2021

Northcott's Theorem: An Adventure in Preperiodic Points of Rational Functions

Evan Michael David
Carleton College

Ian Klein
Carleton College

Benjamin Colby Richardson
Carleton College

Sameer Swarup
Carleton College

Northcott's Theorem: An Adventure in Preperiodic Points of Rational Functions

Evan David , Ian Klein , Ben Richardson , and Sameer Swarup

Advisor: Rafe Jones

February 24, 2021

Abstract

What happens when you repeatedly apply a function to a rational number? Some numbers eventually loop back on themselves; Northcott's theorem states that there are only finitely many of these "preperiodic" numbers for a given rational function. We introduce a notion of height to provide a surprising proof of Northcott's theorem. During this proof, we make an emergency foray into algebra in the form of homogeneous polynomials and polynomial rings. We also introduce the canonical height and notions of good reduction to give insight into when particular points will be preperiodic.

1 Introduction, motivation, definitions

Discrete dynamical systems are systems in which points evolve in ambient space over discrete time intervals. They are of interest in a wide variety of disciplines such as finance, physics and genetics, featuring prominently as a modeling technique for real world phenomena in these fields. We will be exploring discrete dynamical systems generated through iterative maps and, in particular, those generated through rational functions. By 'rational functions,' we refer to the function field $\mathbb{Q}(x)$:

Definition 1. *The field of rational functions with coefficients in $\mathbb{Q}$, that is, functions which can be written as $\frac{p(x)}{q(x)}$, where $p(x), q(x)$ are rational-coefficient polynomials, is referred to as $\mathbb{Q}(x)$.*

Iterating in this function field looks the same as it does elsewhere.

Definition 2. *Let $f(x) \in \mathbb{Q}(x)$, where $\mathbb{Q}(x)$ is the field of rational functions with coefficients in $\mathbb{Q}$. Then*

$$f^n = \underbrace{f \circ f \circ \cdots \circ f}_{n \text{ times}} \quad \text{is the } n\text{th iterate of } f(x).$$

Example 1.1. Let $f(x) = x^2 - 1$. Then $f^2(x) = (x^2-1)^2 - 1 = x^4 - 2x^2$

An important aspect of these systems is the orbits of the inputs. Knowledge of the orbits gives insight into the evolution points undergo, allowing researchers to determine how real-world phenomena may change when discrete dynamical systems are used to model them. For a discrete dynamical system generated by iterating a rational function $f(x) \in \mathbb{Q}(x)$, the orbit of a rational number $q \in \mathbb{Q}$ is:

Definition 3 (Orbit)**.** *For $q \in \mathbb{Q}$, we call $O_f(q) = \{f^n(q) \mid n \in \mathbb{N}\}$ the orbit of q under f.*

Some inputs will wander on forever, never returning to where they were previously, making their orbits infinite. However, there are some inputs for which the orbits are finite. As an example, let's take a look at the rational function $f(x) = x^2$.

Example 1.2. Let

$$f(x) = x^2.$$

- $O_f(2) = \{2, 4, 16, ...\}$ is a wandering orbit as the sequence is strictly increasing.
- $O_f(-1) = \{-1, 1, 1....\}$ is a finite orbit as after the second term, all following terms are equal to 1.
- $O_f\left(\frac{-1+\sqrt{-3}}{2}\right) = \left\{\frac{-1+\sqrt{-3}}{2}, \frac{-1-\sqrt{-3}}{2}, \frac{-1+\sqrt{-3}}{2}...\right\}$ is also a finite orbit as the sequence oscillates between $\frac{-1+\sqrt{-3}}{2}$ and $\frac{-1-\sqrt{-3}}{2}$.

In our example above, there is a key distinction to be made. Although both $O_f(-1)$ and $O_f\left(\frac{-1+\sqrt{-3}}{2}\right)$ are finite orbits, -1 is a preperiodic point whilst $\left(\frac{-1+\sqrt{-3}}{2}\right)$ is a periodic point. This is because at no point in its orbit does -1 come back to equalling itself whilst $f\left(\frac{-1+\sqrt{-3}}{2}\right)$ does. To formalize the difference between a preperiodic and periodic point we refer to the definitions below:

Definition 4 (Preperiodic)**.** *For a given $f \in \mathbb{Q}(x)$, we call $q \in \mathbb{Q}$ preperiodic under f if $O_f(q)$ is finite. Equivalently, $f^n(q) = f^m(q)$ for some $n, m \geq 1$ with $n \neq m$.*

Definition 5 (Periodic)**.** *For a given $f \in \mathbb{Q}(x)$, we call $q \in \mathbb{Q}$ periodic under f if $f^n(q) = q$ for some $n \geq 1$.*

Remark. Observe that once the function has crossed its path, the construction of function iteration demands it will trace that loop indefinitely. This is why the orbit will be finite.

To further consolidate the definition of preperiodic points and illustrate how the number of rational preperiodic points depends on the rational function being investigated, we look at the rational functions $f(x) = x^2 - \frac{133}{144}$ and $g(x) = x^2 + 1$.

Example 1.3. Let

$$f(x) = x^2 - \frac{133}{144}.$$

- One preperiodic point under this function is $\frac{-7}{12}$. After all, $f\left(\frac{-7}{12}\right) = \frac{49}{144} - \frac{133}{144} = \frac{-7}{12}$. Then $O_f\left(\frac{-7}{12}\right) = \left\{\frac{-7}{12}\right\}$.
- Or consider $\frac{11}{12} \to \frac{-1}{12} \to \frac{-11}{12} \to \frac{-1}{12}$ with $O_f\left(\frac{-11}{12}\right) = \left\{\frac{11}{12}, \frac{-1}{12}, \frac{-11}{12}\right\}$.
- An example of a non-preperiodic point is $\frac{5}{12}$; $O_f(\frac{5}{12}) = \left\{\frac{5}{12}, \frac{-3}{4}, \frac{-13}{36}, \frac{-257}{324}, \frac{-7{,}727}{26{,}244}, \ldots\right\}$.

Example 1.4. Alternatively, let

$$g(x) = x^2 + 1.$$

- Consider $\frac{1}{2}$: $O_f(\frac{1}{2}) = \left\{\frac{5}{4}, \frac{41}{16}, \frac{1{,}937}{256}, \frac{3{,}817{,}505}{65{,}536}, \ldots\right\}$.
- Or $\frac{-10}{7}$: $O_f(\frac{-10}{7}) = \left\{\frac{-10}{7}, \frac{149}{49}, \frac{24{,}602}{2{,}401}, \frac{611{,}023{,}205}{5{,}764{,}801}, \ldots\right\}$.

In fact, no points in $\mathbb{Q}$ will be preperiodic over $g(x)$. After all, $x^2 + 1 = x$ has no rational solutions, and after the first iteration, the output must be greater than or equal to one. From there, outputs will be strictly increasing.

One big question we can ask, then, is how many rational preperiodic points a given function will have. More broadly, arithmetic dynamics is a field of discrete dynamical systems that deal with the repeated application of functions with coefficients in fields of number-theoretic interest, applied to points in those same fields. An early and powerful result in the field is Northcott's theorem. As applied to $\mathbb{Q}(x)$, it restricts the quantity of rational preperiodic points of rational functions with rational coefficients to a finite bound. This is somewhat unexpected. After all, if we consider the set of inputs in $\mathbb{C}$, we find infinitely many preperiodic points for any rational function.

Theorem 1 (Northcott)**.** *For any $f(x) \in \mathbb{Q}(x)$ with $\deg(f(x)) \geq 2$, $\{q \in \mathbb{Q} \mid q$ is preperiodic $\}$ is finite.*

You can formulate this alternatively:

> Some fractions loop back whence they grew.
> By Northcott, they are finitely few.
> They dance and they gyre,
> while others perspire—
> those poor $\aleph_0$-many of $\mathbb{Q}$.

This paper presents a proof of Northcott's theorem that we hope will be accessible to an undergraduate audience. We then examine canonical heights and notions of 'good' reduction, which give more insight into finding preperiodic points.

2 From a height bound to Northcott's theorem

Consider our examples from before. How did we know that some of those points would not be preperiodic? Well, you trusted us. But that trust wasn't blind. Let's consider $\frac{-5}{12}$ under f again:

$$O_f(\frac{5}{12}) = \left\{\frac{5}{12}, \frac{-3}{4}, \frac{-13}{36}, \frac{-257}{324}, \frac{-7,727}{26,244}, \ldots\right\}.$$

On an intuitive level, it makes sense that this won't be preperiodic. The 'complexity' of these rational outputs is generally increasing; each output is more unruly than the last. This intuition is our compass towards a proof of Northcott's theorem, but we need to formalize this notion of 'complexity' first.

Definition 6 (Height)**.** *For $\frac{a}{b} \in \mathbb{Q}$ with a, b in lowest terms and $a \neq 0$, we call $H(q) = \max\{|a|, |b|\}$ the height of $\frac{a}{b}$. When $a = 0$, $H(0) = 1$.*

Example 2.1. We can build some intuitions here.

