8-26-2016 12:00 AM

A Sequence of Symmetric Bézout Matrix Polynomials

Leili Rafiee Sevyeri, *The University of Western Ontario*

Supervisor: Rob Corless, *The University of Western Ontario*

A thesis submitted in partial fulfillment of the requirements for the Master of Science degree in Applied Mathematics

Abstract

There are problems concerning the set of root of a sequence of polynomials. A simple question is to ask if the set of roots lies entirely in real numbers. Many approaches to answering this question are known. The main object of this dissertation is to develop new tools for tackling the above problem. In order to be able to apply the ideas we define a specific numerical sequence, and then we consider the sequence of their minimal polynomials over the rational numbers.

The first step is to find a recursive way of defining the sequence of polynomials by using the so-called Bézout matrices, which are a specific family of matrix polynomials. Having the construction of minimal polynomials as the determinant of some Bézout matrix, we interpret the roots of each polynomial as eigenvalues of the corresponding Bézout matrix. Then by using a symmetric linearization of such matrix polynomial we can talk about the real roots.

Keywords: Matrix Polynomials, Linearization, Bézout Matrix.

Acknowledgement

First and foremost my deepest gratitude is to my advisor, Prof. **Rob Corless**. I have been amazingly fortunate to have an advisor who has supported me throughout my thesis with his motivation, enthusiasm, and immense knowledge. His patience and support helped me overcome many crisis situations and finish this dissertation. The door to **Rob**'s office was always open whenever I ran into a trouble spot or had a question about my research or writing. One simply could not wish for a better or friendlier supervisor. I hope that one day I would become as good an advisor to my students as Rob has been to me.

Besides my advisor, I would like to thank Prof. **David Jeffery** for his constant positive attitude and for all the encouragements he gave throughout the process. I am proud, honoured and extremely grateful to have been able to work with such amazingly talented researchers, but most of all, who are also extremely kind human beings. **Rob**, **David**.

For having the opportunity to study in the Western University, I want to huge thank Prof. **Lindi Wahl** to introduce me to Rob and helped me make a right decision early at the beginning of my Master program.

This thesis would have never been possible without my loving husband, **Armin**. Words would never say how grateful I am to you. I am truly appreciative of all you have done for me over the years. I could not have reached my goals without your help and support. You really gave me the reason to continue. You have made me feel special even when times were tough for me. You have been my pillar of strength through all my ups and downs. I feel blessed to be a part of your life, and am fortunate to be married to the most kind-hearted and generous person I have ever known. I love you more and more everyday. You can take all the credit for much of what I have achieved and what I will achieve in the future.

Finally, I acknowledge the people who mean a lot to me, my parents, **Maman** and **Baba**, for showing faith in me and giving me liberty to choose what I desired. I salute you all for the selfless love, care, pain and sacrifice you did to shape my life. Although you hardly understood what I researched on, you were willing to support any decision I made. I would never be able to pay back the love and affection showered upon by my parents. Also I express my thanks to my sisters **Laya** and **Sara**, for their support, valuable prayers, their selfless love, care and dedicated efforts which contributed a lot for completion of my thesis. Even though we are thousand of miles away, you were always there whenever I needed you.

My heart felt regard goes to my parents in law, bothers in law, **Arash** and **Aryo** for their love and moral support.

Thank you, **Allah**, for always being there for me. This thesis is only a beginning of my journey.

Contents

List of Figures

List of Tables

List of Appendices

Chapter 1

Introduction

In this chapter we introduce a sequence of numbers that is given recursively. We briefly describe the origin of this sequence. In order to do so we need to look at an ordinary differential equation which is derived from Torricelli's law for the leaky bucket. Torricelli's law is a model in fluid dynamics which gives the relation between the speed of fluid (flowing out) to the height of the fluid (above the drainage hole) in the bucket.

Assume that we have a cylindrical bucket which contains some fluid. If there is a hole in the bottom of the bucket, eventually the bucket will run out of water. In [3] the authors commented on using the following very well-known model:

$$\frac{dy}{dt} = -\sqrt{y} \tag{1.1}$$

where $y = \frac{h}{H}$, i.e. the ratio of height of fluid to height of bucket. The authors applied the Euler forward method to solve the model equation. Before talking about the outcome of Euler forward method, we just recall the method quickly.

Euler forward method is a method for solving ordinary differential equation. For an ordinary differential equation given by

$$y'(t) = f(t, y(t))$$

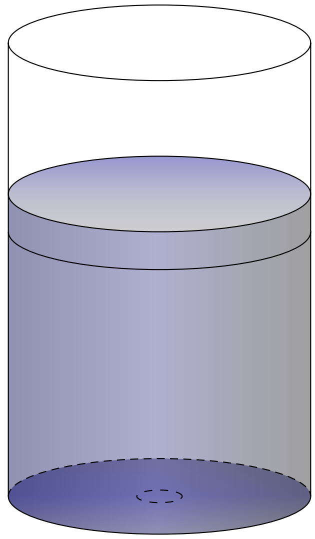

Figure 1.1: Cylindrical bucket with a drainage hole

and the initial condition $y(t_0) = y_0$, the method gives an approximation of the solution at a given point. Choosing a value $\triangle t$ for the size of each step, one can write $t_n = t_0 + n \triangle t$, then one step of the Euler method from t_n to $t_n + 1$ is given by

$$y_{n+1} = y_n + \triangle t f(t_n, y_n)$$

where y_n approximates the solution of the equation at t_n.

Now let us get back to the model (1.1). Using the Euler forward formula and letting the bucket be full in the beginning (or equivalently $y_0 = 1$), we get

$$y_{n+1} = y_n - \triangle t \sqrt{y_n}.$$

For $\triangle t = 1$, we get $y_1 = 0$ which means the bucket empties in just one step. Assuming $\triangle t = \frac{1}{2}$ one can see $y_1 = \frac{1}{2}$ and calculating y_3 gives a negative number which leads to a complex value for y_4. It turns out that for almost all choices of $\triangle t$ except one family, eventually $y_n < \triangle t^2$ for some integer n and therefore $y_{n+1} < 0$ which gives a complex y_{n+2}.

Now letting $\triangle t = \frac{2}{1+\sqrt{5}}$ implies

$$y_1 = y_0 - \triangle t \sqrt{y_0} = 1 - \frac{2}{1+\sqrt{5}} \cdot 1 = \frac{\sqrt{5}-1}{1+\sqrt{5}}$$

$$y_2 = y_1 - \triangle t\sqrt{y_1} = \frac{\sqrt{5}-1}{1+\sqrt{5}} - \frac{2}{1+\sqrt{5}}\sqrt{\frac{\sqrt{5}-1}{1+\sqrt{5}}} = 0$$

which means the bucket empties in two steps. Considering $\triangle t = \frac{2}{1+\sqrt{7+2\sqrt{5}}}$ one can see

$$y_1 = y_0 - \triangle t\sqrt{y_0} = 1 - \frac{2}{1+\sqrt{7+2\sqrt{5}}}.1 = \frac{\sqrt{7+2\sqrt{5}}-1}{\sqrt{7+2\sqrt{5}}+1}$$

and then

$$y_2 = y_1 - \triangle t\sqrt{y_1} = \frac{\sqrt{7+2\sqrt{5}}-1}{\sqrt{7+2\sqrt{5}}+1} - \frac{2}{1+\sqrt{7+2\sqrt{5}}}\sqrt{\frac{\sqrt{7+2\sqrt{5}}-1}{\sqrt{7+2\sqrt{5}}+1}} = 4\left(\sqrt{7+2\sqrt{5}}+1\right)^{-2}$$

finally

$$y_3 = y_2 - \triangle t\sqrt{y_2} = 4\left(\sqrt{7+2\sqrt{5}}+1\right)^{-2} - \frac{2}{1+\sqrt{7+2\sqrt{5}}}\sqrt{4\left(\sqrt{7+2\sqrt{5}}+1\right)^{-2}} = 0$$

Now we define a sequence $\{u_i\}$, as

$$u_1 = 1$$

$$u_{i+1} = 2 + u_i + 2\sqrt{u_i} \quad for \;\; i \geq 2$$

In [3] the authors showed that in fact the above experimental result holds for any n.

Theorem 1.0.1 *[3] Assume $\triangle t = \frac{2}{1+\sqrt{u_n}}$, then the bucket will be empty in n steps.*

Proof We have

$$y_{n+1} = y_n - \triangle t\sqrt{y_n}. \tag{1.2}$$

Now assume $y_n = (\triangle t)^2 f_n$ and $1 = y_0 = (\triangle t)^2 f_0$.

Hence we get $(\triangle t) = \frac{1}{\sqrt{f_0}}$. Now we can reformulate equation (1.2) as

$$(\triangle t)^2 f_{n+1} = (\triangle t)^2 f_n - (\triangle t)\sqrt{(\triangle t)^2 f_n}$$

$$\Longrightarrow (\triangle t)^2 f_{n+1} = (\triangle t)^2 f_n - (\triangle t)^2\sqrt{f_n}$$

Since $(\triangle t) \neq 0$ then

$$f_{n+1} = f_n - \sqrt{f_n} \tag{1.3}$$

solving equation 1.3 for f_n one gets

$$f_n = \left(\frac{1+\sqrt{1+4f_{n+1}}}{2}\right)^2$$

On the other hand, we have $y_n = 0$ if and only if $f_n = 0$. The goal here is to find $\triangle t$ such that $y_{n+1} = 0$ and $y_n \neq 0$. Hence by putting $f_{n+1} = 0$ in 1.3 and calculating f_0 recursively we can get $\triangle t$. For example if we want the bucket to empty in one step then $f_2 = 0$ so

$$f_1 = \left(\frac{1+\sqrt{1}}{2}\right)^2$$

$$f_0 = \left(\frac{1+\sqrt{1+4(\frac{1+\sqrt{1}}{2})^2}}{2}\right)^2 = \left(\frac{1+\sqrt{5}}{2}\right)^2.$$

This implies $\triangle t = \frac{2}{1+\sqrt{5}}$.

Now let u_k be the number appearing under the square root in corresponding $\triangle t$, when bucket is going to be empty in k steps. In order to be able to find the similar number in the case bucket gets empty in $k+1$ steps, we write

$$f_0 = \left(\frac{1+\sqrt{1+4\left(\frac{1+\sqrt{u_k}}{2}\right)^2}}{2}\right)^2 = \left(\frac{1+\sqrt{2+u_k+2\sqrt{u_k}}}{2}\right)^2$$

This completes the proof.

Table 1.1 shows $\triangle t = \frac{2}{1+\sqrt{u_n}}$ for $3 \leq n \leq 8$ which is calculated using the following Maple code (notice that all computation is performed exactly, floating point is used only for display):

```
Rem_Amount:= proc(n)
  local(i, k, t, x, y);
  u || 1 := 1; #The symbol || means concatenation.
    for k from 2 to n do
      u || k := u || (k-1)+2+2*sqrt(u || (k-1));
    od;
  t := 2/(1+sqrt(u || n));
  y := 1;
```

```
    for i from 1 to n do
      x := simplify(y-t*sqrt(y));
      y := x;
    od;
  print(evalf(t));
  print(evalf(u || n));
  y;
end proc:
```

n	u_n	$\triangle t$	y_{n+1}
3	11.47213595	0.4558867802	0
4	20.24624429	0.3636639572	0
5	31.24540965	0.3035012194	0
6	44.42492836	0.2609193850	0
7	59.75533396	0.2290909432	0
8	77.21564883	0.2043476280	0

Table 1.1: Values of y_{n+1} for different values of $\triangle t$

Above results motivated us to investigate more about the sequence of $\{u_i\}$. As the first step, we see some experiments that were done (using Maple and Sage). From the recursive formula of the sequence, one can see the fact that each u_n is algebraic over rational numbers (see definition 2.2.3). Our experiments give us the idea to find the sequence of minimal polynomials of $\{u_i\}$ in a constructive way (see definition 2.2.4). We also noticed that all roots of minimal polynomials are real.

The main *purpose* of the study is not so much the study of the properties of the $\{u_i\}$, but rather developing new tools and applying them: this sequence provides a convenient challenge to develop recursively-constructed eigenvalue problems. We will learn facts about the sequence, but more importantly, we will develop a new method for proving that all the roots of a family of polynomials are real, namely the construction of symmetric Bézout matrix polynomials and using a symmetric linearization of such matrix polynomials expressed in the Lagrange basis.

One purpose of this thesis is to test the utility of the block symmetric linearization of matrix polynomials expressed in a Lagrange basis [1]. This block symmetrization was only published in 2009, and is as yet relatively unused; there is no application yet in the literature of which we are aware.

One simple possibility is that a block symmetric linearization of a symmetric matrix polynomial might allow easy deduction of *definiteness*, *i.e.* that all of the non-linear eigenvalues of the matrix polynomial are positive (or non-negative). Definiteness of a matrix polynomial is an important property for many purposes: see e.g. [16].

We show in this thesis that, yes, the new block symmetric linearization can be helpful to decide definiteness. Specifically, we look at a new recursively defined family of polynomials that arises from a non-linear recurrence relation that generates a sequence containing nested square roots, and show that they are all positive definite by using the aforementioned linearization.

Sequences of nested square roots have been of interest since Viéte (also known by the Latinized form Vieta) gave the formula (in 1593)

$$\frac{2}{\pi} = \frac{\sqrt{2}}{2} \cdot \frac{\sqrt{2+\sqrt{2}}}{2} \cdot \frac{\sqrt{2+\sqrt{2+\sqrt{2}}}}{2} \cdots .$$

Perhaps surprisingly, such sequences are still generating new research results. See for instance [13], [17] and [14].

This thesis proves that a certain sequence of polynomials that define the sequence of nested square roots described in section 4.1 and section 4.4 is positive definite, thus demonstrating the utility of the block symmetric linearization.

This thesis is structured as follows: chapter 2 is devoted to the basic notions of field theory which provides the tools for our calculations. In chapter 3, we provide information about matrix polynomials and the Bézout matrix. We also define the resultant and its applications to solving bivariate system of equations. In the last chapter the main results are presented, namely finding a boundary of $\{u_i\}$, minimal polynomials of $\{u_i\}$, roots of minimal polynomials and finally symmetric linearization of the Bézout matrix.

Chapter 2

Preliminaries

In this chapter we provide the preliminary definitions that will be needed in future chapters. Here we just take a basic knowledge of ring theory and linear algebra for granted and invite the reader to see [7] and [6] if needed.

The material in this chapter is organized such that the first two sections give general information and definitions for Galois theory and the last two sections provide information for roots of a family of polynomials and some algorithms to find them.

2.1 Basic Definitions

We recall that a field is commutative ring with identity with the property that every non zero element has a multiplicative inverse.

Throughout the thesis, we denote the fields of rational numbers, real numbers and complex numbers (resp.) by $\mathbb{Q}$, $\mathbb{R}$ and $\mathbb{C}$.

