CALCULATING MATRIX RANK

USING A GENERALIZATION OF THE WRONSKIAN

A Thesis

Presented to

The Faculty of the Department of Mathematics

California State University, Los Angeles

In Partial Fulfillment

of the Requirements for the Degree

Master of Science

in

Mathematics

By

Jayson Grassi

June 2016

The thesis of Jayson Grassi is approved.

Gary Brookfield, PhD., Committee Chair

Adam Fuller, PhD.

Kristin Webster, PhD.

Grant Fraser, PhD., Department Chair

California State University, Los Angeles

June 2016

ABSTRACT

Calculating Matrix Rank

Using a Generalization of the Wronskian

By

Jayson Grassi

By mapping the rows/columns of a matrix, M, to a set of polynomials, and then calculating the Wronskian of that set we can deterimine whether or not the rows/columns of M are linearly independent. However, if the rows/columns of M are linearly dependent, the Wronskian does not provide any additional insight into the exact dimension of the span of these vectors.

We call the dimension of the span of the rows/columns of a matrix M, the rank of M. In this thesis we define the rank of a two variable polynomial and show that there is a natural mapping between matrices and polynomials of two variables that preserves rank. Using this mapping, we take concepts that apply to single variable polynomials, like Wronskians, and generalize them to two variable polynomials. Finally we show how these tools can be used to calculate the rank of a matrix.

ACKNOWLEDGMENTS

My deepest thanks goes to Dr. Brookfield, without whom this thesis would have never come to fruition. He has been a fantastic advisor who has shown me how to conduct proper mathematical research. The things I've learned throughout the writing of this thesis will stay with me for the rest of my career and I am very thankful to Dr. Brookfield for this experience.

TABLE OF CONTENTS

Abstract iv

Acknowledgments v

List of Tables vii

Chapter

1. Introduction 1
2. Matrices and Determinants 3
3. Wronskians 8
4. Polynomials of Two Variables 19
5. Implications for Matrices 27

References 35

LIST OF TABLES

Table

5.1. The number of equations needed to calculate matrix rank using different methods 33

CHAPTER 1

Introduction

Given an $m \times n$ matrix over a field $\mathbb{F}$, one property that may be of interest to us is the rank of that matrix. The formal definition of matrix rank will be given in Chapter 2 along with some lemmas about matrices, determinants and differential equations that will prove useful throughout this thesis. The reader should already be aware that there are many different methods for calculating matrix rank, each with its own benefits and drawbacks. For example, if one wanted to know the exact rank of a matrix, one could first put it into row echelon form and then calculate the rank of the row equivalent matrix easily. Alternatively, one could calculate all the possible k-minors of the matrix for each possible k and deduce the rank of the matrix to be the largest value k for which there exists a nonzero k-minor. Sometimes one may simply wish to know whether or not the rank of a matrix is the largest it could possibly be for a matrix of that given size, that is to say, whether the matrix has full rank or not. For this, one may make use of the Wronskian. Introduced in 1812 by Józef Hoëne-Wroński (1776-1853) [5], the Wronskian is a type of determinant used mainly in the study of differential equations to test for linear dependence among solution sets. In Chapter 3 we define the Wronskian as well as the alternant [4], another type of determinant that could be considered a relative of the Wronskian. By mapping the rows/columns of a matrix into polynomials and then calculating the Wronskian of those polynomials one can determine whether or not the rows/columns of a matrix form a linearly dependent set, thus determining if the matrix has full rank. The Wronskian is a quick and efficient tool, so it is easy to see why we would like to generalize the Wronskian into an operator that can be used to calculate the rank of a matrix rather than just determine if it is full rank or not. In Chapter 4 of this thesis, we construct this operator in a way that reduces back to the Wronskian method when the matrix has full rank.

The concept of generalizing the Wronskian was first introduced by Ostrowski [1]. However, because our intent is to generalize the Wronskian so that we may calculate the rank of a matrix, Ostrowski's definition of the generalized Wronskian will not suffice. Instead, we must first make a connection between the Wronskian and the alternant for a set of polynomials. We see in Chapter 3 that the two share the property that they are zero if and only if the set of polynomials is linearly dependent. Then in Chapter 4 we introducing the reader to the rank of a two variable polynomial and easily generalize the alternant to work on a two variable polynomial. In this way we can construct a generalization of the Wronskian that will satisfy lemmas that mirror those from Chapter 3 about the Wronskian and the alternant. This thesis then concludes with Theorem 5.3 in Chapter 5 tying together the concepts of rank of a matrix, the rank of a polynomial in two variables, the generalized alternant, and the generalized Wronskian.

CHAPTER 2

Matrices and Determinants

We begin this thesis with a chapter covering some of the more useful properties of matrices and their determinants. As with any discussion involving the Wronskian, differential equations are an integral part of the underlying mathematics at work here, and so we also include Lemma 2.7 whose proof is rarely introduced at the undergraduate level and therefore often unknown to many graduate readers. These facts and lemmas are used throughout the remainder of this thesis, and so we provide proofs that are in the same spirit of what is to come later on. For a more in depth look at the matrial in this chapter we refer the reader to graduate texts in linear algebra and differential equations like "Advanced Linear Algebra" [6] and "Theory of Ordinary Differential Equations" [7].

All matrices are assumed to be over a field $\mathbb{F}$ of characteristic zero unless otherwise stated. This is to make use of the property that for any polynomial $f(x) \in \mathbb{F}[x]$, $f(a) = 0$ for all $a \in \mathbb{F}$ if and only if $f(x) \equiv 0$. For a matrix A, we write $\operatorname{col} A$ for the column space of A and $\operatorname{row} A$ for the row space of A.

Lemma 2.1. [2] *Let A be an $m \times n$ matrix. Then* $\dim(\operatorname{col} A) = \dim(\operatorname{row} A)$.

We define the rank of A, $\operatorname{rank} A$, to be $\dim(\operatorname{col} A) = \dim(\operatorname{row} A)$.

Theorem 2.2. [3] *Let A be an $m \times n$ matrix and $k \in \mathbb{N}$. The following are equivalent:*

(1) $\operatorname{rank} A \leq k$.

(2) $A = CB$ *for some* $m \times k$ *matrix* C *and* $k \times n$ *matrix* B.

(3) $A = C_1 + C_2 + \cdots + C_k$ *for some* $m \times n$ *rank one matrices* $C_1, C_2, \ldots, C_k$.

Proof. By definition, $\operatorname{rank} A \leq k$ if and only if there exists k column vectors $\mathbf{c}_1, \ldots, \mathbf{c}_k$ of size m such that for each column $\mathbf{v}_j$ of A we have $\mathbf{v}_j = \sum_{i=1}^{k} b_{ij}\mathbf{c}_i$ for some $b_{ij} \in \mathbb{F}$. Let C be the matrix whose columns are $\mathbf{c}_1, \ldots, \mathbf{c}_k$ and $B = \left[b_{ij}\right]$. Then $A = CB$.

Conversely, if $A = CB$ for some $m \times k$ matrix C and $k \times n$ matrix B, then $\operatorname{col} A \subseteq \operatorname{col} C$ which implies that $\operatorname{rank} A \leq \operatorname{rank} C \leq k$. So we have (1) if and only if (2).

Now (1) holds if and only if there are column vectors $\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_n$ and scalars b_{ij} such that A can be written as follows,

$$\begin{aligned} A &= [\mathbf{v}_1|\mathbf{v}_2|\cdots|\mathbf{v}_n] \\ &= \left[\sum_{i=1}^{k} b_{i1}\mathbf{c}_i \middle| \sum_{i=1}^{k} b_{i2}\mathbf{c}_i \middle| \cdots \middle| \sum_{i=1}^{k} b_{in}\mathbf{c}_i\right] \\ &= \sum_{i=1}^{k} [b_{i1}\mathbf{c}_i|b_{i2}\mathbf{c}_i|\cdots|b_{in}\mathbf{c}_i] \\ &= \sum_{i=1}^{k} C_i \end{aligned}$$

where each $C_i = [b_{i1}\mathbf{c}_i|b_{i2}\mathbf{c}_i|\cdots|b_{in}\mathbf{c}_i]$ is a $m \times n$ matrix with rank 1 since each column of C_i is a scalar multiple of $\mathbf{c}_i$. So then (1) holds if and only if (3) holds. $\square$

This theorem can be read as saying that the rank of a matrix A is the smallest number k such that A can be written as a sum of k rank one matrices. Later we use this idea to define the rank of a polynomial in two variables.

A useful property of determinants is the fact that they are multilinear functions of the columns/rows, and we will make use of this in later proofs. The following lemma formalizes this property.

Lemma 2.3. [8] *Let $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n, \mathbf{w}$ be n column vectors in $\mathbb{F}^n$. Then*

$$\det([\mathbf{v}_1|\mathbf{v}_2|\ldots|\mathbf{v}_i + \mathbf{w}|\ldots|\mathbf{v}_n]) = \det([\mathbf{v}_1|\mathbf{v}_2|\ldots|\mathbf{v}_i|\ldots|\mathbf{v}_n]) + \det([\mathbf{v}_1|\mathbf{v}_2|\ldots|\mathbf{w}|\ldots|\mathbf{v}_n])$$

The idea of solution spaces play a large part in differential equations and, because differential equations is the birthplace of the Wronskian, it is necessary that we introduce some terminology from differential equations along with a few lemmas that may seem out of place right now, but will be useful in Chapter 3 when we begin working with the Wronskian.

Definition 2.4. *We define the differentiation operator*

$$D : \mathbb{F}[x] \to \mathbb{F}[x]$$

by $D(f(x)) = f^{'}(x)$, *and use the notation* $D^n(f(x)) := D^{n-1}(D(f(x)))$.

Definition 2.5. *A differential operator* $\mathcal{P}(D) = c_n D^n + c_{n-1} D^{n-1} + \cdots + c_1 D + c_0$, *with coefficients* $c_i \in \mathbb{F}[x]$, *is a function*

$$\mathcal{P}(D) : \mathbb{F}[x] \to \mathbb{F}[x]$$

defined by $\mathcal{P}(D)(y) = c_n D^n(y) + c_{n-1} D^{n-1}(y) + \cdots + c_1 D(y) + c_0 y$ *for all* $y \in \mathbb{F}[x]$.