- As we want the larger of the numerator and denominator, $H\left(\frac{1}{3}\right) = 3$.
- As we need to be considering lowest terms, $H\left(\frac{8}{14}\right) = H\left(\frac{4}{7}\right) = 7$.
- As we consider the absolute value of the numerator and denominator, $H\left(\frac{-21}{13}\right) = 21$.

Example 2.2. This allows us to observe that something really was going on earlier. The height of the orbit (or orbit of heights),

$$H\left(O_f\left(\frac{5}{12}\right)\right) = \{12, 4, 36, 324, 26244, 172186884, \ldots\}$$

is generally increasing and does so strictly after $f^2\left(\frac{5}{12}\right)$.

Theorem 2. *For a bound $B \in \mathbb{N}$, there are finitely-many $q \in \mathbb{Q}$ with $H(q) \leq B$. In particular, there are*

$$4\left(\sum_{i=1}^{B} \varphi(i)\right) - 1 \text{ points with height less than or equal to } B.$$

Proof. As $H\left(\frac{r}{s}\right) = \max\{|r|, |s|\} \leq B$, it must be the case that $-B \leq r, s, \leq B$. Observe that for $B > 1$ there will be exactly $4 \cdot \varphi(B)$ points with height of B. After all, we want all fractions $\pm(\frac{j}{B})^{\pm 1}$ where $0 < j < B$ and $(j, B) = 1$. The theorem follows by induction on B. □

Ruling out points whose height increases under iteration is one way to rule out candidates for status as a preperiodic point. If we can rule out enough of them this way, we'll be left with a finite quantity which could possibly be preperiodic. From this, Northcott's Theorem follows.

One way to think about this is 'guaranteed growth.' We'd like a lower bound on the height of a rational function under iteration such that the bound grows as n grows. Actually finding this bound will be quite difficult.

Example 2.3. For

$$k(x) = x^2 - 1,$$

it is true for any $q \in \mathbb{Q}$ that

$$H(k(q)) \geq \frac{1}{2}H(q)^2.$$

- For instance, if $q = \frac{1}{3}$, $H(k(q)) = 9$ while $\frac{1}{2}H(q)^2 = \frac{9}{2}$.
- Or if $q = \frac{10}{3}$, $H(k(q)) = 91$ while $\frac{1}{2}H(q)^2 = 50$.
- If $q = \frac{-17}{29}$, $H(k(q)) = 841$ while $\frac{1}{2}H(q)^2 = \frac{841}{2}$.

Example 2.4. And for

$$f(x) = x^2 - \frac{133}{144},$$

it is true for any $q \in \mathbb{Q}$ that

$$H(f(q)) \geq \left(\frac{1}{134 \cdot 144}\right) H(q)^2.$$

- For instance, if $q = \frac{11}{12}$, $H(f(q)) = 12$ while $\left(\frac{1}{134 \cdot 144}\right) H(q)^2 = \frac{1}{134}$.
- Or if $q = \frac{5}{12}$, $H(f(q)) = 4$ while $\left(\frac{1}{134 \cdot 144}\right) H(q)^2 = \frac{1}{134}$.
- If $q = \frac{1}{2}$, $H(f(q)) = 144$ while $\left(\frac{1}{134 \cdot 144}\right) H(q)^2 = \frac{1}{134 \cdot 36}$.

The goal of this section is merely to prove that *given* a lower bound on the height of a rational function, Northcott's Theorem follows. Sections three through five will produce this bound.

Lemma 1. Let $f \in \mathbb{Q}(x)$ be a rational function and $q \in \mathbb{Q}$ be a preperiodic point of f. Then $\lim_{n \to \infty} H(f^n(q))$ is bounded.

Proof. By hypothesis, q is a preperiodic point. Its orbit, then, must be finite. Consider $H(O_f(q)) = \{H(f^n(q)) \mid n \in \mathbb{N}\}$, which must be finite as well. Any finite set of integers will have a maximum, so there is some $k \in \mathbb{N}$ for which $H(f^k(q)) = \max H(O_f(q))$. This $H(f^k(q))$ is a bound on $\lim_{n \to \infty} H(f^n(q))$. □

Theorem 3. *Let $f \in \mathbb{Q}(x)$ have degree $d \geq 2$. Suppose there exists $C_L > 0$ such that for all $q \in \mathbb{Q}$, $H(f(q)) \geq C_L H(q)^d$. Then for all prepreriodic points $x \in \mathbb{Q}$, $H(x) \leq C_L^{\frac{-1}{d-1}}$.*

Proof. Suppose $q \in \mathbb{Q}$ and $H(f((\frac{a}{b})) \geq C_L H(\frac{a}{b})^d$ for some constant $C_L > 0$. In almost all cases, we will have $C_L < 1$, but if we don't, we choose a smaller C_L below 1 and the inequality will still hold. Then we can deduce:

$$H(f^2((q)) \geq C_L H(q)^d \geq (C_L)^{1+d} (H(q))^{d^2}.$$

This in turn will let us deduce that

$$H\left(f^3(q)\right) \geq C_L H(f^2((q))^2 \geq (C_L)^{1+d} (H(f(q)))^{d^2} \geq (C_L)^{1+d+d^2} (H(q))^{d^3}.$$

We'll prove the generalization by induction on n. We've already shown that the base case holds. Suppose that for some $n \in \mathbb{N}$,

$$H(f^n(q)) \geq C_L^{1+d+d^2+\ldots+d^{n-1}} H(q)^{d^n}.$$

Then

$$\begin{aligned}
H(f^{n+1}(q)) &= H(f(f^n(q))) \\
&\geq C_L H(f^n(q))^d \\
&\geq C_L \left(C_L^{1+d+d^2+\ldots+d^{n-1}} H(q)^{d^n}\right)^d \\
&= C_L \left(C_L^{d+d^2+d^3+\ldots+d^n} H(q)^{d^{n+1}}\right) \\
&= C_L^{1+d+d^2+d^3+\ldots+d^n} H(q)^{d^{n+1}}.
\end{aligned}$$

And, as the statement holds for $n+1$, the statement must hold for all natural numbers. Now observe that

$$\begin{aligned}
\lim_{n \to \infty} H\left(f^n(q)\right) &\geq \lim_{n \to \infty} C_L^{1+d+d^2+\ldots+d^{n-1}} H(q)^{d^n} \qquad (1) \\
&\geq \lim_{n \to \infty} \left(C_L^{\frac{1+d+d^2+\ldots+d^{n-1}}{d^n}} H(q)\right)^{d^n} \\
&\geq \lim_{n \to \infty} \left(C_L^{\frac{1}{d-1}} H(q)\right)^{d^n} \\
&\geq \lim_{n \to \infty} \left(\left(C_L^{\frac{1}{d-1}} H(q)\right)^d\right)^n \qquad (2)
\end{aligned}$$

The second-to-last step follows because

$$\frac{1+d+\ldots+d^{n-1}}{d^n}\left(\frac{d-1}{d-1}\right)=\frac{d^n-1}{d^n(d-1)}=\frac{d^n-1}{d^n}\left(\frac{1}{d-1}\right)<\frac{1}{d-1},$$

for any $d \geq 2$, $n \geq 1$, and we already know that $C_L < 1$. A constant below 1 raised to a greater power will be smaller.

But notice that n is the only moving piece in that last limit. This is then a geometric sequence, which will diverge to infinity $\iff \left(C_L^{\frac{1}{d-1}}H(q)\right)^d > 1 \iff C_L^{\frac{1}{d-1}}H(q) > 1$. And if (2) diverges to infinity, the left side of (1) will as well.

Hence, if q is preperiodic, $H(f^n(q))$ cannot diverge to infinity by Lemma 1. If $H(f^n(q))$ does not diverge to infinity, then neither will the right side above and so $C_L^{\frac{1}{d-1}}H(q) \leq 1$. This is equivalent to saying that $H(q) \leq C_L^{\frac{-1}{d-1}}$. □

Example 2.5. For

$$f(x)=x^2-\frac{133}{144}, \text{ we know that } C_L=\frac{1}{133\cdot 144}.$$

Hence, any preperiodic point of f will have height less than or equal to $C_L^{\frac{-1}{2-1}} = 19,152$. By the output of my crude Java function,

$$4\left(\sum_{i=1}^{19,152}\varphi(i)\right)-1=445,979,919.$$

Checking all 446 million of these points for preperiodicity is a miserable task, but a finite one. Actually, there will only end up being eight preperiodic points:

$$\pm\frac{1}{12},\pm\frac{7}{12},\pm\frac{11}{12}, \text{ and } \pm\frac{19}{12}.$$

The combination of theorems 2 and 3 give the result of Northcott's. However, we have not yet satisfied its condition: that there exists this $C_L > 0$ with $H(f^n(q)) \geq C_L H(q)^d$. This is a hard bound to find. It must, after all, account both for the behavior of f itself—plug in just the right q and the numerator polynomial might be zero—and for the possible cancellation on top and bottom. The rest of this paper establishes the bound.