Theorem 2.1.1 *If p is a prime, then the ring $\mathbb{Z}_p$ is a field which is finite.*

Example 2.1.1 *Assume K is a field and x is an indeterminate. The rational function field of the polynomial ring $K[x]$, is denoted by $K(x)$ and consists of all $f(x)/g(x)$ such that $g(x) \neq 0$ and $f(x), g(x) \in K[x]$.*

In general for a given polynomial $p(x) \in F(x)$, where F is a field, it is not always the case that all roots of $p(x)$ lie in F. In order to study the roots of $p(x)$ we need to find a field such that it contains the roots. To see this better, we note that the roots of $x^2 - 2 \in \mathbb{Q}(x)$ are $\pm\sqrt{2}$ which are not rational numbers. However, there exists a field such that it contains both roots and $\mathbb{Q}$. The smallest field with this property is denoted by $\mathbb{Q}(\sqrt{2})$. One can extend this idea to the following definition.

Definition 2.1.1 *Let F be a field and K be another field such that $F \subseteq K$. Let A be a subset of K. Then the smallest subfield (with respect to inclusion) of K containing A and F, which is the intersection of all subfields containing A and F, is denoted by $F(A)$ and is called the field generated by F and A. If $A = \{a_1, \cdots, a_n\}$ be a finite set we denote $F(A)$ by $F(a_1, \cdots, a_n)$ and call it finitely generated. Moreover if $A = \{a\}$ we write $F(a)$ instead of $F(A)$ and we call it a simple extension of F.*

The above definition does not give $F(a)$ concretely. There is a concrete description of $F(A)$ in general. We describe the idea for $A = \{a\}$ and we mention that a similar argument can be applied to any finite $A \subseteq K$ case. We claim that

$$F(a) = \{\frac{f(a)}{g(a)} : f(x), g(x) \in F[x], g(a) \neq 0\}.$$

We note that $F(a)$ is indeed the quotient field of $F[a]$. So it is enough to show that

$$F[a] = \{f(a) : f(x) \in F[x]\}.$$

In order to do so we define $\varphi_a : F[x] \longrightarrow K$ such that $\varphi_a(f(x)) = f(a)$. One can verify that φ is indeed a ring homomorphism which is called the evaluation homomorphism at a. It is clear that $Im(\varphi_a) = \{f(a) : f(x) \in F[x]\}$ and hence it is a subring of K. Obviously if $R \subseteq K$ be a subring which contains F and a, then it contains $f(a)$. Therefore we can see that any subring of K which has F and a inside, should contain $\{f(a) : f(x) \in F[x]\}$. So

$$F[a] = \{f(a) : f(x) \in F[x]\}.$$

Since $F(a)$ is the quotient field of $f[a]$, we have proved our claim.

Theorem 2.1.2 *Let F and K be fields, $K \subseteq F$ and $\{a_1, \cdots, a_n\} \subseteq F$, then*

$$K[a_1, \cdots, a_n] = \{f(a_1, \cdots, a_n) : f \in F[x_1, \cdots, x_n]\}$$

and

$$K(a_1, \cdots, a_n) = \{\frac{f(a_1, \cdots, a_n)}{g(a_1, \cdots, a_n)} : f, g \in F[x_1, \cdots, x_n], g \neq 0\}$$

2.2 Minimal Polynomials and Algebraic Extensions

We begin this section by providing the definition of field extensions and then we look at a specific family of field extensions which is called algebraic extensions.

Definition 2.2.1 *If $K \subseteq F$ are fields, we say F is a field extension of K. We denote the extension by F/K and we call K the base field.*

It is well known that any field F is a vector space over any subfield of itself. So the following definition will work.

Definition 2.2.2 *If F/K is a field extension, then we call* $\dim_K(F)$ *(as vector space), the degree of the extension and we denote it by $[F : K]$. F is called a finite extension if the degree of F over K is finite, otherwise it is called an infinite extension.*

In this section we are interested in those extensions F/K for which any $a \in F$ satisfies a polynomial in $K[x]$, i.e. there exists a polynomial $f(x) \in K[x]$ such that $f(a) = 0$. In order to understand these extensions better, we need some more formal definitions.

Definition 2.2.3 *Assume K/F is a field extension and $a \in F$. We say a is algebraic over K if there exists a polynomial in $K[x]$ that satisfies at a. If a is not algebraic over K, it is called transcendental over K. Moreover, if every element of F is algebraic over K, we say F is an algebraic extension of K and F/K is said to be an algebraic extension.*

If $a \in F$ is algebraic over K, it satisfies a polynomial over K, but it is not necessarily the smallest (with respect to degree) polynomial with such a property.

Definition 2.2.4 *Assume a is an algebraic element over K. The minimal polynomial of a over K is the monic polynomial with smallest degree in $K[x]$ that satisfies at a.*

Example 2.2.1 *It is clear that $\sqrt{2}$ satisfies $x^4 - 4 \in \mathbb{Q}[x]$, so it is algebraic over $\mathbb{Q}$. However, $x^4 - 4$ is not the minimal polynomial of $\sqrt{2}$ since it satisfies $x^2 - 2$ which is of least degree (it is irreducible over $\mathbb{Q}$). As another example, $i = \sqrt{-1}$ is algebraic over $\mathbb{Q}$ with the minimal polynomial, $x^2 + 1$. On the other hand i is algebraic over $\mathbb{C}$ with minimal polynomial, $x - i$.*

Example 2.2.2 *It is not trivial that e is transcendental over $\mathbb{Q}$. The first proof was given by Hermite in 1873. In 1882, Lindemann showed that π is transcendental over $\mathbb{Q}$. One can find proofs for both cases in [12].*

Theorem 2.2.1 *[12] Let K/F be a field extension and $\alpha \in K$ be algebraic over F. Then*

1. *The minimal polynomial of α over F is irreducible over F.*
2. *If $f(x), g(x) \in F[x]$, such that $f(x)$ be the minimal polynomial of α over F, then $g(\alpha) = 0$ if and only if $f(x)$ divides $g(x)$.*
3. *If the degree of minimal polynomial of α over F is n, then $\{1, \alpha, \alpha^2, \cdots, \alpha^{n-1}\}$ is a basis for $F(\alpha)$ over F . Hence $[F(\alpha) : F] = \deg(p)$ where p is the minimal polynomial of α over F.*

Example 2.2.3 *Let's see if $\sqrt[3]{5}$ is algebraic over $\mathbb{Q}[x]$ or not. Assume $\alpha = \sqrt[3]{5}$, now we have*

$$\alpha^3 = 5$$

Hence

$$\alpha^3 - 5 = 0$$

This means α satisfies $p(x) = x^3 - 5 \in \mathbb{Q}$. On the other hand, $x^3 - 5$ is irreducible, since otherwise it must have a root in $\mathbb{Q}$. If $\frac{a}{b} \in \mathbb{Q}$ and $\gcd(a, b) = 1$ be a root of $p(x)$, then

$$(\frac{a}{b})^3 - 5 = 0 \Longrightarrow \frac{a^3}{b^3} - 5 = 0$$

then

$$a^3 = 5b^3 \Longrightarrow 5 \mid a^3 \Longrightarrow 5 \mid a$$

where $a \mid b$ means a divides b. Let $a = 5t$, then

$$(5t)^3 = 5b^3 \Longrightarrow 125t^3 = 5b^3 \Longrightarrow 5 \mid b$$

which is absurd. Irreducibility of $p(x)$ shows that, it is the minimal polynomial of $\sqrt[3]{5}$ over $\mathbb{Q}$. The above argument can be applied in general to any prime q to see $x^n - q$ is irreducible over rationals (alternatively one can apply the Eisenstein criterion [6] to see this). Hence

$$[\mathbb{Q}(\sqrt[n]{q}) : \mathbb{Q}] = n.$$

Theorem 2.2.2 *Assume F/K is a finite extension. Then F is finitely generated and algebraic over K.*

Proof Let $\{\alpha_1, \cdots, \alpha_n\}$ be a basis for F over K. Clearly any $a \in F$ can be written as

$$\sum_{i=1}^{n} c_i\alpha_i \quad for\ some \quad c_i \in K$$

hence

$$F = K(\alpha_1, \cdots, \alpha_n).$$

If $a \in F$, then $\{1, a, \cdots, a^n\}$ is a dependent set over K (recall that $\dim_K(F) = n$). This implies, there exists $u_1, \cdots, u_n$ such that

$$\sum_{i=1}^{n} u_i a^i = 0$$

If

$$f(x) = \sum_{i=1}^{n} u_i x^i = 0 \in K[x]$$

then $f(a) = 0$. Therefore a is algebraic over K.

The following theorem enables us to state the degree of a field extension in terms of extensions of middle fields. Moreover, having this machinery, one can prove the converse of the previous theorem.

Theorem 2.2.3 *Assume $F \subseteq L \subseteq K$ are fields and all extensions are finite. Then*

$$[K : F] = [K : L] \times [L : F].$$

Proof Assume $\{\alpha_i\}_{i=1}^n$ and $\{\beta_j\}_{j=1}^m$ are respectively bases for K/L and L/F. We claim that

$$B = \{\alpha_i\beta_j : 1 \leq i \leq n,\ 1 \leq j \leq m\}$$

is a basis for K/F. As the first step we show B spans K over F. For any element a in K, we can write a as a combination of α_i, i.e.

$$a = \sum_{i=1}^{n} a_i\alpha_i \quad for\ some \quad a_i \in L.$$

On the other hand, any element of L can be written as a linear combination of the β_is. In particular each a_i can be written in such a way. Hence

$$a = \sum_{i=1,j=1}^{n} c_{ij}\alpha_i\beta_j \quad for\ some \quad c_{ij} \in F.$$

Now assume

$$\sum_{i=1,j=1}^{n} d_{ij}\alpha_i\beta_j = 0$$

for some d_{ij}s in F. The linear independency of β_js implies that

$$\sum_{i=1,j=1}^{n} d_{ij}\alpha_i = 0$$

for any j. Moreover the linear independency of α_is implies $d_{ij} = 0$ for all i, j. So B is a basis of K over F, therefore

$$[K : F] = |B| = n \times m = [K : L] \times [L : F].$$

At this point we want to present the converse of theorem 2.2.2. One can find a proof of it which is based on induction, in [12].

Theorem 2.2.4 *Let K/F be a field extension. If each α_i be algebraic over F, then $F(\alpha_1, \cdots, \alpha_n)$ is a finite field extension of F. Moreover we have*

$$[F(\alpha_1, \cdots, \alpha_n) : F] \leq \prod_{i=1}^{n} [F(\alpha_i) : F].$$

One good property of algebraic extensions is the transitivity in the sense of presence of a middle extension.

Theorem 2.2.5 *Assume $F \subseteq L \subseteq K$ are fields. If K/L is algebraic and L/F is algebraic, then K/F is an algebraic extension.*

Proof We have to show any α in K satisfies a polynomial $p(x)$ in $F[x]$. Since α is algebraic over L, $f(\alpha) = 0$ for some

$$f(x) = \sum_{i=0}^{n} a_i x^i$$

in $L[x]$. Now consider the field $L' = F(a_1, \cdots, a_n)$ over F which is finite by theorem 2.2.4. We note that $f(x) \in L'[x]$, so α is algebraic over L'. Thus

$$[L'(\alpha) : F] = [L'(\alpha) : L'] \cdot [L' : F] < \infty$$

On the other hand we have $F(\alpha) \subseteq L'(\alpha)$ which implies finiteness of $[F(\alpha) : F]$ and this means α is algebraic over F. This argument holds for any element of K. Hence K is algebraic over F.

Definition 2.2.5 *A field F is called algebraically closed if any non-constant polynomial $f(x)$ in $F[x]$ has a root in F.*

Remark 2.2.1 *The above definition is saying that a field F is algebraically closed if any polynomial over F factors completely over F. There are equivalent definitions for algebraic closed fields. One can also say that F is algebraically closed if for any algebraic extension K of F we have $F = K$.*

Definition 2.2.6 *A field extension of F is called an algebraic closure of F if it is an algebraic extension of F and algebraically closed. We denote the algebraic closure of F by $\bar{F}$.*

Example 2.2.4 *It is well-known that $\mathbb{C}$ is an algebraic closure of $\mathbb{R}$. However, One should note that $\mathbb{C}$ is not an algebraic closure of $\mathbb{Q}$, since it is not algebraic over the field of rational numbers.*

One can use Zorn's lemma (see [6]) to prove the following important theorem.

Theorem 2.2.6 *Every field F has an algebraic closure.*

Although the algebraic closure of a field is not unique in general, it is unique up to isomorphism, i.e any two algebraic closures of a field F are isomorphic.

2.3 Roots of Polynomials

We all learned how to solve a quadratic equation in high school. Although no one taught us how to do the same thing for cubic or degree four equations at high school, but there we finally learned that similar methods exist for solving these equations. For a long time, people were wondering whether is it possible to extend the same idea for higher degree equations or not. Before getting in to details about roots of polynomials let us provide formal definitions.

Definition 2.3.1 *Assume K/F be a field extension and $p(x_1, \cdots, x_n)$ be a polynomial in $F[x_1, \cdots, x_n]$. $\alpha_1, \cdots, \alpha_n \in K$ is called a root of p if $p(\alpha_1, \cdots, \alpha_n) = 0$.*

By solving a polynomial, we mean finding the set of roots of it. A univariate polynomial is called solvable by radicals if it can be solved by a combination of the following operations:

- Operations of the field.
- Using n-th roots in the field for some positive integer (e.g. square root, third root and etc.)

Ruffini was the first one who argued the fact that a general univariate polynomial of degree five or higher is not solvable by radicals. Although his proof was not complete, one should admit it was a great achievement at that time. Later in 1824, Abel gave a complete proof. However the most prominent and inspiring proof was given by Galois. He not only proved the above statement, but also he could solve some other geometric problems which were unsolved for centuries. The results of his work these days is called Galois theory.

Knowing that degree five (and higher) univariate polynomials are (in general) unsolvable with radicals, people were looking for other methods to solve the problem. Indeed there are different approaches for any problem. In this special case one can use numerical methods to find approximations of roots or use symbolic computations to find roots. It is worth mentioning that there are methods combining these two approaches which are called hybrid methods.

As an example we can mention Gröbner Basis and the Buchberger algorithm (see [6] and

[8]) as a method of solving polynomial equations, by reducing them to univariate polynomials or to linear system. As a matter of fact, in many applications, in practice we just need an approximation of a root. However, there are some applications which need exact solutions.

All we have discussed until now is about solving one single equation. What if we have a system of non-linear equations and we are interested in the set of solutions?

Though modern algebraic geometry uses the language of schemes ([10]), the classic algebraic geometry define varieties as the set of solution of a family of polynomials. Algebraic geometers are interested in studying varieties and their properties. In some cases people want to calculate the set of points of a variety which is equivalent of solving a system of non-linear equations. There are relations between the number of variables in the non-linear systems and the dimension of the corresponding geometric object. Thus, in dimensions higher than one we want to be able to solve a system of multivariate polynomials over a field.

Again Gröbner basis is a powerful tool in solving a non-linear system of multivariate polynomials. In the next section we will describe the related application of resultant in solving a bivariate system of polynomials.

2.4 Solving Bivariate Polynomials

In this section we want to see a method which can be applied to solve a system of bivariate equations. We do not claim that this is the most efficient one. In fact this method will be used in the last chapter in order to calculate the desired minimal polynomials.

