Rose-
Hulman
Undergraduate
Mathematics
Journal

Rational Function Decomposition of Polynomials

Steven Carter[a]

Volume 13, No. 2, Fall 2012

Sponsored by

Rose-Hulman Institute of Technology
Department of Mathematics
Terre Haute, IN 47803
Email: mathjournal@rose-hulman.edu
http://www.rose-hulman.edu/mathjournal

[a]Birmingham-Southern College

Rose-Hulman Undergraduate Mathematics Journal
Volume 13, No. 2, Fall 2012

Rational Function Decomposition of Polynomials

Steven Carter

Abstract. We determine conditions under which an arbitrary polynomial can be expressed as the composition of two rational functions, generalizing the work of J. Rickards on the decomposition into two polynomials. We show that a polynomial can be expressed non-trivially as a composition of two rational functions if and only if it can be so decomposed into two polynomials.

Acknowledgements: Special acknowledgement to Birmingham-Southern College Mathematics Department for encouragement and support.

1 Introduction

While it is simple to show that the composition of two polynomials is always a polynomial, when usefully defined, the converse is not always true. J. Rickards discovered a set of necessary and sufficient conditions on the roots of a polynomial under which the polynomial can be written as the composition of other polynomials [1]. Let us refer to the functions which compose to yield a desired polynomial *constituent* functions. We extended the work of Rickards by allowing the constituent functions to be rational functions.

In [1], Rickards noted that if linear constituent functions were allowed then every polynomial is expressible as the composition of two polynomials. Thus, he defined constituent polynomials to be trivial if their degree was less than 2. The key property of a constituent function that Rickard's needed to exclude was invertibility within the set of polynomials. Of course the only functions within the set of polynomials that have inverses which are also polynomials are linear. Thus we will define a polynomial to have a proper rational decomposition if and only if it can be written as the composition of non-trivial rational functions, where a rational function is considered trivial if it possess an inverse which is also a rational function.

We proceed by formalizing a notion of equivalence amongst rational functions and making precise the operation of composition on them. We then identify the inherent algebraic structure given by the set of rational functions under composition and identify all the invertible elements. We define a proper rational function decomposition as one without the use of invertible elements. A necessary and sufficient factorization condition is determined for the existence of a proper rational function decomposition of an arbitrary polynomial and we show that this condition is equivalent to the one given by Rickards in [1]. Finally, we introduce structure for future work to determine if rational function decomposition of polynomials is unique.

2 Preliminaries

A *polynomial in one variable* is an expression of finite length constructed from real constants and non-negative powers of the variable via addition and multiplication. A *rational function* is the quotient of two polynomials, where the denominator is not the zero polynomial. Much like for rational numbers, we wish to identify certain rational functions as the same. We define a relation on the set of all rational functions by saying two rational functions are related if their values differ on at most a finite number of points. Thus in essence we say two rational functions are related if one can be algebraically manipulated into the other. This relation is clearly reflexive, symmetric and transitive and so is an equivalence relation on the set of rational functions.

We wish to define the operation of composition on the set of non-constant equivalence classes induced on the set of rational functions by the equivalence relation. Here an equivalence class

is non-constant if all members of the class contain some non-zero power of the variable, that is, no member of the class is a constant function. We do so by picking a representative from each equivalence class and compose them. Let $f(x)$ and $g(x)$ be the representatives given as:

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

Here we assume $p(x)$ and $q(x) \neq 0$ are polynomials. We then see that:

$$\begin{aligned} (f \circ g)(x) &= \frac{\sum_{i=0}^{n} a_i \left(\frac{p(x)}{q(x)}\right)^i}{\sum_{j=0}^{m} A_j \left(\frac{p(x)}{q(x)}\right)^j} \\ &= \left(\frac{q(x)^{max(m,n)-n}}{q(x)^{max(m,n)-m}}\right) \frac{\sum_{i=0}^{n} a_i p(x)^i q(x)^{n-i}}{\sum_{j=0}^{m} A_j p(x)^j q(x)^{m-j}}. \end{aligned} \tag{1}$$

