Prime Rational Functions and Integral Polynomials

Jesse Larone, Bachelor of Science

Mathematics and Statistics

Submitted in partial fulfillment
of the requirements for the degree of

Master of Science

Faculty of Mathematics and Science, Brock University
St. Catharines, Ontario

© 2014

Abstract. Let $f(x)$ be a complex rational function. In this work, we study conditions under which $f(x)$ cannot be written as the composition of two rational functions which are not units under the operation of function composition. In this case, we say that $f(x)$ is prime. We give sufficient conditions for complex rational functions to be prime in terms of their degrees and their critical values, and we derive some conditions for the case of complex polynomials. We consider also the divisibility of integral polynomials, and we present a generalization of a theorem of Nieto. We show that if $f(x)$ and $g(x)$ are integral polynomials such that the content of g divides the content of f and $g(n)$ divides $f(n)$ for an integer n whose absolute value is larger than a certain bound, then $g(x)$ divides $f(x)$ in $\mathbb{Z}[x]$. In addition, given an integral polynomial $f(x)$, we provide a method to determine if f is irreducible over $\mathbb{Z}$, and if not, find one of its divisors in $\mathbb{Z}[x]$.

Key words and phrases : Prime polynomials; Prime rational functions; Critical Values; Resultant; Integral polynomials.

Acknowledgement.

I would like to extend my deepest gratitude to my supervisor Dr. Omar Kihel for his continued support of my studies and his contributions to this research. I truly appreciate his expertise and knowledge, and I have benefited greatly from our conversations, as well as from his comments, suggestions, and advice.

In addition to my supervisor, I would also like to thank the other members of my thesis committee: Dr. Bill Ralph, Dr. Babak Farzad, and Dr. Sheridan Houghten, for their comments, encouragement, and difficult questions.

I thank Dr. Mohamed Ayad for his results and contributions. Our discussions were also of great importance in the furthering of this research.

My thanks also goes to the Natural Sciences and Engineering Research Council (NSERC) for their financial assistance.

I would lastly like to thank my family for their constant support and encouragement throughout my graduate program and my life.

Contents

1. Introduction

Given a set of elements or a single element of a set, it is natural to seek a way to simplify the given object. One of the motivating examples is the set of prime numbers. This concept can be extended to that of irreducible elements in integral domains. These considerations deal with the operation of multiplication, where we look to take elements of a ring and write them as products of irreducible elements of that ring. This naturally leads us to consider similar approaches for different operations.

Let $f(x)$ be a non-constant polynomial. Ayad's paper [1] and Beardon's paper [2] deal with the possibility of expressing $f(x)$ as the composition of two polynomials $g(x)$ and $h(x)$ with degrees at least 2. In this case $f(x)$ is said to be composite, otherwise $f(x)$ is said to be prime. We will look to extend this definition to rational functions.

In Chapter 2, we recall some definitions and results concerning rings, fields, and vector spaces. We present the material required to define polynomial rings and fields of fractions, which are two important concepts that will be required for many of the results found in this work. In Chapter 3, we present a few of Ayad's results from [1] which will serve as the foundation for many of the definitions and main results presented in the following chapter. In Chapter 4, we motivate the definitions of prime and composite rational functions. We make use of the set of units under function composition to show that the multiplicities of a rational function's zeros and poles are useful to determine if that rational function is prime or composite. Among other results, we show that a complex rational function $f(x)$ of degree n is prime if it has a zero or a pole whose multiplicity is divisible by a prime number $p > d$, where d is the greatest proper divisor of n. We also extend the definition of the resultant to rational functions, which is then used to define the multiplicity of a critical value of a rational function. The multiplicities of a rational function's critical values, as well as the number of critical values it possesses, can also provide more prime rational functions. In particular, a rational function $f(x)$ of degree n is prime if it has at least $2d$ non-zero critical values with multiplicity 1, where d is the greatest proper divisor of n. In Chapter 5, we extend the work of Nieto [7] on the divisibility of integral polynomials, who showed that if $f(x)$ and $g(x)$ are polynomials over $\mathbb{Z}$ such that $cont(g)$ divides $cont(f)$ and $g(n)$ divides $f(n)$ for infinitely many integers n, then the polynomial $g(x)$ divides $f(x)$. We improve upon this result by showing that it is sufficient for $g(n)$ to divide $f(n)$ for any integer n greater than a certain bound. We also provide a method to determine if an integral polynomial $f(x)$ is irreducible over $\mathbb{Z}$, and if not, find one of its divisors in $\mathbb{Z}[x]$.

2. Rings, Fields, and Vector Spaces

In this chapter we establish results regarding the structures of rings and fields. We begin on the topic of rings, and we present the necessary results to motivate and introduce unique factorization domains and fields, which are rings which possess certain particular and desirable properties. We then introduce vector spaces. Among other results, we define a basis for a vector space and use this concept as motivation to define the dimension of a vector space. We lastly discuss extension fields. The results developed for rings and vector spaces lead to the definition of extension degrees, which are useful in the justification of results regarding algebraic field extensions.

2.1. Rings.

The material presented here covers the theorems we will require on rings. We recall the definitions and some properties of integral domains, polynomial rings, principal ideal domains, unique factorization domains, ideals, and fields of fractions. Unless otherwise stated, these results can be found in [3].

Definition 2.1. *A* ring R *is a set with two binary operations called addition and multiplication, denoted by* $a + b$ *and* ab *respectively, such that the following properties hold for all* $a, b, c \in R$*:*

i) $a + b = b + a$*;*
ii) $(a + b) + c = a + (b + c)$*;*
iii) there exists an element $0 \in R$ *such that* $a + 0 = a$ *for all* $a \in R$*;*
iv) for each $a \in R$ *there exists an element* $-a$ *such that* $a + (-a) = 0$*;*
v) $a(bc) = (ab)c$*;*
vi) $a(b + c) = ab + ac$ *and* $(b + c)a = ba + ca$*.*

A *commutative ring* is a ring where multiplication is commutative. We note that there is no requirement in our definition that a ring must possess an identity element under the operation of multiplication. A non-zero element of a ring that is an identity under multiplication is called a *unity*, and we denote it by 1. Similarly, there is no requirement that a non-zero element of a commutative ring with unity must have a multiplicative inverse. A non-zero element of a ring which does possess a multiplicative inverse is called a *unit*. If a and b are two elements of a commutative ring and a is non-zero, we say that a is a *divisor* or a *factor* of b if there exists an element c in the ring such that $b = ac$.

It will be useful to consider rings with properties which mimic the familiar properties of the integers. This important class of rings is described with the following definition.

Definition 2.2. *A* zero-divisor *is a non-zero element* a *of a commutative ring* R *such that there exists a non-zero element* $b \in R$ *with* $ab = 0$*. An* integral domain *is a commutative ring with unity and no zero-divisors.*

Definition 2.3. *Elements a and b of an integral domain R are said to be* associates *if $a = ub$ for some unit $u \in R$. A non-zero element a of an integral domain R is called an* irreducible *if a is not a unit and, whenever $a = bc$ for elements $b, c \in R$, then b or c is a unit. A non-zero element a of an integral domain R is called a* prime *if a is not a unit and if a divides b or a divides c whenever a divides bc for $b, c \in R$.*

Prime and irreducible elements of an integral domain are related by the following result.

Theorem 2.1. *Let R be an integral domain, and let $p \in R$ be prime. Then p is irreducible in R.*

Proof. Suppose that $p = ab$ for some $a, b \in R$. Since p is prime, it must divide either a or b. Assume without loss of generality that p divides a, then $a = pc$ for some $c \in R$. Then $p = ab = (pc)b = pcb$ and thus $1 = cb$. Then b is a unit, so p is irreducible in R. □

Another important class of rings are given in the next definition.

Definition 2.4. *A* field *is a commutative ring with unity in which every non-zero element is a unit.*

Definition 2.5. *A subset K of a field F is called a* subfield *of F if K is a field with the operations of F.*

Given a field F and a subset K of F, the following result provides a convenient way to determine if K is itself a field without needing to verify all of the required properties of a field.

Theorem 2.2. *Let F be a field and let K be a subset of F with at least two elements. Then K is a subfield of F if $a - b, ab^{-1} \in K$ for all $a, b \in K$ where $b \neq 0$.*

Proof. Since the operation of addition of K is the same as the operation of addition of F, it is both associative and commutative. Since K is non-empty, we choose an element $x \in K$ and let $a = x$ and $b = x$ so that $0 = x - x = a - b \in K$ by our hypothesis. Letting $a = 0$ and $b = x$, we have $-x = 0 - x = a - b \in K$ as well. To show that addition is closed, we let $x, y \in K$ and show that $x + y \in K$. Since $-y \in K$, we set $a = x$ and $b = -y$ so that $x + y = x - (-y) = a - b \in K$.

Since the operation of multiplication of K is the same as the operation of multiplication of F, it is associative, commutative, and multiplication distributes over addition. Since K has at least two elements, we choose $x, y \in K$ such that $y \neq 0$ and let $a = y$ and $b = y$ so that $1 = yy^{-1} = ab^{-1} \in K$ by our hypothesis. Letting $a = 1$ and $b = y$, we have $y^{-1} = 1y^{-1} = ab^{-1} \in K$ as well. To show that multiplication is closed, we let $x, y \in K$ such that $y \neq 0$ and show that $xy \in K$. Since $y^{-1} \in K$, we set $a = x$ and $b = y^{-1}$ so that $xy = x(y^{-1})^{-1} = ab^{-1} \in K$. Therefore K is a field with the operations of F. □

In the same way that we have considered subfields of fields, it will be of use to consider particular subsets of rings in general.

Definition 2.6. *A subset A of a ring R is called a* subring *of R if A is a ring with the operations of R. A subset A of a ring R is called a two-sided* ideal *of R if*

i) A is a subring of R;

ii) $ra \in A$ and $ar \in A$ for every $r \in R$ and every $a \in A$.

We will call a two-sided ideal of a ring R simply an ideal *of R.*

Given a ring R and a subset A of R, the following result provides a way to determine if A is an ideal of R.

Theorem 2.3. *A non-empty subset A of a ring R is an ideal of R if*

i) $a - b \in A$ for all $a, b \in A$;

ii) $ra, ar \in A$ for every $a \in A$ and every $r \in R$.

Proof. Since the operation of addition of A is the same as the operation of addition of R, it is both associative and commutative. Since A is non-empty, we choose an $x \in A$ and let $a = x$ and $b = x$ so that $0 = x - x = a - b \in A$ by our hypothesis. Letting $a = 0$ and $b = x$, we have $-x = 0 - x = a - b \in A$ as well. To show that addition is closed, we let $x, y \in A$ and show that $x + y \in A$. Since $-y \in A$, we set $a = x$ and $b = -y$ so that $x + y = x - (-y) = a - b \in A$. Lastly, since the operation of multiplication of A is the same as the operation of multiplication of R, it is both associative and distributive over addition. Then for all $x, y \in A$, we have $y \in R$ so that $xy \in A$ by our hypothesis. Therefore A is a subring of R. The second condition of the theorem is then sufficient to conclude that A is an ideal of R. □

Let R be a commutative ring with unity and let $a \in R$. We note that the set $\langle a \rangle = \{ra \mid r \in R\}$ is an ideal of R, and we call such an ideal a *principal ideal* of R generated by the element a.

Let R be a ring and let A be a subset of R. For all $r \in R$, we define the set $r + A$ by

$$r + A = \{r + a \mid a \in A\}.$$

If A is an ideal of R, we have the following result.

Lemma 2.1. *Let A be an ideal of a ring R and let $a \in R$. Then $a + A = A$ if and only if $a \in A$.*

Proof. Suppose that $a + A = A$. Then $a = a + 0 \in a + A = A$, thus $a \in A$. Now assume that $a \in A$. Since A is closed under addition, we have $a + A \subseteq A$. To show that $A \subseteq a + A$, we let $b \in A$. We then have $b - a \in A$ and $b = 0 + b = (a - a) + b = a + (b - a) \in a + A$. Therefore $A = a + A$. □

For two ideals A_1 and A_2 of R, the sets

$$A_1 + A_2 = \{a_1 + a_2 \mid \ a_1 \in A_1, a_2 \in A_2\},$$

$$A_1A_2 = \{a_1b_1 + a_2b_2 \mid a_1, a_2 \in A_1, b_1, b_2 \in A_2\}$$

are ideals of R respectively called the sum and product of the ideals A_1 and A_2.

We will now present some notation and the Chinese Remainder Theorem. Let R be a ring, let $x_1, x_2 \in R$, and let I be an ideal of R. We write $x_1 \equiv x_2 \pmod{I}$ to denote $x_1 + I = x_2 + I$. Similarly, if $a \in R$, then we write $x_1 \equiv x_2 \pmod{a}$ to denote $x_1 + \langle a \rangle = x_2 + \langle a \rangle$.

Theorem 2.4. [5] *Let R be a ring with unity and let $A_1, ..., A_n$ be ideals of R such that $A_i + A_j = R$ for all $i \neq j$. Given elements $r_1, ..., r_n \in R$, there exists an element $r \in R$ such that $r \equiv r_i \pmod{A_i}$ for all $i = 1, ..., n$.*

Proof. We first prove the result for $n = 2$. Since $A_1 + A_2 = R$, we have $a_1 + a_2 = 1$ for some elements $a_1 \in A_1$ and $a_2 \in A_2$. We let $r = r_2a_1 + r_1a_2$ and the result holds.

We now prove the general result. For each $i = 2, ..., n$, we choose elements $a_i \in A_1$ and $b_i \in A_i$ such that $a_i + b_i = 1$. We let $x = \prod_{i=2}^{n}(a_i + b_i)$ and $I = \prod_{i=2}^{n} A_i$, where $x = 1$ and $x \in A_1 + I$. Then $A_1 + I = R$, so we can find an element $y_1 \in R$ such that $y_1 \equiv 1 \pmod{A_1}$ and $y_1 \equiv 0 \pmod{I}$. We repeat this argument to find elements $y_2, ..., y_n \in R$ such that $y_j \equiv 1 \pmod{A_j}$ and $y_j \equiv 0 \pmod{A_k}$ for $k \neq j$. We then set $r = r_1y_1 + r_2y_2 + \cdots + r_ny_n$ and the result holds. □

Given a ring R, the ideals of R can be used to construct new rings.

Theorem 2.5. *Let R be a ring and let A be a subring of R. The set $R/A = \{r + A \mid r \in R\}$ is a ring under the operations $(s + A) + (t + A) = s + t + A$ and $(s + A)(t + A) = st + A$ if and only if A is an ideal of R.*

Proof. We show first that addition is well defined if A is an ideal of R. Let $s + A = s' + A$ and $t + A = t' + A$. Then $s = s' + a$ and $t = t' + b$ for some elements $a, b \in A$. We obtain

$$\begin{aligned} s' + t' + A &= s + a + t + b + A \\ &= s + a + t + A \\ &= s + a + A + t \\ &= s + A + t \\ &= s + t + A. \end{aligned}$$

We now show that multiplication is well defined if and only if A is an ideal of R. Suppose that A is an ideal and let $s + A = s' + A$ and $t + A = t' + A$. Then $s = s' + a$ and $t = t' + b$ for some elements $a, b \in A$. We obtain

$$st = (s' + a)(t' + b) = s't' + at' + s'b + ab$$

so that

$$st + A = s't' + at' + s'b + ab + A = s't' + A.$$

Thus multiplication is well defined when A is an ideal.

Next we suppose that A is a subring of R that is not an ideal of R. Then there exist elements $a \in A$ and $r \in R$ such that $ar \notin A$ or $ra \notin A$. Consider the elements $a + A = A$ and $r + A$. If $ar \notin A$, then $(a + A)(r + A) = ar + A$ but $(0 + A)(r + A) = A \neq ar + A$. We apply the same argument to the case $ra \notin A$. Therefore multiplication is not well defined when A is not an ideal.

It remains to show that the properties of a ring are satisfied. Let $a, b, c \in R$. $(a+A)+(b+A) = (a+b)+A = (b+a)+A = (b+A)+(a+A)$ so that addition is commutative. $(a+b+A)+(c+A) = (a+b+c)+A = (a+A)+(b+c+A)$ so that addition is associative. $0 + A$ is the additive identity and $-a + A$ is the additive inverse of $a+A$. $(a+A)\big((b+A)(c+A)\big) = (a+A)(bc+A) = a(bc)+A = (ab)c+A = (ab+A)(c+A) = \big((a+A)(b+A)\big)(c+A)$ so that multiplication is associative. $(a+A)(b+c+A) = a(b+c)+A = ab+ac+A = (ab+A)+(ac+A)$ and $(b+c+A)(a+A) = (b+c)a+A = ba+ca+A = (ba+A)+(ca+A)$ so that multiplication distributes over addition. This shows that the set R/A is indeed a ring. □

Given an ideal A of a ring R, it is sometimes possible to determine information about the structure of R/A if we know the structure of A. The following definition is necessary to provide the condition under which R/A is a field.

Definition 2.7. *An ideal A of a commutative ring R is a* maximal ideal *of R if it is a proper ideal of R such that, for any ideal B of R satisfying $A \subseteq B \subseteq R$, either $B = A$ or $B = R$.*

Theorem 2.6. *Let R be a commutative ring with unity and let A be an ideal of R. Then R/A is a field if and only if A is maximal.*

Proof. Suppose that R/A is a field and B is an ideal of R that properly contains A. Let $b \in B$, where $b \notin A$. Then $b + A$ is a non-zero element of R/A and there exists an element $c + A$ such that $(b + A)(c + A) = 1 + A$. Since $b \in B$, we have $bc \in B$. We obtain

$$1 + A = (b + A)(c + A) = bc + A,$$

so that $1 - bc \in A \subset B$ and thus $1 = (1 - bc) + bc \in B$. Then $B = R$ and A is maximal.

Suppose that A is maximal. Since R/A is a ring, it suffices to show that multiplication is commutative, R/A has a unity, and that all non-zero elements of R/A are units. Let $r, s \in R$, then since R is a commutative ring we have $(r + A)(s + A) = rs + A = sr + A = (s + A)(r + A)$ so that multiplication is commutative. $1 + A$ is the unity. Let $b \in R$ where $b \notin A$. It remains to show that $b+A$ has a multiplicative inverse. We consider the set $B = \{br+a \mid r \in R, a \in A\}$. B is an ideal of R that properly contains A, thus $B = R$ since A is maximal. It follows that $1 \in B$, so we write $1 = br_0 + a_0$ where $r_0 \in R$ and $a_0 \in A$. We then have

$$1 + A = br_0 + a_0 + A = br_0 + A = (b + A)(r_0 + A)$$

so that $r_0 + A$ is the desired multiplicative inverse of $b + A$. □

While the structure of a ring R can be quite complicated, the ideals of R can provide some information about the structure of R since they are subsets of the ring R itself. Another useful method to determine information about R is to compare R to other rings with similar structures. This motivates the following definition.

Definition 2.8. *A* ring homomorphism ϕ *from a ring R to a ring S is a mapping from R to S such that*

$$\phi(a+b) = \phi(a) + \phi(b) \quad \textit{and} \quad \phi(ab) = \phi(a)\phi(b)$$

for all $a, b \in R$. A ring homomorphism that is also a bijection is called a ring isomorphism*.*

Given a ring homomorphism ϕ from a ring R to a ring S, the following subset of R is of particular importance.

Definition 2.9. *Let ϕ be a ring homomorphism from a ring R to a ring S, then we define*

$$\ker \phi = \{r \in R \mid \phi(r) = 0\}.$$

Lemma 2.2. *Let ϕ be a homomorphism from a ring R to a ring S. Then the following properties hold:*

i) $\phi(0) = 0$*;*
ii) $\phi(-a) = -\phi(a)$ *for all* $a \in R$*.*
iii) $\phi(a) = \phi(b)$ *if and only if* $a + \ker \phi = b + \ker \phi$ *for all* $a, b \in R$*.*

Proof. We have $0 = \phi(0) - \phi(0) = \phi(0+0) - \phi(0) = \phi(0) + \phi(0) - \phi(0) = \phi(0)$, so that the first property holds. We let $a \in R$, then $\phi(a) - \phi(a) = 0 = \phi(0) = \phi(a-a) = \phi(a) + \phi(-a)$. Thus $-\phi(a) = \phi(-a)$ and the second property holds. To prove the third, we let $a, b \in \ker \phi$. Then

$$\begin{aligned}
\phi(a) = \phi(b) &\Leftrightarrow 0 = \phi(a) - \phi(b) \\
&\Leftrightarrow 0 = \phi(a) + \phi(-b) = \phi(a-b) \\
&\Leftrightarrow a - b \in \ker \phi \\
&\Leftrightarrow a - b + \ker \phi = \ker \phi \\
&\Leftrightarrow a + \ker \phi = b + \ker \phi.
\end{aligned}$$

□

Lemma 2.3. *Let ϕ be an onto homomorphism from an integral domain R to an integral domain S. Then the following properties hold:*

i) $\phi(1) = 1$*;*
ii) $\phi(a^{-1}) = \big(\phi(a)\big)^{-1}$ *for all units* $a \in R$ *such that* $\phi(a) \neq 0$*.*

Proof. Suppose that R and S are integral domains and that ϕ is onto. If $\phi(1) = 0$, then $\phi(r) = \phi(r \cdot 1) = \phi(r)\phi(1) = \phi(r) \cdot 0 = 0$ for all $r \in R$. Thus $\phi(R) = \{0\} = S$ since ϕ is onto. This contradicts S being an integral domain, since an integral domain must have a unity. Thus $\phi(1) \neq 0$. Then $1 \cdot \phi(1) = \phi(1) = \phi(1 \cdot 1) = \phi(1)\phi(1)$ so that $\phi(1) = 1$, proving the first property. Next, we let $a \in R$ be a unit

such that $\phi(a) \neq 0$. Then there exists $a^{-1} \in R$ and we have $\phi(a)\big(\phi(a)\big)^{-1} = 1 = \phi(1) = \phi(aa^{-1}) = \phi(a)\phi(a^{-1})$ so that $\phi(a^{-1}) = \big(\phi(a)\big)^{-1}$. □

The following result regarding the kernel of a ring homomorphism links the notions of ring homomorphisms and ideals.

Theorem 2.7. *Let ϕ be a ring homomorphism from a ring R to a ring S, then* $\ker\phi$ *is an ideal of R.*

Proof. $\ker\phi$ is a non-empty subset of R since $\phi(0) = 0$. For all $a, b \in \ker\phi$ we have $\phi(a-b) = \phi(a) - \phi(b) = 0 - 0 = 0$ so that $a - b \in \ker\phi$. For all $a \in \ker\phi$ and all $r \in R$ we have $\phi(ra) = \phi(r)\phi(a) = \phi(r)\cdot 0 = 0$ and $\phi(ar) = \phi(a)\phi(r) = 0\cdot\phi(r) = 0$ so that $ra, ar \in \ker\phi$. Therefore $\ker\phi$ is an ideal of R. □

Given an ideal A of a ring R, it is of interest to observe the structure of R/A. The following important result, the *First Isomorphism Theorem for Rings*, considers the ring R/A where the ideal A is the kernel of a ring homomorphism.

