

Dr. Raymond Wells, JR.
Professor of Mathematics Science

Houston, Texas

April, 2000
Topics in Resultants and Implicitization

Ming Zhang

Abstract

Resultants are computational tools for determining whether or not a system of polynomials has a common root without actually solving for the roots of these equations. Resultants can also be used to solve for the common roots of polynomial systems.

Classical resultants are typically represented as determinants whose entries are polynomials in the coefficients of the original polynomials in the system. The work in this dissertation on classical resultants focuses on bivariate polynomials. It is shown that bivariate resultants can be represented as determinants in a variety of innovative ways and that these various formulations are interrelated. Remarkable internal structures in these resultant matrices are exposed. Based on these structures, efficient computational algorithms for calculating the entries of these resultant matrices are developed.

Sparse resultants are used for solving systems of sparse polynomials, where classical resultants vanish identically and hence fail to give any useful information about the common roots of the sparse polynomials. Nevertheless, sparse polynomial systems frequently appear in surface design. Sparse resultants are usually represented as GCDs of a collection of determinants. These GCDs are extremely awkward for symbolic computation. Here a new way is presented to construct sparse resultants as single determinants for a large collection of sparse systems of bivariate polynomials.

An important application of both classical and sparse resultants in geometric modeling is implicitization. Implicitization is the process of converting surfaces from parametric form into algebraic form. Classical resultant methods fail when a rational
surface has base points. The method of moving quadrics, first introduced by Professor Tom Sederberg at Brigham Young University, is known empirically to successfully implicitize rational surfaces with base points. But till now nobody has ever been able to give a rigorous proof of the validity of this technique. The first proof of the validity of this method when the surfaces have no base points is provided in this dissertation.
Acknowledgments

I am most indebted to my thesis advisor Professor Ronald Goldman. His sharp understanding, broad and deep knowledge in the field inspired me for my thesis research. He tirelessly read many drafts of every part of this thesis: his thoughtful comments and insights guided me all the way through my research. I gratefully thank him for his excellent direction, guidance, and encouragement.

I would also like to sincerely thank Dr. Eng-Wee Chionh for being such a rich source of help. Countless talks with him helped me shape my ideas and improve them. His perseverance and clever approaches to problems always amazed me and set the example for me.

Professor David Cox helped greatly on the validity proof in Chapter 6. I thank him for kindly bringing me to the theory of syzygies.

Professor Joe Warren has urged me to pursue the application aspects of the research in this thesis. He gave me valuable advice and encouragement through many formal and informal conversations. My sincere thanks go to him.

Collin McCurday, Henrik Weimer and other graduate students have helped me a lot during my stay in the department. Dr. Chaoming Zhang, Jingsong He, and Jianliang Qian have been very good mentors and friends for the last four years. I greatly thank all of them for their help.

I thank Professor Joe Warren and Professor Raymond Wells, Jr. for kindly serving on my thesis committee and providing valuable comments and criticism.

I would also like to thank the national science foundation for partially supporting the research of this thesis financially.

My parents, my sister, and my brother-in-law have always been very helpful and
supportive. This thesis is the result of their continuous encouragement and support.
To my wife Yue. I can not thank her enough for her love, sacrifice, and patience. It is to her that this thesis is dedicated.
Contents

Abstract ii
Acknowledgments iv
List of Illustrations xi

1 Introduction 1
1.1 Motivation ........................................... 1
1.2 Background ........................................ 2
1.2.1 Resultants ....................................... 2
1.2.2 Implicitization ................................... 4
1.3 Outline and Main Results .................................. 6

I Resultants 9

2 Univariate Resultants 10
2.1 The Sylvester and Bézout Resultants ....................... 11
2.2 Exact Division by Truncated Formal Power Series ........ 14
2.3 The Transformation Matrix From Sylvester to Bézout .... 15
2.4 Fast Computation of the Entries of the Bézout Resultant .................................................. 16
2.5 Computational Complexity .................................. 20
2.6 Hybrid Resultants ........................................ 21
2.7 Non-Homogeneous Bézout Matrices .......................... 24

3 Bivariate Tensor Product Resultants 29
3.1 The Three Dixon Resultants ................................ 31
3.1.1 The Sylvester Resultant $Syl(f, g, h)$ ........................................ 31
3.1.2 The Cayley Resultant $Cay(f, g, h)$ ........................................ 32
3.1.3 The Mixed Cayley-Sylvester Resultant $Mix(f, g, h)$ ................. 33
3.1.4 Notation .................................................................................... 34
3.1.5 Review and Preview ................................................................. 36

3.2 The Transformation from $Syl(f, g, h)$ to $Mix(f, g, h)$ .................. 36
3.2.1 The Conversion Matrix $G(f, g, h)^*$ ....................................... 37
3.2.2 The Entries of $G(f, g, h)^*$ .................................................... 40
3.2.3 Properties of $G(f, g, h)^*$ ....................................................... 40

3.3 The Transformation from $Syl(f, g, h)$ to $Cay(f, g, h)$ ................. 43
3.3.1 The Conversion Matrix $F(f, g, h)$ .......................................... 43
3.3.2 The Entries of $F(f, g, h)$ ....................................................... 45
3.3.3 Properties of $F(f, g, h)$ .......................................................... 47

3.4 The Transformation from $Mix(f, g, h)$ to $Cay(f, g, h)$ ................. 51
3.4.1 The Conversion Matrix $E(f, g, h)$ .......................................... 51
3.4.2 The Entries of $E(f, g, h)$ ....................................................... 53
3.4.3 Properties of $E(f, g, h)$ .......................................................... 54

3.5 The Block Structure of the Three Dixon Resultants ...................... 55
3.5.1 The Block Structure of $Syl(f, g, h)$ ....................................... 55
3.5.2 The Block Structure of $Mix(f, g, h)$ ....................................... 57
3.5.3 The Block Structure of $Cay(f, g, h)$ ....................................... 60

3.6 Convolution Identities ................................................................. 60
3.6.1 Interleaving in $S_n$ and $M_n$ ............................................... 61
3.6.2 The Convolution Identities ..................................................... 63

3.7 Fast Computation of the Entries of $Cay(f, g, h)$ ......................... 65

3.8 Computational Complexity .......................................................... 67

3.9 Hybrids of the Three Dixon Resultants ........................................ 71
3.9.1 Hybrids of the Sylvester and Cayley Resultants ....................... 72
3.9.2 Hybrids of the Sylvester and Mixed Cayley-Sylvester Resultants 75
3.9.3 Hybrids of the Mixed Cayley-Sylvester and Cayley Resultants 80

4 Sparse Resultants 84
4.1 Construction of Sylvester $A$-resultants 86
  4.1.1 Bi-degree Sylvester Resultant Matrices 86
  4.1.2 Rectangular Corner Cut Sylvester $A$-resultant Matrices 87
4.2 Sylvester $A$-resultants 90
  4.2.1 Only the Upper Right Corner is Cut Off 91
  4.2.2 Cut Off One More Corner: The Upper Left Corner 95
  4.2.3 Cut Off One More Corner: The Lower Right Corner 96
  4.2.4 Cut Off All Four Corners 96
4.3 Remarks on Sylvester $A$-resultants 99
4.4 Construction of Dixon $A$-resultants 100
  4.4.1 Bi-degree Dixon Resultant Matrices 102
  4.4.2 Corner Cut Dixon $A$-resultant Matrices 103
4.5 Dixon $A$-resultants 106
4.6 Comparison of the Sylvester and Dixon $A$-resultants 107
4.7 Implicitization by Sylvester $A$-resultants 107
4.8 Implicitization by Dixon $A$-resultants 110
  4.8.1 Hirzebruch Surfaces 110
  4.8.2 Diamond Cyclides 111
  4.8.3 One-Horned Cyclides 112
  4.8.4 Hexagons 113
  4.8.5 Pentagons 116

II Implicitization 121
5 Implicitization Using Moving Curves

5.1 The Method of Moving Curves ........................................ 124
5.2 Implicitizing Curves Using Moving Lines .......................... 128
   5.2.1 Even Degree Rational Curves ................................. 128
   5.2.2 Odd Degree Rational Curves ................................. 132
   5.2.3 Anti-Annihilation by $\mu$-Basis ............................. 133
5.3 Implicitizing Curves Using Moving Conics ........................ 135
   5.3.1 Moving Line and Moving Conic Coefficient Matrices .......... 136
   5.3.2 $|ML|$ Factors $|MCw|$ ..................................... 138
   5.3.3 Moving Conics and Curves of Odd Degrees .................. 143

6 Implicitization Using Moving Surfaces ............................ 145

6.1 The Method of Moving Surfaces .................................... 147
6.2 Implicitizing Surfaces Using Moving Planes ...................... 151
6.3 Resultants and Syzygies .......................................... 157
   6.3.1 Triangular Polynomials .................................... 158
   6.3.2 Tensor Product Polynomials ................................ 159
6.4 Implicitizing Tensor Product Surfaces Using Moving Quadrics ... 161
   6.4.1 Moving Plane and Moving Quadric Coefficient Matrices .... 161
   6.4.2 The Validity of Moving Quadrics for Tensor Product Surfaces 162
6.5 Implicitizing Triangular Surfaces Using Moving Quadrics ........ 167
   6.5.1 Moving Plane and Moving Quadric Coefficient Matrices .... 167
   6.5.2 The Method of Moving Quadrics for Triangular Surfaces ... 169
   6.5.3 The Validity of Moving Quadrics for Triangular Surfaces ... 170

III Open Questions .................................................... 176

7 Open Questions .................................................... 177
Illustrations

4.1 The monomial support $A_{m,n}$ and the multiplying set $C_{m,n}$. 87
4.2 Cut rectangular corners off $A_{m,n}$ and $C_{m,n}$. 88
4.3 The Minkowski sum $A + C_A$. 89
4.4 An example of $A$ and $C_A$. 89
4.5 The sets $A, C_A, A + C_A$ when one corner is cut off. 92
4.6 The sets $A, C_A, A + C_A$ when two corners are cut off. 95
4.7 The sets $A, C_A, A + C_A$ when three corners are cut off. 97
4.8 The sets $A, C_A, A + C_A$ when all four corners are cut off. 98
4.9 Another choice for the multiplying set. 101
4.10 Two multiplying sets for the Hirzebruch surface. 101
4.11 The monomial support $A$, row support $R_A$, and column support $C_A$. 105
4.12 A Hirzebruch surface. 119
4.13 A diamond cyclide. 119
4.14 A one-horn cyclide. 119
4.15 A hexagon surface. 119
4.16 A pentagon surface. 120

7.1 Staircase corner cut off from $A_{m,n}$. 180
Chapter 1

Introduction

1.1 Motivation

For computational problems, people usually build models using polynomial systems. For example, many computer aided design (CAD) and computer aided manufacturing (CAM) systems use Bézier-Bernstein splines to model mechanical components and assemblies. Even if a model is not a polynomial system, often it may be reduced to a polynomial system or approximated with a polynomial system. For instance, when a system involves the transcendental functions sine or cosine, we may not be able to solve the system directly. Instead we try to replace these trigonometry functions with rational functions or approximate them with polynomials, say with finite Taylor expansions. The reason for these reductions and approximations is that a lot is known about working with polynomial systems.

Research on polynomial systems has a wide range of applications in such diverse areas as algebraic geometry, automated geometric theorem proving, CAD/CAM, computer aided drug design, computer aided geometric design, computer graphics, computer vision, computer algebra, solid modeling, robotics, and virtual reality. For example, in robotics, when a robot moves, it needs to detect whether it will collide with an obstacle. Both the robot and the obstacles may be represented as polynomials or polynomial systems. Collision detection is then reduced to solving a polynomial system. Also in CAD/CAM systems, which are very popular among industrial scientists and engineers, one of the core packages is the polynomial system. Many products manufactured on assembly lines such as automobiles, ships, and airplanes are designed using Bézier, B-spline, or NURBS (non-uniform rational B-splines) techniques.
that is, by polynomial or piecewise polynomial systems. Or to consider another example: pharmacologists now design new drugs by screening a large database of small molecules. The small molecules are checked to see whether they can fit into a docking position in a large protein so that the small molecule and the large protein may interact with each other to achieve the desired pharmaceutical effect. The small molecules and the docking places can all be represented as polynomial systems. Thus the process of choosing the right molecule for new drugs is reduced to analyzing and solving polynomial systems.

To analyze and solve various polynomial systems, mathematicians have developed many effective tools. Resultants are one of the most powerful of these computational techniques [2] [6] [14] [21] [25] [29] [31] [32] [39] [44] [50] [56].

1.2 Background

1.2.1 Resultants

For a system of $n$ polynomials in $n - 1$ variables, the resultant is a polynomial in the coefficients of the original $n$ polynomials. The vanishing of the resultant is a necessary and sufficient condition for the $n$ polynomials in the system to have a common root.

Resultants are used to determine whether or not a system of polynomials has a common root without explicitly solving for the roots of these equations. Sometimes actually solving for the common roots of a polynomial system might be very expensive or even impossible when the number of common roots is infinite. When the number of roots is finite and we are really interested in the numerical values of the common roots, resultants can also be used to solve for common roots of polynomial systems.

The history of classical resultants can be traced back two hundred years [2]. Since people had to do all the computations by hand, resultants were eclipsed in the first half of the twentieth century by more theoretical, less constructive techniques. Research on classical resultants revived in the second half of the twentieth century when computers became a great help to human computation and CAD/CAM developed
into such a powerful tool for industrial scientists and engineers [32] [36] [50]. With greatly increased computational ability, people again picked up the constructive tool of resultants for various modeling, design, and analysis tasks.

Classical resultants are typically represented as determinants whose entries are polynomials in the coefficients of the original polynomials in the system. For two univariate polynomials, there are two well-known determinant representations: Sylvester resultants and Bézout resultants. The Sylvester resultant of two degree \( n \) polynomials is a \( 2n \times 2n \) determinant constructed using Sylvester's dialytic method. This determinant has lots of zero entries, and the non-zero entries are simply the coefficients of the two original polynomials. The Bézout resultant, generated from Cayley's expression, is an \( n \times n \) determinant whose entries are more complicated polynomials in the coefficients of the two original polynomials [15] [32] [46] [58] — see too Chapter 2.

For three bi-degree \((m, n)\) polynomials, Dixon describes three distinct determinant formulations for the resultant [25]. The first formulation is constructed from the dialytic method and the size of the determinant is \( 6mn \times 6mn \). The second formulation is the determinant of a \( 2mn \times 2mn \) matrix found by using an extension of Cayley's determinant device for generating the Bézout resultant for two univariate polynomials. The third representation is the determinant of a \( 3mn \times 3mn \) matrix generated by combining Cayley's determinant device with Sylvester's dialytic method. The first formulation has large determinant dimension, but simple entries. The second and the third formulations have smaller determinant dimensions but much more complicated entries. The differences between these three formulations are analogous to the differences between the Sylvester resultant and the Bézout resultant in the univariate setting [15] [61] — see too Chapter 3.

For three bivariate polynomials of total degree \( n \), the most popular resultant is the Macaulay quotient [39]. Using the dialytic method, Macaulay generates a determinant of dimension \( 4n^2 - n \). The resultant is then expressed as the quotient of this determinant and a sub-determinant [6] [21] [40]. A more compact formulation for
the resultant of three bivariate polynomials of total degree $n$ can be given in terms of a single determinant of dimension $2n^2 - n$ [21] [31].

In general, multipolynomial resultants for systems of total degree polynomials are constructed as Macaulay quotients, or expressed as the greatest common divisors (GCDs) of a collection of determinants [21] [31].

Sparse resultants, developed using more advanced geometric tools, emerged in 1970's and are used for solving systems of sparse polynomials, i.e. polynomials with lots of zero coefficients. Sparse resultants extend the theory of the classical resultants to detect nontrivial common roots in a system of sparse polynomial equations [21] [29] [31] [56]. Classical resultants vanish identically for sparse polynomials and hence fail to give any useful information about the common roots of these sparse polynomials. Nevertheless, sparse polynomial systems frequently appear in surface design, for example in multi-sided Bézier patches [59]. One standard approach to construct sparse resultants is to introduce a set of monomials — a multiplying set — to multiply the polynomials in the original system. If the multiplying set is carefully chosen, the determinant of the coefficient matrix of the resulting polynomials after multiplication may be a nonzero multiple of the resultant of the original polynomials. To eliminate the extraneous factors and to extract the sparse resultant from this determinant, one takes the GCD of a collection of such determinants. Research on sparse resultants has been very active in the last decade [5] [8] [17] [21] [27] [29] [31] [57] [60] [65].

1.2.2 Implicitization

In geometric modeling and computer aided design systems, curves and surfaces can be represented either parametrically or implicitly. For example, the unit sphere written as $x^2 + y^2 + z^2 - 1 = 0$ is in implicit form, and written as

$$
x = \frac{2s}{1 + s^2 + t^2}, \quad y = \frac{2t}{1 + s^2 + t^2}, \quad z = \frac{1 - s^2 - t^2}{1 + s^2 + t^2}
$$

is in parametric form. Implicit representations of curves and surfaces provide a straightforward way to detect whether or not a point lies on a curve or surface. Para-
metric representations of curves and surfaces are very convenient for shape control and for rendering. Both the parametric form and the implicit form are important, but many curve and surface design systems start from parametric representations. Implicitization is the process of converting curves and surfaces from parametric form into implicit form.

Free-form curves and surfaces are often represented parametrically as ratios of polynomials. Curves and surfaces represented by ratios of polynomials are called rational curves or rational surfaces. Both classical resultants and sparse resultants can be used to implicitize rational curves and surfaces [6] [17] [21] [32] [50] [51] – see too Chapters 4, 5, 6.

The implicitization problem for rational curves has been completely and successfully solved using resultants. However, the surface implicitization problem is still far from a complete and satisfactory solution. Classical resultants fail when a rational surface has base points, which is one of the reasons that sparse resultants were introduced. A rational surface is said to have a base point if its numerators and denominators vanish simultaneously at some parameter values.

Professor Tom Sederberg at Brigham Young University first introduced the method of moving curves and moving surface to solve the implicitization problem for rational curves and surfaces. For rational curves, the method of moving lines is equivalent to the resultant method. But the method of moving conics generates the implicit equation of a rational curve in a much smaller determinantal expression than the resultant method [23] [53] [54] [61]. Similarly, for rational surfaces, the method of moving planes is equivalent to the method of bivariate resultants. Moreover, the method of moving quadrics seems to successfully implicitize rational surfaces with base points using a much smaller matrix than the resultant matrix. Many examples have shown the range and power of this new method [52].
1.3 Outline and Main Results

This thesis investigates resultants and implicitization, with an emphasis on bivariate resultants and the implicitization problem of rational surfaces.

The work in this thesis on classical resultants focuses on bivariate polynomials. Bivariate polynomials are used widely in surface design and bivariate resultants are one of the most powerful tools for analyzing such surfaces. A variety of innovative ways are introduced for representing bivariate resultants as single determinants and inter-relationships between these various formulations are investigated. Remarkable internal structures are found in these resultant matrices by grouping the entries of these matrices into specific blocks, making it much easier to understand and manipulate these rather large matrices. Based on these block structures, efficient computational algorithms for calculating the entries of certain resultant matrices are developed.

Sparse resultants are usually represented as GCDs of a collection of determinants. These GCDs are extremely awkward for symbolic computation, since these determinants usually expand to hundreds or even thousands of terms. A new way is discovered here to construct sparse resultants as single determinants for a large collection of interesting sparse systems of bivariate polynomials.

The method of moving quadrics has been shown to be very successful in implicitizing rational surfaces, but till now nobody has ever been able to give a rigorous proof of the validity of this technique. The first proof of the validity of this method when the surfaces have no base points is provided in this thesis. It is shown that the method works in almost all cases and it is determined as well exactly when the method fails.

This thesis is organized as follows:

Chapter 2 discusses univariate resultants. A variety of mathematical connections between the Sylvester resultant matrices and the Bézout resultant matrices are explored, and a simple, block structured matrix that transforms the Sylvester matrix into the Bézout matrix is derived. This matrix transformation captures the essence
of the mathematical relationships between these two resultant matrices. From this transformation, new and efficient algorithms for computing the entries of the Bézout resultant matrices are developed, and hybrid resultants of the Sylvester and Bézout matrices are constructed.

Chapter 3 explores bi-degree resultants. Transformations between the three Dixon resultant representations are investigated. Block structures are imposed on the entries of the transformation matrices as well as on the three Dixon resultant matrices. Polynomials are associated to the blocks so that these blocks are related with a set of convolution identities. Based on these block structures, fast algorithms to compute the resultant matrix entries are developed, along with hybrid resultants of the three Dixon resultant matrices.

Chapter 4 shows how particular multiplying sets can help to construct resultant matrices for a large collection of sparse bivariate polynomials. The single determinants of the coefficient matrices of the polynomials after multiplication are proven to give the sparse resultants. The technique developed by EngWee Chionh to extract sparse resultant matrices from the Dixon resultant matrices is also discussed. As an application, the sparse resultant method is used to implicitize the surfaces provided by Zube in [65].

Chapter 5 investigates the implicitization problem for rational curves by applying the method of moving curves. It is shown in this chapter that the implicit equation of a rational curve can be obtained by taking the Sylvester resultant of two specific moving lines: this construction generates a novel implicitizing matrix in the style of the Sylvester resultant and the size of the Bézout resultant. Some new perspective on the method of moving conics is also provided. By deriving a factorization relation between the moving line coefficient matrix and a submatrix of the moving conic coefficient matrix, a new proof is provided for the theorem that the method of moving conics almost always successfully implicitizes rational curves. This derivation is generalized to the method of moving quadrics for implicitizing rational surfaces in the
next chapter.

Chapter 6 is devoted to the method of moving surfaces for implicitizing rational surfaces. It is shown in this chapter that generically the implicit equation of a rational bi-degree \((m, n)\) surface can be obtained from \(2m\) specific moving planes: this implicitization matrix has compact size, but also has the Sylvester style, mimicking the moving lines method for implicitizing rational curves. The first proof that the method of moving quadrics almost always successfully produces the implicit equation for rational surfaces is also provided.

Chapter 7 concludes this thesis with some interesting unresolved questions concerning classical resultants, sparse resultants, and the method of moving quadrics.
Part I

Resultants
Chapter 2

Univariate Resultants

Two univariate polynomials have a common root if and only if their resultant is zero. For two univariate polynomials of the same degree, there are two standard determinant representations for the resultant: the Sylvester resultant [58] and the Bézout resultant [32] [46]. Typically these two resultants are constructed in isolation, without any regard for how one matrix is related to the other. In this chapter, we will explore a variety of mathematical connections between the resultant matrices of Sylvester and Bézout.

We shall derive a simple, block symmetric, block upper triangular matrix that transforms the Sylvester matrix into the Bézout matrix. This matrix transformation captures the essence of the mathematical relationships between these two resultant matrices. We shall then apply this transformation matrix to:

- derive an explicit formula for each entry of the Bézout matrix;
- develop an efficient recursive algorithm for computing all the entries of the Bézout matrix;
- construct a sequence of hybrid resultant matrices that provide a natural transition from the Sylvester matrix to the Bézout matrix, each hybrid consisting of some columns from the Sylvester matrix and some columns from the Bézout matrix;
- extend the Bézout construction to two univariate polynomials of different degrees.
Except for the recursive algorithm, all of these results are already known [21] [32] [54]. What is new here in addition to the recursive algorithm is our approach: deriving all these properties in a unified fashion based on the structure of the transformation matrix. Equally important, these seemingly straightforward features of the resultants for two univariate polynomials of degree \( n \) anticipate similar, but much more intricate, properties and interrelationships between the resultants for three bivariate polynomials of bidegree \((m, n)\), which we will discuss in the next chapter.

This chapter is organized in the following fashion. In Section 2.1 we briefly review the standard constructions for the Sylvester and Bézout resultants. In Section 2.2 we introduce the method of truncated formal power series, and in Section 2.3 we apply this method to generate the matrix transformation from the Sylvester matrix to the Bézout matrix. Section 2.4 is devoted to developing a new, efficient, recursive algorithm for computing all the entries of the Bézout resultant. Section 2.5 discusses the complexity of this new algorithm as well as the complexity of the standard method for computing the entries of the Bézout matrix. Hybrid resultants are discussed in Section 2.6. We close this chapter in Section 2.7 with a derivation of the Bézout resultant for two univariate polynomials of different degree.

### 2.1 The Sylvester and Bézout Resultants

Let

\[
    f(t) = \sum_{i=0}^{n} a_i t^i, \quad g(t) = \sum_{j=0}^{n} b_j t^j
\]

be two polynomials of degree \( n \), and let \( L(t) = [f(t) \ g(t)] \). The determinant of the \( 2n \times 2n \) coefficient matrix of the \( 2n \) polynomials \( t^r L \), \( 0 \leq r \leq n - 1 \), is known as the Sylvester resultant of \( f \) and \( g \). To fix the order of the rows and columns of the Sylvester matrix \( Syl(f, g) \), we define

\[
    \begin{bmatrix}
        L & tL & \cdots & t^{n-1} L
    \end{bmatrix} =
    \begin{bmatrix}
        1 & \cdots & t^{2n-1}
    \end{bmatrix} Syl(f, g).
\]  

(2.1)
Let \( L_i = [a_i, b_i] \); then
\[
Syl(f, g) = \begin{bmatrix}
L_0 \\
: \\
L_{n-1} \\
L_n \\
L_{n-1} \\
L_n \\
: \\
L_0
\end{bmatrix}
= \begin{bmatrix}
a_0 & b_0 \\
: & : \\
a_{n-1} & b_{n-1} \\
a_n & b_n \\
a_{n-1} & b_{n-1} \\
a_n & b_n \\
: & : \end{bmatrix}.
\]

Note that we have adopted a somewhat nonstandard form for the Sylvester matrix by grouping together the polynomials \( f \) and \( g \) (compare to [58]). This grouping will simplify our work later on and will change at most the sign of the Sylvester determinant. Observe too that with this ordering of the columns the top (bottom) half of the Sylvester matrix is striped, block lower (upper) triangular, and block symmetric with respect to the northeast-southwest diagonal.

To obtain the Bézout matrix for \( f \) and \( g \), we consider the Cayley expression
\[
\Delta_n(t, \lambda) = \begin{vmatrix}
f(t) & g(t) \\
f(\lambda) & g(\lambda)
\end{vmatrix} / (\lambda - t).
\]
(2.2)

Notice that the numerator of \( \Delta_n(t, \lambda) \) is zero when \( \lambda = t \). Since the denominator exactly divides the numerator in \( \Delta_n(t, \lambda) \), this rational Cayley expression is really a degree \( n - 1 \) polynomial in \( t \) and in \( \lambda \), so we can write
\[
\Delta_n(t, \lambda) = \sum_{v=0}^{n-1} D_v(t) \lambda^v
\]
(2.3)
where \( D_v(t), 0 \leq v \leq n - 1 \), are polynomials of degree \( n - 1 \) in \( t \). The determinant of the \( n \times n \) coefficient matrix of the \( n \) polynomials \( D_0(t), \ldots, D_{n-1}(t) \) is known as the Bézout resultant of \( f \) and \( g \).

To fix the order of the rows and columns of the Bézout matrix \( Bez(f, g) \), we define
\[
\sum_{v=0}^{n-1} D_v(t) t^v = \begin{bmatrix} 1 & \cdots & t^{n-1} \end{bmatrix} Bez(f, g) \begin{bmatrix} 1 \\
\vdots \\
\lambda^{n-1}
\end{bmatrix}.
\]
(2.4)
Thus the rows of $Bez(f, g)$ are indexed by $1, t, \ldots, t^{n-1}$ and the columns by $1, \beta, \ldots, \beta^{n-1}$.

It is well-known that the Bézout matrix $Bez(f, g)$ is symmetric [32]. Here we provide an elementary, high level, proof of this fact without computing entry formulas for $Bez(f, g)$. Since

$$
\Delta_n(t, \beta) = \begin{vmatrix} f(t) & g(t) \\ f(\beta) & g(\beta) \end{vmatrix} = \begin{vmatrix} f(\beta) & g(\beta) \\ f(t) & g(t) \end{vmatrix} = \Delta_n(\beta, t) \quad (2.5)
$$

we have

$$
\begin{bmatrix} 1 & \cdots & \beta^{n-1} \end{bmatrix} Bez(f, g)^T \begin{bmatrix} 1 \\ \vdots \\ \beta^{n-1} \end{bmatrix}
$$

$$
= \begin{bmatrix} 1 & \cdots & \beta^{n-1} \end{bmatrix} Bez(f, g)^T \begin{bmatrix} 1 \\ \vdots \\ \beta^{n-1} \end{bmatrix}^T
$$

$$
= \begin{bmatrix} 1 & \cdots & \beta^{n-1} \end{bmatrix} Bez(f, g) \begin{bmatrix} 1 \\ \vdots \\ \beta^{n-1} \end{bmatrix}
$$

$$
= \Delta_n(t, \beta)
$$

$$
= \Delta_n(\beta, t) \quad \text{[by Equation (2.5)]}
$$

$$
= \Delta_n(t, \beta) \begin{bmatrix} 1 \\ \vdots \\ \beta^{n-1} \end{bmatrix}
$$

where $Bez(f, g)^T$ is the transpose of $Bez(f, g)$. Hence $Bez(f, g)^T = Bez(f, g)$, so $Bez(f, g)$ is symmetric.

The columns of the Sylvester matrix $Syl(f, g)$ represent polynomials of degree $2n - 1$, whereas the columns of the Bézout matrix $Bez(f, g)$ represent polynomials of degree $n - 1$. Since, in general, $|Syl(f, g)| \neq 0$, the columns of $Syl(f, g)$ are linearly independent. Hence the $2n$ polynomials represented by these columns span the space
of polynomials of degree $2n - 1$. In particular, we can express the polynomials represented by the columns of the Bézout matrix as linear combinations of the polynomials represented by the columns of the Sylvester matrix. Thus there must be a matrix that transforms the Sylvester matrix into the Bézout matrix. In the next section we recall a mathematical technique that will help us to find explicit formulas for the entries of this transformation matrix.

2.2 Exact Division by Truncated Formal Power Series

Let $\mu(x, y) = \sum_{i=0}^{n} a_i(y)x^i$ be a polynomial in $x$ and $y$. If the rational expression $\mu(x, y)/(x - y)$ is actually a polynomial — that is, if $x - y$ exactly divides $\mu(x, y)$ — we can convert the division to multiplication by replacing each monomial $x^i$ in the numerator by the sum $\sum_{k=0}^{i-1} y^{i-1-k}x^k$. That is,

$$\sum_{i=0}^{n} a_i(y)x^i/(x - y) = \sum_{i=0}^{n} a_i(y) \sum_{k=0}^{i-1} y^{i-1-k}x^k. \quad (2.6)$$

where the vacuous sum $\sum_{k=0}^{i-1}$ is taken to be zero. This identity is motivated by the observation that formally we have

$$\frac{1}{x - y} = \sum_{k=0}^{\infty} \frac{y^k}{x^{k+1}}.$$

Thus

$$x^i \sum_{k=0}^{\infty} \frac{y^k}{x^{k+1}} = x^i \sum_{k=0}^{i-1} \frac{y^k}{x^{k+1}} + x^i \sum_{k=i}^{\infty} \frac{y^k}{x^{k+1}}$$

$$= \sum_{k=0}^{i-1} y^{i-1-k}x^k + \text{terms involving negative powers of } x.$$

Since, by assumption, the quotient on the left hand side of Equation (2.6) is a polynomial, terms involving negative powers of $x$ must cancel; therefore, these terms can be ignored.

To illustrate this division by truncated formal power series technique, consider the example:

$$\mu(x, y)/(x - y) = \frac{4x + 7x^2 - 4y - 7y^2}{x - y}.$$
Using the method, we replace the monomial $x$ by the sum \( \sum_{k=0}^{1-1} y^{1-1-k} x^k \equiv 1 \) and the monomial $x^2$ by the sum \( \sum_{k=0}^{2-1} y^{2-1-k} x^k \equiv y + x \). The quotient is then
\[
\frac{4x + 7x^2 - 4y - 7y^2}{x - y} = 4 \times 1 + 7 \times (y + x)
\]
as expected.

### 2.3 The Transformation Matrix From Sylvester to Bézout

If we perform the division in the Cayley expression (2.2) using the method of truncated formal power series and delay expanding $f(t)$ and $g(t)$ as sums, we obtain a relationship between the Sylvester matrix $Syl(f, g)$ and the Bézout matrix $Bez(f, g)$. Recall that $L = [f(t) \quad g(t)]$ and let $R_j = [b_j \quad -a_j]^T$. Then
\[
\Delta_n(t, \beta) = \begin{vmatrix} f(t) & g(t) \\ f(\beta) & g(\beta) \end{vmatrix} / (\beta - t) \\
= (f(t) \sum_{j=0}^{n} b_j \beta^j - g(t) \sum_{j=0}^{n} a_j \beta^j) / (\beta - t) \\
= \sum_{j=0}^{n} LR_j, \beta^j / (\beta - t) \\
= \sum_{j=0}^{n} LR_j \sum_{\nu=0}^{j-1} t^{j-1-\nu} \beta^\nu \\
= \sum_{\nu=0}^{n-1} \sum_{j=\nu+1}^{n} LR_{j, \beta^{j-1-\nu} \beta^\nu}.
\]
(2.7)

Hence there is a $2n \times n$ matrix $T(f, g)$ such that
\[
\Delta_n(t, \beta) = \begin{bmatrix} L & tL & \cdots & t^{n-1}L \end{bmatrix} T(f, g) \begin{bmatrix} 1 \\ \vdots \\ t^{2n-1} \end{bmatrix}
= \begin{bmatrix} 1 & \cdots & t^{2n-1} \end{bmatrix} Syl(f, g) \cdot T(f, g) \begin{bmatrix} 1 \\ \vdots \\ t^{n-1} \end{bmatrix}
\]
(2.8)
Comparing Equation (2.8) with Equations (2.3) and (2.4), we see that

\[ Syl(f, g) \cdot T(f, g) = \begin{bmatrix} Brz(f, g) \\ 0_{n \times n} \end{bmatrix}. \]

Note that \( n \) rows of zeros are appended below \( Brz(f, g) \) because, unlike \( Syl(f, g) \), \( Brz(f, g) \) does not involve the monomials \( t^\tau, n \leq \tau \leq 2n - 1 \).

To find the entry of \( T(f, g) \) indexed by \( (t^r, L, \beta^r) \), let \( \tau = j - 1 - v \). Then \( j = \tau + v + 1 \). Hence by Equation (2.7) the entry of \( T(f, g) \) indexed by \( (t^r, L, \beta^r) \) is simply

\[ R_{\tau+v+1}, \quad (2.9) \]

for \( \tau + v + 1 \leq n \) and zero otherwise. From this entry formula, we easily see that:

\[ T(f, g) = \begin{bmatrix} R_1 & R_2 & \cdots & R_{n-1} & R_n \\ R_2 & R_3 & \cdots & R_n \\ \vdots & \vdots & \ddots & \vdots \\ R_{n-1} & R_n \\ R_n \end{bmatrix}, \quad (2.10) \]

since the entries are constant along the diagonals \( \tau + v = constant \) and zero below the diagonal \( \tau + v = n - 1 \). Though the dimension of \( T(f, g) \) is \( 2n \times n \), \( T(f, g) \) can be treated as a square matrix of dimension \( n \times n \) if its entries are viewed as column vectors \( R_j \). From this perspective, we see that \( T(f, g) \) is striped, symmetric, and upper triangular.

### 2.4 Fast Computation of the Entries of the Bézout Resultant

An explicit entry formula for the Bézout matrix is provided in [32]. In this section, we derive this formula from the transformation matrix of Section 2.3. We also present a new, more efficient method for computing the entries of Bézout matrices and sums of Bézout matrices.
From Section 2.1 and Section 2.3, we know that

$$Syl(f, g) = \begin{bmatrix} L_0 & \cdots & \cdots & L_0 \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ L_{n-1} & \cdots & \cdots & L_n \end{bmatrix}, \quad T(f, g) = \begin{bmatrix} R_1 & \cdots & \cdots & R_n \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ R_n & \cdots & \cdots & R_n \end{bmatrix}.$$ 

and

$$Syl(f, g) \cdot T(f, g) = \begin{bmatrix} B_{e_2}(f, g) \\ \vdots \\ \vdots \\ 0_{n \times n} \end{bmatrix}.$$ 

Hence

$$B_{e_2}(f, g) = \begin{bmatrix} L_0 & \cdots & \cdots & L_0 \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ L_{n-1} & \cdots & \cdots & L_n \end{bmatrix} \begin{bmatrix} R_1 & \cdots & \cdots & R_n \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ R_n & \cdots & \cdots & R_n \end{bmatrix}.$$ \hspace{1cm} (2.11)

Write

$$B_{e_2}(f, g) = \begin{bmatrix} B_{0,0} & \cdots & B_{0,n-1} \\ \vdots & \ddots & \vdots \\ B_{n-1,0} & \cdots & B_{n-1,n-1} \end{bmatrix}.$$ \hspace{1cm} (2.12)

Then by Equation (2.11) and Equation (2.12),

$$B_{i,j} = \sum_{l=0}^{\min(t, n-1-j)} L_{t-l} \cdot R_{j+l+1}, \quad 0 \leq i, j \leq n - 1.$$ \hspace{1cm} (2.13)

Since $L_t \cdot R_j + L_j \cdot R_t = 0$, it is easy to verify that

$$\sum_{l=0}^{\max(0, t-1)} L_{t-l} \cdot R_{j+l+1} = 0.$$ 

Therefore, we can rewrite Equation (2.13) as

$$B_{i,j} = \sum_{l=\max(0, t-j)}^{\min(t, n-1-j)} L_{t-l} \cdot R_{j+l+1}, \quad 0 \leq i, j \leq n - 1.$$ \hspace{1cm} (2.14)
Equation (2.14) is equivalent to the explicit entry formula for the Bézout matrix given in [32]. By comparing the entry formulas in Equation (2.14), it follows that $B_{i,j} = B_{j,i}$. Thus we see again that $Bez(f, g)$ is a symmetric matrix.

From Equation (2.14), it is easy to recognize recursion along the diagonals $i+j = k$ for $0 \leq k \leq 2n - 2$. In fact,

$$B_{i,j} = B_{i-1,j+1} + L_i \cdot R_{j+1}. \quad (2.15)$$

Equation (2.15) can be applied to compute the entries of $Bez(f, g)$ very efficiently using the following three-step algorithm:

(i) Initialization:

$$Bez(f, g)_{init} = \begin{bmatrix}
L_0 \cdot R_1 & \cdots & L_0 \cdot R_n \\
\vdots & & \vdots \\
L_{n-1} \cdot R_{n-1} & & L_n \cdot R_n
\end{bmatrix}.$$

That is,

$$(B_{i,j})_{init} = L_i \cdot R_{j+1} = a_i b_{j+1} - a_{j+1} b_i, \quad 0 \leq i \leq j \leq n - 1. $$

(ii) Recursion:

for $i = 1$ to $n - 2$

for $j = n - 2$ to $i$ (step -1)

$$B_{i,j} \leftarrow B_{i,j} + B_{i-1,j+1}.$$

That is, we add along the diagonals marching from upper right to lower left, so schematically.
\[
\begin{bmatrix}
B_{0,0} & \cdots & B_{0,n-1} \\
\vdots & & \vdots \\
B_{n-1,n-1}
\end{bmatrix}
\]

\[
= \begin{bmatrix}
L_0 \cdot R_1 & L_0 \cdot R_2 & L_0 \cdot R_3 & \cdots & L_0 \cdot R_{n-3} & L_0 \cdot R_{n-2} & L_0 \cdot R_{n-1} \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\
L_{n-3} \cdot R_{n-1} & L_{n-2} \cdot R_{n-1} & L_{n-1} \cdot R_{n-1}
\end{bmatrix}
\]

(iii) Symmetry: \( B_{i,j} = B_{j,i}, \quad i > j \).

That is, the entries below the diagonal \( i = j \) are obtained via symmetry from the entries above the diagonal.

This method can also be applied to compute sums of Bézout matrices of the same size, say \( n \times n \), efficiently. Instead of computing each Bézout matrix separately, and then adding them together, we can save time and space by adopting the following strategy of initialization and marching:

- initialize the entries in this \( n \times n \) matrix with indices \( 0 \leq i \leq j \leq n - 1 \) using the sums of the initializations for each Bézout matrix—the previous initialize-and-march method assigns an initialization for each Bézout matrix in the sum:

- march to the southwest adding along the diagonals \( 1 \leq i + j \leq 2n - 3 \).

- generate the entries with indices \( i > j \) via symmetry.

In this way, we save not only the space for storing the individual Bézout matrices, but also we march along the diagonals just once: hence this approach greatly reduces the computational complexity of the algorithm. This fast computation of
sums of Bézout matrices has applications in the efficient computation of resultants for bivariate polynomials [14] [17] (see Chapter 3).

2.5 Computational Complexity

Since $Bez(f, g)$ is symmetric, we need to compute only the entries $B_{i,j}$ where $i \leq j$. Using the algorithm developed in Section 2.4, we initialize and march southwest updating these $(n^2 + n)/2$ entries as we go:

\[
\begin{bmatrix}
    L_0 R_1 & \cdots & L_0 R_n \\
    \vdots & \ddots & \vdots \\
    L_{n-1} R_n
\end{bmatrix}
\begin{bmatrix}
    * & * & \cdots & * & * \\
    \vdash & \cdots & \vdash & * \\
    \vdots & \ddots & \vdots & \vdots & \vdash \\
    \vdash & \cdots & \vdash & * \\
    * & * & \cdots & \cdots & *
\end{bmatrix}
\]

**Initialization**

**Marching**

During initialization, each of these entries requires two multiplications and one addition. Thus there are $(n^2 + n)$ multiplications and $(n^2 + n)/2$ additions in this initialization. As we march southwest, each entry above the diagonal $i = j$ except for the entries in the first row or the last column needs one addition. Thus there are $(n^2 - n)/2$ more additions. Therefore, to compute $Bez(f, g)$ by the new algorithm, we need to perform $(n^2 + n)$ multiplications and $n^2$ additions.

On the other hand, to compute $Bez(f, g)$ in the standard way, i.e. computing each entry separately by the entry formula (free of cancellation)

\[
B_{i,j} = \sum_{k=\max(0,i-j)}^{\min(i,n-i-j)} (a_{i-k} h_{j+1+k} - b_{i-k} a_{j+1+k}).
\]

we must compute the entries $\{B_{i,j}\}$ for all $i \leq j$:

\[
B_{i,j} = \begin{cases} 
    \sum_{k=0}^{i} (a_{i-k} h_{j+1+k} - b_{i-k} a_{j+1+k}), & i + j \leq n - 1; \\
    \sum_{k=0}^{n-i-j} (a_{i-k} h_{j+1+k} - b_{i-k} a_{j+1+k}), & i + j > n - 1.
\end{cases}
\]
Each summand requires two multiplications and one addition. When $n$ is odd, the total number of multiplications to compute all the $B_{i,j}$, $i \leq j$, is
\[
2 \cdot \sum_{i=0}^{(n-1)/2} (n-2i) \cdot (i+1) + 2 \cdot \sum_{j=(n+1)/2}^{n-1} (n-j) \cdot (2j-n+1)
= \frac{2n^3 + 9n^2 + 10n + 3}{12}.
\]
and the total number of additions is
\[
\sum_{i=0}^{(n-1)/2} (n-2i) \cdot (2i+1) + \sum_{j=(n+1)/2}^{n-1} (2n-2j-1) \cdot (2j-n+1)
= \frac{2n^3 + 3n^2 + 4n + 3}{12}.
\]
Similar results hold when $n$ is even. Therefore, the standard method needs $O(n^4)$ multiplications and additions to compute all the entries of $Brz(f, g)$, while our new technique requires only $O(n^2)$ additions and multiplications:

<table>
<thead>
<tr>
<th></th>
<th>Standard method</th>
<th>New algorithm</th>
</tr>
</thead>
<tbody>
<tr>
<td># of mul.</td>
<td>$O(n^3)$</td>
<td>$n^2 + n$</td>
</tr>
<tr>
<td># of add.</td>
<td>$O(n^3)$</td>
<td>$n^2$</td>
</tr>
</tbody>
</table>

Table 2.1: Computing the Bézout resultant matrix

### 2.6 Hybrid Resultants

A collection of $n-1$ hybrids of the Sylvester resultant and the Bézout resultant for $f$ and $g$ are constructed in [54] [60]. These hybrid resultants are composed of some columns from the Sylvester resultant and some columns from the Bézout resultant. In this section, we will generate these hybrids based on the transformation in Section 2.3 from the Sylvester to the Bézout resultant.
Let $H_j$, $j = 0, \ldots, n$, be the $j$-th hybrid resultant matrix
\[
\begin{bmatrix}
  L_0 & B_{0,0} & \cdots & B_{0,j-1} \\
  \vdots & \ddots & \ddots & \vdots \\
  L_{n-j-1} & \cdots & L_0 & B_{n-j-1,0} & \cdots & B_{n-j-1,j-1} \\
  \vdots & \cdots & \ddots & \ddots & \ddots & \vdots \\
  L_{n-1} & \cdots & L_j & B_{n-1,0} & \cdots & B_{n-1,j-1} \\
  L_n & \cdots & L_{j+1} \\
  \vdots \\
  L_n
\end{bmatrix}.
\]

That is, $H_j$ contains the first $2(n - j)$ truncated columns of $Syl(f, g)$, and the first $j$ columns of $Brz(f, g)$. Note that $H_j$ is a square matrix of order $2n - j$: moreover, $H_0 = Syl(f, g)$ and $H_n = Brz(f, g)$. Below we will show that $|H_j| = \pm |H_{j+1}|$, $j = 0, \ldots, n - 1$. Since $Syl(f, g)$ and $Brz(f, g)$ are known to be resultants for $f$ and $g$, it follows that $H_j$ is a resultant of $f$ and $g$ for $j = 0, \ldots, n$.