Here we know the denominator is not identically zero since $g(x)$ is assumed non-constant and thus can not take the value of a root of the denominator of $f(x)$ except possibly at a finite number of points. The possibility of generating an expression with the denominator being identically zero is why we must exclude equivalence classes containing constant functions from the definition of composition. Note we have systematically produced a rational function from the two representatives. This motivates us to define composition of two non-constant equivalence classes by:

$$[f(x)] \circ [g(x)] = [(f \circ g)(x)].$$

If is not difficult to show that this definition of composition of two equivalence classes is well-defined. Note, (1) also gives an algorithm for simplifying the standard prescription of the composition of two rational functions into a form that is clearly a rational function.

We define a rational function to be in *lowest terms* if there is no non-constant function of x that can be factored from both the numerator and denominator (and thus be canceled). Next we show that if both f and g are in lowest terms, then so is the composition written as in (1). This means that there can be no surprise cancelations between the numerator and denominator of the composition.

Lemma 2.1 *If both $f(x)$ and $g(x)$ are in lowest terms, then $(f \circ g)(x)$ computed as in (1) is in lowest terms.*

Proof: Let $f(x)$ and $g(x)$ be rational functions written in lowest terms. Using the Fundamental Theorem of Algebra, we can write:

$$f(x) = \frac{a\Pi_{i=1}^{n}(x - r_i)}{\Pi_{j=1}^{m}(x - s_j)}$$

where the r_i and s_j may be complex, and n and m are integers. Now if $g(x) = \frac{p(x)}{q(x)}$ then we have by (1) the equation:

$$(f \circ g)(x) = \left(\frac{q(x)^{max(m,n)-n}}{q(x)^{max(m,n)-m}}\right) \frac{a\Pi_{i=1}^{n} a_i(p(x) - r_i q(x))}{\Pi_{j=1}^{m}(p(x) - s_j q(x))}.$$

If the composition is not in lowest terms, then there is a t, possibly complex, so that $x - t$ is a factor of both the numerator and the denominator. This leads to three cases:

1. $(x - t)|q(x)$ and for some i we have $(x - t)|(p(x) - r_i q(x))$,
2. $(x - t)|q(x)$ and for some j we have $(x - t)|(p(x) - s_j q(x))$, or
3. $(x - t)|(p(x) - r_i q(x))$ for some i and $(x - t)|(p(x) - s_j q(x))$ for some j.

Note cases (1) and (2) are identical by exchanging r_i for s_j, so we only consider cases (1) and (3). In case (1), we see that $q(x) = (x - t)q'(x)$ for some polynomial $q'(x)$. Similarly, there is a polynomial $u(x)$ so that

$$(x - t)u(x) = p(x) - r_i q(x) = p(x) - r_i(x - t)q'(x).$$

Rearranging we see that $x - t$ divides $p(x)$. But this contradicts $g(x)$ being in lowest terms.

Now assume case (3) holds. Then there are polynomials $u(x)$ and $v(x)$ so that

$$(x - t)u(x) = p(x) - r_i q(x)$$

and

$$(x - t)v(x) = p(x) - s_j q(x).$$

Solving for $p(x)$ above and equating leads to the equation:

$$(s_j - r_i)q(x) = (x - t)(u(x) - v(x)).$$

From this we see that either $s_j = r_i$ in which case $f(x)$ is not in lowest terms, or $x - t$ divides $q(x)$ which leads us back to case (1). □

Now let R be the set of all non-constant equivalence classes induced on the set of rational functions by the equivalence relation and let $\mathcal{R}$ be the potential algebraic structure built

from R under composition as defined in (1). We have already noted that the operation of composition on equivalence classes is well-defined. Consider the composition of arbitrary elements of R. We are free to choose any representative function for a given equivalence class, so let these non-constant equivalence classes be represented by rational functions in lowest terms. Then they can be written as f and g in (1) and that calculation shows that the composition is a rational function. Since the output is in lowest terms by Lemma 2.1, it is not equivalent to a constant and composition is a binary operation on $\mathcal{R}$. It is trivial to show that $\circ$ is associative in $\mathcal{R}$ and that the identity equivalence class contains the identify function $e(x) = x \in R$. Therefore, it follows that $\mathcal{R}$ is a monoid.