Definition 2.4.1 *Assume two non-zero polynomials $f(x)$ and $g(x)$ are given in $R[x]$, where R is a UFD*[1]*. If*

$$f(x) = \sum_{i=0}^{n} a_i x^i \quad and \quad g(x) = \sum_{i=0}^{m} b_i x^i$$

[1]Unique Factorization Domain, for a definition see [6].

then the Sylvester matrix of f and g which we denote by $\mathrm{Syl}_x(f,g)$ *is the* $m+n$ *by* $m+n$ *matrix*

$$\begin{bmatrix} a_n & a_{n-1} & \cdots & a_1 & a_0 & & \\ & a_n & a_{n-1} & \cdots & a_1 & a_0 & \\ & & \cdots & \cdots & \cdots & \cdots & \\ & & & a_n & \cdots & \cdots & a_0 \\ b_m & b_{m-1} & \cdots & b_1 & b_0 & & \\ & b_m & b_{m-1} & \cdots & b_1 & b_0 & \\ & & \cdots & \cdots & \cdots & \cdots & \\ & & & b_m & \cdots & \cdots & b_0 \end{bmatrix}$$

Definition 2.4.2 *For* $f(x), g(x) \in R[x]$, *where R is a UFD, the resultant of f and g with respect to x which is denoted by* $\mathrm{res}_x(f,g)$, *is the determinant of the Sylvester matrix.* $\mathrm{res}_x(0, f) = 0$ *for non-zero* $f \in R[x]$ *and by definition,* $\mathrm{res}_x(f,g) = 1$ *for non-zero constants in R.*

Example 2.4.1 *Using Maple built in functions, SylvesterMatrix and resultant, we calculate them for two polynomials. We can find the resultant directly or by forming the Sylvester matrix and computing its determinant.*

```
> with(LinearAlgebra):
```

```
> f:=x^4-6*x^2+1;
```

$$f := x^4 - 6\,x^2 + 1$$

```
> g:=x^5+2*x^3-x;
```

$$g := x^5 + 2\,x^3 - x$$

```
> A:=SylvesterMatrix(f,g,x);
```

$$A := \begin{bmatrix} 1 & 0 & -6 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & -6 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & -6 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & -6 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & -6 & 0 & 1 \\ 1 & 0 & 2 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 2 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 2 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 2 & 0 & -1 & 0 \end{bmatrix}$$

```
> Determinant(A);
```

$$784$$

```
> resultant(f,g,x);
```

$$784$$

One can see that the determinant of Sylvester matrix of two polynomials is equal to resultant of them.

The resultant is also helpful in computing the greatest common divisors of two polynomials, but here we are interested in its applications to solving system of equations.

We just recall that a solution of a system of n polynomials in m variables

$$f_1(x_1, \cdots, x_m) = 0$$

$$f_2(x_1, \cdots, x_m) = 0$$

$$\vdots$$

$$f_n(x_1, \cdots, x_m) = 0$$

over a field F is an m-tuple $(\alpha_1, \cdots, \alpha_m)$ in some field extension F such that it satisfies all equations, i.e. for each i we have $f_i(\alpha_1, \cdots, \alpha_m) = 0$.

Theorem 2.4.1 *(Sylvester's Criterion) [8] Let $f(x), g(x) \in R[x]$ where R is a UFD. Then $f(x)$ and $g(x)$ have a non-trivial common factor if and only if* $\mathrm{res}_x(f, g) = 0$.

Example 2.4.2 *In this example we can see that for two polynomials f and g which have a common factor $x-1$, the resultant of these two is zero.*

```
> f := x^5-3*x^4+2*x^3+x-1;
```

$$f := x^5 - 3\,x^4 + 2\,x^3 + x - 1$$

```
> g := x^6+x^5-2*x^4-x^2+x;
```

$$g := x^6 + x^5 - 2\,x^4 - x^2 + x$$

```
> resultant(f,g,x);
```

$$0$$

One can have the following general fact about resultants for multivariate polynomial systems.

Theorem 2.4.2 *(Fundamental Theorem of Resultants) [8] Let $\bar{F}$ be an algebraically closed field, and let*

$$f(x_1, \cdots, x_n) = \sum_{i=0}^{n} a_i(x_2, \cdots, x_n)x_1^i,$$

$$g(x_1, \cdots, x_n) = \sum_{i=0}^{m} b_i(x_2, \cdots, x_n)x_1^i$$

be elements of $\bar{F}[x_1, \cdots, x_n]$ of positive degrees in x_1. Then if $(\alpha_1, \cdots, \alpha_n)$ is a common zero of f and g, their resultant with respect to x_1 satisfies

$$\mathrm{res}_{x_1}(f, g)(\alpha_2, \cdots, \alpha_n) = 0$$

Conversely if the above resultant vanishes at $(\alpha_1, \cdots, \alpha_n)$, then at least one of the following holds:

1. $a_n(\alpha_2, \cdots, \alpha_n) = \cdots = a_0(\alpha_2, \cdots, \alpha_n) = 0$;

2. $b_m(\alpha_2, \cdots, \alpha_n) = \cdots = b_0(\alpha_2, \cdots, \alpha_n) = 0$;

3. $a_n(\alpha_2, \cdots, \alpha_n) = b_m(\alpha_2, \cdots, \alpha_n) = 0$;

4. *There exists* $\alpha_1 \in \bar{F}$ *such that* $(\alpha_1, \cdots, \alpha_n)$ *is a common zero of* f *and* g.

The above result gives us a way to reduce some system of bivariate polynomials into solving a univariate polynomial. In fact let $f(x, y)$ and $g(x, y)$ be two bivariate polynomials. Assume there exists a solution for their system which we call (α_1, α_2). By the above theorem we have $\mathrm{res}_x(f, g)(\alpha_2) = 0$. Thus in order to find (α_1, α_2) we can calculate $\mathrm{res}_x(f, g)$ and find its roots. For each root of the resultant, we can plug it in to the system and see if they have a common root. Obviously this works if we have an efficient way to solve a univariate equation. The following example illustrates this idea.

Example 2.4.3 *Assume that polynomials* $f(x) = x^2 - y^2 - 1$ *and* $g(x) = x^2 + yx - 7$ *are given. We want to find all* (α, β) *such that it satisfies both* f *and* g. *In order to do so, we first compute the resultant* $\mathrm{res}_x(f, g)$.

```
> with(LinearAlgebra):

> f := proc (x, y) options operator, arrow; x^2-y^2-1 end proc;
```

$$f := (x, y) \mapsto x^2 - y^2 - 1$$

```
> g := proc (x, y) options operator, arrow; x^2+y*x-7 end proc;
```

$$g := (x, y) \mapsto x^2 + yx - 7$$

```
> R := resultant(f(x, y), g(x, y), x);
```

$$R := -13\,y^2 + 36$$

Having the resultant, we can find its roots which gives us all candidates for β.

```
> ysolve := [solve(R)];
```

$$ysolve := [-\frac{6\sqrt{13}}{13}, \frac{6\sqrt{13}}{13}]$$

It is time to check if we can find an α *which completes* β *as a solution. By plugging in the values of resultant in* f, *we find candidates for* α.

```
> f(x, ysolve[1]);
```

$$x^2 - \frac{49}{13}$$

```
> xsolve := [solve(f(x, ysolve[1]))];
```

$$xsolve := [\frac{7\sqrt{13}}{13}, -\frac{7\sqrt{13}}{13}]$$

The final step is to check the pairs formed from candidates for α and β, to find which pair satisfies both equations.

```
> f(xsolve[2], ysolve[1]);
```

$$0$$

```
> g(xsolve[2], ysolve[1]);
```

$$0$$

```
> f(xsolve[1], ysolve[2]);
```

$$0$$

```
> g(xsolve[1], ysolve[2]);
```

$$0$$

This approach is combinatorially expensive in the number of variables. We will see a better technique using eigenvectors.

Chapter 3

Matrix Polynomials

3.1 Basic Definition

3.1.1 Matrix Polynomial

Matrices are very useful in mathematics and its applications [9]. There are various types of matrices. Among those, there are a family of matrices which instead of having numbers as their entries, they have polynomials as entries. One of the applications of this family is finding common roots of polynomials. We want to apply this method, so we will convert rootfinding of a system of polynomial equations in two variables into a so-called nonlinear eigenvalue problem and solve that by linearization. All of these concepts will be defined in this chapter.

Definition 3.1.1 *A univariate matrix polynomial $\boldsymbol{M}(x)$ of degree n is defined as*

$$\boldsymbol{M}(x) = \sum_{i=0}^{n} \boldsymbol{M}_i x^i \tag{3.1}$$

where $\boldsymbol{M}_i$ are numerical matrices, $i \geq 0$.

It is easy to see that any matrix polynomial is a matrix with univariate polynomials as entries. This duality is simple, but is not useless.

Example 3.1.1 *In this example we provide a matrix polynomial which is represented by a matrix with polynomial entries. We use Maple to rewrite it as a polynomial with matrix coefficients.*

```
> with(LinearAlgebra):
```

```
> with(MatrixPolynomialAlgebra):
```

```
> M:=<<x^2-2*x|x-1>,<x+3|x^3-1>>;
```

$$M := \begin{bmatrix} x^2 - 2x & x - 1 \\ x + 3 & x^3 - 1 \end{bmatrix}$$

```
> add(x^i*A[i], i = 0..3);
```

$$x^3 A_3 + x^2 A_2 + x A_1 + A_0$$

where

```
> A[3]:= coeff(M,x,3);
```

$$A[3] := \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

```
> A[2]:= coeff(M,x,2);
```

$$A[2] := \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$

```
> A[1]:= coeff(M,x,1);
```

$$A[1] := \begin{bmatrix} -2 & 1 \\ 1 & 0 \end{bmatrix}$$

```
> A[0]:= coeff(M,x,0);
```

$$A[0] := \begin{bmatrix} 0 & -1 \\ 3 & -1 \end{bmatrix}$$

3.1.2 Generalized Eigenvalues and Eigenvalue Problem

We recall that the eigenvalue problem for a matrix $\mathbf{A}$ is finding scalars λ and corresponding vectors v such that they satisfy the following (if they exist):

$$\mathbf{A}v = \lambda v.$$

One can easily convert above expression to

$$(\lambda I - \mathbf{A})v = 0.$$

In general one can form the determinant of $(xI - \mathbf{A})$, which is a univariate polynomial in x, and the roots of this polynomials are the eigenvalues. In practice finding the eigenvalues by above method needs polynomial factorization which is problematic. There are numerical methods that one can apply to get the eigenvalues. The most famous one among these methods is the QR algorithm; see [4].

If we replace the identity matrix above with another matrix $\mathbf{B}$, we may still be interested in finding λ and v such that

$$\mathbf{A}v = \lambda \mathbf{B}v.$$

Such a λ is called a generalized eigenvalue. There is another version of this notion, where a pair (α, β) (that are not both zero) of field elements, is called a generalized eigenvalue of a matrix pair $(\mathbf{A}, \mathbf{B})$ if there exists a non-zero vector v such that

$$\alpha \mathbf{A}v = \beta \mathbf{B}v.$$

We may take $|\alpha|^2 + |\beta|^2 = 1$ and $\mathrm{Re}(\alpha) \geq 0$, without loss of generality. If $\beta \neq 0$, $\lambda = \dfrac{\alpha}{\beta}$. If $\beta = 0$, we say $(\mathbf{A}, \mathbf{B})$ has an infinite eigenvalue.

There is a similar (to QR) algorithm for solving the generalized eigenvalue problem numerically, which is called the QZ algorithm; see [4].

Definition 3.1.2 *A pair of matrices $(\boldsymbol{A}, \boldsymbol{B})$ where $\boldsymbol{A}, \boldsymbol{B} \in \mathbb{C}^{n\times n}$, is called a matrix pencil if they define a generalized eigenvalue problem*

$$(\lambda \boldsymbol{B} - \boldsymbol{A})v = 0$$

and λ is called a generalized eigenvalue for the pencil $(\boldsymbol{A}, \boldsymbol{B})$ or equivalently (α, β) with $|\alpha|^2 + |\beta|^2 = 1$ and $\mathrm{Re}(\alpha) \geq 0$.
It is not hard to see t is a generalized eigenvalue of the pencil $(\boldsymbol{A}, \boldsymbol{B})$ if and only if t is a root of $\det(y\boldsymbol{B} - \boldsymbol{A})$. So one can restate the generalized eigenvalue problem as finding roots of the determinant of $(y\boldsymbol{B} - \boldsymbol{A})$.

The following definition presents formally the above notions and introduces some notations.

Definition 3.1.3 *Assume $\boldsymbol{A}$ and $\boldsymbol{B}$ are two matrices. Then the set of eigenvalues of $\boldsymbol{A}$ which is denoted by $\Lambda(\boldsymbol{A})$, is the set*

$$\Lambda(\boldsymbol{A}) = \{z : \det(zI - \boldsymbol{A}) = 0\}$$

The set of generalized eigenvalues of the pair $(\boldsymbol{A}, \boldsymbol{B})$, is denoted by $\Lambda(\boldsymbol{A}, \boldsymbol{B})$ is

$$\Lambda(\boldsymbol{A}, \boldsymbol{B}) = \{z : \det(z\boldsymbol{B} - \boldsymbol{A}) = 0\}$$

and finally the set of nonlinear eigenvalues of a matrix polynomial $\boldsymbol{P}(z) = z^d \boldsymbol{A}_d + \cdots + z\boldsymbol{A}_1 + \boldsymbol{A}_0$ is denoted by $\Lambda(\boldsymbol{P})$ and is given by

$$\Lambda(\boldsymbol{P}) := \{z : \det(\boldsymbol{P}(z)) = 0\}.$$

A matrix polynomial, $\mathbf{M}$, is called regular if $\det(\mathbf{M})$ is not identically zero and is called singular otherwise. One should note that there exist pencils which are singular. As an example for $\mathbf{A} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$ and $\mathbf{B} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$ then

$$\det(z\mathbf{A} - \mathbf{B}) = \det \begin{bmatrix} 0 & 0 \\ 0 & z - 1 \end{bmatrix} = 0.$$

3.1.3 The Companion Pencil

One of the simplest and most useful matrix polynomials is the companion pencil, which we introduce in this section. For simplicity we first define the companion pencil for an ordinary polynomial.

Definition 3.1.4 *Let $p(x)$ be a univariate polynomial over complex numbers. Writing $p(x)$ as*

$$p(x) = \sum_{i=0}^{n} a_i x^i$$

we define the pencil $(\boldsymbol{C}_0, \boldsymbol{C}_1) \in (\mathbb{C}^{n\times n})^2$ as

$$\boldsymbol{C}_0 = \begin{bmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ -a_0 & -a_1 & \cdots & \cdots & -a_{n-1} \end{bmatrix}$$

$$\boldsymbol{C}_1 = \begin{bmatrix} 1 & & & & \\ & 1 & & & \\ & & \ddots & & \\ & & & 1 & \\ & & & & a_n \end{bmatrix}$$

Note that if $a_n = 1$ (the polynomial is monic) then this is also called the Frobenius companion pencil. There are others (Fiedler, [5]).

If $(\boldsymbol{C}_0, \boldsymbol{C}_1)$ is the companion pencil of $p(x)$, a simple calculation shows that

$$p(x) = \det(\boldsymbol{C}_1 x - \boldsymbol{C}_0).$$

We have similar definition for a matrix polynomial $\mathbf{P}(x)$.