To proceed with our proof, let
\[
R_j^t = \begin{bmatrix}
  R_n \\
  \vdots \\
  R_j
\end{bmatrix}, \quad J = \begin{bmatrix}
  0 \\
  1
\end{bmatrix}.
\]

and consider the $(2n - j) \times (2n - j)$ matrix
\[
T_j = \begin{bmatrix}
  I_{2(n-j-1)} & 0 & R_{n-1}^J \\
  0 & 0 & R_n \\
  0 & I_j & 0
\end{bmatrix}.
\]

Multiplying $H_j$ by $T_j$, we get
\[ H_j \cdot T_j \]

\[
\begin{bmatrix}
L_0 & B_{0,0} & \cdots & B_{0,j-1} \\
\vdots & \ddots & \ddots & \vdots \\
L_{n-j-1} & L_0 & B_{n-j-1,0} & \cdots & B_{n-j-1,j-1} \\
\vdots & \ddots & \ddots & \vdots & \ddots \\
L_{n-1} & L_j & B_{n-1,0} & \cdots & B_{n-1,j-1} \\
L_n & L_{j+1} & & & \ddots \\
& & & L_n & \\
\end{bmatrix}
\]

\[
= \begin{bmatrix}
I_{2(n-j-1)} & 0 & R_{n-i}^{j+1} & 0 \\
0 & 0 & R_n & J \\
0 & I_j & 0 & 0 \\
\end{bmatrix}
\]

\[
\begin{bmatrix}
L_0 & B_{0,0} & \cdots & B_{0,j} \\
\vdots & \ddots & \ddots & \vdots \\
L_{n-j-2} & L_0 & B_{n-j-2,0} & \cdots & B_{n-j-2,j} \\
\vdots & \ddots & \ddots & \vdots & \ddots & b_0 \\
L_{n-1} & L_j & B_{n-1,0} & \cdots & B_{n-1,j} \\
L_n & L_{j+1} & & & \ddots & b_{n-1} \\
& & & L_n & b_n \\
\end{bmatrix}
\]

(2.16)

Notice that the top-left \((2n - j - 1) \times (2n - j - 1)\) submatrix is exactly \(H_{j+1}\), so Equation (2.16) can be rewritten as

\[ H_j \cdot T_j = \begin{bmatrix} H_{j+1} & \ast \\ 0 & b_n \end{bmatrix}. \]

(2.17)

Therefore,

\[ |H_j| \cdot |T_j| = |H_{j+1}| \cdot b_n. \]

But by construction,

\[ |T_j| = \pm |[ R_n \ J ]| = \pm \begin{bmatrix} b_n & 0 \\ -a_n & 1 \end{bmatrix} = \pm b_n. \]
\[ |H_j| = \pm |H_{j+1}|. \] (2.18)

Thus we generate a sequence of resultants that are hybrids of the Sylvester and Bézout resultants. Since \(H_0 = Syl(f, g)\) and \(H_n = Bez(f, g)\), we obtain in this fashion a direct proof that
\[ |Syl(f, g)| = \pm |Bez(f, g)| \] (2.19)
without appealing to any specific properties of resultants.

2.7 Non-Homogeneous Bézout Matrices

When the coefficients \(a_i, b_i, 0 \leq i \leq n\), of \(f\) and \(g\) are treated as formal symbols, the Bézout matrix \(Bez(f, g)\) is homogeneous in the sense that each entry is quadratic in the \(a_i\)'s and \(b_i\)'s.

If the degree of \(g\) is \(m\) and \(m < n\), then we have \(b_{m+1} = \cdots = b_n = 0\). In this situation, \(|Syl(f, g)|\) is not exactly the resultant of \(f, g\) — it has an extraneous factor of \((a_n)^{n-m}\). Rather the correct Sylvester matrix \(Syl(f, g)_{n,m}\), whose determinant is the exact resultant, is the coefficient matrix of \(n + m\) polynomials:
\[
\begin{bmatrix}
L & tL & \cdots & t^{m-1}L & t^mg & \cdots & t^{n-1}g
\end{bmatrix} =
\begin{bmatrix}
1 & \cdots & t^{n+m-1}
\end{bmatrix} Syl(f, g)_{n,m}. \tag{2.20}
\]

Note that \(Syl(f, g)\) is of order \(2n\), while \(Syl(f, g)_{n,m}\) is of order \(n + m\) and
\[ |Syl(f, g)| = \pm (a_n)^{n-m} \cdot |Syl(f, g)_{n,m}|. \] (2.21)

Similarly, the Bézout resultant matrix \(Bez(f, g)\) obtained from the Cayley expression \(\Delta_n(t, \beta)\) also contains the extraneous factor \((a_n)^{n-m}\). It is pointed out in [19] that in this case the correct Bézout resultant for \(f\) and \(g\) can be written as the determinant of a non-homogeneous matrix \(Bez(f, g)_{n,m}\) of size \(n \times n\). The matrix \(Bez(f, g)_{n,m}\) has \(m\) columns of quadratic entries consisting of coefficients of both \(f\) and \(g\), and \(n - m\) columns of linear entries consisting of coefficients of \(g\). Below we present a simple alternative proof of this fact.
Recall from Section 2.3 that
\[
\begin{bmatrix}
B_{cz}(f \cdot g) \\
0
\end{bmatrix}
= Syl(f \cdot g) \cdot T(f \cdot g)
\]
\[
= Syl(f \cdot g) \cdot \begin{bmatrix}
R_1 & \cdots & R_n \\
\vdots & \ddots & \vdots \\
R_n
\end{bmatrix}
\]
\[
= Syl(f \cdot g) \cdot \begin{bmatrix}
R_1 & \cdots & R_n \\
\vdots & \ddots & \vdots \\
R_{n-m} & \cdots & R_n \\
R_n
\end{bmatrix}
Syl(f \cdot g) \cdot \begin{bmatrix}
R_{n+1} & \cdots & R_n \\
\vdots & \ddots & \vdots \\
R_n & \cdots & 0 \\
0 & \cdots & 0
\end{bmatrix}
\]
(2.22)

Let us focus on the second matrix product on the right hand side of Equation (2.22).

With \( h_{n+1} = \cdots = h_n = 0 \).

\[
Syl(f \cdot g) \cdot \begin{bmatrix}
R_{n+1} & \cdots & R_n \\
\vdots & \ddots & \vdots \\
0 & \cdots & 0
\end{bmatrix}
= \begin{bmatrix}
L_0 \\
\vdots \\
L_{n-m} & \cdots & L_0 \\
\vdots & \ddots & \vdots \\
L_{n-1} & \cdots & L_m \\
L_n & \cdots & L_{m+1} \\
L_n \\
0_{m \times 2}
\end{bmatrix}
\cdot
\begin{bmatrix}
0 & 0 \\
-\alpha_{n+1} & \cdots & -\alpha_n \\
\vdots & \ddots & \vdots \\
0 \\
-\alpha_n
\end{bmatrix}
\]
\[
\begin{bmatrix}
 b_0 \\
 \vdots \\
 b_{n-m-1} & \cdots & b_0 \\
 \vdots & \vdots & \vdots \\
 b_{n-1} & \cdots & b_m \\
 b_n & \cdots & b_{m+1} \\
 \vdots \\
 b_n \\
 0_{m \times 1}
\end{bmatrix}
\cdot
\begin{bmatrix}
 -a_{m+1} & \cdots & -a_n \\
 \vdots & \ddots & \vdots \\
 -a_n & & \ddots \\
 -a_n & & & \ddots 
\end{bmatrix}
\].
\tag{2.23}
\]

Note that the term \(0_{m \times 2}\) in the first factor on the right hand side of Equation (2.23) represents an \(m \times 2\) matrix of zero entries, since \(Syl(f, g)\) has \(2n\) rows. Let
\[
A_{m,n} = \begin{bmatrix}
 -a_{m+1} & \cdots & -a_n \\
 \vdots & \ddots & \vdots \\
 -a_n & & \ddots 
\end{bmatrix}
\]

Then putting together Equations (2.22) and (2.23), and using the notation of Equation (2.12), we obtain

\[
Bez(f, g)
= \begin{bmatrix}
 B_{0,0} & \cdots & B_{0,m-1} \\
 \vdots & \ddots & \vdots \\
 B_{n-1,0} & \cdots & B_{n-1,m-1}
\end{bmatrix}
\cdot
\begin{bmatrix}
 b_0 \\
 \vdots \\
 b_m & \cdots & b_0 \\
 \vdots \\
 b_m \\
 b_n
\end{bmatrix}
\cdot
\begin{bmatrix}
 I_m \\
 A_{m,n}
\end{bmatrix}
\].
\tag{2.24}
\]

Let us denote the \(n \times n\) matrix in the first factor on the right hand side of Equation (2.24) by \(Bez(f, g)_{n,m}\). Then we can rewrite Equation (2.24) as

\[
Bez(f, g) = Bez(f, g)_{n,m} \cdot \begin{bmatrix}
 I_m \\
 A_{m,n}
\end{bmatrix}
\].
\tag{2.25}
Hence

\[ |B rz(f, g)| = \pm (-a_n)^{n-m} \cdot |B rz(f, g)_{n,m}|. \]  \hspace{1cm} (2.26)

Now it follows from Equations (2.19), (2.21), (2.26) that

\[ |Syl(f, g)_{n,m}| = \pm |B rz(f, g)_{n,m}|. \]  \hspace{1cm} (2.27)

That is, the coefficient matrix of the following \( n \) polynomials is the Bézout matrix for \( f \) and \( g 
\[
\begin{array}{cccc}
D_0 & D_1 & \cdots & D_{m-1} \\
0.1 & 0.2 & \cdots & g \\
0.2 & 0.3 & \cdots & t^{n-1}g \\
0.3 & 0.4 & \cdots & \\
0.4 & 0.5 & \cdots & \\
0.5 & 1.5 & \\
\end{array}
\]  \hspace{1cm} (2.28)

where the polynomials \( D_j, j = 0, \ldots, m - 1 \), are defined in Equation (2.3).

For example, we have, when \( n = 5, m = 2 \).

\[
B rz(f, g)_{5,2} = \\
\begin{bmatrix}
0.1 & 0.2 & b_0 \\
0.2 & 0.3 + 1.2 & b_1 & b_0 \\
0.3 & 0.4 + 1.3 & b_2 & b_1 & b_0 \\
0.4 & 0.5 + 1.4 & b_2 & b_1 \\
0.5 & 1.5 & b_2 \\
\end{bmatrix}
\]

and when \( n = 5, m = 3 \).

\[
B rz(f, g)_{5,3} = \\
\begin{bmatrix}
0.1 & 0.2 & 0.3 & b_0 \\
0.2 & 0.3 + 1.2 & 0.4 + 1.3 & b_1 & b_0 \\
0.3 & 0.4 + 1.3 & 0.5 + 1.4 + 2.3 & b_2 & b_1 & b_0 \\
0.4 & 0.5 + 1.4 & 1.5 + 2.4 & b_2 & b_1 \\
0.5 & 1.5 & 2.5 & b_2 \\
\end{bmatrix}
\]

where \( i, j = L_i \cdot R_j = a_i b_j - a_j b_i, \ 0 \leq i, j \leq n \).

In summary, there exist matrices \( \tilde{T}(f, g)_{n,m} \) and \( T(f, g)_{n,m} \) such that

\[
Syl(f, g) \cdot \tilde{T}(f, g)_{n,m} = \\
\begin{bmatrix}
B rz(f, g)_{n,m} \\
0 \\
\end{bmatrix} = Syl(f, g)_{n,m} \cdot T(f, g)_{n,m}.
\]
In fact, let $J = \begin{bmatrix} 0 & 1 \end{bmatrix}^T$; then

$$
\tilde{T}(f, g)_{n,m} = \begin{bmatrix} 
R_1 & \cdots & R_m & J \\
R_2 & \cdots & R_{m+1} & J \\
\vdots & \ddots & \vdots & \vdots \\
R_{n-m} & \cdots & R_{n-1} & J \\
R_{n-m+1} & \cdots & R_n & \\
\vdots & \ddots & \vdots & \vdots \\
R_n & & & 
\end{bmatrix}.
$$

Moreover, when $n - m > m$, 

$$
T(f, g)_{n,m} = \begin{bmatrix} 
R_1 & \cdots & R_m & J \\
\vdots & \ddots & \vdots & \vdots \\
R_m & \cdots & R_{2m-1} & J \\
-a_{m+1} & \cdots & -a_{2m} & 1 \\
\vdots & \ddots & \vdots & \vdots \\
-a_{n-m} & \cdots & -a_{n-1} & 1 \\
-a_{n-m+1} & \cdots & -a_n & \\
\vdots & \ddots & \vdots & \vdots \\
-a_n & & & 
\end{bmatrix}.
$$

and when $n - m \leq m$, 

$$
T(f, g)_{n,m} = \begin{bmatrix} 
R_1 & \cdots & R_{n-m} & \cdots & R_m & J \\
\vdots & \ddots & \vdots & \ddots & \vdots & \vdots \\
R_{n-m} & \cdots & R_{2n-2m-1} & \cdots & R_{n-1} & J \\
R_{n-m+1} & \cdots & R_{2n-2m} & \cdots & R_n & \\
\vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\
R_n & \cdots & R_{n-1} & \\
-a_{m+1} & \cdots & -a_n & \\
\vdots & \ddots & \vdots & \vdots \\
-a_n & & & 
\end{bmatrix}.
$$
Chapter 3

Bivariate Tensor Product Resultants

Dixon [25] describes three distinct determinant expressions for the resultant of three bivariate polynomials of bidegree \((m, n)\). The first formulation is the determinant of a \(6mn \times 6mn\) matrix generated by Sylvester’s dialytic method. We shall refer to this determinant as the *Sylvester resultant*. The second expression is the determinant of a \(2mn \times 2mn\) matrix found by using an extension of Cayley’s determinant device for generating the Bézout resultant for two univariate polynomials. This \(2mn \times 2mn\) matrix has come to be known as the Dixon resultant [6] [42] [49] [50] [51]. However, to distinguish this determinant expression from Dixon’s other two formulations for the resultant, we shall call this determinant the *Cayley resultant*. The third representation is the determinant of a \(3mn \times 3mn\) matrix generated by combining Cayley’s determinant device with Sylvester’s dialytic method. We shall refer to this determinant as the *mixed Cayley-Sylvester resultant*.

Dixon introduces these three representations for the resultant independently, without attempting to relate one to the other. In this chapter, we will derive connections between these three distinct representations for the resultant. In particular, we will prove that the polynomials represented by the columns of the Cayley and the mixed Cayley-Sylvester resultants are linear combinations of the polynomials represented by the columns of the Sylvester resultant. Thus there are matrices relating the Sylvester resultant to the Cayley and mixed Cayley-Sylvester resultants. We shall show that these transformation matrices all have similar, simple, upper triangular, block symmetric structures and the blocks themselves are either Sylvester-like or symmetric. In addition, we shall provide straightforward formulas for the entries of these matrices.
Reexamining the three Dixon resultant matrices, we find that they too have simple block structures compatible with the block structures of the transformation matrices. Indeed, we shall see that the blocks for the transformation matrices are essentially transposes of the blocks for these resultant matrices, although the arrangements of the blocks within these matrices differ. We shall also show that the blocks of the Sylvester and mixed Cayley-Sylvester matrices are related to the Sylvester and Bézout resultants of certain univariate polynomials.

The entries of the Sylvester resultant are just the coefficients of the original polynomials, but the entries of the Cayley and mixed Cayley-Sylvester resultants are more complicated expressions in these coefficients. We shall show how to take advantage of the block structure of the resultant matrices together with the block structure of the transformation matrices to simplify the calculation of the entries of the Cayley and mixed Cayley-Sylvester resultants and to remove redundant computations. These techniques for the efficient computation of the entries of the Cayley and mixed Cayley-Sylvester resultants are extensions to the bivariate setting of similar results on the efficient computation of the entries of the Bézout resultant of two univariate polynomials [Ch 2] [15].

We will also derive a new class of representations for the bivariate resultant in terms of determinants of hybrid matrices formed by mixing and matching appropriate columns from the Sylvester, Cayley, or mixed Cayley-Sylvester matrices. The derivation of these hybrid resultants is also based on the block structure of the Dixon resultants and the accompanying transformation matrices. Again these new hybrid resultants are generalizations to the bivariate setting of known hybrids between the Sylvester and Bézout resultants for two univariate polynomials [Ch 2] [15] [54] [60].

This chapter is structured in the following manner. We begin in Section 3.1 by establishing our notation and reviewing Dixon's formulations of the Sylvester, Cayley, and mixed Cayley-Sylvester resultants. Next, in Sections 3.2 and 3.3 we derive explicit formulas for the entries of the matrices relating the Sylvester resultant
to the Cayley and mixed Cayley-Sylvester resultants and discuss the block structure and symmetry properties of these transformation matrices. We devote Section 3.4 to deriving similar results for the transformation relating the mixed Cayley-Sylvester and Cayley resultants. In Section 3.5 we analyze the block structures of the three Dixon resultant matrices and in Section 3.6 we illustrate how the blocks of the Sylvester resultant matrices and the mixed Cayley-Sylvester resultant matrices are related by a set of convolution identities. We show in Section 3.7 how to take advantage of the block structures of the Sylvester resultant matrices and the transformation matrices to simplify the calculation of the entries of the Cayley resultants and to make the calculations more efficient by removing redundant computations. A comparison of the computational complexity of this new method versus the standard method for computing the entries of the Cayley resultant is provided in Section 3.8. In Section 3.9, we show how to construct hybrid resultant matrices from the Sylvester, Cayley, and mixed Cayley-Sylvester resultant matrices.

3.1 The Three Dixon Resultants

Consider three bivariate polynomials of bidegree \((m, n)\):

\[
\begin{align*}
    f(s, t) &= \sum_{i=0}^{m} \sum_{j=0}^{n} a_{ij} s^i t^j, \\
    g(s, t) &= \sum_{k=0}^{m} \sum_{l=0}^{n} b_{kl} s^k t^l, \\
    h(s, t) &= \sum_{p=0}^{m} \sum_{q=0}^{n} c_{pq} s^p t^q.
\end{align*}
\]

Dixon outlines three methods for constructing a resultant for \(f, g, h\) [25]. In this section, we briefly review the construction of each of Dixon's three resultants.

3.1.1 The Sylvester Resultant \(\text{Syl}(f, g, h)\)

The Sylvester resultant for \(f(s, t), g(s, t), h(s, t)\) is constructed using Sylvester's dialytic method. Consider the \(6mn\) polynomials \(\{s^\sigma t^\tau f, s^\sigma t^\tau g, s^\sigma t^\tau h \mid \sigma = 0, \ldots, 2m - 1; \tau = 0, \ldots, n - 1\}\). Let \(L(s, t) = \begin{bmatrix} f(s, t) & g(s, t) & h(s, t) \end{bmatrix}\); then this system of
polynomials can be written in matrix notation as

\[
\begin{bmatrix}
L & \ldots & L & s^t L & s^t_{t+1} L & \ldots & s^{2m-1} L & 1
\end{bmatrix}
\begin{bmatrix}
1 \\
\vdots \\
s^t \\
s^t_{t+1} \\
\vdots \\
s^{m-1}_{2n-1}
\end{bmatrix}^T
= Syl(f, g, h), \quad (3.1)
\]

where the rows are indexed lexicographically with \( s > t \). That is, the monomials are ordered as 1, \( t, \ldots, t^{2n-1} \), \( s^t, s^t_{t+1}, \ldots, s^{3m-1}, \ldots, s^{3m-1}_{2n-1} \).

Notice that the coefficient matrix \( Syl(f, g, h) \) is a square matrix of order \( 6mn \). The Sylvester resultant for \( f, g, h \) is simply \( |Syl(f, g, h)| \).

### 3.1.2 The Cayley Resultant \( Cay(f, g, h) \)

The Cayley resultant for \( f, g, h \) can be derived from the Cayley expression

\[
\Delta_{m,n}(s, t, \alpha, \beta) = \begin{vmatrix}
f(s, t) & g(s, t) & h(s, t) \\
f(\alpha, t) & g(\alpha, t) & h(\alpha, t) \\
f(\alpha, \beta) & g(\alpha, \beta) & h(\alpha, \beta)
\end{vmatrix} / (\alpha - s)(\beta - t). \quad (3.2)
\]

Since the numerator vanishes when \( \alpha = s \) or \( \beta = t \), the numerator is divisible by \( (\alpha - s)(\beta - t) \). Hence \( \Delta_{m,n}(s, t, \alpha, \beta) \) is a polynomial in \( s, t, \alpha, \beta \), so

\[
\Delta_{m,n}(s, t, \alpha, \beta) = \sum_{u=0}^{2m-1} \sum_{v=0}^{n-1} \sum_{s=0}^{m-1} \sum_{r=0}^{2n-1} c_{r} s^t \alpha^u \beta^v.
\]

In matrix form,

\[
\Delta_{m,n}(s, t, \alpha, \beta) =
\begin{bmatrix}
1 \\
\vdots \\
s^t \\
s^t_{t+1} \\
\vdots \\
s^{m-1}_{2n-1}
\end{bmatrix}
\begin{bmatrix}
1 \\
\vdots \\
\alpha^u \\
\alpha^u_{t+1} \\
\vdots \\
\alpha^{2m-1}_{2n-1}
\end{bmatrix}
= Cay(f, g, h), \quad (3.3)
\]
The rows and columns of $Cay(f, g, h)$ are indexed lexicographically by the pairs $(s^a t^r, \alpha^n \beta^n)$ with $s > t, \alpha > \beta$. Notice that the coefficient matrix $Cay(f, g, h)$ is again a square matrix but of order $2mn$. The Cayley resultant for $f, g, h$ is $|Cay(f, g, h)|$.

### 3.1.3 The Mixed Cayley-Sylvester Resultant $Mix(f, g, h)$

The mixed Cayley-Sylvester resultant can be derived by combining Cayley’s determinant device with Sylvester’s dialytic method. Consider the expression

$$
\phi(g, h) = \begin{vmatrix}
g(s, t) & h(s, t) \\
g(s, \beta) & h(s, \beta)
\end{vmatrix} / (\beta - t).
$$

(3.4)

Since the numerator is always divisible by the denominator, $\phi(g, h)$ is a polynomial in $s, t, \beta$, with degree $2m$ in $s, n - 1$ in $t$, and $n - 1$ in $\beta$. Collecting the coefficients of $\beta^j, j = 0, \ldots, n - 1$, we get $n$ polynomials $\overline{f}_j(s, t)$ such that $\phi(g, h) = \sum_{j=0}^{n-1} \overline{f}_j(s, t) \beta^j$.

Multiplying these $n$ polynomials by the $m$ monomials $1, s, \ldots, s^{m-1}$ yields $mn$ polynomials $s^i \overline{f}_j(s, t), 0 \leq i \leq m - 1, 0 \leq j \leq n - 1$, in the $3mn$ monomials $s^i \beta^j, 0 \leq i \leq 3m - 1, 0 \leq j \leq n - 1$.

Now do the same for

$$
\phi(h, f) = \begin{vmatrix}
h(s, t) & f(s, t) \\
h(s, \beta) & f(s, \beta)
\end{vmatrix} / (\beta - t), \quad \phi(f, g) = \begin{vmatrix}
f(s, t) & g(s, t) \\
f(s, \beta) & g(s, \beta)
\end{vmatrix} / (\beta - t).
$$

Altogether we get $3mn$ polynomials $s^i \overline{g}_j, s^i \overline{f}_j, s^i \overline{h}_j, j = 0, \ldots, n - 1, i = 0, \ldots, m - 1$, in $3mn$ monomials, where

$$
\phi(g, h) = \sum_{j=0}^{n-1} \overline{f}_j(s, t) \beta^j, \quad \phi(h, f) = \sum_{j=0}^{n-1} \overline{g}_j(s, t) \beta^j, \quad \phi(f, g) = \sum_{j=0}^{n-1} \overline{h}_j(s, t) \beta^j.
$$

Let $\nu(s, \alpha) = \sum_{n=0}^{m-1} s^n \alpha^n$ and $\overline{L}_k = [\overline{f}_k \quad \overline{g}_k \quad \overline{h}_k]$. Then
\[
\begin{align*}
\psi(s, \alpha) \left[ \begin{array}{ccc}
\phi(g, h) & \phi(h, f) & \phi(f, g) \\
\end{array} \right] \\
= \left[ \begin{array}{cccc}
\bar{L}_0 & \cdots & \bar{L}_{n-1} & \cdots & s^{m-1}\bar{L}_0 & \cdots & s^{m-1}\bar{L}_{n-1} \\
\end{array} \right] \left[ \begin{array}{c}
1 \\
\vdots \\
\alpha^{n-j} \\
\vdots \\
\alpha^{m-1-j} \\
\end{array} \right]
\end{align*}
\]

\[
\begin{align*}
&= \begin{bmatrix}
1 \\
\vdots \\
\alpha^{n-j} \\
\vdots \\
\alpha^{m-1-j}
\end{bmatrix}^T \text{Mix}(f, g, h) \begin{bmatrix}
1 \\
\vdots \\
\alpha^{n-j} \\
\vdots \\
\alpha^{m-1-j}
\end{bmatrix}.
\end{align*}
\]

(3.5)

where \(\text{Mix}(f, g, h)\) is the coefficient matrix of these \(3mn\) polynomials. Note that in this notation, \(\text{Mix}(f, g, h)\) has \(3mn\) rows and \(mn\) columns, but each entry of \(\text{Mix}(f, g, h)\) is actually a \(1 \times 3\) matrix. Thus the real size of \(\text{Mix}(f, g, h)\) is \(3mn \times 3mn\); hence \(\text{Mix}(f, g, h)\) is a square matrix. The mixed Cayley-Sylvester resultant is \(|\text{Mix}(f, g, h)|\).

3.1.4 Notation

We plan to derive explicit formulas for the entries of the conversion matrices between \(\text{Syl}(f, g, h)\), \(\text{Cay}(f, g, h)\) and \(\text{Mix}(f, g, h)\), and then to apply these formulas to elucidate the structure and simplify the computation of the entries of the three Dixon resultants. For convenience, we adopt the following notation:
\[ |i, j; k, l; p, q| = \begin{vmatrix} a_{i,j} & b_{i,j} & c_{i,j} \\ a_{k,l} & b_{k,l} & c_{k,l} \\ c_{p,q} & b_{p,q} & c_{p,q} \end{vmatrix} \]

\[ A_{k,l,p,q} = \begin{vmatrix} b_{k,l} & c_{k,l} \\ b_{p,q} & c_{p,q} \end{vmatrix}, \quad B_{k,l,p,q} = \begin{vmatrix} c_{k,l} & a_{k,l} \\ c_{p,q} & a_{p,q} \end{vmatrix}, \quad C_{k,l,p,q} = \begin{vmatrix} a_{k,l} & b_{k,l} \\ a_{p,q} & b_{p,q} \end{vmatrix}. \]

\[ L_{i,j} = \begin{bmatrix} a_{i,j} & b_{i,j} & c_{i,j} \end{bmatrix}, \quad R_{k,l,p,q} = \begin{bmatrix} A_{k,l,p,q} & B_{k,l,p,q} & C_{k,l,p,q} \end{bmatrix}. \]

\[ X_{p,q} = \begin{bmatrix} 0 & -c_{p,q} & b_{p,q} \\ c_{p,q} & 0 & -a_{p,q} \\ -b_{p,q} & a_{p,q} & 0 \end{bmatrix}. \]

**Example 3.1** \((m = n = 1)\)

\[ Syl(f, g, h) = \begin{bmatrix} a_{0,0} & b_{0,0} & c_{0,0} & 0 & 0 & 0 \\ a_{0,1} & b_{0,1} & c_{0,1} & 0 & 0 & 0 \\ a_{1,0} & b_{1,0} & c_{1,0} & a_{0,0} & b_{0,0} & c_{0,0} \\ a_{1,1} & b_{1,1} & c_{1,1} & a_{0,1} & b_{0,1} & c_{0,1} \\ 0 & 0 & 0 & a_{1,0} & b_{1,0} & c_{1,0} \\ 0 & 0 & 0 & a_{1,1} & b_{1,1} & c_{1,1} \end{bmatrix} = \begin{bmatrix} L_{0,0} & 0 \\ L_{0,1} & 0 \\ L_{1,0} & L_{0,0} \\ L_{1,1} & L_{0,1} \\ 0 & L_{1,0} \\ 0 & L_{1,1} \end{bmatrix}. \]

\[ Cay(f, g, h) = \begin{bmatrix} 0, 0; 1, 0; 0, 1 & -1, 0; 0, 0; 1, 1 \\ 0, 0; 1, 1; 0, 1 & -1, 0; 0, 1; 1, 1 \end{bmatrix}. \]

\[ Mix(f, g, h) = \begin{bmatrix} A_{0,0;0,1} & B_{0,0;0,1} & C_{0,0;0,1} \\ A_{0,0;1,1} + A_{1,0;0,1} & B_{0,0;1,1} + B_{1,0;0,1} & C_{0,0;1,1} + C_{1,0;0,1} \\ A_{1,0;1,1} & B_{1,0;1,1} & C_{1,0;1,1} \end{bmatrix} = \begin{bmatrix} R_{0,0;0,1} \\ R_{0,0;1,1} + R_{1,0;0,1} \\ R_{1,0;1,1} \end{bmatrix}. \]
3.1.5 Review and Preview

In this section, we have introduced the three Dixon determinants for the resultant of three bivariate polynomials of bidegree \((m, n)\). In Sections 3.2, 3.3, 3.4, we will investigate relationships between the three Dixon resultant matrices \(Syl(f, g, h)\), \(Cay(f, g, h)\), and \(Mix(f, g, h)\). It follows by inspection from Equations (3.1), (3.3), (3.5) that the columns of Dixon’s three resultant matrices represent bivariate polynomials in \(s, t\). In particular, the columns of \(Syl(f, g, h)\), \(Mix(f, g, h)\), \(Cay(f, g, h)\) are respectively bivariate polynomials of bidegree \((3m - 1, 2n - 1)\), \((3m - 1, n - 1)\), \((m - 1, 2n - 1)\). Since the columns of \(Syl(f, g, h)\) are linearly independent, the \(6mn\) polynomials represented by these columns span the space of bivariate polynomials of bidegree \((3m - 1, 2n - 1)\). Therefore the polynomials represented by the columns of \(Mix(f, g, h)\) and \(Cay(f, g, h)\) must be linear combinations of the polynomials represented by the columns of \(Syl(f, g, h)\). Sections 3.2, 3.3 are devoted to deriving explicit formulas for the transformations from \(Syl(f, g, h)\) to \(Mix(f, g, h)\) and \(Cay(f, g, h)\) compatible with the polynomial indexing of the rows and columns.

3.2 The Transformation from \(Syl(f, g, h)\) to \(Mix(f, g, h)\)

Recall that the columns of \(Syl(f, g, h)\) and \(Mix(f, g, h)\) represent respectively bivariate polynomials of bidegree \((3m - 1, 2n - 1)\) and \((3m - 1, n - 1)\). In order to derive a \(6mn \times 3mn\) conversion matrix \(G(f, g, h)\) such that

\[
Syl(f, g, h) \cdot G(f, g, h) = \begin{bmatrix} Mix(f, g, h) \\ 0 \end{bmatrix}.
\]

(3.6)

that is, in order to insure that the \(3mn\) zero rows are not interleaved with the \(3mn\) non-zero rows, we have to abandon the default lexicographic orders \(s > t, \alpha > \beta\) and impose instead the lexicographic orders \(t > s, \beta > \alpha\). Matrices with this order of rows and columns will be indicated by an "*" to differentiate them from those matrices with the default lexicographic orders. In particular, \(Syl(f, g, h)^*\) and \(Mix(f, g, h)^*\)
are defined by

\[
\begin{bmatrix}
L & \cdots & t^r s^\sigma L & t^r s^{\sigma+1} L & \cdots & t^{n-1} s^{2m-1} L
\end{bmatrix} =
\begin{bmatrix}
1 \\
\vdots \\
1^r s^\sigma \\
1^r s^{\sigma+1} \\
\vdots \\
1^{2n-1} s^{3m-1}
\end{bmatrix}
\quad Syl(f, g, h)^*.
\]

(3.7)

and

\[
\omega(s, \alpha) \begin{bmatrix}
\phi(g, h) & \phi(h, f) & \phi(f, g)
\end{bmatrix}
= 
\begin{bmatrix}
1 \\
\vdots \\
1^r s^\sigma \\
1^r s^{\sigma+1} \\
\vdots \\
1^{n-1} s^{3m-1}
\end{bmatrix}^T \begin{bmatrix}
1 \\
\vdots \\
\beta^r \alpha^h \\
\beta^r \alpha^{u+1} \\
\vdots \\
\beta^{n-1} \alpha^{m-1}
\end{bmatrix} = \text{Mix}(f, g, h)^*,
\]

(3.8)

The derivation below of the transformation matrix produces an explicit entry formula that reveals the simple elegant block structure of \( G(f, g, h)^* \).

### 3.2.1 The Conversion Matrix \( G(f, g, h)^* \)

To find \( G(f, g, h)^* \), we expand the numerator on the right hand side of Equation (3.4) by cross-multiplying and writing \( g(s, \jmath) \), \( h(s, \jmath) \) as explicit sums of monomials to obtain:

\[
\phi(g, h) = \sum_{p=0}^{m} \sum_{q=0}^{n} (g(s, t)c_{p,q} - h(s, t)b_{p,q}) s^p \jmath^q / (\jmath - t).
\]
Applying the technique of truncated formal power series from Chapter 2, Section 2.2 with the numerator as a polynomial in \( \beta \) yields:

\[
\phi(g, h) = \sum_{p=0}^{m} \sum_{q=0}^{n} (g(s, t)c_{p,q} - h(s, t)b_{p,q}) s^p \sum_{v=0}^{q-1} t^{q-1-v} \alpha^v.
\]

Rearranging the range of \( q \) and \( v \) results in:

\[
\phi(g, h) = \sum_{r=0}^{n-1} \sum_{p=0}^{m} \sum_{q=r+1}^{n} (g(s, t)c_{p,q} - h(s, t)b_{p,q}) s^p t^{q-1-v} \alpha^v.
\]

Substituting \( L = \begin{bmatrix} f & g & h \end{bmatrix} \) leads to:

\[
\phi(g, h) = \sum_{r=0}^{n-1} \sum_{p=0}^{m} \sum_{q=r+1}^{n} L \begin{bmatrix} 0 \\ c_{p,q} \\ -b_{p,q} \end{bmatrix} s^p t^{q-1-v} \alpha^v.
\]

Hence

\[
u(s, \alpha) \phi(g, h) = \sum_{u=0}^{n-1} s^u \alpha^u \phi(g, h)
\]

\[
= \sum_{u=0}^{n-1} \sum_{v=0}^{m-1} \sum_{p=0}^{n} \sum_{q=v+1}^{n} L \begin{bmatrix} 0 \\ c_{p,q} \\ -b_{p,q} \end{bmatrix} s^p t^{q-1-v} \alpha^u \alpha^v
\]

Recalling that

\[
X_{p,q} = \begin{bmatrix} 0 & -c_{p,q} & b_{p,q} \\ c_{p,q} & 0 & -a_{p,q} \\ -b_{p,q} & a_{p,q} & 0 \end{bmatrix},
\]

we have

\[
u(s, \alpha) \begin{bmatrix} \phi(g, h) & \phi(h, f) & \phi(f, g) \end{bmatrix} = \sum_{u=0}^{n-1} \sum_{v=0}^{m-1} \sum_{p=0}^{m} \sum_{q=v+1}^{n} LX_{p,q} s^{p+u} t^{q-1-v} \alpha^u \alpha^v. \tag{3.9}
\]

Since \( 0 \leq p + u \leq 2m - 1 \) and \( 0 \leq q - 1 - v \leq n - 1 \), this equation can be written as
\[ \psi(s, \alpha) \begin{bmatrix} \phi(g, h) & \phi(h, f) & \phi(f, g) \end{bmatrix} \]

\[ = \begin{bmatrix} L^T \vspace{1em} \\ \vdots \vspace{1em} \\ L^T \vspace{1em} \end{bmatrix}^T \begin{bmatrix} 1 \vspace{1em} \\ \vdots \vspace{1em} \\ 1 \end{bmatrix} \begin{bmatrix} \beta_v \alpha^u \vspace{1em} \\ \vdots \vspace{1em} \\ \beta_v \alpha^{u+1} \vspace{1em} \end{bmatrix} G(f, g, h)^* \]

\[ = \begin{bmatrix} L^T \vspace{1em} \\ \vdots \vspace{1em} \\ L^T \vspace{1em} \end{bmatrix}^T \begin{bmatrix} 1 \vspace{1em} \\ \vdots \vspace{1em} \\ \beta_v \alpha^{m-1} \vspace{1em} \end{bmatrix} \]

where \( G(f, g, h)^* \) is the \( 2mn \times mn \) matrix whose entries are the \( 3 \times 3 \) matrices \( X_{p,q} \).

Thus the actual dimensions of \( G(f, g, h)^* \) are \( 6mn \times 3mn \).

It follows by the definition of \( \text{Syl}(f, g, h)^* \) (Equation (3.7)) that

\[ \psi(s, \alpha) \begin{bmatrix} \phi(g, h) & \phi(h, f) & \phi(f, g) \end{bmatrix} \]

\[ = \begin{bmatrix} 1 \vspace{1em} \\ \vdots \vspace{1em} \\ \beta_v \alpha^{m-1} \vspace{1em} \end{bmatrix} \begin{bmatrix} 1 \vspace{1em} \\ \vdots \vspace{1em} \\ \beta_v \alpha^{m-1} \end{bmatrix} \begin{bmatrix} \beta_v \alpha^u \vspace{1em} \\ \vdots \vspace{1em} \\ \beta_v \alpha^{u+1} \vspace{1em} \end{bmatrix} \]

\[ = \begin{bmatrix} L^T \vspace{1em} \\ \vdots \vspace{1em} \\ L^T \vspace{1em} \end{bmatrix}^T \begin{bmatrix} 1 \vspace{1em} \\ \vdots \vspace{1em} \\ \beta_v \alpha^{m-1} \end{bmatrix} \begin{bmatrix} \beta_v \alpha^u \vspace{1em} \\ \vdots \vspace{1em} \\ \beta_v \alpha^{u+1} \vspace{1em} \end{bmatrix} \]

Therefore, by the definition of \( \text{Mix}(f, g, h)^* \) (Equation (3.8)), we see immediately that

\[ \text{Syl}(f, g, h)^* \cdot G(f, g, h)^* = \begin{bmatrix} \text{Mix}(f, g, h)^* \vspace{1em} \\ 0 \end{bmatrix} \]
3.2.2 The Entries of $G(f.g.h)^*$

By solving $p+u = \sigma, q-1-v = \tau$ in Equation (3.9), we obtain $p = \sigma-u, q = \tau+v+1$. Consequently the $3 \times 3$ entry $X_{p,q}$ of $G(f.g.h)^*$ indexed by $(t^\sigma s^q, f^v \alpha^n)$ is

$$X_{\sigma-u, \tau+v+1}$$

(3.13)

for $0 \leq \sigma - u \leq m$ and $\tau + v + 1 \leq n$; the entries are zero everywhere else.

3.2.3 Properties of $G(f.g.h)^*$

The structure of the $6mn \times 3mn$ conversion matrix $G(f.g.h)^*$ is most revealing if we view it as a $2mn \times mn$ matrix whose entries are the $3 \times 3$ matrices $X_{p,q}$. Let $B_{\tau,v}^*$ be the $2m \times m$ submatrix of $G(f.g.h)^*$ indexed by $(t^\sigma s^q, f^v \alpha^n), 0 \leq \sigma \leq 2m-1, 0 \leq u \leq m-1$. That is, $B_{\tau,v}^*$ contains the matrices $X_{\sigma-u, \tau+v+1}$ such that $0 \leq \sigma \leq 2m-1, 0 \leq u \leq m-1$. Using the entry formula (3.13), we readily observe the following properties.

- $G(f.g.h)^*$ is block upper triangular.

The blocks $B_{\tau,v}^*$ along the lower-left/top-right diagonal have $\tau + v = n - 1$. The blocks $B_{\tau,v}^*$ below this diagonal have $\tau + v + 1 > n$ and hence are zero by the entry formula (3.13).

- $G(f.g.h)^*$ is block symmetric and striped.

By the entry formula (3.13) it follows easily that $B_{\tau,v}^* = B_{v,\tau}^*$ and $B_{\tau+1,v-1}^* = B_{v-1,\tau+1}^*$. Hence $G(f.g,h)^*$ is block symmetric and $G_{\tau+1,v}^* = B_{\tau,v}^*$ is well-defined. Consequently,

$$G(f.g.h)^* = \begin{bmatrix}
  G_0^* & G_1^* & \cdots & G_{n-1}^* \\
  G_1^* & \ddots & \ddots & \ddots \\
  \vdots & \ddots & \ddots & \ddots \\
  G_{n-1}^*
\end{bmatrix}.$$  

(3.14)
The lower-left and upper-right corners of $G_i^\ast$ are zeros.

The $(m-1)m/2$ entries at the lower-left and the $(m-1)m/2$ entries at the upper-right corners of $G_i^\ast$ are indexed by $(\sigma, u)$ with $\sigma > u + m$ and $u > \sigma$ respectively; hence they are zero by the entry formula (3.13).

- The columns of $G_i^\ast$ march southeasterly.

By the entry formula (3.13) we have $G_i^\ast[\sigma, u] = G_i^\ast[\sigma + 1, u + 1]$.

- $G_i^\ast$ is Sylvester-like.

Combining the two previous properties, the block $G_i^\ast$ can be expressed as

$$G_i^\ast = \begin{bmatrix}
X_{0,0+1} & \cdots & \cdots & \cdots \\
\vdots & \ddots & & \\
X_{m-1,m+1} & \cdots & X_{0,0+1} \\
X_{m+1} & \cdots & X_{1,1-1} \\
\vdots & & \ddots & \\
X_{m,4+1}
\end{bmatrix}.$$  \hspace{1cm} (3.15)

**Example 3.2** ($m = 1$ and $n = 2$)

$$Mix(f, g, h)^\ast = \begin{bmatrix}
-R_{0,1;0,0} & -R_{0,2;0,0} \\
-R_{1,1;0,0} + R_{1,0;0,1} & -R_{1,2;0,0} + R_{1,0;0,1} \\
-R_{1,1;0,0} & -R_{1,2;1,0} \\
-R_{0,2;0,0} & -R_{0,2;0,1} \\
-R_{1,2;0,0} + R_{1,0;0,2} & -R_{1,2;0,1} + R_{1,1;0,2} \\
-R_{1,2;1,0} & -R_{1,2;1,1}
\end{bmatrix}.$$
\[
Syl(f, g, h)^* \cdot G(f, g, h)^* = \begin{bmatrix}
L_{0,0} & 0 & 0 & 0 \\
L_{1,0} & L_{0,0} & 0 & 0 \\
0 & L_{1,0} & 0 & 0 \\
L_{0,1} & 0 & L_{0,0} & 0 \\
L_{1,1} & L_{0,1} & L_{1,0} & L_{0,0} \\
0 & L_{1,1} & 0 & L_{1,0} \\
L_{0,2} & 0 & L_{0,1} & 0 \\
L_{1,2} & L_{0,2} & L_{1,1} & L_{0,1} \\
0 & L_{1,2} & 0 & L_{1,1} \\
0 & 0 & L_{0,2} & 0 \\
0 & 0 & L_{1,2} & L_{0,2} \\
0 & 0 & 0 & L_{1,2}
\end{bmatrix}
\begin{bmatrix}
X_{0,1} & X_{0,2} \\
X_{1,1} & X_{1,2} \\
X_{0,2} & 0 \\
X_{1,2} & 0
\end{bmatrix}
\]

= \begin{bmatrix}
-R_{0,1;0,0} & -R_{0,2;0,0} \\
-R_{1,1;0,0} + R_{1,0;0,1} & -R_{1,2;0,0} + R_{1,0;0,2} \\
-R_{1,1;1,0} & -R_{1,2;1,0} \\
-R_{0,2;0,0} & -R_{0,2;0,1} \\
-R_{1,2;0,0} + R_{1,0;0,2} & -R_{1,2;0,1} + R_{1,1;0,2} \\
-R_{1,2;1,0} & -R_{1,2;1,1} \\
0 & 0 \\
0 & 0 \\
0 & 0 \\
0 & 0 \\
0 & 0 \\
0 & 0
\end{bmatrix}
3.3 The Transformation from $Syl(f, g, h)$ to $Cay(f, g, h)$

In this section we revert to the default lexicographic orders $s > t$ and $\alpha > \beta$. We shall now derive a $6mn \times 2mn$ conversion matrix $F(f, g, h)$ such that

$$Syl(f, g, h) \cdot F(f, g, h) = \begin{bmatrix} Cay(f, g, h) \\ 0 \end{bmatrix}. \quad (3.16)$$

This result is very satisfying because it extends the following well-known expansion by cofactors property of $3 \times 3$ determinants

$$[i, j; k, l; p, q] = \begin{bmatrix} a_{i,j} & b_{i,j} & c_{i,j} \\ A_{k,l,p,q} & B_{k,l,p,q} & C_{k,l,p,q} \end{bmatrix} = L_{i,j} R_{k,l,p,q}^T.$$

The derivation below produces an explicit entry formula that reveals the simple elegant block structure of $F(f, g, h)$.