At this point we will dispense with the language of equivalence classes. It will be understood that when speaking of a rational function, we mean the equivalence class represented by the rational function. In particular, when we write $f(x) \in \mathcal{R}$ we mean the equivalence class represented by $f(x)$ is a member of $\mathcal{R}$. Furthermore, we will freely choose the most appropriate representative of the equivalence class for the particular task at hand.

The presence of a monoid allows the possibility of inverses. We say an element of $\mathcal{R}$ is a *unit* if it possesses an inverse in the monoid. If units are are considered legitimate constituent functions then every rational function would be decomposable into the composition of two rational functions trivially as follows:

$$f = (f \circ u) \circ u^{-1} = u^{-1} \circ (u \circ f).$$

To avoid these trivial decompositions, we need to identify all of the units in $\mathcal{R}$. We begin by identifying the ratios of linear functions that are constant when written in lowest terms.

Lemma 2.2 *Let*

$$f(x) = \frac{ax + b}{Ax + B}$$

where a, b, A, and B are real numbers with A and B not both zero. Then $f(x)$ is equivalent to a constant function if and only if $aB - Ab = 0$.

Proof: Suppose $aB - Ab = 0$. Then either $aB = Ab = 0$ or $\frac{B}{A} = \frac{b}{a}$. The former implies $a = b = 0$, $a = A = 0$, or $B = b = 0$, all of which clearly imply $f(x)$ is constant. The latter also implies $f(x)$ is a constant as follows:

$$f(x) = \frac{ax + b}{Ax + B} = \frac{a(x + \frac{b}{a})}{A(x + \frac{B}{A})} = \frac{a}{A}.$$

Next suppose that $f(x)$ is equivalent to a constant function, say $f(x) = c$ except possibly at one point. If $B = 0$, then a direct calculation shows that $b = 0$ implying $aB - Ab = 0$ as desired. If $B \neq 0$, then $f(0) = \frac{b}{B} = c$. Also

$$\lim_{x \to \infty} f(x) = \frac{a}{A} = c.$$

We conclude that $\frac{a}{A} = \frac{b}{B}$ and so $aB - Ab = 0$. $\square$

We now identify some units in the monoid by recognizing that non-constant functions of the form

$$f(x) = \frac{ax + b}{Ax + B}$$

are Möbius transformations with real coefficients (which satisfy the proper conditions by 2.2). These functions form a group isomorphic to $SL_2(\mathbb{R})$, a subgroup of the Möbius group [2]. Therefore, these rational functions are units in $\mathcal{R}$.

We are now ready to classify all of the units in $\mathcal{R}$.

Theorem 2.3 *The only units in $\mathcal{R}$ are of the form:*

$$f(x) = \frac{ax + b}{Ax + B} \tag{2}$$

with $aB - Ab \neq 0$.

Proof: Rational functions of this form are invertible in $\mathcal{R}$ by the previous argument. All that remains to be shown is that there are no other units in $\mathcal{R}$. Let $f(x) \in \mathcal{R}$ be in lowest terms and assume there is a rational function $g(x) \in \mathcal{R}$, also in lowest terms, which when composed with $f(x)$ gives a function equivalent to the identity function. Using the Fundamental Theorem of Algebra, we can write

$$f(x) = \frac{a\Pi_{i=1}^{n}(x - r_i)}{\Pi_{j=1}^{m}(x - s_j)}$$

where r_i and s_j may be complex. Suppose $g(x) = \frac{p(x)}{q(x)}$. Then by equation (1) we have

$$(f \circ g)(x) = \left(\frac{q(x)^{max(m,n)-n}}{q(x)^{max(m,n)-m}}\right) \frac{a\Pi_{i=1}^{n}(p(x) - r_i q(x))}{\Pi_{j=1}^{m}(p(x) - s_j q(x))}. \tag{3}$$

By Lemma 2.1, this expression is in lowest terms. Furthermore, by assumption this expression is equivalent to the identity function. We conclude that (3) is equal to x. Since (3) is equal to x, the denominator must be constant which implies $q(x)^{max(m,n)-m}$ is constant. We conclude that either $q(x)$ is constant or $m \geq n$.