Lemma 2.6. *Let* $\mathcal{P}(D) = c_n(x) D^n + c_{n-1}(x) D^{n-1} + \cdots + c_1(x) D + c_0(x)$ *be a differential operator with coefficients* $c_n, c_{n-1}, \ldots, c_0 \in \mathbb{F}[x]$ *such that* $c_n(0) \neq 0$. *If* $y = f(x)x^n \in \mathbb{F}[x]$ *is a solution of the differential equation* $\mathcal{P}(D)(y) = 0$ *for some* $f \in \mathbb{F}[x]$, *then* $f = 0$.

Proof. Suppose, to the contrary, that $f \neq 0$. Let $a_k x^k$ be the lowest degree term of f with $a_k \neq 0$. Since $y = f(x)x^n$ is a solution of the given homogeneous differential equation, it is also a solution of all derivatives of this equation, in particular, it is a solution of its k^{th} derivative. Using the product rule repeatedly, we see that the k^{th} derivative of $\mathcal{P}(D)(y) = 0$ has the form

$$D^k(\mathcal{P}(D)(y)) = c_n(x) D^{n+k} y + \text{terms of lower order in } D$$

Because of the form of the solution $y = f(x)x^n$, if we apply the differential operator $D^k(\mathcal{P}(D)(y))$ and set $x = 0$ we get

$$c_n(x)D^{n+k}(f(x)x^n)\big|_{x=0} = c_n(x)D^{n+k}(a_k x^{n+k})\big|_{x=0} = (n+k)!c_n(0)a_k$$

But, with our assumptions, this expression is nonzero, which means that $y = f(x)x^n$ cannot be a solution of $\mathcal{P}(D)(y) = 0$ when f is nonzero. $\square$

Lemma 2.7. *Let*

$$\mathcal{P}(D) = c_n(x)D^n + c_{n-1}(x)D^{n-1} + \cdots + c_1(x)D + c_0(x)$$

be a differential operator with coefficients $c_n, c_{n-1}, \ldots, c_0 \in \mathbb{F}[x]$ such that $c_n \neq 0$. Then the dimension of the space of polynomial solutions of the differential equation $\mathcal{P}(D)(y) = 0$ is at most n.

Proof. First assume that $c_n(0) \neq 0$. Let $f_1 = g_1 + x^n h_1,\ f_2 = g_2 + x^n h_2,\ \ldots,\ f_{n+1} = g_{n+1} + x^n h_{n+1} \in \mathbb{F}[x]$ be polynomial solutions of $\mathcal{P}(D)(y) = 0$ with $\deg(g_i) < n$. The space of polynomials of degree less than n has dimension n, so the set $\{g_1, g_2, \ldots, g_{n+1}\}$ is linearly dependent. Thus there are $a_1, a_2, \ldots, a_{n+1} \in \mathbb{F}$, not all zero, such that $\sum\limits_{i=1}^{n+1} a_i g_i = 0$. Then

$$\sum_{i=1}^{n+1} a_i f_i = \sum_{i=1}^{n+1} a_i(g_i + x^n h_i) = \sum_{i=1}^{n+1} a_i g_i + \sum_{i=1}^{n+1} a_i x^n h_i = 0 + \left(\sum_{i=1}^{n+1} a_i h_i\right) x^n = f(x)x^n$$

is a solution of $\mathcal{P}(D)(y) = 0$, for some $f = \sum_{i=1}^{n+1} a_i h_i \in \mathbb{F}[x]$. Since $f(x)x^n$ is a solution of the differential equation, Lemma 2.6 implies that $f = 0$. Thus the polynomials $f_1, f_2, \ldots, f_{n+1}$ are linearly dependent over $\mathbb{F}$. This implies that the dimension of the space of polynomial solutions of the differential equation $\mathcal{P}(D) = 0$ is at most n.

In general, $c_n(0)$ may be zero so the above argument does not work. But since c_n is a nonzero polynomial, there is some $a \in \mathbb{F}$ such that $c_n(a) \neq 0$. Then, once all polynomials have been shifted by a, the above argument does apply. $\square$

CHAPTER 3

Wronskians

Definition 3.1. *Let $\mathbb{F}$ be a field of characteristic zero. Let $f_1, f_2, f_3, \ldots$ be a sequence of polynomials in $\mathbb{F}[x]$. For $n \in \mathbb{N}$, define $V_n \in \mathbb{F}[x_1, x_2, \ldots, x_n]$ by*

$$V_n = V(f_1, f_2, \ldots, f_n, x_1, x_2, \ldots, x_n) = \begin{vmatrix} f_1(x_1) & f_2(x_1) & \cdots & f_n(x_1) \\ f_1(x_2) & f_2(x_2) & \cdots & f_n(x_2) \\ \vdots & \vdots & \ddots & \vdots \\ f_1(x_n) & f_2(x_n) & \cdots & f_n(x_n) \end{vmatrix}$$

It will sometimes be convenient to use the following notation. We write

$$\mathbf{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{F}^n \quad \mathbf{f}_n = \{f_1, f_2, \ldots, f_n\} \quad \mathbf{f}_n(x) = \{f_1(x), f_2(x), \ldots, f_n(x)\}.$$

The matrix above, whose determinant is V_n, is called the **alternant matrix** for $f_1, f_2, \ldots, f_n$, and we denote it as $\mathbf{f}_n(\mathbf{x})$. So $V_n = V(\mathbf{f}_n, \mathbf{x}) = |\mathbf{f}_n(\mathbf{x})|$. V_n is the **alternant determinant** (or just the **alternant**) of $\mathbf{f}_n$. [4]

Example 3.2. Examples of alternant matrices include the Vandermonde matrices, named after Alexandre-Théophile Vandermonde, where $f_i(x) = x^{i-1}$. For $n = 3$ we have

$$\mathbf{f}_3(\mathbf{x}) = \begin{bmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ 1 & x_3 & x_3^2 \end{bmatrix}$$

To calculate the alternant $V(\mathbf{f}_3, \mathbf{x})$, first let $x_1 = x$, $x_2 = a_2$ and $x_3 = a_3$. Then we have,

$$V(\mathbf{f}_3, x, a_2, a_3) = \begin{vmatrix} 1 & x & x^2 \\ 1 & a_2 & a_2^2 \\ 1 & a_3 & a_3^2 \end{vmatrix}$$

Evaluating along the top row we see that this is a polynomial $f(x)$ of degree 2 with roots $x = a_2$ and $x = a_3$. Thus $f(x) = C(x - a_2)(x - a_3)$ where C is the coefficient of x^2. Which is easily seen to be $\begin{vmatrix} 1 & a_2 \\ 1 & a_3 \end{vmatrix} = (a_3 - a_2)$. Thus $f(x) = (a_3 - a_2)(x - a_2)(x - a_3)$.

Rearranging terms and making the substitutions $x = x_1$, $a_2 = x_2$ and $a_3 = x_3$, we get $V(\mathbf{f}_3, \mathbf{x}) = (x_3 - x_2)(x_3 - x_1)(x_2 - x_1)$.

The result generalizes to arbitrary Vandermonde matrices. So, for $f_i(x) = x^{i-1}$, we have

$$V(\mathbf{f}_n, \mathbf{x}) = \prod_{1 \leq i < j \leq n} (x_j - x_i).$$

Notice that the Vandermonde determinant is alternating in the entries, meaning an odd permutation of the indeterminants changes the sign, while an even permutation of the indeterminants does not change the value of the determinant. ■

Lemma 3.3. *If* $V_{n-1} \neq 0$ *and* $V_n = 0$*, then* $\mathbf{f}_n$ *is linearly dependent.*

Proof. Because $\mathbb{F}$ has characteristic zero and $V_{n-1} \neq 0$, there exist $u_1, u_2, \ldots, u_{n-1} \in \mathbb{F}$ such that $V_{n-1}(u_1, u_2, \ldots, u_{n-1}) \in \mathbb{F}$ is nonzero. Expanding the determinant for V_n along the bottom row, setting $x_1 = u_1, x_2 = u_2, \ldots, x_{n-1} = u_{n-1}$ and $x_n = x$, we get

$$0 = V_n = a_1 f_1(x) + a_2 f_2(x) + \cdots + a_n f_n(x)$$

where $a_1, a_2, \ldots, a_n \in \mathbb{F}$ are $(n-1) \times (n-1)$-minors of the matrix $\mathbf{f}_n(\mathbf{x})$. Since $a_n = V_{n-1}(u_1, u_2, \ldots, u_{n-1}) \neq 0$, the linear combination is nontrivial and $\mathbf{f}_n$ is linearly dependent. □

Lemma 3.4. $\mathbf{f}_n$ *is linearly dependent if and only if* $V(\mathbf{f}_n, \mathbf{x}) = 0$.

Proof. Suppose that $\{f_1, f_2, \ldots, f_n\}$ is linearly dependent. Without loss of generality we can assume that f_1 is a linear combination of $f_2, f_3, \ldots, f_n$. This implies that the first column of the matrix defining V_n is a linear combination of the other columns. Thus V_n, the determinant of this matrix, is zero.

Now suppose that $V_n = 0$. If $V_1 = f_1(x_1)$ is zero, then the claim is obviously true. Otherwise we have $V_1 \neq 0$ and $V_n = 0$, so there must be some $1 < k \leq n$ such that $V_{k-1} \neq 0$ and $V_k = 0$. By Lemma 3.3, $\{f_1, f_2, \ldots, f_k\}$ is linearly dependent. This implies that $\{f_1, f_2, \ldots, f_n\}$ is also linearly dependent. $\square$

Lemma 3.5. *$\mathbf{f}_n$ is linearly independent if and only if there is $\mathbf{u} = (u_1, u_2, \ldots, u_n) \in \mathbb{F}^n$ such that the matrix $M = \mathbf{f}_n(\mathbf{u})$ is invertible.*

Proof. Let $\mathbf{u} \in \mathbb{F}^n$ be such that $M = \mathbf{f}_n(\mathbf{u})$ is invertible. Let $a_1, a_2, \ldots, a_n \in \mathbb{F}$ such that

$$a_1 f_1(x) + a_2 f_2(x) + \cdots + a_n f_n(x) = 0$$

for all x. Then we have the system of equations

$$\begin{aligned}
&a_1 f_1(u_1) + a_2 f_2(u_1) + \cdots + a_n f_n(u_1) = 0 \\
&a_1 f_1(u_2) + a_2 f_2(u_2) + \cdots + a_n f_n(u_2) = 0 \\
&\vdots \\
&a_1 f_1(u_n) + a_2 f_2(u_n) + \cdots + a_n f_n(u_n) = 0
\end{aligned}$$

which corresponds to the matrix equation

$$\begin{bmatrix} f_1(u_1) & f_2(u_1) & \cdots & f_n(u_1) \\ f_1(u_2) & f_2(u_2) & \cdots & f_n(u_2) \\ \vdots & \vdots & \ddots & \vdots \\ f_1(u_n) & f_2(u_n) & \cdots & f_n(u_n) \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Because the matrix $\mathbf{f}_n(\mathbf{u})$ is invertible we must have $a_i = 0$ for all $i = 1, 2, \ldots, n$, and thus $\mathbf{f}_n$ is linearly independent.