3 Conditions for producing a height bound

The previous section demonstrated that, given a lower bound on the height of a rational function, Northcott's Theorem holds. The rest of this paper works to establish that height bound.

In order to do so, a natural question is to ask how we expect the height of a rational number to increase upon the application of a function, and what can go 'wrong', i.e. increase height by less than is expected.

If we apply $f(x) = x^d$ to a rational number in lowest terms $\frac{a}{b}$, we know that $\left(\frac{a}{b}\right)^d = \frac{a^d}{b^d}$, and since a and b are relatively prime, a^d and b^d are relatively prime. Therefore, $\max\{|a^d|, |b^d|\} = (\max\{|a|, |b|\})^d$. This means that

$$H\left(f\left(\frac{a}{b}\right)\right)=H\left(\frac{a}{b}\right)^d.$$

So, we have an idea of the expected increase in height upon application of a rational function. How can this go wrong? If there is cancellation upon application of our function, then the height will not increase by the

expected amount. Here is an example. Let $g(x) = x^2 - \frac{133}{144}$. Considering the rational number $-\frac{1}{24}$, we have that

$$f\left(-\frac{1}{24}\right) = \frac{144(-1/24)^2 - 133}{144} = \frac{1/4 - 133}{144} = \frac{-531/4}{144} = -\frac{59}{64}.$$

So, some cancellation is present in the application of this function, but it is very unclear how this cancellation comes about. 64 does not divide 144, so the cancellation is more complicated than simply taking the gcd of the numerator and denominator. The culprit here is the $\frac{1}{4}$ which appears in the numerator. In order to deal with this problem, we will take a detour through the world of homogeneous polynomials.

Definition 7 (Homogeneous polynomial)**.** *A polynomial $p(a, b)$ in two variables is called homogeneous if the sum of the degrees of each term are equal—that is, if there exists an $n \in \mathbb{N}$ such that $p(a, b) = \sum_{i=0}^{n} c_i a^{n-i} b^n$.*

For instance, $ab + 2b^2$ and $a^3 - 2a^2b + 3ab^2 - b^3$ are homogeneous, while $a^3 + ab$ is not. As may be apparent, there is a close relationship between polynomials of a single variable and certain homogeneous polynomials.

Definition 8 (Homogenization)**.** *For a single-variable polynomial $P(x) = \sum_{i=0}^{n} c_i x^i$, the homogenization of $P(x)$ is $P(a, b) = \sum_{i=0}^{n} c_i A^i B^{n-i}$.*

Why do we care about homogenizing polynomials? Well, if we look at the degree d homogenization $h(a, b)$ of a degree d polynomial $h(x)$, when considering $\frac{a_0}{b_0} \in \mathbb{Q}$ a rational number in lowest terms, observe that

$$h(a_0, b_0) = b_0^d h\left(\frac{a_0}{b_0}\right).$$

So, when considering a rational function of degree d, $f(x) = \frac{N(x)}{D(x)}$ where $N(x)$ and $D(x)$ have no common roots, we define the homogenization of $f(x)$ to be

$$f(a, b) = \frac{N(a, b)}{D(a, b)},$$

where $N(a, b)$ and $D(a, b)$ are the degree d homogenizations of $N(x)$ and $D(x)$ respectively. We then know that for a given rational number $\frac{a_0}{b_0}$ in lowest terms,

$$f(a_0, b_0) = \frac{N(a_0, b_0)}{D(a_0, b_0)} = \frac{b_0^d N\left(\frac{a_0}{b_0}\right)}{b_0^d D\left(\frac{a_0}{b_0}\right)} = f\left(\frac{a_0}{b_0}\right).$$

So, in a sense we can homogenize rational functions without changing the value of the rational function. Since we can write rational functions as having integer coefficients, instead of considering applying the function to a rational number $\frac{a_0}{b_0}$ in lowest terms, we can instead consider applying the homogenized rational function to (a_0, b_0). Since the numerator and denominator will contain integer polynomials, and a_0, b_0 are also integers, we will have integers in the numerator and the denominator, which allows us to consider the gcd when looking at cancellation.

For an example, here is the problematic case that we initially discussed. Given

$$g(x) = x^2 - \frac{133}{144} = \frac{144x^2 - 133}{144},$$

homogenize this function to get

$$g(a, b) = \frac{144a^2 - 133b^2}{144b^2}.$$

Then, look at

$$g(-1,24) = \frac{144(-1)^2 - 133(24)^2}{144(24)^2} = -\frac{76464}{82944} = -\frac{59}{64}.$$

After homogenizing, cancellation behaves much nicer. For instance, we know that since $24 \mid 144$, we see that 24 divides the numerator and denominator and the 24 in the denominator cancels against the part of the 144 in the numerator. This homogenization allows us to see this sort of cancellation much easier than we previously could. Now that we have some motivation behind homogenization, what can we do with it?

3.1 An Upper Bound on Cancellation

We have established that homogenization is important for understanding how much fractions can cancel upon application of a function. But why? and in what way? This theorem shows us how homogenization of polynomials gives us an upper bound on the amount of cancellation which can occur, given an assumption that will be proved later:

Theorem 4. *Let $f(x) \in \mathbb{Q}(x)$, and let $N(x), D(x)$ be polynomials with integer coefficients and no common roots such that $f(x) = \frac{N(x)}{D(x)}$. Let $N(a,b)$ and $D(a,b)$ denote the homogenizations of N and D. Furthermore, assume that there exist homogeneous polynomials with rational coefficients $Q_{1,2}(a,b)$ and $R_{1,2}(a,b)$ with the property that*

$$a^j = Q_1(a,b)N(a,b) + R_1(a,b)D(a,b)$$
$$b^k = Q_2(a,b)N(a,b) + R_2(a,b)D(a,b).$$

where j and k are positive integer. Then for any relatively prime integers a_0 and b_0, there exists some $M > 0$ such that $H(f(\frac{a_0}{b_0})) \geq \frac{\max\{N(a_0,b_0),D(a_0,b_0)\}}{M}$.

Proof. Let $d_{Q_1}, d_{Q_2}, d_{R_1}, d_{R_2}$ be the degrees of $Q_1(a,b), Q_2(a,b), R_1(a,b), R_2(a,b)$ respectively. Then, write that

$$Q_1(a,b) = \textstyle\sum_{j=1}^{d_{Q_1}} q_{1j}a^j b^{d_{Q_1}-j} \quad Q_2(a,b) = \sum_{j=1}^{d_{Q_2}} q_{2j}a^j b^{d_{Q_2}-j}$$
$$R_1(a,b) = \textstyle\sum_{j=1}^{d_{R_1}} r_{1j}a^j b^{d_{R_1}-j} \quad R_2(a,b) = \sum_{j=1}^{d_{R_2}} r_{2j}a^j b^{d_{R_2}-j}.$$

Next, define L_1 as the least common multiple of the denominators of the coefficients of $Q_1(a,b)$ and $R_1(a,b)$ in lowest terms, and define L_2 as the least common multiple of the denominators of the coefficients of $Q_2(a,b)$ and $R_2(a,b)$ in lowest terms. We then know that

$$L_1Q_1(a,b), L_2Q_2(a,b), L_1R_1(a,b), L_2R_2(a,b) \in \mathbb{Z}[a,b],$$

which is to say that these are now homogeneous polynomials with integer coefficients Which satisfy the following system of equations:

$$L_1a^j = L_1Q_1(a,b)N(a,b) + L_1R_1(a,b)D(a,b)$$
$$L_2b^k = L_2Q_2(a,b)N(a,b) + L_2R_2(a,b)D(a,b).$$

Now, if we consider relatively prime integers a_0 and b_0, we know that

$$L_1a_0^j = L_1Q_1(a_0,b_0)N(a_0,b_0) + L_1R_1(a_0,b_0)D(a_0,b_0)$$
$$L_2b_0^k = L_2Q_2(a_0,b_0)N(a_0,b_0) + L_2R_2(a_0,b_0)D(a_0,b_0).$$

What can we do with this? Well, now we know that $L_1a_0^j$ and $L_2b_0^k$ can be written as sums of products of integers, so now we can talk about divisibility. We know that since $\gcd(N(a_0,b_0), D(a_0,b_0))$ divides both $N(a_0,b_0)$ and $D(a_0,b_0)$, looking at our system of equations, we can see that

$$\gcd(N(a_0,b_0), D(a_0,b_0)) \mid L_1a_0^j$$
$$\gcd(N(a_0,b_0), D(a_0,b_0)) \mid L_2b_0^k.$$

Since a_0 and b_0 are relatively prime, a_0^j and b_0^k are also relatively prime, which means that if any integer divides both $L_1 a_0^j$ and $L_2 b_0^k$, it must be the case that that integer divides $\text{lcm}(L_1, L_2)$. Therefore, we know that

$$\gcd(N(a_0, b_0), D(a_0, b_0)) \leq \text{lcm}(L_1, L_2),$$

where L_1 and L_2 do not depend on a_0 or b_0. Now, we also know that

$$H\left(f\left(\frac{a_0}{b_0}\right)\right) = \max\left\{\frac{N(a_0,b_0)}{\gcd(N(a_0,b_0),D(a_0,b_0))}, \frac{D(a_0,b_0)}{\gcd(N(a_0,b_0),D(a_0,b_0))}\right\} \geq \max\left\{\frac{N(a_0,b_0)}{\text{lcm}(L_1,L_2)}, \frac{D(a_0,b_0)}{\text{lcm}(L_1,L_2)}\right\}.$$

Taking $M = \text{lcm}(L_1, L_2)$, this proves our theorem.