Definition 3.1.5 *Let $\boldsymbol{P}(x)$ be a matrix polynomial. That is, suppose*

$$\boldsymbol{P}(x) = \sum_{i=0}^{n} \boldsymbol{A}_i x^i$$

where for each i, $\boldsymbol{A}_i \in \mathbb{C}^{n\times n}$. Then we define the companion pencil (also known as a "linearization" of the matrix polynomial) of $\boldsymbol{P}(x)$ as $(\boldsymbol{C}_0, \boldsymbol{C}_1) \in (\mathbb{C}^{n\times n})^2$ where

$$
\boldsymbol{C}_0 = \begin{bmatrix} 0 & \boldsymbol{I} & & & \\ & 0 & \boldsymbol{I} & & \\ & & \ddots & \ddots & \\ & & & \ddots & \boldsymbol{I} \\ -\boldsymbol{A}_0 & -\boldsymbol{A}_1 & \cdots & \cdots & -\boldsymbol{A}_{n-1} \end{bmatrix}
$$

$$
\boldsymbol{C}_1 = \begin{bmatrix} \boldsymbol{I} & & & & \\ & \boldsymbol{I} & & & \\ & & \ddots & & \\ & & & \boldsymbol{I} & \\ & & & & \boldsymbol{A}_n \end{bmatrix}
$$

One can see that

$$
\det(\mathbf{P}(x)) = \det(x\mathbf{C}_1 - \mathbf{C}_0)
$$

We will see this equality in section 3.4. Note that there are other linearizations.

3.2 Bézout Matrix

In this section we want to introduce the Bézout matrix for two polynomials. Let $f(x)$ and $g(x)$ be two polynomials of degree m and n respectively:

$$
f(x) = \sum_{i=0}^{m} a_i x^i
$$

$$
g(x) = \sum_{i=0}^{n} c_i x^i
$$

The Cayley quotient of f and g is defined as follows, where $d = \max(m, n)$:

$$
C(x, y) = \frac{f(x)g(y) - g(x)f(y)}{x - y}
$$

Since $x = y$ obviously makes

$$
f(x)g(y) - g(x)f(y) = 0
$$

then $x - y$ is a factor of the numerator.

Hence, we can write the Cayley quotient as:

$$\sum_{i=1}^{d}\sum_{j=1}^{d} b_{ij}x^{i-1}y^{j-1}$$

The Bézout matrix is the following $d \times d$ matrix:

$$\mathbf{B} = (b_{ij})_{d\times d}$$

Then

$$C(x,y) = \begin{bmatrix} 1 & x & x^2 & \cdots & x^{d-1} \end{bmatrix} \mathbf{B} \begin{bmatrix} 1 \\ y \\ y^2 \\ \vdots \\ y^{d-1} \end{bmatrix}$$

Example 3.2.1 *Let's construct the Bézout matrix in different ways. For a given pair of polynomials we form the Cayley quotient and then construct the Bézout matrix from it. On the other hand we can use the built in function for the Bézout matrix in Maple with different options.*

```
> with(LinearAlgebra):
```

```
> f := randpoly(x, degree = 2);
```

$$f := -7\,x^2 + 22\,x - 55$$

```
> g := randpoly(x, degree = 3);
```

$$g := -94\,x^3 + 87\,x^2 - 56\,x$$

```
> C := simplify(expand(f*subs(x = y, g)-subs(x = y, f)*g)/(x-y));
```

$$658\,x^2y^2 - 2068\,x^2y - 2068\,xy^2 + 5170\,x^2 + 6692\,xy + 5170\,y^2 - 4785\,x - 4785\,y + 3080$$

```
> d := max(degree(f, x), degree(g, x));
```

$$d := 3$$

```
> Bez := Matrix(d, d, (i, j) -> coeff(coeff(C, x,d-i), y, d-j));
```

$$B2 := \begin{bmatrix} -658 & 2068 & -5170 \\ 2068 & -6692 & 4785 \\ -5170 & 4785 & -3080 \end{bmatrix}$$

```
> B1:=BezoutMatrix(f,g);
```

$$B1 := \begin{bmatrix} 4778 & -3553 & 0 \\ -1459 & 4778 & 0 \\ -7 & 22 & -55 \end{bmatrix}$$

```
> B2:=BezoutMatrix(f,g, method = symmetric);
```

$$B2 := \begin{bmatrix} -658 & 2068 & -5170 \\ 2068 & -6692 & 4785 \\ -5170 & 4785 & -3080 \end{bmatrix}$$

Remark Obviously the above example shows different matrices for the corresponding Bézout matrix to a fixed pair of polynomials. However, one can easily verify that the determinants of the matrices are equal up to a constant.

We note that we are looking at Cayley quotient in monomial basis. We are able to write the Cayley quotient in the Lagrange basis and construct a similar Bézout matrix. This will be investigated in the next section (see [15]).

3.3 Bézout Matrix in Lagrange Basis

As we mentioned in the previous section, it is possible to write the Cayley quotient in a Lagrange basis. This has some computational benefits. In [15] Shakoori described the Bézout matrix in Lagrange basis. What we are presenting here is a brief discussion of her paper and we encourage the reader to see [15] for more details.

We defined the Bézout matrix using the Cayley quotient.

$$C(x,y) = \frac{f(x)g(y) - f(y)g(x)}{x-y}$$

Recall that $d = \max(\deg(f), \deg(g))$. It is not hard to see that $C(x, y)$ is a polynomial in terms of x and also in y of degree at most $d - 1$. Indeed one can write the Cayley quotient as:

$$C(x,y) = \begin{bmatrix} 1 & x & x^2 & \cdots & x^{d-1} \end{bmatrix} \mathbf{B} \begin{bmatrix} 1 \\ y \\ y^2 \\ \vdots \\ y^{d-1} \end{bmatrix}$$

where $\mathbf{B}$ is the corresponding Bézout matrix for f and g.

A closer look gives the idea of interpolating the polynomials x^i and y^i for $0 \le i \le d-1$. In order to do so we set two sets of nodes, namely $X = \{x_1, \cdots, x_d\}$ and $Y = \{y_1, \cdots, y_d\}$. We also write $L_i(x)$ ($L_i(y)$ resp.) i.e.

$$L_i(x) = \prod_{\substack{0 \le j \le d \\ i \ne j}} \frac{x - x_j}{x_i - x_j}$$

for the i-th Lagrange polynomial for x_i (y_i resp). So after all, we can have interpolation for each x^i and y^i. One can write the interpolations as follow:

$$\begin{bmatrix} 1 \\ x \\ x^2 \\ \vdots \\ x^{d-1} \end{bmatrix} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_d \\ x_1^2 & x_2^2 & \cdots & x_d^2 \\ \vdots & \vdots & \cdots & \vdots \\ x_1^{d-1} & x_2^{d-1} & \cdots & x_d^{d-1} \end{bmatrix} \begin{bmatrix} L_1(x) \\ L_2(x) \\ L_3(x) \\ \vdots \\ L_d(x) \end{bmatrix}$$

In the above expression the middle matrix is the transpose of well-known Vandermonde matrix. One can write a similar expression for row matrix of x^is:

$$\begin{bmatrix} 1 & x & x^2 & \cdots & x^{d-1} \end{bmatrix} = \begin{bmatrix} L_1(x) & L_2(x) & L_3(x) & \cdots & L_d(x) \end{bmatrix} \begin{bmatrix} 1 & x_1 & \cdots & x_1^{d-1} \\ 1 & x_2 & \cdots & x_2^{d-1} \\ 1 & x_3 & \cdots & x_3^{d-1} \\ \vdots & \vdots & \cdots & \vdots \\ 1 & x_d & \cdots & x_d^{d-1} \end{bmatrix}$$

Now we can get back to write the Cayley quotient in the Lagrange basis, by replacing the appropriate matrices we wrote above. Hence

$$C(x,y) = \begin{bmatrix} L_1(x) & L_2(x) & L_3(x) & \cdots & L_d(x) \end{bmatrix} \mathbf{V}_X \mathbf{B} \mathbf{V}_Y^T \begin{bmatrix} L_1(y) \\ L_2(y) \\ L_3(y) \\ \vdots \\ L_d(x) \end{bmatrix}$$

where

$$\mathbf{V}_X = \begin{bmatrix} 1 & x_1 & \cdots & x_1^{d-1} \\ 1 & x_2 & \cdots & x_2^{d-1} \\ 1 & x_3 & \cdots & x_3^{d-1} \\ \vdots & \vdots & \cdots & \vdots \\ 1 & x_d & \cdots & x_d^{d-1} \end{bmatrix}$$

We call $\mathbf{B}_L = \mathbf{V}_X \mathbf{B} \mathbf{V}_Y^T$ the Bézout matrix in Lagrange basis. We do not compute it this way, however, except as demonstration.

Example 3.3.1 *The following Maple code outputs the Bézout matrix for a pair of polynomials* (p, f) *where* p *is a univariate polynomial in* x*, and* f *is a polynomial derived from the relation between* u_i *and* u_{i+1} *such that* $f(u_i, u_{i+1}) = 0$*. In chapter* 4 *we will talk about* f *in more detail.*

```
with(LinearAlgebra):
##This function returns Bezout matrix based on monomial basis, for the
    input p which is a polynomial and f (which is fixed).
bez :=proc(p)
  local f, B, n, c, i, j;
  f := (y-2)^2-2*x*y+x^2;
  n := max(degree(f), degree(p));
  B := Matrix(n); ##It is a blank square matrix of size n.
  c := simplify(expand((f*(eval(p, x = s))-p*(eval(f, x = s)))/(x-s)));
##This loop assigns coefficients of Cayley quotient to B.
   for i from 0 to n-1 do
     for j from 0 to n-1 do
        B[i+1, j+1] := coeff(coeff(c, x, i), s, j);
```

```
    od;
  od;
 B;
end proc:
```

```
BezL := proc (p, r)
 local v, vT, BM, BL, m, n;
 BM := bez(p);
 m := degree(p);
   if nops(r) = m then
     v := VandermondeMatrix(<r>);
     vT := Transpose(v);
     BL := Multiply(v, Multiply(BM, vT));
   else
   return (0);
   end if;
end proc:
```

```
> BezL(390*x^2-1508*x+1409, [2, 4]);
```

$$\begin{bmatrix} -52\,y^2+510\,y-604 & -832\,y^2+6750\,y-6938 \\ -832\,y^2+6750\,y-6938 & -1612\,y^2+16110\,y-19304 \end{bmatrix}$$

In [15] using above discussion, Shakoori proved the following efficient way of computing the Bézout matrix in Lagrange basis.

Theorem 3.3.1 *[15] For two bivariate polynomials f and g by their values at some known points, the Bézout matrix in Lagrange basis is defined by*

$$\boldsymbol{B}^L_{ij} = \frac{f(x_i)g(y_j) - f(y_j)g(x_i)}{x_i - y_j} \qquad if \;\; i \neq j \;\; and \quad 0 \leq i, j \leq d$$

$$\boldsymbol{B}^L_{ii} = f_i' g_i - f_i g_i' \qquad 0 \leq i \leq d.$$

where $f_i = f(x_i)$, $g_i = g(x_i)$, $f_i' = f'(x_i)$ and $g_i' = g'(x_i)$. Moreover, the null vectors of $\boldsymbol{B}^L$ may be parametrized by the Lagrange basis polynomials evaluated at the common roots of f and g.

Shakoori also introduced an effective method for calculating the values of the derivatives of f and g using their values on a set of points. Here we just briefly present her results.

Assume we have a polynomial $p(x)$ of degree n and set of distinct nodes $\{x_0, \cdots, x_n\}$. If we denote the values of p on x_i by p_i and the values of p' by p'_i, then we can compute the values of the derivative using the following formula:

$$\mathbf{D}\begin{bmatrix} p_0 \\ p_1 \\ \vdots \\ p_n \end{bmatrix} = \begin{bmatrix} p'_0 \\ p'_1 \\ \vdots \\ p'_n \end{bmatrix}$$

where

$$\mathbf{D}_{ii} = \sum_{i \neq j} \frac{1}{x_i - x_j}$$

$$\mathbf{D}_{ij} = \prod_{\substack{k \neq j \\ k \neq i}} (x_i - x_k) / \prod_{k \neq j} (x_j - x_k)$$

3.4 Linearization of a Matrix polynomial

We already introduced matrix polynomials and their generalized eigenvalues. We are interested in studying the set of non-linear eigenvalues which is called the spectrum of the matrix polynomial. Normally we have no control on the degree of the given matrix polynomial. It would be helpful if we could use a linear matrix polynomial, i.e. a matrix polynomial of the form $\mathbf{A}x - \mathbf{B}$ (where $\mathbf{A}$ and $\mathbf{B}$ are numerical matrices) such that the spectra of $\mathbf{A}x - \mathbf{B}$ and the original matrix polynomial are the same.

Definition 3.4.1 *[9] Let*

$$\boldsymbol{P}(x) = \sum_{i=0}^{n} \boldsymbol{A}_i x^i$$

be a regular (i.e. its determinant in not identically zero) $m \times m$ matrix polynomial. An $mn \times mn$ linear matrix polynomial $\boldsymbol{S}_1 x + \boldsymbol{S}_0$ is called a linearization of $\boldsymbol{P}(x)$ if

$$\begin{bmatrix} \boldsymbol{P}(x) & 0 \\ 0 & \boldsymbol{I}_{m(n-1)} \end{bmatrix} = \boldsymbol{E}(x)(\boldsymbol{S}_1 x + \boldsymbol{S}_0)\boldsymbol{F}(x)$$

for some $mn \times mn$ matrix polynomial $\boldsymbol{E}(x)$ and $\boldsymbol{F}(x)$ with non-zero constant determinants.