Now assume that for all $\mathbf{u} \in \mathbb{F}^n$ the matrix $\mathbf{f}_n(\mathbf{u})$ is singular. We show, by induction on n that the set of polynomials $\mathbf{f}_n$ is linearly dependent.

Let $n = 1$. Then the matrix $\mathbf{f}_n(\mathbf{u})$ is just $[f_1(u_1)]$ which is singular for all $u_1 \in \mathbb{F}$ only if $f_1 = 0$, which, of course, means that $\mathbf{f}_n$ is linearly dependent. Now we assume that the lemma holds for $n-1$ polynomials and consider the set $\mathbf{f}_n = \{f_1, f_2, \ldots, f_n\}$. Recall $\mathbf{f}_{n-1}$ denotes the set $\{f_1, f_2, \ldots, f_{n-1}\}$. We may assume that there exists an element $\mathbf{u}_{n-1} = (u_1, u_2, \ldots, u_{n-1}) \in \mathbb{F}^{n-1}$ such that the matrix $\mathbf{f}_{n-1}(\mathbf{u}_{n-1})$ is nonsingular, else by the inductive hypothesis $\mathbf{f}_{n-1}$ is linearly dependent and then so is $\mathbf{f}_n$. Consider the matrix $\mathbf{f}_n(u_1, u_2, \ldots, u_{n-1}, x)$. We know, by assumption, that for all $x \in \mathbb{F}$

$$\begin{vmatrix} f_1(u_1) & f_2(u_1) & \cdots & f_n(u_1) \\ f_1(u_2) & f_2(u_2) & \cdots & f_n(u_2) \\ \vdots & \vdots & \ddots & \vdots \\ f_1(u_{n-1}) & f_2(u_{n-1}) & \cdots & f_n(u_{n-1}) \\ f_1(x) & f_2(x) & \cdots & f_n(x) \end{vmatrix} = 0$$

Expanding the determinant along the bottom row we get the nontrivial linear combination $c_1 f_1(x) + c_2 f_2(x) + \cdots + c_n f_n(x) = 0$, where $c_n = |\mathbf{f}_{n-1}(\mathbf{u}_{n-1})| \neq 0$. Therefore $\mathbf{f}_n$ is linearly dependent. $\square$

Definition 3.6. [5] *The **Wronskian** of $f_1, f_2, \ldots, f_n \in \mathbb{F}[x]$ is defined by*

$$W_n = W(\mathbf{f}_n, x) = \begin{vmatrix} f_1(x) & f_1'(x) & \cdots & f_1^{(n)}(x) \\ f_2(x) & f_2'(x) & \cdots & f_2^{(n)}(x) \\ \vdots & \vdots & \ddots & \vdots \\ f_n(x) & f_n'(x) & \cdots & f_n^{(n)}(x) \end{vmatrix}$$

Example 3.7. Let $f_1(x) = x^3,\ f_2(x) = x^2,\ f_3(x) = x$. Then

$$\begin{aligned} W(\mathbf{f}_3, x) &= \begin{vmatrix} x^3 & x^2 & x \\ 3x^2 & 2x & 1 \\ 6x & 2 & 0 \end{vmatrix} \\ &= 0 \begin{vmatrix} x^3 & x^2 \\ 3x^2 & 2x \end{vmatrix} - 1 \begin{vmatrix} x^3 & x^2 \\ 6x & 2 \end{vmatrix} + x \begin{vmatrix} 3x^2 & 2x \\ 6x & 2 \end{vmatrix} \\ &= -(2x^3 - 6x^3) + x(6x^2 - 12x^2) \\ &= -2x^3 \end{aligned}$$

$\blacksquare$

Notice that $W_n \in \mathbb{F}[x]$, and when each $f_i(x)$ is a monomial, as in the example above, then $W(\mathbf{f}_n, x)$ is also a monomial.

Lemma 3.8. *Let f and g be polynomials over $\mathbb{R}$. Then $\{f, g\}$ is linearly dependent over $\mathbb{R}$ if and only if $W(f, g) = 0$.*

Proof. If $\{f, g\}$ is linearly dependant, then without loss of generality, $g = cf$ for some constant $c \in \mathbb{R}$. Then $Dg = cDf$ and so $W(f, g) = 0$.

The converse is obviously true if $f = g = 0$, so we suppose that $W(f, g) = 0$ and $f \neq 0$. Since f is nonzero and has at most finitely many roots, there is an open interval I of the real line on which f is never zero. Then g/f is a real differentiable function on I. The derivative of g/f on I is

$$\frac{d}{dx}\left(\frac{g}{f}\right) = \frac{fDg - gDf}{f^2} = \frac{W(f, g)}{f^2} = 0.$$

This implies that g/f is constant on I, or, equivalently, there is a constant $c \in \mathbb{R}$ such that $cf(x) - g(x) = 0$ for all $x \in I$. Since $cf(x) - g(x)$ is a polynomial that is identically zero on an open interval $I \subset \mathbb{R}$ we must have $cf(x) - g(x) = 0$ for all $x \in \mathbb{R}$. Which implies that $g = cf$ and $\{f, g\}$ is linearly dependant. $\square$

The condition that the functions in Lemma 3.8 be polynomials has been generalized to sets of analytic functions on $\mathbb{R}$ and $\mathbb{C}$, however it is known that there are some examples of linearly independent functions, which are not analytic, whose Wronskian is nonzero. The most famous example of this was first provided by Peano [9] who observed that the functions $f(x) = x$ and $g(x) = x|x|$ defined on $\mathbb{R}$ are linearly independent, but their Wronskian vanishes identically. Later Bôcher[12] proved more generally that there exist families of infinitely differentiable real functions sharing the same property.

The following lemma is a generalization of the previous result into the realm of fields of characteristic zero.

Lemma 3.9. [10] *A set of polynomials* $\{f_1, f_2, \ldots, f_n\} \subseteq \mathbb{F}[x]$ *is linearly dependent if and only if* $W_n = 0$.

Proof. Let $\{f_1, f_2, \ldots, f_n\}$ be a linearly dependent set of polynomials in $\mathbb{F}[x]$. Then there exists $c_1, c_2, \ldots, c_2 \in \mathbb{F}$, not all zero, such that $c_1 f_1 + c_2 f_2 + \cdots + c_n f_n = 0$. Because the derivative is a linear operator we have $c_1 f_1^{(i)} + c_2 f_2^{(i)} + \cdots + c_n f_n^{(i)} = 0$ for all $i \in \mathbb{N}$. So the columns of the Wronskian matrix are linearly dependent which implies $W(\mathbf{f}_n, x) = 0$.

Now assume that $W(\mathbf{f}_n, x) = 0$. If $f_i(x) = 0$ for all $i = 1, 2, \ldots, n$ then the claim is trivially true. So we may assume, without loss of generality, that $f_1(x)$ is not identically zero. So $W(f_1) \neq 0$. Then there exists an integer $m \leq n$ such that $W_{m-1} \neq 0$ and

$$W(f_1, f_2, \ldots, f_m) = \begin{vmatrix} f_1(x) & f_2(x) & \cdots & f_m(x) \\ f_1'(x) & f_2'(x) & \cdots & f_m'(x) \\ \vdots & \vdots & \ddots & \vdots \\ f_1^{(m-1)}(x) & f_2^{(m-1)}(x) & \cdots & f_m^{(m-1)}(x) \end{vmatrix} = 0$$

Evaluating along the last column we get

$$G_{m-1}(x) f_m^{(m-1)} + G_{m-2}(x) f_m^{(m-2)} + \cdots + G_0(x) f_m = 0$$

where each $G_i(x) \in \mathbb{F}[x]$ and $G_{m-1}(x) = W_{m-1} \neq 0$. So f_m is a solution of the $(m-1)$-order linear differential equation

$$G_{m-1}(x) y^{(m-1)} + G_{m-2}(x) y^{(m-2)} + \cdots + G_0(x) y = 0 \tag{3.1}$$

For each $i = 1, 2, \ldots, m-1$, $f_i(x)$ is a solution to (3.1). Because (3.1) is a linear differential equation with coefficients in $\mathbb{F}[x]$ and order $(m-1)$ we know, by Lemma 2.7, that it has at most $m-1$ linearly independent solutions. Thus $\{f_1, f_2, \ldots, f_m\}$ is linearly dependent, and therefore so is $\{f_1, f_2, \ldots, f_n\}$. $\square$

Example 3.10. Let us employ Lemma 3.9 to verify the linear independence of the following polynomials in $\mathbb{C}[x]$,

$$f_1(x) = 2x^2 + (2i-1)x + 1, \qquad f_2(x) = x^2 + ix - 1, \qquad f_3(x) = 2x - i$$

We have

$$\begin{aligned} W(f_1, f_2, f_3) &= \begin{vmatrix} (2x^2 + (2i-1)x + 1) & (x^2 + ix - 1) & (2x - i) \\ (4x + 2i - 1) & 2x + i & 2 \\ 4 & 2 & 0 \end{vmatrix} \\ &= 4 \begin{vmatrix} (x^2 + ix - 1) & (2x - i) \\ 2x + i & 2 \end{vmatrix} - 2 \begin{vmatrix} (2x^2 + (2i-1)x + 1) & (2x - i) \\ (4x + 2i - 1) & 2 \end{vmatrix} \\ &= -12 + 2i \neq 0 \end{aligned}$$