□

Now that we have a grasp on how cancellation might affect the increase of height of a rational number under application of a rational function, what's next? Is that everything that can 'go wrong', i.e. is that the only factor that inhibits height growth? Not quite. If we think about the function $h(x) = x^2 - 3$, notice that $h(2) = 1$. Applying this quadratic function decreased the height of our rational input! This largely has to do with the degree 0 term of -3. Is this issue one we can easily account for?

It turns out, yes. If we put together our concerns about cancellation and our concerns about the coefficients of the terms in our function, and add the same assumption we made about our homogeneous polynomials in the previous theorem, we can produce a lower bound on the amount that height increases upon application of a rational function.

Theorem 5. *Let $f(z) \in \mathbb{Q}(z)$ with $N(z), D(z)$ such that $f(z) = \frac{N(z)}{D(z)}$ and $N(z), D(z)$ have no common roots and the greatest common divisor of their coefficients is 1. Then suppose that there exist homogeneous polynomials $Q_{1,2}(a,b), R_{1,2}(a,b)$ of degree m such that*

$$a^{d+m} = Q_1(a,b)N(a,b) + R_1(a,b)D(a,b) \qquad \text{and} \qquad b^{d+m} = Q_2(a,b)N(a,b) + R_2(a,b)D(a,b), \tag{3}$$

where $d = \max\{\deg(N(z)), \deg(D(z))\}$. For

$$Q_1(a,b) = \sum_{i=0}^{m} q_i a^i b^{d-i}, \qquad \text{let} \qquad C_{Q_1} = \sum_{i=0}^{m} |q_i|$$

and similarly for $C_{Q_2}, C_{R_1}, C_{R_2}$. Additionally suppose that there exists $M \in \mathbb{N}$ such that for all $\frac{a_0}{b_0} \in \mathbb{Q}$,

$$M \geq \gcd(N(a_0, b_0), D(a_0, b_0)).$$

It follows that, for all $\frac{a_0}{b_0} \in \mathbb{Q}$ with a_0, b_0 relatively prime integers,

$$H\left(\frac{a_0}{b_0}\right)^d \leq H\left(f\left(\frac{a_0}{b_0}\right)\right) \cdot \max\left\{C_{Q_1} + C_{R_1}, C_{Q_2} + C_{R_2}\right\} \cdot M.$$

Proof. Observe that

$$H\left(\frac{a_0}{b_0}\right)^{d+m} = \max\{|a^{d+m}|, |b^{d+m}|\} \tag{4}$$

$$= \max\left\{\left|Q_1(a_0,b_0)N(a_0,b_0) + R_1(a_0,b_0)D(a_0,b_0)\right|, \left|Q_2(a_0,b_0)N(a_0,b_0) + R_2(a_0,b_0)D(a_0,b_0)\right|\right\}, \tag{5}$$

where the second equality follows from (3). Applying the triangle inequality twice, we find that this is less than or equal to

$$\max\Big\{|Q_1(a_0,b_0)| \cdot |N(a_0,b_0)| + |R_1(a_0,b_0)| \cdot |D(a_0,b_0)|, \tag{6}$$
$$|Q_2(a_0,b_0)| \cdot |N(a_0,b_0)| + |R_2(a_0,b_0)| \cdot |D(a_0,b_0)|\Big\}.$$

Ex hyposthesi, we know that $\frac{\max\{N(a_0,b_0),D(a_0,b_0)\}}{M} \leq H\left(f\left(\frac{a_0}{b_0}\right)\right)$. Hence, we can substitute to find that the previous expression is less than or equal to

$$\max\Big\{\Big(|Q_1(a_0,b_0)|+|R_1(a_0,b_0)|\Big)\Big(\Big|H\left(f\left(\frac{a_0}{b_0}\right)\right)\cdot M\Big|\Big),\Big(|Q_2(a_0,b_0)|+|R_2(a_0,b_0)|\Big)\Big(\Big|H\left(f\left(\frac{a_0}{b_0}\right)\right)\cdot M\Big|\Big)\Big\} \tag{7}$$

$$= \Big|H\left(f\left(\frac{a_0}{b_0}\right)\right)\cdot M\Big|\cdot\max\Big\{\Big(|Q_1(a_0,b_0)|+|R_1(a_0,b_0)|\Big),\ \ \Big(|Q_2(a_0,b_0)|+|R_2(a_0,b_0)|\Big)\Big\}. \tag{8}$$

At this point, one may observe that for any $\frac{a_0}{b_0}\in\mathbb{Q}$ and $0\leq i\leq m$,

$$|a_0^i b_0^{m-i}| \leq \max\{|a_0|^m,|b_0|^m\} = H\left(\frac{a_0}{b_0}\right)^m.$$

It therefore follows that

$$|Q_1(a_0,b_0)| = \left|\sum_{i=0}^{m} q_i a_0^i b_0^{m-i}\right| \leq \sum_{i=0}^{m}|q_i a_0^i b_0^{m-i}| \leq \sum_{i=0}^{m}|q_i|H\left(\frac{a_0}{b_0}\right)^m = C_{Q_1}\cdot H\left(\frac{a_0}{b_0}\right)^m.$$

Hence, (8) is less than or equal to

$$\Big|H\left(f\left(\frac{a_0}{b_0}\right)\right)\cdot M\Big|\cdot\max\Big\{\Big(C_{Q_1}+C_{R_1}\Big)H\left(\frac{a_0}{b_0}\right)^m,\ \ \Big(C_{Q_2}+C_{R_2}\Big)H\left(\frac{a_0}{b_0}\right)^m\Big\} \tag{9}$$

$$= H\left(\frac{a_0}{b_0}\right)^m\cdot H\left(f\left(\frac{a_0}{b_0}\right)\right)\cdot M\cdot\max\Big\{\Big(C_{Q_1}+C_{R_1}\Big),\ \ \Big(C_{Q_2}+C_{R_2}\Big)\Big\}. \tag{10}$$

As $(4)=(5)\leq(6)\leq(7)=(8)\leq(9)=(10)$, we have that

$$H\left(\frac{a_0}{b_0}\right)^{d+m} \leq H\left(\frac{a_0}{b_0}\right)^m\cdot H\left(f\left(\frac{a_0}{b_0}\right)\right)\cdot M\cdot\max\Big\{\Big(C_{Q_1}+C_{R_1}\Big),\ \ \Big(C_{Q_2}+C_{R_2}\Big)\Big\}.$$

The result follows. □

To summarize, we now are able to construct a lower bound constant, here

$$\left(M\cdot\max\{(C_{Q_1}+C_{R_1}),\ (C_{Q_2}+C_{R_2})\}\right)^{-1}.$$

We have already shown that we can set $M=\mathrm{lcm}(L_1,L_2)$ as defined in the previous proof, so once we have these certain polynomials $Q_{1,2}(a,b),R_{1,2}(a,b)$ always do exist, then we have proven Northcott's Theorem. The following section seeks to prove this claim.

4 Powers of a and b as linear combinations of homogeneous polynomials

In order to prove Northcott's theorem, for a variety of reasons, we will want to be able to find, for any two rational homogeneous polynomials $N(a,b)$ and $D(a,b)$, two other rational polynomials $Q_1(a,b)$ and $R_1(a,b)$ such that

$$a^n = N(a,b)Q_1(a.b)+D(a,b)R_1(a,b),$$

and two more polynomials $Q_2(a,b)$ and $R_2(a,b)$ such that

$$b^n = N(a,b)Q_2(a.b)+D(a,b)R_2(a,b).$$

In this section we will prove this fact using an algebraic proof, followed by a variety of examples.

4.1 Proving existence and motivation

So far, we have assumed an ability to write a to some power and b to some power as a linear combination of $N(a,b)$ and $D(a,b)$, the homogenizations of $N(z)$ and $D(z)$ given some rational function $f(z) = N(z)/D(z)$. Not only does previous work in this paper rely on this fact, but it will become a useful tool for finding M, the assumed upper bound on $\gcd(N(a,b), D(a,b))$ mentioned in theorem 5. Explicitly, the theorem is

Theorem 6. *Let $f(z) \in \mathbb{Q}(z)$. Write $f(z) = N(z)/D(z)$ with $N, D \in \mathbb{Q}[z]$ having no common roots in $\mathbb{C}$ and $N(z)$ is monic. Then, there exist homogeneous $Q_1(a,b), R_1(a,b), Q_2(a,b), R_2(a,b) \in \mathbb{Q}[a,b]$ of degree $m \geq 0$ such that*

$$a^{n+m} = Q_1(a,b)N(a,b) + R_1(a,b)D(a,b) \qquad \text{and} \qquad b^{n+m} = Q_2(a,b)N(a,b) + R_2(a,b)D(a,b),$$

where $N(a,b)$, $D(a,b)$ are homogenizations of $N(z)$ and $D(z)$ and $n = \max\{\deg(N(z)), \deg(D(z))\}$.