For the first case, assume $q(x)$ is constant. Then since the denominator of (3) is constant we can conclude that either $p(x)$ is also constant or $m = 0$ so that the product in the denominator of (3) is empty. If $p(x)$ is constant then $g(x)$ is constant which is a contradiction. Thus we must have $m = 0$ and so both $f(x)$ and $g(x)$ are polynomials. In this case equation (3) simplifies to:

$$(f \circ g)(x) = aq^{-n}\Pi_{i=1}^{n}(p(x) - r_i q) = x.$$

We conclude that $p(x)$ is a linear function and $n = 1$. Thus $f(x)$ is in the form (2).

Next suppose $q(x)$ is not constant so that necessarily $m \geq n$. Now the product in the denominator of (3) must be nonempty. If not, then $m = n = 0$ and $f(x)$ is constant. We conclude that there must be constants, c_l so that $p(x) - s_l q(x) = c_l$ for all $1 \leq l \leq m$. Solving for $p(x)$ and equating with each other shows that all of the s_j must be equal, say to s. Then we have $p(x) = c + sq(x)$ and the denominator of (3) equates to c^m. Next, moving to the numerator of (3) and using the equation for $p(x)$ above, we see that

$$q(x)^{m-n} a \Pi_{i=1}^{n} (c + (s - r_i) q(x)) = c^m x.$$

We must have $m = n$ and $n = 1$ (as well as $q(x)$ linear), which shows that $f(x)$ is of the form (2). □

Theorem 2.3 completely describes the units within the monoid. We are interested in non-trivial decompositions of the polynomials within $\mathcal{R}$. Using Theorem 2.3, we can now identify when a given composition is trivial. In the next section, we characterize which polynomials can be decomposed nontrivially into the composition of two rational functions.

3 Main Results

In [1], the author gives conditions on the roots of a polynomial, which if satisfied, allow that polynomial to be nontrivially decomposed into the composition of two polynomials. Moreover, Rickards gives an algorithm for finding the two polynomial constituent functions. Following in this vein, we next give a condition on the roots of a polynomial under which the polynomial can be written as the composition of two rational function in a nontrivial way. Our conditions on the roots of the polynomial take the form of the existence of a particular kind of factorization.

Theorem 3.1 *A polynomial $h(x)$ has a non-trivial, non-polynomial rational function decomposition if and only if there exists a polynomial $q(x)$ of degree at least two, such that $h(x)$ can be written in the form:*

$$h(x) = a' q(x)^{m-n} \Pi_{i=1}^{n} (1 - r_i' q(x)) \tag{4}$$

where a' is real, r_i' are possibly complex, and for the integers m and n we require $m \geq n$ and $m \geq 2$.

Proof: Let $h(x)$ be a polynomial and suppose $h(x) = (f \circ g)(x)$ for non-unit, non-polynomial rational functions $f(x)$ and $g(x)$ in $\mathcal{R}$ written in lowest terms. Write

$$f(x) = \frac{a \Pi_{i=1}^{n} (x - r_i)}{\Pi_{j=1}^{m} (x - s_i)}$$

and let $g(x) = p(x)/q(x)$ as before. Then the composition, $h(x)$, is given by

$$\left(\frac{q(x)^{max(m,n)-n}}{q(x)^{max(m,n)-m}}\right)\frac{a\Pi_{i=1}^{n}(p(x)-r_iq(x))}{\Pi_{j=1}^{m}(p(x)-s_jq(x))} \tag{5}$$

and is in lowest terms by Lemma 2.1. We conclude $\Pi_{j=1}^{m}(p(x)-s_iq(x))$ is constant since $h(x)$ is a polynomial. Thus either $m = 0$, or we have constants c_l such that $p - s_lq = c_l$ for all $1 \leq l \leq m$.