Thus, by Lemma 3.9, the three polynomials are linearly independent. ■

Lemma 3.11. *Let $f_1, f_2, \ldots, f_n \in \mathbb{F}[x]$ be polynomials of degree $\leq m$, with $m \geq n-1$. Then*

$$\deg(W_n) \leq mn - n^2 + n$$

Proof. Let $f_i(x) = \sum_{k=0}^{m} a_{i,k} x^k$ for some $a_{i,k} \in \mathbb{F}$. Then

$$\begin{aligned} W(\mathbf{f}_n) &= W\left(\sum_{k=0}^{m} a_{1,k} x^k, \sum_{k=0}^{m} a_{2,k} x^k, \ldots, \sum_{k=0}^{m} a_{n,k} x^k\right) \\ &= \sum_{k_1=0}^{m} \sum_{k_2=0}^{m} \cdots \sum_{k_n=0}^{m} W\left(a_{1,k_1} x^{k_1}, a_{2,k_2} x^{k_2}, \ldots, a_{n,k_n} x^{k_n}\right) \end{aligned}$$

By Lemma 3.9, $W\left(a_{1,k_1}x^{k_1}, a_{2,k_2}x^{k_2}, \ldots, a_{n,k_n}x^{k_n}\right) = 0$ whenever $k_i = k_j$ for some $i \neq j$. This implies that $\deg\left(W_n\right) \leq \deg W\left(x^m, x^{m-1}, \ldots, x^{m-(n-1)}\right)$

$$= \deg \begin{vmatrix} x^m & x^{m-1} & \cdots & x^{m-(n-1)} \\ (m)_1 x^{m-1} & (m-1)_1 x^{m-2} & \cdots & (m-n+1)_1 x^{m-n} \\ \vdots & \vdots & \ddots & \vdots \\ (m)_{n-1} x^{m-(n-1)} & (m-1)_{n-1} x^{m-n} & \cdots & (m-n+1)_{n-1} x^{m-2(n-1)} \end{vmatrix}$$

$$= m + (m-2) + (m-4) + \cdots + (m - 2(n-1))$$

$$= \sum_{i=0}^{n-1} m - 2i = mn - n^2 + n$$

The coefficients in the above determinant are falling factorials

$$(m)_k = m(m-1)\cdots(m-(k-1)) \qquad \square$$

Note that in the previous theorem, if $m < n-1$, the polynomials are necessarily linearly dependent and therefore $\deg(W_n) = 0$.

If we apply Lemma 3.11 to Example 3.10 where $n = 3$ and $m = 2$ we get $\deg\left(W(f_1, f_2, f_3)\right) \leq 2(3) - 3^2 + 3 = 0$ which is verified by our calculation that $W(f_1, f_2, f_3) = -12 + 2i$.

We now make a connection between the Wronskian and the alternant that is rarely observed in undergraduate or graduate courses. This is a connection that we preserve when crafting analogues of the alternant and the Wronskian for two variable polynomials in Chapter 4.

Define the function $\delta(\mathbf{x}) = \prod_{j<k}(x_k - x_j)$, and $\delta(x_1) = 1$. So, for example, if $\mathbf{x} = (1, 2, \ldots, n) \in \mathbb{R}^n$ then we have

$$\delta(\mathbf{x}) = \delta(1, 2, \ldots, n) = \prod_{1 \leq j < k \leq n} (k - j) = \prod_{k=1}^{n-1} k!$$

Lemma 3.12. *Let* $f_1, f_2, \ldots, f_n \in \mathbb{F}[x]$*,* $W_n = W(\mathbf{f}_n)$ *and* $V_n = V(\mathbf{f}_n)$ *be as in Definition* 3.6 *and Definition* 3.1 *respectively. Then*

$$W_n = \left.\frac{\delta(1,2,\ldots,n)V_n}{\delta(\mathbf{x})}\right|_{x_1=x_2=\cdots=x_n=x}$$

Proof. Let $x_i = x + y_i$, where x and y_i are indeterminants in $\mathbb{F}$. Then $\delta(\mathbf{x}) = \delta(\mathbf{y})$, and expressing f_j as a Taylor series we get $f_j(x_i) = \sum\limits_{k=0}^{\infty} \frac{1}{k!} f_j^{(k)}(x) y_i^k$. So then,

$$\frac{V_n}{\delta(\mathbf{x})} = \frac{1}{\delta(\mathbf{y})} \begin{vmatrix} \sum \frac{1}{k!} f_1^{(k)}(x) y_1^k & \sum \frac{1}{k!} f_2^{(k)}(x) y_1^k & \cdots & \sum \frac{1}{k!} f_n^{(k)}(x) y_1^k \\ \sum \frac{1}{k!} f_1^{(k)}(x) y_2^k & \sum \frac{1}{k!} f_2^{(k)}(x) y_2^k & \cdots & \sum \frac{1}{k!} f_n^{(k)}(x) y_2^k \\ \vdots & \vdots & \ddots & \vdots \\ \sum \frac{1}{k!} f_1^{(k)}(x) y_n^k & \sum \frac{1}{k!} f_2^{(k)}(x) y_n^k & \cdots & \sum \frac{1}{k!} f_n^{(k)}(x) y_n^k \end{vmatrix}$$

Making use of Lemma 2.3, which describes the multilinearity of determinants, we get

$$\frac{V_n}{\delta(\mathbf{x})} = \frac{1}{\delta(\mathbf{y})} \sum_{m=1}^{\infty} \begin{vmatrix} \frac{1}{m!} f_1^{(m)}(x) y_1^m & \frac{1}{m!} f_2^{(m)}(x) y_1^m & \cdots & \frac{1}{m!} f_n^{(m)}(x) y_1^m \\ \sum \frac{1}{k!} f_1^{(k)}(x) y_2^k & \sum \frac{1}{k!} f_2^{(k)}(x) y_2^k & \cdots & \sum \frac{1}{k!} f_n^{(k)}(x) y_2^k \\ \vdots & \vdots & \ddots & \vdots \\ \sum \frac{1}{k!} f_1^{(k)}(x) y_n^k & \sum \frac{1}{k!} f_2^{(k)}(x) y_n^k & \cdots & \sum \frac{1}{k!} f_n^{(k)}(x) y_n^k \end{vmatrix}$$

Let $x_1 = x$, so $y_1 = 0$. Then

$$\left.\frac{V_n}{\delta(\mathbf{x})}\right|_{x_1=x} = \frac{1}{\delta(y_2,\ldots,y_n) \prod\limits_{j=2}^{n} y_j} \begin{vmatrix} f_1(x) & f_2(x) & \cdots & f_n(x) \\ \sum \frac{1}{k!} f_1^{(k)}(x) y_2^k & \sum \frac{1}{k!} f_2^{(k)}(x) y_2^k & \cdots & \sum \frac{1}{k!} f_n^{(k)}(x) y_2^k \\ \vdots & \vdots & \ddots & \vdots \\ \sum \frac{1}{k!} f_1^{(k)}(x) y_n^k & \sum \frac{1}{k!} f_2^{(k)}(x) y_n^k & \cdots & \sum \frac{1}{k!} f_n^{(k)}(x) y_n^k \end{vmatrix}$$

Distributing $\frac{1}{y_2}$ from the product in the denominator to the second row, and again making use of the multilinearity of determinants we get,

$$\left.\frac{V_n}{\delta(\mathbf{x})}\right|_{x_1=x} = \frac{1}{\delta(y_2,\ldots,y_n) \prod_{j=3}^{n} y_j} \sum_{m=0}^{\infty} \begin{vmatrix} f_1(x) & \cdots & f_n(x) \\ \frac{1}{m!} f_1^{(m)}(x) y_2^{m-1} & \cdots & \frac{1}{m!} f_n^{(m)}(x) y_2^{m-1} \\ \sum \frac{1}{k!} f_1^{(k)}(x) y_3^k & \cdots & \sum \frac{1}{k!} f_n^{(k)}(x) y_3^k \\ \vdots & \ddots & \vdots \\ \sum \frac{1}{k!} f_1^{(k)}(x) y_n^k & \cdots & \sum \frac{1}{k!} f_n^{(k)}(x) y_n^k \end{vmatrix}$$

Notice that for the $m = 0$ term of the sum, the determinant vanishes because row 1 and row 2 are linearly dependent. So this becomes a sum from $m = 1$ to ∞. Next let

$x_2 = x$ so $y_2 = 0$. Then we have,

$$\frac{V_n}{\delta(\mathbf{x})}\bigg|_{\substack{x_1=x\\x_2=x}} = \frac{1}{\delta(y_3, y_4, \ldots, y_n)\prod_{j=3}^{n}(y_j)^2}\begin{vmatrix} f_1(x) & \cdots & f_n(x) \\ f_1'(x) & \cdots & f_n'(x) \\ \sum \frac{1}{k!}f_1^{(k)}(x)y_3^k & \cdots & \sum \frac{1}{k!}f_n^{(k)}(x)y_3^k \\ \vdots & \ddots & \vdots \\ \sum \frac{1}{k!}f_1^{(k)}(x)y_n^k & \cdots & \sum \frac{1}{k!}f_n^{(k)}(x)y_n^k \end{vmatrix}$$

Distributing $\frac{1}{(y_3)^2}$ from the product in the denominator to the third row, and using multilinearity we get,

$$\frac{V_n}{\delta(\mathbf{x})}\bigg|_{\substack{x_1=x\\x_2=x}} = \frac{1}{\delta(y_3, y_4, \ldots, y_n)\prod_{j=4}^{n}(y_j)^2}\sum_{m=0}^{\infty}\begin{vmatrix} f_1(x) & \cdots & f_n(x) \\ f_1'(x) & \cdots & f_n'(x) \\ \frac{1}{m!}f_1^{(m)}(x)y_3^{m-2} & \cdots & \frac{1}{m!}f_n^{(m)}(x)y_3^{m-2} \\ \vdots & \ddots & \vdots \\ \sum \frac{1}{k!}f_1^{(k)}(x)y_n^k & \cdots & \sum \frac{1}{k!}f_n^{(k)}(x)y_n^k \end{vmatrix}$$

Again notice that for the $m = 0$ and $m = 1$ terms of the sum, the determinant vanishes by linear dependence. So this becomes a sum from $m = 2$ to ∞.