Assume $\mathbf{P}(x)$ is a given matrix polynomial of degree n. So we can write $\mathbf{P}(x)$ as:

$$\mathbf{P}(x) = \mathbf{A}_n x^n + \cdots + \mathbf{A}_1 x + \mathbf{A}_0.$$

In the simplest case if $\mathbf{A}_n$ is a non-singular matrix, we can form the companion matrix for

$$\mathbf{A}_n^{-1}\mathbf{A}_n x^n + \cdots + \mathbf{A}_n^{-1}\mathbf{A}_1 x + \mathbf{A}_n^{-1}\mathbf{A}_0$$

i.e.

$$\mathbf{C}_p = \begin{bmatrix} 0 & \mathbf{I} & 0 & \cdots & 0 \\ 0 & 0 & \mathbf{I} & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & \mathbf{I} \\ -\mathbf{A}_n^{-1}\mathbf{A}_0 & -\mathbf{A}_n^{-1}\mathbf{A}_1 & -\mathbf{A}_n^{-1}\mathbf{A}_2 & \cdots & -\mathbf{A}_n^{-1}\mathbf{A}_{n-1} \end{bmatrix}$$

Having $\mathbf{C}_p$ we can verify that $x\mathbf{I} - \mathbf{C}_p$ is a linearization for $\mathbf{P}(x)$, using

$$\mathbf{F}(x) = \begin{bmatrix} \mathbf{I} & 0 & 0 & \cdots & 0 \\ -x\mathbf{I} & \mathbf{I} & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{I} & 0 \\ 0 & 0 & \cdots & -x\mathbf{I} & \mathbf{I} \end{bmatrix}$$

and

$$\mathbf{E}(x) = \begin{bmatrix} \mathbf{B}_{n-1}(x) & \mathbf{B}_{n-2}(x) & \cdots & & \mathbf{B}_0(x) \\ -\mathbf{I} & 0 & \cdots & & 0 \\ 0 & -\mathbf{I} & \cdots & & \vdots \\ \vdots & \vdots & \ddots & & \vdots \\ 0 & 0 & \cdots & -\mathbf{I} & 0 \end{bmatrix}$$

where $\mathbf{B}_0(x) = \mathbf{I}$ and $\mathbf{B}_{t+1}(x) = x\mathbf{B}_t(x) + \mathbf{A}_{n-t-1}$ for $0 \leq t \leq n-2$ and forming

$$\begin{bmatrix} \mathbf{P}(x) & 0 \\ 0 & \mathbf{I}_{m(n-1)} \end{bmatrix} \mathbf{F}(x) = \begin{bmatrix} \mathbf{A}_n & 0 \\ 0 & \mathbf{I}_{m(n-1)} \end{bmatrix} \mathbf{E}(x)(\mathbf{I}x - \mathbf{C}_p).$$

Obviously the linearization of a matrix is not unique. More precisely having a linearization $\mathbf{A}x - \mathbf{B}$ we can find many more linearizations by applying transformations. In other words for any non-singular $\mathbf{C}$ and $\mathbf{D}$, $\mathbf{C}(\mathbf{A}x - \mathbf{B})\mathbf{D}$ is a linearization again. In particular, if we use the following transformations:

$$\mathbf{C} = \begin{bmatrix} \mathbf{I} & 0 & \cdots & 0 & 0 \\ 0 & \mathbf{I} & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \mathbf{I} & 0 \\ 0 & 0 & \cdots & 0 & \mathbf{A}_n \end{bmatrix}$$

And $\mathbf{D} = \mathbf{I}$ (of appropriate size) then

$$\mathbf{C}(\mathbf{I}x - \mathbf{C}_p)\mathbf{D} = x\begin{bmatrix} \mathbf{I} & 0 & \cdots & 0 & 0 \\ 0 & \mathbf{I} & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{I} & 0 \\ 0 & 0 & \cdots & 0 & \mathbf{A}_n \end{bmatrix} - \begin{bmatrix} 0 & \mathbf{I} & 0 & \cdots & 0 \\ 0 & 0 & \mathbf{I} & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & \mathbf{I} \\ -\mathbf{A}_0 & -\mathbf{A}_1 & \cdots & -\mathbf{A}_{n-2} & -\mathbf{A}_{n-1} \end{bmatrix}$$

The above matrix polynomial is the companion pencil of $\mathbf{P}(x)$ which we already introduced in section 3.1.3

Example 3.4.1 *For the given matrix polynomial we calculate $\boldsymbol{C}_p$ and then compare the determinants of p and $x\boldsymbol{I}-\boldsymbol{C}_p$. We also do the same thing for $\boldsymbol{C}_1x-\boldsymbol{C}_0$.*

```
> with(LinearAlgebra):

> A2 := RandomMatrix(3, 3, generator = 1 .. 10);
```

$$A2 := \begin{bmatrix} 10 & 9 & 10 \\ 1 & 7 & 6 \\ 6 & 1 & 6 \end{bmatrix}$$

```
> Determinant(A2);
```



```
> A1 := RandomMatrix(3, 3, generator = 1 .. 10);
```

$$A1 := \begin{bmatrix} 7 & 1 & 7 \\ 3 & 3 & 8 \\ 6 & 10 & 7 \end{bmatrix}$$

```
> A0 := RandomMatrix(3, 3, generator = 1 ..  2);
```

$$A0 := \begin{bmatrix} 1 & 1 & 1 \\ 2 & 1 & 1 \\ 2 & 2 & 1 \end{bmatrix}$$

```
> P := A2*x^2+A1*x+A0;
```

$$P := \begin{bmatrix} 10x^2+7x+1 & 9x^2+x+1 & 10x^2+7x+1 \\ x^2+3x+2 & 7x^2+3x+1 & 6x^2+8x+1 \\ 6x^2+6x+2 & x^2+10x+2 & 6x^2+7x+1 \end{bmatrix}$$

```
> Cp :=<<ZeroMatrix(3)|IdentityMatrix(3)>
> ,<Multiply(-A2^(-1), A0)| Multiply(-A2^(-1), A1)>>>;
```

$$Cp := \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ \frac{21}{55} & 2/11 & \frac{6}{55} & -\frac{6}{55} & \frac{64}{55} & \frac{53}{55} \\ \frac{7}{22} & \frac{7}{22} & 1/11 & \frac{9}{22} & \frac{47}{22} & \frac{7}{11} \\ -\frac{169}{220} & -\frac{25}{44} & -\frac{16}{55} & -\frac{211}{220} & -\frac{701}{220} & -\frac{123}{55} \end{bmatrix}$$

```
> Determinant(P);
```

$$220x^6+46x^5-594x^4-373x^3-30x^2+10x+1$$

```
> D := Determinant(x*IdentityMatrix(6)-Cp);
```

$$D := x^6+\frac{23x^5}{110}-\frac{27x^4}{10}-\frac{373x^3}{220}-\frac{3x^2}{22}+x/22+\frac{1}{220}$$

```
> 220*D;
```

$$220\,x^6 + 46\,x^5 - 594\,x^4 - 373\,x^3 - 30\,x^2 + 10\,x + 1$$

```
> C0 :=<<ZeroMatrix(3)|IdentityMatrix(3)>,<-A0|-A1>>>;
```

$$C0 := \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ -1 & -1 & -1 & -7 & -1 & -7 \\ -2 & -1 & -1 & -3 & -3 & -8 \\ -2 & -2 & -1 & -6 & -10 & -7 \end{bmatrix}$$

```
> C1 := DiagonalMatrix([IdentityMatrix(3), A2]);
```

$$C1 := \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 10 & 9 & 10 \\ 0 & 0 & 0 & 1 & 7 & 6 \\ 0 & 0 & 0 & 6 & 1 & 6 \end{bmatrix}$$

```
> Determinant(C1*x-C0);
```

$$220\,x^6 + 46\,x^5 - 594\,x^4 - 373\,x^3 - 30\,x^2 + 10\,x + 1$$

As we can see in the previous example, the linearization and the original matrix have the same determinant and this is not a coincidence. It is always the case.
For more details about linearizations see [9] or [11].

3.5 Linearization with Lagrange Basis

So far we defined linearizations in the monomial basis. In this section we want to talk about a linearization in a Lagrange basis.

Suppose $\mathbf{P}(x)$ to be a matrix polynomial of degree n which is given in monomial basis.

$$\mathbf{P}(x) = \mathbf{A}_n x^n + \cdots + \mathbf{A}_1 x + \mathbf{A}_0$$

Let's have $n+1$ sample nodes for writing $\mathbf{P}(x)$ in Lagrange basis. Assuming $x_0, \cdots, x_n$ to be distinct points, the values of $\mathbf{P}(x)$ over these points will be shown by $\mathbf{P}_i$, i.e. $\mathbf{P}_i = \mathbf{P}(x_i)$. Now Lagrange polynomials are presented as

$$L_i(x) = \prod_{j=0, j\neq i}^{n} \frac{x - x_j}{x_i - x_j}$$

It is convenient to define

$$\omega_i = \prod_{j=0, j\neq i}^{n} \frac{1}{x_i - x_j}.$$

Then $\mathbf{P}(x)$ in Lagrange polynomial will be

$$\mathbf{P}(x) = \sum_{i=0}^{n} L_i(x)\mathbf{P}_i$$

Amiraslani in his Ph.D. dissertation [2], calculated the companion pencil for $\mathbf{P}(x)$ in Lagrange basis:

$$x\mathbf{C}_1 - \mathbf{C}_0 = \begin{bmatrix} (x-x_0)\mathbf{I} & 0 & \cdots & 0 & -\mathbf{P}_0 \\ 0 & (x-x_1)\mathbf{I} & \cdots & 0 & -\mathbf{P}_1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & (x-x_n)\mathbf{I} & -\mathbf{P}_n \\ \omega_0\mathbf{I} & \omega_1\mathbf{I} & \cdots & \omega_n\mathbf{I} & 0 \end{bmatrix}$$

which is a linearization for $\mathbf{P}(x)$ (see [1]).

Amiraslani et al. in [1] introduced a matrix $\mathbf{A}$ as

$$\mathbf{A} = \begin{bmatrix} \omega_0^{-1}\mathbf{P}_0 & 0 & \cdots & 0 & 0 \\ 0 & \omega_1^{-1}\mathbf{P}_1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \omega_n^{-1}\mathbf{P}_n & 0 \\ 0 & 0 & \cdots & 0 & -\mathbf{I} \end{bmatrix}$$

such that

$$(x\mathbf{C}_1 - \mathbf{C}_0)\mathbf{A} = \begin{bmatrix} \frac{x-x_0}{\omega_0}\mathbf{P}_0 & 0 & \cdots & 0 & \mathbf{P}_0 \\ 0 & \frac{x-x_1}{\omega_1}\mathbf{P}_1 & \cdots & 0 & \mathbf{P}_1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \frac{x-x_n}{\omega_n}\mathbf{P}_n & \mathbf{P}_n \\ \mathbf{P}_0 & \mathbf{P}_1 & \cdots & \mathbf{P}_n & 0 \end{bmatrix}$$

The resulting matrix is block symmetric. We use this fact in the next chapter.

Chapter 4

A Sequence of Symmetric Bézout Matrix Polynomials

4.1 Introduction

In this chapter we state the main results. The first observation for the sequence of u_is is that all of the terms are algebraic over $\mathbb{Q}$. This fact is stated formally in section 4.2. In section 4.3 a lower and upper bound for u_is is presented. Section 4.4 is devoted to a constructive approach for finding the sequence of minimal polynomials. Having minimal polynomials, showing that "the roots of minimal polynomials are all real" is the subject of section 4.5 and finally section 4.6 is devoted to finding symmetric linearization of the Bézout matrix.

Before getting in to details let us recall the definition of the sequence of u_is.

$$u_1 = 1$$

$$u_{n+1} = 2 + u_n + 2\sqrt{u_n} \tag{4.1}$$

It is easy to see that the first few terms in the sequence are:

$$u_1 = 1$$

$$u_2 = 5$$

$$u_3 = 7 + 2\sqrt{5}$$

$$u_4 = 9 + 2\sqrt{5} + 2\sqrt{7 + 2\sqrt{5}} \tag{4.2}$$

By looking at these numbers, one might guess the degree of minimal polynomial of each u_i over $\mathbb{Q}$. We will find the degrees by considering a tower of field extensions over rational numbers, in section 4.3.

4.2 $\{u_i\}$ is Algebraic over the set of Rational Numbers

In this section we want to show that all u_is are algebraic over $\mathbb{Q}$. For this reason we can look at field extensions generated by u_i over $\mathbb{Q}$ for all $i \geq 2$.

Theorem 4.2.1 *For each i, u_i is algebraic over $\mathbb{Q}$.*

Proof The proof is by induction on i. We first show that u_3 is algebraic over rational numbers. We can do as below:

$$u_3 = 7 + 2\sqrt{5} \Longrightarrow u_3 - 7 = 2\sqrt{5} \Longrightarrow (u_3 - 7)^2 = 20$$

Hence

$$u_3^2 - 14u_3 + 29 = 0$$

So u_3 satisfies

$$x^2 - 14x + 29 = 0.$$

Now assume that u_i is algebraic over $\mathbb{Q}$ for some $i \geq 3$. Consider the extensions below:

$$\begin{array}{c} \mathbb{Q}(u_i)(u_{i+1}) = \mathbb{Q}(u_i, u_{i+1}) \\ | \\ \mathbb{Q}(u_i) \\ | \\ \mathbb{Q} \end{array}$$

By our assumption, we know that $\mathbb{Q}(u_i)/\mathbb{Q}$ is an algebraic extension.

On the other hand, we have

$$u_{i+1} = 2 + u_i + 2\sqrt{u_i}$$

$$\implies (u_{i+1} - 2 - u_i)^2 = 4u_i$$

$$\implies ((u_{i+1} - 2) - u_i)^2 = 4u_i$$

$$\implies (u_{i+1} - 2)^2 - 2(u_{i+1} - 2)u_i + u_i^2 - 4u_i = 0$$

$$\implies (u_{i+1} - 2)^2 - 2u_{i+1}u_i + u_i^2 = 0$$

So u_{i+1} satisfies $(x-2)^2 - 2xu_i + u_i^2 = 0$. Thus u_{i+1} is algebraic over $\mathbb{Q}(u_i)$. By induction assumption $\mathbb{Q}(u_i)/\mathbb{Q}$ is algebraic. Now Theorem 2.2.5 tells us u_{i+1} is algebraic over $\mathbb{Q}$.

$\natural$

4.3 Boundary of $\{u_i\}$

In this section we want to find the upper and lower bound for u_n.

Theorem 4.3.1 *If $n \geq 5$, then*

$$u_n \geq (n+1)^2 - 5.$$

Proof We can prove this statement by using induction on n.
Let $n = 5$. One can see that

$$u_5 = 31.24 \geq (5+1)^2 - 5 = 36 - 5 = 31$$

which holds for $i = 5$.

Now assume the statement is true for some $k \geq 5$. i.e.

$$u_k \geq (k+1)^2 - 5.$$

We want to show that

$$u_{k+1} \geq (k+2)^2 - 5.$$

For showing this, we have:

$$u_{k+1} = 2 + u_k + 2\sqrt{u_k} \geq 2 + (k+1)^2 - 5 + 2\sqrt{(k+1)^2 - 5}.$$

Now we want to show that the last term is greater that $(k+2)^2 - 5$

$$2 + (k+1)^2 - 5 + 2\sqrt{(k+1)^2 - 5} - (k+2)^2 + 5 \geq 0$$

$$2 + (k^2 + 2k + 1) - 5 + 2\sqrt{(k+1)^2 - 5} - (k^2 + 4k + 4) + 5 \geq 0$$

$$2\sqrt{(k+1)^2 - 5} - 2k - 2 \geq 0 \tag{4.3}$$

Since $2\sqrt{(k+1)^2 - 5} - 2k - 2 = 0$ has only one positive root which is less than 5 and it is ascending, equation 4.3 holds for $k \geq 5$.

♮

Theorem 4.3.2 *If $n \geq 5$, then*

$$u_n \leq \frac{5}{4}n^2.$$

Proof The proof is by induction on n.

Let $n = 5$. One can see that

$$u_5 = 31.24 \leq \frac{5}{4}(5)^2 = 31.25$$

Assume that the statement is true for some $k \geq 5$. i.e.

$$u_k \leq \frac{5}{4}k^2.$$

We want to show that

$$u_{k+1} \leq \frac{5}{4}(k+1)^2.$$

We know that

$$u_{k+1} = 2 + u_k + 2\sqrt{u_k} \leq 2 + \frac{5}{4}k^2 + 2k\sqrt{\frac{5}{4}}.$$

Now we claim that

$$2 + \frac{5}{4}k^2 + 2k\sqrt{\frac{5}{4}} \leq \frac{5}{4}(k+1)^2.$$

We have:

$$\frac{5}{4}(k+1)^2 - 2 - \frac{5}{4}k^2 - 2k\sqrt{\frac{5}{4}} \geq 0$$

$$\frac{5}{4}(k+1)^2 - \frac{5}{4}(\frac{2}{\frac{5}{4}} + k^2 + \frac{2k}{\sqrt{\frac{5}{4}}}) \geq 0$$

$$\frac{5}{4}((k+1)^2 - \frac{8}{5} - k^2 - \frac{4k}{\sqrt{5}}) \geq 0$$

$$k^2 + 2k + 1 - \frac{8}{5} - k^2 - \frac{4k}{\sqrt{5}} \geq 0$$

So

$$-\frac{3}{5} - k(\frac{4}{\sqrt{5}} - 2) \geq 0 \tag{4.4}$$

We easily can see that 2.84 is the root of equation 4.4. So for $k \geq 3$, equation 4.4 is true and we are done.