Proof. Although this is only a result for $\mathbb{Q}[a,b]$, we will begin our proof in integral domain of $\mathbb{Q}[z]$. Since $\mathbb{Q}$ is a field, we know that $\mathbb{Q}[z]$ is a Euclidean domain, and thus a principal ideal domain. It follows then that

$$\langle N(z), D(z)\rangle = \{A_1(z)N(z) + A_2(z)D(z) \mid A_1, A_2 \in \mathbb{Q}[z]\} \tag{11}$$

is a principal ideal. By the definition of a principal ideal there must then exist some $T(z) \in \mathbb{Q}[z]$ such that $\langle T(z)\rangle = \langle N(z), D(z)\rangle$. If in Eq. 11 we set $A_1 = 1$ and $A_2 = 0$ and then $A_1 = 0$ and $A_2 = 1$ we see that $N(z), D(z) \in \langle N(z), D(z)\rangle$. This means that

$$N(z) = B_1(z)T(z) \text{ and } D(z) = B_2(z)T(z)$$

where $B_1(z), B_2(z) \in \mathbb{Q}[z]$. Now, since N and D have no common roots in $\mathbb{C}$, $\deg(T(z)) = 0$. This implies that $T(z)$ is a unit and

$$\langle T(z)\rangle = \langle N(z), D(z)\rangle = \mathbb{Q}[z].$$

This gives us the result that $1 \in \langle N(z), D(z)\rangle$, which means that there exist $Q_2, R_2 \in \mathbb{Q}[z]$ such that

$$1 = Q_2(z)N(z) + R_2(z)D(z). \tag{12}$$

In general, we can write some polynomial $f(z) \in \mathbb{Q}[z]$ as $f(z) = \frac{1}{b^d} f(a,b)$ where $f(a,b) \in \mathbb{Q}[a,b]$ is the degree d homogenization of $f(z)$. Because of this, if we say that m is the degree of the homogenization $Q_2(a,b)$ and $R_2(a,b)$ and n is the degree of the homogenization of $N(a,b)$ and $D(a,b)$, we can write Eq. 12 as

$$1 = \frac{1}{b^m}Q_2(a,b)\frac{1}{b^n}N(a,b) + \frac{1}{b^m}R_2(a,b)\frac{1}{b^n}D(a,b),$$

and then as

$$b^{n+m} = Q_2(a,b)N(a,b) + R_2(a,b)D(a,b).$$

We have now proved half of the theorem.

Now we define two new polynomials in $\mathbb{Q}[z']$

$$N'(z') = \frac{1}{a^n}N(a,b) \qquad \text{and} \qquad D'(z') = \frac{1}{a^n}D(a,b)$$

where $z' = \frac{b}{a}$.

It is worth pausing here to make sure that $\frac{1}{a^n}N(a,b)$ truly is a polynomial in z'. Let's say

$$N(a,b) = \alpha_0 a^n + \alpha_1 a^{n-1}b + ... + \alpha_n b^n.$$

Then we can write that

$$\begin{aligned}\frac{1}{a^n}N(a,b) &= \alpha_0\frac{1}{a^n}a^n + \alpha_1\frac{1}{a^n}a^{n-1}b + ... + \alpha_j\frac{1}{a^n}a^{n-j}b^j + ... + \alpha_n\frac{1}{a^n}b^n \\ &= \alpha_0 + \alpha_1\frac{1}{a}b + ... + \alpha_j\frac{1}{a^j}b^j + ... + \alpha_n\frac{1}{a^n}b^n \\ &= \alpha_0 + \alpha_1 z' + ... + \alpha_j(z')^j + ... + \alpha_n(z')^n,\end{aligned}$$

which is a polynomial in z'. The same argument can be used for $D'(z')$.

We should check to make sure that z' is not a common factor of $N'(z')$ and $D'(z')$. We know that $N(a,b)$ and $D(a,b)$ have no common factors in $\mathbb{Q}[a,b]$. This implies that either $N(a,b)$ or $D(a,b)$ has a non-zero coefficient in front of the a^n term. If neither did, the homogeneous polynomials would share a factor of b. This then implies that either N' or D' has a non-zero coefficient in front of the $(z')^0$ term, and therefore does not have zero as a root.

This allows us to say, in a parallel argument to the first half of the theorem, that

$$\langle N'(z), D'(z)\rangle = \mathbb{Q}[z'] \qquad \Rightarrow \qquad 1 = Q_1(z')N'(z') + R_1(z')D(z').$$

where once again we have Q_1 and R_1 are degree m. We can now write that

$$a^{n+m} = a^m Q_1(z')a^n N'(z') + a^m R_1(z')a^n D'(z')$$

which can be rewritten as

$$a^{n+m} = Q_1(a,b)N(a,b) + R_1(a,b)D(a,b).$$

This concludes the proof. □

4.2 Examples of Theorem 6

Great, now we have a proof of the existence of our desired linear combinations. This proof is somewhat abstract, and involves some confusing notation. So, here are two concrete examples following the heuristic established in the proof of theorem 6.

Example 1: Let $f(z) = \frac{z^2+2}{z}$. We then observe that $n = 2$, $N(z) = z^2+2$, $N(a,b) = a^2+2b^2$, $D(z) = z$, and $D(a,b) = ab$. We also have that

$$(\frac{1}{2})(z^2+2) + (-\frac{1}{2}z)(z) = 1,$$

so $Q_2(z) = \frac{1}{2}$ and $R_2(z) = -\frac{1}{2}z$. Furthermore, when substituting $z' = \frac{b}{a}$, we see that $N'(z') = 2z'^2 + 1$, and $D'(z') = z'$. observe that

$$(1)(2z'^2+1) + (-2z')(z') = 1,$$

so $Q_1(z') = 1$, $R_1(z') = -2z'$, and $m = 1$. This gives us the homogenizations

$$\begin{aligned} Q_1(a,b) &= a \\ R_1(a,b) &= -2b \\ Q_2(a,b) &= \frac{1}{2}b \\ R_2(a,b) &= -\frac{1}{2}a. \end{aligned}$$

Here, observe that

$$\begin{aligned} a(a^2+2b^2) - 2b(ab) &= a^3 \\ \frac{1}{2}b(a^2+2b^2) - \frac{1}{2}a(ab) &= b^3. \end{aligned}$$

We have thus expressed a power of a and a power of b as linear combinations of homogeneous polynomials involving the homogenizations of $N(z) = z^2+2$ and $D(z) = z$.

Example 2: Now, we will examine a slightly more complex example. Let $g(z) = \frac{z^3+z+1}{z^2-1}$. We observe that $n = 3$, $N(z) = z^3 + z + 1$, $N(a, b) = a^3 + ab^2 + b^3$, $D(z) = z^2 - 1$, and $D(a, b) = a^2b - b^3$. Following the Euclidean algorithm, we arrive at the result that

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

This gives us $Q_2(z) = -\frac{1}{3}(1 - 2z)$, and $R_2(z) = -\frac{1}{3}(2z^2 - z + 4)$. Now, letting $z' = \frac{b}{a}$, we see that

$$N'(z') = z'^3 + z'^2 + 1, D'(z') = -z'^3 + z'.$$

One again, applying the Euclidean algorithm yields the result that

$$\frac{1}{3}(-z'^2 - z' + 3)(z'^3 + z'^2 + 1) + \frac{1}{3}(-z'^2 - 2z' + 1)(-z'^3 + z') = 1.$$

So, $Q_1(z') = \frac{1}{3}(-z'^2-z'+3)$, $R_1(z') = \frac{1}{3}(-z'^2-2z'+1)$, and $m = 2$. These results give us the homogenizations

$$\begin{aligned} Q_1(a, b) &= \frac{1}{3}(-b^2 - ab + 3a^2) \\ R_1(a, b) &= \frac{1}{3}(-b^2 - 2ab + a^2) \\ Q_2(a, b) &= -\frac{1}{3}(b^2 - 2ab) \\ R_2(a, b) &= -\frac{1}{3}(2a^2 - ab + 4b^2) \end{aligned}$$

This gives us the result that

$$\begin{aligned} \frac{1}{3}(-b^2 - ab + 3a^2)(a^3 + ab^2 + b^3) + \frac{1}{3}(-b^2 - 2ab + a^2)(a^2b - b^3) &= a^5 \\ -\frac{1}{3}(b^2 - 2ab)(a^3 + ab^2 + b^3) - \frac{1}{3}(2a^2 - ab + 4b^2)(a^2b - b^3) &= b^5 \end{aligned}$$

We have thus expressed a power of a and a power of b as linear combinations of homogeneous polynomials involving the homogenizations of $N(z)$ and $D(z)$.