3.3.1 The Conversion Matrix $F(f, g, h)$

To find $F(f, g, h)$, we expand the numerator on the right hand side of Equation (3.2) by cofactors of the first row to obtain

$$\Delta_{m,n} = \begin{pmatrix} g(\alpha, t) & h(\alpha, t) \\ g(\alpha, \beta) & h(\alpha, \beta) \end{pmatrix} \left( f(s, t) \sum_{k=0}^{m} \sum_{l=0}^{n} \sum_{p=0}^{m} \sum_{q=0}^{n} \frac{A_{k,l,p,q} \alpha^{k+p} \beta^{l+q}}{(\beta - t)} + \cdots \right) \frac{1}{(\alpha - s)}. \quad (3.17)$$

Expanding the $2 \times 2$ numerator determinant yields

$$\Delta_{m,n} = \left( f(s, t) \sum_{k=0}^{m} \sum_{l=0}^{n} \sum_{p=0}^{m} \sum_{q=0}^{n} \frac{A_{k,l,p,q} \alpha^{k+p} \beta^{l+q}}{(\beta - t)} + \cdots \right) \frac{1}{(\alpha - s)}.$$

Applying the technique of truncated formal power series from Chapter 2, Section 2.2 by regarding the numerator as a polynomial in $\beta$ leads to

$$\Delta_{m,n} = \left( f(s, t) \sum_{k=0}^{m} \sum_{l=0}^{n} \sum_{p=0}^{m} \sum_{q=0}^{n} A_{k,l,p,q} \alpha^{k+p} \sum_{v=0}^{q-i} l^{i+q-i-v} j^v + \cdots \right) \frac{1}{(\alpha - s)}.$$
Rearranging the range of \( q \) and \( v \) yields
\[
\Delta_{m,n} = \left( f(s,t) \sum_{k=0}^{m} \sum_{l=0}^{n} \sum_{p=0}^{m} \sum_{q=0}^{m} A_{k,l,p,q \alpha^{k+p-l+q-1-v,j^v}} \right) \frac{1}{(\alpha - s)}.
\]
But the degree in \( t \) of the sum is at most \( n - 1 \): hence the summation range of \( q \) can be made more precise:
\[
\Delta_{m,n} = \left( f(s,t) \sum_{r=0}^{n-1} \sum_{k=0}^{m} \sum_{l=0}^{n} \sum_{p=0}^{m} \sum_{q=r+1}^{\min(n,n+r-l)} A_{k,l,p,q \alpha^{k+p-l+q-1-v,j^v}} \right) \frac{1}{(\alpha - s)}.
\]
Again applying the technique of truncated formal power series by regarding the numerator as a polynomial in \( \alpha \) leads to
\[
\Delta_{m,n} = \sum_{r=0}^{n-1} \sum_{k=0}^{m} \sum_{l=0}^{n} \sum_{p=0}^{m} \sum_{q=r+1}^{\min(n,n+r-l)} f(s,t) A_{k,l,p,q} \alpha^{k+p-l+q-1-v,j^v} \sum_{u=0}^{k+p-l-u} \alpha^u j^v + \ldots.
\]
Rearranging the range of \( p \) and \( u \) yields
\[
\Delta_{m,n} = \sum_{u=0}^{2m-1} \sum_{r=0}^{n-1} \sum_{k=0}^{m} \sum_{l=0}^{n} \sum_{p=\max(0,n+1-k)}^{m} \sum_{q=r+1}^{\min(n,n+r-l)} \alpha^{k+p-l-u+q-1-v} \alpha^u j^v + \ldots.
\]
Replacing the \( \ldots \) with the corresponding expressions for \( g(s,t) \) and \( h(s,t) \) produces
\[
\Delta_{m,n} = \sum_{u=0}^{2m-1} \sum_{r=0}^{n-1} \sum_{k=0}^{m} \sum_{l=0}^{n} \sum_{p=\max(0,n+1-k)}^{m} \sum_{q=r+1}^{\min(n,n+r-l)} (f A_{k,l,p,q} + g B_{k,l,p,q} + h C_{k,l,p,q}) \alpha^{k+p-l-u+q-1-v} \alpha^u j^v.
\]
Writing \( L = \begin{bmatrix} f & g & h \end{bmatrix} \) leads to
\[
\Delta_{m,n} = \sum_{u=0}^{2m-1} \sum_{r=0}^{n-1} \sum_{k=0}^{m} \sum_{l=0}^{n} \sum_{p=\max(0,n+1-k)}^{m} \sum_{q=r+1}^{\min(n,n+r-l)} \alpha^{k+p-l-u+q-1-v} \alpha^u j^v L^T R_{k,l,p,q} \alpha^u j^v.
\]
(3.18)

Since \( k \leq m, u + 1 - k \leq p \leq m, u \geq 0, v + 1 \leq q \leq n + v - l \), we have
\[0 \leq k + p - 1 - u \leq 2m - 1 \text{ and } 0 \leq l + q - 1 - v \leq n - 1. \]
Hence there is a \( 6mn \times 2mn \)
matrix \( F(f, g, h) \) such that \( \Delta_{m,n} \) can be written in matrix form as

\[
\Delta_{m,n} = \begin{bmatrix}
LT \\
\vdots \\
s^\sigma f^\tau LT \\
s^\sigma f^\tau+1 LT \\
\vdots \\
s^{2m-1}f^nLT
\end{bmatrix}^T \begin{bmatrix}
1 \\
\vdots \\
\alpha^n f^\nu \\
\alpha^n f^\nu+1 \\
\vdots \\
\alpha^{2m-1} f^{n-1}
\end{bmatrix}. \quad (3.19)
\]

It follows by Equation (3.19) and the definition of \( Syl(f, g, h) \) (Equation (3.1)) that

\[
\Delta_{m,n} = \begin{bmatrix}
1 \\
\vdots \\
s^\sigma f^\tau \\
s^\sigma f^\tau+1 \\
\vdots \\
s^{3m-1}f^{2n-1}
\end{bmatrix}^T Syl(f, g, h) \cdot F(f, g, h) \begin{bmatrix}
1 \\
\vdots \\
\alpha^n f^\nu \\
\alpha^n f^\nu+1 \\
\vdots \\
\alpha^{2m-1} f^{n-1}
\end{bmatrix}. \quad (3.20)
\]

Thus by the definition of \( Cay(f, g, h) \) (Equation (3.3)), we see that

\[
Syl(f, g, h) \cdot F(f, g, h) = \begin{bmatrix}
Cay(f, g, h) \\
0
\end{bmatrix}. \quad (3.21)
\]

Notice that 4mn rows of zeros must be appended after the rows of \( Cay(f, g, h) \) because the row indices \( s^\sigma \) of \( Syl(f, g, h) \) run from 0 to \( 3m - 1 \) but the row indices \( s^\tau \) of \( Cay(f, g, h) \) run from 0 to \( m - 1 \).

### 3.3.2 The Entries of \( F(f, g, h) \)

Solving \( k + p - 1 - u = \sigma \) and \( l + q - 1 - v = \tau \) in Equation (3.18) gives

\[
p = \sigma + 1 + u - k.
\]

\[
q = \tau + 1 + v - l.
\]
The constraints

\[ 0 \leq k \leq m. \]
\[ 0 \leq l \leq n. \]
\[ \max(0, u + 1 - k) \leq p \leq m. \]
\[ r + 1 \leq q \leq \min(n, n + v - l). \]

lead to

\[ \max(0, \sigma + 1 + u - m) \leq k \leq \min(m, \sigma + 1 + u). \]
\[ \max(0, \tau + 1 + v - n) \leq l \leq \tau. \]

Therefore the entry of the cofactor matrix \( F(f, g, h) \) indexed by \( (s^\sigma l^\tau, \alpha^n f^v) \) is the sum

\[ \sum_{k=\max(0, \sigma + 1 + u - m)}^{\min(m, \sigma + 1 + u)} \sum_{l=\max(0, \tau + 1 + v - n)}^{\tau} R_{k,l,\sigma+1+u-k,\tau+1+v-l}. \]  

(3.22)

for \( 0 \leq \sigma, u \leq 2m - 1, \ 0 \leq \tau, v \leq n - 1 \). But observe that the above sum contains mutually canceling cofactors of the form \( R_{k,l,p,q}^T + R_{p,q;k,l}^T = 0 \) because \( A_{k,l,p,q} + A_{p,q;k,l} = 0 \), \( B_{k,l,p,q} + B_{p,q;k,l} = 0 \), \( C_{k,l,p,q} + C_{p,q;k,l} = 0 \). To find the canceling partner of \( R_{k,l,\sigma+1+u-k,\tau+1+v-l}^T \) let

\[ k = \sigma + 1 + u - k'. \]
\[ l = \tau + 1 + v - l'. \]
\[ \sigma + 1 + u - k = k'. \]
\[ \tau + 1 + v - l = l'. \]

A quick check reveals that when \( v + 1 \leq l \), we have

\[ \max(0, \sigma + 1 + u - m) \leq k' \leq \min(m, \sigma + 1 + u), \]
\[ \max(0, \tau + 1 + v - n) \leq l' \leq \tau. \]
That is, the range of $l$ need not run beyond $r$ because terms out of that range mutually cancel. Consequently, the sum for the entry formula (3.22) of $F(f, g, h)$ will be free of mutually canceling terms $R_{k,l;p,q}^r$ if it is written as

$$
\sum_{k=\max(0,\sigma+1+u-m)}^{\min(m,\sigma+1+u)} \sum_{l=\max(0,\tau+1+v-n)}^{\min(\tau,v)} R_{k,l;\sigma+1+u-k,\tau+1+v-l}^r,
$$

(3.23)

for $0 \leq \sigma, u \leq 2m - 1, 0 \leq \tau, v \leq n - 1$.

### 3.3.3 Properties of $F(f, g, h)$

The structure of $F(f, g, h)$ is most revealing if we view this $6mn \times 2mn$ matrix as a square $2mn \times 2mn$ matrix by considering its entries as sums of the $3 \times 1$ matrices $R_{k,l;p,q}^r$.

- $F(f, g, h)$ is symmetric.

  From the entry formula (3.23), we see that the entries indexed by $(s^\sigma \ell^\tau, \alpha^u \beta^v)$ and $(s^\tau \ell^\sigma, \alpha^u \beta^v)$ are the same. In other words, the matrix $F(f, g, h)$ is symmetric.

- $F(f, g, h)$ is block upper triangular.

  Let the $n \times n$ submatrix of $F(f, g, h)$ indexed by $(s^\sigma \ell^\tau, \alpha^u \beta^v), 0 \leq \tau, v \leq n - 1$, be $B_{\sigma,u}$. When $\sigma + u > 2m - 1$ in the entry formula (3.23), we have

$$
\sum_{k=\max(0,\sigma+u+1-m)}^{\min(m,\sigma+u+1)} = \sum_{k=\sigma+u+1-m}^{m},
$$

(3.24)

and clearly this summation range is vacuous. Thus the sum is zero. Consequently the conversion matrix $F(f, g, h)$ is block upper triangular in the sense that the blocks $B_{\sigma,u}$ with $\sigma + u > 2m - 1$ are all zero.

- $F(f, g, h)$ is block symmetric and striped.

  From the entry formula (3.23), we easily see that

$$
B_{\sigma-1,u+1} = B_{\sigma,u} = B_{\sigma+1,u-1}.
$$
Therefore, the notation $F_{\sigma,u} = B_{\sigma,u}$ is well-defined, and

$$F(f, g, h) = \begin{bmatrix}
F_0 & F_1 & \cdots & F_{2m-1} \\
F_1 & \ddots & & \\
\vdots & & \ddots & \\
F_{2m-1} & & & \\
\end{bmatrix}.$$  \hspace{1cm} (3.25)

- $F_k$'s are symmetric.

Notice that $\tau$ and $\upsilon$ are symmetric in formula (3.23); thus

$$B_{\sigma,u} = B_{\sigma,u}^T.$$  \hspace{1cm} (3.26)

That is, each of the blocks $F_k$ is symmetric.

- Three levels of symmetry.

In summary, $F(f, g, h)$ has the following symmetry properties:

- $F(f, g, h)$ is symmetric if we view $F(f, g, h)$ as a $2mn \times 2mn$ square matrix, whose entries are $3 \times 1$ submatrices, i.e. sums of the $R_{k,l,p,q}^T$;

- $F(f, g, h)$ is also symmetric if we view $F(f, g, h)$ as a $2m \times 2m$ square matrix, whose entries are $3n \times n$ matrices, i.e. the $F_k$'s;

- the submatrices $F_k$ are symmetric if we view each $F_k$ as an $n \times n$ matrix, whose entries are $3 \times 1$ submatrices, i.e. sums of the $R_{k,l,p,q}^T$.

- $F_k$'s are Bézoutian.

Let

$$f_i(t) = \sum_{j=0}^{n} a_{i,j} t^j, \quad g_i(t) = \sum_{j=0}^{n} b_{i,j} t^j, \quad h_i(t) = \sum_{j=0}^{n} c_{i,j} t^j.$$  \hspace{1cm} (3.27)

Using the technique of truncated formal power series from Section 2.2, we can write Equation (3.17) as
\[
\Delta_{m,n}(s, t, \alpha, \beta) = \left[ f(s, t) \left( \sum_{i=0}^{m} g_i(t) \alpha^i \sum_{j=0}^{m} h_j(\beta) \alpha^j - \sum_{i=0}^{m} g_i(\beta) \alpha^i \sum_{j=0}^{m} h_j(t) \alpha^j \right) \right] \frac{1}{\beta - t} + \cdots \\
= \sum_{i=0}^{m} \sum_{j=0}^{m} \sum_{u=0}^{i+j-1} \left( s^{i+j-1-u} \alpha^u f(s, t) \frac{g_i(t) h_j(\beta) - g_i(\beta) h_j(t)}{\beta - t} \right) + \cdots 
\]

Therefore, the entries of the submatrix \( B_{\sigma,u} (= F_{\sigma+u}) \) come from the coefficients of

\[
\sum_{i+j-1=\sigma+u} \frac{g_i(t) h_j(\beta) - g_i(\beta) h_j(t)}{\beta - t} \\
\sum_{i+j-1=\sigma+u} \frac{h_i(t) f_j(\beta) - h_i(\beta) f_j(t)}{\beta - t} \quad (3.26) \\
\sum_{i+j-1=\sigma+u} \frac{f_i(t) g_j(\beta) - f_i(\beta) g_j(t)}{\beta - t}
\]

Each term in these three sums is a Cayley expression for two univariate polynomials \( \{g_i, h_j\}, \{h_i, f_j\}, \{f_i, g_j\} \): hence each term generates a Bézout matrix when written in matrix form \([15] [32] [46]\) — see too Chapter 2, Section 2.1. Therefore, each block \( F_{\sigma+u} = B_{\sigma,u} \) contains three summations of Bézout matrices, where the three matrices interleave row by row. In particular, if \( \sigma + u = 2m - 1 \), then \( i = j = m \) in expression (3.26). so \( F_{2m-1} \) is the matrix obtained by interleaving the rows of the three Bézout matrices: \( \text{Bez}(g_m, h_m) \), \( \text{Bez}(h_m, f_m) \), \( \text{Bez}(f_m, g_m) \).

Explicit formulas for the entries of the Bézout matrix of two univariate polynomials are given in [32]. More efficient methods for computing the entries of a Bézout resultant or the entries of a sum of Bézout resultants are described in [Ch 2] [15]. Since the blocks \( F_k \) are Bézoutian, we can adopt those methods.
here to provide an efficient algorithm for computing the entries of $F_k$. Fast computation of the Bézoutiants $F_k$ is important because, as we shall see in Section 3.7, we can use the blocks $F_k$ to speed up the computation of the entries of the Cayley matrix $Cay(f, g, h)$.

**Example 3.3 (n = 1 and n = 2)**

\[
Cay(f, g, h) = \begin{bmatrix}
1.0:0.1:0.0 & 1.0:0.2:0.0 \\
1.0:0.2:0.0 & 1.1:0.1:0.0 & 1.1:0.2:0.0 \\
1.2:0.1:0.0 & 1.1:0.2:0.0 & 1.1:0.2:0.0 & 1.2:0.2:0.0 \\
-1.1:1.0:0.0 & -1.2:1.0:0.0 \\
-1.2:1.0:0.0 & -1.1:1.0:0.1 & -1.2:1.0:0.1 & -1.2:1.1:0.0 \\
-1.2:1.0:0.1 & -1.1:1.0:0.2 & -1.2:1.1:0.1 & -1.2:1.0:0.2 \\
-1.2:1.0:0.2 & -1.2:1.1:0.2
\end{bmatrix},
\]

\[
Syl(f, g, h) \cdot F(f, g, h) = \begin{bmatrix}
L_{0,0} & 0 & 0 & 0 \\
L_{0,1} & L_{0,0} & 0 & 0 \\
L_{0,2} & L_{0,1} & 0 & 0 \\
0 & L_{0,2} & 0 & 0 \\
L_{1,0} & 0 & L_{0,0} & 0 \\
L_{1,1} & L_{1,0} & L_{0,1} & L_{0,0} \\
L_{1,2} & L_{1,1} & L_{0,2} & L_{0,1} \\
0 & L_{1,2} & 0 & L_{0,2} \\
0 & 0 & L_{1,0} & 0 \\
0 & 0 & L_{1,1} & L_{1,0} \\
0 & 0 & L_{1,2} & L_{1,1} \\
0 & 0 & 0 & L_{1,2}
\end{bmatrix}
\begin{bmatrix}
-R_{1,1,0,0} + R_{1,0,0,1} & -R_{1,2,0,0} + R_{1,0,0,2} & -R_{1,1,1,0} & -R_{1,2,1,0} \\
-R_{1,2,0,0} + R_{1,0,0,2} & -R_{1,2,0,1} + R_{1,1,0,2} & -R_{1,2,1,0} & -R_{1,2,1,1} \\
-R_{1,1,1,0} & -R_{1,2,1,0} & 0 & 0 \\
-R_{1,2,1,0} & -R_{1,2,1,1} & 0 & 0
\end{bmatrix}.
\]
Now it is straightforward to check that indeed
\[
Syl(f, g, h) \cdot F(f, g, h) = \begin{bmatrix}
Cay(f, g, h) \\
0
\end{bmatrix}.
\]

### 3.4 The Transformation from $Mix(f, g, h)$ to $Cay(f, g, h)$

We have shown that the polynomials represented by the columns of the Cayley resultant and the mixed Cayley-Sylvester resultant are linear combinations of the polynomials represented by the columns of the Sylvester resultant. There is, however, no way to represent the polynomials represented by the columns of the Cayley resultant as linear combinations of the polynomials represented by the columns of the mixed Cayley-Sylvester resultant because the Cayley columns represent polynomials of bidegree $(m - 1, 2n - 1)$ whereas the Cayley-Sylvester columns represent polynomials of bidegree $(3m - 1, n - 1)$. Thus while the Cayley matrix is smaller than the Cayley-Sylvester matrix, the polynomials in the column space of the Cayley matrix do not form a subspace of the polynomials in the column space of the Cayley-Sylvester matrix. On the other hand, by Equation (3.3), the rows of the Cayley matrix represent polynomials of bidegree $(2m - 1, n - 1)$. Therefore since the columns of the Cayley-Sylvester matrix are linearly independent, the polynomials represented by the rows of the Cayley matrix can be expressed as linear combinations of the polynomials represented by the columns of the Cayley-Sylvester matrix. Below we will derive the conversion matrix $E(f, g, h)$ that transforms $Mix(f, g, h)$ to $Cay(f, g, h)^T$.

#### 3.4.1 The Conversion Matrix $E(f, g, h)$

To find $E(f, g, h)$, we expand the determinant on the right hand side of Equation (3.2) with respect to the first row. Recall from Section 3.1.3 that
\[
\begin{vmatrix}
g(\alpha, t) & h(\alpha, t) \\
g(\alpha, \beta) & h(\alpha, \beta)
\end{vmatrix}
/ (\beta - t) = \begin{vmatrix}
g(\alpha, \beta) & h(\alpha, \beta) \\
g(\alpha, t) & h(\alpha, t)
\end{vmatrix}
/ (t - \beta) = \sum_{e=0}^{n-1} f_e(\alpha, \beta) t^e.
\]
Therefore,

\[
\Delta_{m,n}(s, t, \alpha, \beta) = - \left( f(s, t) \sum_{n=0}^{n-1} \bar{f}_n(\alpha, \beta) t^n + g(s, t) \sum_{n=0}^{n-1} \bar{g}_n(\alpha, \beta) t^n + h(s, t) \sum_{n=0}^{n-1} \bar{h}_n(\alpha, \beta) t^n \right) \frac{1}{s-\beta}.
\]  

(3.27)

Using the notation of Section 3.1.3, write \( \bar{L}_n = [\bar{f}_n \ \bar{g}_n \ \bar{h}_n] \). Since the numerator is divisible by the denominator in Equation (3.27), we can apply the technique of truncated formal power series — Chapter 2, Section 2.2 — by treating the numerator as a polynomial in \( s \) to write \( \Delta_{m,n}(s, t, \alpha, \beta) \) as

\[
\Delta_{m,n}(s, t, \alpha, \beta) = - \sum_{i=0}^{m} \sum_{j=0}^{n} \sum_{\sigma=0}^{i-1} \sum_{v=0}^{n-1} \bar{L}_v \cdot \begin{bmatrix} a_{i,j} \\ h_{i,j} \\ c_{i,j} \end{bmatrix} \alpha^{i-1-\sigma} s^\sigma t^{j+v}
\]

\[
= \begin{bmatrix} \bar{L}_0 \\ \vdots \\ \alpha^u \bar{L}_v \\ \alpha^u \bar{L}_{v+1} \\ \vdots \\ \alpha^{m-1-1} \bar{L}_{n-1} \end{bmatrix} \begin{bmatrix} 1 \\ \vdots \\ E(f, g, h) \\ s^\sigma t^\tau \\ s^\sigma t^{\tau+1} \\ \vdots \\ s^{m-1} t^{2n-1} \end{bmatrix}.
\]  

(3.28)

for some coefficient matrix \( E(f, g, h) \). Hence by the definition of \( Mix(f, g, h) \),

\[
\Delta_{m,n}(s, t, \alpha, \beta) = \begin{bmatrix} 1 \\ \vdots \\ \alpha^u t^v \\ \alpha^u t^{v+1} \\ \vdots \\ \alpha^{m-1-n} t^{n-1} \end{bmatrix} \begin{bmatrix} 1 \\ \vdots \\ Mix(f, g, h) \cdot E(f, g, h) \\ s^\sigma t^\tau \\ s^\sigma t^{\tau+1} \\ \vdots \\ s^{m-1} t^{2n-1} \end{bmatrix}.
\]
On the other hand, by Equation (3.3),

\[
\Delta_{m,n}(s, t, \alpha, \beta) = \begin{bmatrix}
1 \\
\vdots \\
\alpha^{n \cdot \beta} \\
\alpha^{n \cdot \beta + 1} \\
\vdots \\
\alpha^{2m-1 \cdot \beta n-1}
\end{bmatrix}^T \cdot C^a y(f, g, h) \cdot \begin{bmatrix}
1 \\
\vdots \\
\alpha^{n \cdot \beta} \\
\alpha^{n \cdot \beta + 1} \\
\vdots \\
\alpha^{2m-1 \cdot \beta n-1}
\end{bmatrix} = \begin{bmatrix}
1 \\
\vdots \\
\alpha^{n \cdot \beta} \\
\alpha^{n \cdot \beta + 1} \\
\vdots \\
\alpha^{2m-1 \cdot \beta n-1}
\end{bmatrix}^T \cdot C^a y(f, g, h)^T \cdot \begin{bmatrix}
1 \\
\vdots \\
\alpha^{n \cdot \beta} \\
\alpha^{n \cdot \beta + 1} \\
\vdots \\
\alpha^{2m-1 \cdot \beta n-1}
\end{bmatrix}
\]

Therefore,

\[
Mix(f, g, h) \cdot E(f, g, h) = \begin{bmatrix}
C^a y(f, g, h)^T \\
0_{mn \times 2mn}
\end{bmatrix}.
\]  

(3.29)

### 3.4.2 The Entries of \(E(f, g, h)\)

From Equation (3.28), the \(3 \times 1\) submatrix of \(E(f, g, h)\) indexed by \((\alpha^{i-1 - \sigma}, \sum_{i \tau + v})\) is \(-L_{i,j}^T\). By solving

\[
u = i - 1 - \sigma.
\]

\[
\tau = j + v.
\]
in Equation (3.28), we get

\[ i = \sigma + u + 1, \]
\[ j = \tau - v. \]

Therefore the $3 \times 1$ submatrix of $E(f, g, h)$ indexed by $(e^w\bar{L}_n, s^\sigma t^\tau)$ is

\[ -L_{\sigma+u+1,\tau-v}. \]  

when $\sigma + u + 1 \leq m$, $0 \leq \tau - v \leq u$, and zero otherwise.

### 3.4.3 Properties of $E(f, g, h)$

From the above entry formula (3.30), arguments similar to those of Section 3.2.3 or 3.3.3 show that $E(f, g, h)$ has the following striped block symmetric structure:

\[
E(f, g, h) = \begin{bmatrix}
E_0 & E_1 & \cdots & E_{m-1} \\
E_1 & \ddots & \ddots & \\
\vdots & \ddots & \ddots & \\
E_{m-1} & & & \\
\end{bmatrix}.
\]

where

\[
E_i =
\begin{bmatrix}
-L_{t+1,0}^T & -L_{t+1,1}^T & \cdots & -L_{t+1,n-1}^T & -L_{t+1,n}^T & 0 & \cdots & 0 \\
0 & -L_{t+1,0}^T & \cdots & -L_{t+1,n-2}^T & -L_{t+1,n-1}^T & -L_{t+1,n}^T & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & -L_{t+1,0}^T & -L_{t+1,1}^T & -L_{t+1,2}^T & \cdots & -L_{t+1,n}^T \\
\end{bmatrix}.
\]  

(3.31)

is of size $n \times 2n$ with each entry a $3 \times 1$ submatrix.
Example 3.4 (m = 1 and n = 2)

\[ Mix(f, g, h) \cdot E(f, g, h) \]

\[
= \begin{bmatrix}
- R_{0,1,0,0} & - R_{0,2,0,0} \\
- R_{0,2,0,0} & - R_{0,2,0,1} \\
- R_{1,1,0,0} + R_{1,0,0,1} & - R_{1,2,0,0} + R_{1,0,0,2} \\
- R_{1,2,0,0} + R_{1,0,2,0} & R_{1,1,0,2} - R_{1,2,0,1} \\
- R_{1,1,1,0} & - R_{1,2,1,0} \\
- R_{1,2,1,0} & - R_{1,2,1,1}
\end{bmatrix}
\begin{bmatrix}
-L^T_{1,0} & -L^T_{1,1} & -L^T_{1,2} & 0 \\
0 & -L^T_{1,0} & -L^T_{1,1} & -L^T_{1,2}
\end{bmatrix}
\]

\[ = \begin{bmatrix}
Cayl(f, g, h)^T \\
0
\end{bmatrix}.\]

3.5 The Block Structure of the Three Dixon Resultants

The natural block structures of the transformation matrices prompt us to reexamine once again the three Dixon resultant matrices to seek natural block structures compatible with the block structures of the transformation matrices.

3.5.1 The Block Structure of Syl(f, g, h)

We can impose a natural block structure on the entries of Syl(f, g, h). Let

\[ f_i(t) = \sum_{j=0}^{n} a_{i,j} t^j, \quad g_i(t) = \sum_{j=0}^{n} b_{i,j} t^j, \quad h_i(t) = \sum_{j=0}^{n} c_{i,j} t^j. \]

and let \( S_i \) be the \( 2n \times 3n \) coefficient matrix for the polynomials \( (t^r f_i, t^r g_i, t^r h_i) \), \( r = 0, \ldots, n - 1 \). Then

\[ S_i = \begin{bmatrix}
 a_{i,0} & b_{i,0} & c_{i,0} \\
 \vdots & \vdots & \vdots \\
 a_{i,n-1} & b_{i,n-1} & c_{i,n-1} \\
 a_{i,n} & b_{i,n} & c_{i,n} \\
 \vdots & \vdots & \vdots \\
 a_{i,1} & b_{i,1} & c_{i,1} \\
 \vdots & \vdots & \vdots \\
 c_{i,n} & b_{i,n} & c_{i,n}
\end{bmatrix}. \quad (3.32)\]
Here the rows are indexed by the monomials \(1, \ldots, t^{2n-1}\), and the columns are indexed by the polynomials \(t^0(f_i, g_i, h_i), \ldots, t^{n-1}(f_i, g_i, h_i)\). Note that the matrix \(S_i\) is Sylvester-like in the sense that if we drop the \(f_i\)-columns from \(S_i\), then we get the univariate Sylvester matrix of \(g_i\) and \(h_i\); dropping the \(g_i\)-columns yields the univariate Sylvester matrix of \(h_i\) and \(f_i\); dropping the \(h_i\)-columns yields the univariate Sylvester matrix of \(f_i\) and \(g_i\).

It follows from Equations (3.1) and (3.32) that

\[
S_{yl}(f, g, h) = \begin{bmatrix}
S_0 \\
S_1 \\
S_2 \\
\vdots \\
S_{m-1} & \cdots & S_0 \\
S_{m-2} & \cdots & S_1 \\
S_{m-3} & \cdots & S_2 \\
\vdots & \ddots & \vdots \\
S_m & \cdots & S_{m-1} \\
S_m & \cdots & S_0
\end{bmatrix}.
\]  

(3.33)

Notice from Equations (3.31) and (3.32) that

\[
E_i = -(S_{i+1})^T, \quad 0 \leq i \leq m - 1.
\]  

(3.34)

Similar results hold when we impose the order \(t > s\). Define

\[
f_j^*(s) = \sum_{i=0}^{m} a_{i,j} s^i, \quad g_j^*(s) = \sum_{i=0}^{m} b_{i,j} s^i, \quad h_j^*(s) = \sum_{i=0}^{m} c_{i,j} s^i.
\]

Now let \(S_j^*\) be the \(3m \times 6m\) coefficient matrix for the polynomials \((s^n f_j^*, s^n g_j^*, s^n h_j^*)\).
\( v = 0, \ldots, 2m - 1 \). Then \( S_j^* \) can be written as

\[
\begin{bmatrix}
L_{0,j} & \cdots & \cdots & \cdots \\
\vdots & \ddots & \vdots & \vdots \\
L_{m-1,j} & L_{0,j} & \cdots & \cdots \\
L_{m,j} & \cdots & L_{1,j} & L_{0,j} \\
& \ddots & \vdots & \vdots \\
L_{m,j} & L_{m-1,j} & \cdots & L_{0,j} \\
L_{m,j} & \cdots & L_{1,j} & \cdots \\
& \cdots & \ddots & \vdots \\
& & & \ddots \\
& & & & \ddots \\
L_{m,j} & & & & \\
\end{bmatrix}
\]

and it follows from Equation (3.7) that

\[
S_{yl}(f, g, h)^* = \begin{bmatrix}
S_0^* \\
\vdots \\
S_{n-1}^* \\
S_n^* \\
S_{n-1}^* \\
\vdots \\
S_1^* \\
S_n^* 
\end{bmatrix} \quad (3.35)
\]

### 3.5.2 The Block Structure of Mix(f, g, h)

Recall from Section 3.1.3 that \( \mathbf{L}_v = [\mathbf{f}_v \quad \mathbf{g}_v \quad \mathbf{h}_v] \), \( v = 0, \ldots, n - 1 \), where \( \mathbf{f}_v, \mathbf{g}_v, \mathbf{h}_v \) are polynomials of degree \( 2m \) in \( s \) and degree \( n - 1 \) in \( t \). Let \( M_i : i = 0, \ldots, 3m - 1 \) be the \( n \times 3n \) coefficient matrix of the monomials \( \{s^i, \ldots, s^i t^{n-1}\} \) of the polynomials \( \mathbf{L}_0, \ldots, \mathbf{L}_{n-1} \). Note that \( M_i = 0 \) when \( i > 2m \) because every \( \mathbf{L}_v \) is of degree \( 2m \) in \( s \).

Since the columns of \( \text{Mix}(f, g, h) \) are indexed by \( s^i \mathbf{L}_0, \ldots, s^i \mathbf{L}_{n-1} \), \( i = 0, \ldots, m - 1 \), and the rows by \( s^i, \ldots, s^i t^{n-1} \), \( i = 0, \ldots, 3m - 1 \), it is easy to see from Equation (3.5)
that $Mix(f, g, h)$ has the following Sylvester-like (shifted block) structure:

$$
Mix(f, g, h) = \begin{bmatrix}
M_0 \\
\vdots \\
M_{m-1} & \cdots & M_0 \\
\vdots & \vdots & \vdots \\
M_{2m} & \cdots & M_{m+1} \\
\vdots & \vdots \\
M_{2m}
\end{bmatrix}.
$$

Write $\phi(g, h), \phi(h, f), \phi(f, g)$ as

$$
\phi(g, h) = \sum_{i=0}^{m} \sum_{j=0}^{m} \frac{g_i(t)h_j(t) - g_j(t)h_i(t)}{t - i - j} s^{i+j},
$$

$$
\phi(h, f) = \sum_{i=0}^{m} \sum_{j=0}^{m} \frac{h_i(t)f_j(t) - h_j(t)f_i(t)}{t - i - j} s^{i+j},
$$

$$
\phi(f, g) = \sum_{i=0}^{m} \sum_{j=0}^{m} \frac{f_i(t)g_j(t) - f_j(t)g_i(t)}{t - i - j} s^{i+j}.
$$

Now notice that the block $M_k$ consists of the coefficients of $s^k$ in these expressions. That is, $M_k$ contains three matrices with their columns interleaved. Just like the expressions in Equation (3.26), the coefficients of $s^k$ are sums of Cayley expressions. Hence each of these three matrices is the sum of the Bézout matrices of the polynomials $\{g_i, h_j\}, \{h_i, f_j\}, \{f_i, g_j\}$, for $i + j = k$. In particular, $M_0$ contains the three interleaved Bézout matrices: $Bz(g_0, h_0), Bz(h_0, f_0), Bz(f_0, g_0)$; similarly, $M_{2m}$ contains three interleaved Bézout matrices: $Bz(g_m, h_m), Bz(h_m, f_m), Bz(f_m, g_m)$.

Notice that $M_k$ and $F_{k-1}$ [Section 3.3.3], $k = 1, \ldots, 2m$, contain the same three Bézoutian matrices; however, $M_k$ is the $n \times 3n$ matrix where the three Bézoutian matrices interleave column by column, whereas $F_{k-1}$ is the $3n \times n$ matrix where the three Bézoutian matrices interleave row by row. Since each Bézout matrix is symmetric [32] [Chapter 2. Section 2.1], it is easy to see that

$$
F_{k-1} = (M_k)^T, \quad 1 \leq k \leq 2m.
$$

(3.36)
The matrix $Mix(f, g, h)^*$ has a different block structure. Recall from Equation (3.35) and Equation (3.14) that $Syl(f, g, h)^*$ and $G(f, g, h)^*$ can be written as

$$Syl(f, g, h)^* = \begin{bmatrix} S_0^* & \cdots & \cdots & \cdots \\ \vdots & \ddots & \ddots & \ddots \\ \vdots & \ddots & S_0^* & \cdots \\ S_n^* & \cdots & \cdots & S_1^* \\ \vdots & \cdots & \cdots & \vdots \\ S_n^* & \cdots & \cdots & S_n^* \end{bmatrix}, \quad G(f, g, h)^* = \begin{bmatrix} G_0^* & \cdots & \cdots & G_{n-1} \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ G_{n-1} \end{bmatrix},$$

where each $S_i^*$ is of size $3m \times 6m$ and each $G_j^*$ is of size $6m \times 3m$. But by Equation (3.12),

$$Syl(f, g, h)^* \cdot G(f, g, h)^* = \begin{bmatrix} Mix(f, g, h)^* \\ 0 \end{bmatrix},$$

so

$$Mix(f, g, h)^* = \begin{bmatrix} S_0^* & \cdots & \cdots & \cdots \\ \vdots & \ddots & \ddots & \ddots \\ \vdots & \ddots & S_0^* & \cdots \\ S_n^* & \cdots & \cdots & S_n^* \end{bmatrix} \cdot \begin{bmatrix} G_0^* & \cdots & \cdots & G_{n-1} \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ G_{n-1} \end{bmatrix}.$$ (3.37)

Write

$$Mix(f, g, h)^* = \begin{bmatrix} K_{0,0}^* & \cdots & K_{0,n-1}^* \\ \vdots & \ddots & \vdots \\ K_{n-1,0}^* & \cdots & K_{n-1,n-1}^* \end{bmatrix},$$

where each block $K_{i,j}^*$ is of size $3m \times 3m$. By construction, each column of blocks

$$\begin{bmatrix} K_{0,j}^* \\ \vdots \\ K_{n-1,j}^* \end{bmatrix}$$

consists of the coefficients of the polynomials $[1, \cdots, s^{m-1}] \cdot \overline{L}_j, j = 0, \cdots, n-1$. Now by Equation (3.37),

$$K_{i,j}^* = \sum_{l=0}^{\min(i, n-1-j)} S_{i-l}^* \cdot G_{j+l}^*, \quad 0 \leq i, j \leq n - 1. \quad (3.38)$$
It follows from Equation (3.38) that
\[ K_{i,j}^* = S_i^* \cdot G_j^* + K_{i-1,j+1}^*. \] (3.39)

### 3.5.3 The Block Structure of \( \text{Cay}(f,g,h) \)

As for the block structure of \( \text{Cay}(f,g,h) \), recalling that
\[ \Delta_{m,n}(s,t,\alpha,\beta) \begin{bmatrix} 1 \\ \vdots \\ s^t I^r \\ s^t I^{r+1} \\ \vdots \\ s^{m-1} I^{2n-1} \end{bmatrix} \begin{bmatrix} 1 \\ \vdots \\ \alpha^u J^v \\ \alpha^u J^{v+1} \\ \vdots \\ \alpha^u J^{n-1} \end{bmatrix}. \] (3.40)

we simply group the entries with respect to the indices \( s^t, u^a \) and write
\[ \text{Cay}(f,g,h) = \begin{bmatrix} C_{0,0} & \cdots & C_{0,2m-1} \\ \vdots & \vdots & \vdots \\ C_{m-1,0} & \cdots & C_{m-1,2m-1} \end{bmatrix}. \] (3.41)

where each block \( C_{i,j} \) is of size \( 2n \times n \). The reason why we impose this particular block structure on the entries of \( \text{Cay}(f,g,h) \) will become clear shortly in Section 3.7.

### 3.6 Convolution Identities

In this section, we are going to derive the following convolution identities:
\[ \sum_{u+v=i} S_u \cdot M_v^T = 0, \quad 0 \leq i \leq 3m. \] (3.42)

which relate the blocks of the Sylvester resultant matrices and the Cayley-Sylvester mixed resultant matrices. We shall proceed in the following manner. Note that
\[ f(s,t) \cdot \phi(g,h) + g(s,t) \cdot \phi(h,f) + h(s,t) \cdot \phi(f,g) \]

\[
\begin{align*}
\begin{vmatrix}
  f(s,t) & g(s,t) & h(s,t) \\
  g(s,\beta) & h(s,\beta) & f(s,\beta) \\
  f(s,\beta) & g(s,\beta) & h(s,\beta)
\end{vmatrix}
+ g(s,t) & h(s,t) & f(s,t) \\
  h(s,\beta) & f(s,\beta) & g(s,\beta) \\
  f(s,\beta) & g(s,\beta) & h(s,\beta)
\end{vmatrix}
+ h(s,t) & f(s,t) & g(s,t) \\
  f(s,\beta) & g(s,\beta) & h(s,\beta)
\end{vmatrix}
\]

\[ \beta - t \]

\[ \begin{vmatrix}
  f(s,t) & g(s,t) & h(s,t) \\
  g(s,\beta) & h(s,\beta) & f(s,\beta) \\
  f(s,\beta) & g(s,\beta) & h(s,\beta)
\end{vmatrix}
\]

\[ \equiv 0. \quad (3.43) \]

To prove the convolution identities (3.42), we will interpret the columns of the matrices on the left hand side of Equation (3.42) as the coefficients of certain monomials on the left hand side of Equation (3.43). But first we need some preliminary observations.

### 3.6.1 Interleaving in \( S_n \) and \( M_n \)

For convenience, we adopt the following notation: Let \( p_i(t) \), \( 1 \leq i \leq k \), be polynomials of degree \( d_i \), and let \( d = \max(d_1, d_2, \cdots, d_k) \). We shall write

\[
\begin{bmatrix}
  p_1(t), & p_2(t), & \cdots, & p_k(t)
\end{bmatrix}^C
\]

to denote the \((d+1) \times k\) coefficient matrix of the polynomials \( p_i(t) \) whose rows are indexed by \( 1, \ldots, t^d \). That is,

\[
\begin{bmatrix}
  p_1(t), & \cdots, & p_k(t)
\end{bmatrix} = (1, \ldots, t^d) \cdot \begin{bmatrix}
  p_1(t), & p_2(t), & \cdots, & p_k(t)
\end{bmatrix}^C
\]

For example,

\[
\begin{bmatrix}
  (1 + 2t), & (1 + 2t)t^3, & (4 + 5t + 6t^2)
\end{bmatrix}^C = \begin{bmatrix}
  1 & 0 & 4 \\
  2 & 0 & 5 \\
  0 & 0 & 6 \\
  0 & 1 & 0 \\
  0 & 2 & 0
\end{bmatrix}
\]
By construction,

\[ S_u = [f_u, g_u, h_u, \ldots, t^{n-1}f_u, t^{n-1}g_u, t^{n-1}h_u]^T. \]

Let

\[ S_u^f = [f_u, \ldots, t^{n-1}f_u]^T, \quad S_u^g = [g_u, \ldots, t^{n-1}g_u]^T, \quad S_u^h = [h_u, \ldots, t^{n-1}h_u]^T. \]

Then \( S_u \) is the matrix generated by interleaving \( S_u^f, S_u^g, S_u^h \) column by column.

Similarly by construction,

\[ M_v = [\bar{f}_{0,v}, \bar{g}_{0,v}, \bar{h}_{0,v}, \ldots, \bar{f}_{n-1,v}, \bar{g}_{n-1,v}, \bar{h}_{n-1,v}]^T. \]

Let

\[ M_v^f = [\bar{f}_{0,v}, \ldots, \bar{f}_{n-1,v}]^T, \quad M_v^g = [\bar{g}_{0,v}, \ldots, \bar{g}_{n-1,v}]^T, \quad M_v^h = [\bar{h}_{0,v}, \ldots, \bar{h}_{n-1,v}]^T. \]

Then \( M_v \) is the matrix generated by interleaving \( M_v^f, M_v^g, M_v^h \) column by column.