♮

4.4 Minimal Polynomials of $\{u_i\}$

In the previous sections we showed that u_is are algebraic over $\mathbb{Q}$. In this section we want to explicitly find their minimal polynomials. More precisely, we will construct the minimal polynomials recursively. We show that the recursive minimal polynomial of u_{i+1} is the resultant of minimal polynomial of u_i and an auxiliary polynomial. By 4.1 we can construct $f(x, y) = (y-2)^2 - 2yx + x^2$ such that (u_i, u_{i+1}) satisfies it.

Now we want to apply the fundamental theorem of resultants (theorem 2.4.2) assuming $F = \mathbb{Q}$, then $\bar{F} = \bar{\mathbb{Q}}$ (a fixed algebraic closure of $\mathbb{Q}$) and

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

and

$$g(x, y) = p_i(x, y)$$

and $(\alpha_1, \alpha_2) = (u_i, u_{i+1})$, then applying the theorem gives us

$$\mathrm{res}_x(f(x, y), p_i(x))(u_{i+1}) = 0 \Longrightarrow q(u_{i+1}) = 0$$

where $q(y)$ is the resultant of f and p_i with respect to x.

We recall that resultant is the the determinant of the Bézout matrix (see section 3.2).

Since $q(u_{i+1}) = 0$, p_{i+1} must divide $q(y)$. On the other hand, we know

$$\deg(q(y)) \leq \deg(f) \cdot \deg(p_i) = 2 \cdot \deg(p_i)$$

We claim that, $p_{i+1} = Cq(y)$ where C is a constant. In order to show this, we only need to show

$$\deg(p_{i+1}) = \deg(q(y))$$

As we have seen above, $p_{i+1}|q(y)$, which means

$$\deg(p_{i+1}) \leq \deg(q(y)) \leq 2 \cdot \deg(p_i)$$

So, it is enough to show

$$\deg(p_{i+1}) = 2 \cdot \deg(p_i)$$

Consider the diagram of field extensions, below:

$$\begin{tikzcd}
& \mathbb{Q}(u_i, u_{i+1}) \arrow[dl, dash, "2"'] \arrow[dr, dash] & \\
\mathbb{Q}(u_i) \arrow[dr, dash, "n"'] & & \mathbb{Q}(u_{i+1}) \arrow[dl, dash] \\
& \mathbb{Q} &
\end{tikzcd}$$

It is not hard to see that $\mathbb{Q}(u_i, u_{i+1}) = \mathbb{Q}(\sqrt{u_i})$, as follows:

We know that

$$u_i = \sqrt{u_i}^2 \in \mathbb{Q}(\sqrt{u_i})$$

and

$$u_{i+1} = 2 + u_i + 2\sqrt{u_i} \in \mathbb{Q}(\sqrt{u_i})$$

So we can see that $\mathbb{Q}(u_i, u_{i+1}) \subseteq \mathbb{Q}(\sqrt{u_i})$

Now for the other inclusion:

$$\sqrt{u_i} = \frac{(u_{i+1} - 2 - u_i)}{2} = \frac{1}{2}u_{i+1} - \frac{1}{2}u_i - 1 \in \mathbb{Q}(u_i, u_{i+1})$$

So

$$\mathbb{Q}(\sqrt{u_i}) \subseteq \mathbb{Q}(u_i, u_{i+1}).$$

Hence we can write the above diagram as

$$\begin{tikzcd}
& \mathbb{Q}(\sqrt{u_i}) \arrow[dl, dash, "2"'] \arrow[dr, dash] & \\
\mathbb{Q}(u_i) \arrow[dr, dash, "n"'] & & \mathbb{Q}(u_{i+1}) \arrow[dl, dash] \\
& \mathbb{Q} &
\end{tikzcd}$$

We know that for any simple extension[1] (See definition 2.1.1) of $\mathbb{Q}$ such as $\mathbb{Q}(u_i)$, $[\mathbb{Q}(u_i) : \mathbb{Q}]$ is equal to degree of minimal polynomial of u_i over $\mathbb{Q}$. Also we know, $\mathbb{Q}(u_i, u_{i+1}) = \mathbb{Q}(u_i)(u_{i+1})$ is a simple extension of $\mathbb{Q}(u_i)$, since the minimal polynomial of u_{i+1} over $\mathbb{Q}(u_i)$ is $f(u_i, y)$, which is of degree 2, we have

$$[\mathbb{Q}(u_i, u_{i+1}) : \mathbb{Q}(u_i)] = 2.$$

Let $\deg(p_i) = n$, then $[\mathbb{Q}(u_i) : \mathbb{Q}] = n$. From chapter 2 we know

$$[\mathbb{Q}(\sqrt{u_i}) : \mathbb{Q}] = [\mathbb{Q}(u_i, u_{i+1}) : \mathbb{Q}] = [\mathbb{Q}(u_i, u_{i+1}) : \mathbb{Q}(u_i)] \cdot [\mathbb{Q}(u_i) : \mathbb{Q}] = 2n$$

[1] a field extension with one generator.

Form the other side of the diagram,

$$2n = [\mathbb{Q}(\sqrt{u_i}) : \mathbb{Q}] = [\mathbb{Q}(\sqrt{u_i}) : \mathbb{Q}(u_{i+1})] \cdot [\mathbb{Q}(u_{i+1}) : \mathbb{Q}] \tag{4.5}$$

Assume $[\mathbb{Q}(u_{i+1}) : \mathbb{Q}] = t$. This implies that the degree of p_{i+1} is t. By equation (4.5), t divides $2n$. We want to show that $t = 2n$.

Assume the contrary statement i.e. $t < 2n$, so t is at most n. Now if

$$p_{i+1}(x) = x^t + c_{t-1}x^{t-1} + ... + c_0$$

then

$$p_{i+1}(u_{i+1}) = (u_{i+1})^t + c_{t-1}(u_{i+1})^{t-1} + ... + c_0 = (2 + u_i + 2\sqrt{u_i})^t + c_{t-1}(2 + u_i + 2\sqrt{u_i})^{t-1} + ... + c_0 = 0 \tag{4.6}$$

On the other hand, $A = \{1, \sqrt{u_i}, \sqrt{u_i}^2, ..., \sqrt{u_i}^{2n-1}\}$ is a basis for $\mathbb{Q}(\sqrt{u_i})$ (as a $\mathbb{Q}$ vector space). After simplifying equation 4.6, (and replacing u_i^t by lower powers if necessary) we obtain a linear combination of some elements of A which is equal to zero, and the leading term has a non zero constant coefficient, which is in contradiction with linearly in-dependency of A, so we are done.

The above argument is a constructive proof for the following theorem which gives us a recursive formula for the sequence of $\{p_i\}$ for $i \geq 2$.

Theorem 4.4.1 *for each $i \geq 4$ the minimal polynomial of u_i, p_i is given as*

$$p_i(y) = \text{res}_x(f(x, y), p_{i-1}(x)).$$

Moreover the degree of each p_i is 2^{i-2} if $i \geq 3$.

4.5 Roots of Minimal Polynomials

According to our experimental results for minimal polynomials of u_is in Maple, we see that all roots of minimal polynomials of u_is are real (we checked up to degree 512), see Table 4.1. We

i	u_i	Degree of p_i	$\sharp$ of Real Roots of p_i	Largest Root of p_i
1	1	1	1	1
2	5	1	1	5
3	11.47213595	2	2	11.47213595
4	20.24624429	4	4	20.24624429
5	31.24540965	8	8	31.24540965
6	44.42492836	16	16	44.42492836
7	59.75533396	32	32	59.75533396
8	77.21564883	64	64	77.21564883
9	96.79013589	128	128	96.79013589
10	118.4665315	256	256	118.4665315
11	142.2349978	512	512	142.2349978

Table 4.1: Experimental results for $\{p_i\}$ for $i \leq 11$

also can see the roots of minimal polynomials from p_3 to p_6 in the Figure 4.1 as an observation to see how the roots are. In this section, we will take more careful look at this observation.

Assume that $p(y) = p_i$ is the minimal polynomial of u_i. We have

$$p(y) = \mathrm{res}_x(f, p_{i-1})$$

where f is the auxiliary polynomial which we defined in previous section.
On the other hand,

$$\mathrm{res}_x(f, p_{i-1}) = C\det(\mathrm{BezoutMatrix}_x(f, p_{i-1}))$$

So the roots of $p(y)$ are the roots of determinant of a Bézout matrix. Hence, studying the roots of the determinant gives sufficient information about roots of p.
We need to take a closer look at BezoutMatrix(f, p_{i-1}). It is easy to see that BezoutMatrix(f, p_{i-1}) which we will show that by $\mathbf{B}(y)$, has the following construction:

$$\mathbf{B}(y) = \mathbf{B}_2 y^2 + \mathbf{B}_1 y + \mathbf{B}_0,$$

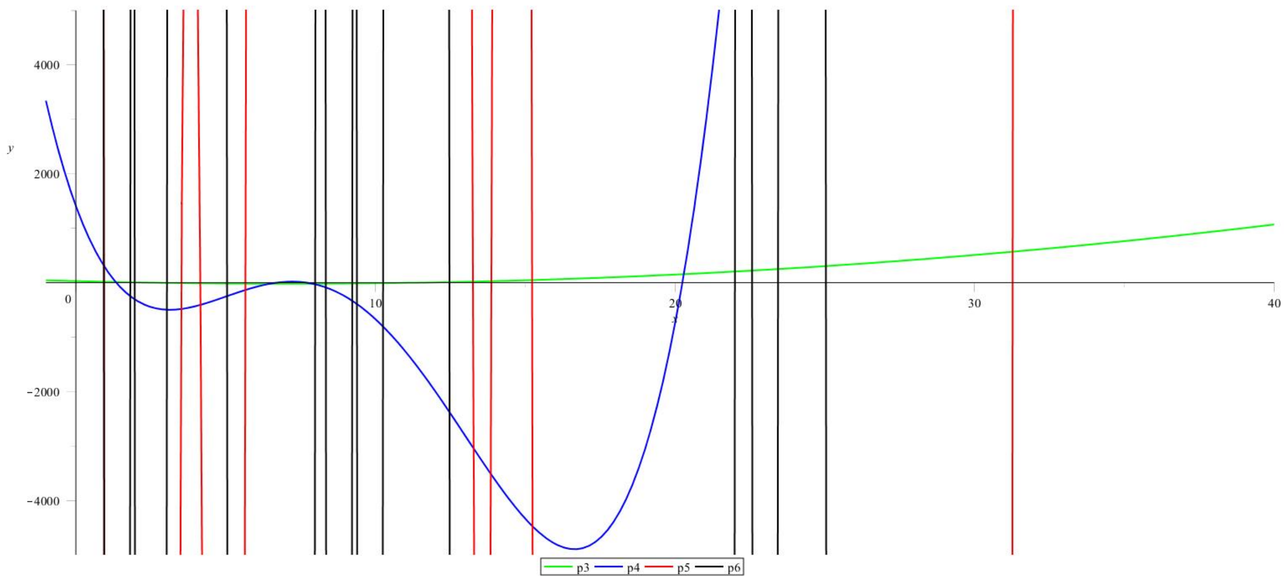

Figure 4.1: Roots of minimal polynomials of the sequence of $\{u_i\}$

where $\mathbf{B}_i$s are real numerical matrices.

If we find a linearization, $\mathbf{A}_1 y - \mathbf{A}_0$, for $\mathbf{B}(y)$, their determinant is just different by a constant coefficient. Hence, $\det(\mathbf{A}_1 y - \mathbf{A}_0)$ and $\det(\mathbf{B}(y))$ have exactly the same roots. So it is a good idea to work with a linearization of $\mathbf{B}(y)$ instead of $\mathbf{B}(y)$ itself.

The simplest linearization can be achieved when $\mathbf{B}_2$ is invertible. If that is the case, then we have the following linearization:

$$y\mathbf{I} - \mathbf{C}_p$$

where

$$\mathbf{C}_p = \begin{bmatrix} 0 & \mathbf{I} \\ -\mathbf{B}_2^{-1}\mathbf{B}_0 & -\mathbf{B}_2^{-1}\mathbf{B}_1 \end{bmatrix}$$

Theorem 4.5.1 *For each p_i the corresponding Bézout matrix has an invertible leading coefficient.*

Proof Let us fix p_i and $\mathbf{B}(y)$ be its corresponding Bézout matrix which is of the form

$$\mathbf{B}(y) = \mathbf{B}_2 y^2 + \mathbf{B}_1 y + \mathbf{B}_0$$

We want to show that $\mathbf{B}_2$ is invertible. In the previous section, we showed that the degree of p_i is 2^{i-2} and size of the Bézout matrix is $2^{(i-2)} \times 2^{(i-2)}$. Now assume $\mathbf{B}_2$ is not invertible. Then it is not a full rank matrix. This is equivalent with saying that $\mathbf{B}_2$ has a row which is a linear combination of the other rows, so with elementary row operations we can make a zero row in $\mathbf{B}_2$ and this will not change the degree of the determinant of $\mathbf{B}(y)$. On the other hand, if there is a row without y^2 in $\mathbf{B}(y)$, then

$$\deg(\det(\mathbf{B}(y)) \leq 2 \times 2^{i-2} - 1$$

which is a contradiction. So $\mathbf{B}_2$ is invertible.

The above discussion says that det(BezoutMatrix) and $\det(y\mathbf{I} - \mathbf{C}_p)$ have the same roots or equivalently roots of p_i are roots of $\det(y\mathbf{I} - \mathbf{C}_p)$ (for appropriate $\mathbf{C}_p$) which are eigenvalues of $\mathbf{C}_p$.

Up to now, we have reduced our original problem to the problem of showing eigenvalues of $\mathbf{C}_p$ are real. However, $\mathbf{C}_p$ is not a symmetric matrix. So our next step is finding another linearization which is symmetric, so that we can talk about eigenvalues or generalized eigenvalues easier.

4.6 Symmetric Linearization of the Bézout Matrix

So far, we used Bézout matrix in monomial basis for our linearization, but let us find a linearization of Bézout matrix using Lagrange basis. In this basis we can see the Bézout matrix becomes a symmetric matrix.

Let us use y_0,y_1,y_2 as nodes, then $\mathbf{B}'_i = \mathbf{B}(y_i)$ for $0 \leq i \leq 2$. Assume $\ell_i(y)$s for $0 \leq i \leq 2$ are Lagrange polynomials for y_is, see section 3.3. Then we have

$$\mathbf{B}(y) = \ell_2(y)\mathbf{B}'_2 + \ell_1(y)\mathbf{B}'_1 + \ell_0(y)\mathbf{B}'_0$$

According to [1], we have:

$$y\mathbf{C}_1 - \mathbf{C}_0 = \begin{bmatrix} (y-y_0)\mathbf{I} & 0 & 0 & -\mathbf{B}'_0 \\ 0 & (y-y_1)\mathbf{I} & 0 & -\mathbf{B}'_1 \\ 0 & 0 & (y-y_2)\mathbf{I} & -\mathbf{B}'_2 \\ \omega_0\mathbf{I} & \omega_1\mathbf{I} & \omega_2\mathbf{I} & 0 \end{bmatrix}$$

where

$$\omega_j = \prod_{k=0,k\neq j}^{2} \frac{1}{y_j - y_k}$$

as a linearization for $\mathbf{B}(y)$.