Continuing this process, setting each $x_i = 0$ and then distributing $\dfrac{1}{(y_i)^{i-1}}$ into the i^{th} row of the determinant we get,

$$\frac{V_n}{\delta(\mathbf{x})}\bigg|_{x_1=x_2=\cdots=x} = \begin{vmatrix} f_1(x) & \cdots & f_n(x) \\ f_1'(x) & \cdots & f_n'(x) \\ \frac{1}{2!}f_1''(x) & \cdots & \frac{1}{2!}f_n''(x) \\ \vdots & \ddots & \vdots \\ \frac{1}{n-1!}f_1^{(n-1)}(x) & \cdots & \frac{1}{n-1!}f_n^{(n-1)}(x) \end{vmatrix}$$

$$= \frac{1}{\delta(1, 2, \ldots, n)}\begin{vmatrix} f_1(x) & \cdots & f_n(x) \\ f_1'(x) & \cdots & f_n'(x) \\ f_1''(x) & \cdots & f_n''(x) \\ \vdots & \ddots & \vdots \\ f_1^{(n-1)}(x) & \cdots & f_n^{(n-1)}(x) \end{vmatrix}$$

$$= \frac{1}{\delta(1, 2, \ldots, n)}W(f_1, f_2, \ldots, f_n) \qquad \square$$

Example 3.13. Let $f_1 = x^3,\ f_2 = x^2,\ f_3 = x$. Then

$$V_3 = \begin{vmatrix} x_1^3 & x_1^2 & x_1 \\ x_2^3 & x_2^2 & x_2 \\ x_3^3 & x_3^2 & x_3 \end{vmatrix}$$

Using the same trick we used in Example 2.2, we let $x_1 = x,\ x_2 = a_2$ and $x_3 = a_3$ getting a third degree polynomial

$$f(x) = \begin{vmatrix} x^3 & x^2 & x \\ a_2^3 & a_2^2 & a_2 \\ a_3^3 & a_3^2 & a_3 \end{vmatrix}$$

which has roots $x = 0,\ x = a_2$ and $x = a_3$ and leading coefficient $C = (a_2^2a_3 - a_2a_3^2) = a_2a_3(a_2 - a_3)$. So $f(x) = a_2a_3(a_2 - a_3)(x - 0)(x - a_2)(x - a_3)$. A quick flip of the terms $(a_2 - a_3),\ (x - a_2),$ and $(x - a_3)$ in order to better align with the terms of $\delta(\mathbf{x})$, and reversing the substitions we made gives us

$$V_3 = -x_3x_2x_1(x_3 - x_2)(x_3 - x_1)(x_2 - x_1)$$

$$= -x_1x_2x_3\delta(\mathbf{x})$$

Now we evaluate

$$\left.\frac{\delta(1,2,3)V_3}{\delta(\mathbf{x})}\right|_{x_1=\cdots=x} = -2x^3$$

$$= W(f_1, f_2, f_3)$$

as we saw in Example 2.7. ■

CHAPTER 4

Polynomials of Two Variables

We now take the ideas from Chapter 2 and Chapter 3, and introduce analogues of them for the ring of polynomials $\mathbb{F}[x, y]$. Let $\mathbb{F}$ be a field of characteristic zero and $F \in \mathbb{F}[x, y]$ a polynomial in x and y. We begin by defining the rank of F in a way that bears a resemblance to Theorem 2.2. Then we create analogues of V_n and W_n that apply to $\mathbb{F}[x, y]$.

Definition 4.1. *The* ***rank*** *of* $F \in \mathbb{F}[x, y]$*, denoted* $\operatorname{rank} F$*, is the least* $n \in \mathbb{N}$ *such that*

$$F(x, y) = g_1(x)h_1(y) + g_2(x)h_2(y) + \cdots + g_n(x)h_n(y) \tag{4.1}$$

for some $g_1, g_2, \ldots, g_n \in \mathbb{F}[x]$ *and* $h_1, h_2, \ldots, h_n \in \mathbb{F}[y]$*. If* $F = 0$*, then we define* $\operatorname{rank} F = 0$.

This is a particular application of the more general definition of tensor rank. [11]

Notice that if F has y-degree k (written $\deg_y F = k$), then

$$F(x, y) = g_0(x) + g_1(x)y + g_2(x)y^2 + \cdots + g_k(x)y^k$$

for some $g_0, g_1, \ldots, g_k \in \mathbb{F}[x]$, so $\operatorname{rank} F \le k + 1$. Since a similar argument works for the x-degree of F, we have $\operatorname{rank} F \le \min(\deg_x F, \deg_y F) + 1$.

Example 4.2. Let $F(x, y) = 1+x+y+xy$. Then $\operatorname{rank}(F) \le 4$ since each term in F is already of the form $g_i(x)h_i(y)$. Furthermore, we can rewrite $F(x, y) = (1+y)+(1+y)x$ so that

$$\operatorname{rank}(F) \le \min\{x\text{-degree}, y\text{-degree}\} + 1.$$

In fact, for this example $F(x, y) = (1 + y)(1 + x)$, so $\operatorname{rank}(F) = 1$. ■

The following lemma formalizes the concepts in the previous example.

Lemma 4.3. *Suppose that $F(x,y)$ has the form* (4.1) *for some $g_1, g_2, \dots g_n \in \mathbb{F}[x]$ and $h_1, h_2, \dots, h_n \in \mathbb{F}[y]$. If either $\{g_1, g_2, \dots, g_n\}$ or $\{h_1, h_2, \dots, h_n\}$ is linearly dependent, then $F(x,y)$ can be written in the form of* (4.1) *with $n-1$ terms, in particular,* $\operatorname{rank} F < n$.

Proof. Without loss of generality, suppose that $g_n = a_1 g_1 + \cdots + a_{n-1} g_{n-1}$ for some $a_1, a_2, \dots, a_{n-1} \in \mathbb{F}$. Then

$$F(x,y) = g_1(x)h_1^*(y) + g_2(x)h_2^*(y) + \cdots + g_{n-1}(x)h_{n-1}^*(y)$$

where $h_k^*(y) = h_k(y) + a_k h_n(y)$ for $k = 1, 2, \dots, n-1$. $\square$

For a fixed $u \in \mathbb{F}$, $F(u,y)$ is an element of $\mathbb{F}[y]$ and so the subset of $\mathbb{F}[y]$ spanned by all such polynomials, namely $\operatorname{span}\{F(u,y) \mid u \in \mathbb{F}\}$, is a subspace of $\mathbb{F}[y]$. Similarly, $\operatorname{span}\{F(x,v) \mid v \in \mathbb{F}\}$ is a subspace of $\mathbb{F}[x]$. A connection can easily be drawn between this observation and the notions of the column space and row space of a matrix. We formalize this connection with the following lemma which serves as an analogue to Lemma 2.1.

Lemma 4.4.

$$\operatorname{rank} F = \dim\left(\operatorname{span}\{F(x,y) \mid x \in \mathbb{F}\}\right) = \dim\left(\operatorname{span}\{F(x,y) \mid y \in \mathbb{F}\}\right)$$

Proof. Let $\operatorname{rank} F = n$. Then (4.1) holds with $\{g_1, g_2, \dots, g_n\}$ and $\{h_1, h_2, \dots, h_n\}$ both linearly independent, by Lemma 4.3. This implies

$$\dim(\operatorname{span}\{g_1, g_2, \dots, g_n\}) = \dim(\operatorname{span}\{h_1, h_2, \dots, h_n\}) = n$$

and it suffices to show that $\text{span}\{F(x,y) \mid y \in \mathbb{F}\} = \text{span}\{g_1(x), g_2(x), \ldots, g_n(x)\}$ and $\text{span}\{F(x,y) \mid x \in \mathbb{F}\} = \text{span}\{h_1(y), h_2(y), \ldots, h_n(y)\}$.

For any $y \in \mathbb{F}$, $F(x,y)$ is a linear combination of $\{g_1(x), g_2(x), \ldots, g_n(x)\}$ and so $\text{span}\{F(x,y) \mid y \in \mathbb{F}\} \subseteq \text{span}\{g_1(x), g_2(x), \ldots, g_n(x)\}$. For the opposite inclusion, since $\mathbf{h}_n = \{h_1, h_2, \ldots, h_n\}$ is linearly independent, by Lemma 3.5 we can choose $\mathbf{v} = \{v_1, v_2, \ldots, v_n\} \subseteq \mathbb{F}$ such that the matrix $H = \mathbf{h}_n(\mathbf{v})$ is invertible. Then

$$\begin{bmatrix} F(x,v_1) \\ F(x,v_2) \\ \vdots \\ F(x,v_n) \end{bmatrix} = H \begin{bmatrix} g_1(x) \\ g_2(x) \\ \vdots \\ g_n(x) \end{bmatrix}$$

and the equation

$$\begin{bmatrix} g_1(x) \\ g_2(x) \\ \vdots \\ g_n(x) \end{bmatrix} = H^{-1} \begin{bmatrix} F(x,v_1) \\ F(x,v_2) \\ \vdots \\ F(x,v_n) \end{bmatrix}$$

implies that each $g_1(x), g_2(x), \ldots, g_n(x)$ is a linear combination of $F(x,v_1), \ldots, F(x,v_n)$ and are in $\text{span}\{F(x,y) \mid y \in \mathbb{F}\}$. Thus

$$\text{span}\{g_1(x), g_2(x), \ldots, g_n(x)\} \subseteq \text{span}\{F(x,y) \mid y \in \mathbb{F}\}$$

Similarly

$$\text{span}\{F(x,y) \mid x \in \mathbb{F}\} = \text{span}\{h_1(y), h_2(y), \ldots, h_n(y)\} \qquad \square$$

Definition 4.5. *For $n \in \mathbb{N}$, $F \in \mathbb{F}(x,y)$ define*

$$\Lambda_n = \Lambda_n(\mathbf{x},\mathbf{y}) = \begin{vmatrix} F(x_1,y_1) & F(x_1,y_2) & F(x_1,y_3) & \cdots & F(x_1,y_n) \\ F(x_2,y_1) & F(x_2,y_2) & F(x_2,y_3) & \cdots & F(x_2,y_n) \\ F(x_3,y_1) & F(x_3,y_2) & F(x_3,y_3) & \cdots & F(x_3,y_n) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ F(x_n,y_1) & F(x_n,y_2) & F(x_n,y_3) & \cdots & F(x_n,y_n) \end{vmatrix} \tag{4.2}$$

Λ_n is a polynomial in $\mathbb{F}[\mathbf{x},\mathbf{y}]$.