To compute the product \( S_u \cdot M_v^T \), let us examine \( M_v^f, M_v^g, M_v^h \) in more detail.

\[ \phi(g, h) = \begin{vmatrix} g(s, t) & h(s, t) \\ g(s, t) & h(s, t) \end{vmatrix}_{\lambda - \mu} = \left( \sum_{k=0}^{m} \sum_{l=0}^{m} g_k(t) s^k \cdot h_l(t) s^l - \sum_{k=0}^{m} \sum_{l=0}^{m} g_k(t) s^k \cdot h_l(t) s^l \right) / (\lambda - \mu) \]

\[ = \sum_{k=0}^{m} \sum_{l=0}^{m} \begin{vmatrix} g_k(t) & h_l(t) \\ g_k(t) & h_l(t) \end{vmatrix} \cdot s^{k+l} \]

\[ = \sum_{r=0}^{2m} \left( \sum_{k+l=r} \begin{vmatrix} g_k(t) & h_l(t) \\ g_k(t) & h_l(t) \end{vmatrix} \right) \cdot s^r. \quad (3.44) \]

But the coefficient matrix of \[ \begin{vmatrix} g_k(t) & h_l(t) \\ g_k(t) & h_l(t) \end{vmatrix} / (\lambda - \mu) \] on the right hand side of Equation (3.44) is the Bézout resultant matrix of \( g_k \) and \( h_l \) [32] [46] [Chapter 2. Section 2.1].
Since $M_v^T$, $0 \leq v \leq 2m$, denotes the coefficients of $s^v$ in $\phi(g, h)$. $M_v^T$ is a sum of Bézout matrices:

$$M_v^T = \sum_{k+l=v} B_v(z)g_k h_l. \quad (3.45)$$

Now recall that Bézout matrices are symmetric [15] [32] [Chapter 2, Section 2.1], so $M_v^T$ is symmetric. Similarly, $M_v^T$, $M_v^h$ are symmetric. Since $M_v$ is generated by interleaving $M_v^T$, $M_v^g$, $M_v^h$ column by column, $M_v^T$ is the matrix generated by interleaving $M_v^T$, $M_v^g$, $M_v^h$ row by row. Therefore

$$S_u \cdot M_v^T = \left(\text{interleaving } S_u^f, S_u^g, S_u^h \text{ column by column}\right)$$

$$\cdot \left(\text{interleaving } M_v^T, M_v^g, M_v^h \text{ row by row}\right)$$

$$= S_u^f \cdot M_v^T + S_u^g \cdot M_v^g + S_u^h \cdot M_v^h. \quad (3.46)$$

### 3.6.2 The Convolution Identities

To prove the convolution identities (3.42), we now investigate the right hand side of Equation (3.46). By Equation (3.45),

$$S_u^f \cdot M_v^T = S_u^f \cdot \sum_{k+l=v} B_v(z)g_k h_l$$

$$= \left[f_u, \ldots, t^{n-1} f_u\right]^C \cdot \sum_{k+l=v} B_v(z)g_k h_l$$

$$= \sum_{k+l=v} \left[f_u, \ldots, t^{n-1} f_u\right]^C \cdot B_v(z)g_k h_l. \quad (3.47)$$

Recall that each column in a Bézout matrix of order $n$ represents a polynomial of degree $n - 1$ [32] [54] [Chapter 2, Section 2.1]. Let the polynomials represented by the columns of $B_v(z)g_k h_l$ be $p_j^{k,l}(t)$, $0 \leq j \leq n - 1$. Then, by Equation (3.47), we have

$$S_u^f \cdot M_v^T = \sum_{k+l=v} \left[f_u, \ldots, t^{n-1} f_u\right]^C \cdot \left[p_0^{k,l}, \ldots, p_{n-1}^{k,l}\right]^C. \quad (3.48)$$

Let us compute the product $\left[f_u, \ldots, t^{n-1} f_u\right]^C \cdot \left[p_0^{k,l}, \ldots, p_{n-1}^{k,l}\right]^C$. 
By construction,
\[
\begin{bmatrix}
1 & \cdots & t^{2n-1}
\end{bmatrix}_{1 \times 2n} \cdot \left[ f_u \cdots t^{n-1} f_u \right]^{C} \cdot \left[ p_{0}^{k,l} \cdots p_{n-1}^{k,l} \right]^{C} \\
= [f_u \cdots t^{n-1} f_u]_{1 \times n} \cdot \left[ p_{0}^{k,l} \cdots p_{n-1}^{k,l} \right]^{C} \\
= f_u \cdot \begin{bmatrix}
1 & \cdots & t^{n-1}
\end{bmatrix}_{1 \times n} \cdot \left[ p_{0}^{k,l} \cdots p_{n-1}^{k,l} \right]^{C} \\
= f_u \cdot \begin{bmatrix}
p_{0}^{k,l} & \cdots & p_{n-1}^{k,l}
\end{bmatrix}_{1 \times n} \\
= [f_u \cdot p_{0}^{k,l} \cdots f_u \cdot p_{n-1}^{k,l}]_{1 \times n} \\
= \begin{bmatrix}
1 & \cdots & t^{2n-1}
\end{bmatrix}_{1 \times 2n} \cdot \left[ f_u \cdot p_{0}^{k,l} \cdots f_u \cdot p_{n-1}^{k,l} \right]^{C}
\end{bmatrix}_{2n \times n}.
\tag{3.49}
\]

It follows from Equation (3.49) that
\[
\begin{bmatrix}
[f_u \cdots t^{n-1} f_u]^{C} \cdot \left[ p_{0}^{k,l} \cdots p_{n-1}^{k,l} \right]^{C}
\end{bmatrix} = [f_u \cdot p_{0}^{k,l} \cdots f_u \cdot p_{n-1}^{k,l}]^{C}.
\tag{3.50}
\]

Therefore, by Equations (3.48) and (3.50)
\[
S_{u}^{l} \cdot M_{v}^{I} = \left[ f_u \cdot \sum_{k+l=v} p_{0}^{k,l} \cdots f_u \cdot \sum_{k+l=v} p_{n-1}^{k,l} \right]^{C}.
\tag{3.51}
\]

On the other hand, by definition
\[
\begin{vmatrix}
g_{k}(t) & h_{l}(t) \\
g_{k}(s) & h_{l}(s)
\end{vmatrix} \cdot [(s - t)^{j}] = \sum_{j=0}^{n-1} p_{j}^{k,l}(t) s^{j}.
\]

Thus by Equation (3.44),
\[
\varphi(g, h) = \sum_{j=0}^{n-1} \sum_{v=0}^{2m} \left( \sum_{k+l=v} p_{j}^{k,l}(t) \right) s^{v} s^{j}.
\]

so \(\sum_{k+l=v} p_{j}^{k,l}(t)\) is the coefficient of \(s^{v} s^{j}\) in \(\varphi(g, h)\). Therefore, \(\sum_{u+v=1} f_{u} \cdot \sum_{k+l=v} p_{j}^{k,l}\) denotes the coefficient of \(s^{v} s^{j}\) in \(f(s, t) \cdot \varphi(g, h)\). Hence from Equation (3.51), we see that the \((j + 1)\)st column of
\[
\sum_{u+v=1} S_{u}^{l} \cdot M_{v}^{I}
\]
is precisely the coefficient of \(s^{v} s^{j}\) in \(f(s, t) \cdot \varphi(g, h)\).
Similar results hold for $S^g_u \cdot M^g_v$ and $S^h_u \cdot M^h_v$: the $(j+1)$st columns of

$$\sum_{n+r=1} S^g_u \cdot M^g_v \quad \text{and} \quad \sum_{n+r=1} S^h_u \cdot M^h_v$$

denote the coefficients of $s^j \cdot \beta^j$ in $g(s,t) \cdot \phi(h,f)$ and $h(s,t) \cdot \phi(f,g)$ respectively. Hence, by Equation (3.46), the polynomial represented by the $(j+1)$st column of $\sum_{n+r=1} S_u \cdot M_v^r$ is exactly the coefficient of $s^j \cdot \beta^j$ in

$$f(s,t) \cdot \phi(g,h) + g(s,t) \cdot \phi(h,f) + h(s,t) \cdot \phi(f,g).$$

But, by Equation (3.43), we know that

$$f(s,t) \cdot \phi(g,h) + g(s,t) \cdot \phi(h,f) + h(s,t) \cdot \phi(f,g) \equiv 0;$$

therefore,

$$\sum_{n+r=1} S_u \cdot M_v^r = 0, \quad 0 \leq i \leq 3m.$$

### 3.7 Fast Computation of the Entries of $Cay(f,g,h)$

In this section we will derive an efficient algorithm based on the block structures of $Syl(f,g,h)$. $Cay(f,g,h)$, and the transformation matrix $F$ from $Syl(f,g,h)$ to $Cay(f,g,h)$ to compute the entries of the Cayley resultant matrix $Cay(f,g,h)$.
From Equations (3.33), (3.41), (3.21), and (3.25), we have

\[
\begin{bmatrix}
C_{\text{ay}}(f, g, h) \\
0_{4mn \times 2mn}
\end{bmatrix}
= 
\begin{bmatrix}
S_0 \\
\vdots \\
S_{m-1} & S_0 & \cdots & S_0 \\
S_m & \vdots & \cdots & \vdots \\
S_{m-1} & S_0 & \cdots & S_0 \\
S_m & \vdots & \cdots & \vdots \\
S_m & \vdots & \cdots & \vdots \\
S_m
\end{bmatrix}
\begin{bmatrix}
F_0 & F_1 & \cdots & F_{2m-1} \\
F_1 & \cdots & \cdots & \cdots \\
\vdots & \vdots & \ddots & \vdots \\
F_{2m-1}
\end{bmatrix}.
\tag{3.52}
\]

Recall that each block \( S_i \) is of size \( 2n \times 3n \), each block \( F_j \) is of size \( 3n \times n \), and each block \( C_{i,j} \) is of size \( 2n \times n \). Considering just the first \( 2mn \) rows on both sides of Equation (3.52), we have

\[
\begin{bmatrix}
C_{0,0} & \cdots & C_{0,2m-1} \\
\vdots & \ddots & \vdots \\
C_{m-1,0} & \cdots & C_{m-1,2m-1}
\end{bmatrix}
= 
\begin{bmatrix}
S_0 \\
\vdots \\
S_{m-1} & S_0
\end{bmatrix}
\begin{bmatrix}
F_0 & \cdots & F_{m-1} & \cdots & F_{2m-1} \\
\vdots & \ddots & \ddots & \ddots & \vdots \\
F_{m-1} & \cdots & F_{2m-1}
\end{bmatrix}.
\tag{3.53}
\]

Therefore,

\[
C_{i,j} = \sum_{k=0}^{\min(i,2m-1-j)} S_{i-k} \cdot F_{j+k}, \quad 0 \leq i \leq m - 1, \quad 0 \leq j \leq 2m - 1.
\tag{3.54}
\]

Notice the similarity of Equation (3.54) for \( C_{i,j} \) to Equation (2.14) for \( B_{i,j} \). Again, we observe that

\[
C_{i,j} = C_{i-1,j+1} + S_i \cdot F_j, \quad 1 \leq i \leq m - 1, \quad 0 \leq j \leq 2m - 2,
\]

which leads to the following fast algorithm for computing the entries of \( C_{\text{ay}}(f, g, h) \):
1. Initialization:

\[
\text{Cay}(f, g, h)_{\text{init}} = \begin{bmatrix}
S_0 \cdot F_0 & \cdots & S_0 \cdot F_{2m-1} \\
\vdots & \ddots & \vdots \\
S_{m-1} \cdot F_0 & \cdots & S_{m-1} \cdot F_{2m-1}
\end{bmatrix} .
\]

That is,

\[
(C_{i,j})_{\text{init}} = S_i \cdot F_j, \quad 0 \leq i \leq m - 1, \quad 0 \leq j \leq 2m - 1.
\]

2. Recursion:

for \( i \) from 1 to \( m - 1 \)

for \( j \) from 0 to \( 2m - 2 \)

\[
C_{i,j} \leftarrow C_{i,j} + C_{i-1,j+1}
\]

\[
\text{Cay}(f, g, h) = \begin{bmatrix}
S_0 \cdot F_0 & \cdots & S_0 \cdot F_{m-1} & S_0 \cdot F_m & \cdots & S_0 \cdot F_{2m-1} \\
\vdots & \ddots & \ddots & \ddots & \ddots & \ddots \\
\vdots & \ddots & \ddots & \ddots & \ddots & \ddots \\
\vdots & \cdots & \ddots & \ddots & \ddots & \ddots \\
S_{m-2} \cdot F_{m-1} & \cdots & S_{m-1} \cdot F_{2m-1} \\
S_{m-1} \cdot F_{m-1} & \cdots & S_{m-1} \cdot F_{2m-1}
\end{bmatrix} .
\]

Notice that this method for computing the entries of \( \text{Cay}(f, g, h) \) eliminates lots of redundant calculations: after the initialization, we need only update the blocks along the southwest diagonals by one (block) addition per block. This approach is much faster than the standard method of calculating each entry independently \([6] [50]\), especially when we compute the Bézoutians \( F_k \) efficiently [see Section 2.5]. In the next section, we present a more detailed analysis of the computational complexity of this algorithm.

### 3.8 Computational Complexity

To compute the entries of \( \text{Cay}(f, g, h) \) by the method developed in Section 3.7, we need the following three steps:
1) Preprocessing: compute $F_u$, $0 \leq u \leq 2m - 1$:

2) Initialization: calculate $(C_{i,u})_{init} = S_i \cdot F_u$, $0 \leq i \leq m - 1$, $0 \leq u \leq 2m - 1$:

3) Recursion: march along the southwest diagonals, and update the entries of $C_{i,u}$.

**Step 1:**

Each $F_u$ contains the following three summations of Bézout matrices:

$$\sum_{i+j=u+1} B_{ez}(g_i, h_j), \sum_{i+j=u+1} B_{ez}(h_i, f_j), \sum_{i+j=u+1} B_{ez}(f_i, g_j).$$

We can use the technique developed in Section 2.4 to efficiently compute the sum

$$\sum_{i+j=u+1} B_{ez}(g_i, h_j).$$

First we initialize each Bézout matrix $B_{ez}(g_i, h_j)_{init}$; then we add these initializations together; and last we march along the southwest diagonals updating the entries as we go. Altogether, we need

$$\frac{(n^2 + n) \cdot \min(u + 2, m + 1)}{\text{initialization}}$$

and

$$\frac{n^2 + n}{2} \cdot \min(u + 2, m + 1) + \frac{n^2 + n}{2} \cdot \min(u + 1, m) + \frac{n^2 - n}{2} \text{ additions.}$$

Similarly, we can compute

$$\sum_{i+j=u+1} B_{ez}(h_i, f_j), \sum_{i+j=u+1} B_{ez}(f_i, g_j)$$

efficiently. Therefore, in Step 1.

The number of multiplications

$$\# \text{ of multiplications} = \sum_{u=0}^{2m-1} 3 \cdot (n^2 + n) \cdot \min(u + 2, m + 1)$$

$$= 3(n^2 + n) \left[ \sum_{u=0}^{m-1} (u + 2) + \sum_{u=m}^{2m-1} (m + 1) \right]$$

$$= 3(n^2 + n) \frac{3m^2 + 5m}{2} = \frac{3}{2} (n^2 + n)(3m^2 + 5m).$$
and
\[
\# \text{ of additions} = \sum_{u=0}^{2m-1} 3 \cdot \left[ \frac{n^2 + n}{2} \min(u + 2, m + 1) + \frac{n^2 + n}{2} \min(u + 1, m) + \frac{n^2 - n}{2} \right]
\]
\[
= 3 \cdot \frac{n^2 + n}{2} \cdot \frac{3m^2 + 5m}{2} + 3 \cdot \frac{n^2 + n}{2} \cdot \frac{3m^2 + m}{2} + 3m(n^2 - n)
\]
\[
= \frac{9}{2}(n^2 + n)(m^2 + m) + 3m(n^2 - n).
\]

**Step 2:**

By Equations (3.32) and (3.25),
\[
S_t \cdot F_n = \begin{bmatrix}
    a_{t,0} \\
    \vdots \\
    a_{t,n-1} \\
    \vdots \\
    a_{t,n}
\end{bmatrix} \cdot \sum_{j+k=u+1} \text{Bez}(g_j, h_k) +
\begin{bmatrix}
    b_{t,0} \\
    \vdots \\
    b_{t,n-1} \\
    \vdots \\
    b_{t,n}
\end{bmatrix} \cdot \sum_{j+k=u+1} \text{Bez}(h_j, f_k) +
\begin{bmatrix}
    c_{t,0} \\
    \vdots \\
    c_{t,n-1} \\
    \vdots \\
    c_{t,n}
\end{bmatrix} \cdot \sum_{j+k=u+1} \text{Bez}(f_j, g_k).
\]

Each one of these three matrix multiplications requires
\[
2 \left( n + 2n + \cdots + n^2 \right) = n^3 + n^2 \quad \text{multiplications}.
\]
and
\[ 2 \left[ n \cdot 0 + n \cdot 1 + \cdots + n(n - 1) \right] = n^3 - n^2 \quad \text{additions}. \]

Adding these three \(2n \times n\) matrices together requires \(4n^2\) more additions. Therefore, since there are \(2m^2\) products \(S_i \cdot F_u\), Step 2 requires \(6m^2 \cdot n^2(n + 1)\) multiplications and \(2m^2 \cdot (3n^3 + n^2)\) additions.

**Step 3:**

Since each Cayley block \(C_{i,j}\) is of size \(2n \times n\), marching along the southwest diagonals requires only \((m - 1)(2m - 1) \cdot 2n^2 = 2n^2(2m^2 - 3m + 1)\) additions and no multiplication.

Altogether, we perform
\[ \frac{3}{2} (3m^2 + 5m)(n^3 + n) + 6m^2(n^3 + n^2) \quad \text{multiplications}. \]

and
\[ \frac{9}{2} (m^2 + m)(n^2 + n) + 3m(n^3 - n) + 2m^2(3n^3 + n^2) + 2n^2(2m^2 - 3m + 1) \quad \text{additions}. \]

Notice that the main bottleneck is step 2, which requires \(\mathcal{O}(m^2n^3)\) multiplications and additions.

On the other hand, the standard way \([6][50]\) to compute the entries of \(\text{Cay}(f, g, h)\) needs to compute
\[ \frac{1}{6} m(m + 1)^2(m + 2) \cdot n(n + 1)^2(n + 2) \]

\(3 \times 3\) determinants, each of which has 6 terms and hence requires 12 multiplications and 5 additions. So if we compute all these \(3 \times 3\) intermediate determinants just once and store them, we need to perform at least
\[ 2m(m + 1)^2(m + 2) \cdot n(n + 1)^2(n + 2) \quad \text{multiplications} \]

and
\[ \frac{5}{6} m(m + 1)^2(m + 2) \cdot n(n + 1)^2(n + 2) \quad \text{additions}. \]
A summary of the complexity of computing the entries of the Cayley resultant matrices using the standard method vs. the new fast algorithm is given in Table 3.1:

Parallel computation can be used to speed up the new algorithm even further. For example, we can compute each of the blocks $F_u$, $0 \leq u \leq 2m - 1$, and perform the initialization of the blocks $(C_{t,u})_{ini} = S_t \cdot F_u$ in parallel. Moreover, the recursions along different diagonals are independent, so these steps can also be done in parallel.

Finally notice that the time complexity $O(m^2n^3)$ for computing the entries of the matrix $Cay(f, g, h)$ is not symmetric in $m$ and $n$. The reason for this asymmetry is that the Sylvester resultant matrix $Syl(f, g, h)$ and the Cayley resultant matrix $Cay(f, g, h)$ are not symmetric in $s, t$ -- hence not symmetric in $m, n$. If we reverse the roles of $s, t$ during the construction of $Syl(f, g, h)$ and $Cay(f, g, h)$, and impose appropriate block structures on these matrices, the time complexity for computing the entries of $Cay(f, g, h)$ is $O(m^3n^2)$. (The difference between the two Cayley resultant matrices $Cay(f, g, h)$ is that one is the transpose of the other.) Therefore, we can actually compute the entries of $Cay(f, g, h)$ in time $O(\min(m^2n^3, m^3n^2))$.

### 3.9 Hybrids of the Three Dixon Resultants

Below we are going to derive hybrid resultant matrices from $Syl(f, g, h)$, $Mix(f, g, h)$ and $Cay(f, g, h)$.

To proceed, we adopt the following matrix notation: $I_k$ is the identity matrix of
order $k$, and

$$U_k = \begin{bmatrix} 0 & 0 & 0 & 0 & \cdots \\ 1 & 0 & 0 & 0 & \cdots \\ 0 & 1 & 0 & 0 & \cdots \\ 0 & 0 & 0 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}, \quad V_k = \begin{bmatrix} 0 & 0 & \cdots \\ 0 & 0 & \cdots \\ 1 & 0 & \cdots \\ 0 & 0 & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix}, \quad W_k = \begin{bmatrix} 1 & 0 & \cdots \\ 0 & 0 & \cdots \\ 0 & 0 & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix}.$$ 

That is, $U_k$ is generated by inserting a row of zeros before every two rows in $I_{2k}$; $V_k$ is generated by inserting two rows of zeros before every row of $I_k$; $W_k$ is generated by inserting two rows of zeros after every row of $I_k$. Thus, $U_k$ is of size $3k \times 2k$, while $V_k$ and $W_k$ are of size $3k \times k$.

### 3.9.1 Hybrids of the Sylvester and Cayley Resultants

Recall that

$$Syl(f, g, h) \cdot F(f, g, h) = \begin{bmatrix} Cay(f, g, h) \\ 0 \end{bmatrix},$$

and $F(f, g, h)$ has a simple block symmetric structure [see Section 3.3]. Let

$$F_j = \begin{bmatrix} F_1 \\ \vdots \\ F_j \end{bmatrix}.$$ 

and form the $6mn \times 6mn$ matrix

$$T_i = \begin{bmatrix} I_{6mn-3n} & F_{2m-2}^0 & 0 \\ 0 & F_{2m-1} & U_n \end{bmatrix},$$

where $n$.
Multiplying $Syl(f, g, h)$ by $T_1$, we get

$$
Syl(f, g, h) \cdot T_1 = \begin{bmatrix}
S_0 & \ldots & C_{0,0} \\
\vdots & \ddots & \vdots \\
S_{m-1} & \ldots & C_{m-1,0} \\
S_m & \ldots & \\
\vdots & \ddots & \\
\vdots & & \\
0 & & S_0 \cdot U_n \\
& & \vdots \\
& & S_{m-1} \cdot U_n \\
& & S_m \cdot U_n \\
\end{bmatrix}
$$

(3.55)

where the left part contains the first $2m - 1$ columns (of blocks $S_i$) of $Syl(f, g, h)$, while the middle part consists of the first column (of blocks $C_{i,j}$) of $Cay(f, g, h)$, and the right part consists of the product of the last column of $Syl(f, g, h)$ with $U_n$. Note, in particular, that the lower-right corner $2n \times 2n$ submatrix $S_m \cdot U_n$ is $S_m$ with the $f_m$-columns dropped. Thus $S_m \cdot U_n$ is exactly the Sylvester matrix of the polynomials $g_m, h_m$; that is, $S_m \cdot U_n = Syl(g_m, h_m)$.

Let $H_1$ be the top-left corner submatrix of the right hand side of Equation (3.55):

$$
H_1 = \begin{bmatrix}
S_0 & \ldots & C_{0,0} \\
\vdots & \ddots & \vdots \\
S_{m-1} & \ldots & C_{m-1,0} \\
S_m & \ldots & \\
\vdots & \ddots & \\
\vdots & & \\
0 & & S_0 \\
& & \vdots \\
& & S_{m-1} \\
& & S_m \\
\end{bmatrix}
$$

We are now going to show that $|H_1|$ is the resultant of the original polynomials $f(s, t), g(s, t), h(s, t)$. Recall that each $S_i$ is of size $2n \times 2n$, and each $C_{i,j}$ is of size $2n \times n$. Therefore $H_1$ is a square matrix of order $6mn - 2n$. Thus we can rewrite
Equation (3.55) as

\[ Syl(f, g, h) \cdot T_1 = \begin{bmatrix} H_1 & * \\ 0 & Syl(g_m, h_m) \end{bmatrix}. \]

It follows that

\[ |Syl(f, g, h)| \cdot |T_1| = |H_1| \cdot |Syl(g_m, h_m)|. \tag{3.56} \]

But by construction, \(|T_1| = [[F_{2m-1} \ U_n]].\) Recall from Section 3.5 that \(F_{2m-1} = M_{2n}^x\) is a Bézoutiant. Now by the construction of \(U_n\), we need only consider the rows of \(F_{2m-1}\) that do not overlap with the non-zero rows of \(U_n\) when computing \(\|[F_{2m-1} \ U_n]\|\). These rows of \(F_{2m-1}\) form exactly the Bézout matrix of \(g_m, h_m\) [see Section 3.3.3]. Therefore

\[ |T_1| = \|[F_{2m-1} \ U_n]\| = \pm |Bez(g_m, h_m)|. \tag{3.57} \]

Since we know that the Bézout resultant and the Sylvester resultant of two univariate polynomials are the same up to sign, we conclude from Equations (3.56) and (3.57) that \(|Syl(f, g, h)|\) and \(|H_1|\) can only differ by at most a sign. Thus replacing the last column (of blocks \(S_i\)) of \(Syl(f, g, h)\) by the first column (of blocks \(C_{i,\gamma}\)) of \(Cay(f, g, h)\) gives us a hybrid resultant \(H_1\) of the Sylvester resultant and the Cayley resultant of order \(6mn - 2n\).

Examining the block structure of \(F(f, g, h)\), we see that the first column of blocks of \(Cay(f, g, h)\) depends on all the columns of \(Syl(f, g, h)\), but the second column of blocks of \(Cay(f, g, h)\), i.e. the coefficients of \([\alpha^i, \cdots, \alpha^i, \beta^{n-1}]\), depends only on the first \(2m - 1\) columns of blocks of \(Syl(f, g, h), Ls^i[1, \cdots, t^{n-1}], i = 0, \cdots, 2m - 2\). Therefore we can continue the previous process, replacing the column of \(H_1\) with index \((2m - 2)\), i.e. \(Ls^{2m-2}[1, \cdots, t^{n-1}]\), by the second column of \(Cay(f, g, h)\), i.e. the coefficients of \([\alpha^i, \cdots, \alpha^i, \beta^{n-1}]\). That is, there exists a transformation matrix \(T_2\).

\[ T_2 = \begin{bmatrix} I_{6mn-6n} & 0 & F_{2m-2}^i \\ 0 & 0 & F_{2m-1} \ U_n \\ 0 & I_n & 0 & 0 \end{bmatrix}. \tag{3.58} \]
such that

$$H_1 \cdot T_2 = \begin{bmatrix} H_2 & \ast \\ 0_{2n \times (6mn-4n)} & Syl(g_m, h_m) \end{bmatrix}.$$  \hspace{1cm} (3.59)

where $H_2$ is the square matrix of order $6mn-4n$, consisting of the first $2m-2$ columns of blocks of $Syl(f, g, h)$ and the first 2 columns of blocks of $Cay(f, g, h)$. Notice that by Equation (3.58) we again have

$$|T_2| = |[F_{2m-1}, T_n]| = \pm |Bvz(g_m, h_m)|.$$

Since $|H_1| \cdot |T_2| = |H_2| \cdot |Syl(g_m, h_m)|$ by Equation (3.59), it follows that $|H_2| = \pm |H_1| = \pm |Syl(f, g, h)|$. Continuing in this fashion, we obtain $2m - 1$ hybrids of the Sylvester and Cayley resultants.

After $2m$ repetitions of this process, we arrive at the Cayley matrix. This gives a direct proof that $|Syl(f, g, h)| = \pm |Cay(f, g, h)|$ without appealing to the properties of resultants. By the way, a fourth Dixon resultant matrix [25] is similar to but not the same as the $m-$th hybrid in this sequence: the $m-$th hybrid has $mn$ columns from $Cay(f, g, h)$ which represent polynomials of degree $m - 1$ in $s$ and $2n - 1$ in $t$, whereas Dixon's fourth resultant has $mn$ columns that represent polynomials of degree $2m - 1$ in $s$ and $2n - 1$ in $t$.

### 3.9.2 Hybrids of the Sylvester and Mixed Cayley-Sylvester Resultants

In this section, we are going to derive hybrids of the Sylvester and the mixed Cayley-Sylvester resultants. We will replace each column (of blocks of $S^*_i$) of $Syl(f, g, h)^*$ by one column (of blocks of $M^*_{i,j}$) from $Mix(f, g, h)^*$. and decrease the order of the resulting matrix by $3m$ without changing the determinant (up to sign).

Recall that

$$Syl(f, g, h)^* \cdot G(f, g, h)^* = \begin{bmatrix} Mix(f, g, h)^* \\ 0 \end{bmatrix},$$
and $G(f, g, h)^*$ has a simple block symmetric structure [c.f. Section 3.2.3]. Let

$$G_j^* = \begin{bmatrix} G_i^* \\ \vdots \\ G_j^* \end{bmatrix}.$$ 

and form the $6mn \times 6mn$ matrix

$$T_1 = \begin{bmatrix} I_{6mn-6n} & G_{n-2}^* & 0 \\ 0 & G_{n-1}^* & \begin{bmatrix} V_m & 0 \\ 0 & U_m^* \end{bmatrix} \end{bmatrix}_{6mn \times 6mn}.$$ 

Multiplying $Syl(f, g, h)^*$ by $T_1$, we get

$$Syl(f, g, h)^* \cdot T_1 = \begin{bmatrix} S_0^* & M_{0,0}^* \\ \vdots & \vdots & \vdots \\ S_{n-2}^* & S_0^* & M_{n-2,0}^* \\ S_{n-1}^* & S_1^* & M_{n-1,0}^* & S_0^* & \begin{bmatrix} V_m & 0 \\ 0 & U_m^* \end{bmatrix} \\ \vdots & \vdots & \vdots \\ S_n^* & S_{n-1}^* & \begin{bmatrix} V_m & 0 \\ 0 & U_m^* \end{bmatrix} \\ 0 & S_n^* & \begin{bmatrix} V_m & 0 \\ 0 & U_m^* \end{bmatrix} \end{bmatrix}.$$

where the left part consists of the first $n - 1$ columns (of blocks $S_i^*$) of $Syl(f, g, h)^*$. while the middle part contains the first column of blocks of $Mix(f, g, h)^*$, and the right part consists of the product of the last column of blocks of $Syl(f, g, h)^*$ with $\begin{bmatrix} V_m & 0 \\ 0 & U_m^* \end{bmatrix}$. Let $H_1$ be the top-left corner submatrix of the right hand side of Equa-
tion (3.60):

\[
H_1 = \begin{bmatrix}
S_0^* & M_{0,0}^* \\
S_1^* & S_0^* & M_{n-2,0}^* \\
S_{n-1}^* & S_1^* & M_{n-1,0}^* \\
\vdots & \vdots & \ddots & \vdots \\
S_n^* & S_{n-2}^* & \cdots & S_0^* & M_{n-0,0}^* \\
\end{bmatrix}
\]

Since each \( S_i^* \) is of size \( 3m \times 6m \), and each \( M_{i,j}^* \) is of size \( 3m \times 3m \), \( H_1 \) is a square matrix of order \( 6mn - 3m \). Now we can rewrite Equation (3.60) as

\[
Syl(f, g, h)^* \cdot T_1 = H_1 \cdot S_n^* \cdot \begin{bmatrix}
V_m & 0 \\
0 & U_m
\end{bmatrix}.
\]

It follows that

\[
|Syl(f, g, h)^*| \cdot |T_1| = |H_1| \cdot |S_n^*| \cdot \begin{bmatrix}
V_m & 0 \\
0 & U_m
\end{bmatrix}.
\]  \hspace{1cm} (3.61)

But recall that \( S_n^* \) consists of the coefficients of the monomials \([s^* f_n^*, s^* g_n^*, s^* h_n^*]\); \( r = 0, \ldots, 2m - 1 \); that is.

\[
S_n^* = \begin{bmatrix}
L_{0,n} \\
\vdots \\
L_{m-1,n} & \cdots & L_{0,n} \\
L_{m,n} & \cdots & L_{1,n} & L_{0,n} \\
\vdots & \vdots & \ddots & \vdots \\
L_{m,n} & L_{m-1,n} & \cdots & L_{0,n} \\
L_{m,n} & \cdots & L_{1,n} \\
\vdots \\
L_{m,n}
\end{bmatrix}.
\]
where \( L_{i,j} = [a_{i,j} \ b_{i,j} \ c_{i,j}] \). Therefore,

\[
S_n^* \begin{bmatrix} V_m & 0 \\ 0 & U_m \end{bmatrix} = \begin{bmatrix} c_{0,n} \\ \vdots \\ \vdots \\ c_{m-1,n} & \cdots & c_{0,n} \\ c_{m,n} & \cdots & c_{1,n} & b_{0,n} & c_{0,n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \cdots \\ c_{m,n} & b_{m-1,n} & c_{m-1,n} & \cdots & b_{0,n} & c_{0,n} \\ b_{m,n} & c_{m,n} & \cdots & b_{1,n} & c_{1,n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \cdots \\ b_{m,n} & c_{m,n} \end{bmatrix},
\]

so it is easy to see that

\[
|S_n^* \begin{bmatrix} V_m & 0 \\ 0 & U_m \end{bmatrix}| = (c_{0,n})^m \cdot |SyI(g_n^*, h_n^*)|.
\]

Hence by Equation (3.61), we have

\[
|SyI(f, g, h)^*| \cdot |T_1| = \pm |H_1| \cdot (c_{0,n})^m \cdot |SyI(g_n^*, h_n^*)|.
\]

(3.62)

Now notice that \(|T_1| = \left| \begin{bmatrix} G_{n-1}^* & \begin{bmatrix} V_m & 0 \\ 0 & U_m \end{bmatrix} \end{bmatrix} \right|\). Moreover, by the construction of \( V_m \) and \( U_m \), we need only consider the rows of \( G_{n-1}^* \) that do not overlap with the nonzero rows of \( V_m \) and \( U_m \) when we compute the determinant of \( G_{n-1}^* \begin{bmatrix} V_m & 0 \\ 0 & U_m \end{bmatrix} \). That is, we only care about the rows \( 3j, 3j+1, j = 0, \cdots, m \leq 1 \) and the rows \( 3m+3k, k = 0, \cdots, m - 1 \) of \( G_{n-1}^* \). Collecting these
3m rows of $G_{n-1}^*$, we get the following $3m \times 3m$ matrix [c.f. Equation (3.15)]:

\[
\begin{bmatrix}
0 & -c_{0,n} & b_{0,n} & 0 & 0 & 0 & \cdots & 0 & 0 \\
c_{0,n} & 0 & -a_{0,n} & 0 & 0 & 0 & \cdots & 0 & 0 \\
0 & -c_{1,n} & b_{1,n} & 0 & -c_{0,n} & b_{0,n} & \cdots & 0 & 0 \\
c_{1,n} & 0 & -a_{1,n} & c_{0,n} & 0 & -a_{0,n} & \cdots & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
0 & -c_{m,n} & b_{m,n} & 0 & -c_{m-1,n} & b_{m-1,n} & \cdots & -c_{1,n} & b_{1,n} \\
0 & 0 & 0 & 0 & -c_{m,n} & b_{m,n} & \cdots & -c_{2,n} & b_{2,n} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & 0 & 0 & 0 & 0 & \cdots & -c_{m,n} & b_{m,n}
\end{bmatrix}
\]

Rearrange the rows and columns of this matrix as follows: group the columns indexed by $3j$, $j = 0, \ldots, m-1$, together and collect the rows indexed by $2j+1$, $j = 0, \ldots, m-1$, together, and put them at the top-left corner. Then we get the matrix

\[
\begin{bmatrix}
\begin{array}{ccccccccc}
c_{0,n} & 0 & \cdots & 0 & -a_{0,n} & 0 & 0 & 0 & 0 \\
c_{1,n} & c_{0,n} & \cdots & 0 & 0 & -a_{1,n} & 0 & -a_{0,n} & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
c_{m-1,n} & c_{m-2,n} & \cdots & c_{0,n} & 0 & -a_{m-1,n} & 0 & -a_{m-2,n} & \cdots & 0 \\
0 & 0 & \cdots & 0 & -c_{0,n} & b_{0,n} & 0 & 0 & \cdots & 0 \\
0 & 0 & \cdots & 0 & -c_{1,n} & b_{1,n} & -c_{0,n} & b_{0,n} & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & 0 & -c_{m-1,n} & b_{m-1,n} & -c_{m-2,n} & b_{m-2,n} & \cdots & -c_{1,n} \\
0 & 0 & \cdots & 0 & -c_{m,n} & b_{m,n} & -c_{m-1,n} & b_{m-1,n} & \cdots & -c_{1,n} \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & \cdots & -c_{m,n} \\
\end{array}
\end{bmatrix}
\]

It is easy to see that the determinant of this matrix is equal to $\pm (c_{0,n})^m \cdot |Syl(g_n^*, h_n^*)|$. Therefore by Equation (3.62)

\[
|Syl(f, g, h)^*| \cdot (c_{0,n})^m \cdot |Syl(g_n^*, h_n^*)| = \pm |H_1| \cdot (c_{0,n})^m \cdot |Syl(g_n^*, h_n^*)|.
\]

Hence $|Syl(f, g, h)^*|$ and $|H_1|$ are the same up to sign. Thus, we get a hybrid resultant matrix $H_1$ of the Sylvester and Cayley-Sylvester resultants with order $6mn - 3m$. 
Checking the block structure of $G(f, g, h)^*$, we see that the column of blocks of $\text{Mix}(f, g, h)^*$ with index $(n - 1 - j)$ depends on the columns of blocks of $Syl(f, g, h)^*$ with indices $[0, \ldots, j]$. Therefore, we can continue this process replacing the column of blocks of $Syl(f, g, h)^*$ indexed by $(n - 1 - j)$, i.e., $Lt^{n-1-j}[1 \cdots s^{2m-1}]$, by the column of blocks of $\text{Mix}(f, g, h)^*$ indexed by $j$, i.e., $\bar{T}_j[1 \cdots s^{m-1}]$, for $j = 0, \ldots, n - 1$. For example, repeating the previous process, i.e., replacing the $(n - 1)$st column of blocks of $H_1$, i.e., $Lt^{n-2}[1 \cdots s^{2m-1}]$ by the second column of blocks of $\text{Mix}(f, g, h)^*$, i.e. $\bar{T}_1[1 \cdots s^{m-1}]$, we get a second hybrid resultant matrix $H_2$ of the Sylvester and Cayley-Sylvester resultants with order $6mn - 6m$.

After $n$ repetitions of this process, we arrive at the mixed Cayley-Sylvester resultant. Again this provides a direct proof that $|Syl(f, g, h)^*| = |\text{Mix}(f, g, h)^*|$ without appealing to the properties of resultants.

### 3.9.3 Hybrids of the Mixed Cayley-Sylvester and Cayley Resultants

In this section we present the construction of hybrid resultants that consist of some columns of the mixed Cayley-Sylvester matrix and some columns of the transposed Cayley matrix. The derivation in this section is very similar to those of Section 3.9.2 and Section 3.9.1.

Recall that

$$\text{Mix}(f, g, h) \cdot E(f, g, h) = \begin{bmatrix} \text{Cay}(f, g, h)^T \\ 0 \end{bmatrix}.$$

and $E(f, g, h)$ has a simple block symmetric structure [c.f. Section 3.4.3]. Let

$$E_j^i = \begin{bmatrix} E_1 \\ \vdots \\ E_j \end{bmatrix}.$$

and write

$$\text{Mix}(f, g, h) = \begin{bmatrix} K_0 & K_1 & \cdots & K_{m-1} \end{bmatrix}.$$
where each $K_j$, $0 \leq j \leq m - 1$, consists of a column of blocks $M_i$. Let

$$K_j = \begin{bmatrix} K_1 & \cdots & K_j \end{bmatrix}.$$ 

Since $E(f, g, h)$ is block symmetric and upper triangular,

$$Mix(f, g, h) \cdot E(f, g, h) = \begin{bmatrix} K_0^{m-1} E_{m-1}^0 & K_0^{m-2} E_{m-1}^1 & \cdots & K_0^0 E_{m-1}^{m-1} \end{bmatrix}$$

$$= \begin{bmatrix} B_{0,0}^T & B_{1,0}^T & \cdots & B_{m-1,0}^T \\ \vdots & \vdots & \cdots & \vdots \\ B_{0,2m-1}^T & B_{1,2m-1}^T & \cdots & B_{m-1,2m-1}^T \\ 0 & 0 & \cdots & 0 \end{bmatrix}.$$ 

Note that the bottom $(m - 1 - j) \cdot n$ rows of $K_j^0$ are all zero. Moreover, $K_{m-1-j} E_{m-1}^j$, $j = 0, \cdots, m - 1$, has at most $2mn$ non-zero rows, since

$$K_{m-1-j}^0 E_{m-1}^j = \begin{bmatrix} B_{j,0}^T \\ \vdots \\ B_{j,2m-1}^T \\ 0 \end{bmatrix}.$$ 

The $j$-th hybrid matrix $H_j$ is given by

$$\begin{bmatrix} H_j \\ 0 \end{bmatrix} = \begin{bmatrix} K_{m-1-j}^0 \ K_{m-1}^0 E_{m-1}^0 \ K_{m-2}^0 E_{m-1}^1 \ \cdots \ K_{m-j}^0 E_{m-1}^{j-1} \\ 0 & B_{0,0}^T & B_{1,0}^T & \cdots & B_{j-1,0}^T \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & B_{0,2m-1}^T & B_{1,2m-1}^T & \cdots & B_{j-1,2m-1}^T \end{bmatrix}$$

for $0 \leq j \leq m$. That is, $H_j$ consists of the first $m - j$ columns of blocks of $Mix(f, g, h)$ and the first $j$ columns of blocks of $Cay(f, g, h)^T$. With the bottom $n \cdot j$ rows of zeros truncated. Note that the matrix $H_j$ has $(3m - j) \cdot n$ rows and $3n(m - j) + 2n \cdot j = n(3m - j)$ columns. Hence $H_j$ is a square matrix. In particular, $H_0$ is $Mix(f, g, h)$, and $H_m$ is $Cay(f, g, h)^T$. 
To show that \(|H_j|\) and \(|H_{j+1}|\) are equal up to sign, \(0 \leq j \leq m - 1\), we construct the matrix

\[
T_j = \begin{bmatrix}
I_{\mathbb{H}(m-j-1)n} & 0 & E_{m-2}^j & 0 \\
0 & 0 & E_{m-1} & W_n \\
0 & I_{2jm} & 0 & 0
\end{bmatrix},
\]

where \(W_n\) is the \(3n \times n\) matrix defined at the beginning of Section 3.9. It is easy to see that

\[
|T_j| = \pm |[E_{m-1} \ W_n]|.
\]

By the construction of \(W_n\), we need only consider the rows of \(E_{m-1}\) that do not overlap with the non-zero rows of \(W_n\) to compute \(|[E_{m-1} \ W_n]|\). Since \(E_{m-1} = -S_m^T\) [c.f. Equation (3.31), (3.32)], and since \(S_m\) is Sylvester-like (Section 3.1.1), \(|[E_{m-1} \ W_n]|\) is exactly the Sylvester resultant of the univariate polynomials \(g_m, h_m\) defined in Section 3.1.1. Therefore,

\[
|T_j| = \pm Syl(g_m, h_m).
\]

Moreover,

\[
\begin{align*}
\left[ \begin{array}{c}
H_j \\
0
\end{array} \right] \cdot T_j \\
= \left[ \begin{array}{cccc}
K_{m-1-j}^{0} & K_{m-1}^{0} & E_{m-1}^{0} & \cdots & K_{m-j-1}^{0} & E_{m-1}^{j-1}
\end{array} \right] \\
= \left[ \begin{array}{c}
I \\\n0
\end{array} \right] E_{m-1}^{0} \cdots K_{m-j-1}^{0} E_{m-1}^{j-1} \left[ \begin{array}{c}
0 \\
W_n
\end{array} \right].
\end{align*}
\]

But

\[
K_{m-1-j}^{0} \left[ \begin{array}{c}
I_{\mathbb{H}(m-1-j)n} \\
0
\end{array} \right] = \left[ \begin{array}{cc}
K_{m-2-j}^{0} & K_{m-1-j}^{0}
\end{array} \right] \left[ \begin{array}{c}
I_{\mathbb{H}(m-1-j)n} \\
0
\end{array} \right] = K_{m-2-j}^{0}.
\]
and
\[ K_{m-1-j}^0 \begin{bmatrix} 0 \\ W_n \end{bmatrix} = \left[ K_{m-2-j}^0 \ K_{m-1-j}^1 \right] : \begin{bmatrix} 0 \\ W_n \end{bmatrix} = K_{m-1-j} \cdot W_n. \]

Hence,
\[
\begin{bmatrix} H_j \\ 0 \end{bmatrix} \cdot T_j
= \left[ K_{m-2-j}^0 \ K_{m-1}^0 E_{m-1}^0 \ \cdots \ K_{m-j-1}^0 E_{m-1}^{j-1} \ K_{m-j-1}^0 E_{m-1}^j \ K_{m-j-1} W_n \right]. \tag{3.66}
\]

Now let us focus on the top \( n \cdot (3m - j) \) rows of the right hand side of Equation (3.66), since all the rows below are zero. Notice that the top-left corner \([n \cdot (3m - j - 1)] \times [n \cdot (3m - j - 1)]\) submatrix is exactly \( H_{j+1} \). Furthermore, the last \( n \) non-zero rows contain all zeros except for \( M_{2m} \cdot W_n \), which arises from the last \( n \) columns of \( K_{m-j-1} W_n \). Since \( M_{2m} \) is a Bézoutian [c.f. Section 3.5], by the construction of \( W_n \), the product \( M_{2m} \cdot W_n \) is the Bézout matrix of \( g_m, h_m \). Therefore, Equation (3.66) implies that
\[ H_j \cdot T_j = \begin{bmatrix} H_{j+1} & * \\ 0 & Bcz(g_m, h_m) \end{bmatrix}. \]

Hence
\[ |H_j| \cdot |T_j| = |H_{j+1}| \cdot |Bcz(g_m, h_m)|. \tag{3.67} \]

By Equation (3.64) and Equation (3.67) we conclude that
\[ |H_j| = \pm |H_{j+1}|, \quad j = 0, \ldots, m - 1. \]

Thus we generate a sequence of \( m - 1 \) hybrid resultants of the Cayley and the mixed Cayley-Sylvester resultants. Since \( H_0 = Mix(f, g, h) \) and \( H_m = Cay(f, g, h)^T \), we obtain in this fashion a direct proof that \( |Mix(f, g, h)| = \pm |Cay(f, g, h)| \) without appealing to the properties of resultants.
Chapter 4

Sparse Resultants

While classical resultants have a history of more than two hundred years [2], modern resultants — sparse resultants or $A$-resultants — developed using more advanced geometric tools emerged only recently in the 1970’s. Sparse resultants extend the theory of classical resultants to detect nontrivial common roots in systems of sparse polynomial equations [21] [29] [31] [56]. For sparse polynomials, classical resultants vanish identically and fail to give any useful information about the common roots.

One standard approach to sparse resultants (equivalent to the dialytic method for classical resultants) is to introduce a set of monomials — a multiplying set — to multiply the polynomials in the original system. If the multiplying set is carefully chosen, the determinant of the coefficient matrix of the resulting polynomials after multiplication may be a non-trivial multiple of the resultant of the original polynomials. To eliminate the extraneous factors and to extract the sparse resultant from this determinant is a difficult and painful process. Unlike classical resultants, which are typically represented as single determinants, sparse resultants are generally represented as GCDs (greatest common divisors) of a collection of determinants. Computing these huge GCDs is extremely awkward using symbolic computation.

Zube [65] proposed two necessary conditions on the multiplying set for constructing bi-variate sparse resultants for three polynomials $f(s, t), g(s, t), h(s, t)$ as single determinants. Define the monomial support of a polynomial (or a system of polynomials) to be the set of integer lattice points representing the exponents. If $A \subset \mathbb{Z}^2$ is the monomial support of $f(s, t), g(s, t), h(s, t)$, then the multiplying set $C_A \subset \mathbb{Z}^2$
needs to satisfy

\[ \#(A + C_A) = 3(#C_A), \quad 2 \cdot \text{Area(Conv}(A)) = #C_A, \]  

(4.1)

where Conv(A) denotes the convex hull (in \(\mathbb{R}^2\)) of \(A\), \#C_A denotes the number of points in \(C_A\), and \(A + C_A\) denotes the Minkowski sum of \(A\) and \(C_A\) [21]. The number of polynomials after multiplication is \(3(#C_A)\) and the number of monomials in all these \(3(#C_A)\) polynomials is \(#(A + C_A)\). So the first condition guarantees that the coefficient matrix of the polynomials after multiplication is a square matrix. On the other hand, the sparse resultant of \(f(s, t), g(s, t), h(s, t)\) is known to be an irreducible homogeneous polynomial in the coefficients of \(f(s, t), g(s, t), h(s, t)\), and the degree of the sparse resultant in the coefficients of each of the polynomials \(f(s, t), g(s, t), h(s, t)\) is known to be \(2 \cdot \text{Area(Conv}(A))\) [21]. Thus the second condition guarantees that the determinant of the coefficient matrix of the polynomials after multiplication has the correct degree in the coefficients of \(f(s, t), g(s, t), h(s, t)\). If this coefficient matrix is generically non-singular, then the determinant of this coefficient matrix is guaranteed to be a sparse resultant for the original system, since clearly this determinant vanishes when \(f(s, t), g(s, t), h(s, t)\) have a common root.

For many interesting rational surfaces, the monomial support \(A\) of the defining parametric polynomials is a rectangle modified by cutting off rectangular corners. For example, multi-sided Bézier patches can be generated from such monomial supports [59]. Additional interesting examples can be found in [65] — see too Section 4.4. In this chapter we show how to construct a multiplying set \(C_A\) that satisfies (4.1) for such a monomial support \(A\). Moreover, we prove that the resulting coefficient matrix, which we call the Sylvester A-resultant matrix, is generically non-singular. Thus the determinant of this coefficient matrix is indeed a sparse resultant for polynomials with monomial support \(A\). This sparse resultant can be readily adapted to implicitize those surfaces discussed in [59] [65].

Recently, Chionh [7] gave an alternative construction for the resultant of three bivariate polynomials with rectangular corners cut from their monomial support by
modifying the Dixon/Cayley construction. Starting from the classical Dixon resultant matrices [25] — see too Chapter 3, Section 3.1.2 — Chionh identifies certain submatrices that form the Dixon $A$-resultant for three such polynomials. Chionh's construction leads to smaller and more compact matrices with more complicated entries. We will discuss the Dixon $A$-resultant in the second part of this chapter. As an application of these sparse resultants, we shall use sparse resultants to implicitize the surfaces discussed in [65].

4.1 Construction of Sylvester $A$-resultants

In this section, we first briefly review the Sylvester resultant matrix for three bivariate tensor product polynomials of bi-degree $(m, n)$ — see too Chapter 3, Section 3.1.1. We then introduce the Sylvester $A$-resultant for three tensor product polynomials with rectangular corners cut from their monomial support.