Example 3: Let $f(z) = \frac{z^5-2z+4}{z^3+4z^2-6}$. We can rewrite $f(z)$in the form of $f(z) = \frac{N(a,b)}{D(a,b)}$ where

$$N(a, b) = a^5 - 2ab^4 + 4b^5$$

and

$$D(a, b) = a^3b^2 + 4a^2b^3 - 6b^5.$$

$$a^{10} = (a^5+\frac{1098}{269}a^2b^3+\frac{3276}{269}ab^4-\frac{5076}{269}b^5)N(a,b)+(-\frac{1098}{269}a^4b+\frac{1654}{269}a^3b^2-\frac{2616}{269}a^2b^3+\frac{3876}{269}ab^4-\frac{3384}{269}b^5)D(a,b)$$

We now are almost finished with the proof of Northcott's theorem. All that remains is showing the existence of an upper bound on $(N(a, b), D(a, b))$ which is independent of a and b.

4.3 A Holistic Example

We have frequently discussed $f(x) = x^2 - \frac{133}{144}$ as a rational function, so it is about time that we prove, as in example 2.5, that for $f(x)$, we have that $C_L = \frac{1}{134\cdot144}$ and

$$H(f^n(q)) \geq C_L H(q)^d.$$

First, rewrite $x^2 - \frac{133}{144}$ as $\frac{144x^2-133}{144}$. Then, homogenize the numerator and denominator to degree 2 polynomials to get $\frac{144a^2-133b^2}{144b^2}$. As per theorem 6, we can find $Q_{1,2}(a,b)$ and $R_{1,2}(a,b)$ as

$$\begin{array}{ll} Q_1(a,b) = \frac{1}{144} & R_1(a,b) = \frac{133}{144} \\ Q_2(a,b) = 0 & R_2(a,b) = \frac{1}{144}, \end{array}$$

and notice that

$$a^2 = Q_1(a,b)(144a^2 - 133b^2) + R_1(a,b)(144b^2)$$
$$b^2 = Q_2(a,b)(144a^2 - 133b^2) + R_2(a,b)(144b^2).$$

So, we have, by theorem 4, that $\gcd(144a_0^2 - 133b_0^2, 144b_0^2) \leq 144^2$ for all relatively prime integers a_0, b_0. This is because 144 is the least common multiple of the denominators of the coefficients of Q_1 and R_1, and it is also the least common multiple of the denominators of the coefficients of Q_2 and R_2.

Furthermore, as in theorem 5, we can say that $C_{Q_1} = \frac{1}{144}$, $C_{Q_2} = 0$, $C_{R_1} = \frac{133}{144}$, and $C_{R_2} = \frac{1}{144}$. With this information, we are able to establish this C_L as $(144^2 \cdot \frac{134}{144})^{-1}$, or

$$C_L = \frac{1}{144 \cdot 134}.$$

5 Canonical height: definitions, existence, properties, and examples

Northcott's Theorem guarantees some limited number of preperiodic points, but it doesn't tell us exactly how many there will be or where to find them. This section develops canonical height as a metric for 'how close' to preperiodicity a point is. The next section on good reduction provides tools for finding particular preperiodic points.

To put it simply, we proved Northcott's Theorem by showing that the height of almost all rational points increased too fast for them to be preperiodic. Canonical height measures how fast, on average, a point's height increases.

5.1 Definitions and Properties of the Canonical Height

Let's begin by defining the canonical height.

Definition 9 (Canonical Height)**.** *For a function $f \in \mathbb{Q}(x)$, we call $\hat{H}_f(q) = \lim_{n\to\infty} H(f^n(q))^{\frac{1}{d^n}}$ the canonical height of q under f.*

It is not trivial that this limit actually exists. What follows is a proof of the sequence's convergence. First, though, we need an upper bound on $H(f^n(q))$'s behavior, just like the lower bound we relied on throughout Northcott's theorem.

Lemma 2. For a rational function f, there exists $C_U \in \mathbb{N}$ such that for all q, $H(f(q)) \leq C_U H(q)^d$.

Proof. As f is a rational function, it can be expressed as the ratio of polynomials with integer coefficients: $\frac{c_d x^d + c_{d-1}x^{d-1} + \ldots + c_1 x + c_0}{g_d y^d + g_{d-1} y^{d-1} + \ldots + g_1 y + g_0}$. Then

$$H(f(q)) \leq \max\left\{\left|c_d x^d + c_{d-1}x^{d-1} + \ldots + c_1 x + c_0\right|, \left|g_d y^d + g_{d-1} y^{d-1} + \ldots + g_1 y + g_0\right|\right\}$$
$$\leq \max\left\{x^d \sum_{i=0}^{d} |c_i|, y^d \sum_{i=0}^{d} |g_i|\right\} \leq H(q)^d \max\left\{\sum_{i=0}^{d} |c_i|, \sum_{i=0}^{d} |g_i|\right\}$$

. □

This aside, we can now state and prove the convergence of the canonical height. We do so by proving that it is a Cauchy sequence.

Theorem 7. *The limit* $\lim_{n\to\infty}(H\circ f^n(q))^{\frac{1}{d^n}}$ *exists for all rational functions* f *of degree* $d\geq 2$.

Proof. A sequence converges as a Cauchy sequence if, for any $\epsilon>0$, there exists some N where, for all n, m greater than N, $|a_n-a_m|<\epsilon$. Without loss of generality, we will say that $n>m$ and $n-m=l$. In this case $a_n=(H\circ f^l\circ f^m(q))^{\frac{1}{d^n}}$ and $a_m=(H\circ f^m(q))^{\frac{1}{d^m}}$. If we simplify $r=f^m(q)$ we can rewrite this as

$$a_n=(H\circ f^l(r))^{\frac{1}{d^n}}$$

and

$$a_m=H(r)^{\frac{1}{d^m}}.$$

Altogether the inequality we need to prove becomes $|(H\circ f^l(r))^{\frac{1}{d^n}}-H(r)^{\frac{1}{d^m}}|<\epsilon$.

For $\epsilon>0$, we let

$$N\geq\max\left\{\frac{1}{\log_{C_U}\left(1+\frac{\epsilon}{C_UH(q)}\right)},\frac{1}{\log_{C_L}\left(1+\frac{\epsilon}{C_LH(q)}\right)}\right\}.$$

This is a very generous bound, but it works. Consider the following three cases.

Case 1: $a_n>a_m$. We have previously shown that, for any f of degree d, there exists some C_U such that $H\circ f(q)\leq C_UH(q)^d$ and further by Theorem 3 that $H\circ f^l(q)\leq C_U^{1+d+d^2+\ldots+d^{l-1}}H(q)^{d^l}$. Therefore

$$a_n-a_m=(H\circ f^l(r))^{\frac{1}{d^n}}-H(r)^{\frac{1}{d^m}}\leq(C_U^{1+d+d^2+\ldots+d^{l-1}}H(r)^{d^l})^{\frac{1}{d^n}}-H(r)^{\frac{1}{d^m}},$$

which simplifies further to

$$a_n-a_m\leq(C_U^{\frac{1+d+d^2+\ldots+d^{l-1}}{d^ld^m}}-1)H(r)^{\frac{1}{d^m}}.$$

Observe that by the properties of geometric series with $d\geq 2$,

$$\begin{aligned}\frac{1+d+d^2+\ldots+d^{l-1}}{d^l}&=\left(\frac{1+d+d^2+\ldots+d^{l-1}}{d^l}\right)\left(\frac{1-d}{1-d}\right)\\&=\frac{1-d^l}{d^l(1-d)}\\&=\frac{d^l-1}{d^l}\left(\frac{1}{d-1}\right)\\&\leq 1.\end{aligned}$$

Hence, $(C_U^{\frac{1+d+d^2+\ldots+d^{l-1}}{d^ld^m}}-1)H(r)^{\frac{1}{d^m}}\leq(C_U^{\frac{1}{d^m}}-1)H(r)^{\frac{1}{d^m}}$. As $m\geq N\geq\frac{2}{\log_{C_U}\left(1+\frac{\epsilon}{C_U^2H(q)}\right)}$ and $d^m\geq m$,

$$\left(C_U^{\frac{1}{d^m}}-1\right)H(r)^{d^m}\leq\left(C_U^{\log_{C_U}(1+\frac{\epsilon}{C_UH(q)})}-1\right)H(r)^{d^m}=\left(\frac{\epsilon}{C_UH(q)}\right)H(r)^{\frac{1}{d^m}}.$$

Expanding r, we find that $H(r)^{\frac{1}{d^m}}=H\left(f^m(q)\right)^{\frac{1}{d^m}}\leq C_U^{\frac{1+d+\ldots+d^{m-1}}{d^m}}H(q)^{\frac{d^m}{d^m}}\leq C_UH(q)$. It follows that $|a_n-a_m|<\left(\frac{\epsilon}{C_UH(q)}\right)(C_UH(q))=\epsilon$.