If $m = 0$, then by (5), we see

$$h(x) = \frac{a\Pi_{i=1}^{n}(p-r_iq)}{q^n},$$

which implies q is a constant, or $n = 0$. However, $m = n = 0$ is a contradiction of the assumption that $f(x)$ is not a constant. If $q(x)$ is constant and $m = 0$, then $f(x)$ and $g(x)$ are both polynomials, another contradiction.

Therefore, assume that $m > 0$ and $p(x) - s_lq(x) = c_l$ for all $1 \leq l \leq m$. Solving for $p(x)$ and equating we see that $s_l = s_k$ for all $1 \leq l, k, \leq m$ (or $q(x)$ is constant which implies $g(x)$ is constant, another contradiction). We call this common value s and let c be the common value for the c_l. Thus we can write $p(x) = sq(x) + c$ and therefore $g(x)$ must have the form:

$$g(x) = \frac{sq(x)+c}{q(x)}.$$

Note that since $g(x)$ is in lowest terms (and not constant), we can conclude $c \neq 0$. Since $g(x)$ is not a unit, by Theorem 2.3 we know that the degree of $q(x)$ is at least two. Returning to (5), we note that $q^{max(n,m)-m}$ must be constant. Since $q(x)$ is not constant, we must have $m \geq n$. This implies, since $f(x)$ is also not a unit, that $m \geq 2$.

Returning to $f(x)$ and using what we have shown, we conclude $f(x)$ must have the form:

$$f(x) = \frac{a\Pi_{i=1}^{n}(x-r_i)}{(x-s)^m}.$$

Next let $u = cx + s$. By Lemma 2.2 we know u is a unit. Moreover a simple calculation will show $u^{-1} = \frac{1}{c}x - \frac{s}{c}$. Let $f' = f \circ u$ and $g' = u^{-1} \circ g$, then $h(x) = (f' \circ g')(x)$ as well. Using the form for $f(x)$ and $g(x)$ above we find that:

$$f'(x) = \frac{ac^{n-m}\Pi_{i=1}^{n}(x-\frac{r_i-s}{c})}{x^m}.$$

This motivates us to define $a' = ac^{n-m}$ and $r_i' = \frac{r_i-s}{c}$. Also observe

$$g'(x) = \frac{1}{q(x)}.$$

Returning to (3) and using f' and g' as the constituent functions we have,

$$h(x) = a'q(x)^{m-n}\Pi_{i=1}^{n}(1 - r_i'q(x))$$

as desired.

Now suppose that $h(x)$ can be written in the form (4) where $m \geq n$, $m \geq 2$ and the degree of $q(x)$ is at least 2. Define $g'(x)$ and $f'(x)$ as above. The conditions imply neither $g'(x)$ nor $f'(x)$ are units by Theorem 2.3. Furthermore both $f'(x)$ and $g'(x)$ are non-polynomial rational constituent functions of $h(x)$. $\square$

A necessary and sufficient condition has been found under which polynomials possess a non-trivial, non-polynomial rational function decomposition. An arbitrary polynomial possesses a proper rational function decomposition if and only if it satisfies the conditions given in [1] for a proper polynomial decomposition or if it can be factored in the special form given in Theorem 3.1. Determining if a polynomial satisfies the factorization in the preceding theorem is difficult, but next we show that this condition is equivalent to the condition for the existence of a proper polynomial decomposition given in [1].

Theorem 3.2 *A polynomial $h(x)$ possesses a non-trivial polynomial decomposition if and only if $h(x)$ possesses a non-trivial rational function decomposition.*

Proof: Since all polynomials are also rational, half of the above statement holds vacuously. However, given a polynomial decomposition we can produce a non-polynomial decomposition as follows. Suppose $h(x) = (f \circ g)(x)$ where both $f(x)$ and $g(x)$ are non-linear polynomials. Write $f(x)$ in standard form as

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

Let $u(x) = \frac{1}{x}$ and note u is its own inverse. We define $f'(x) = (f \circ u)(x)$ and $g'(x) = (u \circ g)(x)$. Then $h(x) = (f' \circ g')(x)$ and furthermore

$$f'(x) = \frac{\sum_{i=0}^{n} a_i x^{n-i}}{x^n},$$

while

$$g'(x) = \frac{1}{g(x)}.$$

If both $f(x)$ and $g(x)$ are non-linear, then $f'(x)$ and $g'(x)$ are both not units in $\mathcal{R}$. We have constructed a non-polynomial decomposition of $h(x)$.