4.1.1 Bi-degree Sylvester Resultant Matrices

For three bivariate polynomials of bi-degree $(m, n)$

\[
f(s, t) = \sum_{i=0}^{m} \sum_{j=0}^{n} a_{i,j} s^i t^j, \quad g(s, t) = \sum_{i=0}^{m} \sum_{j=0}^{n} b_{i,j} s^i t^j, \quad h(s, t) = \sum_{i=0}^{m} \sum_{j=0}^{n} c_{i,j} s^i t^j.
\]

the Sylvester resultant matrix for $f(s, t), g(s, t), h(s, t)$ is constructed using Sylvester's dialytic method. Consider the $6mn$ polynomials

\[
\{s^i t^j, s^i t^j g, s^i t^j h \mid 0 \leq i \leq 2m - 1, \ 0 \leq j \leq n - 1\}.
\]

This system of polynomials can be written in matrix notation as

\[
\begin{bmatrix}
f & g & h & \cdots & t^{n-1}(f & g & h) & \cdots & s^{2m-1}(f & g & h) & \cdots & s^{3n-1}t^{n-1}(f & g & h) \\
1 & \cdots & t^{2n-1} & \cdots & s^{3m-1} & \cdots & s^{3m-1}t^{2n-1}
\end{bmatrix} \cdot Syl(f, g, h).
\]

The coefficient matrix $Syl(f, g, h)$ is a square matrix of order $6mn$, and the Sylvester resultant for $f, g, h$ is simply the determinant $|Syl(f, g, h)|$ [14] [25].
The monomial support of \( f(s, t), g(s, t), h(s, t) \) is the rectangular lattice

\[
A_{m,n} = \{0, \ldots, m\} \times \{0, \ldots, n\} \subset \mathbb{Z}^2.
\]

The monomials that multiply \( f(s, t), g(s, t), h(s, t) \) are \( s^it^j \), \( 0 \leq i \leq 2m - 1, 0 \leq j \leq n - 1 \). The exponents of these monomials form another rectangular lattice

\[
C_{m,n} = \{0, 1, \ldots, 2m - 1\} \times \{0, 1, \ldots, n - 1\}
\]

We shall call both the lattice \( C_{m,n} \) and the monomials \( \{s^it^j \mid (i, j) \in C_{m,n}\} \) the multiplying set; our exact meaning will be clear from the context.

In the following figure we represent both the monomial support \( A_{m,n} \) and the multiplying set \( C_{m,n} \). Note that this figure is really only a schematic representation — \( A_{m,n} \) and \( C_{m,n} \) actually contain only the lattice points inside these rectangles. Many of the subsequent figures in the remainder of this chapter are similar schematic representations.

![Diagram of \( A_{m,n} \) and \( C_{m,n} \)](image)

Figure 4.1: The monomial support \( A_{m,n} \) and the multiplying set \( C_{m,n} \).

### 4.1.2 Rectangular Corner Cut Sylvester \( A \)-resultant Matrices

When rectangular corners \( S_1, S_2, S_3, S_4 \) are cut off from \( A_{m,n} \) (Figure 4.2) — that is, when the monomial support of \( f(s, t), g(s, t), h(s, t) \) is \( A = A_{m,n} \setminus (S_1 \cup S_2 \cup S_3 \cup S_4) \) — the Sylvester resultant \( |Syl(f, g, h)| \equiv 0 \). If, however, \( S_1, S_2, S_3, S_4 \) are disjoint, then we can still construct a Sylvester \( A \)-resultant matrix such that the determinant gives the resultant of the three original polynomials.
Let $\mathcal{C}_A = \mathcal{C}_{m,n} \setminus (\mathcal{S}_1 \cup \mathcal{S}_2 \cup \mathcal{S}_3 \cup \mathcal{S}_4)$, where

$\mathcal{S}_2' = (m - 1, 0) + \mathcal{S}_2$, \hspace{1em} $\mathcal{S}_3' = (m - 1, -1) + \mathcal{S}_4$, \hspace{1em} $\mathcal{S}_4' = (0, -1) + \mathcal{S}_4$.

as illustrated in Figure 4.2:

![Diagram](image)

Figure 4.2: Cut rectangular corners off $\mathcal{A}_{m,n}$, and cut the same rectangles at the corresponding corners off $\mathcal{C}_{m,n}$.

If we multiply $f(s,t), g(s,t), h(s,t)$ by the monomials $s^i t^j$, $(i, j) \in \mathcal{C}_A$, then we get $3(\#\mathcal{C}_A)$ polynomials. The number of monomials in these polynomials is $\#(\mathcal{A} + \mathcal{C}_A)$, where $\mathcal{A} + \mathcal{C}_A$ is the Minkowski sum of $\mathcal{A}$ and $\mathcal{C}_A$. Let $Syl_A(f, g, h)$ be the coefficient matrix of these polynomials — each column of $Syl_A(f, g, h)$ consists of the coefficients of one of the $3(\#\mathcal{C}_A)$ polynomials. The sum $\mathcal{A} + \mathcal{C}_A$ is illustrated in Figure 4.3.

From Figure 4.3 it is easy to see that

$$\#(\mathcal{A} + \mathcal{C}_A) = \#(\mathcal{A}_{m,n} + \mathcal{C}_{m,n}) - 3(\#\mathcal{S}_1) - 3(\#\mathcal{S}_2) - 3(\#\mathcal{S}_3) - 3(\#\mathcal{S}_4)$$

$$= 3(\#\mathcal{C}_{m,n} - \#\mathcal{S}_1 - \#\mathcal{S}_2 - \#\mathcal{S}_3 - \#\mathcal{S}_4)$$

$$= 3(\#\mathcal{C}_A).$$

Therefore, $Syl_A(f, g, h)$ is a square matrix of order $3(\#\mathcal{C}_A)$.

Example 4.1

When $m = n = 2$, and

$$f(s,t) = a_{0,1} t + a_{1,0} s + a_{1,1} st + a_{1,2} st^2 + a_{2,1} s^2 t,$$

$$g(s,t) = b_{0,1} t + b_{1,0} s + b_{1,1} st + b_{1,2} st^2 + b_{2,1} s^2 t,$$

$$h(s,t) = c_{0,1} t + c_{1,0} s + c_{1,1} st + c_{1,2} s t^2 + c_{2,1} s^2 t.$$
then the monomial support is

\[ \mathcal{A} = \{(0.1), (1.0), (1.1), (1.2), (2.1)\} = \mathcal{A}_{2,2} \setminus \{(0.0), (2.0), (2.2), (0.2)\}. \]

and the multiplying set is

\[ \mathcal{C}_A = \{(1.0), (1.1), (2.0), (2.1)\} = \mathcal{C}_{2,2} \setminus \{(0.0), (3.0), (3.1), (0.1)\}. \]

The sets \( \mathcal{A} \) and \( \mathcal{C}_A \) are illustrated in Figure 4.4.

\[ \begin{array}{ccc}
\times & \bullet & \times \\
\bullet & \bullet & \bullet \\
\times & \bullet & \times \\
& \times & \bullet & \times
\end{array} \]

\[ \mathcal{A} \quad \mathcal{C}_A \]

Figure 4.4: An example of the monomial support \( \mathcal{A} \) and the multiplying set \( \mathcal{C}_A \). In the diagram "\( \bullet \)" represents elements in the sets \( \mathcal{A} \) and \( \mathcal{C}_A \) and "\( \times \)" represents elements of the corners removed from \( \mathcal{A}_{2,2} \) and \( \mathcal{C}_{2,2} \).
The Sylvester $\mathcal{A}$-resultant matrix $\text{Syl}_{\mathcal{A}}(f,g,h)$ is

$$
\begin{bmatrix}
    a_{0,1} & b_{0,1} & c_{0,1} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
    0 & 0 & 0 & a_{0,1} & b_{0,1} & c_{0,1} & 0 & 0 & 0 & 0 & 0 & 0 \\
    a_{1,0} & b_{1,0} & c_{1,0} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
    a_{1,1} & b_{1,1} & c_{1,1} & a_{1,0} & b_{1,0} & c_{1,0} & 0 & 0 & 0 & 0 & 0 & 0 \\
    a_{1,2} & b_{1,2} & c_{1,2} & a_{1,1} & b_{1,1} & c_{1,1} & 0 & 0 & 0 & 0 & 0 & 0 \\
    0 & 0 & 0 & a_{1,2} & b_{1,2} & c_{1,2} & 0 & 0 & 0 & 0 & 0 & 0 \\
    0 & 0 & 0 & 0 & 0 & a_{1,2} & b_{1,2} & c_{1,2} & 0 & 0 & 0 & 0 \\
    a_{2,1} & b_{2,1} & c_{2,1} & 0 & 0 & 0 & a_{1,1} & b_{1,1} & c_{1,1} & a_{1,0} & b_{1,0} & c_{1,0} \\
    0 & 0 & 0 & a_{2,1} & b_{2,1} & c_{2,1} & a_{1,1} & b_{1,1} & c_{1,1} & 0 & 0 & 0 \\
    0 & 0 & 0 & 0 & 0 & 0 & a_{1,2} & b_{1,2} & c_{1,2} & 0 & 0 & 0 \\
    0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & a_{2,1} & b_{2,1} & c_{2,1} \\
    0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & a_{2,1} & b_{2,1} & c_{2,1}
\end{bmatrix}
$$


4.2 Sylvester $\mathcal{A}$-resultants

In this section, we prove the following

**Theorem 4.1** Let $f(s,t)$, $g(s,t)$, $h(s,t)$ be three bi-degree polynomials whose monomial support is $\mathcal{A} = \mathcal{A}_{m,n}\setminus (S_1 \cup S_2 \cup S_3 \cup S_4)$. If $S_1$, $S_2$, $S_3$, $S_4$ are disjoint, then $|\text{Syl}_{\mathcal{A}}(f,g,h)|$ is the resultant of the polynomials $f(s,t)$, $g(s,t)$, $h(s,t)$.

It is known that the resultant of $f(s,t), g(s,t), h(s,t)$ is an irreducible homogeneous polynomial in the coefficients of $f(s,t), g(s,t), h(s,t)$, and the degree of the resultant in the coefficients of $f(s,t)$ (or $g(s,t), h(s,t)$) is $2 \cdot \text{Area(Conv(A))}$ [21]. Therefore, to prove $|\text{Syl}_{\mathcal{A}}(f,g,h)|$ is indeed the resultant of $f(s,t), g(s,t), h(s,t)$, we need to show that:

- $|\text{Syl}_{\mathcal{A}}(f,g,h)|$ vanishes when $f(s,t), g(s,t), h(s,t)$ have a common zero in $(\mathbb{C}^*)^2$. where $\mathbb{C}^* = \mathbb{C}\setminus \{0\}$, so $|\text{Syl}_{\mathcal{A}}(f,g,h)|$ is a multiple of the resultant:
• $|Syl_A(f, g, h)|$ is homogeneous of degree $2 \cdot \text{Area}(\text{Conv}(A))$ in the coefficients of $f(s, t)$ (or $g(s, t), h(s, t)$), so $|Syl_A(f, g, h)|$ has the same degree as the resultant:

• $|Syl_A(f, g, h)|$ does not vanish identically, so $|Syl_A(f, g, h)|$ is the resultant (up to a non-zero multiple).

It is easy to see that the first bullet is true because the rows of the $Syl_A(f, g, h)$ are linearly dependent when $f(s, t), g(s, t), h(s, t)$ have a common root. For the second bullet, observe that since

$$2 \cdot \text{Area}(\text{Conv}(A)) = 2mn - \#S_1 - \#S_2 - \#S_3 - \#S_4 = \#C_A,$$

$|Syl_A(f, g, h)|$ is indeed homogeneous of degree $2 \cdot \text{Area}(\text{Conv}(A))$ in the coefficients of $f(s, t)$ (or $g(s, t), h(s, t)$). Thus to complete the proof, we need only show that $|Syl_A(f, g, h)|$ is not identically zero. We now proceed to show this is so by cutting off one corner at a time.

### 4.2.1 Only the Upper Right Corner is Cut Off

We will start with the monomial support $A$ where only the upper right corner is cut off. In this case, $A, C_A$ and $A + C_A$ are illustrated in Figure 4.5.

Consider the polynomials

$$f^*(s, t) = f(s, t) - x, \quad g^*(s, t) = g(s, t) - y, \quad h^*(s, t) = h(s, t) + \lambda s^m t^n - z.$$

where $x, y, z$ are three indeterminates. The role of $x, y, z$ will become clear soon.

If we regard $x, y, z$ as constants, then the convex hull of the monomial support of $f^*(s, t), g^*(s, t), h^*(s, t)$ is the rectangular lattice $A_{m,n}$. The Sylvester resultant
Figure 4.5: The sets $\mathcal{A}, \mathcal{C}_A, \mathcal{A} + \mathcal{C}_A$ when the upper right corner is cut off.

The matrix $Syl(f^*, g^*, h^*)$ can be written as

$$
\begin{pmatrix}
\{f^*, g^*, h^*\} \cdot \mathcal{C}_A & \{f^*, g^*, h^*\} \cdot \mathcal{S}_3' \\

\end{pmatrix}
$$

$$
\begin{pmatrix}
\mathbb{I} L & * \\

\end{pmatrix}
$$

$$
\begin{pmatrix}
\Omega_1 & 0 0 \lambda I 0 0 & * * * \\
\Omega_2 & 0 0 0 \lambda I 0 & * * * \\
\Omega_3 & (f^*) \cdot (g^*) \cdot (h^*) & (f^*) \cdot (g^*) \cdot (h^*)
\end{pmatrix}
$$

where $I$ is the identity matrix of order $\#S_3$. In the rows indexed by $\Omega_1, \Omega_2, \Omega_3$, the first column of zeros comes from the coefficients of $f^* \cdot \mathcal{C}_A$, the second column of zeros
from the coefficients of \(g^* \cdot \mathcal{C}_A\), and the next three columns from the coefficients of \(h^* \cdot \mathcal{C}_A\). The last three columns come from the coefficients of \(f^* \cdot \mathcal{S}_3\), \(g^* \cdot \mathcal{S}_3\), \(h^* \cdot \mathcal{S}_3\) respectively.

Notice that the upper left block \(UL\) of \(Syl(f^*, g^*, h^*)\) (indexed by \(\{f^*, g^*, h^*\} \cdot \mathcal{C}_A\)
and \(\mathcal{A} + \mathcal{C}_A\)) is exactly \(Syl_A(f - x, g - y, h - z)\) if \(\lambda\) is set to zero. Further, notice that the entries marked with \(*\) in the lower right block (indexed by \(\{f^*, g^*, h^*\} \cdot \mathcal{S}_3\)
and \(\Omega_1, \Omega_2\)) do not involve \(\lambda\).

Now we know that \(|Syl(f^*, g^*, h^*)|\) is not identically zero as a polynomial in \(\lambda\).
In fact, choosing \(\lambda = 1, x = y = z = 0\) and \(f^*(s, t) = g^*(s, t) = s^m + t^n\),
\(h^*(s, t) = s^m t^n\), the absolute value of \(|Syl(f^*, g^*, h^*)|\) is 1.

From the last \#\(\mathcal{S}_3\) rows (indexed by \(\Omega_3\)) of \(Syl(f^*, g^*, h^*)\), we see that the determinant \(|Syl(f^*, g^*, h^*)|\) contains \(\lambda^{\#\mathcal{S}_3}\) as a factor. We can write \(|Syl(f^*, g^*, h^*)|\) as a polynomial in \(\lambda\):

\[|Syl(f^*, g^*, h^*)| = c_d \cdot \lambda^d + c_{d+1} \cdot \lambda^{d+1} + \cdots.\]

where \(d\) is the smallest degree in \(\lambda\), so \(d \geq \#\mathcal{S}_3\). The coefficients \(c_d, c_{d+1}, \ldots\) are polynomials in \(x, y, z\) and the coefficients of \(f(s, t), g(s, t), h(s, t)\). In particular, \(c_d\) is
of degree \(2mn - d \leq 2mn - \#\mathcal{S}_3\) in the coefficients of \(h(s, t) - z\), since \(|Syl(f^*, g^*, h^*)|\)
is homogeneous of degree \(2mn\) in the coefficients of \(h^*(s, t)\).

On the other hand, the implicit equation of the surface

\[x = f(s, t), \quad y = g(s, t), \quad z = h(s, t) + \lambda s^m t^n.\] (4.3)
is exactly \(P(x, y, z) \equiv |Syl(f^*, g^*, h^*)| = 0.\)

If we write

\[c_d(x, y, z) = \sum T_{\alpha, \beta, \gamma} x^\alpha y^\beta z^\gamma.\]

then

\[P(f(s, t), g(s, t), h(s, t) + \lambda s^m t^n)\]

\[= c_d(f, g, h + \lambda s^m t^n) \lambda^d + \cdots\]

\[= \sum T_{\alpha, \beta, \gamma} f^\alpha g^\beta (h + \lambda s^m t^n)^\gamma \lambda^d + \cdots.\] (4.4)
Since all the "\ldots" terms on the right hand side of Equation (4.4) have degree greater than \(d\) in \(\lambda\), the lowest degree of \(\lambda\) in \(P(f(s,t), g(s,t), h(s,t) + \lambda s^m t^n)\) is \(d\). Moreover, the coefficient of \(\lambda^d\) in \(P(f(s,t), g(s,t), h(s,t) + \lambda s^m t^n)\) is

\[c_d(f(s,t), g(s,t), h(s,t)) = \sum T_{\alpha, \beta} \cdot f^{\alpha} g^{\beta} h^\gamma.\]

Since \(P(x, y, z) = 0\) is the implicit equation of surface (4.3),

\[P(f(s,t), g(s,t), h(s,t) + \lambda s^m t^n) \equiv 0.\]

Therefore,

\[c_d(f(s,t), g(s,t), h(s,t)) \equiv 0.\]

That is, \(c_d\) contains the implicit equation of the surface

\[x = f(s,t), \quad y = g(s,t), \quad z = h(s,t),\]

as a factor. Hence, \(c_d\) contains the resultant of \(f(s,t) - x, g(s,t) - y, h(s,t) - z\) as a factor. Since this resultant is of degree \(2 \cdot \text{Area}(	ext{Conv}(A)) = 2mn - \#S_3\) in the coefficients of \(h(s,t) - z\) [21], it follows that the degree of \(c_d\) in the coefficients of \(h(s,t) - z\) is no less than \(2mn - \#S_3\). Therefore, \(d = \#S_3\) and \(c_d\) is the same as \(|Syl(f - x, g - y, h - z)|\) up to a nonzero constant. As a consequence, \(c_{\neq S_3}\) is not identically zero.

Expanding \(|Syl(f^*, g^*, h^*)|\) with respect to the last \(3(\#S_3)\) rows, we see that the terms with the lowest degree \(\#S_3\) in \(\lambda\) come from the upper left sub-determinant \(|U^\prime L|\) and the bottom right sub-determinant (of order \(3(\#S_3)\)). Since the bottom right sub-determinant contains \(\lambda^{\#S_3}\) as a factor, \(|U^\prime L|\) must have a nonzero constant when viewed as a polynomial in \(\lambda\). Thus, \(|U^\prime L|\) is nonzero when \(\lambda = 0\). That is, \(|Syl_{\lambda}(f - x, g - y, h - z)|\) does not vanish identically. Regarding \(x, y, z\) as constants, we observe that \(f(s,t) - x, g(s,t) - y, h(s,t) - z\) also have the monomial support \(A = A_{m,n} \setminus S_3\). Therefore, \(|Syl_{\lambda}(f, g, h)|\) is not identically zero when the monomial support of \(f(s,t), g(s,t), h(s,t)\) is \(A = A_{m,n} \setminus S_3\).
4.2.2 Cut Off One More Corner: The Upper Left Corner

Now assume that $A = A_{m,n} \setminus (S_3 \cup S_4)$ is the monomial support of $f(s,t), g(s,t)$, $h(s,t)$, and choose $C_A = C_{m,n} \setminus (S_3' \cup S_4')$ (Figure 4.6).

![Figure 4.6: The sets $A$, $C_A$, $A + C_A$ when the upper right and upper left corners are cut off.]

Consider the polynomials

$$f^*(s,t) = f(s,t) - x, \quad g^*(s,t) = g(s,t) - y, \quad h^*(s,t) = h(s,t) + \lambda t^n - z.$$ 

We know that the resultants of $f^*, g^*, h^*$ with respect to the monomial support $(A_{m,n} \setminus (S_3 \cup S_4)) \cup \{(0,n)\}$ or the monomial support $A_{m,n} \setminus S_3$ are the same, because these two monomial supports have the same convex hulls. Therefore, by Section 4.2.1, the resultant of $f^*, g^*, h^*$ is given by the Sylvester $A$-resultant $Syl_A(f^*, g^*, h^*)$. Thus $|Syl_A(f^*, g^*, h^*)| = 0$ is the implicit equation of the surface

$$x = f(s,t), \quad y = g(s,t), \quad z = h(s,t) + \lambda t^n.$$
Write $Syl_{\mathcal{A}}(f^*, g^*, h^*)$ explicitly as

$$
\mathcal{A} + \mathcal{C}_{\mathcal{A}} \begin{bmatrix}
\{f^*, g^*, h^*\} \\
\{f^*, g^*, h^*\}
\end{bmatrix}
\begin{bmatrix}
U L \\
* 
\end{bmatrix}
\begin{bmatrix}
\Omega_1 \\
\Omega_2 \\
\Omega_3
\end{bmatrix}

By an argument similar to the proof in Section 4.2.1, we can show that the Sylvester $\mathcal{A}$-resultant $|Syl_{\mathcal{A}}(f - x, g - y, h - z)|$ is not identically zero. Again, regarding $x, y, z$ as constants, we see that $|Syl_{\mathcal{A}}(f, g, h)|$ does not vanish identically when $\mathcal{A} = \mathcal{A}_{m,n} \setminus (\mathcal{S}_1 \cup \mathcal{S}_4)$.

4.2.3 Cut Off One More Corner: The Lower Right Corner

Now assume that the monomial support is $\mathcal{A} = \mathcal{A}_{m,n} \setminus (\mathcal{S}_2 \cup \mathcal{S}_3 \cup \mathcal{S}_4)$, and choose $\mathcal{C}_{\mathcal{A}} = \mathcal{C}_{m,n} \setminus (\mathcal{S}_2 \cup \mathcal{S}_3 \cup \mathcal{S}_4)$ (Figure 4.7).

Consider the polynomials

$$
f^*(s, t) = f(s, t) - x, \quad g^*(s, t) = g(s, t) - y, \quad h^*(s, t) = h(s, t) + \lambda s^m - z.
$$

Proceeding as in Section 4.2.2, we can again show that the Sylvester $\mathcal{A}$-resultant $|Syl_{\mathcal{A}}(f, g, h)|$ is not identically zero when $\mathcal{A} = \mathcal{A}_{m,n} \setminus (\mathcal{S}_2 \cup \mathcal{S}_3 \cup \mathcal{S}_4)$.

4.2.4 Cut Off All Four Corners

Finally assume that the monomial support is $\mathcal{A} = \mathcal{A}_{m,n} \setminus (\mathcal{S}_1 \cup \mathcal{S}_2 \cup \mathcal{S}_3 \cup \mathcal{S}_4)$, and choose $\mathcal{C}_{\mathcal{A}} = \mathcal{C}_{m,n} \setminus (\mathcal{S}_1 \cup \mathcal{S}_2 \cup \mathcal{S}_3 \cup \mathcal{S}_4)$ (Figure 4.8).

Consider the polynomials

$$
f^*(s, t) = f(s, t) - x, \quad g^*(s, t) = g(s, t) - y, \quad h^*(s, t) = h(s, t) - z.
$$
Then the monomial support of $f^*, g^*, h^*$ is contained in $\mathcal{A}_{m,n} \setminus (S_2 \cup S_3 \cup S_4)$.

Note that this case is different from the previous three cases because the monomial support of $f(s,t) - x \cdot g(s,t) - y \cdot h(s,t) - z$ is different from the monomial support of $f(s,t), g(s,t), h(s,t)$. We shall, therefore, need other tools to prove that $|Syl\mathcal{A}(f, g, h)|$ is not identically zero even when we know that $|Syl\mathcal{A}(f^*, g^*, h^*)|$ is not identically zero.

To proceed, notice that the resultant of $f^*, g^*, h^*$ with respect to the monomial support $(\mathcal{A}_{m,n} \setminus (S_1 \cup S_2 \cup S_3 \cup S_4)) \cup \{(0,0)\}$ is given by the Sylvester $\mathcal{A}$-resultant $|Syl\mathcal{A}(f^*, g^*, h^*)|$ with respect to the monomial support $\mathcal{A}_{m,n} \setminus (S_2 \cup S_3 \cup S_4)$ because these two monomial supports have the same convex hulls. Thus

$$|Syl\mathcal{A}(f^*, g^*, h^*)| = 0$$

is the implicit equation of the surface

$$x = f(s,t), \quad y = g(s,t), \quad z = h(s,t).$$

(4.5)
Figure 4.8: The sets $\mathcal{A}, \mathcal{C}_\mathcal{A}, \mathcal{A} - \mathcal{C}_\mathcal{A}$ when all four corners are cut off.

The intersections of surface (4.5) and a generic line

\[
\begin{cases} 
A_1 x + B_1 y + C_1 z = 0, \\
A_2 x + B_2 y + C_2 z = 0.
\end{cases}
\] (4.6)

that passes through the origin can be obtained by solving the system

\[
\begin{cases} 
A_1 \cdot f(s, t) + B_1 \cdot g(s, t) + C_1 \cdot h(s, t) = 0, \\
A_2 \cdot f(s, t) + B_2 \cdot g(s, t) + C_2 \cdot h(s, t) = 0.
\end{cases}
\] (4.7)

The number of generic solutions of this system (intersections at the origin) — counting multiplicities — is exactly $\# S_1$ [59].

Write the implicit equation $|Syl_A(f^*, g^*, h^*)| = 0$ of surface (4.5) as

\[
|Syl_A(f^*, g^*, h^*)| = \sum_{i \geq d} P_i,
\]

where $P_i$ is the homogeneous component of degree $i$ in $x, y, z$ and $d$ is the lowest degree of all the components. Then the number of intersections (at the origin) of
surface (4.5) and a generic line that passes through the origin is \(d\). It follows that \(d = \#S_1\).

The Sylvester \(A\)-resultant matrix \(Syl_A(f^*, g^*, h^*)\) can be written explicitly as:

\[
\begin{pmatrix}
\{f^* \cdot g^* \cdot h^*\} \cdot C_A \\
\{f^* \cdot g^* \cdot h^*\} \cdot S_1
\end{pmatrix}
\]

\[
A + C_A \begin{pmatrix}
\begin{pmatrix}
-\varepsilon I & 0 & 0 & -yI & 0 & 0 & -zI & 0 & 0 \\
0 & -\varepsilon I & 0 & 0 & -yI & 0 & 0 & -zI & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & -\varepsilon I & -yI & -zI
\end{pmatrix}
\end{pmatrix}
\]

Note that the upper left submatrix \(UL\) is \(Syl_A(f, g, h)\) when \(x = y = z = 0\). Also notice that the entries marked by \(\ast\) in the lower right block (indexed by \(\{f^* \cdot g^* \cdot h^*\} \cdot S_1\) and \(\Omega_1, \Omega_2\)) do not involve \(x, y, z\).

Expanding this determinant with respect to the last \(3(\#S_1)\) rows (indexed by \(\Omega_1, \Omega_2, \Omega_3\)), we see that each term in \(|Syl_A(f^*, g^*, h^*)|\) is of degree \(\geq \#S_1\) in \(x, y, z\). Furthermore, the only terms that are of degree exactly \(\#S_1\) in \(x, y, z\) come from the upper left sub-determinant \(|UL|\) and the lower right sub-determinant (of order \(3(\#S_1)\)). Since the lower right sub-determinant is homogeneous in \(x, y, z\) and of degree \(\#S_1\), it follows that \(|UL|\) (regarded as a polynomial in \(x, y, z\)) must have a non-zero constant term. Thus, when \(x, y, z\) are set to 0, \(|UL|\) is nonzero. That is, \(|Syl_A(f, g, h)|\) is not identically zero when all four corner rectangles are cut off.

The proof of Theorem 4.1 at the beginning of this section is now complete.

### 4.3 Remarks on Sylvester \(A\)-resultants

We have just shown how to construct the Sylvester \(A\)-resultant matrix for three bivariate tensor product polynomials with rectangular corners cut off from their monomial
support. We have also proved that the determinant of this $\mathcal{A}$-resultant matrix does indeed give the resultant of the three bivariate polynomials. Looking back over these results, we can now make the following additional observations.

1) The proof of the Theorem 4.1 proceeded as the corner rectangles were cut off in a particular order. In fact, the corner rectangles are independent. The theorem is also valid when any one, two, or three corner rectangles are cut off.

2) Points and lines are also rectangles. In fact, lines and points are rectangles with one or both dimensions set to zero. Rectangular monomial supports with points or lines cut off at the corners are frequently seen in surface design [65]. Thus the Sylvester $\mathcal{A}$-resultant could be very useful in geometric modeling.

3) The choice of the multiplying set $\mathcal{C}_A$ is not unique. We chose $\mathcal{C}_A$ to be $\mathcal{C}_{m,n} = \{0, \ldots, 2m - 1\} \times \{0, \ldots, n - 1\}$ with the corresponding rectangular corners cut off. We could also have started with $\mathcal{R}_{m,n} = \{0, \ldots, m - 1\} \times \{0, \ldots, 2n - 1\}$ and chosen $\mathcal{R}_A$ to be $\mathcal{R}_{m,n}$ with the corresponding rectangular corners cut off (Figure 4.9). The proof is much the same. Note that the sizes of $\mathcal{R}_A$ and $\mathcal{C}_A$ are the same, but their shapes can be quite different (see Figure 4.10). Indeed although the shapes of $\mathcal{R}_{m,n}$ and $\mathcal{C}_{m,n}$ look symmetric, the shapes of the multiplying sets $\mathcal{R}_A$ from $\mathcal{R}_{m,n}$ and $\mathcal{C}_A$ from $\mathcal{C}_{m,n}$ need not be symmetric. For example, the two multiplying sets given by Zube [65] for the Hirzebruch surface look very different as we can see in Figure 4.10.

4.4 Construction of Dixon $\mathcal{A}$-resultants

In this section we describe the construction of Dixon $\mathcal{A}$-resultant matrices. These resultants are obtained from the classical Dixon resultants by omitting an equal number of rows and columns of zero entries. The main observation is that just like Sylvester $\mathcal{A}$-resultant, Dixon $\mathcal{A}$-resultants arise when the monomial support $\mathcal{A}$ is obtained by
removing disjoint rectangular sub-supports from each of the corners of a rectangular bi-degree support. For example, the following support \( A \)

\[
\begin{array}{c}
  \bullet \\
  \bullet \\
  \bullet \\
  \bullet \\
  \bullet
\end{array}
\]

admits a Dixon \( A \)-resultant, as well as a Sylvester \( A \)-resultant, because it is obtained by removing the four corner "rectangular" sub-supports \{\{(0,0)\}, \{(3,0)\}, \{(3,3)\}, \{(0,3)\}\} from the rectangular support \( A_{3,3} = \{0,1,2,3\} \times \{0,1,2,3\} \). Note that a
rectangle can degenerate to a single point or to a horizontal or vertical line segment of points.

### 4.4.1 Bi-degree Dixon Resultant Matrices

We first briefly review the construction of the classical Dixon resultant for three polynomial equations in two variables [Ch 3] [14] [25].

Let $\mathcal{A}$ be the monomial support of the polynomials

$$
\begin{align*}
  f(s, t) &= \sum_{(i, j) \in \mathcal{A}} a_{i,j}s^i t^j, \\
  g(s, t) &= \sum_{(i, j) \in \mathcal{A}} b_{i,j}s^i t^j, \\
  h(s, t) &= \sum_{(i, j) \in \mathcal{A}} c_{i,j}s^i t^j.
\end{align*}
$$

Then the Cayley expression for $f, g, h$ is

$$
\Delta_\mathcal{A}(f(s, t), g(\alpha, t), h(\alpha, \beta)) = \frac{
\begin{vmatrix}
  f(s, t) & g(s, t) & h(s, t) \\
  f(\alpha, t) & g(\alpha, t) & h(\alpha, t) \\
  f(\alpha, \beta) & g(\alpha, \beta) & h(\alpha, \beta)
\end{vmatrix}
}{(s - \alpha)(t - \beta)}.
\tag{4.9}
$$

Note that the denominator divides the numerator because the numerator vanishes when $s = \alpha$ or $t = \beta$. The expression $\Delta_\mathcal{A}$ can be viewed as a polynomial in either $s, t$ or $\alpha, \beta$ to suit the need. The monomial supports of $\Delta_\mathcal{A}$ in $s, t$ and $\alpha, \beta$ are denoted by $\mathcal{R}_\mathcal{A}$ and $\mathcal{C}_\mathcal{A}$ respectively. We can also write $\Delta_\mathcal{A}$ in matrix form with row indices \([s^\sigma t^\tau \mid (\sigma, \tau) \in \mathcal{R}_\mathcal{A}]\) and column indices \([\alpha^a \beta^b \mid (a, b) \in \mathcal{C}_\mathcal{A}]\) as

$$
\Delta_\mathcal{A} = \begin{bmatrix}
\vdots \\
 s^\sigma t^\tau \\
\vdots
\end{bmatrix}^T \begin{bmatrix}
\vdots \\
Dix_\mathcal{A}(f, g, h) \\
\vdots
\end{bmatrix} \begin{bmatrix}
\vdots \\
\alpha^a \beta^b \\
\vdots
\end{bmatrix}.
\tag{4.10}
$$

$Dix_\mathcal{A}(f, g, h)$ is the Dixon resultant matrix of $f, g, h$. Instead of calling $\mathcal{R}_\mathcal{A}$ and $\mathcal{C}_\mathcal{A}$ the monomial supports of $\Delta_\mathcal{A}$, we shall call $\mathcal{R}_\mathcal{A}$ the row support of $Dix_\mathcal{A}(f, g, h)$ and $\mathcal{C}_\mathcal{A}$ the column support of $Dix_\mathcal{A}(f, g, h)$. 

Since the numerator in (4.9)

\[
\begin{vmatrix}
  f(s, t) & g(s, t) & h(s, t) \\
  f(\alpha, t) & g(\alpha, t) & h(\alpha, t) \\
  f(\alpha, \beta) & g(\alpha, \beta) & h(\alpha, \beta)
\end{vmatrix}
= \sum_{(i, j), (k, l), (p, q) \in A} |i, j; k, l; p, q| s^{i+1} t^{l+1} \alpha^{k+p} \beta^{q+q}
\]

(4.11)

where

\[
|i, j; k, l; p, q| = \begin{vmatrix}
  a_{i,j} & b_{i,j} & c_{i,j} \\
  a_{k,l} & b_{k,l} & c_{k,l} \\
  a_{p,q} & b_{p,q} & c_{p,q}
\end{vmatrix}
\]

(4.12)

the entries of \(Dis_A(f, g, h)\) are linear in the coefficients of each of \(f\), \(g\), and \(h\).

For economy of space, we abbreviate \(|i, j; k, l; p, q|\) to \(ijklpq\) when there is no danger of confusion. For example,

\[
000110 = |0, 0; 0, 1; 1, 0| = \begin{vmatrix}
  a_{0,0} & b_{0,0} & c_{0,0} \\
  a_{0,1} & b_{0,1} & c_{0,1} \\
  a_{1,0} & b_{1,0} & c_{1,0}
\end{vmatrix}
\]

Let \(A_{m,n} = \{0, \ldots, m\} \times \{0, \ldots, n\}\) be the bi-degree \((m, n)\) monomial support. When \(A = A_{m,n}\) we write \(\Delta_A\), \(Dis_A(f, g, h)\), \(R_A\), \(C_A\) as \(\Delta_{m,n}\), \(Dis_{m,n}(f, g, h)\), \(R_{m,n}\), \(C_{m,n}\) respectively. Clearly \(\Delta_{m,n}\) is of degree \(m - 1\) in \(s\), \(2n - 1\) in \(t\), \(2m - 1\) in \(\alpha\), and \(n - 1\) in \(\beta\). Thus

\[
\begin{align*}
R_{m,n} &= \{0, \ldots, m - 1\} \times \{0, \ldots, 2n - 1\}.
\end{align*}
\]

(4.13)

Consequently, \(Dis_{m,n}(f, g, h)\) is a square matrix of order \(2mn\) — that is, \(#R_{m,n} = #C_{m,n} = 2mn\). The determinant \(|Dis_{m,n}(f, g, h)|\) is the classical Dixon resultant of \(f, g, h\) when \(A = A_{m,n}\).

### 4.4.2 Corner Cut Dixon \(A\)-resultant Matrices

This section gives the row and column supports \(R_A\) and \(C_A\) for a monomial support \(A\) that is a rectangle missing some disjoint corner sub-rectangles. In fact, we shall
see that $C_A$ and $R_A$ are the same as for the Sylvester $A$-resultant. In the following discussion, we assume $A \subseteq A_{m,n}$.

Let $Dix_{m,n}^A(f, g, h)$ be $Dix_{m,n}(f, g, h)$ specialized with respect to $A$. That is, $Dix_{m,n}^A(f, g, h)$ is obtained from $Dix_{m,n}(f, g, h)$ by setting $a_{i,j} = b_{i,j} = c_{i,j} = 0$ for $(i, j) \in A_{m,n} \backslash A$.

When $A_{m,n} \backslash A$ is a union of disjoint rectangular sub-supports at the corners of $A_{m,n}$, the following theorem identifies $Dix_A(f, g, h)$ as a square sub-matrix of $Dix_{m,n}^A(f, g, h)$ by giving explicitly its row support $R_A$ and column support $C_A$.

**Theorem 4.2** Let $A$ be the monomial support of $f, g, h$: let $R_A$ be the row support of $Dix_A(f, g, h)$, and $C_A$ be the column support of $Dix_A(f, g, h)$. Suppose that $S = S_1 \cup S_2 \cup S_3 \cup S_4$, where $S_1, S_2, S_3, S_4$ are respectively rectangular sub-supports at the bottom-left, bottom-right, top-right, top-left corners of $A_{m,n}$ that are possibly empty and are disjoint from each other. If $A \subseteq A_{m,n}$ and $A_{m,n} \backslash A = S$, then

\begin{align*}
C_A &= C_{A_{m,n}} \backslash (S_1 \cup S_2 \cup S_3 \cup S_4). \quad (4.14) \\
R_A &= R_{A_{m,n}} \backslash (S_1 \cup S_2 \cup S_3 \cup S_4). \quad (4.15)
\end{align*}

where

\begin{align*}
S_1 &= (m - 1, 0) + S_2, \quad S_2 = (m - 1, -1) + S_3, \quad S_3 = (0, -1) + S_4, \\
S_4 &= (-1, 0) + S_2, \quad S_3' = (-1, n - 1) + S_3, \quad S_4' = (0, n - 1) + S_4.
\end{align*}

That is, $Dix_A(f, g, h)$ is the square sub-matrix of $Dix_{m,n}^A(f, g, h)$ of order $\#C_A$ (or $\#R_A$) obtained by removing $\#S$ rows of zeros and $\#S$ columns of zeros from $Dix_{m,n}^A(f, g, h)$ [Figure 4.11].

**Proof:** A proof of this theorem can be found in [8].

For example, when $m = n = 2$, consider removing four corner monomials $1, s^2, s^2t^2, t^2$ from the nine bi-degree $(2, 2)$ monomials $s^it^j$: $i, j = 0, 1, 2$. Pictorially the
Figure 4.11: The monomial support $A$, row support $R_A$, and column support $C_A$. Compare to Figure 4.2 for Sylvester $A$-resultants.

resulting monomial, row, and column supports are

\[
\begin{array}{ccc}
  \times & \times & \\
  \times & \bullet & \times \\
  \bullet & \bullet & \times \\
  \times & \times & \bullet \\
  \end{array}
\begin{array}{ccc}
  \bullet & \bullet & \times \\
  \bullet & \bullet & \times \\
  \times & \times & \bullet \\
  \end{array}
\begin{array}{ccc}
  \times & \times & \times \\
  \times & \times & \times \\
  \times & \times & \times \\
  \end{array}
\tag{4.16}
\]

respectively. In this diagram "•" represents elements of the supports and "×" represents elements of the corner sub-supports. Indeed, by computing the Dixon polynomial (4.9) directly, or by using an entry formula [7], or by applying the recursive technique in Chapter 3, Section 3.7, the specialized Dixon matrix $D_{\text{ext}}(f, g, h)$ is found to be

\[
\begin{bmatrix}
  0 & 0 & 0 & 0 & 0 & 0 & 0 \\
  0 & 110110 & 101201 & 210110 & 0 & 0 & 0 \\
  0 & 120110 & 111201 & 0 & 211201 & 0 & 0 \\
  0 & 0 & 0 & 0 & 0 & 0 & 0 \\
  0 & 0 & 0 & 0 & 0 & 0 & 0 \\
  0 & 210110 & 0 & 211110 & 211210 & 0 & 0 \\
  0 & 0 & 0 & 211201 & 211210 & 211211 & 0 & 0 \\
  0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\end{bmatrix}
\]

with row indices (from top to bottom) $1, t, t^2, t^3, s, st, st^2, st^3$ and column indices
(from left to right) $1, 3, \alpha, \alpha^2, \alpha^3, \alpha^2, \alpha^3, \alpha^2, \alpha^3$. After removing the zero rows and zero columns, we obtain

\[
\text{Dir}_A(f, g, h) = \begin{bmatrix}
110110 & 101201 & 210110 & 0 \\
120110 & 111201 & 0 & 211201 \\
210110 & 0 & 211110 & 211210 \\
0 & 211201 & 211210 & 211211
\end{bmatrix}.
\]

(4.17)

Notice how this matrix compares in size and the complexity of its entries to the matrix $Syl_A(f, g, h)$ at the end of Section 4.1.

### 4.5 Dixon $A$-resultants

The following theorem establishes that $|\text{Dir}_A(f, g, h)|$ is indeed the resultant.

**Theorem 4.3** Let $A$ be the monomial support of $f, g, h$. If $A$ satisfies the conditions stated in Theorem 4.2, then the determinant $|\text{Dir}_A(f, g, h)|$ is the $A$-resultant of the polynomials $f, g, h$.

**Proof**: A proof of this theorem can be found in [8].

For example, if the monomial support $A$ is that of (4.16), the Dixon $A$-resultant for $f, g, h$ is

\[
|\text{Dir}_A(f, g, h)| = \begin{vmatrix}
110110 & 101201 & 210110 & 0 \\
120110 & 111201 & 0 & 211201 \\
210110 & 0 & 211110 & 211210 \\
0 & 211201 & 211210 & 211211
\end{vmatrix}.
\]

(4.18)

An $A$-resultant for this particular monomial support $A$ is also presented in [65] as the determinant of

\[
\text{Dir}_A(f, g, h)' = \begin{bmatrix}
1 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1 \\
0 & 1 & 0 & 0
\end{bmatrix} \quad \text{Dir}_A(f, g, h) = \begin{bmatrix}
1 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 \\
0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0
\end{bmatrix}.
\]
4.6 Comparison of the Sylvester and Dixon $A$-resultants

The Sylvester $A$-resultant has simple entries, while the entries of the Dixon $A$-resultant are very complicated polynomials in the coefficients of the original polynomials. On the other hand, the Sylvester $A$-resultant matrix is nine times the size of the Dixon $A$-resultant matrix, since the Sylvester $A$-resultant is of order $3(#C_A)$ whereas the Dixon $A$-resultant is of order $#C_A$. This comparison mimics the situation for bivariate tensor product polynomials of bi-degree $(m, n)$, where the Sylvester resultant matrix is of order $6mn$ and the Dixon resultant matrix is of order $2mn$. The advantage of the Dixon $A$-resultant is its compact size. The advantage of the Sylvester $A$-resultant is the simplicity of its entries, which are just the coefficients of the original polynomials. The simplicity of the entries of Sylvester-like resultant matrices may be helpful for theoretical analysis: this is certainly the case in the univariate setting [39] [54]. On the other hand, the compact size of the Dixon $A$-resultant makes it most useful for numerical computations.

4.7 Implicitization by Sylvester $A$-resultants

To implicitize the rational surface

$$
X = \frac{x(s, t)}{w(s, t)}, \quad Y = \frac{y(s, t)}{w(s, t)}, \quad Z = \frac{z(s, t)}{w(s, t)},
$$

we construct the three auxiliary equations

$$
\begin{align*}
f(s, t) &= x(s, t) - X \cdot w(s, t), \\
g(s, t) &= y(s, t) - Y \cdot w(s, t), \\
h(s, t) &= z(s, t) - Z \cdot w(s, t).
\end{align*}
$$
Let \( w_{i,j}, x_{i,j}, y_{i,j}, z_{i,j} \) be the coefficients of \( s^2 t \) in \( w(s,t), x(s,t), y(s,t), z(s,t) \) respectively. Then \( a_{i,j} = x_{i,j} - X w_{i,j}, b_{i,j} = y_{i,j} - Y w_{i,j}, c_{i,j} = z_{i,j} - Z w_{i,j} \) and consequently

\[
| i, j; k, l; p, q | = \begin{vmatrix} 1 & X & Y & Z \\ w_{i,j} & x_{i,j} & y_{i,j} & z_{i,j} \\ w_{k,l} & x_{k,l} & y_{k,l} & z_{k,l} \\ w_{p,q} & x_{p,q} & y_{p,q} & z_{p,q} \end{vmatrix}.
\]

Theorem 3.13 of [21] (page 312) ensures that, in general, the \( \mathcal{A} \)-resultant of the auxiliary equations of the parametrization gives the implicit equation of the rational surface.