Theorem 2.8. *Let ϕ be a ring homomorphism from R to S. Then the mapping from $R/\ker\phi$ to $\phi(R)$ given by $r + \ker\phi \mapsto \phi(r)$ is an isomorphism.*

Proof. Let $\psi : R/\ker\phi \to \phi(R)$ be the mapping given by $r + \ker\phi \mapsto \phi(r)$. It follows from the third property of Lemma 5.1 that ψ is well-defined and one-to-one. For every $\phi(r) \in \phi(R)$, the element $r + \ker\phi \in R/\ker\phi$ is such that $\psi(r+\ker\phi) = \phi(r)$ so that ψ is also onto. It only remains to show that ψ preserves the operations of addition and multiplication:

$$\begin{aligned}\psi\big((x+\ker\phi)+(y+\ker\phi)\big) &= \psi\big((x+y)+\ker\phi\big)\\ &= \phi(x+y)\\ &= \phi(x)+\phi(y)\\ &= \psi\big(x+\ker\phi\big)+\psi\big(y+\ker\phi\big)\end{aligned}$$

and

$$\begin{aligned}\psi\big((x+\ker\phi)(y+\ker\phi)\big) &= \psi\big((xy)+\ker\phi\big)\\ &= \phi(xy)\\ &= \phi(x)\phi(y)\\ &= \psi\big(x+\ker\phi\big)\psi\big(y+\ker\phi\big).\end{aligned}$$

Therefore ψ is an isomorphism from $R/\ker\phi$ to $\phi(R)$. □

Ring isomorphisms can also be useful in the construction of fields. We begin with the following definition of an equivalence relation, which generalizes the concept of equality.

Definition 2.10. *An* equivalence relation *on a set S is a set $\sim$ of ordered pairs of elements of S such that the following properties hold:*

i) $(a, a) \in\sim$ *for all* $a \in S$*;*
ii) $(a, b) \in\sim$ *implies* $(b, a) \in\sim$*;*

iii) $(a,b) \in\sim$ *and* $(b,c) \in\sim$ *imply that* $(a,c) \in\sim$.

When $\sim$ *is an equivalence relation on a set* S*, we write* $a \sim b$ *to denote* $(a,b) \in\sim$*, and the set* $[a] = \{x \in S \mid x \sim a\}$ *is called the* equivalence class *of* S *containing* a*.*

The following result now allows us to construct a field F, starting from an integral domain R, in such a way that a ring isomorphic to R exists in F.

Theorem 2.9. *Let* R *be an integral domain. Then there exists a field* F*, called the* field of fractions of R*, that contains a subring isomorphic to* R*.*

Proof. Let $S = \{(a,b) \mid a,b \in R,\ b \neq 0\}$. We define an equivalence relation $\sim$ on S by $(a,b) \sim (c,d)$ if $ad - bc = 0$. We let F be the set of equivalence classes under the relation $\sim$ and denote the equivalence class that contains (a,b) by a/b. We define addition and multiplication on F by

$$a/b + c/d = (ad + bc)/(bd) \quad \text{and} \quad a/b \cdot c/d = (ac)/(bd).$$

Since R is an integral domain, we have $bd \neq 0$ when $b \neq 0$ and $d \neq 0$, thus addition and multiplication are closed.

We must now show that the operations are well defined. Suppose that $a/b = a'/b'$ and $c/d = c'/d'$ so that $ab' = a'b$ and $cd' = c'd$. Then

$$\begin{aligned}(ad + bc)b'd' &= adb'd' + bcb'd' \\ &= (ab')dd' + (cd')bb' \\ &= (a'b)dd' + (c'd)bb' \\ &= a'd'bd + b'c'bd \\ &= (a'd' + b'c')bd,\end{aligned}$$

so that $(ad + bc)/(bd) = (a'd' + b'c')/(b'd')$. Therefore addition is well defined. Similarly, we have

$$\begin{aligned}(ac)(b'd') &= (ab')(cd') \\ &= (a'b)(c'd) \\ &= (a'c')(bd),\end{aligned}$$

so that $(ac)/(bd) = (a'c')/(b'd')$. Therefore multiplication is well defined. Since multiplication is commutative in R, it follows that multiplication is also commutative in F. Let 1 denote the unity of R, then $0/1$ is the additive identity of F. The additive inverse of a/b is $-a/b$, and the multiplicative inverse of a non-zero element a/b is b/a.

We let the mapping $\phi : R \to F$ be given by $x \mapsto x/1$. For all $x, y \in R$, $\phi(x) = \phi(y)$ implies $x/1 = y/1$ so that $x \cdot 1 = y \cdot 1$ and $x = y$. We also have $\phi(x + y) = (x + y)/1 = (x \cdot 1 + y \cdot 1)/(1 \cdot 1) = x/1 + y/1 = \phi(x) + \phi(y)$ and $\phi(xy) = (xy)/1 = (xy)/(1 \cdot 1) = (x/1)(y/1) = \phi(x)\phi(y)$ for all $x, y \in R$. Then ϕ is one-to-one and it preserves both operations, so that ϕ is a ring isomorphism from R to $\phi(R)$. $\square$

Given a ring R, we have shown that it is possible to construct new rings by considering the ideals of R or the field of fractions of R. It is also possible to construct new rings by considering the polynomials with coefficients in R.

Definition 2.11. *Let R be a commutative ring. The set of formal sums*

$$R[x] = \left\{ \sum_{i=0}^{n} a_i x^i \mid a_i \in R,\ n \text{ is a non-negative integer} \right\}$$

is called the ring of polynomials over R in the indeterminate x. *The two elements*

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

of $R[x]$ are considered equal if and only if $a_i = b_i$ for all non-negative integers i, where we define $a_i = 0$ when $i > n$ and $b_i = 0$ when $i > m$. We define addition and multiplication by

$$f(x) + g(x) = \sum_{i=0}^{\max\{n,m\}} (a_i + b_i)x^i$$

and

$$f(x)g(x) = \sum_{i=0}^{n+m} c_i x^i \quad \textit{where} \quad c_j = \sum_{k=0}^{j} a_k b_{j-k}$$

for $j = 0, ..., n + m$.

If $f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$ where $a_n \neq 0$ and $f(x) \neq 0$, we say that $f(x)$ has *degree* n and denote this by $\deg f = n$. The term a_n is called the *leading coefficient* of $f(x)$, and if the leading coefficient is the unity of R, we say that $f(x)$ is *monic*.

The polynomial ring $R[x]$ can share some properties with the ring R. The following result is an important example of one such property.

Theorem 2.10. *If R is an integral domain, then $R[x]$ is an integral domain.*

Proof. We must show that the ring $R[x]$ is commutative, possesses a unity, and has no zero-divisors. $R[x]$ is commutative since R is commutative. If 1 is the unity of R, then $f(x) = 1$ is the unity of $R[x]$. Suppose that

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

where $a_n \neq 0$ and $b_m \neq 0$. Then $f(x)g(x)$ has leading coefficient $a_n b_m \neq 0$ since R is an integral domain. □

Another very useful property, frequently called the division algorithm, can be observed for polynomial rings $F[x]$ where F is a field. We present first the Division Algorithm for $\mathbb{Z}$, and we then consider the similar result for $F[x]$.

Theorem 2.11. *Let $a, b \in \mathbb{Z}$ where $b > 0$. Then there exists unique integers q and r, respectively called the* quotient *and* remainder *in the division of a by b, such that $a = qb + r$ and $0 \leq r < b$.*

Proof. Let $S = \{a - bk \mid k \in \mathbb{Z} \text{ and } a - bk \geq 0\}$. If $0 \in S$, then the result holds with $q = a/b$ and $r = 0$. Assume then that $0 \notin S$. Since S is non-empty, there exists a smallest element $r = a - qb$ of S. Then $a = qb + r$ and $r \geq 0$. If $r \geq b$, then $0 \leq r - b = a - qb - b = a - b(q + 1) \in S$ where $a - b(q + 1) < a - bq$. This contradicts r being the smallest element of S, thus $r < b$.

Let q_1, q_2 and r_1, r_2 be integers such that $a = q_1 b + r_1$ and $a = q_2 b + r_2$ where $0 \leq r_1, r_2 < b$. Assume without loss of generality that $r_1 \leq r_2$, then $r_2 - r_1 = (q_1 - q_2)b$ so that b divides $r_2 - r_1$ and $0 \leq r_2 - r_1 \leq r_2 < b$. It follows that $r_2 - r_1 = 0$ so that $r_1 = r_2$ and $q_1 = q_2$. □

We now state and prove the Division Algorithm for $F[x]$.

Theorem 2.12. *Let F be a field and let $f(x), g(x) \in F[x]$ where $g(x) \neq 0$. Then there exist unique polynomials $q(x), r(x) \in F[x]$, respectively called the* quotient *and the* remainder *in the division of $f(x)$ by $g(x)$, such that $f(x) = q(x)g(x) + r(x)$ and either $\deg r(x) < \deg g(x)$ or $r(x) = 0$.*

Proof. Let $n = \deg f$ and $m = \deg g$. If $f(x) = 0$ or $n < m$, we set $q(x) = 0$, and $r(x) = f(x)$. We now assume that $n \geq m$, and let $f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$ and $g(x) = b_m x^m + b_{m-1} x^{m-1} + \cdots + b_1 x + b_0$. We prove the existence of the polynomials $q(x)$ and $r(x)$ by induction.

We first prove the basis step. If $n = 0$, then $m = 0$ so we set $q(x) = a_0 b_0^{-1}$ and $r(x) = 0$.

Assume now that the result holds when we divide all polynomials in $F[x]$ of degree less than n by $g(x)$. We let $p(x) = f(x) - a_n {b_m}^{-1} x^{n-m} g(x)$, then $p(x) = 0$ or $\deg p < n$. If either $p(x) = 0$ or $\deg p < \deg g$, we set $q(x) = a_n {b_m}^{-1} x^{n-m}$, and $r(x) = p(x)$ so that

$$f(x) = q(x)g(x) + r(x).$$

Then $q(x), r(x) \in F[x]$ where $r(x) = 0$ or $\deg r < \deg g$ as required. If $\deg g \leq \deg p < n$, then by the induction hypothesis there exist polynomials $\overline{q}(x)$ and $\overline{r}(x)$ in $F[x]$ such that

$$p(x) = \overline{q}(x)g(x) + \overline{r}(x),$$

where either $\overline{r}(x) = 0$ or $\deg \overline{r} < \deg g$. We obtain

$$\begin{aligned} f(x) &= a_n {b_m}^{-1} x^{n-m} g(x) + p(x) \\ &= a_n {b_m}^{-1} x^{n-m} g(x) + \overline{q}(x)g(x) + \overline{r}(x) \\ &= \big(a_n {b_m}^{-1} x^{n-m} + \overline{q}(x)\big)g(x) + \overline{r}(x). \end{aligned}$$

We set $q(x) = a_n {b_m}^{-1} x^{n-m} + \overline{q}(x)$ and $r(x) = \overline{r}(x)$, so that $q(x), r(x) \in F[x]$ and either $r(x) = 0$ or $\deg r < \deg g$ as required.

To prove the uniqueness of the polynomials $q(x)$ and $r(x)$, we suppose that $f(x) = q_1(x)g(x) + r_1(x)$ and $f(x) = q_2(x)g(x) + r_2(x)$ where $r_1(x) = 0$ or $\deg r_1 < \deg g$ and $r_2(x) = 0$ or $\deg r_2 < \deg g$. We then have $r_2(x) - r_1(x) = \big(q_1(x) - q_2(x)\big)g(x)$. Thus either $r_2(x) - r_1(x) = 0$ or $\deg\big(r_2(x) - r_1(x)\big) \geq \deg g(x)$. Since $\deg\big(r_2(x) - r_1(x)\big) \leq \max\{\deg r_1(x), \deg r_2(x)\} < \deg g(x)$, we must have $r_2(x) - r_1(x) = 0$ so that $r_1(x) = r_2(x)$ and $q_1(x) = q_2(x)$. □

One of the most important aspects of a polynomial are its zeros. We make the following definitions.

Definition 2.12. *Let F be a field and let $f(x) \in F[x]$. An element $a \in F$ is called* a zero of multiplicity k, *where $k \geq 1$, if $(x-a)^k$ is a factor of $f(x)$ but $(x-a)^{k+1}$ is not.*

Let $F(x)$ be the field of fractions of $F[x]$. Let $f(x) \in F(x)$, then $f(x) = f_1(x)/f_2(x)$ for some $f_1(x), f_2(x) \in F[x]$. The zeros of $f_2(x)$ are called the poles *of $f(x)$, and the set F excluding the poles of $f(x)$ is called the* domain of definition *of $f(x)$.*

The Division Algorithm for $F[x]$ allows us to obtain the following results regarding the zeros of a polynomial.

Corollary 2.1. *Let F be a field, $a \in F$, and $f(x) \in F[x]$. Then $f(a)$ is the remainder in the division of $f(x)$ by $x - a$.*

Proof. By the division algorithm for $F[x]$, there exist unique polynomials $q(x), r(x) \in F[x]$ such that $f(x) = q(x)(x-a) + r(x)$ and either $r(x) = 0$ or $\deg r(x) < 1$. It follows that $r(x)$ must be a constant, so that $f(x) = q(x)(x-a) + b$ for some $b \in F$. Then $f(a) = q(a)(a-a) + b = b$ so that $f(x) = q(x)(x-a) + f(a)$. □

Corollary 2.2. *Let F be a field, $a \in F$, and $f(x) \in F[x]$. Then a is a zero of $f(x)$ if and only if $x - a$ is a factor of $f(x)$.*

Proof. Suppose that a is a zero of $f(x)$. Since $f(a)$ is the remainder in the division of $f(x)$ by $x - a$, there exists a unique polynomial $q(x) \in F[x]$ such that $f(x) = q(x)(x-a) + f(a)$. Then $f(x) = q(x)(x-a)$ since $f(a) = 0$, so that $(x-a)$ is a factor of $f(x)$.

Suppose that $x - a$ is a factor of $f(x)$. Then $f(x) = g(x)(x-a)$ for some polynomial $g(x) \in F[x]$ and it follows that $f(a) = g(a)(a-a) = 0$ so that a is a zero of $f(x)$. □

The division algorithms for $\mathbb{Z}$ and $F[x]$ can also be used to determine properties of the ideals of these two rings. We present a definition which motivates the results that follow.

Definition 2.13. *A* principal ideal domain *is an integral domain R in which every ideal is a principal ideal.*

Theorem 2.13. *$\mathbb{Z}$ is a principal ideal domain.*

Proof. Let I be an ideal of $\mathbb{Z}$. If $I = \{0\}$ then $I = \langle 0 \rangle$. If $I \neq \{0\}$, let n be the smallest positive integer in I. We have $\langle n \rangle \subseteq I$ since $n \in I$. Let $m \in I$. There exist integers q and r such that $m = qn + r$ where $r < n$. Since I is an ideal, we have $r = m - qn \in I$ where n is the smallest positive integer in I, thus we must have $r = 0$. Then $m = qn$ so that $I \subseteq \langle n \rangle$. Therefore $I = \langle n \rangle$. □

Theorem 2.14. *Let F be a field, then $F[x]$ is a principal ideal domain. Moreover, for any nonzero ideal I in $F[x]$ and any element $g(x)$ of $F[x]$, $I = \langle g(x) \rangle$ if and only if $g(x)$ is a nonzero polynomial of minimum degree in I.*

Proof. Let I be an ideal of the integral domain $F[x]$. If $I = \{0\}$, then $I = \langle 0 \rangle$. If $I \neq \{0\}$, let $g(x) \in I$ be an element of minimum degree among all of the elements of I. We have $\langle g(x) \rangle \subseteq I$ since $g(x) \in I$. Let $f(x) \in I$. By the division algorithm, we write $f(x) = q(x)g(x) + r(x)$ where either $\deg r(x) < \deg g(x)$ or $r(x) = 0$. Since $r(x) = f(x) - q(x)g(x) \in I$, we conclude that $r(x) = 0$ since $g(x) \in I$ was chosen to have minimal degree. Thus $f(x) \in \langle g(x) \rangle$ and $I \subseteq \langle g(x) \rangle$. Therefore $I = \langle g(x) \rangle$. □

In integral domains, a prime is always an irreducible. In principal ideal domains, the notions of prime and irreducible coincide.

Theorem 2.15. *Let R be a principal ideal domain. Then $a \in R$ is prime if and only if it is irreducible.*

Proof. Since R is an integral domain, it follows that a is irreducible if it is prime. To show the converse, we suppose that a is an irreducible element of R such that a divides bc. We consider the ideal $I = \{ax + by \mid x, y \in R\}$, where $I = \langle r \rangle$ for some element $r \in R$ since R is a principal ideal domain. Since $a \in I$ is irreducible, we have $a = rs$ where either r or s is a unit. If r is a unit, then $I = R$ and $1 = ax + by$ for some $x, y \in R$ so that $c = acx + bcy$. Then a divides c since it divides acx and bcy. If s is a unit, then $\langle a \rangle = \langle r \rangle = I$. Since $b \in I$, there exists $t \in R$ such that $b = at$, thus a divides b. Therefore a is prime. □

We have seen that maximal ideals are useful in the construction of fields. For polynomial rings over fields, we have the following characterization for maximal ideals.

Theorem 2.16. *Let F be a field and let $p(x) \in F[x]$. Then $\langle p(x) \rangle$ is a maximal ideal in $F[x]$ if and only if $p(x)$ is irreducible over F.*

Proof. Suppose that $\langle p(x) \rangle$ is a maximal ideal in $F[x]$. Since neither $\{0\}$ nor $F[x]$ is a maximal ideal in $F[x]$, $p(x)$ is neither the zero polynomial nor a unit in $F[x]$. If $p(x) = g(x)h(x)$ is a factorization of $p(x)$ over F, then $\langle p(x) \rangle \subseteq \langle g(x) \rangle \subseteq F[x]$. Thus $\langle p(x) \rangle = \langle g(x) \rangle$ or $F[x] = \langle g(x) \rangle$. The first case yields $\deg p(x) = \deg g(x)$. The second case yields $\deg g(x) = 0$ and thus $\deg p(x) = \deg h(x)$. Therefore $p(x)$ is irreducible over F.

Suppose that $p(x)$ is irreducible over F, and let I be any ideal of $F[x]$ such that $\langle p(x) \rangle \subseteq I \subseteq F[x]$. Since $F[x]$ is a principal ideal domain, $I = \langle g(x) \rangle$ for some

$g(x) \in F[x]$. Then $p(x) \in \langle g(x) \rangle$ and $p(x) = g(x)h(x)$, where $h(x) \in F[x]$. Since $p(x)$ is irreducible over F, either $g(x)$ or $h(x)$ is constant. If $g(x)$ is constant, then $I = F[x]$. If $h(x)$ is constant, then $\langle p(x) \rangle = \langle g(x) \rangle = I$. Therefore $\langle p(x) \rangle$ is maximal in $F[x]$. □

Corollary 2.3. *Let F be a field and let $p(x)$ be an irreducible polynomial over F, then $F[x]/\langle p(x) \rangle$ is a field*

Proof. If $p(x)$ is an irreducible polynomial over F, then $\langle p(x) \rangle$ is a maximal ideal of $F[x]$ and $F[x]/\langle p(x) \rangle$ is a field. If $F[x]/\langle p(x) \rangle$ is a field, then $\langle p(x) \rangle$ is a maximal ideal of $F[x]$ and $p(x)$ must be an irreducible polynomial over F. □

We now present the definition for one last important class of rings.

Definition 2.14. *An integral domain R is a* unique factorization domain *if*

i) every non-zero element of R that is not a unit can be written as a product of irreducible elements of R;

ii) the factorization into irreducible elements is unique up to associates and the order in which the factors are written.

It can be shown that principal ideal domains are unique factorization domains. To this end, we prove here the ascending chain condition for ideals in a principal ideal domain.

Lemma 2.4. *In a principal ideal domain, any strictly increasing chain of ideals $I_1 \subset I_2 \subset \cdots$ must be finite in length.*

Proof. Let $I_1 \subset I_2 \subset \cdots$ be a chain of strictly increasing ideals of an integral domain R, and let $I = \bigcup I_k$ be the union of the ideals I_k over every index k. We claim that I is an ideal of R. Indeed, $0 \in I_k$ for every index k, thus I is a nonempty subset of R. For all $a, b \in I$, there exist ideals I_{k_1} and I_{k_2} in the given chain such that $a \in I_{k_1}$ and $b \in I_{k_2}$. We let $k_0 = \max\{k_1, k_2\}$ so that $a, b \in I_{k_0}$. Then for all $r \in R$, since I_{k_0} is an ideal of R we have $a - b \in I_{k_0} \subseteq I$ and $ra, ar \in I_{k_0} \subseteq I$. Thus I is an ideal of R.

Since R is a principal ideal domain, there exists an element $a \in R$ such that $I = \langle a \rangle$. Then $a \in I = \bigcup I_k$, so that $a \in I_n$ for some ideal I_n of the chain. Then for any ideal I_m of the chain, we have $I_m \subseteq I = \langle a \rangle \subseteq I_n$ so that I_n is the last ideal of the chain. □

We are now able to prove the claim that principal ideal domains are unique factorization domains. Two important examples of unique factorization domains follow the result.

Theorem 2.17. *Every principal ideal domain is a unique factorization domain.*

Proof. Let R be a principal ideal domain and let a_1 be a nonzero element of R which is not a unit. We first show that a_1 has an irreducible factor. If a_1 is irreducible, then a_1 is an irreducible factor of a_1. We now assume that $a_1 = b_2 a_2$, where a_2 is nonzero and neither a_2 nor b_2 is a unit. If a_2 is not irreducible, we

write $a_2 = b_3a_3$ where a_3 is nonzero and neither a_3 nor b_3 is a unit. We repeat this process to obtain the sequence $b_2, b_3, ...$ of elements which are not units of R and the sequence $a_1, a_2, ...$ of nonzero elements of R where $a_i = b_{i+1}a_{i+1}$ for each i. If $a_i \in \langle b_i a_i \rangle$, then $a_i = xb_i a_i$ for some $x \in R$. Since $a_i \neq 0$, we obtain $1 = xb_i$ so that b_i is a unit. We conclude that for each i, we have the strict inclusion $\langle b_i a_i \rangle \subset \langle a_i \rangle$. Then $\langle a_1 \rangle \subset \langle a_2 \rangle \subset \cdots$ is a strictly increasing chain of ideals, which must be finite by the ascending chain condition. We let $\langle a_1 \rangle \subset \langle a_2 \rangle \subset \cdots \subset \langle a_j \rangle$ be the finite chain, so that a_j is an irreducible factor of a_1.