Next suppose $h(x)$ allows a non-trivial decomposition into rational functions. By Theorem 3.1 we can write $h(x)$ in the form (4) for some polynomial $q(x)$ of degree at least 2

where we have the appropriate conditions on m and n as outlined in the theorem. This motivates us to define

$$f(x) = a'x^{m-n}\Pi_{i=1}^{n}(1 - r_i'x).$$

Since $m \geq n$ and $m \geq 2$ we see that $f(x)$ is not a unit. We then trivially see that $h(x) = (f \circ q)(x)$ and we have written $h(x)$ nontrivially as the composition of two polynomials. $\square$

In [1], Rickards gives an algorithm for determining if a given polynomial can be decomposed into two nontrivial constituent polynomials and a means of finding such a decomposition. We note that Theorem 3.2, or rather the proof given, provides a simple way of taking a decomposition into polynomials and building a decomposition into rational functions.

4 Future Work

We defined a polynomial to possess a proper rational decomposition only if it could be written as the composition of non-unit constituents of $\mathcal{R}$. Determining if this decomposition is unique up to unit factors requires the concept of irreducible constituents. Let a rational function in $\mathcal{R}$ be *irreducible* if it is not a unit and it does not possess a proper rational decomposition in $\mathcal{R}$. For example $f(x) = x^p$ with p prime is irreducible. This function clearly can not be decomposed into the composition of two polynomials and therefore can not be decomposed into the composition of two rational functions by Theorem 3.2. At present, without complete knowledge of the irreducible elements of $\mathcal{R}$ it is impossible to tell whether decomposition in $\mathcal{R}$ is unique. Given distinct non-unit rational functions r_1, r_2, and r_3 in $\mathcal{R}$, $h = r_1 \circ (r_2 \circ r_3)$ and $h = (r_1 \circ r_2) \circ r_3$ would both give what appears to be legitimate, nonequivalent decompositions. To discuss uniqueness, we introduce a bit more structure.

Let f and g in $\mathcal{R}$ be *associate* if $f = u \circ g \circ v$ for units u and v in $\mathcal{R}$. It can be shown that the notion of associate is an equivalence relation on the set of rational functions. A *fundamental rational decomposition* of $h(x)$ is an expression for h of the form

$$h(x) = f_1 \circ f_2 \circ ... \circ f_n$$

where f_1 through f_n are irreducible in $\mathcal{R}$. The two fundamental rational decompositions

$$h(x) = f_1 \circ f_2 \circ ... \circ f_n$$
$$h(x) = g_1 \circ g_2 \circ ... \circ g_m$$

are *similar* if $n = m$ and f_i and g_i are associate for all indices i. We then say that $\mathcal{R}$ is a *unique decomposition domain* if for every element in $\mathcal{R}$ all fundamental rational decompositions are similar.

The author of [1] considered unique decomposition of polynomials to not exist and provided a counterexample. However, in the structure we developed here, this counterexample no longer stands as the constituent pairs used in [1] are associates.

Finally, since composition is not commutative in general, if f and g are not associate, constituent pairs $f \circ g$ and $g \circ f$ must be considered distinct. Therefore, a counterexample to unique decomposition can be constructed with prime $p, q \geq 2$; $x^{pq} = x^p \circ x^q = x^q \circ x^p$. Therefore, given the structure above, $\mathcal{R}$ is not a unique decomposition domain. Presently, this is our only counterexample to $\mathcal{R}$ being a unique factorization domain, so it remains unclear if the factorization could be considered unique given a new definition of similarity to ignore commutative constituent pairs.

References

[1] Rickards, J., *When Is a Polynomial a Composition of Other Polynomials?*, The American Mathematical Monthly 118, **No. 4** (2011) pp. 358–363.

[2] Tòth, G., *Finite Möebius Groups, Minimal Immersions of Spheres, and Moduli*, 1st edition, Springer, 2001.