**Example 4.2**

For the rational surface

\[
\begin{align*}
X &= \frac{2s^2 t + st^2 + st + s}{s^2 t + st^2 + st + s + t}, \\
Y &= \frac{s^2 t + 2st^2 + st + t}{s^2 t + st^2 + st + s + t}, \\
Z &= \frac{10st}{s^2 t + st^2 + st + s + t},
\end{align*}
\]

the auxiliary polynomials \( f, g, h \) are

\[
\begin{align*}
f(s,t) &= (X - 2)s^2 t + (X - 1)st^2 + (X - 1)st + (X - 1)s + Xt, \\
g(s,t) &= (Y - 1)s^2 t + (Y - 2)st^2 + (Y - 1)st + Ys + (Y - 1)t, \\
h(s,t) &= Zs^2 t + Zst^2 + (Z - 10)st + Zs + Zt.
\end{align*}
\]

Thus the monomial support \( \mathcal{A} \) of \( f, g, h \) is the same as in Example 4.1. Hence we can
construct the Sylvester resultant of \( f, g, h \) as

\[
\begin{bmatrix}
X & Y - 1 & Z & 0 & 0 & 0 \\
0 & 0 & 0 & X & Y - 1 & Z \\
X - 1 & Y & Z & 0 & 0 & 0 \\
X - 1 & Y - 1 & Z - 10 & X - 1 & Y & Z \\
X - 1 & Y - 2 & Z & X - 1 & Y - 1 & Z - 10 \\
0 & 0 & 0 & X - 1 & Y - 2 & Z \\
0 & 0 & 0 & 0 & 0 & 0 \\
X - 2 & Y - 1 & Z & 0 & 0 & 0 \\
0 & 0 & 0 & X - 2 & Y - 1 & Z \\
0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}
= 0.
\]

For additional examples of implicitization using the Sylvester \( A \)-resultant, see Zube [65].
4.8 Implicitization by Dixon $\mathcal{A}$-resultants

In this section we use Dixon $\mathcal{A}$-resultants to implicitize the examples given in Zube [65]. All of these examples have rectangular corners cut from the tensor product monomial supports. Some of these monomial supports can be used to generate multi-sided Bézier patches [59]. Note that the matrix generated by the Dixon approach is one-ninth the size of the matrix build by the Sylvester method and the entries of the Dixon matrix are in compact bracket form. Furthermore, there is no need to employ a multiplying set. The small size of the Dixon $\mathcal{A}$-resultant matrices makes them particularly useful for numerical computations.

4.8.1 Hirzebruch Surfaces

The monomial support of the Hirzebruch surface $F_n$ is

$$\mathcal{A} = \{(0, 0), (1, 0), \ldots, (n + 1, 0), (0, 1), (1, 1)\}$$

where $n \geq 0$. For $n = 2$, the monomial, row, and column supports of $F_2$ are respectively

$$\begin{array}{ccc}
\bullet & \bullet & \times & \times \\
\bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \times & \times \\
\mathcal{A} & \mathcal{R}_\mathcal{A} & \mathcal{C}_\mathcal{A}
\end{array}$$

The corresponding matrix $Di(x, y, h)$ with appropriate coefficients set to zero is

$$
\begin{bmatrix}
100100 & 101100 + 200100 & 201100 + 300100 & 301100 & 0 & 0 \\
110100 & 110110 & 110120 & 110130 & 0 & 0 \\
200100 & 200110 + 201100 + 300100 & 201110 + 300110 + 301100 & 301110 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 \\
300100 & 300110 + 301100 & 300120 + 301110 & 301120 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}.
$$
Thus the implicit equation of $F_2$ is

\[
\begin{bmatrix}
100100 & 101100 + 200100 & 201100 + 300100 & 301100 \\
110100 & 110110 & 110120 & 110130 \\
200100 & 200110 + 201100 + 300100 & 201110 + 300110 + 301110 & 301110 \\
300100 & 300110 + 301100 & 300120 + 301110 & 301120
\end{bmatrix} = 0.
\]

For example, the parametrization

\[
\begin{bmatrix}
& u(s, t) \\
& x(s, t) \\
& y(s, t) \\
& z(s, t)
\end{bmatrix} = \sum_{(i,j) \in A} \begin{bmatrix}
1 \\
\times i \\
\times j \\
\times i^2 + j^2
\end{bmatrix} s^i t^j
\]

has the implicit equation

\[
\begin{bmatrix}
X + Y - Z & 5X + 3Y - 3Z & 13X + 3Y - 5Z & 9X - 3Y - 3Z \\
4X + 2Y - 2Z & 16X + 6Y - 6Z - 2 & 20X + 6Y - 6Z - 8 & 8X + 2Y - 2Z - 6 \\
9X + 3Y - 3Z & 17X + 5Y - 5Z - 6 & 13X + 9Y - 3Z - 12 & 5X + 3Y - Z - 6
\end{bmatrix} = 0.
\]

This surface is shown in Figure 4.12.

### 4.8.2 Diamond Cyclides

The monomial, row, and column supports of a "diamond" cyclide are given in (4.16). Its implicit equation is then given by setting $|Dir_A(f, g, h)|$ in (4.17) to zero.

For example, the parametrization

\[
\begin{bmatrix}
& u(s, t) \\
& x(s, t) \\
& y(s, t) \\
& z(s, t)
\end{bmatrix} = \sum_{(i,j) \in A} \begin{bmatrix}
1 \\
\times i \\
\times j \\
\times 0
\end{bmatrix} s^i t^j + \begin{bmatrix}
0 \\
0 \\
0 \\
10
\end{bmatrix} st
\]
has the implicit equation

\[
\begin{vmatrix}
10 - 10X - 10Y + Z & 2Z & 2Z & 0 \\
-2Z & 10 + 10X - 10Y - Z & 0 & -2Z \\
-2Z & 0 & 10 - 10X + 10Y - Z & -2Z \\
0 & -2Z & -2Z & 30 - 10X - 10Y - Z
\end{vmatrix} = 0.
\]

This surface is shown in Figure 4.13.

### 4.8.3 One-Horned Cyclides

The monomial support of a one-horned cyclide is

\[\mathcal{A} = \{ (0, 0), (1, 0), (2, 0), (1, 1), (1, 2) \}.\]

The monomial, row, and column supports are respectively

\[
\begin{array}{ccc}
\times & \times & \\
\times & \bullet & \times \\
\times & \bullet & \times \\
\bullet & \bullet & \bullet \\
\bullet & \bullet & \bullet \\
\end{array}
\quad
\begin{array}{ccc}
\times & \times & \\
\times & \times & \\
\bullet & \bullet & \bullet \\
\bullet & \bullet & \bullet \\
\end{array}
\]

\[\mathcal{A} \quad \mathcal{R}_\mathcal{A} \quad \mathcal{C}_\mathcal{A}\]

The corresponding matrix \[Di_{2,2}(f, g, h)\] with the appropriate coefficients set to zero is

\[
\begin{bmatrix}
0 & 0 & 101100 & 101200 & 201100 & 201200 & 0 & 0 \\
0 & 0 & 101200 & 111200 & 201200 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 201100 & 201200 & 201110 & 201210 & 0 & 0 \\
0 & 0 & 201200 & 0 & 201210 & 201211 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}.
\]
Thus the implicit equation for the one-horned cyclides is

\[
\begin{vmatrix}
101100 & 101200 & 201100 & 201200 \\
101200 & 111200 & 201200 & 0 \\
201100 & 201200 & 201110 & 201210 \\
201200 & 0 & 201210 & 201211 \\
\end{vmatrix}
= 0.
\]

For example, the parametrization

\[
\begin{bmatrix}
w(s,t) \\
x(s,t) \\
y(s,t) \\
z(s,t) \\
\end{bmatrix}
= \sum_{(i,j) \in \mathcal{A}} \begin{bmatrix} 1 \\ i \\ j \\ 0 \end{bmatrix} s^i t^j + \begin{bmatrix} 0 \\ 0 \\ 0 \\ 20 \end{bmatrix}
\]

has the implicit equation

\[
\begin{vmatrix}
20Y - Z & -2Z & 40Y - 2Z & -4Z \\
-2Z & 40X - 20Y - Z & -4Z & 0 \\
40Y - 2Z & -4Z & 20Y - Z & -2Z \\
-4Z & 0 & -2Z & 80 - 40X - 20Y - Z \\
\end{vmatrix}
= 0.
\]

This surface is shown in Figure 4.14.

### 4.8.4 Hexagons

Let \( \mathcal{A} = \{(0,0),(1,0),(0,1),(1,1),(2,1),(1,2),(2,2)\} \). The monomial, row, and columns supports are respectively

\[
\begin{array}{c}
\times \\
\times \\
\times \\
\times \\
\times \\
\times \\
\times \\
\end{array}
\quad
\begin{array}{c}
\bullet \\
\bullet \\
\bullet \\
\bullet \\
\bullet \\
\bullet \\
\bullet \\
\end{array}
\quad
\begin{array}{c}
\times \\
\times \\
\times \\
\times \\
\times \\
\times \\
\times \\
\end{array}
\]

\[
\begin{array}{c}
\mathcal{A} \\
\mathcal{R}_\mathcal{A} \\
\mathcal{C}_\mathcal{A} \\
\end{array}
\]
The corresponding matrix $Dix_{2,2}(f, g, h)$, with appropriate coefficients set to zero, is

\[
\begin{bmatrix}
100100 & 0 & 101100 & 101200 \\
110100 & 0 & 101200 + 110110 + 210100 & 101201 + 111200 \\
120100 & 0 & 120110 + 220100 & 111201 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
210100 & 0 & 210110 + 211100 & 211200 \\
220100 & 0 & 220110 + 221100 + 211200 & 211201 + 221200 \\
0 & 0 & 221200 & 221201 \\
0 & 0 & 102100 & 102200 \\
102200 + 210110 & 102201 + 112200 + 211200 & 0 & 212200 \\
220110 & 112201 + 211201 & 0 & 212201 \\
0 & 0 & 0 & 0 \\
0 & 0 & 211110 & 211210 + 212200 \\
211110 + 211210 & 211211 + 212201 + 221210 & 0 & 212211 \\
221210 & 221211 & 0 & 221221
\end{bmatrix}
\]

Thus the implicit equation is

\[
\begin{bmatrix}
100100 & 101100 & 101200 \\
110100 & 101200 + 110110 + 210100 & 101201 + 111200 \\
120100 & 120110 + 220100 & 111201 \\
210100 & 210110 + 211100 & 211200 \\
220100 & 220110 + 221100 + 211200 & 211201 + 221200 \\
0 & 221200 & 221201
\end{bmatrix}
\]
\[
\begin{pmatrix}
102100 & 102200 & 0 \\
102200 + 210110 & 102201 + 112200 + 211200 & 212200 \\
220110 & 112201 + 211200 & 212201 \\
211110 & 211210 + 212200 & 212210 \\
221110 + 211210 & 211211 + 212201 + 221210 & 212211 \\
221210 & 221211 & 221221
\end{pmatrix} = 0.
\]

For example, the parametrization
\[
\begin{pmatrix}
w(s,t) \\
x(s,t) \\
y(s,t) \\
z(s,t)
\end{pmatrix} = \sum_{(i,j) \in A} \begin{pmatrix} 1 \\ i \\ j \\ 0 \end{pmatrix} s^i t^j + \begin{pmatrix} 0 \\ 0 \\ 0 \\ 30 \end{pmatrix} s t
\]
has the implicit polynomial
\[
\begin{vmatrix}
-Z & 30Y' - Z & -2Z \\
30X - Z & 30X + 30Y' - 5Z - 30 & 60X - 30Y' - 3Z \\
-Z & -4Z & 30X - 30Y' - Z + 30 \\
-2Z & -30X + 60Y' - 3Z & -3Z \\
-2Z & -60X + 60Y' - 6Z & -4Z \\
0 & -2Z & -Z \\
-Z & -2Z & 0 \\
-4Z & 60X - 60Y' - 6Z & -2Z \\
-3Z & 30X - 60Y' - 3Z + 60 & -2Z \\
-30X + 30Y' - Z + 30 & -4Z & -Z \\
-60X + 30Y' - 3Z + 60 & -30X - 30Y' - 5Z + 30 & -30X - Z + 60 \\
-2Z & -30Y' - Z + 60 & -Z
\end{vmatrix}.
\]

This surface is shown in Figure 4.15.
4.8.5 Pentagons

Let \( A = \{(0.0), (1.0), (2.0), (0.1), (1.1), (2.1), (0.2), (1.2)\} \). The monomial, row, and column supports are respectively

\[
\begin{array}{c}
\bullet \times \\
\bullet & \bullet \\
\bullet & \bullet & \bullet \\
\bullet & \bullet & \bullet & \bullet & \times \\
A & \mathcal{R}_A & \mathcal{C}_A
\end{array}
\]

The corresponding matrix \( D_{i x_{2,2}}(f, g, h) \) with appropriate coefficients set to zero is

\[
\begin{bmatrix}
100100 & 100200 & 101100 + 200100 \\
100200 + 110100 & 100201 + 110200 & 101200 + 110110 + 200200 + 210100 \\
120100 + 110200 & 110201 + 120200 & 120110 + 110210 + 210200 \\
120200 & 120201 & 120210 \\
200100 & 200200 & 200110 + 201100 \\
200200 + 210100 & 200201 + 210200 & 200210 + 201200 + 210110 + 211100 \\
210200 & 210201 & 210210 + 211200 \\
0 & 0 & 0
\end{bmatrix}
\]

\[
\begin{bmatrix}
101200 + 200200 & 102100 + 201100 \\
101201 + 111200 + 200201 + 210200 & 110120 + 201200 + 210110 \\
111201 + 120210 + 210201 & 120120 + 110220 + 210210 \\
120211 & 120220 \\
200210 + 201200 & 201110 + 202100 \\
200211 + 201201 + 210210 + 211200 & 201210 + 210120 + 211110 \\
210211 + 211201 & 210220 + 211210 \\
0 & 0
\end{bmatrix}
\]
Thus the implicit equation is the minor with respect to the entry \((8,8)\) set to zero.

For example, the parametrization

\[
\begin{bmatrix}
w(s,t) \\
x(s,t) \\
y(s,t) \\
z(s,t)
\end{bmatrix} = \sum_{(i,j) \in \mathcal{A}} \begin{bmatrix} 1 \\
i \\
j \\
0
\end{bmatrix}s^i t^j + \begin{bmatrix} 0 \\
0 \\
0 \\
10
\end{bmatrix} s t
\]

has the implicit polynomial

\[
\begin{array}{cccc}
-Z & -2Z & 10Y^{} - 3Z & -6Z \\
10X - 3Z & 20X - 3Z & 10X + 10Y^{} - 9Z - 10 & 20X - 10Y^{} - 9Z \\
20X - 3Z & 10X - 3Z & 20X + 10Y^{} - 7Z - 20 & 10X - 10Y^{} - 5Z + 10 \\
-Z & -2Z & -10Y^{} - Z + 20 \\
-2Z & -4Z & 20Y^{} - 3Z & -6Z \\
-6Z & -6Z & -10X + 20Y^{} - 9Z & -20X - 20Y^{} - 9Z + 40 \\
-4Z & -2Z & -6Z & -10X - 20Y^{} - 3Z + 40
\end{array}
\]
\[ \begin{array}{ccc}
20Y - 3Z & -4Z & -2Z \\
10X + 20Y - 7Z - 20 & -6Z & -2Z \\
20X + 20Y - 6Z - 40 & -4Z & -2Z \\
-2Z & -Z & 0 \\
10Y - 3Z & -2Z & -Z \\
-10X + 10Y - 5Z + 10 & -20X - 10Y - 3Z + 40 & -10X - Z + 20 \\
-4Z & -10X - 10Y - Z + 30 & -Z \\
\end{array} \]

This surface is shown in Figure 4.16.
Figure 4.12: A Hirzebruch surface.

Figure 4.13: A diamond cyclide.

Figure 4.14: A one-horn cyclide.

Figure 4.15: A hexagon surface.
Figure 4.16: A pentagon surface.
Part II

Implicitization
Chapter 5

Implicitization Using Moving Curves

In Computer Aided Geometric Design (CAGD), curves and surfaces have two standard representations: parametric and implicit. The parametric representation is convenient for rendering curves and surfaces, whereas the implicit representation is useful for checking whether or not a point lies on a curve or surface. Both representations are important. But many curve and surface design systems start from parametric representations. Implicitization is the process of converting curves and surfaces from parametric form into implicit form.

Resultants are an effective tool for solving this problem for rational curves and surfaces [32] [43] [46] -- see too Chapter 4. To recall why resultants arise naturally in implicitization, consider a rational curve

\[ X = \frac{x(t)}{w(t)}, \quad Y = \frac{y(t)}{w(t)}, \]  

where \( x(t), y(t), w(t) \) are polynomials. To obtain the implicit representation

\[ F(X, Y) = 0. \]

for curve (5.1), introduce two auxiliary polynomials (in \( t \))

\[ X \cdot w(t) - x(t), \quad Y \cdot w(t) - y(t). \]  

By definition, the resultant of these two polynomials vanishes if and only if they have a common root, i.e. if and only if the point \((X, Y)\) satisfies the two equations

\[ X \cdot w(t) - x(t) = 0, \quad Y \cdot w(t) - y(t) = 0. \]
for some value of $t$. Thus, $(X, Y)$ makes the resultant of $X \cdot w(t) - x(t), Y \cdot w(t) - y(t)$ vanish if and only if $(X, Y)$ is on curve (5.1). So setting the resultant to zero yields the implicit equation of the parametric curve.

The major drawback of the resultant method for implicitization is that resultants vanish identically in the presence of base points [6] [42]. A rational curve or surface is said to have a base point if its numerators and denominators simultaneously vanish at some parameter value. Professor Tom Sederberg first introduced the method of moving algebraic curves and surfaces to solve the implicitization problem for rational curves and surfaces with base points [53] [54].

In this chapter, we will focus on the implicitization of rational curves. Implicitizing rational surfaces is the subject of the next chapter. We begin by reviewing Sederberg's method of moving curves. Then we discuss implicitizing rational curves using moving lines. We show that the implicit equation can be obtained by taking the Sylvester resultant (see too Chapter 2) of two particular moving lines. This construction generates a novel implicitization matrix in the style of the Sylvester resultant — striped and shifted columns and lots of zeros — and the size of the Bézout resultant. In Section 5.3 we present a novel perspective on the method of moving conics for rational curve implicitization. Based on the observation that the determinant of the moving line coefficient matrix factors a sub-determinant of the moving conic coefficient matrix, we present a new proof that when there are no low degree moving lines following the rational curve, then the method of moving conics successfully implicitizes the rational curve. This new proof is important because, as we shall see in Chapter 6, it generalizes to rational surfaces where previously there was no known proof for the validity of this method.
5.1 The Method of Moving Curves

Among moving algebraic curves, moving lines and moving conics are the most important. A moving line of degree \( d \)

\[
\sum_{i=0}^{d} (A_i x + B_i y + C_i w) t^i = 0 \tag{5.3}
\]

is a one parameter family of implicitly defined lines, with one line corresponding to each parameter \( t \). Similarly, a moving conic of degree \( d \)

\[
\sum_{i=0}^{d} (A_i x^2 + B_i y^2 + C_i x y + D_i x w + E_i y w + F_i w^2) t^i = 0 \tag{5.4}
\]

is a one parameter family of implicitly defined conics. For each fixed \( t \), Equation (5.4) is the implicit equation of a different conic.

A moving line (5.3) or a moving conic (5.4) is said to follow a rational curve (5.1) if

\[
\sum_{i=0}^{d} (A_i x(t) + B_i y(t) + C_i w(t)) t^i \equiv 0; \tag{5.5}
\]

or

\[
\sum_{i=0}^{d} (A_i x^2(t) + B_i y^2(t) + C_i x(t) y(t)) + D_i x(t) w(t) + E_i y(t) w(t) + F_i w^2(t) t^i \equiv 0. \tag{5.6}
\]

For example, if the rational curve (5.1) is of degree \( n \), then the equations

\[
x \cdot w(t) - w \cdot x(t) = 0, \quad y \cdot w(t) - w \cdot y(t) = 0.
\]

or equivalently,

\[
X \cdot w(t) - x(t) = 0, \quad Y \cdot w(t) - y(t) = 0.
\]

are two moving lines of degree \( n \) that follow the rational curve (5.1). Thus the standard way to find the implicit equation of the rational curve (5.1) is to compute the resultant of these two moving lines of degree \( n \) that follow the curve. So in
this sense, we shall see that the method of moving curves is a generalization of the standard resultant method for implicitizing rational curves.

By setting the coefficients of all monomials \( t^k \) in Equation (5.5) (or Equation (5.6)) to zero, we generate a linear system with unknowns \( \{A_i, B_i, C_i, D_i\} \) (or \( \{A_i, \cdots, F_i\} \)). Any solution of this linear system is a moving line (or conic) that follows the rational curve (5.1). The method of moving lines (moving conics) constructs the implicit equation of a rational curve by taking the determinant of the coefficient matrix of a set of independent moving lines (moving conics) that follow the curve. Here independence means not just the linear independence of the solutions of the linear system generated from Equation (5.5) (or Equation (5.6)), but rather independence of the moving lines (or conics). Thus, for example, \( k \) moving lines

\[
l_1 \equiv (A_{1,0}x + B_{1,0}y + C_{1,0}w) + \cdots + (A_{1,d}x + B_{1,d}y + C_{1,d}w)t^d = 0.
\]

\[
\vdots
\]

\[
l_k \equiv (A_{k,0}x + B_{k,0}y + C_{k,0}w) + \cdots + (A_{k,d}x + B_{k,d}y + C_{k,d}w)t^d = 0
\]

are said to be independent if the matrix

\[
\begin{bmatrix}
A_{1,0}x + B_{1,0}y + C_{1,0}w & \cdots & A_{k,0}x + B_{k,0}y + C_{k,0}w \\
\vdots & \vdots & \vdots \\
A_{1,d}x + B_{1,d}y + C_{1,d}w & \cdots & A_{k,d}x + B_{k,d}y + C_{k,d}w
\end{bmatrix}
\]

is of rank \( k \).

**Example 5.1 (Moving Lines)**

Find the implicit equation of the rational curve

\[
x(t) = 1 - t^2.
\]

\[
y(t) = 2t.
\]

\[
w(t) = 1 + t^2.
\]
A degree 1 moving line has the form
\[(A_0 x + B_0 y + C_0 w) + (A_1 x + B_1 y + C_1 w)t = 0.\]

The linear system generated by the method is
\[
\begin{align*}
A_0 + C_0 &= 0, \\
A_1 + C_1 + 2B_0 &= 0, \\
-A_0 + 2B_1 + C_0 &= 0, \\
-A_1 + C_1 &= 0.
\end{align*}
\]

Solving this system, we obtain two independent moving lines:
\[
\begin{align*}
(-y) + (x + w)t &= 0, \\
(-x + w) + (-y)t &= 0.
\end{align*}
\]

And, as expected, the implicit equation is
\[
\begin{vmatrix}
-x + w \\
x + w \\
-y
\end{vmatrix} \equiv x^2 + y^2 - w^2 = 0.
\]

Example 5.2 (Moving Conics)

Find the implicit equation of the rational curve
\[
\begin{align*}
x(t) &= t^4 - 2, \\
y(t) &= t^4 + t^2 + 1, \\
w(t) &= t^3 + 1.
\end{align*}
\]

A degree 1 moving conic has the form
\[
\sum_{i=0}^{t} (A_i x^2 + B_i y^2 + \cdots + F_i w^2) t^i = 0.
\]
The linear system generated from this moving conic is

\[
4A_0 + B_0 - 2C_0 + E_0 + F_0 - 2D_0 = 0.
\]
\[
4A_1 + B_1 - 2C_1 + E_1 + F_1 - 2D_1 = 0.
\]
\[
2B_0 - 2C_0 + E_0 = 0.
\]
\[
2B_1 - 2C_1 + E_0 + E_1 + 2F_0 - 2D_0 = 0.
\]
\[-4A_0 + 3B_0 - C_0 + E_0 + E_1 + 2F_1 + D_0 - 2D_1 = 0.
\]
\[-4A_1 + 3B_1 - C_1 + E_0 + E_1 + D_1 = 0.
\]
\[2B_0 + C_0 + E_1 + F_0 = 0.
\]
\[2B_1 + C_1 + E_0 + F_1 + D_0 = 0.
\]
\[A_0 + B_0 + C_0 + E_1 + D_1 = 0.
\]
\[A_1 + B_1 + C_1 = 0.
\]

Two independent moving conics are obtained by solving this system:

\[
\frac{-2wx}{t} - \frac{8wy}{t} + \frac{4y^2}{t} + t \left( w^2 + \frac{4wx}{t} - \frac{8wy}{t} - \frac{3xy}{t} + \frac{3y^2}{t} \right) = 0.
\]

\[
u^2 + \frac{6wx}{t} - \frac{4wy}{t} - xy - \frac{5y^2}{t} + t \left( \frac{2wx}{t} + \frac{10wy}{t} + \frac{2xy}{t} - \frac{2y^2}{t} \right) = 0.
\]

The implicit equation is

\[
\begin{vmatrix}
-\frac{2wx}{t} - \frac{8wy}{t} + \frac{4y^2}{t} & w^2 + \frac{6wx}{t} - \frac{4wy}{t} - xy - \frac{3y^2}{t} \\
\frac{4wx}{t} - \frac{8wy}{t} - \frac{3xy}{t} + \frac{3y^2}{t} & \frac{2wx}{t} + \frac{10wy}{t} + \frac{2xy}{t} - \frac{2y^2}{t}
\end{vmatrix}
\equiv \frac{1}{t} \left( -7w^4 - 10w^3x - 4w^2x^2 + 12w^3y + 14w^2xy + 6wx^2y - 14w^2y^2 \\
- 10wxy^2 - 3x^2y^2 + 4wy^3 + 2xy^3 + y^4 \right)
\]
\equiv 0.
5.2 Implicitizing Curves Using Moving Lines

In this section, we consider first rational curves of even degree. We show that there are always at least two moving lines of degree $m$ that follow a rational curve of degree $2m$. The $m \times m$ Bézout determinant of these two moving lines has been used by previous authors to establish the efficacy of implicitization by the method of moving conics [23] [54]. Here we prove that the $2m \times 2m$ Sylvester determinant of these two moving lines is an implicit expression for the rational curve if and only if there are no moving lines of degree $< m$ that follow the curve. This construction generates an implicitization matrix in the style of Sylvester with the order of Bézout. At the end of this section, we develop similar results for rational curves of odd degree.

5.2.1 Even Degree Rational Curves

A rational curve of degree $2m$ can be written as $(x(t) : y(t) : w(t))$, where

\[ x(t) = \sum_{i=0}^{2m} a_i t^i, \quad y(t) = \sum_{i=0}^{2m} b_i t^i, \quad w(t) = \sum_{i=0}^{2m} c_i t^i \]  \hspace{1cm} (5.7)

and $\gcd(x(t), y(t), w(t)) = 1$. We can think of a rational curve as the track of a moving point.

Let

\[ (A_0 x + B_0 y + C_0 w) + \cdots + (A_d x + B_d y + C_d w) t^d = 0. \]  \hspace{1cm} (5.8)

be a moving line. The moving line (5.8) follows the rational curve (5.7) if and only if

\[ (A_0 x(t) + B_0 y(t) + C_0 w(t)) + \cdots + (A_d x(t) + B_d y(t) + C_d w(t)) t^d \equiv 0. \]  \hspace{1cm} (5.9)

By equating the coefficients of the powers of $t$ in Equation (5.9) to zero, we obtain $2m + d + 1$ equations in $3d + 3$ unknowns. The $3d + 3$ unknowns $A_0, B_0, C_0, \cdots, A_d, B_d, C_d$ of the moving line (5.8) can be found by solving the $(2m + d + 1) \times (3d + 3)$ linear system

\[
\text{Coeff}(x(t), y(t), w(t), \cdots, x(t)t^d, y(t)t^d, w(t)t^d)
\]
where "Coeff" stands for the matrix whose columns are the coefficients of the given polynomials. When \( d = m \), the dimension of the linear system is \((3m + 1) \times (3m + 3)\). Consequently, there are at least two linearly independent solutions \( p(x, y, w: t) \) and \( q(x, y, w: t) \).

The \( 2m \times 2m \) Sylvester matrix \( Syl(p, q) \) obtained by eliminating \( t \) from \( p \) and \( q \) can be written as

\[
Syl(p, q) = \text{Coeff} (p, q, pt, qt, \ldots, p t^{m-1}, q t^{m-1})
\]

(see too Chapter 2).

**Theorem 5.1** |\( Syl(p, q) \)| = 0 is the implicit equation of the rational curve (5.1) when there are no moving lines of degree < \( m \) that follow curve (5.1).

**Proof:** Since the implicit equation of a rational curve of degree \( 2m \) is represented by an irreducible polynomial of degree \( 2m \) [54], we need only establish three facts:

1) |\( Syl(p, q) \)| \( \neq 0 \):

2) |\( Syl(p, q) \)| is of degree at most \( 2m \); and

3) |\( Syl(p, q) \)| vanishes on \((x(t) : y(t) : w(t))\).

From the properties of resultants (Chapter 2), we know that |\( Syl(p, q) \)| \( \equiv 0 \) if and only if \( p \) and \( q \) have a common factor \( g(t) \) of degree \( \geq 1 \). Since \( p \) and \( q \) are of degree 1 in \( x, y, w \), one of \( g \) and \( p/g \) would be of degree 1 in \( x, y, w \), i.e. a moving line of degree < \( m \) that follows the curve. But by assumption there are no such moving lines, so |\( Syl(p, q) \)| cannot vanish identically.

Since |\( Syl(p, q) \)| is the determinant of a \( 2m \times 2m \) matrix with linear entries in \( x, y, w \), obviously the degree of |\( Syl(p, q) \)| is at most \( 2m \) in \( x, y, w \).
Finally, \( p(x, y, w; t) \) and \( q(x, y, w; t) \) follow the rational curve, so

\[
p(x(t_0), y(t_0), w(t_0); t_0) \equiv 0, \quad q(x(t_0), y(t_0), w(t_0); t_0) \equiv 0.
\]

for any parameter \( t_0 \). That is, the two polynomials

\[
p(x(t_0), y(t_0), w(t_0); t), \quad q(x(t_0), y(t_0), w(t_0); t)
\]

have a common root \( t_0 \). Hence, the resultant

\[
|\text{Syl}(p(x(t_0), y(t_0), w(t_0); t), q(x(t_0), y(t_0), w(t_0); t))| = 0.
\]

Therefore, \( |\text{Syl}(p, q)| \) vanishes on \( (x(t) : y(t) : w(t)) \). \( \square \)

In summary, we have shown that for a degree \( 2m \) rational curve, the \( 2m \times 2m \) Sylvester determinant of two degree \( m \) moving lines is the implicit equation of the curve if there are no moving lines of degree \( < m \) that follow the curve. The existence of a moving line of degree \( m - 1 \) that follows the curve is equivalent to the vanishing of the \( 3m \times 3m \) determinant

\[
|\text{Coeff}(x(t), y(t), w(t), \ldots, t^{m-1}x(t), t^{m-1}y(t), t^{m-1}w(t))|.
\]

This determinant is a polynomial in the coefficients of \( x(t), y(t), w(t) \) and therefore almost never vanishes. However, in case such lower degree moving lines do exist, the desired Sylvester determinant can be salvaged by finding a \( \mu \)-basis for the curve (see Section 5.2.3).

**Example 5.3**

Consider the rational sextic curve

\[
x(t) = 1 + 2t^2 + 2t^5, \quad y(t) = 2 + t^6, \quad w(t) = 1 + t + 2t^2 + 2t^3 + t^4 + t^6.
\]

To use the standard method to implicitize this curve, we introduce two auxiliary polynomials

\[
X \cdot w(t) - x(t), \quad Y \cdot w(t) - y(t).
\]
Their Sylvester resultant is the $12 \times 12$ determinant

$$
\begin{bmatrix}
-1 + X & -2 + Y & 0 & 0 & \cdots & 0 & 0 \\
X & Y & -1 + X & -2 + Y & \cdots & 0 & 0 \\
-2 - 2X & 2Y & X & Y & \cdots & 0 & 0 \\
2X & 2Y & -2 + 2X & 2Y & \cdots & 0 & 0 \\
X & Y & 2X & 2Y & \cdots & 0 & 0 \\
-2 & 0 & X & Y & \cdots & -1 + X & -2 + Y \\
X & -1 + Y & -2 & 0 & \cdots & X & Y \\
0 & 0 & X & -1 + Y & \cdots & -2 + 2X & 2Y \\
0 & 0 & 0 & 0 & \cdots & 2X & 2Y \\
0 & 0 & 0 & 0 & \cdots & X & Y \\
0 & 0 & 0 & 0 & \cdots & -2 & 0 \\
0 & 0 & 0 & 0 & \cdots & X & -1 + Y
\end{bmatrix}
$$

where the six columns in the middle have been omitted. The Bézout resultant is the $6 \times 6$ determinant \[17\] \[32\]

$$
\begin{bmatrix}
2X - Y & -1 + 4X & 4X - 2Y & 2X - Y & -4 + 2Y & 1 + X - Y \\
-4 + 4X & 4X & 2X - Y & -4 + 2Y & 1 + X + Y & -X \\
4X - 2Y & 2X - Y & -4 - 2Y & 1 + X - Y & -X + 4Y & 2 - 2X - 2Y \\
2X - Y & -4 + 2Y & 1 + X - Y & -X + 4Y & 2 - 2X + 2Y & -2X \\
-4 + 2Y & 1 + X + Y & -X + 4Y & 2 - 2X + 2Y & -2X + 2Y & -X \\
1 + X - Y & -X & 2 - 2X - 2Y & -2X & -X & 2 - 2Y
\end{bmatrix}
$$

On the other hand, using linear algebra, it is easy to calculate two degree 3 moving lines following this curve:

$$(855w + 31x - 443y) + t(77y - 778w - 231x)$$

$$+ t^2(338x - 666y) + t^3(25w + 333x - 25y) = 0.$$ 

$$(780w - 413y + 46x) + t(-748w + 82y - 196x)$$

$$+ t^2(25w - 631y + 333x) + t^3(303x) = 0.$$
The new method computes the implicit equation for this curve by taking the $6 \times 6$ Sylvester determinant of these two moving lines:

\[
\begin{vmatrix}
-443y + 31x + 855w & 46x - 413y + 780w & 0 \\
-778w - 231x + 77y & -748w + 82y - 196x & -443y + 31x + 855w \\
338x - 666y & 25w - 631y + 333x & -778w - 231x + 77y \\
25w + 333x - 25y & 303x & 338x - 666y \\
0 & 0 & 25w + 333x - 25y \\
0 & 0 & 0
\end{vmatrix}
\]

Using Mathematica, we verified that all three methods produce the correct implicit equation for the given rational curve. Notice that the determinant generated by the new method has the structure of the Sylvester resultant but the order of the Bézout resultant.

5.2.2 Odd Degree Rational Curves

For a rational curve of degree $2m + 1$, there is always at least one non-zero moving line of degree $m$ and at least 3 linearly independent moving lines of degree $m + 1$ that follow the curve. Therefore, there always exists a moving line $p$ of degree $m$ and a moving line $q$ of degree $m + 1$, where $q$ is not a multiple of $p$, that follow the rational curve. Suppose there is no moving line of degree $< m$ that follows the curve. Then by an argument similar to the case of even degrees, the Sylvester resultant of $p$ and $q$ is the determinant of a $(2m + 1) \times (2m + 1)$ matrix that represents the implicit equation of the rational curve.
5.2.3 Anti-Annihilation by $\mu$-Basis

The implicitization method in Section 5.2.1 works when there are no low degree moving lines that follow the curve. In the rare cases when there do exist low degree moving lines following the curve, the Sylvester resultant used in Section 5.2.1 generally vanishes identically [52]. In order to circumvent this difficulty—that is, to counter the annihilation effect of low degree moving lines—and show how the desired Sylvester-style/Bézout-size determinant can still be constructed, we need the notion of a $\mu$-basis [23].

Consider a degree $n$ rational curve $(x(t) : y(t) : w(t))$. By solving an $(n + d + 1) \times (3d + 3)$ linear system, we find that the number of linearly independent degree $d$ moving lines that follow this curve is at least $(3d + 3) - (n + d + 1) = 2d + 2 - n$. Thus the system always has solutions when $3d + 3 > n + d + 1$ or $d > n/2 - 1$. Hence if $\mu$ is the lowest degree in $t$ of all the moving lines that follow the curve, then $\mu \leq \lfloor n/2 \rfloor$. Let $p$ be a moving line with the lowest degree $\mu$ that follows the curve.

By our previous analysis, there are at least $2(n - \mu) + 2 - n = n + 2 - 2\mu$ linearly independent moving lines of degree $n - \mu$ that follow the curve. Not all of them can be multiples of $p$ because $p$ can only generate $n + 1 - 2\mu$ independent moving lines of degree at most $n - \mu$: $p, \cdots, pt^{n-2\mu}$. Hence there is a degree $n - \mu$ moving line $q$ that is not a multiple of $p$.

The two moving lines $p$ and $q$ that we just constructed have the following nice property:

**Theorem 5.2** Any degree $d$ moving line $l$ that follows the curve $(x(t) : y(t) : w(t))$ can be written uniquely as $Ap + Bq$, where $A$ is a polynomial in $t$ of degree at most $d - \mu$, and $B$ is a polynomial in $t$ of degree at most $d + \mu - n$ [23].

**Proof:** [23] presents a proof of this result based on ideal theory. An alternative simpler proof using only linear algebra is provided in [16]. \qed
The two moving lines $p$ and $q$ in Theorem 5.1 are called a $\mu$–basis of the curve $(x(t) : y(t) : w(t))$.

We have shown in Section 5.2.1 that when there are no moving lines of degree $< m$ following a rational curve of degree $n = 2m$, there will be two moving lines of degree $m$ following the curve and their Sylvester determinant gives the implicit equation. Clearly these two moving lines are simply $p$ and $q$ in Theorem 1 with $\mu = m = n - \mu$. Theorem 1 also tells us that for $\mu \leq d < n - \mu$, a degree $d$ moving line $l$ has the form

$$l = c_0p + c_1pt + \cdots + c_{d-\mu}pt^{d-\mu},$$

where $c_i$ are constants because $l$ is of the form $Ap + Bq$ with $B = 0$ due to the degree constraints on $A$ and $B$. Consequently, the Sylvester determinant of any two of these degree $d$ moving lines vanishes as both are multiples of $p$: furthermore, the number of such linearly independent degree $d$ moving lines is $N_d = d - \mu + 1$. In particular, when there are moving lines of degree $< m$ that follow the curve, we have $\mu < m < n - \mu$, so the Sylvester resultant of any two degree $m$ moving lines vanishes and there are $N_m = m - \mu + 1 \geq 2$ degree $m$ moving lines following the curve. Note that we can find $\mu$ in terms of $N_m$:

$$\mu = m - N_m + 1,$$

$$N_m = 3m + 3 - \text{Rank of } \text{Coeff}(x(t), y(t), w(t), \ldots, t^mx(t), t^my(t), t^mw(t)).$$

In general then, for a degree $n$ rational curve $(x(t) : y(t) : w(t))$, we can obtain the $\mu$–basis functions $p$ and $q$ by straightforward linear algebra. Since $p$ is irreducible (by degree minimality) and $q$ is not a multiple of $p$, they have no common factors. Hence their Sylvester resultant

$$\text{Syl}(p, q) = \text{Coeff}(p, pt, \ldots, pt^{n-\mu-1}, q, qt, \ldots, qt^{n-\mu-1})$$

is a matrix of size $n \times n$ whose determinant does not vanish identically. By the divisibility and degree argument of Section 5.2.1, we see that this Sylvester determinant
gives an implicit expression for the rational curve \((x(t) : y(t) : w(t))\) in the style of Sylvester with the order Bézout.

As an example, consider the degree \(n\) rational curve

\[
(x(t) : y(t) : w(t)) = (1 : t^{n-1} : 1 + t^n).
\]  \hspace{1cm} (5.10)

Simple calculations reveal that \(p = x + ty - w\) and \(q = t^{n-1}x - y\). The Sylvester determinant

\[
\begin{vmatrix}
  p & \cdots & p^{n-2} & q \\
  x - w & \cdots & -y \\
  y & \cdots & \vdots \\
  \vdots & \cdots & x - w \\
  y & x 
\end{vmatrix} = (-1)^{n-1}y^n + x(x - w)^{n-1}
\]

is easily seen to represent the implicit equation of this rational curve.

5.3 Implicitizing Curves Using Moving Conics

It is known that the method of moving lines always successfully produces the implicit equation of a rational curve [54]. Moving lines have been thoroughly studied [54], and clever applications of moving lines are presented in [16] [23]. However, the moving line method requires computing a large determinant to generate the implicit representation. The method of moving conics also always successfully implicizes rational curves. Moreover, the moving conic method computes the implicit equation of a rational curve by taking a determinant of much smaller size than the determinant generated by the method of moving lines.

This section presents a new perspective on the methods of moving lines and moving conics. For a rational curve \(X = \frac{x(t)}{w(t)}, Y = \frac{y(t)}{w(t)}\) of degree \(2m\), consider the linear system generated by substituting \(x(t), y(t), w(t)\) into the moving line (5.5) (moving conic (5.6)) and equating the coefficients of all monomials (in \(t\)) to zero. The solutions of this linear system are moving lines (moving conics) that follow the rational curve. We call the coefficient matrix of this linear system the moving line coefficient matrix.
(moving conic coefficient matrix). The goal of this section is to derive a relationship between the moving line coefficient matrix and the moving conic coefficient matrix. In particular we shall show that for moving lines and moving conics of degree \( m - 1 \), the determinant of the moving line coefficient matrix actually factors a sub-determinant of the moving conic coefficient matrix. Based on this observation, we present a new proof that when there are no low degree moving lines following a rational curve, the method of moving conics successfully produces the implicit equation of the rational curve. Unlike the work in [54], this new proof will generalize naturally to the method of moving quadrics for rational surfaces [Ch 6] [22].

**Notation**

For convenience, as we did in Section 3.6, we adopt the following notation: Let \( p_i(t) \), \( 1 \leq i \leq k \), be polynomials of degree \( d_i \), and let \( d = \max(d_1, d_2, \ldots, d_k) \). We shall write \( \begin{bmatrix} p_1(t) & p_2(t) & \cdots & p_k(t) \end{bmatrix}^C \) to denote the \((d+1) \times k\) coefficient matrix of the polynomials \( p_i(t) \) whose rows are indexed by \( 1, t, \ldots, t^d \). That is,

\[
\begin{bmatrix} p_1(t) & \cdots & p_k(t) \end{bmatrix} = \begin{bmatrix} 1 & \cdots & t^d \end{bmatrix} \cdot \begin{bmatrix} p_1(t) & p_2(t) & \cdots & p_k(t) \end{bmatrix}^C.
\]

If \( \begin{bmatrix} p_1(t) & p_2(t) & \cdots & p_k(t) \end{bmatrix}^C \) is a square matrix, we denote its determinant by

\[
\begin{vmatrix} p_1(t) & \cdots & p_k(t) \end{vmatrix}.
\]

**5.3.1 Moving Line and Moving Conic Coefficient Matrices**

In Sections 5.3.1 and 5.3.2, we will study only rational curves with even degrees and no base points. Consider then a degree \( 2m \) rational curve in homogeneous form

\[
\begin{bmatrix} x(t) : y(t) : w(t) \end{bmatrix},
\]

where

\[
x(t) = \sum_{i=0}^{2m} a_i t^i, \quad y(t) = \sum_{i=0}^{2m} b_i t^i, \quad w(t) = \sum_{i=0}^{2m} c_i t^i.
\]

and \( \text{GCD}(x(t), y(t), w(t)) = 1 \).
To find a degree $d$ moving line that follows curve (5.11), we consider the linear system generated by equating the coefficients of all monomials (in $t$) in Equation (5.5) to zero. We can write this system as

$$
\begin{bmatrix}
A_0 \\
B_0 \\
C_0 \\
\vdots \\
A_d \\
B_d \\
C_d
\end{bmatrix}
\begin{bmatrix}
x, 
y, 
w, \cdots, 
t^d x, 
t^d y, 
t^d w \end{bmatrix}^T \overset{\mathcal{C}}{\rightarrow} (5.12)
$$

The coefficient matrix $[x, y, w, \cdots, t^d x, t^d y, t^d w]^T$ is of size $(2m + d + 1) \times (3d + 3)$. Similarly, to find degree $d$ moving conics that follow curve (5.11), we consider the linear system generated by equating the coefficients of all monomials (in $t$) in Equation (5.6) to zero. We can write this system as

$$
\begin{bmatrix}
A_0 \\
B_0 \\
C_0 \\
D_d \\
E_d \\
F_d
\end{bmatrix}
\begin{bmatrix}
x^2, 
y^2, 
xy, 
xw, 
yw, 
w^2, \cdots, 
t^d x w, 
t^d y w, 
t^d w^2 \end{bmatrix}^T \overset{\mathcal{C}}{\rightarrow} (5.13)
$$

The coefficient matrix $[x^2, y^2, xy, xw, yw, w^2, \cdots, t^d x w, t^d y w, t^d w^2]^T$ is of size $(4m + d + 1) \times (6d + 6)$.

It is well known that the implicit form of curve (5.11) is a polynomial of degree $2m$ in $X, Y$. The method of moving conics finds a set of $m$ independent moving conics of degree $m - 1$ that follow the curve and then generates the implicit representation from the determinant of the coefficient matrix of this set of moving conics. Consider
then \( m \) independent moving conics of degree \( m - 1 \) that follow the curve:

\[
p_0(x, y, w; t) = c_{0,0}(x, y, w) + c_{0,1}(x, y, w)t + \cdots + c_{0,m-1}(x, y, w)t^{m-1},
\]

\[
\vdots
\]

\[
p_{m-1}(x, y, w; t) = c_{m-1,0}(x, y, w) + c_{m-1,1}(x, y, w)t + \cdots + c_{m-1,m-1}(x, y, w)t^{m-1},
\]

where \( c_{i,j}(x, y, w), 0 \leq i, j \leq m - 1 \), are quadratic in \( x, y, w \). It is known that such independent conics exist and that

\[
\begin{bmatrix}
c_{0,0}(x, y, w) & \cdots & c_{m-1,0}(x, y, w) \\
\vdots & \ddots & \vdots \\
c_{0,m-1}(x, y, w) & \cdots & c_{m-1,m-1}(x, y, w)
\end{bmatrix} = 0
\]

is the implicit equation of curve (5.11), when there are no moving lines of degree \( m - 1 \) that follow the curve [54]. We will therefore consider the case where \( d = m - 1 \) in Equations (5.12) and (5.13). Then \( [x, y, w, \cdots, t^d x, t^d y, t^d w]^C \) is a square matrix of order \( 3m \) — denote this \( 3m \times 3m \) matrix by \( ML \). The matrix \( [x^2, y^2, xy, xw, yw, w^2, \cdots, t^d x^2, t^d y^2, t^d w^2]^C \) is of size \( 5m \times 6m \) — denote this \( 5m \times 6m \) matrix by \( MC \). Our goal is to find a relationship between \( ML \) and \( MC \), and then to apply this relationship to show that the method of moving conics works whenever there are no low degree moving lines that follow the curve.