The authors also introduced a matrix $\mathbf{A}$ as follows:

$$\mathbf{A} = \begin{bmatrix} \omega_0^{-1}\mathbf{B}'_0 & 0 & 0 & 0 \\ 0 & \omega_1^{-1}\mathbf{B}'_1 & 0 & 0 \\ 0 & 0 & \omega_2^{-1}\mathbf{B}'_2 & 0 \\ 0 & 0 & 0 & -\mathbf{I} \end{bmatrix}$$

such that $(y\mathbf{C}_1 - \mathbf{C}_0)\mathbf{A}$ is the following block symmetric matrix:

$$(y\mathbf{C}_1 - \mathbf{C}_0)\mathbf{A} = \begin{bmatrix} \frac{y-y_0}{\omega_0}\mathbf{B}'_0 & 0 & 0 & \mathbf{B}'_0 \\ 0 & \frac{y-y_1}{\omega_1}\mathbf{B}'_1 & 0 & \mathbf{B}'_1 \\ 0 & 0 & \frac{y-y_2}{\omega_2}\mathbf{B}'_2 & \mathbf{B}'_2 \\ \mathbf{B}'_0 & \mathbf{B}'_1 & \mathbf{B}'_2 & 0 \end{bmatrix}$$

Having $y\mathbf{C}_1 - \mathbf{C}_0$ as linearization we can write:

$$\det(\mathbf{B}(y)) = \det(y\mathbf{C}_1 - \mathbf{C}_0)$$

$$p(y) = \det(\mathbf{B}(y)) = \det(y\mathbf{C}_1 - \mathbf{C}_0)$$

Let

$$h(y) = p(y).\det(\mathbf{A})$$

So h and p have the same roots, then

$$h(y) = p(y).\det(\mathbf{A}) = \det(y\mathbf{C}_1 - \mathbf{C}_0).\det(\mathbf{A}) = \det((y\mathbf{C}_1 - \mathbf{C}_0)\mathbf{A})$$

$(y\mathbf{C}_1 - \mathbf{C}_0)A$ is block symmetric and if the blocks are symmetric, then it will be a symmetric matrix. We can have symmetric blocks if we construct the Bézout matrix from the beginning using Lagrange basis.

So from now on, we will assume $(y\mathbf{C}_1 - \mathbf{C}_0)\mathbf{A}$ is symmetric.

$$(y\mathbf{C}_1 - \mathbf{C}_0)\mathbf{A} = \begin{bmatrix} \frac{y-y_0}{\omega_0}\mathbf{B}_0' & 0 & 0 & \mathbf{B}_0' \\ 0 & \frac{y-y_1}{\omega_1}\mathbf{B}_1' & 0 & \mathbf{B}_1' \\ 0 & 0 & \frac{y-y_2}{\omega_2}\mathbf{B}_2' & \mathbf{B}_2' \\ \mathbf{B}_0' & \mathbf{B}_1' & \mathbf{B}_2' & 0 \end{bmatrix} =$$

$$\begin{bmatrix} \frac{y}{\omega_0}\mathbf{B}_0' & 0 & 0 & \mathbf{B}_0' \\ 0 & \frac{y}{\omega_1}\mathbf{B}_1' & 0 & \mathbf{B}_1' \\ 0 & 0 & \frac{y}{\omega_2}\mathbf{B}_2' & \mathbf{B}_2' \\ \mathbf{B}_0' & \mathbf{B}_1' & \mathbf{B}_2' & 0 \end{bmatrix} - \begin{bmatrix} \frac{y_0}{\omega_0}\mathbf{B}_0' & 0 & 0 & \mathbf{B}_0' \\ 0 & \frac{y_1}{\omega_1}\mathbf{B}_1' & 0 & \mathbf{B}_1' \\ 0 & 0 & \frac{y_2}{\omega_2}\mathbf{B}_2' & \mathbf{B}_2' \\ \mathbf{B}_0' & \mathbf{B}_1' & \mathbf{B}_2' & 0 \end{bmatrix}$$

Which is equal to

$$y\begin{bmatrix} \frac{1}{\omega_0}\mathbf{B}_0' & 0 & 0 & \mathbf{B}_0' \\ 0 & \frac{1}{\omega_1}\mathbf{B}_1' & 0 & \mathbf{B}_1' \\ 0 & 0 & \frac{1}{\omega_2}\mathbf{B}_2' & \mathbf{B}_2' \\ \mathbf{B}_0' & \mathbf{B}_1' & \mathbf{B}_2' & 0 \end{bmatrix} - \begin{bmatrix} \frac{y_0}{\omega_0}\mathbf{B}_0' & 0 & 0 & \mathbf{B}_0' \\ 0 & \frac{y_1}{\omega_1}\mathbf{B}_1' & 0 & \mathbf{B}_1' \\ 0 & 0 & \frac{y_2}{\omega_2}\mathbf{B}_2' & \mathbf{B}_2' \\ \mathbf{B}_0' & \mathbf{B}_1' & \mathbf{B}_2' & 0 \end{bmatrix}$$

After simplifying we have:

$$(y\mathbf{C}_1 - \mathbf{C}_0)\mathbf{A} = y\begin{bmatrix} \hat{\mathbf{B}}_0 & 0 & 0 & \mathbf{B}_0' \\ 0 & \hat{\mathbf{B}}_1 & 0 & \mathbf{B}_1' \\ 0 & 0 & \hat{\mathbf{B}}_2 & \mathbf{B}_2' \\ \mathbf{B}_0' & \mathbf{B}_1' & \mathbf{B}_2' & 0 \end{bmatrix} - \begin{bmatrix} \tilde{\mathbf{B}}_0 & 0 & 0 & \mathbf{B}_0' \\ 0 & \tilde{\mathbf{B}}_1 & 0 & \mathbf{B}_1' \\ 0 & 0 & \tilde{\mathbf{B}}_2 & \mathbf{B}_2' \\ \mathbf{B}_0' & \mathbf{B}_1' & \mathbf{B}_2' & 0 \end{bmatrix} = y\mathbf{P} - \mathbf{Q}$$

where

$$\tilde{\mathbf{B}}_i = \frac{y_i}{\omega_i}\mathbf{B}_i'$$

and

$$\hat{\mathbf{B}}_i = \frac{1}{\omega_i}\mathbf{B}'_i$$

also **P** and **Q** are symmetric.

Now assume that λ is a generalized eigenvalue of $y\mathbf{P} - \mathbf{Q}$. Then we have $\mathbf{Q}u = \lambda\mathbf{P}u$ where u is a non-zero eigenvector.
Taking complex conjugate of both sides of $\mathbf{Q}u = \lambda\mathbf{P}u$, we have:

$$\lambda^*\mathbf{P}u^* = \mathbf{Q}u^* \tag{4.7}$$

Now we pre-multiply $\mathbf{Q}u = \lambda\mathbf{P}u$ by $(u^*)^T$ to obtain:

$$\lambda(u^*)^T\mathbf{P}u = (u^*)^T\mathbf{Q}u$$

$$\lambda(\mathbf{P}^T u^*)^T u = (\mathbf{Q}^T u^*)^T u$$

$$\lambda(\mathbf{P}u^*)^T u = (\mathbf{Q}u^*)^T u$$

$$\lambda(\mathbf{P}u^*)^T u = (\lambda^*\mathbf{P}u^*)^T u = \lambda^*(\mathbf{P}u^*)^T u$$

Thus,

$$(\lambda - \lambda^*)(\mathbf{P}u^*)^T u = 0$$

Now if $(\mathbf{P}u^*)^T u$ is non-zero, then $\lambda = \lambda^*$ which means λ is real. So we have proved the following theorem.

Theorem 4.6.1 *Any generalized eigenvalue λ of $y\boldsymbol{P} - \boldsymbol{Q}$, which has an eigenvector u such that $(u^*)^T\boldsymbol{P}u \neq 0$, is real.*

If the condition of the above theorem holds for the linearization of a Bézout matrix, it means that all roots of the corresponding minimal polynomial are real. In the case that the condition fails, it may be because of having $\mathbf{P}u = 0$. This only happens if u is an infinite eigenvalue. Since we do not care about infinite eigenvalues, this case is OK as well.

Corollary 4.6.2 *All roots of the minimal polynomial of u_k are real.*

Chapter 5

Concluding Remarks

In this dissertation, we defined a sequence. We showed that all elements of the sequence are algebraic over rational numbers. Then we used a sequence of symmetric Bézout matrix in order to get the sequence of minimal polynomials for our original sequence. Furthermore, by interpreting the set of roots of each minimal polynomial to the set of eigenvalues of the corresponding Bézout matrix, we were be able to show that all roots of the minimal polynomials of the original sequence are real.
It seems reasonable to ask what else can be done by using the ideas presented in this dissertation. We briefly state some problems which may be solved by presented methods.

5.1 Future Work

There are still open questions using the method of converting the original question to a non-linear eigenvalue problem. We present some of the interesting one below.

5.1.1 Largest Root of $\{p_i\}$

The first observation about the roots of p_is is that u_n is the largest root of p_n for each n. This looks reasonable to ask if it can be interpreted as a question related to eigenvalue problem. In fact we have no idea why u_is are largest roots at all.

5.1.2 Distribution of Zeros of $\{p_i\}$

Another interesting question is if one can find intervals such that each of them contains a root of a fixed p_i. More precisely let $p = p_i$ for some i be of degree n and $\alpha_1, \cdots \alpha_n$ are its real roots. The goal is to find (a_j, b_j) for $1 \leq j \leq n$ such that $\alpha_j \in (a_j, b_j)$. The interesting part would be the fact if one can convert it to eigenvalue problem or one of its variants.

Assume that all roots of p_i lies in the interval (a, b). One can partition the (a, b) and ask about the distribution of the roots in each partition. We did some experiments for p_i for $5 \leq i \leq 8$ which is presented in Figures 5.1, 5.2, 5.3 and Figure 5.4. According to these figures, surprisingly, we can see that the largest root which is u_i, is somehow isolated from the other roots. This can be an interesting question to work on.

5.1.3 Recursion Between Bézout Matrices

We gave a constructive method to find the sequence of minimal polynomials which uses a sequence of Bézout matrices. Each Bézout matrix is constructed from two polynomial. Thus to get a Bézout matrix one have to calculate the previous ones. Another possibility is that there may be a relation between the Bézout matrices. If one figures out the relation between the sequence of Bézout matrices, then it is feasible to construct Bézout matrices from previous ones .