Case 2: $a_m<a_n$. This case follows in a manner quite similar to the first. We have previously shown that, for any f of degree d, there is also some C_L such that $H\circ f(q)\geq C_LH(q)^d$ and further that $H\circ f^l(q)\geq C_L^{1+d+d^2+\ldots+d^{l-1}}H(q)^{d^l}$. Therefore

$$a_m-a_n=H(r)^{\frac{1}{d^m}}-(H\circ f^l(r))^{\frac{1}{d^n}}\leq H(r)^{\frac{1}{d^m}}-(C_L^{1+d+d^2+\ldots+d^{l-1}}H(r)^{d^l})^{\frac{1}{d^n}},$$

which simplifies to

$$a_m - a_n \leq (1 - C_L^{\frac{1+d+d^2+\ldots+d^{l-1}}{d^l d^m}}) H(r)^{\frac{1}{d^m}}.$$

As we saw before,

$$\frac{1+d+d^2+\ldots+d^{l-1}}{d^l} \leq 1.$$

Hence, $(C_L^{\frac{1+d+d^2+\ldots+d^{l-1}}{d^l d^m}} - 1) H(r)^{\frac{1}{d^m}} \leq (C_L^{\frac{1}{d^m}} - 1) H(r)^{\frac{1}{d^m}}$. As $m \geq N \geq \frac{1}{\log_{C_U}\left(1+\frac{\epsilon}{C_L H(q)}\right)}$ and $d^m \geq m$,

$$\left(C_L^{\frac{1}{d^m}} - 1\right) H(r)^{d^m} \leq \left(C_L^{\log_{C_L}(1+\frac{\epsilon}{C_L H(q)})} - 1\right) H(r)^{d^m} = \left(\frac{\epsilon}{C_L H(q)}\right) H(r)^{\frac{1}{d^m}}.$$

Expanding r, we find that $H(r)^{\frac{1}{d^m}} = H\left(f^m(q)\right)^{\frac{1}{d^m}} \leq C_L^{\frac{1+d+\ldots+d^{m-1}}{d^m}} H(q)^{\frac{d^m}{d^m}} \leq C_L H(q)$. It follows that $|a_m - a_n| < \left(\frac{\epsilon}{C_L H(q)}\right)(C_L H(q)) = \epsilon$.

Case 3: $a_n = a_m$. This is the trivial case in which $|a_n - a_m| = 0 < \epsilon$. □

Canonical height exists. But what is it good for? Earlier we promised that it could measure 'how close' a point is to being preperiodic. Let's look at some examples.

Example 5.1. Consider our familiar function $f(x) = x^2 - \frac{133}{144}$.

- The preperiodic point $\frac{-11}{12}$ behaves like so:
  - At $n = 1 \to 3.464$.
  - At $n = 2 \to 1.861$.
  - At $n = 3 \to 1.364$.
  - At $n = 4 \to 1.168$.
  - At $n = 5 \to 1.080$.
  - At $n = 6 \to 1.039$.

  We see that the limit for $\frac{-11}{12}$ is approaching one, which will be the canonical height of any preperiodic point.
- $\left(H\left(f^n(\frac{1}{2})\right)\right)^{\frac{1}{d^n}}$ goes right to 12.
- $\left(H\left(f^n(\frac{15}{12})\right)\right)^{\frac{1}{d^n}}$ has more complex behavior:
  - At $n = 1 \to 6$.
  - At $n = 2 \to 4.2$.
  - At $n = 3 \to 3.6$.
  - At $n = 4 \to 3.3$.
  - At $n = 5 \to 3.1$.
  - At $n = 6 \to 3.07$.
- Looks like $\hat{H}_f(\frac{15}{12})$ might equal 3? This is to say that $\frac{15}{12}$ seems to be closer to being preperiodic than $\frac{1}{2}$.

Canonical height also displays a number of handy algebraic properties:

Theorem 8. $\hat{H}_f \circ f(q) = [\hat{H}_f(q)]^d$

Proof. By definition

$$\begin{aligned}\hat{H}_f \circ f(q) &= \lim_{n\to\infty}\left[(H\circ f^n\circ f(q))^{\frac{1}{d^n}}\right] \\ &= \lim_{n\to\infty}\left[\left(H\circ f^{n+1}(q)\right)^{\frac{d}{d^{n+1}}}\right]. \\ &= \lim_{n\to\infty}\left[\left(\left(H\circ f^{n+1}(q)\right)^{\frac{1}{d^{n+1}}}\right)^d\right].\end{aligned} \tag{13}$$

Now taking a brief detour, we can use Theorem 7 to write

$$\lim_{n\to\infty}\left[\left(H\circ f^{n+1}(q)\right)^{\frac{1}{d^{n+1}}} - (H\circ f^n(q))^{\frac{1}{d^n}}\right] = 0.$$

The algebraic limit theorem allows us to rewrite this as

$$\lim_{n\to\infty}\left[\left(H\circ f^{n+1}(q)\right)^{\frac{1}{d^{n+1}}}\right] - \lim_{n\to\infty}\left[(H\circ f^n(q))^{\frac{1}{d^n}}\right] = 0,$$

and then

$$\lim_{n\to\infty}\left[\left(H\circ f^{n+1}(q)\right)^{\frac{1}{d^{n+1}}}\right] = \lim_{n\to\infty}\left[(H\circ f^n(q))^{\frac{1}{d^n}}\right].$$

This result can then be plugged into the end of Eq. 13 and then manipulated once again using the algabraic limit theorem as shown:

$$\begin{aligned}\hat{H}_f \circ f(q) &= \lim_{n\to\infty}\left[\left(\left(H\circ f^{n+1}(q)\right)^{\frac{1}{d^{n+1}}}\right)^d\right]. \\ &= \lim_{n\to\infty}\left[\left((H\circ f^n(q))^{\frac{1}{d^n}}\right)^d\right] \\ &= \left[\lim_{n\to\infty}\left((H\circ f^n(q))^{\frac{1}{d^n}}\right)\right]^d \\ &= [\hat{H}_f(q)]^d.\end{aligned}$$

□

Similarly:

Theorem 9. *For a given $f\in\mathbb{Q}(x)$, there exist constants B_1, B_2 such that for any choice of $q\in\mathbb{Q}$,*

$$B_1 H(q) \le \widehat{H}_f(q) \le B_2 H(q).$$

Proof. From Theorem 7 and Lemma 2, we know that there exist $C_L, C_U \in \mathbb{N}$ with

$$C_L^{\frac{1+d+\ldots+d^{n-1}}{d^n}} H(q) \le H(f^n(q))^{\frac{1}{d^n}} \le C_U^{\frac{1+d+\ldots+d^{n-1}}{d^n}} H(q)$$

for any $n\in\mathbb{N}$. In almost all cases, $C_L \le 1 \le C_U$, but if this is not the case, we can redefine C_L, C_U so that both inequalities hold. By the Algebraic Limit Theorem,

$$\lim_{n\to\infty} C_L^{\frac{1+d+\ldots+d^{n-1}}{d^n}} H(q) \le \widehat{H}_f(q) \le \lim_{n\to\infty} C_U^{\frac{1+d+\ldots+d^{n-1}}{d^n}} H(q)$$

so long as the outer limits exist. Note that

$$\frac{1+d+\ldots+d^{n-1}}{d^n}\left(\frac{d-1}{d-1}\right) = \frac{d^n-1}{d^n(d-1)},$$

and it will hold that

$$\frac{d^n-1}{d^n(d-1)} < \frac{1}{d-1}.$$

As $C_L < 1 < C_U$, we can simplify our bounds as such:

$$\lim_{n\to\infty} C_L^{\frac{1}{d-1}} H(q) < \widehat{H}_f(q) < \lim_{n\to\infty} C_U^{\frac{1}{d-1}} H(q).$$

The new outer limits, not depending on n, plainly exist in the limit $n \to \infty$ and converge to their contents. Hence,

$$C_L^{\frac{1}{d-1}} H(q) < \widehat{H}_f(q) < C_U^{\frac{1}{d-1}} H(q)$$

and $C_L^{\frac{1}{d-1}} = B_1$ with $C_U^{\frac{1}{d-1}} = B_2$. □

5.2 Using the Canonical Height to Compare Points that are not Preperiodic

From Def. 9 one can recognize that for any preperiodic point q of a function f we have that $\hat{H}_f(q) = 1$. This is due to the fact that there are finitely many possible values of $H \circ f^n(q)$, so taking the repeated iterations of f will not cause the height to go towards infinity. It might be interesting to compare not preperiodic points under $\hat{H}_f$ to see which behaves most like a preperiodic point. That is to say, for a function f and two not preperiodic points r and s we can say that r behaves more like a preperiodic point than s if $|\hat{H}_f(r) - 1| < |\hat{H}_f(s) - 1|$.