We next show that a_1 is a product of irreducible factors. We let $a_1 = p_2c_2$ where p_2 is irreducible and c_2 is not a unit. If c_2 is not irreducible, then we write $c_2 = p_3c_3$ where p_3 is irreducible and c_3 is not a unit. We repeat this process to obtain another strictly increasing sequence $\langle a_1 \rangle \subset \langle c_2 \rangle \subset \langle c_3 \rangle \subset \cdots$ which must be finite by the ascending chain condition. We let $\langle c_k \rangle$ be the last ideal of the chain, then c_k is irreducible and $a_1 = p_2p_3 \cdots p_kc_k$ where p_i is irreducible for $i = 2, ..., k$. Therefore every nonzero element of R which is not a unit is a product of irreducible elements of R.

We now show that the factorization is unique up to associates and the order in which the factors are written. We let $a \in R$ and write

$$a = p_1p_2 \cdots p_n = q_1q_2 \cdots q_m,$$

where p_i and q_j are irreducible elements of R for $i = 1, ..., n$ and $j = 1, ..., m$. We apply induction on n. If $n = 1$, then a is irreducible and $m = 1$ so that $p_1 = q_1$. We now assume that the factorization of an element into a product of less than n irreducible elements is unique up to associates and the order in which the factors are written. Since p_1 divides $q_1q_2 \cdots q_m$, p_1 must divide an element q_{m_0} for some $1 \leq m_0 \leq m$. Suppose without loss of generality that $m_0 = 1$, then $q_1 = up_1$ for some unit $u \in R$. Now

$$up_1p_2 \cdots p_n = uq_1q_2 \cdots q_m = q_1uq_2 \cdots q_m$$

so that

$$p_2 \cdots p_n = uq_2 \cdots q_m.$$

By the induction hypothesis, these two factorizations are identical up to associates and the order in which the elements are written. Therefore the two factorizations of a are also identical up to associates and the order in which the factors are written. □

Corollary 2.4. *$\mathbb{Z}$ is a unique factorization domain.*

Corollary 2.5. *Let F be a field, then $F[x]$ is a unique factorization domain.*

The following definition can be made for elements in unique factorization domains.

Definition 2.15. *Let R be a unique factorization domain and let $a, b \in R$. We call an element $d \in R$ a* common divisor *of a and b if d is a divisor of both a and b. Let S be the set of all common divisors of a and b, then we define a* greatest

common divisor *of a and b, denoted* $\gcd(a, b)$*, to be an element of S such that d divides* $\gcd(a, b)$ *for all $d \in S$. If* $\gcd(a, b) = 1$*, we say that a and b are* relatively prime.

Relatively prime elements of a principal ideal domain possess the following property.

Theorem 2.18. *Let R be a principal ideal domain. If* $\gcd(a, b) = 1$ *for $a, b \in R$, then there exist $s, t \in R$ such that $sa + tb = 1$.*

Proof. Let $I = \langle a \rangle + \langle b \rangle$. Since R is a principal ideal domain, there exists $c \in R$ such that $I = \langle c \rangle = \langle a \rangle + \langle b \rangle$. Thus c must divide both a and b where a and b are relatively prime. The element c must then be a unit in R. Since $c \in I$, there exist $s', t' \in R$ such that $s'a + t'b = c$. Setting $s = c^{-1}s'$ and $t = c^{-1}t'$, we obtain $sa + tb = 1$. □

2.2. **Vector Spaces.**

We recall the axioms of a vector space. Among other results, we establish the concept of the dimension of a vector space, found in [4], which is useful in the discussion of fields.

Definition 2.16. *A* vector space *V over a field F is a set with two binary operations called addition and scalar multiplication, denoted by $v+u$ and av respectively, such that the following properties hold for all $v, u, w \in V$ and all $a, b \in F$:*

i) $v + u = u + v$;
ii) $(v + u) + w = v + (u + w)$;
iii) there exists an element $0 \in V$ *such that* $v + 0 = v$ *for all* $v \in V$;
iv) for each $v \in V$ *there exists an element* $-v \in V$ *such that* $v + (-v) = 0$;
v) $1v = v$ *for all* $v \in V$*, where* 1 *is the unity of F;*
vi) $(ab)v = a(bv)$;
vii) $a(v + u) = av + au$;
viii) $(a + b)v = av + bv$.

The following are two important examples of vector spaces.

Example 2.1. *Let F be a field. The set*

$$F^n = \{(a_1, ..., a_n) \mid a_1, ..., a_n \in F\}$$

is a vector space over F with the operations of component-wise addition and scalar multiplication.

Example 2.2. *Let n be a non-negative integer and let F be a field. The set $P_n(F)$ consisting of the zero polynomial and all polynomials of degree less than or equal to n with coefficients from F is a vector space over F with the standard operations of addition and scalar multiplication.*

Example 2.3. *Let E be a field and let F be a subfield of E. Then E is a vector space over F. The operations of addition and scalar multiplication are respectively the operations of addition and multiplication of E.*

As with rings and fields, it will be useful to consider subsets of vector spaces which are themselves vector spaces.

Definition 2.17. *Let V be a vector space over a field F. A subset W of V is called a* subspace *of V if W is a vector space over F with the operations of V.*

The following result provides a way to determine when a subset of a vector space is a subspace without verifying all of the axioms of a vector space.

Theorem 2.19. *Let V be a vector space over a field F and let W be a subset of V. Then W is a subspace of V if and only if the following hold:*

i) $0 \in W$;
ii) $v_1 + v_2 \in W$ *for all* $v_1, v_2 \in W$;
iii) $cv \in W$ *for all* $v \in W$ *and all* $c \in F$.

Proof. If W is a subspace of V, then W is a vector space with the operations of V. There must exist a vector $0_W \in W$ such that $v + 0_W = v$ for all $v \in W$, and since $v + 0 = v = v + 0_W$ for all $v \in W$, it follows that $0_W = 0 \in W$.

Next, we must show that W is a subspace of V when the given conditions hold. Since the operations of W are the same as the operations of the vector space V, we know that addition in W is commutative and associative, and that scalar multiplication in W is associative and distributes over addition. We are given that the zero vector is in W and we also have $1v = v$ for all $v \in W$. It only remains to show that for every vector $v \in W$, there exists a vector $-v \in W$ such that $v + (-v) = 0$. We have $(-1)v \in V$ and $v + (-1)v = \big((1 + (-1)\big)v = 0$ so that $(-1)v$ is the desired additive inverse. Thus W is a subspace of V. □

Of particular importance are the notions of spanning sets and linear independence. We present first the definition of the span of a subset S of a vector space V, and we prove that such a set is a subspace of V.

Definition 2.18. *Let S be a non-empty subset of a vector space V. The* span *of S, denoted span(S), is the set of all linear combinations of the vectors of S. We define the span of the empty set to be $\{0\}$. We say that the subset S* generates *V if span(S) $= V$.*

Theorem 2.20. *Let V be a vector space over a field F and let S be a subset of V. Then span(S) is a subspace of V. Any subspace of V that contains S must also contain span(S).*

Proof. If S is empty, then $\{0\}$ is a subspace of the vector space V. If S is non-empty, then there exists $v \in S$ so that $0v = 0 \in \text{span}(S)$. Let $v_1, v_2 \in \text{span}(S)$, then there exist $u_1, ..., u_n, w_1, ..., w_m \in S$ and $a_1, ..., a_n, b_1, ..., b_m \in F$ such that $v_1 = a_1u_1 + \cdots a_nu_n$ and $v_2 = b_1w_1 + \cdots b_mw_m$. Then $v_1 + v_2$ is a linear combination

of vectors in S and cv_1 is a linear combination of vectors in S for any $c \in F$. Thus $v_1 + v_2, cv_1 \in \text{span}(S)$ and $\text{span}(S)$ is a subspace of V.

Let W be a subspace of V such that $S \subseteq W$. If $w \in \text{span}(S)$, then there exists $v_1, ..., v_n \in S$ and $a_1, ..., a_n \in F$ such that $w = a_1v_1 + \cdots a_nv_n$. Then $v_1, ..., v_n \in W$ since $S \subseteq W$, so that $w \in W$ since W is a subspace of V. Thus $\text{span}(S) \subseteq W$. □

We now present the definition of linear independence.

Definition 2.19. *A subset S of a vector space V is called* linearly dependent *if there exists a finite number of distinct vectors $v_1, ..., v_n$ in S and scalars $a_1, ..., a_n$, not all zero, such that $a_1v_1 + \cdots + a_nv_n = 0$. We also say that the vectors of S are linearly dependent. A subset S of a vector space V that is not linearly dependent is called* linearly independent*. We also say that the vectors of S are linearly independent.*

For any set of vectors $v_1, ..., v_n$ of a vector space V, we have $a_1v_1 + \cdots + a_nv_n = 0$ for $a_1 = \cdots = a_n = 0$ and we call this the trivial representation of 0 as a linear combination of $v_1, ..., v_n$. If a set S of a vector space V is linearly dependent, there exists a non-trivial representation of 0 as a linear combination of the vectors of S. A set S of a vector space V is linearly independent if and only if the only representation of 0 as a linear combination of its vectors is the trivial representation.

Definition 2.20. *Let V be a vector space. A* basis β *for V is a linearly independent subset of V that generates V.*

A basis for a vector space V is a very useful concept, since we may generate all of the vectors of V with the vector of the basis in a unique way.

Theorem 2.21. *Let V be a vector space and let $\beta = \{v_1, ..., v_n\}$ be a subset of V. Then β is a basis for V if and only if each element $v \in V$ can be uniquely expressed as a linear combination of vectors of β.*

Proof. Let β be a basis for V. If $v \in V$, then $v \in \text{span}(\beta)$ since $\text{span}(\beta) = V$. Thus v is a linear combination of the vectors of β. Suppose that $v = a_1v_1 + \cdots + a_nv_n$ and $v = b_1v_1 + \cdots + b_nv_n$ are two representations of v. We obtain

$$0 = (a_1 - b_1)v_1 + \cdots + (a_n - b_n)v_n$$

where β is linearly independent. Thus $a_1 - b_1 = \cdots = a_n - b_n = 0$ and $a_1 = b_1$, ..., $a_n = b_n$, so that v can be expressed uniquely as a linear combination of the vectors of β.

Suppose that every element of V can be expressed uniquely as a linear combination of vectors of β. It follows that $V \subseteq \text{span}(\beta)$, thus it only remains to show that β is a linearly independent subset of V. Suppose that $0 = a_1v_1 + \cdots + a_nv_n$ is a representation of $0 \in V$ as a linear combination of vectors of β. One such representation is obtained by taking $a_1 = \cdots = a_n = 0$, and by assumption this representation of $0 \in V$ is unique. Therefore β is linearly independent. □

The number of vectors in a basis for a vector space is unique.

Theorem 2.22. *Let $\beta = \{v_1, ..., v_n\}$ and $\gamma = \{w_1, ..., w_m\}$ be two bases of a vector space V over a field F. Then $n = m$.*

Proof. Suppose that $n \neq m$. Assume without loss of generality that $n < m$ and consider the set $\{w_1, v_1, v_2, ..., v_n\}$. Since the elements of β span V, w_1 is a linear combination of $v_1, ..., v_n$. Let $w_1 = a_1v_1 + \cdots + a_nv_n$ where $a_i \in F$ for $i = 1, ..., n$. There exists at least one $a_i \neq 0$ from $a_1, ..., a_n$. Assume without loss of generality that $a_1 \neq 0$, then

$$v_1 = \frac{1}{a_1}w_1 + \frac{(-a_2)}{a_1}v_2 + \cdots + \frac{(-a_n)}{a_1}v_n$$

so that v_1 is a linear combination of $w_1, v_2, ..., v_n$. It follows that the set $\{w_1, v_2, ..., v_n\}$ spans V.

Next we consider the set $\{w_1, w_2, v_2, ..., v_n\}$. Since $w_2 \in V$ and the set $\{w_1, v_2, ..., v_n\}$ spans V, w_2 must be a linear combination of $w_1, v_2, ..., v_n$. Let $w_2 = b_1w_1 + b_2v_2 + \cdots + b_nv_n$ where $b_i \in F$ for $i = 1, ..., n$. Since the elements of γ are linearly independent, there must exist at least one $b_i \neq 0$ from the elements $b_2, ..., b_n$. Assume without loss of generality that $b_2 \neq 0$, then $\{w_1, w_2, v_3, ..., v_n\}$ spans V.

Repeating the argument, it follows that $\{w_1, ..., w_n\}$ spans V, thus $w_{n+1} \in V$ is a linear combination of $w_1, ..., w_n$. This contradicts the set γ being linearly independent. □

Since the number of vectors in a basis for a vector space is independent of the choice of basis, we have motivation for the following definition.

Definition 2.21. *We say that a vector space V is* finite-dimensional *if it has a basis consisting of a finite number of vectors. The unique number of vectors in a basis for V is called the* dimension *of V and is denoted by* $\dim V$. *The empty set is a basis for the vector space $\{0\}$, and we say that this vector space has dimension 0.*

We briefly recall the Lagrange interpolation formula. Let $c_0, c_1, ..., c_n$ be distinct scalars in an infinite field F and define the Lagrange polynomials

$$f_i(x) = \prod_{k \neq i} \frac{x - c_k}{c_i - c_k}.$$

Each $f_i(x)$ is a polynomial in $P_n(F)$, and as a function $f_i : F \to F$ we have

$$f_i(c_j) = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j. \end{cases}$$

Suppose that, for some scalars $a_0, a_1, ..., a_n$,

$$\sum_{i=0}^{n} a_i f_i = 0,$$

where 0 is the zero function. Then

$$\sum_{i=0}^{n} a_i f_i(c_j) = 0 \qquad \text{and} \qquad \sum_{i=0}^{n} a_i f_i(c_j) = a_j$$

for $j = 0, 1, ..., n$. Thus $a_j = 0$ for $j = 0, 1, ..., n$ so the set $\beta = \{f_0, f_1, ..., f_n\}$ is linearly independent. Since $P_n(F)$ has dimension $n+1$, it follows that β is a basis for $P_n(F)$. Every polynomial function g in $P_n(F)$ is then a linear combination of elements of β. Let $g = b_0 f_0 + b_1 f_1 + \cdots + b_n f_n$. Then

$$g(c_j) = \sum_{i=0}^{n} b_i f_i(c_j) = b_j$$

so that

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

is the unique representation of g as a linear combination of elements of β called the *Lagrange interpolation formula*. Given any scalars $b_0, b_1, ..., b_n$, the Lagrange interpolation formula yields the unique polynomial in $P_n(F)$ such that $g(c_j) = b_j$ for $j = 0, 1, ..., n$.

As with rings, it will be useful to introduce a notion for vector spaces similar to that of ring homomorphisms for rings.

Definition 2.22. *Let V and W be vector spaces over a field F. A function $T : V \to W$ is called a* linear transformation *from V to W if the following properties hold for all $x, y \in V$ and all $a \in F$:*

i) $T(x + y) = T(x) + T(y)$;
ii) $T(ax) = aT(x)$.

Lemma 2.5. *Let V and W be vector spaces and let $T : V \to W$ be linear. Then the following properties hold:*

i) $T(0) = 0$;
ii) $T(v_1 - v_2) = T(v_1) - T(v_2)$ *for all* $v_1, v_2 \in V$.

Proof. The first property follows from $0 = T(0) - T(0) = T(0 + 0) - T(0) = T(0) + T(0) - T(0) = T(0)$, so that $T(0) = 0$. To prove the second property, we let $v \in V$. Then $T(v) - T(v) = 0 = T(0) = T(v - v) = T(v) + T(-v)$ so that $-T(v) = T(-v)$. It then follows that for all $v_1, v_2 \in V$ we have $T(v_1 - v_2) = T(v_1) + T(-v_2) = T(v_1) - T(v_2)$. □

The kernel and image of a linear transformation are of interest, motivating the following definitions and results.

Definition 2.23. *Let V and W be vector spaces and let $T : V \to W$ be linear. The* null space *of T is the set $N(T) = \{v \in V \mid T(v) = 0\}$. The* range *of T is the set $R(T) = \{T(v) \mid v \in V\}$. We call the dimension of $N(T)$ the* nullity *of T,*

denoted by nullity(T)*, and we call the dimension of* $R(T)$ *the* rank *of* T*, denoted by rank*(T)*.*

Theorem 2.23. *Let* V *and* W *be vector spaces and let* $T : V \to W$ *be linear. Then* $N(T)$ *is a subspace of* V*, and* $R(T)$ *is a subspace of* W*.*

Proof. Since $T(0) = 0$, we have $0 \in N(T)$. Let $v_1, v_2 \in N(T)$ and let $c \in F$, then $T(v_1+v_2) = T(v_1)+T(v_2) = 0$ and $T(cv_1) = cT(v_1) = 0$. Thus $v_1+v_2, cv_1 \in N(T)$ so that $N(T)$ is a subspace of V.

Since $T(0) = 0$, we have $0 \in R(T)$. Let $w_1, w_2 \in R(T)$ and let $c \in F$. There exist $v_1, v_2 \in V$ such that $T(v_1) = w_1$ and $T(v_2) = w_2$. Then $T(v_1 + v_2) = T(v_1) + T(v_2) = w_1 + w_2$ and $T(cv_1) = cT(v_1) = cw_1$. Thus $w_1 + w_2, cw_1 \in R(T)$ so that $R(T)$ is a subspace of W. □

Theorem 2.24. *Let* V *and* W *be vector spaces, and let* $T : V \to W$ *be linear. If* $\beta = \{v_1, ..., v_n\}$ *is a basis for* V*, then*

$$R(T) = \mathit{span}\big(T(\beta)\big).$$

Proof. We have $T(v_i) \in R(T)$ for $i = 1, ..., n$. Since $R(T)$ is a subspace of the vector space W, it follows that $\operatorname{span}\big(T(\beta)\big) \subseteq R(T)$.

Now suppose that $w \in R(T)$, then $w = T(v)$ for some $v \in V$. Since β is a basis for V, we have $v = a_1v_1 + \cdots + a_nv_n$ where $a_i \in F$ for $i = 1, ..., n$. Then $w = T(v) = a_1T(v_1) + \cdots + a_nT(v_n) \in \operatorname{span}\big(T(\beta)\big)$ since T is linear. Therefore $R(T) \subseteq \operatorname{span}\big(T(\beta)\big)$. □

The following result relating the dimension of a vector space with those of the null space and range of a linear transformation is called the *Dimension Theorem.*

Theorem 2.25. *Let* V *and* W *be vector spaces and let* $T : V \to W$ *be linear. If* V *is finite-dimensional, then nullity*(T) + *rank*$(T) = \dim V$*.*

Proof. Suppose that $\dim V = n$ and $\dim\big(N(T)\big) = k$. Let $\beta' = \{v_1, ..., v_k\}$ be a basis for the subspace $N(T)$ of V and extend β' to a basis $\beta = \{v_1, ..., v_n\}$ for V. We prove that the set $\gamma = \{T(v_{k+1}), ..., T(v_n)\}$ is a basis for $R(T)$. Since $T(v_i) = 0$ for $i = 1, ..., k$, we have $R(T) = \operatorname{span}\big(\{T(v_1), ..., T(v_n)\}\big) = \operatorname{span}\big(\{T(v_{k+1}), ..., T(v_n)\}\big) = \operatorname{span}(\gamma)$. Suppose that

$$\sum_{i=k+1}^{n} b_iT(v_i) = 0$$

for some $b_{k+1}, ..., b_n \in F$. Since T is linear, we have

$$T\left(\sum_{i=k+1}^{n} b_iv_i\right) = 0$$

and

$$\sum_{i+k+1}^{n} b_iv_i \in N(T).$$

We have that β' is a basis for $N(T)$, so there exist $c_1, ..., c_n \in F$ such that

$$\sum_{i=k+1}^{n} b_i v_i - \sum_{i=1}^{k} c_i v_i = 0.$$

Since β is a basis for V, we must have $b_{k+1} = \cdots = b_n = 0$, thus γ is linearly independent. Since the elements of γ are distinct, it follows that $\text{rank}(T) = n - k$. □

We define here a similar concept to that of a ring isomorphism. We then present results that will allow us to determine exact conditions under which two vector spaces are isomorphic.

Definition 2.24. *Let V and W be vector spaces. We say that V is* isomorphic *to W if there exists a linear transformation $T : V \to W$ that is both one-to-one and onto, and we call T an* isomorphism *from V onto W.*

Theorem 2.26. *Let V and W be vector spaces and let $T : V \to W$ be linear. Then T is one-to-one if and only if $N(T) = \{0\}$.*

Proof. Suppose that T is one-to-one and $v \in N(T)$. Then $T(v) = 0 = T(0)$. Since T is one-to-one, it follows that $v = 0$, thus $N(T) = \{0\}$.

Assume that $N(T) = \{0\}$, and suppose that $T(v_1) = T(v_2)$. Then $0 = T(v_1) - T(v_2) = T(v_1 - v_2)$, thus $v_1 - v_2 = 0$ and $v_1 = v_2$. Therefore T is one-to-one. □

Theorem 2.27. *Let V and W be vector spaces over a field F, and suppose that $\beta = \{v_1, ..., v_n\}$ is a basis for V. For $w_1, ..., w_n \in W$, there exists a unique linear transformation $T : V \to W$ such that $T(v_i) = w_i$ for $i = 1, ..., n$.*

Proof. Let $v \in V$. Since β is a basis for V, we have $v = a_1 v_1 + \cdots a_n v_n$ for unique $a_1, ..., a_n \in F$. Define $T : V \to W$ by $T(v) = a_1 w_1 + \cdots a_n w_n$. To show that T is linear, suppose that we also have $u \in V$ and $c \in F$. Then $u = b_1 v_1 + \cdots b_n v_n$ for some $b_1, ..., b_n \in F$. Thus

$$cv + u = \sum_{i=1}^{n} (ca_i + b_i) v_i$$

and

$$T(cv + u) = \sum_{i=1}^{n} (ca_i + b_i) w_i = c \sum_{i=1}^{n} a_i w_i + \sum_{i=1}^{n} b_i w_i = cT(v) + T(u),$$

where $T(v_i) = w_i$ for $i = 1, ..., n$.

To verify that T is unique, we let $U : V \to W$ be a linear transformation such that $U(v_i) = w_i$ for $i = 1, ..., n$ and let $v \in V$. Since β is a basis for V, we have $v = a_1 v_1 + \cdots a_n v_n$ for unique $a_1, ..., a_n \in F$. Then

$$U(v) = \sum_{i=1}^{n} a_i U(v_i) = \sum_{i=1}^{n} a_i w_i = T(v)$$

so that $T = U$. □

We are now able to state and prove the following important result.