The reader may notice the immediate similarities between $\Lambda_n \in \mathbb{F}[\mathbf{x}, \mathbf{y}]$ and $V_n \in \mathbb{F}[\mathbf{x}]$. Lemma 4.6 shows that the two are related by more than just appearances. In fact, almost every lemma from Chapter 3 regarding the function V_n has an analogous lemma in Chapter 4 about Λ_n.

Lemma 4.6. *If $F(x, y) = g_1(x)h_1(y) + g_2(x)h_2(y) + \cdots + g_n(x)h_n(y)$ for some $g_1, g_2, \ldots, g_n \in \mathbb{F}[x]$ and $h_1, h_2, \ldots, h_n \in \mathbb{F}[y]$, then $\Lambda_n = V(\mathbf{g}_n, \mathbf{x})V(\mathbf{h}_n, \mathbf{y})$.*

Proof.

$$
\begin{aligned}
V(\mathbf{g}_n, \mathbf{x})V(\mathbf{h}_n, \mathbf{y}) &= \begin{vmatrix} g_1(x_1) & g_2(x_1) & \cdots & g_n(x_1) \\ g_1(x_2) & g_2(x_2) & \cdots & g_n(x_2) \\ \vdots & \vdots & \ddots & \vdots \\ g_1(x_n) & g_2(x_n) & \cdots & g_n(x_n) \end{vmatrix} \begin{vmatrix} h_1(y_1) & h_2(y_1) & \cdots & h_n(y_1) \\ h_1(y_2) & h_2(y_2) & \cdots & h_n(y_2) \\ \vdots & \vdots & \ddots & \vdots \\ h_1(y_n) & h_2(y_n) & \cdots & h_n(y_n) \end{vmatrix} \\
&= \left| \begin{bmatrix} g_1(x_1) & g_2(x_1) & \cdots & g_n(x_1) \\ g_1(x_2) & g_2(x_2) & \cdots & g_n(x_2) \\ \vdots & \vdots & \ddots & \vdots \\ g_1(x_n) & g_2(x_n) & \cdots & g_n(x_n) \end{bmatrix} \begin{bmatrix} h_1(y_1) & h_1(y_2) & \cdots & h_1(y_n) \\ h_2(y_1) & h_2(y_2) & \cdots & h_2(y_n) \\ \vdots & \vdots & \ddots & \vdots \\ h_n(y_1) & h_n(y_2) & \cdots & h_n(y_n) \end{bmatrix} \right| \\
&= \begin{vmatrix} F(x_1, y_1) & F(x_1, y_2) & F(x_1, y_3) & \cdots & F(x_1, y_n) \\ F(x_2, y_1) & F(x_2, y_2) & F(x_2, y_3) & \cdots & F(x_2, y_n) \\ F(x_3, y_1) & F(x_3, y_2) & F(x_3, y_3) & \cdots & F(x_3, y_n) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ F(x_n, y_1) & F(x_n, y_2) & F(x_n, y_3) & \cdots & F(x_n, y_n) \end{vmatrix} \\
&= \Lambda_n
\end{aligned}
$$

□

Lemma 4.7. *If $\Lambda_{n-1} \neq 0$ and $\Lambda_n = 0$, then* $\operatorname{rank} F < n$.

Proof. We show that $F(x, y) = g_1(x)h_1(y) + \cdots + g_{n-1}(x)h_{n-1}(y)$ for some $g_1, \ldots, g_{n-1} \in \mathbb{F}[x]$ and $h_1, \ldots, h_{n-1} \in \mathbb{F}[y]$. Since $\Lambda_{n-1} \neq 0$, there are $\mathbf{u}_{n-1} = (u_1, u_2, \ldots, u_{n-1})$ and $\mathbf{v}_{n-1} = (v_1, v_2, \ldots, v_{n-1}) \in \mathbb{F}^{n-1}$ such that $\Lambda_{n-1}(\mathbf{u}_{n-1}, \mathbf{v}_{n-1}) \in \mathbb{F}$ is nonzero (because $\mathbb{F}$ has characteristic zero). If we expand the determinant defining Λ_n in (4.2) along the last column and set $x_1 = u_1,\ x_2 = u_2,\ \ldots,\ x_{n-1} = u_{n-1},\ x_n = x$ and

$y_1 = v_1,\ y_2 = v_2,\ \dots,\ y_{n-1} = v_{n-1},\ y_n = y$ we get

$$0 = G_1(x)F(u_1, y) + G_2(x)F(u_2, y) + \cdots + G_{n-1}(x)F(u_{n-1}, y) + \Lambda_{n-1}(\mathbf{u}_{n-1}, \mathbf{v}_{n-1})\, F(x, y).$$

Here $G_1(x), G_2(x), \dots, G_{n-1}(x)$ are $(n-1)\times(n-1)$-minors of the matrix in which y does not appear. Since $\Lambda_{n-1}(\mathbf{u}_{n-1}, \mathbf{v}_{n-1}) \neq 0$, this equation can be solved for $F(x, y)$ resulting in the claimed form. □

Lemma 4.8. $\Lambda_n = 0$ *if and only if* $\operatorname{rank} F < n$.

Proof. If $\operatorname{rank} F < n$ then $F(x, y) = g_1(x)h_1(y) + \cdots + g_{n-1}(x)h_{n-1}(y)$ for some $g_1, g_2, \dots, g_{n-1} \in \mathbb{F}[x]$ and $h_1, h_2, \dots, h_{n-1} \in \mathbb{F}[y]$. Hence $\Lambda_n = 0$ follows from Lemma 4.6 with $g_n(x) = h_n(y) = 0$.

If $\Lambda_1 = 0$ then $F(x, y) = 0$ and the claim is trivially true. Otherwise we have $\Lambda_1 \neq 0$ and $\Lambda_n = 0$, so there must be some $1 < k \leq n$ such that $\Lambda_{k-1} \neq 0$ and $\Lambda_k = 0$. By Lemma 4.7, $\operatorname{rank} F < k \leq n$. □

Lemma 4.9. *$F \in \mathbb{F}[x, y]$ has rank n if and only if* (4.1) *holds with $\{g_1, g_2, \dots, g_n\}$ and $\{h_1, h_2, \dots, h_n\}$ linearly independent.*

Proof. Suppose that $\operatorname{rank} F = n$. Then (4.1) holds for some $g_1, g_2, \dots, g_n \in \mathbb{F}[x]$ and $h_1, h_2 \dots, h_n \in \mathbb{F}[y]$. By Lemma 4.3, $\{g_1, g_2, \dots, g_n\}$ and $\{h_1, h_2, \dots, h_n\}$ must be linearly independent since otherwise $\operatorname{rank} F < n$.

Now suppose that $\{g_1, g_2, \dots, g_n\}$ and $\{h_1, h_2, \dots, h_n\}$ are linearly independent. Then, by Lemma 4.6 and Lemma 3.4, $\Lambda_n \neq 0$ and $\operatorname{rank} F \geq n$ by Lemma 4.8. On the other hand, since F has the form (4.1), we also have $\operatorname{rank} F \leq n$. Therefore $\operatorname{rank} F = n$. □

We turn our attention towards crafting an analogue of the Wronskian, which we call Δ_n, in such a way that any lemmas about W_n from Chapter 3 will translate into near identical lemmas for Δ_n.

For $i, j \geq 0$, we write $\partial_{ij}F = \dfrac{\partial^{i+j}F}{\partial x^i \partial y^j}$. For example, $\partial_{00}F = F$, $\partial_{10}F = \dfrac{\partial F}{\partial x}$, $\partial_{11}F = \dfrac{\partial^2 F}{\partial x \partial y}$, etc.

Definition 4.10. *Let Δ_n be the determinant of the matrix of partial derivatives of F:*

$$\Delta_n = \Delta_n(x, y) = \begin{vmatrix} \partial_{00}F & \partial_{01}F & \partial_{02}F & \cdots & \partial_{0,n-1}F \\ \partial_{10}F & \partial_{11}F & \partial_{12}F & \cdots & \partial_{1,n-1}F \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \partial_{n-1,0}F & \partial_{n-1,1}F & \partial_{n-1,2}F & \cdots & \partial_{n-1,n-1}F \end{vmatrix}$$

Then Δ_n is a polynomial in $\mathbb{F}[x, y]$.

The reader may recall that in the introduction we discussed a connection, revealed by Lemma 3.12, between the Wronskian and the alternant of a set of polynomials. So far, we have generalized the alternant V_n, to an operator Λ_n that acts on a two variable polynomial. Now we will show that Δ_n is an operator on a two variable polynomial that fits the role of our generalized Wronskian. We do this by first showing that it has a similar connection with Λ_n as the Wronskian had with the alternant, an analogue of Lemma 3.12.

Lemma 4.11.

$$\Delta_n = \left. \frac{(\delta(1, 2, \ldots, n))^2 \Lambda_n}{\delta(\mathbf{x})\delta(\mathbf{y})} \right|_{\substack{x_1 = x_2 = \cdots = x_n = x \\ y_1 = y_2 = \cdots = y_n = y}}$$

Proof. Let $f_1 = \partial_{00}F, f_2 = \partial_{01}F, \ldots, f_n = \partial_{0,n-1}F \in \mathbb{F}(y)[x]$ and $g_1 = f_1(x_1, y), g_2 = f_1(x_2, y), \ldots, g_n = f_1(x_n, y) \in \mathbb{F}(x_1, x_2, \ldots, x_n)[y]$. So each f_j is a polynomial with indeterminant x and each g_i is a polynomial with indeterminant y. Then taking derivatives of f_j with respect to x we get $\dfrac{d^i}{dx^i}f_j = \partial_{i,j-1}F$. So we can write

$$\Delta_n = W(f_1, f_2, \ldots, f_n)$$

Using Lemma 3.12 we get

$$\Delta_n = W(\mathbf{f}_n) = \frac{\delta(1,2,\ldots,n)}{\delta(\mathbf{x})}\left.\begin{vmatrix} f_1(x_1) & f_2(x_1) & \cdots & f_n(x_1) \\ \vdots & \vdots & \ddots & \vdots \\ f_1(x_n) & f_2(x_n) & \cdots & f_n(x_n) \end{vmatrix}\right|_{x_1=x_2=\cdots=x_n=x}$$