To obtain a square submatrix from \( MC \), we delete the columns that represent the polynomials \( t^k w^2, k = 0, \cdots, m - 1 \). This procedure generates a square submatrix of order \( 5m \). We denote this \( 5m \times 5m \) submatrix by \( MC_w \). To summarize:

\[
ML = \begin{bmatrix}
x, & y, & w, & \cdots & t^{m-1}x, & t^{m-1}y, & t^{m-1}w
\end{bmatrix}^C_{3m \times 3m},
\]

\[
MC_w = \begin{bmatrix}
x^2, & y^2, & xy, & xw, & yw, & \cdots, & t^{m-1}(x^2, & y^2, & xy, & xw, & yw)
\end{bmatrix}^C_{5m \times 5m}.
\]

### 5.3.2 \( |ML| \) Factors \( |MC_w| \)

We are going to show that \( |MC_w| = \text{constant} \cdot \text{Res}(x, y) \cdot |ML|^2 \), where \( \text{Res}(x, y) \) denotes the resultant of the polynomials \( x, y \). We begin with a few preliminary lemmas.
Lemma 5.1 $|ML|$ is irreducible in the coefficients of $x(t), y(t), w(t)$ \[54\].

Lemma 5.2 $\text{Res}(x(t), y(t))$ is irreducible in the coefficients of $x(t), y(t)$ \[39\].

Lemma 5.3 If $|ML| = 0$, then $|MC_w| = 0$.

Proof: If $|ML| = 0$, then the columns in $ML$ are linearly dependent. That is, there exist constants $\lambda_i, i = 1, \ldots, 3m$, such that

$$\lambda_1 x(t) + \lambda_2 y(t) + \lambda_3 w(t) + \cdots + \lambda_{3m-2} t^{m-1} x(t) + \lambda_{3m-1} t^{m-1} y(t) + \lambda_{3m} t^{m-1} w(t) \equiv 0.$$  \hspace{1cm} (5.14)

Multiplying both sides of Equation (5.14) by $x(t)$ or $y(t)$, we get a linear relationship between the polynomials $t^k x^2, t^k y^2, t^k x y, t^k x w, t^k y w, k = 0, \ldots, m - 1$. Thus the columns of $MC_w$ are linearly dependent. Therefore, $|MC_w| = 0$. \hfill \Box

Lemma 5.4 If the resultant $\text{Res}(x(t), y(t)) = 0$, then $|MC_w| = 0$.

Proof: When the resultant $\text{Res}(x(t), y(t)) = 0$, the two polynomials $x(t), y(t)$ have a common root $t_0$. Therefore,

$$\begin{bmatrix} 1 & t_0 & \cdots & t_0^{5m-1} \end{bmatrix} \cdot MC_w$$

$$= \begin{bmatrix} x^2(t_0) & \cdots & y(t_0) w(t_0) & \cdots & t_0^{m-1} x^2(t_0) & \cdots & t_0^{m-1} y(t_0) w(t_0) \end{bmatrix}_{1 \times 5m}$$

$$= \begin{bmatrix} 0 & \cdots & 0 \end{bmatrix}_{1 \times 5m}.$$ Thus the rows of $MC_w$ are linearly dependent, so $|MC_w| = 0$. \hfill \Box

Lemma 5.5 If $|MC_w| = 0$, then either $\text{Res}(x(t), y(t)) = 0$ or $|ML| = 0$.

Proof: If $|MC_w| = 0$, then there exist constants $\lambda_i, i = 1, \ldots, 5m$ such that

$$\lambda_1 x^2(t) + \lambda_2 y^2(t) + \lambda_3 x(t) y(t) + \cdots + \lambda_{5m-1} t^{m-1} x(t) w(t) + \lambda_{5m} y(t) w(t) \equiv 0.$$  \hspace{1cm} (5.15)

Collect the coefficients of $x^2, y^2, xy, xw, yw$ and rewrite Equation (5.15) as

$$p_1(t) x^2 + p_2(t) y^2 + p_3(t) xy + p_4(t) xw + p_5(t) yw \equiv 0.$$  \hspace{1cm} (5.16)
where $p_i(t), i = 1, \ldots, 5$ are polynomials in $t$ of degree $m - 1$. Equation (5.16) can also be written as

$$
\left( p_1(t)x + p_3(t)y + p_4(t)w \right) \cdot x \equiv \left( -p_2(t)y - p_5(t)w \right) \cdot y. \tag{5.17}
$$

Now we shall prove that if $R(\mathbb{C}(x(t), y(t))) \neq 0$, then $|\mathcal{M}| = 0$. If $R(\mathbb{C}(x(t), y(t))) \neq 0$, then $x(t)$ and $y(t)$ do not have a common root. Therefore, from Equation (5.17), $y$ must divide $p_1(t)x + p_3(t)y + p_4(t)w$, and $x$ must divide $-p_2(t)y - p_5(t)w$. Without lose of generality, we can assume $\deg(x) = 2m$. Therefore, there exists a polynomial $q(t)$ of degree less than or equal to $m - 1$ such that $q(t)x = -p_2(t)y - p_5(t)w$. That is,

$$
q(t)x + p_2(t)y + p_5(t)w \equiv 0. \tag{5.18}
$$

When $p_2(t)$ or $p_5(t)$ are not identically zero, relationship (5.18) asserts that the columns of $\mathcal{M}$ are linearly dependent, since the degrees of $q(t), p_2(t), p_5(t)$ are all at most $m - 1$. If $p_2(t) \equiv 0$ and $p_5(t) \equiv 0$, then by Equation (5.17) there exist non-zero polynomials $p_1(t), p_3(t), p_4(t)$, such that

$$
p_1(t)x + p_3(t)y + p_4(t)w \equiv 0 \tag{5.19}
$$

with the degrees of $p_1(t), p_3(t), p_4(t)$ all less than or equal to $m - 1$. This relationship again asserts that the columns of $\mathcal{M}$ are linearly dependent. Therefore, in either case, $|\mathcal{M}| = 0$. \hfill \square

With all this preparation, we can now finally prove our main result.

**Theorem 5.3** $|MC_w| = c \cdot R(\mathbb{C}(x(t), y(t))) \cdot |\mathcal{M}|^2$, where $c$ is some non-zero constant.

**Proof:** From Lemma 5.5, we conclude that $|MC_w|$ has only two non-constant factors: $R(\mathbb{C}(x(t), y(t)))$ and $|\mathcal{M}|$. Therefore, there exist positive integers $p, q$ such that

$$
|MC_w| = c \cdot \left[ R(\mathbb{C}(x(t), y(t))) \right]^p \cdot |\mathcal{M}|^q. \tag{5.20}
$$

where $c$ is some non-zero constant.
Let us now examine the degrees of each of these determinants in the coefficients of \(x(t), y(t), w(t)\). First, consider \(|MC_w|\). Since each column of \(MC_w\) contains entries homogeneous in the coefficients of \(x(t), y(t), w(t)\), the determinant \(|MC_w|\) is a homogeneous polynomial in these coefficients. Specifically, \(|MC_w|\) is homogeneous of degree \(4m\) in the coefficients of \(x(t)\) and \(y(t)\), and homogeneous of degree \(2m\) in the coefficients of \(w(t)\). On the other hand, the resultant \(Res\left(x(t), y(t)\right)\) is homogeneous of degree \(2m\) in the coefficients of \(x(t)\) and \(y(t)\), and \(|ML|\) is homogeneous of degree \(m\) in the coefficients of \(x(t), y(t), w(t)\).

Comparing the homogeneous degrees in the coefficients of \(x(t), y(t), w(t)\) on both sides of Equation (5.20), we have the following equalities:

\[
2m \ast p + m \ast q = 4m. \quad \text{(in the coefficients of } x(t)\text{)}
\]
\[
2m \ast p + m \ast q = 4m. \quad \text{(in the coefficients of } y(t)\text{)}
\]
\[
m \ast q = 2m. \quad \text{(in the coefficients of } w(t)\text{)}
\]

It is easy to see that the only solution to these equalities is \(p = 1\) and \(q = 2\). The proof is therefore complete. \(\square\)

Note that in the original moving conic matrix \(MC\) (c. f. Equation (5.13)), if we discard the columns that represent the polynomials \(t^k x^2\) or \(t^k y^2\), \(0 \leq k \leq m - 1\), we can obtain similar results:

\[
|MC_x| = c \cdot Res\left(y(t), w(t)\right) \cdot |ML|^2.
\]
\[
|MC_y| = c \cdot Res\left(w(t), x(t)\right) \cdot |ML|^2.
\]

**Corollary 5.1** If a degree \(2m\) rational curve \(\left(x(t) : y(t) : w(t)\right)\) does not have base points, and there are no degree \(m - 1\) moving lines that follow \(\left(x(t) : y(t) : w(t)\right)\). then the method of moving conics always succeeds in producing the implicit equation for \(\left(x(t) : y(t) : w(t)\right)\).

**Proof:** First we observe that the vanishing of \(Res\left(x(t), y(t)\right)\) does not affect the success of the moving conics method. In fact, if the resultant \(Res\left(x(t), y(t)\right) = 0,\)
then $x(t)$ and $y(t)$ have at least one common root. But for any parameter $t_0$ such that $x(t_0) = 0, y(t_0) = 0$, we know that $w(t_0) \neq 0$ since, by assumption, there is no common root among $x(t), y(t), w(t)$. We can then translate the curve \( \left( x(t) : y(t) : w(t) \right) \) to \( \left( x(t) + \text{constant} \ast w(t) : y(t) : w(t) \right) \) so that $x(t) + \text{constant} \ast w(t)$ and $y(t)$ do not have common roots. If $F(X, Y) = 0$ is the implicit equation of the original curve, then the implicit equation of the translated curve is $F(X - \text{constant}, Y) = 0$. Therefore, to find the implicit equation of the original curve is equivalent to finding the implicit equation of the translated curve. Thus we can always assume, without lose of generality, that the resultant $Res\left( x(t), y(t) \right)$ is not zero.

Second, if there are no degree $m - 1$ moving lines that follow the rational curve \( \left( x(t) : y(t) : w(t) \right) \), then $|ML| \neq 0$. Therefore by Theorem 5.3, $|MC_w| \neq 0$. Write the linear system (5.13) (when $d = m - 1$) as

$$MC_w \begin{bmatrix} A_0 \\ \vdots \\ E_0 \\ \vdots \\ A_{m-1} \\ \vdots \\ E_{m-1} \end{bmatrix} = - [w^2, \ldots, l^{m-1}w^2]^C \begin{bmatrix} F_0 \\ \vdots \\ F_{m-1} \end{bmatrix}. \quad (5.21)$$

The $5m \times m$ matrix $[w^2, \ldots, l^{m-1}w^2]^C$ on the right hand side of Equation (5.21) has full rank $m$, because the columns of this matrix are linearly independent. Therefore, the system (5.21) has $m$ linearly independent solutions. Let $P_i$, $0 \leq i \leq m - 1$, be the solution of system (5.21) corresponding to setting

$$F_j = \begin{cases} 1 & j = i; \\ 0 & j \neq i. \end{cases}$$

Then

$$P_i = t^i w^2 + \text{terms without } w^2, \quad 0 \leq i \leq m - 1.$$
Therefore the coefficient matrix \([ P_0, \ldots, P_{m-1}]^C\) contains \(w^2\) only in the diagonal entries. Hence the determinant \(| P_0, \ldots, P_{m-1}|\) contains the term \(w^{2m}\). Thus, this determinant does not vanish identically. Since each entry in this determinant is quadratic in \(x, y, w\), the total degree of this determinant is at most \(2m\). Moreover, by construction, each column \(P_i, 0 \leq i \leq m-1\), is a moving conic that follows the curve, so for points on the curve, the rows are linearly dependent; hence this determinant is zero for points on the curve. On the other hand, the implicit equation of the degree \(2m\) rational curve \(\left( x(t) : y(t) : w(t) \right)\) is represented by a unique irreducible degree \(2m\) polynomial equation. Therefore, the determinant \(| P_0, \ldots, P_{m-1}|\) must be the implicit equation of the rational curve, so the method of moving conics succeeds when there is no moving line of degree \(m - 1\) that follows the curve. 

5.3.3 Moving Conics and Curves of Odd Degrees

In Section 5.3.1 and Section 5.3.2, we discussed a relationship between the moving line and moving conic coefficient matrices for rational curves of even degrees. For odd degree rational curves, similar propositions hold.

Consider a rational curve \(\left( x(t) : y(t) : w(t) \right)\) of degree \(2m + 1\). The moving line matrix

\[
ML = [ x, y, w, \ldots, t^m x, t^m y, t^m w]^C
\]

is of size \((3m + 2) \times (3m + 3)\). Thus there always exists a moving line of degree \(m\) that follows the rational curve. If there exists only one such independent moving line, then the matrix \(ML\) is of full rank, i.e. \(3m + 2\). That is, there exists one column of \(ML\) that is linearly dependent on the remaining \(3m + 2\) columns. The remaining \(3m + 2\) columns are linearly independent. Without loss of generality, we can assume that this dependent column is \(t^k w\). To get a square submatrix, we discard the column \(t^k w\) from \(ML\), and write the resulting submatrix as

\[
ML_w = [ x, y, w, \ldots, t^k(x, y), \ldots, t^m(x, y, w)]^C_{(3m+2) \times (3m+2)}.
\]
The moving conic matrix

\[ MC = [x^2, y^2, xy, xw, yw, w^2, \ldots, t^m(x^2, y^2, xy, xw, yw, w^2)] \]

is of size \((5m + 3) \times (6m + 6)\). Delete the columns that represent the polynomials \(t^j w^2, 0 \leq j \leq m\), and \(t^k xw, t^k yw\). The resulting square submatrix is

\[ MC_w = [x^2, y^2, xy, xw, yw, \ldots, t^k(x^2, y^2, xy), \ldots, t^m(x^2, y^2, xy, xw, yw)] \]

By an analysis similar to that of Section 5.3, we now have

\[ |MC_w| = c \cdot \text{Resultant}(x(t), y(t)) \cdot |ML_w|^2, \]

where \(c\) is some non-zero constant. It follows from this equation by an argument analogous to that in the proof of Corollary 5.1 that when there are no base points and when there exists only one independent moving line of degree \(m\) that follows the degree \(2m + 1\) curve \((x(t) : y(t) : w(t))\), the method of moving conics successfully generates the implicit equation for the rational curve.
Chapter 6

Implicitization Using Moving Surfaces

The classical method for finding the implicit equation of a rational parametric surface

\[ x = \frac{x(s,t)}{w(s,t)}, \quad y = \frac{y(s,t)}{w(s,t)}, \quad z = \frac{z(s,t)}{w(s,t)} , \]

is to compute the bivariate resultant of the three polynomials:

\[ x(s,t) - x \cdot w(s,t), \quad y(s,t) - y \cdot w(s,t), \quad z(s,t) - z \cdot w(s,t). \]

Unfortunately for many applications, the classical resultant of these three polynomials vanishes identically when the surface has base points - that is, parameter values \((s_0, t_0)\) for which

\[ x(s_0, t_0) = 0, \quad y(s_0, t_0) = 0, \quad z(s_0, t_0) = 0, \quad w(s_0, t_0) = 0. \]

Moreover sparse resultants are not always readily available. Several years ago, Professor Tom Sederberg introduced a new technique for finding the implicit equation of a rational surface which he called the method of moving quadrics [52]. This technique is essentially an extension to surfaces of the method of moving conics discussed in Chapter 5. Extensive experiments reveal that Sederberg’s technique is generally impervious to base points. Thus the method of moving quadrics promised to be a substantial improvement over classical methods based on resultants, which could implicitize rational surfaces with base points only by invoking rather complicated perturbation techniques [42]. Sederberg’s method of moving quadrics uses only elementary linear algebra — solving a system of linear equations — and as an added bonus represents the implicit equation as the determinant of a matrix one-fourth the size of the classical resultant.
Nevertheless, although there is substantial empirical evidence that the method of moving quadrics frequently works in practice, there is no rigorous proof that the method is correct in theory. Nor is there any systematic analysis of precisely under what conditions the method might fail or how to recover gracefully when it does.

This chapter is the first step in such an analysis. Here we show that if the rational surface has no base points, then the method of moving quadrics will succeed provided that there is no moving plane of low degree that follows the surface. Since the existence of such a moving plane of low degree is represented by a polynomial condition, this result establishes that the method of moving quadrics works for almost all rational surfaces without base points. A proof for surfaces where base points are present is still an open problem.

Even in the relatively simple setting of no base points, the proof is not elementary. For curves, the proof requires only some standard linear algebra, generic properties of resultants, and a few simple facts about factoring univariate polynomials [Ch 5] [63]. But for surfaces we need to study syzygies of bivariate polynomials — that is, polynomial relationships between polynomial functions. The algebraic analysis of such relationships requires sophisticated tools from algebraic geometry and commutative algebra, including Cohen-Macaulay rings, R-sequences, Koszul complexes, and sheaf cohomology [22].

In this chapter, we will first review Sederberg's method of moving surfaces. Then we will discuss implicitizing rational surfaces of bi-degree \((m, n)\) using moving planes. We will show that, generically, the implicit expression can be obtained by taking the determinant of a particular collection of \(2mn\) moving planes. This determinant has the style of the Sylvester resultant — striped and shifted columns and lots of zeros — and the size of Dixon/Cayley resultant. The construction is a generalization of the moving line method which we discussed in Section 5.2. The remainder of the chapter is devoted to the method of moving quadrics. In Section 6.3 we discuss resultants and syzygies for tensor product polynomials and triangular polynomials. Two important
propositions are provided in this section. In the final two sections we establish the validity of the method of moving quadrics for implicitizing rational tensor product surfaces and rational triangular surfaces when these surfaces have no base points.

### 6.1 The Method of Moving Surfaces

Now let us review briefly Sederberg’s method of moving surfaces with particular emphasis on moving planes and moving quadrics. Here we focus on rational tensor product surfaces — that is, rational parametric surfaces
\[ \begin{pmatrix} x(s, t) \\ y(s, t) \\ z(s, t) \\ w(s, t) \end{pmatrix}, \]

where
\[ x(s, t) = \sum_{i=0}^{m} \sum_{j=0}^{n} a_{i,j} s^i t^j, \quad y(s, t) = \sum_{i=0}^{m} \sum_{j=0}^{n} b_{i,j} s^i t^j, \]
\[ z(s, t) = \sum_{i=0}^{m} \sum_{j=0}^{n} c_{i,j} s^i t^j, \quad w(s, t) = \sum_{i=0}^{m} \sum_{j=0}^{n} d_{i,j} s^i t^j. \]  

(6.1)

We will only consider surfaces without base points. Recall that a base point for surface (6.1) is a pair of parameters \((s_0, t_0)\) such that
\[ x(s_0, t_0) = y(s_0, t_0) = z(s_0, t_0) = w(s_0, t_0) = 0. \]

It is well known that in projective space the bi-degree patch (6.1) always has \(n^2\) base points at \(s = \infty\) and \(m^2\) base points at \(t = \infty\). So when we say surface (6.1) has no base points, we mean that the surface does not have any additional base points.

A moving plane of bi-degree \((\sigma_1, \sigma_2)\) is an implicit equation of the form
\[ \sum_{i=0}^{\sigma_1} \sum_{j=0}^{\sigma_2} (A_{i,j} x + B_{i,j} y + C_{i,j} z + D_{i,j} w) \cdot s^i t^j = 0. \]  

(6.2)

For each fixed value of \(s\) and \(t\), Equation (6.2) is the implicit equation of a plane in \(\mathbb{C}^3\). Similarly, a moving quadric of bi-degree \((\sigma_1, \sigma_2)\) is an implicit equation of the form
\[ \sum_{i=0}^{\sigma_1} \sum_{j=0}^{\sigma_2} (A_{i,j} x^2 + B_{i,j} y^2 + C_{i,j} z^2 + D_{i,j} xy + E_{i,j} xz + F_{i,j} yz + G_{i,j} xw + H_{i,j} yw + I_{i,j} zw + J_{i,j} w^2) \cdot s^i t^j = 0. \]  

(6.3)
Again, when \( s, t \) are fixed, Equation (6.3) is the implicit equation of a quadric in \( \mathbb{C}^3 \).

The moving plane (6.2) or moving quadric (6.3) is said to follow surface (6.1) if

\[
\sum_{i=0}^{\sigma_1} \sum_{j=0}^{\sigma_2} (A_{i,j} x(s,t) + B_{i,j} y(s,t) + C_{i,j} z(s,t) + D_{i,j} w(s,t)) \cdot s^i t^j \equiv 0, \tag{6.4}
\]
or

\[
\sum_{i=0}^{\sigma_1} \sum_{j=0}^{\sigma_2} (A_{i,j} x^2(s,t) + B_{i,j} y^2(s,t) + \cdots + J_{i,j} w^2(s,t)) \cdot s^i t^j \equiv 0. \tag{6.5}
\]

By equating the coefficients of all the monomials \( s^i t^j \) in Equations (6.4) or (6.5) to zero, we obtain a system of linear equations in the indeterminates \( \{A_{i,j}, B_{i,j}, C_{i,j}, D_{i,j}\} \) or \( \{A_{i,j}, B_{i,j}, C_{i,j}, D_{i,j}, E_{i,j}, F_{i,j}, G_{i,j}, H_{i,j}, I_{i,j}, J_{i,j}\} \), \( 0 \leq i \leq \sigma_1, 0 \leq j \leq \sigma_2 \). Solving this system gives us a collection of moving planes or moving quadrics that follow surface (6.1). For certain special values of \( \sigma_1, \sigma_2 \), it turns out that we can find \( (\sigma_1 + 1) \cdot (\sigma_2 + 1) \) linearly independent moving planes or moving quadrics. The method of moving planes (or moving quadrics) then constructs the implicit equation of the parametric surface (6.1) by taking the determinant of the coefficient matrix of the \( (\sigma_1 + 1) \cdot (\sigma_2 + 1) \) linearly independent moving planes (or moving quadrics) that follow the parametric surface (6.1).

Specifically, for moving planes, we can choose \( \sigma_1 = 2m - 1, \sigma_2 = n - 1 \). Then from Equation (6.4) we obtain a homogeneous linear system of \( 6mn \) equations with \( 8mn \) unknowns. This system has at least \( 2mn \) linearly independent solutions:

\[
L_1 \equiv \sum_{i=0}^{2m-1} \sum_{j=0}^{n-1} (A_{i,j}^l x + B_{i,j}^l y + C_{i,j}^l z + D_{i,j}^l w) s^i t^j = 0.
\]

\[
\vdots
\]

\[
L_{2mn} \equiv \sum_{i=0}^{2m-1} \sum_{j=0}^{n-1} (A_{i,j}^{2mn} x + B_{i,j}^{2mn} y + C_{i,j}^{2mn} z + D_{i,j}^{2mn} w) s^i t^j = 0.
\]

Each of these solutions is a moving plane that follows surface (6.1). The determinant
of the coefficients of \( s^i t^j \) of these \( 2mn \) moving planes, i.e.

\[
\begin{vmatrix}
A_{0,0}^1 x + B_{0,0}^1 y + C_{0,0}^1 z + D_{0,0}^1 w \\
\vdots \\
A_{2m-1, n-1}^1 x + \cdots + D_{2m-1, n-1}^1 w \\
A_{0,0}^{2mn} x + B_{0,0}^{2mn} y + C_{0,0}^{2mn} z + D_{0,0}^{2mn} w \\
\vdots \\
A_{2m-1, n-1}^{2mn} x + \cdots + D_{2m-1, n-1}^{2mn} w
\end{vmatrix}
\]

vanishes whenever \((x, y, z, w)\) lies on the surface. Hence if this determinant does not vanish identically, then this determinant is a multiple of the implicit equation of surface (6.1). Moreover, the degree in \( x, y, z, w \) of this determinant is \( 2mn \), which is the generic degree of surface (6.1). Therefore, this determinant is a good candidate for representing the implicit equation of surface (6.1).

For moving quadrics, we can choose \( \sigma_1 = m - 1, \sigma_2 = n - 1 \). Then from Equation (6.5) we obtain a homogeneous linear system of \( 9mn \) equations with \( 10mn \) unknowns. This system has at least \( mn \) linearly independent solutions

\[
Q_1 \equiv \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} (A_{i, j}^1 x^2 + B_{i, j}^1 y^2 + \cdots + J_{i, j}^1 w^2) s^i t^j = 0.
\]

\[
\vdots
\]

\[
Q_{mn} \equiv \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} (A_{i, j}^{mn} x^2 + B_{i, j}^{mn} y^2 + \cdots + J_{i, j}^{mn} w^2) s^i t^j = 0.
\]

Each solution is a moving quadric following surface (6.1). The determinant

\[
\begin{vmatrix}
A_{0,0}^1 x^2 + B_{0,0}^1 y^2 + \cdots + J_{0,0}^1 w^2 \\
\vdots \\
A_{m-1, n-1}^1 x^2 + \cdots + J_{m-1, n-1}^1 w^2 \\
A_{0,0}^{mn} x^2 + B_{0,0}^{mn} y^2 + \cdots + J_{0,0}^{mn} w^2 \\
\vdots \\
A_{m-1, n-1}^{mn} x^2 + \cdots + J_{m-1, n-1}^{mn} w^2
\end{vmatrix}
\]

of the coefficients of \( s^i t^j \) of these \( mn \) moving quadrics is again a good candidate for representing the implicit equation of surface (6.1) because this determinant vanishes when the point \((x, y, z, w)\) lies on the surface and the degree of this determinant is \( 2mn \) in \( x, y, z, w \).
Example 6.1 (Moving Planes)

Consider the rational surface
\[ x(s, t) = st + 1, \]
\[ y(s, t) = s, \]
\[ z(s, t) = t, \]
\[ w(s, t) = s + t + 1. \]  \hspace{1cm} (6.6)

It is easy to find two moving planes following (6.6):
\[ (w - x - y - z) + s \cdot z = 0, \]
\[ (w - x - 2y - z) + s \cdot (w - y) = 0. \]

The determinant
\[ \begin{vmatrix} w - x & y - z \\ w - x - 2y - z & w - y \end{vmatrix} = w^2 - wx - 2wy + xy + y^2 - 2wz + xz + 3yz + z^2 \]
gives the implicit equation for surface (6.6).

Example 6.2 (Moving Quadrics)

Consider the rational surface
\[ x(s, t) = s^2 t^2 + 1, \]
\[ y(s, t) = s^2 + s + 1, \]
\[ z(s, t) = t + 1, \]
\[ w(s, t) = s^2 + t^2 + 1. \]  \hspace{1cm} (6.7)

Choose \( \sigma_1 = \sigma_2 = 1 \). Using Mathematica we find that there are four moving quadrics of bi-degree \((1, 1)\) that follow surface (6.7):
\[ Q_1 \equiv (-3wy + 3y^2 - wz + 3xz + 2yz - 4z^2) + s (wz - 4yz - 2z^2) + t (-3wy + 3y^2 + 4z^2) + st (2z^2) = 0. \]
\[ Q_2 \equiv (-3w^2 + 6wx + 20wy - 3xy - 19y^2 - 9wz - 12xz - 4yz + 24z^2) \]
\begin{align*}
&\quad + s (3ux - 2uy - 3xy + 2y^2 - 18wz + 3xz + 34yz + 6z^2) \\
&+ t (18w^2y - 18y^2 + 3wz - 6xz - 8yz - 12z^2) \\
&+ st (-3xz + 2yz) = 0, \\
Q_3 &\equiv (-2w^2y + 3y^2 + 3z^2) + s (wy - y^2 + wz - xz - 4yz) \\
&+ t (-2w^2y + 3y^2 + 2z^2) + st (wy - y^2) = 0, \\
Q_4 &\equiv (-5w^2 + 9wx + 34wy - 3xy - 33y^2 - 11wz - 21xz - 12yz + 42z^2) \\
&+ s (-w^2 + 6wx - 2w^2y - 6xy + 3y^2 - 31wz + 6xz + 60yz + 12z^2) \\
&+ t (30w^2y - 30y^2 + 5wz - 9xz - 12yz - 24z^2) + st (wz - 6xz) = 0.
\end{align*}

The implicit equation is obtained by taking the determinant of the coefficients of 1.

**6.2 Implicitizing Surfaces Using Moving Planes**

In this section, we extend the moving line technique to the method of moving planes in order to implicitize rational tensor product surfaces. We begin by reviewing two
resultant formulations for three tensor product polynomials: the \textit{Sylvester resultant} and the \textit{Cayley resultant}. Using moving planes, we then implicitize rational tensor product surfaces by constructing a determinant which has the style of the Sylvester resultant and the size of the Cayley resultant. This result mimics a similar result for rational curves presented in Chapter 5, Section 5.2.

For three polynomials of bi-degree \((m, n)\)

\[
f(s, t) = \sum_{i=0}^{m} \sum_{j=0}^{n} a_{i,j} s^i t^j, \quad g(s, t) = \sum_{i=0}^{m} \sum_{j=0}^{n} b_{i,j} s^i t^j, \quad h(s, t) = \sum_{i=0}^{m} \sum_{j=0}^{n} c_{i,j} s^i t^j.
\]

the resultant has many different representations [13] [25] [61]. Here we review two of these representations — the Sylvester resultant and the Cayley resultant — which are the analogues of the Sylvester and Bézout resultants in the univariate setting — see also Chapter 3.

The Sylvester resultant is the determinant of the coefficient matrix

\[
Syl(f, g, h) = [f, g, h, \ldots, t^{n-1}(f, g, h), \ldots, s^{2m-1}(f, g, h), \ldots, s^{2n-1} t^{n-1}(f, g, h)]^T.
\]

That is.

\[
\begin{bmatrix}
f & g & h & \ldots & s^{2m-1} t^{n-1}(f, g, h)
\end{bmatrix} = \begin{bmatrix}
1 & \ldots & t^{2n-1} & \ldots & s^{3m-1} & \ldots & s^{3m-1} t^{2n-1}
\end{bmatrix} \cdot Syl(f, g, h).
\]

The Cayley resultant is the determinant of the coefficient matrix \(Cay(f, g, h)\), where \(Cay(f, g, h)\) is defined by

\[
\begin{bmatrix}
f(s, t) & g(s, t) & h(s, t) \\
f(\alpha, t) & g(\alpha, t) & h(\alpha, t) \\
f(\alpha, \beta) & g(\alpha, \beta) & h(\alpha, \beta)
\end{bmatrix}_{(\alpha - s)(\beta - t)} = \begin{bmatrix} 1 & \cdots & \alpha^{2m-1} \\
\vdots & \ddots & \vdots \\
\alpha^{2m-1} & \cdots & \alpha^{2m-1} \end{bmatrix} \cdot Cay(f, g, h) \cdot \begin{bmatrix} 1 \\
\vdots \\
\alpha^{2m-1} \\
\vdots \\
\alpha^{2m-1} \end{bmatrix},
\]
The non-zero entries of $\text{Syl}(f, g, h)$ are very simple, since these entries come directly from the coefficients of $f, g, h$. The matrix $\text{Syl}(f, g, h)$ is also sparse, which is useful for computations. But $\text{Syl}(f, g, h)$ is very large — of size $6mn \times 6mn$. On the other hand, the matrix $\text{Cay}(f, g, h)$ is much smaller — only of size $2mn \times 2mn$; but $\text{Cay}(f, g, h)$ has very complicated entries [Ch 3] [17]. Ideally we would like to combine the advantages of both the Sylvester resultant and the Cayley resultant to implicitly rational tensor product surfaces.

Here we will consider only tensor product surfaces $(x(s, t) : y(s, t) : z(s, t) : w(s, t))$

\[
\begin{align*}
x(s, t) &= \sum_{i=0}^{m} \sum_{j=0}^{n} a_{i,j} s^i t^j, \\
y(s, t) &= \sum_{i=0}^{m} \sum_{j=0}^{n} b_{i,j} s^i t^j, \\
z(s, t) &= \sum_{i=0}^{m} \sum_{j=0}^{n} c_{i,j} s^i t^j, \\
w(s, t) &= \sum_{i=0}^{m} \sum_{j=0}^{n} d_{i,j} s^i t^j.
\end{align*}
\]

(6.8)

without base points.

The resultant of $X \cdot w(s, t) - x(s, t)$, $Y \cdot w(s, t) - y(s, t)$, and $Z \cdot w(s, t) - z(s, t)$, if not identically zero, is the implicit equation of surface (6.8). Since both the Sylvester resultant and the Cayley resultant have certain advantages, ideally we would like to construct an implicitization matrix which has the style of the Sylvester resultant $\text{Syl}(f, g, h)$ and the size of the Cayley resultant $\text{Cay}(f, g, h)$. This goal can be achieved by applying the method of moving planes.

Notice that the equations

\[
\begin{align*}
x \cdot w(s, t) - w \cdot x(s, t) &= 0, \\
y \cdot w(s, t) - w \cdot y(s, t) &= 0, \\
z \cdot w(s, t) - w \cdot z(s, t) &= 0,
\end{align*}
\]

or equivalently.

\[
\begin{align*}
X \cdot w(s, t) - x(s, t) &= 0, \quad Y \cdot w(s, t) - y(s, t) = 0, \quad Z \cdot w(s, t) - z(s, t) = 0,
\end{align*}
\]

are three moving planes of bi-degree $(m, n)$ that follow surface (6.8). Thus the standard way to find the implicit equation of a rational tensor product surface of bi-degree
\((m, n)\) is to compute the resultant of three moving planes of bi-degree \((m, n)\) that follow the surface.

Now consider instead the moving planes of bi-degree \((m - 1, n)\)

\[
\sum_{i=0}^{m-1} \sum_{j=0}^{n} (A_{i,j}x + B_{i,j}y + C_{i,j}z + D_{i,j}w) s^i t^j = 0
\]  

(6.9)

that follow the bi-degree \((m, n)\) surface (6.8). Substituting \(x = x(s, t), y = y(s, t), z = z(s, t), w = w(s, t)\) into Equation (6.9), and setting the coefficients of \(s^k t^l\), \(0 \leq k \leq 2m - 1, 0 \leq l \leq 2n\), to zero generates a linear system of \(2m(2n+1)\) equations with \(4m(n+1)\) unknowns. Therefore, there are always at least \(4m(n+1) - 2m(2n+1) = 2m\) linearly independent solutions to this linear system. From these 2m solutions, we can construct 2m moving planes that follow surface (6.8):

\[
p_1 = \sum_{i=0}^{m-1} \sum_{j=0}^{n} (A_{i,j}^1x + B_{i,j}^1y + C_{i,j}^1z + D_{i,j}^1w) s^i t^j = 0.
\]

\[
p_2 = \sum_{i=0}^{m-1} \sum_{j=0}^{n} (A_{i,j}^2x + B_{i,j}^2y + C_{i,j}^2z + D_{i,j}^2w) s^i t^j = 0.
\]

\[
p_{2m} = \sum_{i=0}^{m-1} \sum_{j=0}^{n} (A_{i,j}^{2m}x + B_{i,j}^{2m}y + C_{i,j}^{2m}z + D_{i,j}^{2m}w) s^i t^j = 0.
\]

Let \(M\) be the matrix whose columns consist of the coefficients of \(s^k t^l\), \(0 \leq l \leq 2n - 1, 0 \leq k \leq m - 1\), of the \(2mn\) moving planes \(p_j, \ldots, p_{n-1} p_j, 1 \leq j \leq 2m\). That
is,

\[
M = \begin{bmatrix}
A_{0,0}^0 x + \cdots + D_{0,0}^0 w \\
\vdots \\
A_{0,n-1}^1 x + \cdots + D_{0,n-1}^1 w \\
A_{0,n}^1 x + \cdots + D_{0,n}^1 w \\
A_{0,0}^2 x + \cdots + D_{0,0}^2 w \\
\vdots \\
A_{0,n-1}^2 x + \cdots + D_{0,n-1}^2 w \\
A_{0,n}^2 x + \cdots + D_{0,n}^2 w \\
\vdots \\
A_{m-1,0}^1 x + \cdots + D_{m-1,0}^1 w \\
\vdots \\
A_{m-1,n-1}^1 x + \cdots + D_{m-1,n-1}^1 w \\
A_{m-1,n}^1 x + \cdots + D_{m-1,n}^1 w \\
A_{m-1,0}^2 x + \cdots + D_{m-1,0}^2 w \\
\vdots \\
A_{m-1,n-1}^2 x + \cdots + D_{m-1,n-1}^2 w \\
A_{m-1,n}^2 x + \cdots + D_{m-1,n}^2 w
\end{bmatrix}
\]

Here the rows of \( M \) are indexed by

\[1, \ldots, t^{2n-1}, s, \ldots, s t^{2n-1}, \ldots, s^{m-1}, \ldots, s^{m-1} t^{2n-1}.\]

Thus \( M \) is a square matrix of order \( 2mn \). Moreover, \( M \) is sparse, since there are lots of zero entries. Hence \( M \) is a matrix in the style of the Sylvester resultant with the order of the Cayley resultant.

It is easy to see that \( |M| = 0 \) whenever a point \((x : y : z : w)\) lies on surface (6.8), since each column represents a moving plane that follows surface (6.8). Moreover, \( |M| \) is at most degree \( 2mn \) in \( x, y, z, w \) because the entries of \( M \) are linear in \( x, y, z, w \). Since, by assumption, surface (6.8) has no base points, the irreducible implicit equation of surface (6.8) is of degree \( 2mn \) thus \( |M| = 0 \) must be a constant multiple of the implicit equation if \( |M| \neq 0 \). Below we shall show that \( |M| \) — as a polynomial in \( x, y, z, w \) — is not, in general, identically zero.
By construction,

$$|M| = \sum_{1 + j + k + l = 2mn} f_{i,j,k,l}(a, b, c, d) x^i y^j z^k w^l.$$ 

where the functions \( \{f_{i,j,k,l}(a, b, c, d)\} \) are polynomials in the coefficients \( a_{u,v}, b_{u,v}, c_{u,v}, d_{u,v}, 0 \leq u \leq m, 0 \leq v \leq n \) of \( x, y, z, w \). Thus \( |M| \) can vanish identically only if the polynomials \( \{f_{i,j,k,l}(a, b, c, d)\} \) vanish identically. To show that this does not happen, we need only exhibit a single surface for which \( |M| \neq 0 \).

Consider the surface

$$x = s^m, \quad y = t^n, \quad z = s^m t^n + 1, \quad w = 1. \quad (6.10)$$

Then

$$l_i = (-z + w)s^i + xs^it^n, \quad r_i = ys^i - ws^it^n.$$

\( 0 \leq i \leq m - 1 \), are \( 2m \) linearly independent moving planes of bi-degree \((m - 1, n)\) that follow surface (6.10).

Placing the coefficients of the monomials \( s^k t^l \), \( 0 \leq l \leq 2n - 1, 0 \leq k \leq m - 1 \), of the polynomials \( \ell l_i, \ell r_i, 0 \leq j \leq n - 1, 0 \leq i \leq m - 1 \), into a matrix, we obtain

$$M = \begin{bmatrix}
(-z + w)I_n & yI_n \\
xI_n & -wI_n \\
& \ddots \\
& & (-z + w)I_n & yI_n \\
xI_n & -wI_n
\end{bmatrix},$$

where \( I_n \) is the identity matrix of order \( n \) and the columns of \( M \) are indexed by

\( l_0, \ldots, t^{n-1}l_0, r_0, \ldots, t^{n-1}r_0, \ldots, l_{m-1}, \ldots, t^{n-1}l_{m-1}, r_{m-1}, \ldots, t^{n-1}r_{m-1} \).

Hence in this case, \( |M| = (wz - w^2 - xy)^{mn} \). Thus as a polynomial in \( a_{u,v}, b_{u,v}, c_{u,v}, d_{u,v} \), \( |M| \neq 0 \).
It may still happen that for some specific choices of the coefficients $a_{u,v}$, $b_{u,v}$, $c_{u,v}$, $d_{u,v}$ the polynomials \{$f_{i,j,k,l}$\} all vanish simultaneously and hence, in these cases, $|M|$ is identically zero. These particular coefficients, satisfying the equations $f_{i,j,k,l}(a, b, c, d) = 0$, form a low dimensional variety in the high dimensional space of all rational surface of bi-degree $(m, n)$ with no base points and exactly $2m$ linearly independent moving planes that follow the surface. Since this low-dimensional variety has measure zero, $|M|$ almost never vanishes identically. Hence $|M| = 0$ is almost always the implicit equation of the tensor product surface (6.8). When surface (6.8) has no base points and has exactly $2m$ linearly independent moving planes of bi-degree $(m - 1, n)$ that follow the surface.

Moreover, surface (6.8) contains base points or has more than $2m$ linearly independent moving planes of bi-degree $(m - 1, n)$ that follow the surface, only if the coefficients $a_{u,v}, b_{u,v}, c_{u,v}, d_{u,v}$ satisfy some polynomial conditions; hence these coefficients again form a low dimensional variety of measure zero in the space of all rational surfaces of bi-degree $(m, n)$. Therefore, $|M| = 0$ is almost always the implicit equation of the rational tensor product surface (6.8). Thus we have the following

**Theorem 6.1** The method of moving planes almost always computes the implicit equation of (6.8) from a determinant in the style of the Sylvester resultant with the size of the Cayley resultant.

### 6.3 Resultants and Syzygies

This section will discuss some interesting algebra connected with bivariate resultants.

The propositions presented here will play an important role in the proof of our main theorems on moving quadrics later in this chapter.
6.3.1 Triangular Polynomials

We begin with three triangular polynomials

\[ x(s,t) = \sum_{i=0}^{n} \sum_{j=0}^{n-1} a_{i,j} s^i t^j, \quad y(s,t) = \sum_{i=0}^{n} \sum_{j=0}^{n-1} b_{i,j} s^i t^j, \quad z(s,t) = \sum_{i=0}^{n} \sum_{j=0}^{n-1} c_{i,j} s^i t^j \]  \hspace{1cm} (6.11)

of total degree \( n \). In this situation, there is a multi-variate resultant \( \text{Res}(x,y,z) \), which has the following well known geometric interpretation: \( \text{Res}(x,y,z) \neq 0 \) if and only if it is impossible to find a pair of parameters \((s_0, t_0)\) such that

\[ x(s_0, t_0) = y(s_0, t_0) = z(s_0, t_0) = 0. \]  \hspace{1cm} (6.12)

More precisely, this means that if we homogenize \( x, y, z \) by adding a third variable, then the homogenized equations have no common solutions in projective space.

This explains the geometry of the resultant, but what about the algebra? In other words, what is the algebraic interpretation of \( \text{Res}(x,y,z) \neq 0 \)? The answer is given by the following proposition.

**Proposition 6.1** Suppose that \( x, y, z \) are defined as in Equation (6.11) and that \( \text{Res}(x,y,z) \neq 0 \). Then, whenever we have polynomials \( A, B, C \in \mathbb{C}[s,t] \) satisfying

\[ Ax + By + Cz = 0, \]

there are polynomials \( h_1, h_2, h_3 \in \mathbb{C}[s,t] \) such that

\[ A = h_1 z + h_2 y, \]
\[ B = -h_2 x + h_3 z, \]
\[ C = -h_1 x - h_3 y. \]

Furthermore, if \( k \) is the maximum degree of \( A, B, C \), then \( h_1, h_2, h_3 \) can be chosen so that they have degree at most \( k - n \).

**Proof:** A proof can be found in [22].
Although the proof of Proposition 6.1 is difficult, we can make some simple observations about what this proposition really says. The equation \( Ax + By + Cz = 0 \) is called a syzygy on \( x, y, z \). If we think of a syzygy as a column vector \((A, B, C)^T\) then there are three obvious syzygies:

\[
\begin{align*}
(z, 0, -x)^T \quad & \text{coming from } z \cdot x + 0 \cdot y + (-x) \cdot z = 0, \\
(y, -x, 0)^T \quad & \text{coming from } y \cdot x + (-x) \cdot y + 0 \cdot z = 0, \\
(0, z, -y)^T \quad & \text{coming from } 0 \cdot x + z \cdot y + (-y) \cdot z = 0.
\end{align*}
\]

Furthermore, multiplying the first of these by \( h_1 \), the second by \( h_2 \), and the third by \( h_3 \), and then adding them together, we get the syzygy

\[
h_1(z, 0, -x)^T + h_2(y, -x, 0)^T + h_3(0, z, -y)^T = (h_1z + h_2y, -h_2x + h_3z, -h_1x - h_3y)^T.
\]

Proposition 6.1 tells us that when the resultant doesn’t vanish, all syzygies on \( x, y, z \) are generated from the obvious ones in this way.

### 6.3.2 Tensor Product Polynomials

We now turn our attention to tensor product polynomials. Consider three polynomials

\[
x(s, t) = \sum_{i=0}^{m} \sum_{j=0}^{n} a_{i,j} s^i t^j, \quad y(s, t) = \sum_{i=0}^{m} \sum_{j=0}^{n} b_{i,j} s^i t^j, \quad z(s, t) = \sum_{i=0}^{m} \sum_{j=0}^{n} c_{i,j} s^i t^j \quad (6.13)
\]

of bi-degree \((m, n)\). For these polynomials, the Dixon resultant \( Res(x, y, z) \) is nonzero if and only if it is impossible to find \((s_0, t_0)\) such that

\[
x(s_0, t_0) = y(s_0, t_0) = z(s_0, t_0) = 0. \quad (6.14)
\]

In the projective plane, these equations always have \(m^2\) solutions at \( s = \infty\) and \(n^2\) solutions at \( t = \infty\). So when we say that (6.14) has no solutions, we mean that there are no additional solutions.

In this situation, the analog of Proposition 6.1 would state that if \( Res(x, y, z) \neq 0 \), then any syzygy \((A, B, C)^T\) on \( x, y, z \) would have the form

\[
(A, B, C)^T = h_1(z, 0, -x)^T + h_2(y, -x, 0)^T + h_3(0, z, -y)^T.
\]
as in Proposition 6.1, and furthermore, if \( A, B, C \) had bi-degree at most \((k, l)\), then \( h_1, h_2, h_3 \) could be chosen to have bi-degree at most \((k - m, l - n)\). Unfortunately, the degree bound can fail in the tensor product case. Here is a simple example to show what can go wrong.