Theorem 2.28. *Let V and W be finite-dimensional vector spaces over a field F. Then V is isomorphic to W if and only if* $\dim V = \dim W$.

Proof. Suppose that V is isomorphic to W and that $T : V \to W$ is an isomorphism from V to W. Then T is one-to-one so that $N(T) = \{0\}$ and $\text{nullity}(T) = 0$. Then $\text{rank}(T) = \dim\big(R(T)\big) = \dim W$ and $\dim V = 0 + \dim W = \dim W$.

Next, suppose that $\dim V = \dim W$. Let $\beta = \{v_1, ..., v_n\}$ and $\gamma = \{w_1, ..., w_n\}$ be bases for V and W respectively. Then there exists a linear transformation $T : V \to W$ such that $T(v_i) = w_i$ for $i = 1, ..., n$, thus $R(T) = \text{span}\big(T(\beta)\big) = \text{span}(\gamma) = W$. Then T is onto, so that $\text{rank}(T) = \dim W = \dim V$ and $\text{nullity}(T) = 0$. It follows that $N(T) = \{0\}$ so that T is one-to-one. Therefore T is an isomorphism. □

Corollary 2.6. *Let V be a vector space over F. Then V is isomorphic to F^n if and only if* $\dim V = n$.

Proof. Let $e_1 = (1, 0, 0, ..., 0, 0), e_2 = (0, 1, 0, ..., 0, 0), ..., e_n = (0, 0, 0, ..., 0, 1) \in F^n$. The set $\{e_1, ..., e_n\}$ is linearly independent and it spans F^n, thus it forms a basis for F^n. Then $\dim F^n = n$ and it follows that a vector space V is isomorphic to F^n if and only if $\dim V = n$. □

2.3. **Extension Fields.**

We recall here the definitions and some properties of field extensions which can be found in [3]. We describe the elements of field extensions through ring isomorphisms, and we relate extension fields to vector spaces to utilize the concept of vector space dimension in the context of field extensions. We also introduce the derivative for polynomials and rational functions over fields.

Definition 2.25. *A field E is an* extension field *of a field F if $F \subseteq E$ and the operations of F are those of E restricted to F.*

The following result concerning the zeros of a polynomial in an extension field is of great importance.

Theorem 2.29. *Let F be a field and let $f(x)$ be a non-constant polynomial in $F[x]$. Then there exists an extension field E of F in which $f(x)$ has a zero.*

Proof. Since $F[x]$ is a unique factorization domain, there exists an irreducible factor $p(x) = a_n x^n + \cdots + a_1 x + a_0$ of $f(x)$. Since $p(x)$ is irreducible, $E = F[x]/\langle p(x)\rangle$ is a field. We let $\phi : F \to E$ be the mapping given by $\phi(a) = a + \langle p(x)\rangle$. For all $a, b \in F$, $\phi(a) = \phi(b)$ implies $a + \langle p(x)\rangle = b + \langle p(x)\rangle$ so that $a = b$. We also have $\phi(a + b) = a + b + \langle p(x)\rangle = \big(a + \langle p(x)\rangle\big) + \big(b + \langle p(x)\rangle\big) = \phi(a) + \phi(b)$ and $\phi(ab) = ab + \langle p(x)\rangle = \big(a + \langle p(x)\rangle\big)\big(b + \langle p(x)\rangle\big) = \phi(a)\phi(b)$. Thus ϕ is one-to-one and preserves both operations. Since ϕ maps the identity and unity of F to the identity and unity of E respectively, $\phi(F)$ has at least two elements. We let $x, y \in \phi(F)$ where $y \neq 0$. Then there exist elements $a, b \in F$ such that $b \neq 0$ and

$x = \phi(a), y = \phi(b)$. We then have $x - y = \phi(a) - \phi(b) = \phi(a - b) \in \phi(F)$ and $xy^{-1} = \phi(a)\phi(b)^{-1} = \phi(a)\phi\big(b^{-1}\big) = \phi\big(ab^{-1}\big) \in \phi(F)$ so that $\phi(F)$ is a subfield of E which is isomorphic to F. Since

$$\begin{aligned} p\big(x + \langle p(x)\rangle\big) &= \sum_{i=0}^{n} a_i\big(x + \langle p(x)\rangle\big)^i \\ &= \sum_{i=0}^{n} a_i\big(x^i + \langle p(x)\rangle\big) \\ &= \left(\sum_{i=0}^{n} a_i x^i\right) + \langle p(x)\rangle \\ &= p(x) + \langle p(x)\rangle \\ &= 0 + \langle p(x)\rangle, \end{aligned}$$

it follows that $x + \langle p(x)\rangle$ is a zero of $p(x)$ in E. □

Let F be a field and let $a_1, ..., a_n$ be elements of some extension field E of F. We write $F(a_1, ..., a_n)$ to denote the smallest subfield of E that contains both F and the set $\{a_1, ..., a_n\}$. It will be useful to have a way to describe the elements of $F(a_1, ..., a_n)$.

Theorem 2.30. *Let F be a field and let $p(x) \in F[x]$ be irreducible over F. If a is a zero of $p(x)$ in some extension field E of F, then $F(a)$ is isomorphic to $F[x]/\langle p(x)\rangle$. Furthermore, if $\deg p(x) = n$, then every element of $F(a)$ can be expressed in the form*

$$b_{n-1}a^{n-1} + b_{n-2}a^{n-2} + \cdots + b_1 a + b_0$$

where $b_0, ..., b_{n-1} \in F$.

Proof. We consider the ring homomorphism $\phi : F[x] \to F(a)$ given by $\phi\big(f(x)\big) = f(a)$. We show that $\ker\phi = \langle p(x)\rangle$. Since $p(a) = 0$, we have $\langle p(x)\rangle \subseteq \ker\phi$. We also know that $\langle p(x)\rangle$ is a maximal ideal in $F[x]$ since $p(x)$ is irreducible over F, so we conclude from $\ker\phi \neq F[x]$ that $\ker\phi = \langle p(x)\rangle$. Now $\phi\big(F[x]\big)$ is a subfield of $F(a)$, and $\phi\big(F[x]\big)$ contains both F and a. Since $F(a)$ is the smallest field containing both F and a, we have that $F[x]/\langle p(x)\rangle$ is isomorphic to $\phi\big(F[x]\big) = F(a)$.

We now prove the final claim of the theorem. Every element of $F[x]/\langle p(x)\rangle$ has the form $f(x)+\langle p(x)\rangle$ for some polynomial $f(x) \in F[x]$. By the division algorithm, there exists unique polynomials $q(x), r(x) \in F[x]$ such that $f(x) = q(x)p(x)+r(x)$ where $r(x) = 0$ or $\deg r(x) < \deg p(x) = n$. Thus every element of $F[x]/\langle p(x)\rangle$ can be expressed uniquely in the form

$$b_{n-1}x^{n-1} + \cdots + b_0 + \langle p(x)\rangle,$$

where $b_0, ..., b_{n-1} \in F$. The natural isomorphism from $F[x]/\langle p(x)\rangle$ to $F(a)$ then sends $b_k x^k + \langle p(x)\rangle$ to $b_k a^k$. □

The previous result motivates the following definitions.

Definition 2.26. *Let E be an extension field of a field F and let $a \in E$. We call a* algebraic *over F if a is the zero of some nonzero polynomial in $F[x]$. We call a* transcendental *over F if it is not algebraic over F. An extension E of F is called an* algebraic *extension of F if every element of E is algebraic over F. E is called a* transcendental *extension of F if it is not an algebraic extension of F.*

An extension E of F of the form $E = F(a)$ is called a simple *extension of F, and we call an element a such that $E = F(a)$ a* primitive element *of E.*

We now present a way to describe the elements of algebraic and transcendental extension fields.

Theorem 2.31. *Let E be an extension field of the field F and let $a \in E$. If a is transcendental over F, then $F(a)$ is isomorphic to $F(x)$. If a is algebraic over F, then $F(a)$ is isomorphic to $F[x]/\langle p(x)\rangle$, where $p(x)$ is a polynomial in $F[x]$ of minimum degree such that $p(a) = 0$. Moreover, $p(x)$ is irreducible over F.*

Proof. Consider the homomorphism $\phi : F[x] \to F(a)$ given by $f(x) \mapsto f(a)$. If a is transcendental over F, then $\ker\phi = \{0\}$. We extend ϕ to an isomorphism $\psi : F(x) \to F(a)$ by defining $\psi\big(f(x)/g(x)\big) = f(a)/g(a)$.

If a is algebraic over F, then $\ker\phi \neq \{0\}$, thus there exists a polynomial $p(x)$ in $F[x]$ such that $\ker\phi = \langle p(x)\rangle$ and $p(x)$ has minimum degree among all of the nonzero elements of $\ker\phi$. Then $p(a) = 0$ and $p(x)$ is irreducible over F since it is a polynomial of minimum degree satisfying this property. □

If E is an extension field of F, we observe that E is a vector space over F. This motivates the following definition of extension degrees.

Definition 2.27. *Let E be an extension field of a field F. We say that E has* degree *n over F, denoted by $[E : F] = n$, if E has dimension n as a vector space over F. E is called a* finite extension *of F if $[E : F]$ is finite, otherwise we call E an* infinite extension *of F.*

Theorem 2.32. *If E is a finite extension of F, then E is an algebraic extension of F.*

Proof. Suppose that $[E : F] = n$ and $a \in E$. The set $\{1, a, ..., a^n\}$ is linearly dependent over F, so there are elements $b_0, b_1, ..., b_n \in F$ which are not all zero such that

$$b_n a^n + b_{n-1}a^{n-1} + \cdots + b_1 a + b_0 = 0.$$

Then a is a zero of the polynomial

$$f(x) = b_n x^n + b_{n-1}x^{n-1} + \cdots + b_1 x + b_0.$$

□

Results regarding vector spaces can also be used to prove the following result regarding field extensions.

Theorem 2.33. *Let K be a field extension of F. Then $[K : F] = n$ if and only if K is isomorphic to F^n as vector spaces.*

Proof. K and F^n will be isomorphic as vector spaces if and only if K has dimension n as a vector space over F. It follows that K is isomorphic to F^n as vector spaces if and only if $[K : F] = n$. □

The following important result relates the extension degrees of multiple field extensions.

Theorem 2.34. *Let K be a finite extension field of the field E and let E be a finite extension field of the field F. Then K is a finite extension field of F and $[K : F] = [K : E][E : F]$.*

Proof. Let $[K : E] = n$ and $[E : F] = m$. Then K is isomorphic to E^n and E is isomorphic to F^m as vector spaces. It follows that K is isomorphic to $(F^m)^n$ and thus F^{mn} as vector spaces. Therefore $[K : F] = mn$. □

We present some final properties of algebraic extensions.

Theorem 2.35. *If K is an algebraic extension of E and E is an algebraic extension of F, then K is an algebraic extension of F.*

Proof. Let $a \in K$. Since a is algebraic over E, a is the zero of an irreducible polynomial $p(x) = c_n x^n + \cdots + c_1 x + c_0 \in E[x]$. Let $F_0 = F(c_0)$, $F_1 = F_0(c_1)$, ..., $F_{n-1} = F_{n-2}(c_{n-1})$, and $F_n = F_{n-1}(c_n) = F(c_0, c_1, ..., c_n)$. Thus $p(x) \in F_n[x]$ and $[F_n(a) : F_n] = n$. Since c_i is algebraic over F for $i = 0, 1, ..., n$, we obtain finite values $[F_{i+1} : F_i]$ for $i = 0, 1, ..., n$. Then $[F_n(a) : F] = [F_n(a) : F_n][F_n : F_{n-1}] \cdots [F_1 : F_0][F_0 : F]$ is finite so that a belongs to a finite extension of F. □

Corollary 2.7. *Let E be an extension field of the field F. Then the set of all elements of E that are algebraic over F, called the* algebraic closure *of F in E, is a subfield of E.*

Proof. Suppose that $a, b \in E$ are algebraic over F and $b \neq 0$. We have $[F(a, b) : F] = [F(a, b) : F(b)][F(b) : F]$ where a is algebraic over $F(b)$ since it is algebraic over F. Then $[F(a, b) : F]$ is finite and the elements $a + b, a - b, ab, a/b \in F(a, b)$ are algebraic over F. □

It will be of use to know when a polynomial or a rational function over a field F has a zero of multiplicity greater than 1 in some extension field E of F. The derivative will allow us to determine when this is the case.

Definition 2.28. *Let F be a field and let $f(x) = a_n x^n + \cdots + a_2 x^2 + a_1 x + a_0 \in F[x]$. We call the polynomial $f'(x) = n a_n x^{n-1} + \cdots + 2a_2 x + a_1$ the* derivative *of $f(x)$. Let $f(x) = \frac{f_1(x)}{f_2(x)} \in F(x)$. We call $f'(x) = \frac{f_1'(x) f_2(x) - f_1(x) f_2'(x)}{f_2(x)^2}$ the* derivative *of $f(x)$.*

We present some properties of the derivative in the following results. We consider first properties of the derivative for polynomials.

Lemma 2.6. *Let F be a field, let $f(x), g(x) \in F[x]$, and let $c \in F$. Then the following properties hold:*

i) $\big(cf(x) + g(x)\big)' = cf'(x) + g'(x)$;
ii) $\big(f(x)g(x)\big)' = f'(x)g(x) + f(x)g'(x)$;

Proof. Let

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

and assume without loss of generality that $n \geq m$ and let $b_k = 0$ for $k > m$. Then

$$\begin{aligned}
\big(cf(x) + g(x)\big)' &= \left(\sum_{i=0}^{n} (ca_i + b_i)x^i\right)' \\
&= \sum_{i=1}^{n} i(ca_i + b_i)x^{i-1} \\
&= c\sum_{i=1}^{n} i a_i x^{i-1} + \sum_{j=1}^{n} j b_j x^{j-1} \\
&= cf'(x) + g'(x).
\end{aligned}$$

We prove the second property by induction on the number of terms of the polynomial $f(x)$. For the basis step, we assume that $f(x)$ is a monomial so we let $f(x) = a_n x^n$ and

$$g(x) = \sum_{j=0}^{m} b_j x^j.$$

Then

$$\begin{aligned}
\big(f(x)g(x)\big)' &= \left(\sum_{j=0}^{m} a_n b_j x^{j+n}\right)' \\
&= \sum_{j=1}^{m} (j+n) a_n b_j x^{j+n-1} \\
&= a_n x^n \sum_{j=1}^{m} j b_j x^{j-1} + n a_n x^{n-1} \sum_{j=1}^{m} b_j x^j \\
&= f(x)g'(x) + f'(x)g(x).
\end{aligned}$$

We assume now that the result holds when $f(x)$ has n or less terms, and we show that the result holds when $f(x)$ has $n+1$ terms. Let

$$f(x) = \sum_{i=0}^{n} a_i x^i \quad \text{and} \quad g(x) = \sum_{j=0}^{m} b_j x^j.$$

Let $\overline{f}(x) = f(x) - a_n x^n$, then $\overline{f}(x)$ is a polynomial with at most n terms. By the first proved property and our induction hypothesis, we obtain

$$\begin{aligned}
\big(f(x)g(x)\big)' &= \big(\overline{f}(x)g(x) + a_n x^n g(x)\big)' \\
&= \overline{f}'(x)g(x) + \overline{f}(x)g'(x) + a_n x^n g'(x) + n a_n x^{n-1} g(x) \\
&= \big(a_n x^n + \overline{f}(x)\big)g'(x) + \big(n a_n x^{n-1} + \overline{f}'(x)\big)g(x) \\
&= f'(x)g(x) + f(x)g'(x).
\end{aligned}$$

□

We can extend the properties of the derivative from the previous result to the derivative of rational functions.

Lemma 2.7. *Let F be a field, let $f(x), g(x) \in F(x)$, and let $c \in F$. Then the following properties hold:*

i) $\big(cf(x) + g(x)\big)' = cf'(x) + g'(x)$;
ii) $\big(f(x)g(x)\big)' = f'(x)g(x) + f(x)g'(x)$.

Proof. Let $f(x) = \frac{f_1(x)}{f_2(x)}$ and $g(x) = \frac{g_1(x)}{g_2(x)}$. The properties follow from the definition of the derivative for elements in $F(x)$ and applications of Lemma 2.6:

$$\begin{aligned}
\big(cf(x) + g(x)\big)' &= \left(\frac{cf_1(x)g_2(x) + f_2(x)g_1(x)}{f_2(x)g_2(x)}\right)' \\
&= \Big[f_2(x)g_2(x)\big[cf_1(x){g_2}'(x) + c{f_1}'(x)g_2(x) + {f_2}'(x)g_1(x) \\
&\qquad + f_2(x){g_1}'(x)\big] - \big(cf_1(x)g_2(x) + f_2(x)g_1(x)\big)\big[f_2(x){g_2}'(x) \\
&\qquad + {f_2}'(x)g_2(x)\big]\Big] / \big(f_2(x)g_2(x)\big)^2 \\
&= \Big(c\big({f_1}'(x)f_2(x) - f_1(x){f_2}'(x)\big)g_2(x)^2 \\
&\qquad + \big({g_1}'(x)g_2(x) - g_1(x){g_2}'(x)\big)f_2(x)^2\Big) / \big(f_2(x)g_2(x)\big)^2 \\
&= c\frac{{f_1}'(x)f_2(x) - f_1(x){f_2}'(x)}{f_2(x)^2} + \frac{{g_1}'(x)g_2(x) - g_1(x){g_2}'(x)}{g_2(x)^2} \\
&= cf'(x) + g'(x).
\end{aligned}$$

$$\begin{aligned}
\big(f(x)g(x)\big)' &= \left(\frac{f_1(x)g_1(x)}{f_2(x)g_2(x)}\right)' \\
&= \Big(f_2(x)g_2(x)\big(f_1{}'(x)g_1(x) + f_1(x)g_1{}'(x)\big) \\
&\qquad -f_1(x)g_1(x)\big(f_2{}'(x)g_2(x) + f_2(x)g_2{}'(x)\big)\Big/\big(f_2(x)g_2(x)\big)^2 \\
&= \Big(\big(f_1{}'(x)f_2(x) - f_1(x)f_2{}'(x)\big)g_1(x)g_2(x) \\
&\qquad +f_1(x)f_2(x)\big(g_1{}'(x)g_2(x) - g_1(x)g_2{}'(x)\big)\Big/\big(f_2(x)g_2(x)\big)^2 \\
&= \frac{f_1{}'(x)f_2(x) - f_1(x)f_2{}'(x)}{f_2(x)^2}\frac{g_1(x)}{g_2(x)} + \frac{f_1(x)}{f_2(x)}\frac{g_1{}'(x)g_2(x) - g_1(x)g_2{}'(x)}{g_2(x)^2} \\
&= f'(x)g(x) + f(x)g'(x).
\end{aligned}$$

□

We present one final property of the derivative for rational functions called the Chain Rule, which considers the derivative of a composition of functions.

Proposition 2.1. *Let F be a field and let $f(x), g(x) \in F(x)$. Then $(f \circ g)'(x) = (f' \circ g)(x) \cdot g'(x)$.*

Proof. We first prove the result for the case $f(x) \in F[x]$. Let

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

For any positive integer m, we have

$$\big(g(x)^m\big)' = \sum_{j=1}^{m} g(x)^{m-1}g'(x) = mg(x)^{m-1}g'(x)$$

so that

$$\begin{aligned}
(f \circ g)'(x) &= \left(\sum_{i=0}^{n} a_i g(x)^i\right)' \\
&= \sum_{i=0}^{n} \big(a_i g(x)^i\big)' \\
&= \sum_{i=1}^{n} i a_i g(x)^{i-1}g'(x) \\
&= f'\big(g(x)\big)g'(x).
\end{aligned}$$

To prove the result for $f(x) \in F(x)$, we let $f(x) = f_1(x)/f_2(x)$ and note that

$$\left(\frac{1}{f(x)}\right)' = \left(\frac{f_2(x)}{f_1(x)}\right)' = \frac{-\big(f_1{}'(x)f_2(x) - f_1(x)f_2{}'(x)\big)}{f_1(x)^2} = \frac{-f'(x)}{f(x)^2}.$$

We now obtain the following:

$$\begin{aligned}
(f \circ g)'(x) &= \left(f_1\big(g(x)\big) \cdot \frac{1}{f_2\big(g(x)\big)} \right)' \\
&= \big(f_1\big(g(x)\big)\big)' \cdot \frac{1}{f_2\big(g(x)\big)} + f_1\big(g(x)\big) \cdot \left(\frac{1}{f_2\big(g(x)\big)} \right)' \\
&= f_1{}'\big(g(x)\big)g'(x) \cdot \frac{1}{f_2\big(g(x)\big)} + f_1\big(g(x)\big) \cdot \left(\frac{-f_2{}'\big(g(x)\big)g'(x)}{f_2\big(g(x)\big)^2} \right) \\
&= \Big(f_1{}'\big(g(x)\big) f_2\big(g(x)\big) - f_1\big(g(x)\big) f_2{}'\big(g(x)\big) \Big) g'(x) / f_2\big(g(x)\big)^2 \\
&= f'\big(g(x)\big)g'(x).
\end{aligned}$$

□

The derivative can be used to prove the following result regarding the zeros of multiplicity greater than 1 of polynomials over a field F.

Theorem 2.36. *A polynomial $f(x)$ over a field F has a zero of multiplicity greater than 1 in some extension field E if and only if $f(x)$ and $f'(x)$ have a common factor of positive degree in $F[x]$.*

Proof. Let a be a zero of $f(x)$ of multiplicity greater than 1 in some extension E of F. Then there exists $g(x) \in E[x]$ such that $f(x) = (x-a)^2 g(x)$. Since $f'(x) = (x-a)^2 g'(x) + 2(x-a)g(x)$, we have $f'(a) = 0$ so that $x - a$ is a factor of both $f(x)$ and $f'(x)$ in E. Suppose that $f(x)$ and $f'(x)$ have no common divisor of positive degree in $F[x]$, then there exist $g(x), h(x) \in F[x]$ such that $f(x)g(x) + f'(x)h(x) = 1$. Then $x - a$ is a factor of $f(x)g(x) + f'(x)h(x)$ in $E[x]$, so that $x - a$ is a factor of 1, yielding a contradiction. Therefore $f(x)$ and $f'(x)$ must have a common divisor of positive degree in $F[x]$.

Suppose that $f(x)$ and $f'(x)$ have a common factor of positive degree in $F[x]$ and let a be a zero of that common factor. Since a is a zero of $f(x)$, there exists a polynomial $q(x)$ such that $f(x) = (x-a)q(x)$. Then $f'(x) = q(x) + (x-a)q'(x)$ and a is a zero of $f'(x)$ so that $0 = f'(a) = q(a)$. Thus $x - a$ is a factor of $q(x)$ and a is a zero of $f(x)$ with multiplicity greater than 1. □

3. **Prime Polynomials**

Let $f(x)$ be a non-constant polynomial. If $f(x)$ can be expressed as the composition of two polynomials $g(x)$ and $h(x)$ with degrees at least 2, $f(x)$ is said to be composite. Otherwise, $f(x)$ is said to be prime. In this chapter we present a few of Ayad's results from [1] regarding prime polynomials. We will first require some definitions, so we present them here.