Notice that $f_j(x_i) = \dfrac{d^{j-1}}{dy^{j-1}}f_1(x_i) = \dfrac{d^{j-1}}{dy^{j-1}}g_i$. So we can write

$$\frac{\delta(1,2,\ldots,n)}{\delta(\mathbf{x})}\left.\begin{vmatrix} f_1(x_1) & f_2(x_1) & \cdots & f_n(x_1) \\ \vdots & \vdots & \ddots & \vdots \\ f_1(x_n) & f_2(x_n) & \cdots & f_n(x_n) \end{vmatrix}\right|_{x_1=x_2=\cdots=x_n=x}$$

$$= \frac{\delta(1,2,\ldots,n)}{\delta(\mathbf{x})}W(g_1, g_2, \ldots, g_n)\Big|_{x_1=x_2=\ldots=x_n=x}$$

Using Lemma 3.12 again we get

$$\Delta_n = \frac{\delta(1,2,\ldots,n)}{\delta(\mathbf{x})}W(g_1, g_2, \ldots, g_n)\Big|_{x_1=x_2=\ldots=x_n=x}$$

$$= \frac{(\delta(1,2,\ldots,n))^2}{\delta(\mathbf{x})\delta(\mathbf{y})}\left.\begin{vmatrix} f_1(x_1,y_1) & f_1(x_2,y_1) & \cdots & f_1(x_n,y_1) \\ \vdots & \vdots & \ddots & \vdots \\ f_1(x_1,y_n) & f_1(x_2,y_n) & \cdots & f_1(x_n,y_n) \end{vmatrix}\right|_{\substack{x_1=x_2=\cdots=x_n=x \\ y_1=y_2=\cdots=y_n=y}}$$

$$= \frac{(\delta(1,2,\ldots,n))^2\,\Lambda_n}{\delta(\mathbf{x})\delta(\mathbf{y})}\Bigg|_{\substack{x_1=x_2=\cdots=x_n=x \\ y_1=y_2=\cdots=y_n=y}}$$

□

Lemma 4.12. *If*

$$F(x,y) = g_1(x)h_1(y) + g_2(x)h_2(y) + \cdots + g_n(x)h_n(y)$$

for some $g_1, g_2, \ldots, g_n \in \mathbb{F}[x]$ *and* $h_1, h_2, \ldots, h_n \in \mathbb{F}[y]$*, then*

$$\Delta_n = W(\mathbf{g}_n, x)W(\mathbf{h}_n, y)$$

Proof. By Lemma 4.6, $\Lambda_n = V(\mathbf{g}_n, \mathbf{x})V(\mathbf{h}_n, \mathbf{y})$. Combining this with Lemma 4.11 and Lemma 3.12 gives the desired result. $\square$

Lemma 4.13. *If* $\operatorname{rank} F = n$, *then* $\Delta_n \neq 0$.

Proof. By Lemma 4.9, (4.1) holds with $\{g_1, g_2, \ldots, g_n\}$ and $\{h_1, h_2, \ldots h_n\}$ linearly independent. Because of Lemma 3.9 we have $W(\mathbf{g}_n, x) \neq 0$ and $W(\mathbf{h}_n, y) \neq 0$. So, by Lemma 4.12, $\Delta_n \neq 0$. $\square$

Lemma 4.14. $\Delta_n = 0$ *if and only if* $\operatorname{rank} F < n$.

Proof. If $\operatorname{rank} F < n$ then

$$F(x, y) = g_1(x)h_1(y) + \cdots + g_{n-1}(x)h_{n-1}(y)$$

for some $g_1, g_2, \ldots, g_{n-1} \in F[x]$ and $h_1, h_2, \ldots, h_{n-1} \in \mathbb{F}[y]$. Hence $\Delta_n = 0$ follows from Lemma 4.12 with $g_n = h_n = 0$.

Now assume $\Delta_n = 0$. Then let $f_i = \partial_{0,i-1} F \in \mathbb{F}(y)[x]$. Then $W(f_1, \ldots, f_n) = \Delta_n = 0$. Since $\mathbb{F}(y)$ is a field of characteristic zero, by Lemma 3.9 we know that $\{f_1, f_2, \ldots, f_n\}$ is linearly dependent over $\mathbb{F}(y)$. So $\Delta_m = W(f_1, \ldots, f_n, \ldots, f_m) = 0$ for all $m \geq n$. Thus, by Lemma 4.13, $\operatorname{rank} F \neq m$ for any $m \geq n$. Therefore $\operatorname{rank} F < n$. $\square$

CHAPTER 5

Implications for Matrices

Let $\mathbb{F}^{(m+1)\times(n+1)}$ be the vector space of all $(m+1)\times(n+1)$ matrices with entries in $\mathbb{F}$. Let

$$M = \begin{bmatrix} a_{00} & a_{01} & \cdots & a_{0n} \\ a_{10} & a_{11} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m0} & a_{m1} & \cdots & a_{mn} \end{bmatrix} \in \mathbb{F}^{(m+1)\times(n+1)}$$

and define a linear function $T : \mathbb{F}^{(m+1)\times(n+1)} \to \mathbb{F}[x, y]$ by

$$T(M) = F(x, y) = \sum_{i=0}^{m} \sum_{j=0}^{n} a_{ij} x^i y^j.$$

The following example illustrates this mapping.

Example 5.1.

$$\text{Let } M = \begin{bmatrix} 2 & 0 & 1 & -1 & 0 \\ 1 & 2 & 1 & 0 & 1 \\ -3 & 3 & 0 & 2 & 2 \\ 1 & -4 & 1 & -2 & -1 \end{bmatrix}$$

$$\begin{aligned} \text{Then } T(M) = F(x, y) &= 2 + 0y + 1y^2 - 1y^3 + 0y^4 \\ &+ 1x + 2xy + 1xy^2 + 0xy^3 + 1xy^4 \\ &- 3x^2 + 3x^2y + 0x^2y^2 + 2x^2y^3 + 2x^2y^4 \\ &+ 1x^3 - 4x^3y + 1x^3y^2 - 2x^3y^3 - 1x^3y^4 \end{aligned}$$ ■

Lemma 5.2. *For all* $A = \left[a_i\right] \in \mathbb{F}^{(m+1)\times 1},\ B = \left[b_j\right] \in \mathbb{F}^{1\times(n+1)}$

$$T(AB) = T(A)T(B)$$

Proof. Notice that because of the size of A and B, their product is very easy to work with and we have

$$T(AB) = T\left(\left[a_i b_j\right]\right) = \sum_{i=0}^{m} \sum_{j=0}^{n} a_i b_j x^i y^j = \sum_{i=0}^{m} a_i x^i \sum_{j=0}^{n} b_j y^j = T(A)T(B)$$ □

The reader should observe that in the lemma above, $T(A) \in \mathbb{F}[x]$ and $T(B) \in \mathbb{F}[y]$. Recall that, by Theorem 2.2, we can decompose any rank k matrix M into a sum of k rank 1 matrices, C_i, which can each be written as a product of a column vector A_i and a row vector B_i. Then, using the linear properties of T, Lemma 5.2 gives us

$$T(M) = T\left(\sum_{i=1}^{k} C_i\right) = \sum_{i=1}^{k} T(C_i) = \sum_{i=1}^{k} T(A_i B_i) = \sum_{i=1}^{k} T(A_i)T(B_i).$$

Theorem 5.3. *With M, $F(x,y) = T(M)$, Λ_n and Δ_n as above, the following are equivalent:*

(1) $\operatorname{rank} M < n$.

(2) *M can be written as a sum of $n-1$ rank one matrices.*

(3) $\operatorname{rank} F < n$.

(4) *$F(x,y) = g_1(x)h_1(y) + g_2(x)h_2(y) + \cdots + g_{n-1}(x)h_{n-1}(y)$ for some $g_1, g_2, \ldots, g_{n-1} \in \mathbb{F}[x]$ and $h_1, h_2, \ldots, h_{n-1} \in \mathbb{F}[y]$.*

(5) $\Lambda_n = 0$.

(6) $\Delta_n = 0$.

Proof. Theorem 2.2 shows that (1) holds if and only if (2) does. We have (2) if and only if (4) by Theorem 2.2 and Lemma 5.2. Item (3) is equivalent to (4) by definition. Lemma 4.8 gives us (3) if and only if (5). Then Lemma 4.14 says (3) holds if and only if (6) does. □

Corollary 5.4. *Let $M \in \mathbb{F}^{(m+1)\times(n+1)}$ and $F(x,y) = T(M)$ as defined above. Then The following are equivalent:*

(1) $\operatorname{rank} M = n$.

(2) $\operatorname{rank} F = n$.

(3) $\Lambda_n \neq 0$ *and* $\Lambda_{n+1} = 0$.

(4) $\Delta_n \neq 0$ *and* $\Delta_{n+1} = 0$.

Now that Corollary 5.4 has been established the reader may be wondering why this is important to us. To answer that question we refer the reader to an article by Winfried Bruns and Roland Schwänzl titled "The Number of Equations Defining a Determinantal Variety"[13]. Algebraic varieties are the central object of study in algebraic geometry, and so the content of Bruns and Schwänzl's article is beyond the scope of this thesis. In simple terms, their article proves that there are $mn - t^2 + 1$ equations that can be checked to determine if an $m \times n$ matrix M, with entries in an algebraically closed field $\mathbb{F}$, has $\operatorname{rank} M < t$. They also show that this is the minimum number of equations with this property.

Let's look at an example below to illustrate this idea.

Example 5.5. Consider the 4×6 matrix $M = [a_{ij}]$,

$$M = \begin{bmatrix} 1 & 0 & -1 & 0 & 4 & -3 \\ 0 & 2 & -2 & 1 & 0 & 5 \\ 1 & 1 & 1 & 0 & 0 & -4 \\ 3 & 1 & 2 & -2 & -1 & -3 \end{bmatrix}$$

Determining if $\operatorname{rank} M = 0$ is equivalent to determining if rank < 1, which Bruns and Schwänzl's article tells us requires no less than $4(6) - 1^2 + 1 = 24$ equations. We know that a matrix has rank zero if and only if every entry is zero, and so in this case it is easy to see that the 24 equations needed are $a_{ij} = 0$ for all i and j.