**Example 6.3**

Let \( x = st, y = st + s + t \) and \( z = st + 1 \). One easily checks that these polynomials do not vanish simultaneously, so their Dixon resultant is not zero. Now consider the syzygy

\[
(-s^2 + s + 1) \ast st + (-s) \ast (st + s + t) + s^2 \ast (st + 1) = 0.
\]

The polynomials \( x, y, z \) generate the unit ideal in \( \mathbb{C}[s, t] \), so that as above, Lemma 1 of Section 2 of [23] implies

\[
(-s^2 + s + 1, -s, s^2) = h_1(z, 0, -x)^T + h_2(y, -x, 0)^T + h_3(0, z, -y)^T
\]

for some \( h_1, h_2, h_3 \). However, \((-s^2 + s + 1, -s, s^2)^T\) has bi-degree \((2, 0)\) and \(x, y, z\) have bi-degree \((1, 1)\), so that \(h_1, h_2, h_3\) cannot have bi-degree \((2, 0) - (1, 1) = (1, -1)\).

This example shows that Proposition 6.1 does not hold in the tensor product case. However, the crucial observation is that it does hold for certain special bi-degrees, which are exactly the ones we will use later in Section 6.4. Here is the precise result we will need.

**Proposition 6.2** Suppose that \( x, y, z \) are three tensor product polynomials of bi-degree \((m, n)\) such that the Dixon resultant \(\text{Res}(x, y, z)\) is nonzero. Also assume that \( (A, B, C)^T \) is a syzygy on \( x, y, z \) of bi-degree \((2m - 1, 2n - 1)\). Then there are polynomials \((h_1, h_2, h_3)\) of bi-degree at most \((m - 1, n - 1)\) such that

\[
\begin{align*}
A &= h_1z + h_2y, \\
B &= -h_2x + h_3z, \\
C &= -h_1x + h_3y.
\end{align*}
\]
Proof: A proof is provided in [22].

6.4 Implicitizing Tensor Product Surfaces Using Moving Quadrics

6.4.1 Moving Plane and Moving Quadric Coefficient Matrices

Consider the rational tensor product surface (6.1). Let $MP$ be the coefficient matrix of the polynomials $s^i t^j x, s^i t^j y, s^i t^j z, s^i t^j w, 0 \leq i \leq m - 1, 0 \leq j \leq n - 1:

\[
\begin{bmatrix}
  x & y & z & w & \cdots & s^{m-1} t^{n-1} x & s^{m-1} t^{n-1} y & s^{m-1} t^{n-1} z & s^{m-1} t^{n-1} w \\
  1 & \cdots & t^{2n-1} & \cdots & s^{2m-1} & \cdots & s^{2m-1} t^{2n-1}
\end{bmatrix} \cdot MP.
\]

That is, the columns of $MP$ are indexed by the polynomials

$s^i x, s^i y, s^i z, s^i w, \cdots, s^i t^{n-1} x, s^i t^{n-1} y, s^i t^{n-1} z, s^i t^{n-1} w, 0 \leq i \leq m - 1,$

and the rows are indexed by the monomials

$1, \cdots, t^{2n-1}, \cdots, s^{2m-1}, \cdots, s^{2m-1} t^{2n-1},$

so that the entries of $MP$ are the coefficients with respect to $s, t$ of the polynomials $s^i t^j x, s^i t^j y, s^i t^j z, s^i t^j w, 0 \leq i \leq m - 1, 0 \leq j \leq n - 1$. Then $MP$ is a square matrix of order $4mn$. Note that $MP$ is the coefficient matrix of the linear system generated by the moving planes of bi-degree $(m - 1, n - 1)$ that follow the rational surface (6.1).

Similarly, let $MQ$ be the coefficient matrix of the polynomials $s^i t^j x^2, s^i t^j y^2, \cdots, s^i t^j w^2, 0 \leq i \leq m - 1, 0 \leq j \leq n - 1$. That is,

\[
\begin{bmatrix}
  x^2 & y^2 & z^2 & xy & xz & yz & xw & yw & zw & w^2 & \cdots & s^{m-1} t^{n-1}(x^2 & \cdots & w^2)
\end{bmatrix} \cdot MQ.
\]

Then $MQ$ is a matrix of size $9mn \times 10mn$. Moreover, $MQ$ is the coefficient matrix of the linear system generated by the moving quadrics of bi-degree $(m - 1, n - 1)$ that
follow surface (6.1). Let \( MQ_w \) be the submatrix of \( MQ \) with the coefficients of \( st tw^2 \) deleted: \( MQ_w \) is then a square matrix of order \( 9mn \). When \( MQ_w \) is non-singular, the linear system of moving quadrics of bi-degree \((m - 1, n - 1)\) has exactly \( mn \) linearly independent solutions.

It is easy to see that when \( |MP| \) vanishes, \( |MQ_w| \) also vanishes, because linear dependencies on the columns of \( MP \) generate linear dependencies on the columns of \( MQ_w \). Below we shall show that the converse is also true when \( Res(x, y, z) \neq 0 \).

### 6.4.2 The Validity of the Method of Moving Quadrics for Tensor Product Surfaces

To establish the validity of the method of moving quadrics for implicitizing rational tensor product surfaces, we begin by observing that if surface (6.1) has no base points, then for the purpose of implicitization, we can assume that \( Res(x, y, z) \neq 0 \). The reason is as follows: if \( Res(x, y, z) = 0 \), then \( x, y, z \) have a common root (either affine or at infinity); since \( w \) does not vanish at this common root, the polynomials \( x + cw, y, z \) do not have a common root (either finite or infinite) for some constant \( c \). Thus \( Res(x + cw, y, z) \neq 0 \). The effect of such a transformation is just a simple translation of surface (6.1). It is easy to see that finding the implicit equation of the original surface is equivalent to finding the implicit equation of the shifted surface.

Moreover, if \( Res(x, y, z) \neq 0 \), then at least one of \( x(s, t), y(s, t), z(s, t) \) has a non-zero leading term, i.e. the coefficient of \( s^mt^n \) is not zero. Suppose, without loss of generality, that \( x(s, t) \) has a non-zero leading coefficient; then

\[
y(s, t) + constant \cdot x(s, t), \quad z(s, t) + constant \cdot x(s, t), \quad w(s, t) + constant \cdot x(s, t)
\]

all have non-zero leading coefficients. Again, these transformations only induce simple transformations on the coordinates of the original surface.

Therefore below, we will assume, without loss of generality, that \( Res(x, y, z) \neq 0 \) and that \( x(s, t), y(s, t), z(s, t), w(s, t) \) all have non-zero leading coefficients. Now for
rational tensor product surfaces. we have the following analogue of Lemma 5.5 for rational curves.

**Theorem 6.2** For the tensor product surface (6.1), if the resultant $\text{Res}(x, y, z) \neq 0$, then $|MQ_w| = 0$ implies $|MP| = 0$.

**Proof:** If $|MQ_w| = 0$, then the columns of $MQ_w$ are linearly dependent. Thus there exist $9mn$ scalars such that the linear combination of the $9mn$ columns of $MQ_w$ with these $9mn$ scalars is identically zero. We can write this linear combination as

$$p_1x^2 + p_2y^2 + p_3z^2 + p_4xy + p_5xz + p_6yz + p_7xz + p_8yw + p_9zw \equiv 0. \quad (6.15)$$

where each of the polynomials $p_i(s, t)$ consists of the coefficients of one of $x^2, y^2, \cdots, zw$. Each polynomial $p_i(s, t)$ is of bi-degree $(m - 1, n - 1)$, since $MQ_w$ consists of the coefficients of the polynomials:

$$x^2, y^2, z^2, xy, xz, yz, xw, yw, zw, \cdots, s^{m-1}t^{n-1}(x^2, y^2, z^2, xy, xz, yz, xw, yw, zw).$$

Rewrite Equation (6.15) as

$$(p_1x + p_4y + p_5z + p_7w)x + (p_2y + p_6z + p_8w)y + (p_3z + p_9w)z \equiv 0. \quad (6.16)$$

If $p_3(s, t) \neq 0$ or $p_9(s, t) \neq 0$, we want to prove that

$$p_3 \cdot z + p_9 \cdot w + h_1 \cdot x + h_3 \cdot y \equiv 0,$$

for some bi-degree $(m - 1, n - 1)$ polynomials $h_1(s, t), h_3(s, t)$ (the reason for labeling these polynomials as $h_1$ and $h_3$ will soon become clear). To prove this result, note that in the terminology of Section 6.3, Equation (6.16) is the syzygy on $x, y, z$ given by

$$(A, B, C)^T = (p_1x + p_4y + p_5z + p_7w, p_2y + p_6z + p_8w, p_3z + p_9w)^T.$$

Furthermore, since each $p_i(s, t)$ has bi-degree $(m - 1, n - 1)$, this syzygy has bi-degree $(2m - 1, 2n - 1)$. Since $\text{Res}(x, y, z) \neq 0$, Proposition 6.2 asserts that there are polynomials $h_1, h_3$ of bi-degree $(m - 1, n - 1)$ such that

$$C = -h_1 \cdot x - h_3 \cdot y.$$
Hence

\[ 0 \equiv C + h_1 \cdot x + h_3 \cdot y = p_3 \cdot z + p_5 \cdot w + h_1 \cdot x + h_3 \cdot y, \]

which proves that the columns of \( MP \) are linearly dependent. Therefore \( |MP| = 0 \).

If \( p_3(s, t) \equiv p_5(s, t) \equiv 0 \), then Equation (6.16) becomes

\[(p_2 x + p_4 y + p_5 z + p_7 w) x + (p_2 y + p_6 z + p_8 w) y \equiv 0. \quad (6.17)\]

Since \( \text{Res}(x, y, z) \neq 0 \), \( x, y \) cannot have any common factor. Otherwise, \( x, y, z \) would have common roots other than the \( m^2 + n^2 \) roots at \( s = \infty \) and \( t = \infty \) and hence \( \text{Res}(x, y, z) = 0 \) contrary to assumption. Therefore, if \( p_2 y + p_6 z + p_8 w \neq 0 \), then \( x \) must be a factor of \( p_2 y + p_6 z + p_8 w \). We examine the following two cases:

(i). If \( p_2 \neq 0 \), or \( p_6 \neq 0 \), or \( p_8 \neq 0 \), then Equation (6.17) implies that

\[ p_2 y + p_6 z + p_8 w = h(s, t) \cdot x. \]

for some polynomial \( h(s, t) \). But by assumption, \( x \) has non-zero leading coefficient, so it is easy to see that \( h(s, t) \) is of bi-degree \((m - 1, n - 1)\). It follows that the columns of \( MP \) are linearly dependent. since \( p_2, p_6, p_8, h \) are all of bi-degree \((m - 1, n - 1)\) in \( s, t \).

(ii). Otherwise, \( p_2 \equiv p_6 \equiv p_8 \equiv 0 \). Then by Equation (6.17).

\[ p_1 x + p_4 y + p_5 z + p_7 w = 0. \]

where \( p_1, p_4, p_5, p_7 \) are not all zero polynomials. This again proves that the columns of \( MP \) are linearly dependent. so \( |MP| = 0 \).

\[ \square \]

Since \( |MP| \neq 0 \) is equivalent to the fact that there are no moving planes of bi-degree \((m - 1, n - 1)\) following surface (6.1), we have the following

**Theorem 6.3** Suppose that surface (6.1) has no base points, that the parametrization is generically one-to-one, and that there are no moving planes of bi-degree \((m-1, n-1)\)
following surface (6.1). Then the method of moving quadrics computes the implicit equation of surface (6.1).

**Proof:** Assume without of loss of generality that \( Res(x, y, z) \neq 0 \). A bi-degree \((m - 1, n - 1)\) moving quadric has the form

\[
\sum_{i=0}^{m-1} \sum_{j=0}^{n-1} (A_{i,j}x^2 + B_{i,j}y^2 + C_{i,j}z^2 + D_{i,j}xy + \cdots + J_{i,j}w^2) s^it^j = 0. \quad (6.18)
\]

The moving quadric (6.18) follows surface (6.1) if and only if

\[
\sum_{i=0}^{m-1} \sum_{j=0}^{n-1} (A_{i,j}x^2(s, t) + B_{i,j}y^2(s, t) + C_{i,j}z^2(s, t) + \cdots + J_{i,j}w^2(s, t)) s^it^j \equiv 0. \quad (6.19)
\]

Thus \((A_{0,0}, \ldots, J_{0,0}, \ldots, A_{m-1,n-1}, \ldots, J_{m-1,n-1})^T\) is a solution to the linear system

\[
MQ \cdot (A_{0,0}, \ldots, J_{0,0}, \ldots, A_{m-1,n-1}, \ldots, J_{m-1,n-1})^T = 0. \quad (6.20)
\]

Rewrite Equation (6.20) as

\[
MQ_w \cdot \begin{pmatrix}
A_{0,0} \\
\vdots \\
I_{0,0} \\
\vdots \\
A_{m-1,n-1} \\
\vdots \\
I_{m-1,n-1}
\end{pmatrix} = -\text{Coeff}(w^2, \ldots, s^{m-1}t^{n-1}w^2) \cdot \begin{pmatrix}
J_{0,0} \\
\vdots \\
J_{m-1,n-1}
\end{pmatrix}. \quad (6.21)
\]

where \(\text{Coeff}(w^2, \ldots, s^{m-1}t^{n-1}w^2)\) consists of the coefficients of the polynomials \(s^it^jw^2\). \(0 \leq i \leq m - 1, 0 \leq j \leq n - 1\). That is, move all terms involving \(J_{i,j}\) on the left hand side in Equation (6.20) to the right hand side of Equation (6.21).

By assumption, \(|MP| \neq 0\) because there are no moving planes of bi-degree \((m - 1, n - 1)\) following surface (6.1). Since \(Res(x, y, z) \neq 0\), we know from Theorem 6.2 that \(MQ_w\) is also non-singular. Therefore, we can solve for \((A_{0,0}, \ldots, I_{0,0}, \ldots, J_{0,0}, \ldots, J_{m-1,n-1})\).
\( \ldots, A_{m-1,n-1}, \ldots, I_{m-1,n-1} \) from Equation (6.21). Thus we will get \( mn \) linearly independent solutions when we set \((J_{0,0}, \ldots, J_{m-1,n-1})\) to the standard unit vectors \(e_k = (0, \ldots, 0, 1, 0, \ldots, 0), 0 \leq k \leq mn - 1.\)

Now \( mn \) moving quadrics can be constructed from these \( mn \) solutions:

\[
Q_{\alpha,\beta} = \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} \left( A_{i,j}^{\alpha,\beta} x^2 + \cdots + J_{i,j}^{\alpha,\beta} w^2 \right) s^i t^j = 0, \quad 0 \leq \alpha \leq m-1, 0 \leq \beta \leq n-1.
\]

where \( Q_{\alpha,\beta} \) corresponds to setting

\[
(J_{0,0}, \ldots, J_{m-1,n-1}) = e_{m+\beta}.
\]

Thus

\[
Q_{\alpha,\beta} = w^2 s^\alpha t^\beta + \text{terms without } w^2, \quad 0 \leq \alpha \leq m-1, 0 \leq \beta \leq n-1.
\]

Therefore the coefficients of the monomials \( s^\alpha t^\beta \) from the moving quadrics \( Q_{\alpha,\beta} \) form an \( mn \times mn \) matrix

\[
M = \begin{bmatrix}
w^2 + \cdots \\
w^2 + \cdots \\
\vdots \\
w^2 + \cdots
\end{bmatrix},
\]

where each row consists of the coefficients of the monomials \( s^i t^j \) from a moving quadric \( Q_{\alpha,\beta} \). Our goal is to show that \( |M| = 0 \) is the implicit equation of the rational surface. Note that the off-diagonal entries of \( M \) are quadratic in \( x, y, z, w \) but do not have the term \( w^2 \). Since \( M \) contains \( w^2 \) only in the diagonal entries, \( |M| \) contains the term \( w^{2mn} \), so this determinant is not identically zero. Since each entry in \( M \) is quadratic in \( x, y, z, w \), the total degree of \( |M| \) is \( 2mn \) in \( x, y, z, w \). Moreover, by construction, each row \( Q_{\alpha,\beta}, 0 \leq \alpha \leq m-1, 0 \leq \beta \leq n-1, \) represents a moving quadric that follows surface \( (6.1) \), so for points on the surface, the columns of \( M \) are linearly dependent; hence \( |M| \) vanishes for points on surface \( (6.1) \). On the other hand, the degree of the
(irreducible) implicit equation of surface (6.1) is $2mn$ [23] since the parametrization is generically one-to-one. Therefore, $|M| = 0$ must be the implicit equation of the rational surface.

\[\square\]

### 6.5 Implicitizing Triangular Surfaces Using Moving Quadrics

In this section, we establish the validity of the method of moving quadrics for implicitizing surfaces of total degree $n$ with no base points.

A rational surface \( \begin{pmatrix} x(s, t) \\ y(s, t) \\ z(s, t) \\ w(s, t) \end{pmatrix} \) is of total degree $n$ if

\[
x(s, t) = \sum_{i+j \leq n} a_{i,j}s^it^j, \quad y(s, t) = \sum_{i+j \leq n} b_{i,j}s^it^j, \\
z(s, t) = \sum_{i+j \leq n} c_{i,j}s^it^j, \quad w(s, t) = \sum_{i+j \leq n} d_{i,j}s^it^j.
\]

(6.22)

Just as in the tensor product case, we consider moving planes and moving quadrics that follow this triangular surface. Here we shall be interested in moving planes and moving quadrics of total degree $n - 1$.

#### 6.5.1 Moving Plane and Moving Quadric Coefficient Matrices

A moving plane of degree $n - 1$

\[
\sum_{i+j \leq n-1} (A_{i,j}x + B_{i,j}y + C_{i,j}z + D_{i,j}w)s^it^j = 0
\]

(6.23)

follows surface (6.22) if it vanishes identically when the polynomials $x(s, t)$, $y(s, t)$, $z(s, t)$, $w(s, t)$ are substituted for $x, y, z, w$ in Equation (6.23).

Let $MP$ be the (moving plane) coefficient matrix of the polynomials $s^it^j(x, y, z, w)$, $i + j \leq n - 1$. That is,

\[
\begin{bmatrix} x & y & z & w & \cdots & t^{n-1}(x & y & z & w) & \cdots & s^{n-1}(x & y & z & w) \end{bmatrix} = \begin{bmatrix} 1 & t & s & \cdots & t^{2n-1} & \cdots & s^{2n-1} \end{bmatrix} \cdot MP.
\]
Note that the number of monomials \( s^i t^j \), where \( i + j \leq k \), is \( \binom{k+2}{2} \). Therefore,

\[
\# \text{ columns of } MP = 4 \cdot \binom{n+1}{2} = 2n^2 + 2n
\]

and

\[
\# \text{ rows of } MP = \binom{2n+1}{2} = 2n^2 + n.
\]

Thus there always exist at least \( n \) moving planes of degree \( n-1 \) that follow a rational parametric surface of total degree \( n \).

To obtain a square submatrix of \( MP \), let \( I \subset \{(k, l) \mid 0 \leq k + l \leq n - 1\} \) be a set of indices with \( |I| = n \). Define \( MP_I \) to be the coefficient matrix of the polynomials

\[
\begin{align*}
 s^i t^j (x, y, z), & \quad 0 \leq i + j \leq n - 1, \\
 s^i t^j w, & \quad (i, j) \notin I.
\end{align*}
\]

That is, to obtain \( MP_I \), we remove the \( n \) columns \( s^i t^j w, (i, j) \in I \) from \( MP \). Thus \( MP_I \) has

\[
4 \cdot \binom{n+1}{2} - n = 2n^2 + n
\]
columns and the same number of rows. Therefore \( MP_I \) is a square submatrix of \( MP \) of order \( 2n^2 + n \).

Similarly, a moving quadric of degree \( n - 1 \)

\[
\sum_{i+j \leq n-1} \left( A_{i,j} x^2 + B_{i,j} y^2 + C_{i,j} z^2 + D_{i,j} x y + E_{i,j} x z + F_{i,j} y z + G_{i,j} x w + H_{i,j} y w + I_{i,j} z w + J_{i,j} w^2 \right) s^i t^j = 0 \quad (6.24)
\]

follows surface (6.22) if it vanishes identically when the polynomials \( x(s, t), y(s, t), z(s, t), w(s, t) \) are substituted for \( x, y, z, w \) in Equation (6.24).

Let \( MQ \) be the (moving quadric) coefficient matrix of the polynomials \( s^i t^j (x^2, y^2, z^2, x y, x z, y z, x w, y w, z w, w^2), i + j \leq n - 1 \). That is,

\[
\begin{bmatrix}
 x^2 & y^2 & \ldots & w^2 & \ldots & l^{n-1} (x^2 y^2 \ldots w^2) & \ldots & s^{n-1} (x^2 y^2 \ldots w^2) \\
 1 & t & s & \ldots & l^{3n-1} & \ldots & s^{3n-1}
\end{bmatrix} \cdot MQ
\]
Then
\[
\# \text{ columns of } MQ = 10 \cdot \binom{n + 1}{2} = \frac{10n^2 + 10n}{2}
\]
and
\[
\# \text{ rows of } MQ = \binom{3n + 1}{2} = \frac{9n^2 + 3n}{2}.
\]
Thus there always exist at least \((n^2 + 7n)/2\) moving quadrics of degree \(n - 1\) that follow a parametric surface of total degree \(n\).

To get a square submatrix of \(MQ\), we remove all the \(s^{ij}w^2\) columns, \(0 \leq i + j \leq n - 1\), as well as the \(3n\) columns \(s^{ij}(xw, yw, zw)\), \((i, j) \in I\), and define \(MQ_I\) to be the coefficient matrix of the remaining polynomials
\[
\begin{align*}
\quad & s^{ij}(x^2, y^2, z^2, xy, xz, yz), & 0 \leq i + j \leq n - 1. \\
\quad & s^{ij}(xw, yw, zw), & (i, j) \notin I.
\end{align*}
\]
Thus \(MQ_I\) has
\[
9 \cdot \binom{n + 1}{2} - 3n = \frac{9n^2 + 3n}{2}
\]
columns and the same number of rows. Therefore, \(MQ_I\) is a square submatrix of \(MQ\) of order \((9n^2 + 3n)/2\).

### 6.5.2 The Method of Moving Quadrics for Triangular Surfaces

There are \((n^2 + n)/2\) monomials \(s^{ij}\), \(0 \leq i + j \leq n - 1\), in the expression for a moving quadric of degree \(n - 1\). Therefore, to compute the implicit equation of a degree \(n\) rational triangular surface by the method of moving quadrics in a manner similar to our approach in Section 6.4 for rational tensor product surfaces, we would need to construct a square matrix of order \((n^2 + n)/2\), whose entries consist of the coefficients (of \(s^{ij}\)) of \((n^2 + n)/2\) moving quadrics. Notice, however, that the degree in \(x, y, z, w\) of the determinant of such an \((n^2 + n)/2 \times (n^2 + n)/2\) matrix is \(n^2 + n\) since each entry is quadratic in \(x, y, z, w\), whereas the degree in \(x, y, z, w\) of the implicit equation of surface (6.22) is only \(n^2\). Therefore, the entries of this \((n^2 + n)/2 \times (n^2 + n)/2\) matrix cannot all be the coefficients of moving quadrics.
To lower the degree of this determinant, we shall replace \( n \) moving quadrics by \( n \) moving planes. From Section 6.5.1, we know that there always exist \( n \) linearly independent moving planes of degree \( n - 1 \) that follow surface (6.22), since \( MP \) is of size \((2n^2 + n) \times (2n^2 + 2n)\). Now this \((n^2 + n)/2 \times (n^2 + n)/2\) matrix consists of the coefficients (of \( s^H \)) of \((n^2 - n)/2\) moving quadrics and \( n \) moving planes. Therefore, the determinant of this square matrix is of degree \( n^2 = 2 \times (n^2 - n)/2 + n \times 1 \) in \( x, y, z, w \), which is exactly what we desire.

Nevertheless, we must be careful how we choose these moving quadrics and moving planes to make sure that the determinant is not identically zero. For example, if one of the moving planes is \( p(x, y, z, w) = 0 \) and one of the moving quadrics is \( q(x, y, z, w) \equiv x \cdot p(x, y, z, w) = 0 \), then the determinant will vanish identically. The method of moving quadrics asserts that, in general, it is possible to choose the \((n^2 - n)/2\) moving quadrics and \( n \) moving planes so that this determinant actually is the implicit equation of surface (6.22). Till now these moving planes and moving quadrics have been chosen in an \textit{ad hoc} manner. In the next section, we present a systematic way to choose the right moving quadrics and moving planes.

### 6.5.3 The Validity of the Method of Moving Quadrics for Triangular Surfaces

Proceeding in a manner similar to our approach to tensor product surfaces, below we first explore a relationship between the two coefficient matrices \( MP_1 \) and \( MQ_1 \). Then we shall show how to exploit this relationship to choose the right moving quadrics and moving planes to implicitize a triangular surface.

It is easy to see that if \(|MP_1|\) vanishes, then \(|MQ_1|\) also vanishes, because linear dependencies on the columns of \(MP_1\) generate linear dependencies on the columns of \(MQ_1\). Below we shall see that the converse is also true when \( Res(x, y, z) \neq 0 \).

**Theorem 6.4** For the triangular surface (6.22), if \( Res(x, y, z) \neq 0 \), then \(|MQ_1| = 0\) implies \(|MP_1| = 0\).
Proof: Suppose $|M_1 Q_1| = 0$. Then as in the proof of Theorem 6.2, there exist polynomials $p_1(s, t), \cdots, p_9(s, t)$ such that

$$p_1 x^2 + p_2 y^2 + p_3 z^2 + p_4 x y + p_5 x z + p_6 y z + p_7 x w + p_8 y w + p_9 z w = 0.$$ (6.25)

where $p_1, p_2, p_3, p_4, p_5, p_6, p_7, p_8, p_9$ are of total degree $n - 1$ in $s, t$, but the exponents of the monomials in $p_7, p_8, p_9$ are not in the index set $I$.

Rewrite Equation (6.25) as

$$(p_1 x + p_4 y + p_5 z + p_7 w) x + (p_2 y + p_6 z + p_8 w) y + (p_3 z + p_9 w) z = 0.$$ (6.26)

If $p_3 \neq 0$ or $p_9 \neq 0$, then Equation (6.26) is a syzygy on $x, y, z$. Hence by Proposition 6.1, there exist polynomials $h_1, h_3$ of total degree $n - 1$ such that

$$p_3 z + p_9 w = -h_1 x - h_3 y.$$ 

Since the exponents of the monomials in $p_9$ are not in $I$, it follows that the columns of $M P_1$ are linearly dependent: hence $|M P_1| = 0$.

If $p_3 \equiv p_9 \equiv 0$, but $p_2 \neq 0$ or $p_6 \neq 0$ or $p_8 \neq 0$, then as in the proof of Theorem 6.2,

$$p_2 y + p_6 z + p_8 w = h(s, t) x$$

for some polynomial $h(s, t)$. Without loss of generality, we can assume that the total degree of $x$ is $n$: therefore $h(s, t)$ is of total degree at most $n - 1$. Since the exponents of the monomials in $p_8$ are not in $I$, this again proves that the columns of $M P_1$ are linearly dependent. so $|M P_1| = 0$.

If $p_3 \equiv p_9 \equiv 0$ and $p_2 \equiv p_6 \equiv p_8 \equiv 0$, then by Equation (6.26),

$$p_1 x + p_4 y + p_5 z + p_7 w = 0.$$

where $p_1, p_4, p_5, p_7$ are not all zero polynomials. Since the exponents of the monomials in $p_7$ are not in $I$, the above equation again shows that the columns of $M P_1$ are linearly dependent. i.e. $|M P_1| = 0$.

Next we are going to show that if $M P$ has maximal rank, then there is an index set $I$ for which $M P_1$ is non-singular.
Lemma 6.1 For the triangular surface (6.22), if \( \text{Res}(x, y, z) \neq 0 \), then the columns
\[ s^i t^j x, s^i t^j y, s^i t^j z, \quad 0 \leq i + j \leq n - 1. \]
of \( MP \) are linearly independent.

Proof: Suppose the columns \( s^i t^j(x, y, z) \) are linearly dependent. Then there exist three polynomials \( p_1(s, t), p_2(s, t), p_3(s, t) \) of total degree \( n - 1 \) in \( s, t \) such that
\[ p_1 \cdot x + p_2 \cdot y + p_3 \cdot z = 0. \]
Since \( \text{Res}(x, y, z) \neq 0 \), by Proposition 6.1 of Section 2, the above syzygy implies that
\[ p_1 = h_1 \cdot z + h_2 \cdot y, \]
\[ p_2 = -h_2 \cdot x + h_3 \cdot z, \]
\[ p_3 = -h_1 \cdot x - h_3 \cdot z. \]
for three polynomials \( h_1, h_2, h_3 \) of degree \(-1\). Thus \( h_1 = h_2 = h_3 \equiv 0 \). Hence \( p_1 = p_2 = p_3 \equiv 0 \). Therefore, the columns
\[ s^i t^j x, s^i t^j y, s^i t^j z, \quad 0 \leq i + j \leq n - 1. \]
are linearly independent. \( \square \)

Lemma 6.2 If \( \text{Res}(x, y, z) \neq 0 \), and \( MP \) has maximal rank, then there exists an index set \( I \), \( |I| = n \), such that \( MP_I \) is non-singular.

Proof: By Lemma 6.1, we know that \( s^i t^j(x, y, z), 0 \leq i + j \leq n - 1 \), are linearly independent. Consider the following algorithm:

\[ I = \emptyset \ \text{(empty set)}; \]
\[ \Omega = \{s^i t^j(x, y, z) \mid 0 \leq i + j \leq n - 1\}; \]
\[ \Gamma = \{s^i t^j w \mid 0 \leq i + j \leq n - 1\}; \]
while \( \Gamma \neq \emptyset \)
Select a column $s' \ell w$ from $\Gamma$ and remove it from $\Gamma$:

If $s' \ell w$ is linearly independent from the columns in $\Omega$, then

add $s' \ell w$ to $\Omega$;

Otherwise, add $(i, j)$ to $I$.

If $MP$ has maximal rank $2n^2 + n$, then $2n^2 + n$ out of the $2n^2 + 2n$ columns in $MP$ are linearly independent. That is, the above algorithm terminates with $|\Omega| = 2n^2 + n$ and $|I| = n$. Now by definition $MP_I$ consists of the columns in $\Omega$, and by construction these columns are linearly independent. Therefore, $MP_I$ is non-singular. \qed

From Theorem 6.4, it follows that when $MP_I$ is non-singular and $Res(x, y, z) \neq 0$, $MQ_I$ is non-singular. The next theorem shows that these two conditions guarantee that the method of moving quadrics successfully implicitizes triangular surfaces.

**Theorem 6.5** Suppose that surface (6.22) has no base points, that the parametrization is generically one-to-one, and that there are exactly $n$ linearly independent moving planes of degree $n - 1$ that follow surface (6.22). Then the method of moving quadrics computes the implicit equation of surface (6.22).

**Proof:** Since, by assumption, our surface has no base points, we can assume, without loss of generality, that $Res(x, y, z) \neq 0$ (see the discussion in Section 6.4.2). By Lemma 6.2, the condition that surface (6.22) has exactly $n$ linearly independent moving planes of degree $n - 1$ is equivalent to $|MP_I| \neq 0$ for some index set $|I| = n$.

Hence by Theorem 6.4, $MQ_I$ is non-singular.

The linear system generated from the degree $n - 1$ moving quadrics [Equation (6.24)] is

$$MQ \cdot [A_{i,j} \ B_{i,j} \ \ldots \ I_{i,j} \ J_{i,j} \ \ldots]^T = 0.$$  \hspace{1cm} (6.27)

Since $|MQ_I| \neq 0$, we can solve for $A_{i,j}$, $B_{i,j}$, $C_{i,j}$, $D_{i,j}$, $E_{i,j}$, $F_{i,j}$, $i + j \leq n - 1$ and $G_{i,j}$, $H_{i,j}$, $I_{i,j}$, $(i, j) \notin I$, in terms of the undetermined coefficients: $J_{i,j}$, $i + j \leq n - 1$; $G_{i,j}$, $H_{i,j}$, $I_{i,j}$, $(i, j) \in I$. 

In particular, set \( G_{i,j} = H_{i,j} = I_{i,j} = J_{i,j} = 0 \) for \((i,j) \in I\). Then the linear system (6.27) can be written as

\[
MQ_I \cdot \begin{bmatrix}
  \vdots \\
  A_{i,j} \\
  \vdots \\
  I_{i,j} \\
  \vdots
\end{bmatrix} = -\text{Coeff} \left( s^t \lambda^2 \mid i + j \leq n - 1, (i,j) \not\in I \right) \cdot \begin{bmatrix}
  \vdots \\
  J_{i,j} \\
  \vdots
\end{bmatrix},
\]

(6.28)

where \( \text{Coeff} \left( s^t \lambda^2 \mid i + j \leq n - 1, (i,j) \not\in I \right) \) is the coefficient matrix of the polynomials \( s^t \lambda^2 \). \((i,j) \not\in I\). Setting \((\cdots, J_{i,j}, \cdots)\), \((i,j) \not\in I\), to the standard unit vectors \((0, \cdots, 0, 1, 0, \cdots, 0)\) in \( \mathbb{C}^{(n^2-n)/2} \), we obtain \((n^2-n)/2\) linearly independent solutions.

These solutions generate \((n^2-n)/2\) moving quadrics:

\[
Q_{i,j} = w^2 s^t \lambda^j + \text{terms without } w^2, \quad (i,j) \not\in I.
\]

On the other hand, since \( M P_I \) is non-singular, we can also find, in an analogous manner, \( n \) moving planes

\[
P_{i,j} = w s^t \lambda^j + \text{terms not involving } w s^k \lambda^l \text{ with } (k,l) \in I, \quad (i,j) \in I.
\]

Altogether we now have \((n^2-n)/2\) moving quadrics and \( n \) moving planes. Collecting the coefficients of these moving quadrics and moving planes, we obtain the matrix

\[
M = \begin{bmatrix}
  w^2 + \cdots \\
  & \ddots \\
  & & w^2 + \cdots \\
  & & & w + \cdots \\
  & & & & \ddots \\
  & & & & & w + \cdots
\end{bmatrix},
\]

where the first \((n^2-n)/2\) rows consist of the coefficients (of \( s^t \lambda^j \)) of the \((n^2-n)/2\) moving quadrics \( Q_{i,j}, (i,j) \not\in I \), and the last \( n \) rows consist of the coefficients (of \( s^t \lambda^j \))
of the $n$ moving planes $P_{i,j}$, $(i, j) \in I$. Note that here the columns are indexed by

$$
\begin{array}{c}
1, t, s, t^2, st, s^2, \ldots, s^t, \ldots, t^{n-1}, \ldots, s^{n-1}, \ldots,\\
\mathop{(i,j)}\in I
\end{array}
$$

Our goal is to prove that $|M| = 0$ is the implicit equation of surface (6.22). The determinant of this matrix has the term $u^{n^2}$, which shows that this determinant is not identically zero. Moreover, it is easy to see that this determinant vanishes whenever the point $(x, y, z, w)$ lies on the original surface (6.22) because each row represents a moving plane or a moving quadric that follows surface (6.22). Finally this determinant is of total degree $n^2$, which is the total degree of the implicit equation of surface (6.22) [23] since the parametrization is generically one-to-one. Hence this determinant is indeed the implicit equation of surface (6.22). ☐
Part III

Open Questions
Chapter 7

Open Questions

In this thesis, we have examined classical resultants, sparse resultants, and implicitization procedures. Below we present some interesting unsolved problems concerning resultants and implicitization that arose out of the investigations in this dissertation.

7.1 Classical Resultants

Classical resultants have a long history dating back over two hundred years. So it is rather remarkable that there are still many open questions regarding these resultants. Below we list three open problems that are extensions of the work presented here.

- By inspecting the Sylvester matrix, we see that the resultant of three bivariate polynomials of bidegree \((m, n)\) is homogeneous of degree \(6mn\) in the coefficients of the three polynomials. We also know that there are resultant matrices of order \(6mn, 3mn, 2mn\) with homogeneous entries of degree 1, 2, and 3 respectively [Ch 3]. Can the resultant also be represented by an \(mn \times mn\) determinant with homogeneous entries of degree 6 in the coefficients of the three original polynomials? What are the polynomials that might generate this very compact \(mn \times mn\) determinant? Are there relationships between these polynomials and the polynomials that generate the three known Dixon resultant matrices?

Such an order \(mn\) resultant could be quite useful for rational surface implicitization. For rational curves of degree \(2m\), the method of moving conics computes the implicit equation by taking determinants of order \(m\). This implicit equation can also be generated by taking the Bézout resultant — again of order \(m\) —
of the $\mu$-basis of the curves. For rational tensor product surfaces of bi-degree $(m, n)$, the method of moving quadrics computes the implicit equation by taking determinants of order $mn$. But we still do not know how to generate order $mn$ resultants from the original three auxiliary polynomials (see too Chapter 4, Section 4.7). If such an $mn \times mn$ resultant were known, we might have an easier way to generate the necessary moving quadrics.

- We have seen that we can impose block structures on the univariate and bivariate resultant matrices, and then use these block structures both to develop efficient algorithms and to build hybrid resultant formulations. Can we generalize this technique to multipolynomial resultants? That is, do multipolynomial resultant matrices also have useful block structures on their entries?

The difficulty here comes from the representations of the multipolynomial resultants. The Dixon matrices can be generalized to multi-degree polynomials [49], but these Dixon matrices are not necessarily square or non-singular. Similarly, Sylvester matrices can be generalized to multipolynomials of fixed total degrees. But again these Sylvester matrices are not necessarily square or non-singular [21] [31]. Typically multipolynomial resultants are constructed as Macaulay quotients, or expressed as the greatest common divisors (GCDs) of a collection of determinants. Finding useful block structures in these quotients or GCDs does not appear to be a simple task.

- We have successfully interpreted the columns of the products of the blocks of the Sylvester resultant and the mixed Cayley-Sylvester resultant as the coefficients of certain polynomials. We then used this interpretation to prove certain convolution identities (see Chapter 3, Section 3.6). Associating the entries with specific polynomials makes the block structures of the resultant matrices much clearer. Can the blocks of other resultant matrices be related with the help of certain similar polynomials?
7.2 Sparse Resultants

Research on sparse resultants is still relatively new and very active. We have the following questions arising directly from our research in Chapter 4 on Sylvester $A$-resultants and Dixon $A$-resultants.

- Can we find explicit generic examples of polynomials for which the Sylvester $A$-resultant constructed in Chapter 4 is not identically zero? Such examples might help to simplify the proof of Theorem 4.1.

- For three tensor product polynomials of bi-degree $(m,n)$, the Sylvester and Dixon resultant matrices are intimately related [14]. In fact, we showed in Chapter 3 that the ideal generated by the polynomials represented by the columns of the Sylvester resultant matrix contains the ideal generated by the polynomials represented by the columns of the Dixon resultant matrix. However, for three polynomials with rectangular corners cut from their monomial support, the corresponding ideal for the Sylvester $A$-resultant matrix need not contain the corresponding ideal for the Dixon $A$-resultant matrix.

Example 7.1

Let $f,g,h$ be the following three polynomials

$$f(s,t) = 2s + t, \quad g(s,t) = st + st^2, \quad h(s,t) = s^2t + 2t.$$  

The ideals in $\mathbb{C}[s,t]$ generated by the polynomials represented by the columns of $Syl_A(f,g,h)$ and $Dix_A(f,g,h)$ have Groebner bases $\{st, s^2\}$ and $\{t\}$ respectively. Clearly, the first ideal does not contain the second ideal.

Whether the two polynomial ideals generated by the polynomials represented by the columns of $Syl_A(f,g,h)$ and $Dix_A(f,g,h)$ are related, and what if anything this relation might be, needs further study.
• We can construct the Sylvester $\mathcal{A}$-resultant when a rectangle is cut off at each corner of a rectangular monomial support. Can we succeed when other shapes -- say, triangles -- are cut off at the corners? That is, does a simple Sylvester $\mathcal{A}$-resultant matrix still exist that gives the true resultant when other shapes are removed from the corners of the monomial support?

• Along similar lines, can we succeed in constructing a Sylvester $\mathcal{A}$-resultant when staircase rectangles are cut off at the corners (Figure 7.1)? That is, instead of cutting one rectangle at each corner from the monomial support, can we cut off multiple rectangles that form a staircase at each corner and still construct a simple Sylvester $\mathcal{A}$-resultant matrix?

![Figure 7.1: Staircase corner cut off from $\mathcal{A}_{m,n}$.](image)

• In Chapter 3 we showed how to construct hybrid resultants for three tensor product polynomials from the Sylvester and Dixon resultant matrices. Do hybrid resultants constructed from the Sylvester $\mathcal{A}$-resultant and Dixon $\mathcal{A}$-resultant matrices exist?

• Can the construction of the Sylvester $\mathcal{A}$-resultant and the Dixon $\mathcal{A}$-resultant for three bivariate polynomials where a rectangle is cut off at each corner of the monomial support be generalized to $n$ polynomials in $n - 1$ variables?
7.3 Moving Quadrics

Some important questions regarding the method of moving quadrics also still remain open. Below we describe some of the most important unresolved problems.

- Base points.

Throughout Sections 6.4 and 6.5 we have assumed that the rational surfaces have no base points in order to prove the validity of implicitization by the method of moving quadrics. But the existence of base points can be represented by polynomial conditions. Therefore, in fact, we have showed that the method of moving quadrics works for surfaces in a Zariski open set — that is, implicitization by the method of moving quadrics almost always succeeds. Nevertheless, we would like to establish that the method works even in the presence of base points because in practical industrial design base points show up quite frequently. What happens if the surfaces do have base points? Each base point will, in general, lower the degree of the implicit equation by one. Experiments [52] show that in the presence of base points some moving quadrics can be replaced by moving planes to correctly compute the implicit equation. However, it is still not clear how to make these replacements systematically when surfaces have base points. Moreover, if there are too many base points, we will need to modify the method of moving quadrics more drastically, for when a bi-degree \((m, n)\) surface has more than \(mn\) base points the degree of the implicit equation is less than \(mn\). Thus we cannot use bi-degree \((m - 1, n - 1)\) moving quadrics or moving planes to implicitize the surface. What degrees should we use and how can we make the adjustment automatically in the presence of base points?

- Factoring the determinant of the moving quadric coefficient matrix.

For rational curves, we showed in Chapter 5 that the determinant of the moving conic coefficient matrix \((MC_w)\) can be factored in terms of the determinant of the moving line coefficient matrix \((ML)\) and the resultant of two coordinate
polynomials. Indeed, we showed in Chapter 5 that

$$|MC_w| = |ML|^2 \cdot Res(x, y).$$

Can we similarly factor the determinant $|MQ_w|$ (or $|MQ_t|$) of the moving quadric coefficient matrix in terms of the determinant $|MP|$ (or $|MP_t|$) of the moving plane coefficient matrix and the resultant of three coordinate polynomials? In particular, is it true that

$$|MQ_w| = |MP|^3 \cdot Res(x, y, z) \quad \text{(for tensor product surfaces).} \quad (7.1)$$

and

$$|MQ_t| = |MP_t|^3 \cdot Res(x, y, z) \quad \text{(for triangular surfaces)?} \quad (7.2)$$

We believe that these factorizations are correct for the following two reasons:

First, both the left hand side and the right hand side of Equations (7.1) and (7.2) have the same degrees in the coefficients of $x(s, t), y(s, t), z(s, t), w(s, t)$. Second, in the univariate setting, each linear relation among the columns of $ML$ generates two linear relations among the columns of $MC_w$ (multiplying each relation by $x(t), y(t)$); thus $|ML|$ is a double factor of $|MC_w|$. For rational surfaces, each linear relation between the columns of $MP$ (or $MP_t$) generates three linear relations between the columns of $MQ_w$ (or $MQ_t$) (multiplying each relation by $x(s, t), y(s, t), z(s, t)$). Therefore, we conjecture that $|MP|$ (or $|MP_t|$) is a triple factor of $|MQ_w|$ (or $MQ_t$). One way to prove this result would be to show that $|MP|$ (or $|MP_t|$) is an irreducible polynomial in the coefficients of $x(s, t), y(s, t), z(s, t), w(s, t)$. We have yet to succeed in establishing such a result. An alternative approach to establishing this factorization has recently been proposed by D’Andrea [24].

- Simplifying the proofs.

The proofs of Theorems 6.3 and 6.5 of the validity of implicitization by the method of moving quadrics rely on Theorems 6.2 and 6.4, which, in turn, depend upon Propositions 6.1 and 6.2 in Section 6.3, and the proofs of these two
propositions require advanced knowledge in algebraic geometry and commutative algebra [22]. Is there an elementary proof for Theorems 6.3 and 6.5? Will the factorization discussed in the previous paragraph lead to such a straightforward proof?

- Symbolic computation.

When the coefficients of a rational surface are numerical, the method of moving quadrics produces the implicit equation of the rational surface with a much smaller matrix than the method of resultants: hence for implicitization moving quadrics are more efficient than resultants. However, when all the coefficients of a rational surface are symbolic, the bottleneck in the moving quadrics method is the complexity of solving the linear system to obtain the moving quadrics. Closed form solutions for the moving quadrics — without having to solve a large system of linear equations — would make the method far more efficient. But how to generate an appropriate collection of moving quadrics in closed form systematically is still an open question.

- Toric varieties.

We showed in Chapter 6 how to implicitize rational tensor product surfaces and rational triangular surfaces using moving quadrics. Can we use moving quadrics to implicitize toric varieties whose defining parametric polynomials do not necessarily have rectangular or triangular monomial supports? For example, can we systematically adjust the method of moving quadrics for the rational surfaces where the monomial supports are rectangles with rectangular corners cut off?

- More polynomials in more variables.

The method of moving quadrics is essentially an extension to surfaces of the method of moving conics. Can we generalize the method of moving quadrics
to higher dimensional varieties where more polynomials in more variables are involved? The proof of the validity of the method of moving conics for curves requires only linear algebra, generic properties of resultants, and a few simple facts about factoring univariate polynomials. However, the proof of the validity of the method of moving quadrics for surfaces is much more difficult as sophisticated tools from algebraic geometry and commutative algebra are required. We can only begin to imagine the difficulty of the proof of the validity of the method when generalized to higher dimensional varieties.
Bibliography