Let $f(x)$ be a non-constant complex polynomial and let $x_0 \in \mathbb{C}$. The smallest integer $i \geq 1$ such that $f^{(i)}(x_0) \neq 0$ is called the valency of $f(x)$ at x_0 and is denoted by $v_f(x_0)$. The number x_0 is called a critical point of $f(x)$ provided that $v_f(x_0) \geq 2$. The number $t_0 \in \mathbb{C}$ is a critical value of $f(x)$ if there exists a critical point x_0 of $f(x)$ such that $f(x_0) = t_0$.

Let R be an integral domain and let K be its field of fractions. Let $u(x) = a_n x^n + \cdots + a_1 x + a_0$ and $v(x) = b_m x^m + \cdots + b_1 x + b_0$ be polynomials over R. Let $\alpha_1, ..., \alpha_n$ and $\beta_1, ..., \beta_m$ be the all of the roots of $u(x)$ and $v(x)$ respectively in an algebraic closure of K. The resultant of $u(x)$ and $v(x)$ is given by

$$\mathrm{Res}_x\big(u(x), v(x)\big) = a_n^m b_m^n \prod_{i=1}^{n} \prod_{j=1}^{m} (\alpha_i - \beta_j).$$

We then define the discriminant of the polynomial $u(x)$ by

$$D\big(u(x)\big) = \frac{(-1)^{n(n-1)/2}}{a_n} \mathrm{Res}_x\big(u(x), u'(x)\big).$$

The following properties of the resultant follow from the definition.

i) Let $u(x)$ and $v(x)$ be polynomials as described above. Then

$$\mathrm{Res}_x\big(v(x), u(x)\big) = (-1)^{nm} \mathrm{Res}_x\big(u(x), v(x)\big).$$

ii) Under the same hypotheses as the first property,

$$\mathrm{Res}_x\big(u(x), v(x)\big) = a_n^m \prod_{i=1}^{n} v(\alpha_i).$$

iii) $u(x)$ and $v(x)$ have a zero in common if and only if

$$\mathrm{Res}_x\big(u(x), v(x)\big) = 0.$$

iv) For an additional polynomial $w(x)$ over R, we have

$$\mathrm{Res}_x\big(u(x), v(x)w(x)\big) = \mathrm{Res}_x\big(u(x), v(x)\big) \mathrm{Res}_x\big(u(x), w(x)\big).$$

Let $f(x)$ be a complex polynomial, let $f'(x)$ be the derivative of $f(x)$, and let $\beta_1, ..., \beta_{n-1}$ be all of the zeros of $f'(x)$. Let t be a variable, let b be the leading coefficient of $f'(x)$, and let $n = \deg f(x)$. Consider the discriminant $D(t) =$

$\mathrm{Res}_x\big(f(x)-t, f'(x)\big)$. Using the properties of the resultant, we have the following:

$$\begin{aligned} D(t) &= \mathrm{Res}_x\left(f(x)-t, b\prod_{i=1}^{n-1}(x-\beta_i)\right) \\ &= (-1)^{n(n-1)}b^n\prod_{i=1}^{n-1}\big(f(\beta_i)-t\big). \end{aligned}$$

We remark that since β_i is a zero of $f'(x)$, β_i is a critical point of $f(x)$ and $f(\beta_i)$ is a critical value of $f(x)$ for $i = 1, ..., m$. It immediately follows that $D(t_0) = 0$ if and only if t_0 is a critical value of $f(x)$. We define the multiplicity of the critical value t_0 as the multiplicity of t_0 as a root of $D(t)$, and we call a critical value with multiplicity equal to one a simple critical value.

Under certain conditions, the critical values of a polynomial can be used to identify if a polynomial is a prime polynomial.

Lemma 3.1. *Let $f(x)$ be a composite polynomial of the from $f(x) = g\big(h(x)\big)$ where $g(x), h(x)$ are polynomials of degree m and k respectively. Let $D(t)$ be the discriminant of $f(x) - t$, then there exists $a \in \mathbb{C}$ such that*

$$D(t) = a[D(g(x)-t)]^k \mathrm{Res}_x(f(x)-t, h'(x)).$$

Proof. Write $u(t) \sim v(t)$ to denote that the polynomials $u(t)$ and $v(t)$ are equal up to multiplication by a constant. We have

$$\begin{aligned} D(t) &\sim \mathrm{Res}_x\left(f(x)-t, f'(x)\right) \\ &\sim \mathrm{Res}_x\left(f(x)-t, g'\big(h(x)\big)\right)\cdot \mathrm{Res}_x\left(f(x)-t, h'(x)\right). \end{aligned}$$

We also have

$$\begin{aligned} \mathrm{Res}_x\left(f(x)-t, g'\big(h(x)\big)\right) &= \mathrm{Res}_x\left(g\big(h(x)\big)-t, g'\big(h(x)\big)\right) \\ &= \mathrm{Res}_x\left[\, g(h(x))-t, mb\prod_{i=1}^{m-1}\big(h(x)-\beta_i\big)\,\right], \end{aligned}$$

where b is the leading coefficient of g and $\beta_1, ..., \beta_{m-1}$ are the roots of $g'(x)$. Thus

$$\begin{aligned} \mathrm{Res}_x\left(f(x)-t, g'\big(h(x)\big)\right) &\sim \mathrm{Res}_x\left[\, g\big(h(x)\big)-t, \prod_{i=1}^{m-1}\big(h(x)-\beta_i\big)\,\right] \\ &\sim \prod_{i=1}^{m-1}\mathrm{Res}_x\left(g\big(h(x)\big)-t, h(x)-\beta_i\right) \\ &\sim \left[\prod_{i=1}^{m-1}\big(g(\beta_i)-t\big)\right]^k \\ &\sim [\,\mathrm{Res}_x\big(g(x)-t, g'(x)\big)]^k \\ &\sim [D\big(g(x)-t\big)]^k. \end{aligned}$$

We conclude that $D(t) \sim [D\big(g(x)-t\big)]^k \cdot \mathrm{Res}_x\big(f(x)-t, h'(x)\big)$. □

Corollary 3.1. *Let $f(x)$ be a polynomial of degree n and let $D(t)$ be the discriminant of $f(x)-t$. Suppose that $f(x)$ is composite and has a right composition factor of degree $k \geq 2$, then there exists two non-constant polynomials $A(t)$ and $B(t)$ such that $\deg B = k-1$ and $D(t) = [A(t)]^k B(t)$.*

Proof. We let $f(x) = g\big(h(x)\big)$ where $h(x)$ is a right composition factor of degree $k \geq 2$. By Lemma 3.1, there exists $a \in \mathbb{C}$ such that $D(t) = a[D\big(g(x)-t\big)]^k \cdot \mathrm{Res}_x\big(f(x)-t, h'(x)\big)$. We set $A(t) = D\big(g(x)-t\big)$ and $B(t) = a\mathrm{Res}_x\big(f(x)-t, h'(x)\big)$, and the result follows. □

Theorem 3.1. *Let $f(x)$ be a complex polynomial of degree n and let d be the greatest proper divisor of n. Suppose that $f(x)$ has at least d simple critical values, then $f(x)$ is prime.*

Proof. Suppose by contradiction that $f(x)$ is composite. There exist complex polynomials $g(x)$ and $h(x)$ of degrees $m \geq 2$ and $k \geq 2$ respectively such that $f(x) = g\big(h(x)\big)$. We let $D(t)$ be the discriminant of $f(x)-t$, and we write $D(t) = [A(t)]^k B(t)$ where $\deg B(t) = k-1$. Let δ be the number of simple critical values of $f(x)$. Since these critical values must be roots of the polynomial $B(t)$, we obtain

$$k-1 = \deg B(t) \geq \delta \geq d \geq k$$

which is a contradiction. Therefore $f(x)$ is prime. □

The valencies of the critical points of a polynomial can also be used to determine if a polynomial is a prime polynomial.

Lemma 3.2. *Let $f_1(x)$ and $f_2(x)$ be non-constant polynomials, then there exists a polynomial $l(x)$ of degree 1 such that $f_2 = l \circ f_1$ if and only if for every $x_0 \in \mathbb{C}$, we have $v_{f_2}(x_0) = v_{f_1}(x_0)$.*

Proof. If there exists a polynomial $l(x)$ such that $f_2 = l \circ f_1$, then there exist $a, b \in \mathbb{C}$ such that $f_2(x) = af_1(x) + b$. We then have ${f_2}'(x) = a{f_1}'(x)$, and it follows that $v_{f_1}(x_0) = v_{f_2}(x_0)$ for all $x_0 \in \mathbb{C}$.

Now, let α be any root of $f_1'(x)$ of order $e \geq 1$. There exists a polynomial $q_1(x)$ with $q_1(\alpha) \neq 0$ such that $f_1'(x) = (x-\alpha)^e q_1(x)$. Hence there exists a polynomial $p_1(x)$ with $p_1(\alpha) \neq 0$ such that $f_1(x) - f_1(\alpha) = (x-\alpha)^{e+1} p_1(x)$. We deduce that $v_{f_2}(\alpha) = v_{f_1}(\alpha) = e+1$. This implies that α is a root of $f_2'(x)$ of order e. Since the roles of $f_1'(x)$ and $f_2'(x)$ are symmetric, we conclude that there exists some $a \in \mathbb{C}^*$ such that $f_2'(x) = af_1'(x)$ and then $f_2(x) = af_1(x) + b$ with $b \in \mathbb{C}$. □

Proposition 3.1. *Let $f(x), g(x)$, and $h(x)$ be three non-constant polynomials, then the following assertions are equivalent:*

i) There exists a polynomial $l(x)$ of degree 1 such that $l \circ f = g \circ h$;
ii) for any $x_0 \in \mathbb{C}$ we have $v_f(x_0) = v_h(x_0) \cdot v_g\big(h(x_0)\big)$.

Proof. Suppose that there exists a polynomial $l(x)$ of degree 1 such that $l \circ f = g \circ h$, then $f = l^{-1} \circ g \circ h$. Set

$$(l^{-1} \circ g)'(x) = c(x - \alpha_1)^{e_1} \cdots (x - \alpha_s)^{e_s}$$

where $e_1, ..., e_s$ are positive integers. Then

$$f'(x) = h'(x) \cdot (l^{-1} \circ g)'\big(h(x)\big) = ch'(x)\big(h(x) - \alpha_1\big)^{e_1} \cdots \big(h(x) - \alpha_s\big)^{e_s}.$$

From this we deduce that for any $x_0 \in \mathbb{C}$, we have

$$v_f(x_0) - 1 = v_h(x_0) - 1 + \big(v_h(x_0)\big) \cdot \Big(v_{l^{-1} \circ g}\big(h(x_0)\big) - 1\Big),$$

hence

$$v_f(x_0) = v_h(x_0) v_{l^{-1} \circ g}\big(h(x_0)\big)$$

and $v_f(x_0) = v_h(x_0) v_g\big(h(x_0)\big)$.

Suppose that for any $x_0 \in \mathbb{C}$ we have $v_f(x_0) = v_h(x_0) v_g\big(h(x_0)\big)$ and set $f_1 = g \circ h$. We conclude from the first half of the proof that for any $x_0 \in \mathbb{C}$, we have $v_{f_1}(x_0) = v_h(x_0) v_g\big(h(x_0)\big)$, hence $v_f(x_0) = v_{f_1}(x_0)$ for any $x_0 \in \mathbb{C}$. Then there exists a polynomial $l(x)$ of degree 1 such that $g \circ h = f_1 = l \circ f$. □

Theorem 3.2. *Let $f(x)$ be a complex polynomial of degree n. Let d be the greatest proper divisor of n and suppose that $f(x)$ has a critical point $x_0 \in \mathbb{C}$ such that its valency $v_f(x_0)$ is a prime number $p > d$, then $f(x)$ is prime.*

Proof. Suppose that $f = g \circ h$ where $g(x)$ and $h(x)$ are polynomials of degree at least 2 and that $v_f(x_0) = p$ is a prime number for some $x_0 \in \mathbb{C}$. We consider two cases:

$$v_h(x_0) = p \quad \text{and} \quad v_g\big(h(x_0)\big) = 1$$

or

$$v_h(x_0) = 1 \quad \text{and} \quad v_g\big(h(x_0)\big) = p.$$

In the first, we have

$$d - 1 \geq \deg h - 1 = \sum_{x \in \mathbb{C}} \big(v_h(x) - 1\big) \geq p - 1.$$

Thus $p \leq d$, yielding a contradiction. The same method also yields a contradiction in the second case. Therefore $f(x)$ is prime. □

Corollary 3.2. *Let $f(x) = (x - x_0)^p u(x)$ be a complex polynomial, where p is a prime number, $\deg u(x) < p$, and $x_0 \in \mathbb{C}$ is such that $u(x_0) \neq 0$. Then $f(x)$ is prime.*

Proof. Let $n = \deg f(x)$ and let d be the greatest proper divisor of n. Then

$$2p > p + \deg u(x) = n \geq 2d$$

so that $p > d$. Since $v_f(x_0) = p > d$, the polynomial $f(x)$ is prime. □

4. **Main Results on Prime Rational Functions**

In this chapter[1] we present the main results regarding prime rational functions. This work is motivated by existing work on polynomials, which we will briefly recall here. Beardon proved in [2] that if a polynomial $f(x)$ of degree n has more than $n/2$ critical values, then $f(x)$ is prime. In [1], Ayad defined the multiplicity of a critical value and proved that if a polynomial $f(x)$ of degree n has more than d simple critical values where d is the greatest proper divisor of n, then $f(x)$ is prime. Ayad also provided examples of prime polynomials by considering the valencies of their critical points.

We extend this concept to rational functions as follows. Let $\mathbb{C}[x]$ be the ring of complex polynomials and let $\mathbb{C}(x)$ be its field of fractions. When we refer to the complex rational function $f(x)$, we mean the unique ratio $\frac{f_1(x)}{f_2(x)}$ of complex polynomials $f_1(x)$ and $f_2(x)$ where $f_2(x)$ is monic and no linear factor divides both $f_1(x)$ and $f_2(x)$. We then define the degree of $f(x)$ by

$$\deg f(x) = \max\{\deg f_1(x), \deg f_2(x)\}.$$

Let $f(x)$ be a non-constant complex rational function. We call $f(x)$ composite if there exist complex rational functions $g(x)$ and $h(x)$, both with degrees at least 2, such that $f(x) = g\big(h(x)\big)$. Otherwise, we call $f(x)$ prime.

In the sections that follow, we make use of the set of units under function composition to provide conditions on the multiplicities of the zeros and poles of a rational function $f(x)$ which are sufficient to conclude that $f(x)$ is prime. We define the resultant of two rational functions, and motivated by Ayad's results in [1], we present conditions on the critical values of a rational function $f(x)$ under which $f(x)$ is prime and use these results to provide examples of prime rational functions.

4.1. **Units and composite rational functions.**

Let $f(x)$ be a complex rational function. Then $f(x)$ can be expressed as the ratio of two complex polynomials such that no linear factor divides both of the polynomials in its numerator and its denominator, and we say that $f(x)$ is in its most reduced form. Since such a reduced form is useful when trying to determine the degree of a rational function, we provide here an expression for the reduced form of a composition of two rational functions.

Lemma 4.1. *Let $g(x)$ and $h(x)$ be rational functions in their most reduced forms with*

$$g(x) = \frac{b\prod_{i=1}^{m_1}(x-\alpha_i)}{\prod_{j=1}^{m_2}(x-\beta_j)} \quad \text{and} \quad h(x) = \frac{h_1(x)}{h_2(x)}.$$

[1] *A version of this chapter has been submitted for publication*
Kihel, O., Larone, J., 2014

Then the expression for $g\big(h(x)\big)$ *given by*

$$g\big(h(x)\big) = \frac{bh_2(x)^{\deg g - m_1}\prod_{i=1}^{m_1}\big(h_1(x)-\alpha_i h_2(x)\big)}{h_2(x)^{\deg g - m_2}\prod_{j=1}^{m_2}\big(h_1(x)-\beta_j h_2(x)\big)}$$

is in its most reduced form.

Proof. Let $m = \deg g$. We suppose by contradiction that the given expression for $g\big(h(x)\big)$ is not in its most reduced form, then there exists a linear factor $\ell(x)$ which divides both the numerator and denominator of $g\big(h(x)\big)$. We consider the possible cases:

i) If $\ell(x)$ divides $h_2(x)$ and $h_1(x)-\gamma h_2(x)$ for any $\gamma \in \mathbb{C}$, then $\ell(x)$ divides $\big(h_1(x)-\gamma h_2(x)\big)+\big(\gamma h_2(x)\big) = h_1(x)$ which contradicts $h(x)$ being in its most reduced form. Thus we reject the case where $\ell(x)$ divides $h_2(x)^{m-m_1}$ and $h_1(x)-\beta_j h_2(x)$ for some $j = 1, ..., m_2$ as well as the case where $\ell(x)$ divides $h_2(x)^{m-m_2}$ and $h_1(x)-\alpha_i h_2(x)$ for some $i = 1, ..., m_1$.

ii) If $\ell(x)$ divides $h_1(x)-\alpha_i h_2(x)$ for some i and $h_1(x)-\beta_j h_2(x)$ for some j, then there exist polynomials $q_1(x)$ and $q_2(x)$ such that $h_1(x)-\alpha_i h_2(x) = q_1(x)\ell(x)$ and $h_1(x)-\beta_j h_2(x) = q_2(x)\ell(x)$. Solving these two expressions for $h_1(x)$ yields

$$q_1(x)\ell(x)+\alpha_i h_2(x) = h_1(x) = q_2(x)\ell(x)+\beta_j h_2(x).$$

From this we obtain

$$\big(q_1(x)-q_2(x)\big)\ell(x) = (\beta_j-\alpha_i)h_2(x).$$

Since $g(x)$ is in its most reduced form, we have $\alpha_i \neq \beta_j$, thus $\ell(x)$ must divide $h_2(x)$. This leads us again to case (i), which yields a contradiction.

□

We prove here a proposition which will be essential for the rest of this paper.

Proposition 4.1. *Let* K *be a field and let* $f(x) = \frac{f_1(x)}{f_2(x)}$ *be a rational function over* K *in its most reduced form, then*

$$\deg f = [K(x) : K(f)].$$

Proof. We have $K(f) \subset K(x) = K(f, x)$ where x is a primitive element of $K(x)$ over $K(f)$. Then x is a root of the polynomial

$$F(y) = f_1(y) - f \cdot f_2(y) \in K(f)[y]$$

and $\deg F = \max\{\deg f_1, \deg f_2\}$. F is irreducible in $K[f, y] = K(f)[y]$, thus it is irreducible in $K(f)[y]$. Then $[K(x) : K(f)] = \deg F = \max\{\deg f_1, \deg f_2\} = \deg f$. □

Proposition 4.2. *Let* K *be a field and let* $f(x) = g\big(h(x)\big)$ *where* $f(x), g(x)$, *and* $h(x)$ *are rational functions over* K, *then*

$$\deg f = \deg g \cdot \deg h.$$

Proof. We have $K(f) \subset K(h) \subset K(x)$, $[K(x) : K(f)] = \deg f$, $[K(x) : K(h)] = \deg h$, and $[K(h) : K(f)] = \deg g$. The desired result follows. □

Corollary 4.1. *Let $f(x)$ be a complex rational function of degree p where p is a prime number. Then $f(x)$ is prime.*

We recall that a rational function $\mu(x)$ is a unit under function composition if there exists another rational function $\mu^{-1}(x)$ such that $\mu\big(\mu^{-1}(x)\big) = \mu^{-1}\big(\mu(x)\big) = x$. Then $\deg \mu(x) \cdot \deg \mu^{-1}(x) = \deg x = 1$ and it follows that both $\mu(x)$ and $\mu^{-1}(x)$ must have degree 1. We claim that the complex rational functions of degree 1 form the group of units under function composition, which is the motivation for the requirement that the composition factors of a composite function have degree at least 2. One can verify that the function $\mu(x) = \frac{ax+b}{cx+d}$ has degree 1 if and only if $ad - bc \neq 0$, and in this case it has an inverse given by $\mu^{-1}(x) = \frac{dx-b}{-cx+a}$. When we refer to a unit $\mu(x)$, we mean that $\mu(x)$ is a unit under function composition.

This group of units will be very useful in the study of whether a function is prime due to the following result.

Lemma 4.2. *Let f be a complex rational function and let μ be a unit. Then $f \circ \mu$ and $\mu \circ f$ are composite if and only if f is composite.*

Proof. If $\mu \circ f$ is composite, then $\mu \circ f = g \circ h$ for some complex rational functions g and h with degrees at least 2, so that $f = (\mu^{-1} \circ g) \circ h$ is composite. Similarly, if $f \circ \mu$ is composite, then $f \circ \mu = g \circ h$ for complex rational functions g and h with degrees at least 2, so that $f = g \circ (h \circ \mu^{-1})$ is composite.

Conversely, if f is composite, then $f = g \circ h$ for complex rational functions g and h with degrees at least 2, so that $\mu \circ f = (\mu \circ g) \circ h$ and $f \circ \mu = g \circ (h \circ \mu)$ are both composite. □

The results provided by the two following lemmas will be frequently used in this paper. The first provides a particular pair of composition factors for composite rational functions, and the second relates the numerator and denominator degrees of a composite rational function with those of its composition factors.

Lemma 4.3. *Let $f(x)$ be a complex composite rational function. There exist complex rational functions $g(x)$ and $h(x)$ of degrees at least 2 such that $f(x) = g\big(h(x)\big)$ where the numerator degree of $h(x)$ is larger than its denominator degree.*

Proof. We are given that $f(x)$ is composite, so there exist complex rational functions $G(x)$ and $H(x)$ of degrees at least 2 such that $f(x) = G\big(H(x)\big)$. We let $\mu(x)$ be a complex rational function of degree 1. We consider the expression $\mu\big(H(x)\big)$ explicitly, and we will choose $\mu(x)$ so that $\mu\big(H(x)\big)$ has larger numerator degree than denominator degree. Let $H(x) = \frac{H_1(x)}{H_2(x)}$ and consider two cases.

i) If $\deg H_1 > \deg H_2$ we let $\mu(x) = x$;
ii) if $\deg H_1 \leq \deg H_2$, we write $H_1(x) = aH_2(x) + r(x)$ where $a \in \mathbb{C}$ and $\deg r < \deg H_2$. Then $H(x) = a + \frac{r(x)}{H_2(x)}$ and we let $\mu(x) = \frac{1}{x-a}$.