In general, the $mn - t^2 + 1$ equations described by Bruns and Schwänzl are complicated and involve combinations of minors of different sizes. We wish to find simple equations that we can use to determine the rank of M, and recall that if we wish to know whether or not rank < 2 then we could always calculate all the (2×2) minors of M and check if they are all zero. But there are $\binom{4}{2}\binom{6}{2} = 90$ (2×2) minors here, and we only need at most $4(6) - 2^2 + 1 = 21$ equations to answer this question.

So one may wonder whether this new method of calculating matrix rank using the ideas introduced in Chapter 4, hereafter known as the Δ method, is an improvement or not. In fact, we show now that the Δ method of calculating matrix rank falls directly between using Bruns and Schwänzl's equations and calculating the minors of the matrix, both in terms of the number of equations and the complexity of the equations.

Recall that a matrix M has rank $< t$ if and only if $\Delta_t = 0$, where Δ_t is calculated using the polynomial $T(M)$. Since $\Delta_t \in \mathbb{F}[x, y]$ is a polynomial, this corresponds to checking if each of the coefficients of Δ_t are zero. Our next lemma tells us exactly how many coefficients there are to calculate.

Lemma 5.6. *Let $F(x, y) \in \mathbb{F}[x, y]$ with $\deg_x(F) = m_x$ and $\deg_y(F) = m_y$. Then:*

(1) $\deg_x(\Delta_t) \leq m_x t - t^2 + t$ *and* $\deg_y(\Delta_t) \leq m_y t - t^2 + t$.

(2) Δ_t *has at most* $(m_x t - t^2 + t + 1)(m_y t - t^2 + t + 1)$ *coefficients.*

Proof. (1) Let $f_1(x) = F,\ f_2(x) = F_y,\ f_3(x) = F_{yy}, \ldots,\ f_n(x) = \dfrac{\partial^t F}{\partial y^{t-1}} \in \mathbb{F}[y][x]$. Then $\deg_x(f_i) \leq m_x$ for all $i = 1, \ldots, t$ and $\Delta_t = W_t(f_1(x), f_2(x), \ldots, f_t(x))$. So $\deg_x(\Delta_t) = \deg_x(W_t(f_1, f_2, \ldots, f_t) \leq m_x t - t^2 + t$ by Lemma 3.12.

Similarly, by setting $g_i(y) = \dfrac{\partial^t F}{\partial x^{t-1}} \in \mathbb{F}[x][y]$, we have

$$\deg_y(\Delta_t) = \deg_y(W_t(g_1, g_2, \ldots, g_t) \leq m_y t - t^2 + t.$$

(2) Each term in Δ_t is of the form ax^iy^j where $a \in \mathbb{F},\ 0 \leq i \leq m_x t - t^2 + t$, and $0 \leq j \leq m_y t - t^2 + t$. So there are at most $(m_x t - t^2 + t + 1)(m_y t - t^2 + t + 1)$ coefficients in Δ_t. □

The comparison between the number of coefficients in Δ_t and the number of $(t \times t)$-minors of an $m \times n$ matrix M is not immediately obvious, so the next lemma makes it clear.

Lemma 5.7. *Let $m, n, t \in \mathbb{Z}$ with $0 < t \leq m \leq n$. Then*

$$mn - t^2 + 1 \leq (mt - t^2 + 1)(nt - t^2 + 1) \leq \binom{m}{t}\binom{n}{t}$$

.

Proof. For $t = 1$ equality is obvious throughout. So let $t > 1$.

First we focus on the inequality $(mt - t^2 + 1)(nt - t^2 + 1) \leq \binom{m}{t}\binom{n}{t}$. We show $\binom{m}{t} \geq t(m - t) + 1$. Fix $m > 1 \in \mathbb{Z}$ and notice that the symmetry, $t \longleftrightarrow m - t$, in the binomial coefficient on the left side as well as the product on the right implies that we only need to consider the case $2t \leq m$. We can write

$$\binom{m}{t} = \frac{(m)(m-1)(m-2)\cdots(m-t+1)}{(1)(2)(3)\cdots(t)} \tag{5.1}$$

and notice that $2t \leq m$ implies $t \leq m-t+1$. Since the numerator and the denominator of the right side of (5.1) have the same number of factors, and $1 \leq \dfrac{(m-t+1)+d}{(t)-d}$ for all $d \geq 0$, we can drop all but the first two factors of the numerator and the denominator and get a strict inequality

$$\binom{m}{t} > \frac{(m)(m-1)}{(1)(2)}.$$

Then by the assumption that $t \leq \dfrac{m}{2}$ and the fact that $m-1 > m-t$, we get

$$\binom{m}{t} > t(m-t).$$

Because these are both integer values we can add one on the right side and get the desired result.

Similarly $\binom{n}{t} \geq t(n-t)+1$ for all $t > 1 \in \mathbb{Z}$. Therefore

$$(mt - t^2 + 1)(nt - t^2 + 1) \leq \binom{m}{t}\binom{n}{t}$$

Now we show $mn - t^2 + 1 \leq (mt - t^2 + 1)(nt - t^2 + 1)$. Notice that for $1 < t \leq m \leq n$ we have $1 - t^2 < 0,\ t - n \leq 0,\ m - t \geq 0$. So

$$0 \leq (1-t^2)(t-n)(m-t) = mt - t^2 - mn + nt - mt^3 + t^4 + mnt^2 - nt^3.$$

Adding $mn - t^2 + 1$ to both sides of the inequality gives us

$$mn - t^2 + 1 \leq mt - 2t^2 + nt - mt^3 + t^4 + mnt^2 - nt^3 + 1 = (mt - t^2 + 1)(nt - t^2 + 1)$$

as desired. $\square$

The number $(mt - t^2 + 1)(nt - t^2 + 1)$ in Lemma 5.7 may seem smaller than $(m_y t - t^2 + t + 1)(m_x t - t^2 + t + 1)$ in Lemma 5.6. However, if we have $F(x, y) = T(M)$ for some $m \times n$ matrix M then $m_x \leq m - 1$ and $m_y \leq n - 1$, so a quick substitution reveals that actually $(mt - t^2 + 1)(nt - t^2 + 1) \geq (m_y t - t^2 + t + 1)(m_x t - t^2 + t + 1)$. This shows that the number of equations needed to determine if an $m \times n$ matrix

	m=1	m=2	m=3	m=4	m=5
t=1	10-10-10	20-20-20	30-30-30	40-40-40	50-50-50
t=2	0	17-17-45	27-51-135	37-85-270	47-119-450
t=3	0	0	22-22-120	32-88-480	42-154-1200
t=4	0	0	0	25-25-210	35-125-1050
t=5	0	0	0	0	26-26-252
t=6	0	0	0	0	0
t=7	0	0	0	0	0
t=8	0	0	0	0	0
t=9	0	0	0	0	0
t=10	0	0	0	0	0
	m=6	**m=7**	**m=8**	**m=9**	**m=10**
t=1	60-60-60	70-70-70	80-80-80	90-90-90	100-100-100
t=2	57-153-675	67-187-945	77-221-1260	87-255-1620	97-289-2025
t=3	52-220-2400	62-286-4200	72-352-6720	82-418-10080	92-484-14400
t=4	45-225-3150	55-325-7350	65-425-14700	75-525-26460	85-625-44100
t=5	36-156-1512	46-286-5292	56-416-14112	66-546-31752	76-676-63504
t=6	25-25-210	35-175-1470	45-325-5880	55-475-17640	65-625-44100
t=7	0	22-22-120	32-176-960	42-330-4320	52-484-14400
t=8	0	0	17-17-45	27-153-405	37-289-2025
t=9	0	0	0	10-10-10	20-100-100
t=10	0	0	0	0	1-1-1

Table 5.1: The number of equations needed to calculate matrix rank using different methods

M has rank $< t$, using the Δ method, is bounded below by the number of equations defined by Bruns and Schwänzl, and bounded above the number of minors of M.

The following table demonstrates the number of equations needed to check if an $m \times 10$ matrix has rank less than t. Each cell of the table has three numbers, $a-b-c$, which correspond to the minimum number as described by Bruns and Schwänzl, the number of coefficients in Δ_t, and the number of $t \times t$ minors respectively.

Some things worth pointing out in Table 5.1 are that for $t = 1$ all three methods require the same number of equations. This corresponds to simply checking if every entry of the matrix is zero. Also, along the diagonal $t = m$ we have equality between the number of coefficients in Δ_t and the minimun number of equations necessary. Δ_m is equivalent to mapping the m rows of the matrix into polynomials and then calculating the Wronskian of those polynomials as was described in the introduction of this thesis. Notice how both the minimum number of equations needed and the number of coefficients of Δ_t grow linearly with m, but the number of $t \times t$ minors grows much more rapidly suggesting that there may be significant computational advantages to using the Δ_t method in practice.

REFERENCES

[1] A. Ostrowski, *Über ein Analogon der Wronskischen Determinante bei Funktionen mehrerer Veränderlicher,* Math. Z., **4**, (1919), 223–230.

[2] W.P. Wardlaw, *Row Rank Equals Column Rank,* Math. Mag., **78**, No 4, (Oct., 2005), 316–318.

[3] S. Puntanen, *Matrix Tricks for Linear Statistical Models,* Springer, (2011), 349–350.

[4] T. Muir, *A treatise on the theory of determinants,* Dover Publications, (1960), 321–363.

[5] J. Hoëne-Wroński, *Réfutation de la théorie des fonctions analytiques de Lagrange*, Dédiée a l'institut impérial de France. Paris: Blankenstein, (1812).

[6] S. Roman, *Advanced Linear Algebra; 3rd edition* Springer, (2007).

[7] E.A. Coddington, *Ordinarly Linear Differential Equations,* Krieger Pub Co, (1984).

[8] S. Lang, *Linear Algebra; 2nd Edition,* Addison-Wesley, (1971), 173–191.

[9] G. Peano, *Sur le determinant Wronskien,* Mathesis 9 (1889), 75–76

[10] R.A. Horn, *On the Wronskian Test for Independence,* Amer. Math. Monthly, **77**, No 1, (1970), 65–66.

[11] J.M. Landsberg, *Tensors: Geometry and Applications,* Amer. Math. Soc., **128**, (2012), 250–264

[12] M. Bocher, *On Linear Dependence of Functions of One Variable,* Bull. Amer. Math. Soc. 7, (1900), 120–121.

[13] W. Bruns, R. Schwänzl, *The Number of Equations Defining a Determinantal Variety*, Bull. London Math. Soc. 22, (1990), 439–445.