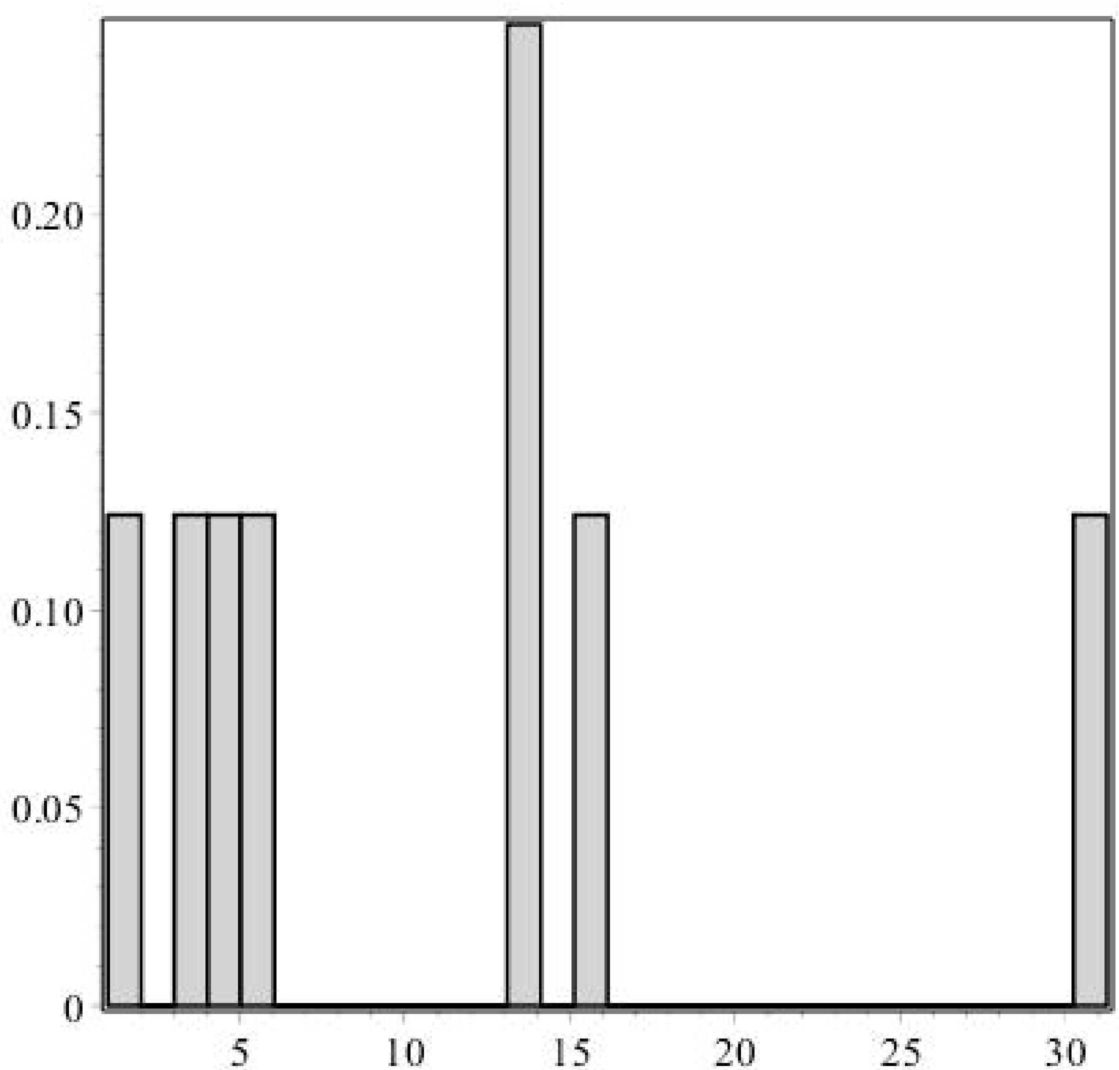

Figure 5.1: Distribution of zeros of p_5

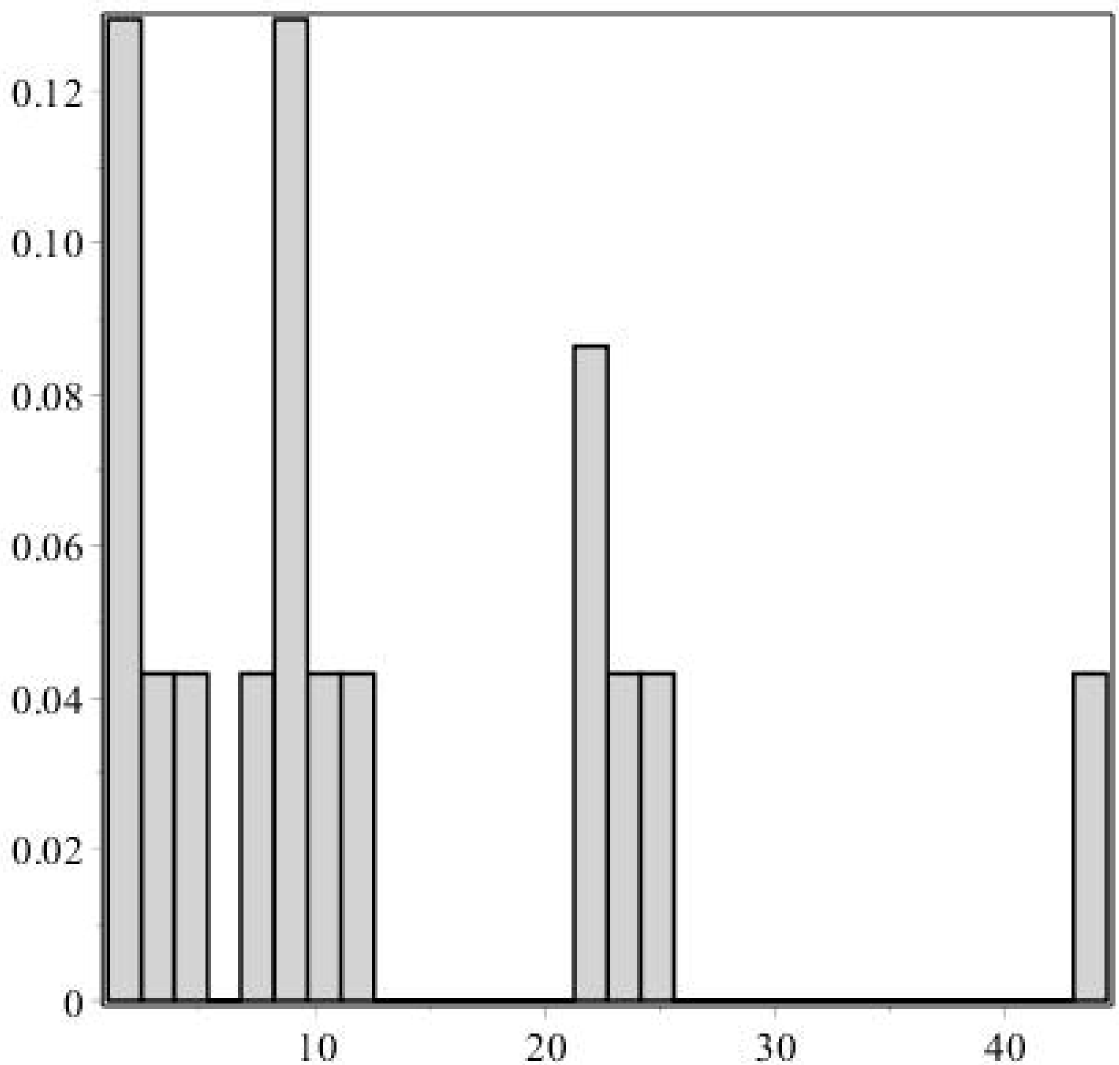

Figure 5.2: Distribution of zeros of p_6

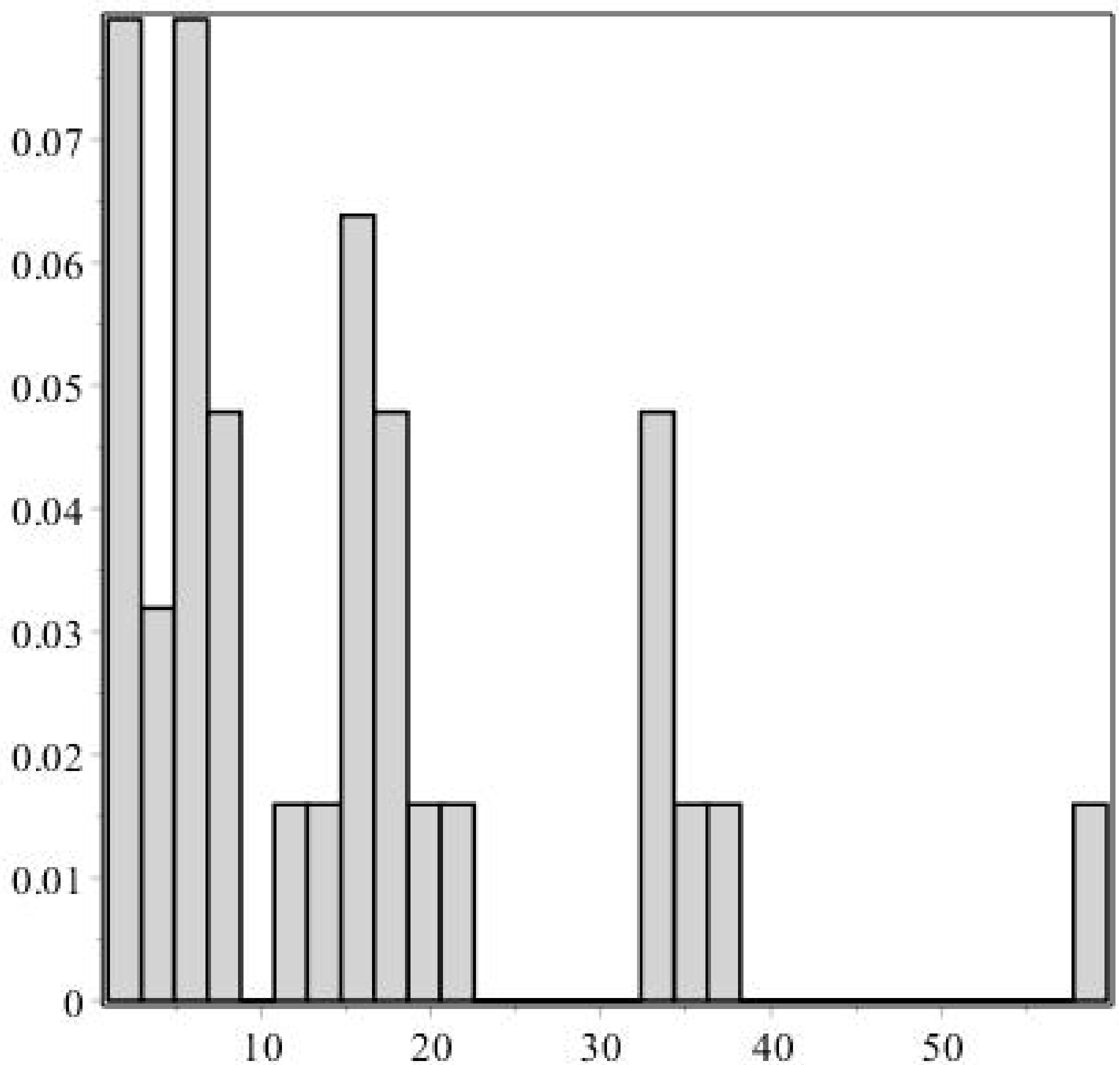

Figure 5.3: Distribution of zeros of p_7

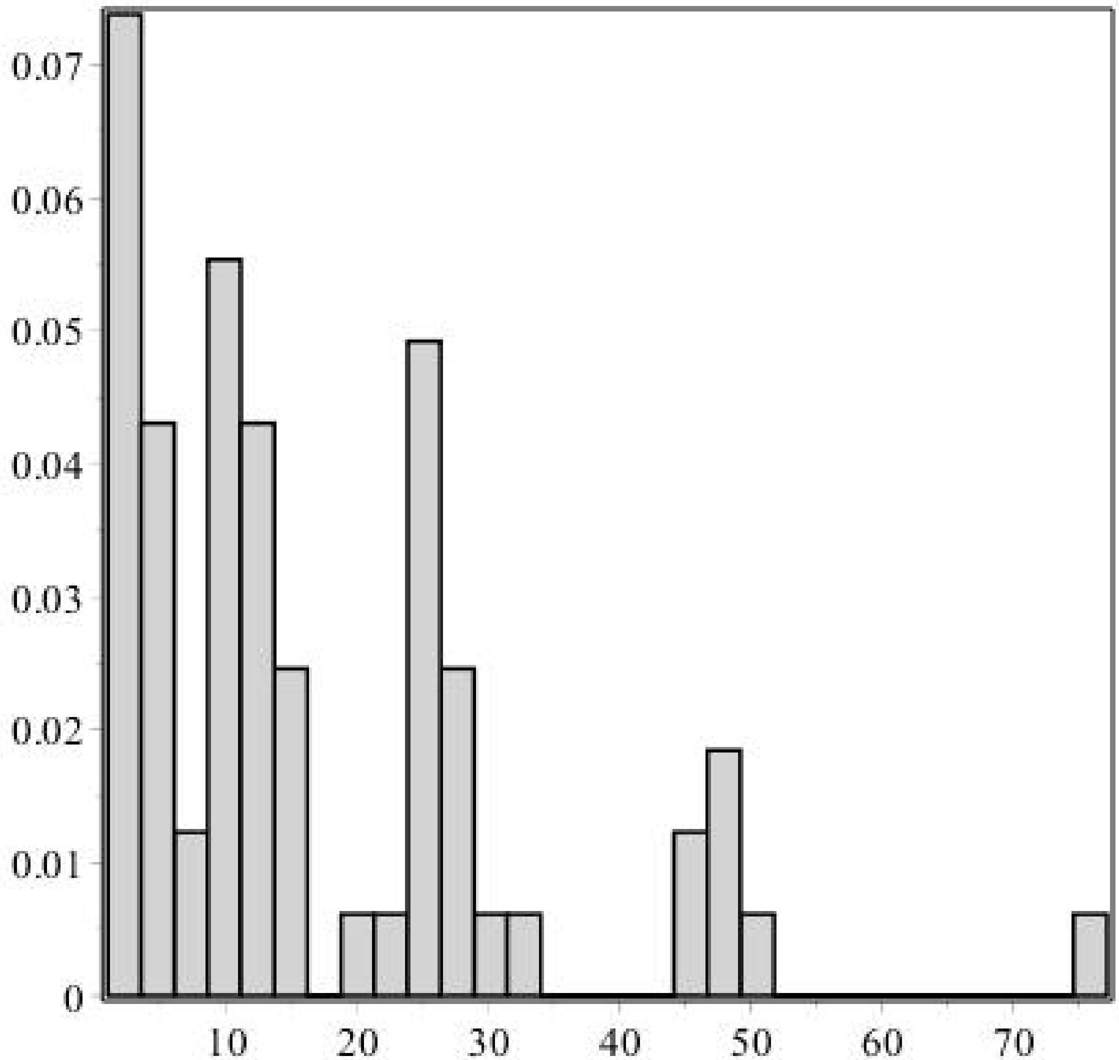

Figure 5.4: Distribution of zeros of p_8

Bibliography

[1] A. Amiraslani, R. M. Corless, and P. Lancaster, *Linearization of matrix polynomials expressed in polynomial bases*, IMA J. Numer. Anal. **29** (2009), no. 1, 141–157. MR 2470944

[2] Amirhossein Amiraslani, *New Algorithms for Matrices, Polynomials and Matrix Polynomials*, Ph.D. thesis, Western University, May 2006.

[3] R. M. Corless and J. E. Jankowski, *Variation on a theme of Euler*, To appear in SIAM Review (2016), `http://www.apmaths.uwo.ca/~rcorless/Bucket/Variations_ThemeOfEuler.pdf`.

[4] Robert M. Corless and Nicolas Fillion, *A graduate introduction to numerical methods*, Springer, New York, 2013. MR 3136200

[5] Fernando de Terán, Froilán M. Dopico, and D. Steven Mackey, *Linearizations of singular matrix polynomials and the recovery of minimal indices*, Electron. J. Linear Algebra **18** (2009), 371–402. MR 2530142

[6] David S. Dummit and Richard M. Foote, *Abstract algebra*, third ed., John Wiley & Sons, Inc., Hoboken, NJ, 2004. MR 2286236 (2007h:00003)

[7] Stephen H. Friedberg, Arnold J. Insel, and Lawrence E. Spence, *Linear algebra*, third ed., Prentice Hall, Inc., Upper Saddle River, NJ, 1997. MR 1434064

[8] K. O. Geddes, S. R. Czapor, and G. Labahn, *Algorithms for computer algebra*, Kluwer Academic Publishers, Boston, MA, 1992. MR 1256483 (96a:68049)

[9] I. Gohberg, P. Lancaster, and L. Rodman, *Matrix polynomials*, Classics in Applied Mathematics, vol. 58, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2009, Reprint of the 1982 original [MR0662418]. MR 3396732

[10] Robin Hartshorne, *Algebraic geometry*, Springer-Verlag, New York-Heidelberg, 1977, Graduate Texts in Mathematics, No. 52. MR 0463157

[11] Peter Lancaster, *Linearization of regular matrix polynomials*, Electron. J. Linear Algebra **17** (2008), 21–27. MR 2373457

[12] Patrick Morandi, *Field and Galois theory*, Graduate Texts in Mathematics, vol. 167, Springer-Verlag, New York, 1996. MR 1410264 (97i:12001)

[13] Samuel G. Moreno and Esther M. García-Caballero, *Chebyshev polynomials and nested square roots*, J. Math. Anal. Appl. **394** (2012), no. 1, 61–73. MR 2926204

[14] M. A. Nyblom, *More nested square roots of 2*, Amer. Math. Monthly **112** (2005), no. 9, 822–825. MR 2179862

[15] A Shakoori, *The Bézout matrix in the Lagrange basis*, (2004), `http://www.apmaths.uwo.ca/~rcorless/frames/PAPERS/PABV/EACA2004Azar.pdf`.

[16] Gilbert Strang, *Introduction to applied mathematics*, Wellesley-Cambridge Press, Wellesley, MA, 1986. MR 870634

[17] Pierluigi Vellucci and Alberto Maria Bersani, *Ordering of nested square roots of 2 according to gray code*, arXiv:1604.000222v1 (April 2016), no. 1.

Curriculum Vitae

Name:	Leili Rafiee Sevyeri
Post-Secondary Education and Degrees:	Western University London, Ontario, Canada 2014-2016 M.Sc.
	University of Guilan Rasht, Guilan, Iran 2009-2011 M.Sc.
	University of Guilan Rasht, Guilan, Iran 2004-2008 B.Sc.
Honours and Awards:	Distinguished Poster Award at International Symposium on Symbolic and Algebraic Computation (ISSAC). July 18-22, Waterloo, Canada, 2016

Related Work Experience:	Teaching Assistant and Research Assistant Western University London, Ontario, Canada 2014 - 2016
	Research Assistant Ontario Research Center for Computer Algebra (ORRCA) London, Ontario, Canada 2015-2016
Publications:	Poster at Computationally Assisted Mathematical Discovery and Experimental Mathematics (ACMES) Western University London, Ontario, Canada, 2016
	Poster at International Symposium on Symbolic and Algebraic Computation (ISSAC). Wilfrid Laurier University Waterloo, Ontario, Canada, 2016