In both cases, $\mu\big(H(x)\big)$ has numerator degree greater than its denominator degree. Since $\mu(x)$ has degree 1, there exists $\mu^{-1}(x)$ such that $\mu^{-1}\big(\mu(x)\big) = x$. We define $g(x) = G\big(\mu^{-1}(x)\big)$ and $h(x) = \mu\big(H(x)\big)$. Then

$$f = G \circ H = G \circ \mu^{-1} \circ \mu \circ H = (G \circ \mu^{-1}) \circ (\mu \circ H) = g \circ h$$

is a decomposition of f such that $f(x) = g\big(h(x)\big)$ where the numerator degree of $h(x)$ is larger than its denominator degree. □

Lemma 4.4. *Let $f(x)$ be a composite complex rational function with $f(x) = g\big(h(x)\big)$. Let n_1, m_1, and k_1 be the numerator degrees of $f(x), g(x)$, and $h(x)$ respectively and let n_2, m_2, and k_2 be the denominator degrees of $f(x), g(x)$, and $h(x)$ respectively. If $k_1 > k_2$, then*

$$(n_1 - n_2) = (m_1 - m_2)(k_1 - k_2).$$

Proof. Let $h(x) = \frac{h_1(x)}{h_2(x)}$ and let

$$g(x) = \frac{b\prod_{i=1}^{m_1}(x - \alpha_i)}{\prod_{j=1}^{m_2}(x - \beta_j)}$$

have degree m. Then we have

$$f(x) = \frac{f_1(x)}{f_2(x)} = \frac{bh_2(x)^{m-m_1}\prod_{i=1}^{m_1}\big(h_1(x) - \alpha_i h_2(x)\big)}{h_2(x)^{m-m_2}\prod_{j=1}^{m_2}\big(h_1(x) - \beta_j h_2(x)\big)}.$$

Since we have $k_1 > k_2$ by assumption, the numerator and denominator degrees of $f(x)$ satisfy $n_1 + (m - m_2)k_2 + m_2k_1 = n_2 + (m - m_1)k_2 + m_1k_1$. It follows that $n_1 - n_2 = (m_1 - m_2)(k_1 - k_2)$ as desired. □

The following property extends the relationship between the degree of a polynomial and that of its derivative to the case of a rational function.

Lemma 4.5. *Let $f(x)$ be a complex rational function with numerator degree n_1 and denominator degree n_2, and let $f'(x)$ have numerator degree ${n_1}'$ and denominator degree ${n_2}'$. If $n_1 - n_2 \neq 0$, then ${n_1}' - {n_2}' = n_1 - n_2 - 1$*

Proof. Let $f(x) = \frac{ax^{n_1} + f_1(x)}{x^{n_2} + f_2(x)}$ where $a \neq 0$, $\deg f_1(x) < n_1$, and $\deg f_2(x) < n_2$. Then the reduced form of $f'(x)$ can be obtained by simplifying the expression

$$F(x) = \frac{(an_1x^{n_1-1} + {f_1}'(x))(x^{n_2} + f_2(x)) - (ax^{n_1} + f_1(x))(n_2x^{n_2-1} + {f_2}'(x))}{(x^{n_2} + f_2(x))^2}.$$

We first expand the numerator and denominator of the previous expression to write it in the form

$$F(x) = \frac{a(n_1 - n_2)x^{n_1+n_2-1} + g_1(x)}{x^{2n_2} + g_2(x)}$$

where $\deg g_1(x) < n_1 + n_2 - 1$ and $\deg g_2 < 2n_2$. The numerator and denominator degrees of $f'(x)$ then satisfy ${n_1}' + 2n_2 = {n_2}' + n_1 + n_2 - 1$, and it follows that ${n_1}' - {n_2}' = n_1 - n_2 - 1$. □

Theorem 4.1. *Let $f(x)$ be a complex rational function with numerator degree n_1 and denominator degree n_2. Let d be the greatest proper divisor of $n = \deg f$. If $|n_1 - n_2| > 0$ is divisible by a prime number $p > d$, then $f(x)$ is prime. If $|n_1 - n_2| > 0$ is divisible by a prime number $p = d$ and $f(x) = g\big(h(x)\big)$ is composite, then either $g(x)$ or $h(x)$ is a polynomial.*

Proof. Suppose that $f(x)$ is composite. There exist complex rational functions $g(x)$ and $h(x)$ of degrees $m, k \geq 2$ respectively such that $f(x) = g\big(h(x)\big)$ and $h(x)$ has larger numerator degree than denominator degree. Let m_1 and k_1 be the numerator degrees of $g(x)$ and $h(x)$ respectively, and let m_2 and k_2 be the denominator degrees of $g(x)$ and $h(x)$ respectively. Assume without loss of generality that $n_1 > n_2$, then $n_1 - n_2 = (m_1 - m_2)(k_1 - k_2)$ and it follows that $m_1 > m_2$.

To prove the first claim, we assume that $p > d$. Since $p|(n_1 - n_2)$ where $n_1 - n_2 = (m - m_2)(k - k_2)$, we have either $p|(m - m_2)$ or $p|(k - k_2)$. Then either $p \leq m - m_2 \leq m \leq d < p$ or $p \leq k - k_2 \leq k \leq d < p$, both cases yielding a contradiction. Therefore $f(x)$ is prime.

To prove the second claim, we assume that $p = d$. Since $p|(n_1 - n_2)$, we have either $p|(m - m_2)$ or $p|(k - k_2)$. Then either $d = p \leq m - m_2 \leq d - m_2$ so that $m_2 = 0$ and $g(x)$ is a polynomial, or $d = p \leq k - k_2 \leq d - k_2$ so that $k_2 = 0$ and $h(x)$ is a polynomial. □

Corollary 4.2. *Let $f(x)$ be a complex rational function of degree n and let d be the greatest proper divisor of n. If $f(x)$ has a zero or a pole whose multiplicity is divisible by a prime number $p > d$, then $f(x)$ is prime.*

Proof. Let $f(x)$ have numerator degree n_1, denominator degree n_2, and let

$$f(x) = \frac{c\prod_{i=1}^{m_1}(x - \alpha_i)^{a_i}}{\prod_{j=1}^{m_2}(x - \beta_j)^{b_j}}.$$

We first consider when $f(x)$ has a zero whose multiplicity is divisible by a prime number $p > d$, and we assume without loss of generality that this zero is α_1 which has multiplicity a_1. We define the unit $\mu(x) = \frac{\alpha_1 x + 1}{x}$ where $\alpha_1 \cdot 0 - 1 \cdot 1 = -1 \neq 0$, then

$$f\big(\mu(x)\big) = \frac{cx^{n-n_1}\prod_{i=1}^{m_1}\big((\alpha_1 x + 1) - \alpha_i x\big)^{a_i}}{x^{n-n_2}\prod_{j=1}^{m_2}\big((\alpha_1 x + 1) - \beta_j x\big)^{b_j}} = \frac{cx^{n-n_1}\prod_{i=2}^{m_1}\big((\alpha_1 - \alpha_i)x + 1\big)^{a_i}}{x^{n-n_2}\prod_{j=1}^{m_2}\big((\alpha_1 - \beta_j)x + 1\big)^{b_j}}$$

has numerator degree N_1 and denominator degree N_2 satisfying $N_1 + (n - n_2) + n_2 = N_2 + (n - n_1) + (n_1 - a_1)$. Then $N_2 - N_1 = a_1$ is divisible by $p > d$, so that $f\big(\mu(x)\big)$ satisfies the conditions of Theorem 4.1 and is prime. Therefore $f(x)$ is also prime.

If $f(x)$ has a pole with multiplicity divisible by $p > d$, we consider the unit $\nu(x) = \frac{1}{x}$. Then $\nu\big(f(x)\big)$ will have a zero with multiplicity divisible by $p > d$, so that $\nu\big(f(x)\big)$ and $f(x)$ are prime. □

The remainder of this section is primarily dedicated to providing examples of prime rational functions. We compose these prime rational functions with units to provide examples of prime polynomials.

Theorem 4.2. *Let $f(x)$ be a complex rational function with numerator degree n_1 and denominator degree n_2, where n_1 and n_2 are relatively prime integers such that $n_1 > n_2$. If the denominator of $f(x)$ is of the form $(x-\gamma)^{n_2}$ for some $\gamma \in \mathbb{C}$, then $f(x)$ is prime.*

Proof. Suppose by contradiction that $f(x)$ is composite. There exist complex rational functions $g(x)$ and $h(x)$ such that $f(x) = g\big(h(x)\big)$, where $g(x)$ is prime and $h(x) = \frac{h_1(x)}{h_2(x)}$ satisfies $\deg h_1(x) > \deg h_2(x)$. Let $k_1 = \deg h_1$ and $k_2 = \deg h_2$, and let

$$g(x) = \frac{c \cdot \prod_{i=1}^{m_1}(x-\alpha_i)}{\prod_{j=1}^{m_2}(x-\beta_j)}.$$

Since $n_1 > n_2$ and $k_1 > k_2$, it follows from Lemma 4.4 that $m_1 > m_2$. Then $f(x)$ is given by the expression

$$f(x) = \frac{c \cdot \prod_{i=1}^{m_1}\big(h_1(x) - \alpha_i h_2(x)\big)}{h_2(x)^{m_1-m_2}\prod_{j=1}^{m_2}\big(h_1(x)-\beta_j h_2(x)\big)}.$$

The denominator of $f(x)$ is $(x-\gamma)^{n_2}$, thus we obtain

$$(x-\gamma)^{n_2} = h_2(x)^{m_1-m_2}\prod_{j=1}^{m_2}\big(h_1(x)-\beta_j h_2(x)\big).$$

The linear factor $(x-\gamma)$ must then divide either $h_2(x)^{m_1-m_2}$ or $\prod_{j=1}^{m_2}\big(h_1(x) - \beta_j h_2(x)\big)$, but this factor cannot divide both as this implies that $(x-\gamma)$ will also divide $h_1(x)$ where $h(x)$ has no linear factor dividing both its numerator and its denominator. Thus we obtain two cases: $(x-\gamma)^{n_2} = h_2(x)^{m_1-m_2}$ and $\prod_{j=1}^{m_2}\big(h_1(x)-\beta_j h_2(x)\big)$ is a non-zero constant, or $(x-\gamma)^{n_2} = \prod_{j=1}^{m_2}\big(h_1(x)-\beta_j h_2(x)\big)$ and $h_2(x)^{m_1-m_2}$ is a non-zero constant.

i) If $h_2(x)^{m_1-m_2}$ is constant, then $h_2(x)$ is constant since $m_1 > m_2$, and $h(x)$ is a polynomial. Then $f(x) = g\big(h(x)\big)$ has numerator degree $n_1 = m_1k_1$ and denominator degree $n_2 = m_2k_1$ contradicting n_1 and n_2 being relatively prime.

ii) If $\prod_{j=1}^{m_2}\big(h_1(x) - \beta_j h_2(x)\big)$ is constant, then $m_2 = 0$ or $h_1(x) - \beta_j h_2(x) = c_j \in \mathbb{C}^*$ for $j = 1, ..., m_2$. We reject $m_2 = 0$, as this would imply that $f(x)$ has numerator degree $n_1 = m_1k_1$ and denominator degree $n_2 = m_1k_2$, contradicting n_1 and n_2 being relatively prime. We now consider the remaining possibility by choosing any two values β_{j_1} and β_{j_2} where $1 \le j_1, j_2 \le m_2$. We solve the expressions $h_1(x) - \beta_{j_1}h_2(x) = c_{j_1}$ and $h_1(x) - \beta_{j_2}h_2(x) = c_{j_2}$ for $h_1(x)$ to obtain

$$c_{j_1} + \beta_{j_1}h_2(x) = c_{j_2} + \beta_{j_2}h_2(x).$$

It follows that $c_{j_1} - c_{j_2} = (\beta_{j_2} - \beta_{j_1})h_2(x)$. Since $h_2(x)$ is not constant, we have $c_{j_1} = c_{j_2}$ and $\beta_{j_1} = \beta_{j_2}$ for every pair j_1 and j_2. We set $\beta_j = \beta$ and

$c_j = c$ for all $j = 1, ..., m_2$. Now $h_1(x) = c + \beta h_2(x)$, and we let

$$\nu(x) = c + \beta x, \quad \mu(x) = \frac{\nu(x)}{x}, \quad \text{and} \quad G(x) = \frac{G_1(x)}{G_2(x)} = g\big(\mu(x)\big)$$

so that $h_1(x) = \nu\big(h_2(x)\big)$ and $f(x) = G\big(h_2(x)\big)$. We note that $\mu(x)$ is a unit since $\beta \cdot 0 - c \cdot 1 = -c \neq 0$.

If $k_2 > 1$, then $f(x)$ has numerator degree $n_1 = \deg G_1 \cdot k_2$ and denominator degree $n_2 = \deg G_2 \cdot k_2$, contradicting n_1 and n_2 being relatively prime integers. If $k_2 = 1$, then $h_2(x)$ is a unit. Since $g(x)$ is prime, it follows that $G(x)$ is prime and therefore $f(x)$ is prime.

All possible cases have been considered, and we conclude that $f(x)$ is prime. □

Corollary 4.3. *Let $f(x) = (x-\alpha_1)^{e_1}(x-\alpha_2)^{e_2}$ be a complex polynomial such that $e_1, e_2 \geq 1$ and $\alpha_1 \neq \alpha_2$. Then $f(x)$ is prime if and only if e_1 and e_2 are relatively prime.*

Proof. Suppose that e_1 and e_1 are not relatively prime. There exists an integer $b \geq 2$ such that $e_1 = a_1 b$ and $e_2 = a_2 b$ for some positive integers a_1 and a_2. We can then write $g(x) = x^b$ and $h(x) = (x-\alpha_1)^{a_1}(x-\alpha_2)^{a_2}$, where both $g(x)$ and $h(x)$ have degree at least 2. Then $f(x) = g\big(h(x)\big)$ is composite.

Conversely, suppose that e_1 and e_2 are relatively prime. Then e_2 and $e_1 + e_2$ are relatively prime as well. We define the units $\nu(x) = \frac{1}{x}$ and $\mu(x) = \frac{\alpha_1 x + 1}{x}$ where $\alpha_1 \cdot 0 - 1 \cdot 1 = -1 \neq 0$, then the function

$$\nu\Big(f\big(\mu(x)\big)\Big) = \nu\left(\frac{\big((\alpha_1 x + 1) - \alpha_1 x\big)^{e_1}\big((\alpha_1 x + 1) - \alpha_2 x\big)^{e_2}}{x^{e_1+e_2}}\right) = \frac{x^{e_1+e_2}}{\big((\alpha_1 - \alpha_2)x + 1\big)^{e_2}}$$

is prime by Theorem 4.2 since e_2 and $e_1 + e_2$ are relatively prime. Therefore $f(x)$ is prime. □

Theorem 4.3. *Let $f(x) = (x-\alpha_1)^{e_1}(x-\alpha_2)^{e_2}(x-\alpha_3)^{e_3}$ be a complex polynomial of degree n such that α_1, α_2 and α_3 are distinct complex numbers and $e_1, e_2, e_3 \geq 1$. If e_1, e_2, and e_3 are pairwise relatively prime integers all relatively prime to n, then $f(x)$ is prime.*

Proof. Suppose by contradiction that $f(x)$ is composite. Then there exist polynomials $g(x)$ and $h(x)$ with degrees at least 2 such that $f(x) = g\big(h(x)\big)$. We write

$$g(x) = \prod_{i=1}^{m} (x - \beta_i)^{b_i}$$

where $\beta_1, ..., \beta_m$ are all of the roots of $g(x)$. Then

$$f(x) = \prod_{i=1}^{m} \big(h(x) - \beta_i\big)^{b_i}.$$

Since $h(x) - \beta_i$ and $h(x) - \beta_j$ do not have any roots in common when $i \neq j$, it follows that $1 \leq m \leq 3$.

If $m = 1$, then $f(x) = \big(h(x) - \beta_1\big)^{b_1}$, and we obtain $h(x) - \beta_1 = (x - \alpha_1)^{r_1}(x - \alpha_2)^{r_2}(x - \alpha_3)^{r_3}$ for some integers r_1, r_2, and r_3. Then $e_1 = r_1 b_1$, $e_2 = r_2 b_1$, and $e_3 = r_3 b_1$ so that b_1 divides the pairwise relatively prime integers e_1, e_2, and e_3. Thus $b_1 = 1$ and $\deg g = 1$, yielding a contradiction.

If $m = 2$, then $f(x) = \big(h(x) - \beta_1\big)^{b_1}\big(h(x) - \beta_2\big)^{b_2}$. We assume without loss of generality that $h(x) - \beta_1 = (x - \alpha_1)^{r_1}$ and $h(x) - \beta_2 = (x - \alpha_2)^{r_2}(x - \alpha_3)^{r_3}$ for some integers r_1, r_2, and r_3. Then $r_1 = \deg h = r_2 + r_3$, $e_1 = r_1 b_1$, $e_2 = r_2 b_2$, and $e_3 = r_3 b_2$ so that b_2 divides the relatively prime integers e_2 and e_3. Thus $b_2 = 1$ and $r_1 = r_2 + r_3 = e_2 + e_3$. It follows that $r_1 = \deg h > 1$ divides both e_1 and $n = e_1 + e_2 + e_3$, yielding a contradiction.

If $m = 3$, then $f(x) = \big(h(x) - \beta_1\big)^{b_1}\big(h(x) - \beta_2\big)^{e_2}\big(h(x) - \beta_3\big)^{e_3}$. We assume without loss of generality that $h(x) - \beta_1 = (x - \alpha_1)^{r_1}$, $h(x) - \beta_2 = (x - \alpha_2)^{r_2}$, and $h(x) - \beta_3 = (x - \alpha_3)^{r_3}$ where $r_1 = r_2 = r_3 = \deg h$. Then $e_1 = r_1 b_1$, $e_2 = r_2 b_2$, and $e_3 = r_3 b_3$, so that $\deg h > 1$ divides the pairwise relatively prime integers e_1, e_2, and e_3, yielding a contradiction.

All of the possible values of m have been rejected, therefore $f(x)$ is prime. □

Theorem 4.4. *Let $f(x)$ be a complex rational function with numerator degree n_1 and denominator degree n_2. Let d be the greatest proper divisor of $n = \deg f$. If $n_2 - n_1 > d$ and $n_2 - n_1$ is relatively prime to n_1 as well as the multiplicities of all zeros of $f(x)$, then $f(x)$ is prime.*

Proof. Suppose by contradiction that $f(x)$ is composite. There exist complex rational functions $g(x)$ and $h(x)$ such that $f(x) = g\big(h(x)\big)$. Let

$$f(x) = \frac{a\prod_{i=1}^{N}(x - a_i)^{e_i}}{f_2(x)}, \qquad g(x) = \frac{b\prod_{i=1}^{m_1}(x - \alpha_i)}{\prod_{j=1}^{m_2}(x - \beta_j)}, \qquad h(x) = \frac{h_1(x)}{h_2(x)}$$

where $k_1 = \deg h_1 > \deg h_2 = k_2$. Since $n_2 - n_1 > d > 0$, we conclude from Lemma 4.4 that $n_2 - n_1 = (m_2 - m_1)(k_1 - k_2)$ so that $m_2 > m_1$, and we obtain

$$\frac{a\prod_{i=1}^{N}(x - a_i)^{e_i}}{f_2(x)} = \frac{bh_2(x)^{m_2 - m_1}\prod_{i=1}^{m_1}\big(h_1(x) - \alpha_i h_2(x)\big)}{\prod_{j=1}^{m_2}\big(h_1(x) - \beta_j h_2(x)\big)}.$$

If $m_2 - m_1 = 1$, then $n_2 - n_1 = k_1 - k_2 \leq k_1 \leq d$ yields a contradiction to $n_2 - n_1 > d$, so we have $m_2 - m_1 \geq 2$. Since n_1 and $n_2 - n_1$ are relatively prime, so are n_1 and n_2. It follows that $h_2(x)$ cannot be constant, since if $h(x)$ is a polynomial, its degree must divide both n_1 and n_2. Then $h_2(x)$ has degree at least 1 and $h_2(x)^{m_2 - m_1}$ divides $a\prod_{i=1}^{N}(x - a_i)^{e_i}$, where $m_2 - m_1$ must then divide e_i for some $i = 1, ..., N$. $m_2 - m_1$ also divides $n_2 - n_1$, which contradicts $n_2 - n_1$ being relatively prime to the multiplicities of all of the zeros of $f(x)$, therefore $f(x)$ is prime. □

The following example shows that the condition $n_2 - n_1$ relatively prime to the multiplicities of all of the zeros of $f(x)$ is necessary.

Example 4.1. Let

$$f(x) = \frac{(x-3)^4(x^3-3x^2+2x+2)}{(x-1)^{15}}.$$

The zeros of x^3-3x^2+2x+2 all have multiplicity 1, thus $n_2-n_1=8$ is relatively prime to all of these multiplicities as well as $n_1=7$. The condition $n_2-n_1>d=5$ is also satisfied. $n_2-n_1=8$ is not relatively prime to 4, and this is sufficient for the above theorem to fail, for $f(x)=g\big(h(x)\big)$ where $g(x)=\frac{x-1}{x^5}$ and $h(x)=\frac{(x-1)^3}{x-3}$.

Corollary 4.4. *Let $f(x)$ be a complex polynomial of degree n with at least two distinct roots and let d be the greatest proper divisor of n. If there exists a root of $f(x)$ with multiplicity $e>d$ such that e is relatively prime to n as well as the multiplicities of all other roots of $f(x)$, then $f(x)$ is prime.*

Proof. Let

$$f(x) = a\prod_{i=1}^{N}(x-\alpha_i)^{e_i}$$

where $N\geq 2$ and assume without loss of generality that α_1 is the root with multiplicity $e_1>d$ which is relatively prime to n and to all other multiplicities. Define the unit $\mu(x)=\frac{\alpha_1 x+1}{x}$ where $\alpha_1\cdot 0-1\cdot 1=-1\neq 0$, then the function

$$f\big(\mu(x)\big)=\frac{a\prod_{i=1}^{N}\big((\alpha_1x+1)-\alpha_ix\big)^{e_i}}{x^n}=\frac{a\prod_{i=2}^{N}\big((\alpha_1-\alpha_i)x+1\big)\big)^{e_i}}{x^n}$$

has numerator degree $n_1=n-e_1$ and denominator degree $n_2=n$. Since e_1 and n are relatively prime, so are n_1 and n_2. Then $n_2-n_1=e_1>d$ and n_2-n_1 is relatively prime to n_1 as well as e_i for all $i=2,...,N$. Then $f\big(\mu(x)\big)$ satisfies the conditions of Theorem 4.4 and is prime. Therefore $f(x)$ is also prime. □

4.2. **Critical values of composite rational functions.**

Let $f(x)$ be a non-constant complex rational function. Let $x_0\in\mathbb{C}$ lie in the domain of the function $f(x)$. The smallest integer $i\geq 1$ such that $f^{(i)}(x_0)\neq 0$ is called the valency of $f(x)$ at x_0 and is denoted by $v_f(x_0)$. x_0 is called a critical point of $f(x)$ provided that $v_f(x_0)\geq 2$. The number $t_0\in\mathbb{C}$ is a critical value of $f(x)$ if there exists a critical point x_0 of $f(x)$ such that $f(x_0)=t_0$.

Theorem 4.5. *Let $f(x)$ be a complex rational function of degree n and let d be the greatest proper divisor of n. Suppose that $f(x)$ has a critical point $x_0\in\mathbb{C}$ such that its valency $v_f(x_0)$ is divisible by a prime number $p>d$, then $f(x)$ is prime.*

Proof. Let $v_f(x_0)=e$ be the valency of some critical point x_0 of $f(x)$ such that e is divisible by a prime number $p>d$. It follows that $f^{(i)}(x_0)=0$ for all $i=1,...,e-1$ and $f^{(e)}(x_0)\neq 0$. Then $f'(x)$ has a zero of order $e-1$ at x_0, so there exists a rational function $q(x)$ such that $f'(x)=(x-x_0)^{e-1}q(x)$ where $q(x_0)\neq 0$. Then there exists a rational function $y(x)$ such that $f(x)-f(x_0)=(x-x_0)^e y(x)$ where $y(x_0)\neq 0$. We define the unit $\mu(x)=x-f(x_0)$, then x_0 is a zero of $\mu\big(f(x)\big)=(x-x_0)^e y(x)$

with multiplicity e divisible by the prime number $p > d$. Thus $\mu\big(f(x)\big)$ is prime by Corollary 4.2 and $f(x)$ is prime as well. $\square$

A useful tool in the study of a polynomial's critical values is the discriminant, which can be described through the resultant of two polynomials. Let R be an integral domain and let K be its field of fractions. Let $u(x) = a_n x^n + \cdots + a_1 x + a_0$ and $v(x) = b_m x^m + \cdots + b_1 x + b_0$ be polynomials over R. Let $\alpha_1, ..., \alpha_n$ and $\beta_1, ..., \beta_m$ be the all of the roots of $u(x)$ and $v(x)$ respectively in an algebraic closure of K. The resultant of $u(x)$ and $v(x)$ is given by

$$\mathrm{Res}_x\big(u(x), v(x)\big) = a_n^m b_m^n \prod_{i=1}^{n} \prod_{j=1}^{m} (\alpha_i - \beta_j).$$

We then define the discriminant of the polynomial $u(x)$ by

$$D\big(u(x)\big) = \frac{(-1)^{n(n-1)/2}}{a_n} \mathrm{Res}_x\big(u(x), u'(x)\big).$$

We extend this concept to rational functions as follows. Let K be a field, and let $u(x) = \frac{u_1(x)}{u_2(x)}$ and $v(x) = \frac{v_1(x)}{v_2(x)}$ be rational functions over K in their most reduced forms, where we assume without loss of generality that $u_2(x)$ and $v_2(x)$ are monic. We then define the resultant of the rational functions $u(x)$ and $v(x)$ by

$$\mathrm{Res}_x\big(u(x), v(x)\big) = \mathrm{Res}_x\big(u_1(x), v_1(x)\big).$$

From this definition, we may obtain information regarding the critical values of rational functions similar to what can be obtained for polynomials from the standard definition of the resultant. We require the following properties, which are analogous to those for the resultant of two polynomials found in [1]. The proof of the properties are omitted.

i) Let $u(x)$ and $v(x)$ be rational functions as described above. Let $u_1(x) = a_n x^n + \cdots + a_1 x + a_0$ and $v_1(x) = b_m x^m + \cdots + b_1 x + b_0$ be polynomials with roots $\alpha_1, ..., \alpha_n$ and $\beta_1, ..., \beta_m$ respectively in an algebraic closure of K. Then

$$\mathrm{Res}_x\big(v(x), u(x)\big) = (-1)^{nm} \mathrm{Res}_x\big(u(x), v(x)\big).$$

ii) Under the same hypotheses as the first property,

$$\mathrm{Res}_x\big(u(x), v(x)\big) = a_n^m \prod_{i=1}^{n} v_1(\alpha_i).$$

iii) $u(x)$ and $v(x)$ have a zero in common if and only if

$$\mathrm{Res}_x\big(u(x), v(x)\big) = 0.$$

iv) For an additional rational function $w(x) = \frac{w_1(x)}{w_2(x)}$ over K, we have

$$\mathrm{Res}_x\big(u(x), v(x)w(x)\big) = \mathrm{Res}_x\big(u_1(x), p(x)\big) \mathrm{Res}_x\big(u_1(x), q(x)\big),$$

where $p(x)$ is the quotient obtained from dividing $v_1(x)$ by the monic greatest common divisor of $v_1(x)$ and $w_2(x)$, and $q(x)$ is the quotient obtained from dividing $w_1(x)$ by the monic greatest common divisor of $w_1(x)$ and $v_2(x)$.

Let $f(x)$ be a complex rational function and let $f'(x)$ be the derivative of $f(x)$. We write $f(x) = \frac{f_1(x)}{f_2(x)}$ and $f'(x) = \frac{\varphi_1(x)}{\varphi_2(x)}$, where we assume without loss of generality that $\varphi_2(x)$ is monic. This expression for $f'(x)$ is the most reduced expression of $F(x) = \frac{f_1{}'(x)f_2(x)-f_1(x)f_2{}'(x)}{f_2(x)^2}$, and it follows that $\varphi_2(x)$ divides $f_2(x)^2$. Since the reduced expression for $f'(x)$ is obtained by simplifying linear factors from the numerator and denominator of $F(x)$, where $f_1(x)$ and $f_2(x)$ share no common linear factors, the only such linear factors which can be simplified must divide both $f_2(x)$ and $f_2{}'(x)$. We conclude that $f_2(x)$ divides $\varphi_2(x)$ and thus $f(x)$ and $f'(x)$ have the same domain.

Let $\beta_1, ..., \beta_m$ be all of the zeros of $f'(x)$, then β_i is in the domain of $f'(x)$, and also the domain of $f(x)$, for $i = 1, ..., m$. Let t be a variable, let b be the leading coefficient of $\varphi_1(x)$, and let $n = \deg f(x)$. Consider the function $R(t) = \mathrm{Res}_x\big(f(x) - t, f'(x)\big)$. Using the properties of the resultant, we have the following:

$$\begin{aligned} R(t) &= \mathrm{Res}_x\left(\frac{f_1(x) - tf_2(x)}{f_2(x)}, \frac{\varphi_1(x)}{\varphi_2(x)}\right) \\ &= \mathrm{Res}_x\left(f_1(x) - tf_2(x), b\prod_{i=1}^{m}(x - \beta_i)\right) \\ &= (-1)^{nm}b^n\prod_{i=1}^{m}\big(f_1(\beta_i) - tf_2(\beta_i)\big) \\ &= (-1)^{nm}b^n\prod_{i=1}^{m}f_2(\beta_i)\prod_{i=1}^{m}\big(f(\beta_i) - t\big) \end{aligned}$$

We remark that since β_i is a zero of $f'(x)$, β_i is a critical point of $f(x)$ and $f(\beta_i)$ is a critical value of $f(x)$ for $i = 1, ..., m$. It immediately follows that $R(t_0) = 0$ if and only if t_0 is a critical value of $f(x)$. Similar to the definition of the multiplicity of a critical value of a polynomial found in [1], we define the multiplicity of the critical value t_0 as the multiplicity of t_0 as a root of $R(t)$, and we call a critical value with multiplicity equal to one a simple critical value.

Lemma 4.6. *Let $f(x)$ be a composite complex rational function of degree n and let d be the greatest proper divisor of n. Let $f(x) = g\big(h(x)\big)$ where $h(x) = \frac{h_1(x)}{h_2(x)}$ satisfies $k = \deg h_1(x) > \deg h_2(x)$, and let n_1 and n_2 be the numerator and denominator degrees of $f(x)$ respectively. Let $R(t)$ be the resultant of $f(x) - t$ and $f'(x)$, then there exists $c \in \mathbb{C}^*$, a non-negative integer ℓ, and a polynomial $p(x)$ dividing the numerator of $h'(x)$ such that*

$$R(t) = ct^{\ell}\Big(\mathrm{Res}_x\big(g(x) - t, g'(x)\big)\Big)^k\mathrm{Res}_x\big(f(x) - t, p(x)\big),$$

where $\ell > 0$ if n_1 and n_2 are relatively prime integers satisfying $n_2 - n_1 > d$.

Proof. We will write $u(t) \sim v(t)$ to denote that the functions $u(t)$ and $v(t)$ are equal up to multiplication by a constant. Let

$$g'(x) = \frac{b\prod_{i=1}^{m_1}(x-\alpha_i)}{\prod_{j=1}^{m_2}(x-\beta_j)}, \quad h'(x) = \frac{h_1{}'(x)h_2(x) - h_1(x)h_2{}'(x)}{h_2(x)^2} = \frac{q_1(x)}{h_2(x)q_2(x)}$$

where $q_1(x)$ and $q_2(x)$ share no common factor, and let $m = \deg g'(x)$. Then

$$f'(x) = \frac{bh_2(x)^{m-m_1}q_1(x)\prod_{i=1}^{m_1}\big(h_1(x)-\alpha_i h_2(x)\big)}{h_2(x)^{m-m_2+1}q_2(x)\prod_{j=1}^{m_2}\big(h_1(x)-\beta_j h_2(x)\big)}.$$

The only linear factors which can be simplified in this expression for $f'(x)$ are shared factors between $h_2(x)^{m-m_1}$ and $h_2(x)^{m-m_2+1}q_2(x)$ or shared factors between $q_1(x)$ and $\prod_{j=1}^{m_2}\big(h_1(x)-\beta_j h_2(x)\big)$. We let $H(x)$ be the quotient obtained from dividing $h_2(x)^{m-m_1}$ by the monic greatest common divisor of $h_2(x)^{m-m_1}$ and $h_2(x)^{m-m_2+1}q_2(x)$, and we let $p(x)$ be the quotient obtained from dividing $q_1(x)$ by the monic greatest common divisor of $q_1(x)$ and $\prod_{j=1}^{m_2}\big(h_1(x)-\beta_j h_2(x)\big)$. Letting $R(t)$ be the resultant of $f(x)-t$ and $f'(x)$, we then have

$$R(t) = \mathrm{Res}_x\left(f_1(x) - tf_2(x), bH(x)p(x)\prod_{i=1}^{m_1}\Big(h_1(x)-\alpha_i h_2(x)\Big)\right).$$

We consider the above expression as a product of three factors. The first factor is

$$\mathrm{Res}_x\left(f_1(x) - tf_2(x), \prod_{i=1}^{m_1}\Big(h_1(x)-\alpha_i h_2(x)\Big)\right).$$

For each $i = 1, ..., m_1$, the equation $h_1(x) - \alpha_i h_2(x)$ has k solutions $s_{i,1}, ..., s_{i,k}$. For any index r, the solution $s_{i,r}$ then satisfies $h_1(s_{i,r}) - \alpha_i h_2(s_{i,r}) = 0$, so that $h(s_{i,r}) = \alpha_i$. Since α_i is a zero of $g'(x)$ for $i = 1, ..., m_1$, each of these zeros must also be in the domain of $g(x)$ and $g(\alpha_i) = g\big(h(s_{i,r})\big) = f(s_{i,r})$ for $i = 1, ..., m_1$ and $r = 1, ..., k$. We then have

$$\begin{aligned}
\mathrm{Res}_x\left(f_1(x) - tf_2(x), \prod_{i=1}^{m_1}\Big(h_1(x)-\alpha_i h_2(x)\Big)\right) &\sim \prod_{i=1}^{m_1}\left(\prod_{r=1}^{k}\Big(f_1(s_{i,r}) - tf_2(s_{i,r}\Big)\right)\\
&\sim \prod_{i=1}^{m_1}\left(\prod_{r=1}^{k}\Big(f(s_{i,r}) - t\Big)\right)\\
&\sim \prod_{i=1}^{m_1}\left(\prod_{r=1}^{k}\Big(g(\alpha_i) - t\Big)\right)\\
&\sim \left(\prod_{i=1}^{m_1}\Big(g(\alpha_i) - t\Big)\right)^k\\
&\sim \Big(\mathrm{Res}_x\big(g(x)-t, g'(x)\big)\Big)^k.
\end{aligned}$$

The second factor is

$$\mathrm{Res}_x\Big(f_1(x) - tf_2(x), H(x)\Big).$$

If $m_1 \geq m_2$, then $H(x)$ is constant and this factor is constant. If $m_2 > m_1$, $H(x)$ will not be constant if $h_2(x)$ is not constant and $m_2 - m_1 > 2$. In this case, we let $\ell = \deg H$ and let $s_1, ..., s_\ell$ be all of the roots of $H(x)$. Since $H(x)$ divides $h_2(x)^{m_2-m_1}$, every such a root s of $H(x)$ satisfies $h_2(s) = 0$ and so $|h(s)|$ is infinite. Since $m_2 > m_1$, the function $f(x) = g\big(h(x)\big)$ has a value of zero at $x = s_r$ for $r = 1, ..., \ell$. Then we have

$$\begin{aligned}\mathrm{Res}_x\Big(f_1(x) - tf_2(x), H(x)\Big) &\sim \prod_{r=1}^{\ell}\Big(f_1(s_r) - tf_2(s_r)\Big)\\ &\sim \prod_{r=1}^{\ell}\Big(f(s_r) - t\Big)\\ &\sim (-t)^\ell.\end{aligned}$$

In particular, if $n_2 - n_1 > d$ where n_1 and n_2 are relatively prime integers, from Lemma 4.4 we have $d < n_2 - n_1 = (\deg g_2 - \deg g_1)(\deg h_1 - \deg h_2) \leq (\deg g_2 - \deg g_1)d$, so that $\deg g_2 - \deg g_1 > 1$. From Lemma 4.5, we then have $m_2 - m_1 = -(\deg g_1 - \deg g_2 - 1) = \deg g_2 - \deg g_1 + 1 > 2$. $h_2(x)$ cannot be constant, as this would imply that $k = \deg h_1$ would divide both n_1 and n_2 yielding a contradiction. It follows that $H(x)$ will not be constant in this case, and by our definition of the function $H(x)$ we have $\ell = \deg H \geq (m_2 - m_1 - 1)k_2 - \deg q_2 \geq (m_2 - m_1 - 2)k_2 = (\deg g_2 - \deg g_1 - 1)k_2$.

The final factor is

$$\mathrm{Res}_x\big(f(x) - t, b \cdot p(x)\big),$$

and we conclude that for some non-zero complex number c we have

$$R(t) = ct^\ell\Big(\mathrm{Res}_x\big(g(x) - t, g'(x)\big)\Big)^k \mathrm{Res}_x\big(f(x) - t, p(x)\big)$$

where ℓ is a non-negative integer such that $\ell > 0$ when n_1 and n_2 are relatively prime integers satisfying $n_2 - n_1 > d$. □

Corollary 4.5. *Let $f(x)$ be a composite complex rational function of degree n which has a right composition factor of degree k. Let $R(t)$ be the resultant of $f(x) - t$ and $f'(x)$. Then there exists a non-negative integer ℓ and polynomials $A(t)$ and $B(t)$ such that $R(t) = t^\ell[A(t)]^k B(t)$ and $\deg B(t) \leq 2k - 1$. Moreover, if d is the greatest proper divisor of n, n_1 and n_2 are the numerator and denominator degrees of $f(x)$ respectively, and n_1 and n_2 are relatively prime integers such that $n_2 - n_1 > d$, then $\ell > 0$.*

Proof. We write $u(t) \sim v(t)$ to denote that the functions $u(t)$ and $v(t)$ are equal up to multiplication by a constant. Since $f(x)$ is composite with right composition factor of degree k, there exist complex rational functions $g(x)$ and $h(x) = \frac{h_1(x)}{h_2(x)}$ such

that $f(x) = g\big(h(x)\big)$ and $k = \deg h_1(x) > \deg h_2(x)$. Then there exists $c \in \mathbb{C}^*$, a non-negative integer ℓ, and a polynomial $p(x)$ which divides the numerator of $h'(x)$, such that $R(t) = ct^\ell \Big(\mathrm{Res}_x\big(g(x)-t, g'(x)\big)\Big)^k \mathrm{Res}_x\big(f(x)-t, p(x)\big)$, and where $\ell > 0$ if n_1 and n_2 are relatively prime integers satisfying $n_2 - n_1 > d$.

Setting $A(t) = \Big(\mathrm{Res}_x\big(g(x)-t, g'(x)\big)\Big)$ and $B(t) = c \cdot \mathrm{Res}_x\big(f(x)-t, p(x)\big)$ yields the desired expression for $R(t)$, so it only remains to show that $\deg B(t) \leq 2k-1$. We let $p(x) = b\prod_{i=1}^{r}(x-\alpha_i)$. Since $p(x)$ divides the numerator of $h'(x)$, it follows that $p(x)$ must divide the numerator of $\frac{h_1{}'(x)h_2(x)-h_1(x)h_2{}'(x)}{h_2(x)^2}$, so that $r \leq \deg h_1(x) + \deg h_2(x) - 1 \leq 2k-1$. Writing $B(t)$ explicitly, we obtain

$$\begin{aligned} B(t) &= c \cdot \mathrm{Res}_x\left(\frac{f_1(x)-tf_2(x)}{f_2(x)}, p(x)\right) \\ &\sim \mathrm{Res}_x\left(f_1(x)-tf_2(x), \prod_{i=1}^{r}(x-\alpha_i)\right) \\ &\sim \prod_{i=1}^{r}\big(f_1(\alpha_i)-tf_2(\alpha_i)\big) \end{aligned}$$

so that $\deg B(t) \leq r \leq 2k-1$. □

The following two results show that the polynomial $R(t)$ obtained by taking the resultant of a complex rational function $f(x)-t$ and its derivative can be useful in determining whether $f(x)$ is prime. The first result considers the non-zero critical values of $f(x)$, and its proof follows the same method as the proof of Theorem 1 from [1]. The second result considers only the critical value zero.

Theorem 4.6. *Let $f(x)$ be a complex rational function of degree n and let d be the greatest proper divisor of n. Suppose that $f(x)$ has at least $2d$ non-zero simple critical values, then $f(x)$ is prime.*

Proof. Suppose by contradiction that $f(x)$ is composite. There exist complex rational functions $g(x)$ and $h(x)$ of degrees $m, k \geq 2$ respectively such that $f(x) = g\big(h(x)\big)$. We let $R(t)$ be the resultant of $f(x)-t$ and $f'(x)$, and we write $R(t) = t^\ell[A(t)]^k B(t)$ where ℓ is a non-negative integer and $\deg B(t) \leq 2k-1$. Let δ be the number of non-zero simple critical values of $f(x)$. Since these critical values must be roots of the polynomial $B(t)$, we obtain

$$2k-1 \geq \deg B(t) \geq \delta \geq 2d \geq 2k$$

which is a contradiction. Therefore $f(x)$ is prime. □

Theorem 4.7. *Let $f(x)$ be a complex rational function of degree n, let d be the greatest proper divisor of n, and let n_1 and n_2 be the numerator and denominator degrees of $f(x)$ respectively. If n_1 and n_2 are relatively prime integers such that $n_2 - n_1 > d$, and if zero is a critical value of $f(x)$ with multiplicity $e < \frac{n_2-n_1-d}{d}$,*

then $f(x)$ is prime. In particular, if zero is not a critical value of $f(x)$, then $f(x)$ is prime.

Proof. Suppose by contradiction that $f(x)$ is composite. There exist complex rational functions $g(x)$ and $h(x)$ such that $f(x) = g\big(h(x)\big)$ and where $h(x)$ has larger numerator degree than denominator degree. Let m_1 and k_1 be the numerator degrees of $g(x)$ and $h(x)$ respectively and let m_2 and k_2 be the denominator degrees of $g(x)$ and $h(x)$ respectively. Since we assume that $k_1 > k_2$ and $n_2 > n_1$, we have $(n_2 - n_1) = (m_2 - m_1)(k_1 - k_2)$. It follows that $m_2 > m_1$ and

$$m_2 - m_1 - 1 = \frac{n_2 - n_1}{k_1 - k_2} - 1 \geq \frac{n_2 - n_1}{k_1} - 1 \geq \frac{n_2 - n_1}{d} - 1 = \frac{n_2 - n_1 - d}{d}.$$

Since n_1 and n_2 are relatively prime, we know that $h(x)$ cannot be a polynomial as this would imply $\deg h$ divides both n_1 and n_2. Then $k_2 \geq 1$ and we obtain $(m_2 - m_1 - 1)k_2 \geq m_2 - m_1 - 1 \geq \frac{n_2 - n_1 - d}{d}$. We now let $R(t)$ be the resultant of $f(x) - t$ and $f'(x)$, and we write $R(t) = t^\ell [A(t)]^{k_1} B(t)$. From the arguments presented in the proof of Lemma 4.6, we have $\ell \geq (m_2 - m_1 - 1)k_2$. It follows that zero is a critical value of $f(x)$ of multiplicity at least $(m_2 - m_1 - 1)k_2$, but by assumption the multiplicity e of this critical value satisfies $e < \frac{n_2 - n_1 - d}{d} \leq (m_2 - m_1 - 1)k_2$ yielding a contradiction. □

The following result provides some examples of prime functions.

Proposition 4.3. *Let $f(x) = \frac{x^n + a}{x^m + b}$ where $a, b \in \mathbb{C}$ are not both zero, let d be the greatest proper divisor of $\deg f$, and let n and m be relatively prime positive integers such that $|n - m| > d$. Then $f(x)$ is prime.*

Proof. We assume without loss of generality that $n \leq m$, and we consider two cases.

Assume first that $a \neq 0$. Suppose by contradiction that $f(x)$ is composite. Since m and n are relatively prime integers, it follows that $n \neq m$ and thus $n < m$. Then zero must be a critical value of $f(x)$ by Lemma 4.6. We show that no critical point of $f(x)$ yields zero as a critical value.

If $b \neq 0$, then

$$f'(x) = \frac{x^{n-1}\big((n-m)x^m + (-am)x^{m-n} + (bn)\big)}{(x^m + b)^2}.$$

Let $\xi_1, ..., \xi_{m+n-1}$ be all of the zeros of $f'(x)$. Then $\xi_1, ..., \xi_{m+n-1}$ are the critical points of $f(x)$, and for each $i = 1, ..., m + n - 1$ we have either $\xi_i^{n-1} = 0$ or $(n-m)\xi_i^m + (-am)\xi_i^{m-n} + (bn) = 0$. A critical point ξ with $\xi^{n-1} = 0$ satisfies $\xi = 0$ and thus $f(\xi) = \frac{a}{b} \neq 0$. For the second case we proceed by contradiction, where we assume a critical point ξ satisfies $(n-m)\xi^m + (-am)\xi^{m-n} + (bn) = 0$. $f(\xi) = 0$ gives $\xi^n + a = 0$, so that $\xi^n = -a \neq 0$ and $(n-m)\xi^m + (m)\xi^{m-n}\xi^n + (bn) = n(\xi^m + b) = 0$. Then $\xi^m + b = 0$ yields a contradiction, since $f(x)$ has no linear factor dividing both its numerator and its denominator.

If $b = 0$, then

$$f'(x) = \frac{(n-m)x^n + (-am)}{x^{m+1}}.$$

Let $\xi_1, ..., \xi_n$ be all of the zeros of $f'(x)$. Then $\xi_1, ..., \xi_n$ are the critical points of $f(x)$, and for each $j = 1, ..., n$ we have $\xi_j^n = \frac{-am}{m-n}$. If $f(\xi_j) = 0$, then $-a = \frac{-am}{m-n}$ yields $m - n = m$ contradicting $n > 0$.

Therefore zero cannot be a critical value of the function $f(x)$ and we conclude that $f(x)$ is prime.

Assume now that $a = 0$. Then by assumption $b \neq 0$, and $f(x) = \frac{x^n}{x^m+b}$ is prime if and only if $F(x) = \frac{x^m+b}{x^n}$ is prime. Since m and n are relatively prime integers such that $m > n$, we conclude by Theorem 4.2 that $F(x)$ is prime and therefore $f(x)$ is prime. □

We conclude this section by providing some examples which show that, in general, knowing whether the numerator and denominator polynomials of a rational function $f(x)$ are prime or composite is not sufficient to conclude whether $f(x)$ is prime or composite.

Example 4.2. Let

$$f(x) = \frac{f_1(x)}{f_2(x)} = \frac{4x^3 + 6x^2 + 4x + 1}{x^4 - 2x^3 - x^2}.$$

Then $f_1(x)$ is prime, $f_2(x)$ is prime by Theorem 1 from [1] since all of its critical values are simple, and $f(x)$ is composite since it is the composition of $g(x) = -\frac{x^2-1}{x-2}$ and $h(x) = \frac{x^2+2x+1}{x^2}$.

Example 4.3. Let

$$f(x) = \frac{f_1(x)}{f_2(x)} = \frac{x^5 + 1}{x^3}.$$

Then $f_1(x)$ and $f_2(x)$ are both prime, and $f(x)$ is prime.

Example 4.4. Let

$$f(x) = \frac{f_1(x)}{f_2(x)} = \frac{x^2 + 1}{x^4}.$$

Then $f_1(x)$ is prime, $f_2(x)$ is composite, and $f(x)$ is composite.

Example 4.5. Let

$$f(x) = \frac{f_1(x)}{f_2(x)} = \frac{x^5 + 1}{x^4}.$$

Then $f_1(x)$ is prime, $f_2(x)$ is composite, and $f(x)$ is prime.

Example 4.6. Let

$$f(x) = \frac{f_1(x)}{f_2(x)} = \frac{x^9 + 1}{x^6}.$$

Then $f_1(x)$, $f_2(x)$, and $f(x)$ are all composite.

Example 4.7. Let

$$f(x) = \frac{f_1(x)}{f_2(x)} = \frac{x^9 + 1}{x^4}.$$

Then $f_1(x)$ and $f_2(x)$ are composite, and $f(x)$ is prime by Theorem 4.2.

5. Main Results on Integral Polynomials

In this chapter [2] we prove a result on the divisibility of integral polynomials. We also provide a way to determine if an integral polynomial is irreducible over $\mathbb{Z}$, and if not, we find one of its divisors. To begin, we recall the work of Nieto in [7]. Let $f, g \in \mathbb{Z}[x]$, then it is proved that g divides f if and only if $cont(g)$ divides $cont(f)$ and $g(n)$ divides $f(n)$ for infinitely many $n \in \mathbb{Z}$. We will show that a similar result holds where $g(n)$ need only divide $f(n)$ for a single integer n whose absolute value larger than a certain bound.

We recall the definitions of the content and the height of a polynomial.

Definition 5.1. *Let $f(x) = a_n x^n + \cdots + a_1 x + a_0 \in \mathbb{Z}[x]$ be a nonzero polynomial. The* content *of f, denoted by $cont(f)$, is the greatest common divisor of the integers $a_n, a_{n-1}, ..., a_0$.*

Definition 5.2. *Let $f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 \in \mathbb{C}[x]$. We define the height of f by*

$$H(f) = \max_{i=0...n}\{|a_i|\}.$$

The following result extends the usual division algorithm for polynomials over a field to polynomials in $\mathbb{Z}[x]$.

Proposition 5.1. *Let $f(x), g(x) \in \mathbb{Z}[x]$ where $g(x) \neq 0$, and let b be the leading coefficient of $g(x)$. Then there exists a non-negative integer $k \leq |\deg f - \deg g| + 1$ and polynomials $q(x)$ and $r(x)$ in $\mathbb{Z}[x]$ such that*

$$b^k f(x) = q(x)g(x) + r(x)$$

and either $r(x) = 0$ or $\deg r < \deg g$. Moreover,

$$H(r) \leq H(f)[2H(g)]^{|\deg f - \deg g|+1}.$$

Proof. Let $n = \deg f$ and $m = \deg g$. If $f(x) = 0$ or $n < m$, we set $k = 0$, $q(x) = 0$, and $r(x) = f(x)$. We now assume that $n \geq m$, and let $f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$ and $g(x) = b_m x^m + b_{m-1}x^{m-1} + \cdots + b_1 x + b_0$. We prove the desired result by induction.

We first prove the basis step. If $n = 0$, then $m = 0$ so we set $k = 1$, $q(x) = a_0$ and $r(x) = 0$.

Assume now that the result holds when we divide all polynomials in $\mathbb{Z}[x]$ of degree less than n by $g(x)$. We let $p(x) = f(x) - a_n {b_m}^{-1} x^{n-m} g(x)$, then $p(x) = 0$ or $\deg p < n$. Thus either $b_m p(x) = 0$ or $b_m p(x) = b_m f(x) - a_n x^{n-m} g(x)$ is a polynomial in $\mathbb{Z}[x]$ of degree less than n. If either $b_m p(x) = 0$ or $\deg p < \deg g$, we set $k = 1, q(x) = a_n x^{n-m}$, and $r(x) = b_m p(x)$ so that

$$b_m f(x) = q(x)g(x) + r(x).$$

[2] *A version of this chapter has been submitted for publication*
Ayad, M., Kihel, O., Larone, J., 2014

Then $q(x), r(x) \in \mathbb{Z}[x]$ where $\deg r < \deg g$ and

$$H(r) \leq H(f)[2H(g)] \leq H(f)[2H(g)]^{|\deg f - \deg g|+1}$$

as required. If $\deg g \leq \deg p < n$, then by the induction hypothesis there exists a non-negative integer $k \leq |\deg p - \deg g| + 1 = \deg p - \deg g + 1$ and polynomials $\overline{q}(x)$ and $\overline{r}(x)$ in $\mathbb{Z}[x]$ such that

$${b_m}^k b_m p(x) = {b_m}^{k+1} p(x) = \overline{q}(x)g(x) + \overline{r}(x),$$

where $H(\overline{r}) \leq |b_m| H(p)[2H(g)]^{\deg p - \deg g + 1}$ and either $\overline{r}(x) = 0$ or $\deg \overline{r} < \deg g$. We obtain

$$\begin{aligned} {b_m}^{k+1} f(x) &= a_n {b_m}^k x^{n-m} g(x) + {b_m}^{k+1} p(x) \\ &= a_n {b_m}^k x^{n-m} g(x) + \overline{q}(x)g(x) + \overline{r}(x) \\ &= \big(a_n {b_m}^k x^{n-m} + \overline{q}(x)\big)g(x) + \overline{r}(x). \end{aligned}$$

We set $q(x) = a_n {b_m}^k x^{n-m} + \overline{q}(x)$ and $r(x) = \overline{r}(x)$, so that $q(x), r(x) \in \mathbb{Z}[x]$ and either $r(x) = 0$ or $\deg r < \deg g$. The non-negative integer $k+1$ then satisfies $k+1 \leq (\deg p + 1) - \deg g + 1 \leq \deg f - \deg g + 1 = |\deg f - \deg g| + 1$ and $H(r)$ satisfies

$$\begin{aligned} H(r) &\leq |b_m| H(p)[2H(g)]^{\deg p - \deg g + 1} \\ &\leq H(f)[2H(g)][2H(g)]^{\deg p - \deg g + 1} \\ &= H(f)[2H(g)]^{(\deg p + 1) - \deg g + 1} \\ &\leq H(f)[2H(g)]^{\deg f - \deg g + 1} \\ &\leq H(f)[2H(g)]^{|\deg f - \deg g| + 1} \end{aligned}$$

as required. □

We will require the following result, due to Cauchy ([6]), which provides a bound on the roots of a polynomial.

Lemma 5.1. *Let $f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$ be a complex polynomial and let β be a root of $f(x)$. Then*

$$|\beta| \leq 1 + \frac{H(f)}{|a_n|}.$$

Proof. The result holds immediately if $|\beta| \leq 1$, thus we suppose that $|\beta| > 1$. Since β is a root of $f(x)$, we have

$$a_n \beta^n + a_{n-1} \beta^{n-1} + \cdots + a_1 \beta + a_0 = 0$$

so that

$$\begin{aligned} |a_n||\beta|^n &= |-a_{n-1}\beta^{n-1} - \cdots - a_1\beta - a_0| \\ &\leq |a_{n-1}\beta^{n-1}| + \cdots + |a_1\beta| + |a_0| \\ &\leq H(f)\big(|\beta|^{n-1} + \cdots + |\beta| + 1\big) \\ &= H(f)\frac{|\beta|^n - 1}{|\beta| - 1}. \end{aligned}$$

Since $|\beta| > 1$, we have

$$|\beta| - 1 \leq \frac{H(f)}{|a_n|}\frac{|\beta|^n - 1}{|\beta|^n} = \frac{H(f)}{|a_n|}\left(1 - \frac{1}{|\beta|^n}\right) \leq \frac{H(f)}{|a_n|}$$

and the result follows. □

We now proceed to state and prove the main result of this chapter.

Theorem 5.1. *Let $f(x), g(x) \in \mathbb{Z}[x]$ where $0 \leq \deg g(x) \leq \deg f(x)$ and cont(g) divides cont(f). If there exists an integer N such that*

$$|N| > 1 + H(g) + H(f)[2H(g)]^{\deg f - \deg g + 1},$$

$g(N)$ divides $f(N)$, and $g(N)f(N) \neq 0$, then $g(x)$ divides $f(x)$ in $\mathbb{Z}[x]$.

Proof. We first consider the case $N > 0$. Let $g(x) = b_m x^m + \cdots + b_1 x + b_0$, and assume without loss of generality that $b_m > 0$. By Proposition 5.1, there exists a non-negative integer $k \leq |\deg f - \deg g| + 1 = \deg f - \deg g + 1$ and polynomials $q(x)$ and $r(x)$ in $\mathbb{Z}[x]$ such that

$$b_m{}^k f(x) = q(x)g(x) + r(x)$$

where $H(r) \leq H(f)[2H(g)]^{\deg f - \deg g + 1}$ and either $r(x) = 0$ or $\deg r < \deg g$. We obtain

$$b_m{}^k f(N) = q(N)g(N) + r(N),$$

thus $g(N)$ must divide $r(N)$ since $g(N)$ divides $f(N)$. It follows that $|g(N)|$ divides $|r(N)|$.

We show that $r(x) = 0$. Assume that $r(x) \neq 0$, then $\deg r < \deg g$. Consider the functions $g(t)$ and $r(t)$ of the complex variable t and let $\varphi_1(t) = g(t) - r(t)$ and $\varphi_2(t) = g(t) + r(t)$. Let $\alpha_1, ..., \alpha_m$ and $\beta_1, ..., \beta_m$ be all of the zeros of $\varphi_1(t)$ and $\varphi_2(t)$ respectively, and let $M_1 = \max\{|\alpha_1|, ..., |\alpha_m|\}$ and $M_2 = \max\{|\beta_1|, ..., |\beta_m|\}$. Then either $\varphi_1(t) > 0$ or $\varphi_1(t) < 0$ for all real values of $t > M_1$, and either $\varphi_2(t) > 0$ or $\varphi_2(t) < 0$ for all real values of $t > M_2$.

Since $\deg g > \deg r$, there exists a number $t_1 > M_1$ such that $g(t) > r(t)$ for all $t \geq t_1$. This implies that $\varphi_1(t) > 0$ for $t \geq t_1$, therefore we must have $\varphi_1(t) > 0$ for all $t > M_1$. Similarly, there exists a number $t_2 > M_2$ such that $g(t) > -r(t)$ for all $t \geq t_2$, so that $\varphi_2(t) > 0$ for $t \geq t_2$ and $\varphi_2(t) > 0$ for all $t > M_2$. Let $M = \max\{M_1, M_2\}$, then we have $\varphi_1(t) > 0$ and $\varphi_2(t) > 0$ for all $t > M$. We obtain $-g(t) < r(t) < g(t)$ for all $t > M$, yielding $g(t) > |r(t)|$ for $t > M$. It follows that for all $x \in \mathbb{Z}$ such that $x > M$, we have $|g(x)| \geq g(x) > |r(x)|$.

By Lemma 5.1, for any zero α_i of $\varphi_1(t)$ we have

$$|\alpha_i| \leq 1 + \frac{H(\varphi_1)}{|b_m|} \leq 1 + H(\varphi_1) = 1 + H(g - r) \leq 1 + H(g) + H(r).$$

It follows that for all zeros α_i, $i = 1, ..., m$, of $\varphi_1(t)$ we have

$$|\alpha_i| \leq 1 + H(g) + H(r) \leq 1 + H(g) + H(f)[2H(g)]^{\deg f - \deg g + 1}.$$

Similarly, for all zeros β_i, $i = 1, ..., m$, of $\varphi_2(t)$ we have

$$|\beta_i| \leq 1 + H(\varphi_2) \leq 1 + H(g) + H(f)[2H(g)]^{\deg f - \deg g + 1}.$$

We then have

$$N > 1 + H(g) + H(f)[2H(g)]^{\deg f - \deg g + 1} \geq M,$$

thus $|g(N)| > |r(N)|$. Since $|g(N)|$ divides $|r(N)|$, we must have $r(x) = 0$ and ${b_m}^k f(x) = q(x)g(x)$. Since $cont(g)$ divides $cont(f)$, we conclude that $g(x)$ divides $f(x)$ as required.

If $N < 0$, we define $G(x) = g(-x)$, $F(x) = f(-x)$, and $N_0 = -N$. We then have $F(x), G(x) \in \mathbb{Z}[x]$ where $0 \leq \deg G(x) \leq \deg F(x)$, $cont(G)$ divides $cont(F)$, and the integer N_0 satisfies

$$N_0 > 1 + H(G) + H(F)[2H(G)]^{\deg F - \deg G + 1}.$$

By the first part of the proof, we conclude that $G(x)$ divides $F(x)$ in $\mathbb{Z}[x]$. Therefore $g(x)$ divides $f(x)$ in $\mathbb{Z}[x]$. □

Remark 5.1. *The assumption in Theorem 5.1 on the contents of f and g cannot be omitted, as it is shown by the example $g(x) = px$ and $f(x) = x(x^p - x)$, where p is a prime number.*

Corollary 5.1. *Let $f(x), g(x) \in \mathbb{Z}[x]$ where $0 \leq \deg g(x) \leq \deg f(x)$. If there exists an integer N such that $|N| > 1 + H(g) + H(f)[2H(g)]^{\deg f - \deg g + 1}$ and $g(N)$ divides $f(N)$, then $g(x)$ divides $f(x)$ in $\mathbb{Q}[x]$.*

Proof. Let $a \in \mathbb{N} \setminus \{0\}$ be the content of $g(x)$, then $g(x) = ag_1(x)$, where $g_1(x) \in \mathbb{Z}[x]$ and $cont(g_1) = 1$. From the assumptions, we conclude that

$$|N| > 1 + H(g_1) + H(f)[2H(g_1)]^{\deg f - \deg g_1 + 1}$$

and $g_1(N)$ divides $f(N)$. Therefore $g_1(x)$ divides $f(x)$ in $\mathbb{Z}[x]$ by Theorem 5.1. It follows that $g(x)$ divides $f(x)$ in $\mathbb{Q}[x]$. □

5.1. **An Application.**

Let $f(x)$ be a polynomial with integral coefficients and let k be a positive integer. We may determine the divisors of $f(x)$ of degree at most k in $\mathbb{Z}[x]$, if any, in the following way.

Select distinct integers $a_0, a_1, ..., a_k$ such that $f(a_i) \neq 0$ for $i \in \{0, 1, ..., k\}$. For any $i = 0, 1, ..., k$, select a divisor b_i of $f(a_i)$ and compute, by using Lagrange's interpolation formula, the unique polynomial $g(x) \in \mathbb{Q}[x]$ of degree at most k

such that $g(a_i) = b_i$ for $i = 0, 1, ..., k$. By varying the values of the $k+1$-tuple $(a_0, a_1, ..., a_k)$, we obtain a set of polynomials of degree at most k with rational coefficients containing those polynomials $g(x) \in \mathbb{Z}[x]$ of degree at most k dividing $f(x)$ in $\mathbb{Z}[x]$.

Using Theorem 5.1, we will prove the following.

Corollary 5.2. *Let $f(x)$ be a polynomial with integral coefficients of degree n. Let $k = \lfloor n/2 \rfloor$ and $N > 1 + \binom{n}{k}\sqrt{n+1}H(f) + H(f)\left[2\binom{n}{k}\sqrt{n+1}H(f)\right]^n$. Let $a_0, a_1, ..., a_{k-1}$ be distinct integers such that $a_i \neq N$ for any $i \in \{0, 1, ..., k-1\}$ and set $a_k = N$. Suppose that $f(a_i) \neq 0$ for $i \in \{0, ..., k\}$.*

Let $E = \{\, g(x) \in \mathbb{Z}[x] : \ 1 \leq \deg g \leq k, \ g(a_i) | f(a_i) \text{ for } i = 0, ..., k \,\}$. Then either f is irreducible over $\mathbb{Z}$, or E is nonempty and any $g(x) \in E$ of minimal height is a divisor of $f(x)$ in $\mathbb{Z}[x]$.

For the proof of this result, we will require the following, whose proof follows from Proposition 2.1.12 of Mignotte and Stefanescu ([8]) and from the definition of the height of a polynomial.

Lemma 5.2. *Let*

$$P(x) = \sum_{i=0}^{d} p_i x^i \qquad \text{and} \qquad Q(x) = \sum_{i=0}^{m} q_i x^i$$

be polynomials with integral coefficients of degree d and m respectively such that $Q(x)$ divides $P(x)$ in $\mathbb{Z}[x]$. Then

$$|q_i| \leq \binom{m}{i}\sqrt{d+1}H(P).$$

Proof of Corollary 5.2. Suppose that $f(x)$ is reducible over $\mathbb{Z}$, then $f(x)$ has an irreducible factor of degree less than or equal to $\lfloor n/2 \rfloor = k$ and greater than or equal to 1. Let $g(x) \in \mathbb{Z}[x]$ be one of its irreducible factors of degree greater than or equal to 1 and less than or equal to k. Then $g(x) \in E$, hence E is nonempty. Let $g_0(x) \in E$ be such that $H(g_0)$ is minimal. Then we have

$$\begin{aligned} N &> 1 + \binom{n}{k}\sqrt{n+1}H(f) + H(f)\left[2\binom{n}{k}\sqrt{n+1}H(f)\right]^n \\ &\geq 1 + H(g) + H(f)[2H(g)]^n \\ &\geq 1 + H(g_0) + H(f)[2H(g_0)]^{\deg f - \deg g_0 + 1}. \end{aligned}$$

Since $H(g_0)$ is minimal, then $cont(g_0) = 1$. By Theorem 5.1, we conclude that $g_0(x)$ divides $f(x)$ in $\mathbb{Z}[x]$. $\square$

6. **Conclusion**

The units under function composition are useful in providing examples of prime rational functions. One may accomplish this by manipulating the zeros and poles of the rational function obtained through the composition of a given rational function and a suitably chosen unit. These units can additionally be useful in providing some examples of prime polynomials.

The resultant is also an effective way to provide examples of prime rational functions. One may consider the resultant of a chosen composite rational function and its derivative to provide conditions on the critical values of the composite function, which present opportunities to consider some results whose proofs proceed by way of contradiction.

One of the most important concepts in the motivation and the proofs of the results found in this work is the degree of a polynomial or of a rational function. As such, it would be of interest to generalize the concept and find other results similar to Proposition 4.1 and Lemma 4.4, both of which can be used to provide a relationship between the degrees of the numerator and denominator of a composite rational function with those of its composition factors. In particular, another mapping $\psi : \mathbb{C}(x) \to \mathbb{Z}$ for which $\psi(g \circ h) = \psi(g) \cdot \psi(h)$ is satisfied for rational functions g and h could potentially provide many more examples of prime functions.

We have also considered polynomials and their reducibility. We have improved upon an existing result regarding the divisibility of integral polynomials by considering a form of the division algorithm for polynomials in $\mathbb{Z}[x]$. Given a polynomial with integral coefficients, we have also provided a method to determine if that polynomial is irreducible over $\mathbb{Z}$, and if not, we have provided one of its divisors in $\mathbb{Z}[x]$.

References

[1] Ayad, M., 2006, Critical points, critical values of a prime polynomial, *Complex Variables and Elliptic Equations: An International Journal*, 51:2, 143-160

[2] Beardon, A.F., 2001, Composition factors of polynomials, *Complex Variables and Elliptic Equations: An International Journal*, 43, 225-239

[3] Gallian, J. A., 2006, *Contemporary Abstract Algebra*, Brooks/Cole

[4] Friedberg, S. H., Insel, A.J., and Spence, L. E., 2003, *Linear Algebra*, Prentice Hall

[5] Lang, S., 1994, *Algebraic Number Theory*, Springer-Verlag

[6] Cauchy, A. L., 1829, *Exercises de Mathématiques* IV Année, De Bure frères, Paris

[7] Nieto, J. H., 2003, Sobre la Divisibilidad de Polinomios con Coeficientes Enteros, *Divulgaciones Matemáticas*, 11:2, 149-152

[8] Mignotte, M., Stefanescu D., 1999, *Polynomials An Algorithmic Approach*, Springer-Verlag