Let's consider an example. Take the function $f(x) = x^2 - \frac{133}{144}$ and let $r = \frac{1}{2}$ and $s = \frac{15}{12}$. Although it is by no means trivial to compute the infinite limits in Def. 9, we can do the next best thing. That is, computing $(f^n(q))^{\frac{1}{d^n}}$ at various values of n. If we look at $(f^n(\frac{1}{2}))^{\frac{1}{d^n}}$, we see that for every value of n, $(f^n(\frac{1}{2}))^{\frac{1}{d^n}} = 12$. More interesting is the behavior of $(f^n(\frac{15}{12}))^{\frac{1}{d^n}}$, shown in Tab. 1.

Iteration	Value
$n = 0$	1.25
$n = 1$	~ 6
$n = 2$	~ 4.2
$n = 3$	~ 3.6
$n = 4$	~ 3.3
$n = 5$	~ 3.1
$n = 6$	~ 3.07

Table 1: The height of n iterations of $\frac{15}{12}$ under $f(x) = x^2 - \frac{133}{144}$ raised to the $\frac{1}{d^n}$ power where $d = 2$. This is the same process as taking the cannonical height of $\frac{15}{12}$ under f, with the exception that we are not taking the limit as $n \to \infty$ but rather finite values of n.

Although we cannot be sure, it seems from this table that $\hat{H}_f(\frac{15}{12}) = 3$ and therefore the intuition we've developed tells us that $\frac{15}{12}$ acts more like a preperiodic point that $\frac{1}{2}$.

6 Finding periodic points via reduction modulo primes

Our goal is to find an algorithm that, given a rational function $f \in \mathbb{Q}(z)$ outputs all of the periodic points of f defined over $\mathbb{Q}$. The issue is, iterating through polynomials to find periodic points quickly becomes unmanageable. Additionally, whilst we know that a rational function f has finitely many rational periodic points, if we wish to find all the rational periodic points of a rational function f, we must have an idea of when to stop looking. To handle with the dynamic behavior of iterating rational functions, we turn to reduction modulo primes to see if we can find some results that allow us to quickly determine the rational periodic points of a rational function.

6.1 Reduction Modulo Prime

To introduce the idea of reduction, we turn to a simple example.

Example 6.1. If you divide 14 by 3, you get: $14 = 3 \cdot 4 + 2$, where 3 is the divisor, 4 is the quotient and 2 is the remainder. The reduction of an integer modulo prime is its remainder in the division by the prime. Thus, the reduction of 14 modulo 3 is 2.

We shall now extend this definition of reduction modulo primes to rational functions.

Definition 10 (Reduction modulo prime)**.** *Let α be $\alpha = \frac{a}{b} \in \mathbb{Q}$ be a rational number and let p be a prime. The reduction of α modulo p is defined by:*

$$\widetilde{\alpha} = \begin{cases} \widetilde{a}\widetilde{b^{-1}} & \text{if } p \nmid b \\ \infty & \text{if } p \mid b \end{cases}$$

Now let $\phi(z) \in \mathbb{Q}$. Then, the reduction of ϕ modulo p is:

$$\widetilde{\phi}(z) = \frac{\widetilde{a_m} z^m + \widetilde{a_{m-1}} z^{m-1} + ...\widetilde{a_0}}{\widetilde{b_n} z^n + \widetilde{b_{n-1}} z^{n-1} + ...\widetilde{b_0}}$$

To further illustrate how reduction modulo primes can be applied to rational functions, we turn to the following example:

Example 6.2. Let $f(x) = \frac{x^2+3x}{6x+1}$ and let $p = 3$. Then the reduction of $f(x)$ modulo 3 =

$$\widetilde{f}(x) = \frac{\widetilde{x^2} + \widetilde{3x}}{\widetilde{6x} + \widetilde{1}} = x^2$$

When evaluating heights of rational points, it seemed as though cancellation was the big blocker of nice results. Without cancellation, applying a function to a rational number over and over would change the height of the number in a consistent way. So, when thinking about "good" reduction, we might also wish to avoid cancellation in our reduction of rational functions.

You might think that this is simple; there is cancellation in $f(x) = \frac{g(x)}{h(x)}$ (a decrease in degree of our rational function upon reduction modulo a prime) only when both leading coefficients are congruent to 0 modulo that prime, since we stipulated that the polynomials $g(x)$ and $h(x)$ share no common roots. There is something a little bit sneaker going on, however. For example, $x - 2$ and $x + 1$ share no common roots, yet when reduced modulo 3, since $-2 \equiv 1(3)$, $\frac{x-2}{x+1} \equiv 1(3)$ for all x. So how do we tell when this sneaky cancellation is happening? This is when we bring in the idea of a resultant.

6.2 Introduction of the Resultant

Definition 11. *Let $g(x) \in \mathbb{Z}[x]$ be defined as $g(x) = \sum_{i=0}^{n} a_i x^i$, and let $h(x) \in \mathbb{Q}[x]$ be defined as $h(x) = \sum_{i=0}^{m} b_i x^i$. Then, the resultant of $g(x)$ and $h(x)$ is defined as the determinant of the $n+m$ by $n+m$ matrix*

$$\begin{pmatrix} a_n & a_{n-1} & ... & a_1 & a_0 & 0 & ... & 0 \\ 0 & a_n & ... & a_1 & a_0 & 0 & ... & 0 \\ 0 & 0 & ... & a_n & a_{n-1} & ... & a_1 & a_0 \\ b_m & b_{m-1} & ... & b_1 & b_0 & 0 & ... & 0 \\ 0 & b_m & ... & b_1 & b_0 & 0 & ... & 0 \\ 0 & 0 & ... & b_m & b_{m-1} & ... & b_1 & b_0 \end{pmatrix}.$$

Similarly, if $f(x) = \frac{g(x)}{h(x)}$, we define the resultant of $f(x)$ as the resultant of $g(x)$ and $h(x)$.

Example 6.3. As an example, let us look at $f(x) = \frac{3x^2+1}{6x^2+2x}$ where $g(x) = 3x^2 + 1$ and $h(x) = 6x^2 + 2x$ Then, the resultant is computed by the following matrix:

$$Res(g,h) = det \begin{pmatrix} 3 & 0 & 1 & 0 \\ 0 & 3 & 0 & 1 \\ 6 & 2 & 0 & 0 \\ 0 & 6 & 2 & 0 \end{pmatrix} = 9$$

6.3 Properties of the Resultant

The resultant has the remarkable property that two polynomials $g(x)$ and $h(x)$ share a common root in $\mathbb{Q}(\curvearrowleft)$ if and only if their resultant is 0.

While this is a very powerful result, it is not exactly what we are looking for. It does not tell us anything about our cancellation upon reduction. In our example with $f(x) = \frac{x-2}{x+1}$, for example, we can compute the resultant of that rational function as 3, which, while nonzero, still led to cancellation upon reduction modulo 3.

This leads us to the important property that makes us care about the resultant so much: Two polynomials $g(x), h(x) \in \mathbb{Q}[x]$ share a common root modulo p if and only if the resultant of $g(x)$ and $h(x)$ is congruent to 0 modulo p.

There is a subtle nuance here: stating that $g(x), h(x)$ share a common root modulo p is different than saying that $\widetilde{g(x)}$ and $\widetilde{h(x)}$ share a common root. Rather, this is saying that, upon reducing the roots of $g(x)$ and $h(x)$ modulo p, there exists a common root of $g(x)$ and $h(x)$. One example of this is $g(x) = 2x^2 - 1$ and $h(x) = 2x^2 + x$. The resultant of these polynomials is $2 \equiv 0(2)$. Furthermore, $\widetilde{g(x)} = -1$, and $\widetilde{h(x)} = x$ upon reduction modulo 2, so the reduced polynomials do not share a root. However, one root of $g(x)$ is $\frac{1}{\sqrt{2}} \equiv \frac{1}{0}(2)$, and one root of $h(x) is \frac{-1}{2} \equiv \frac{1}{0}(2)$. So, $g(x)$ and $h(x)$ do indeed share a common root modulo 2.

7 Definition of Good/Bad Reduction

We will now use this fact about the resultant to define our notion of "good" and "bad" reduction.

Definition 12. *Let $f(x) = \frac{g(x)}{h(x)}$, and let $\widetilde{f}(x) = \frac{\widetilde{g}(x)}{\widetilde{h}(x)}$. Then, the following statements are equivalent for some prime p*

1. $\deg(f) = \deg(\widetilde{f})$
2. *$g(x)$ and $h(x)$ have no common roots modulo p*
3. *Res(f) $\not\equiv 0(p)$*

We call a prime p a prime of good reduction if these conditions are satisfied. Otherwise, we call prime p a prime of bad reduction.

A key reason to exploring the idea of reduction of rational functions is to investigate if reduction can assist us in finding all the periodic points of a rational function. Conveniently, the conditions laid out in the definition of good reduction allow us to state a very important theorem:

Theorem 10. *Let $f(x) = \frac{g(x)}{h(x)}$ be a rational function, and let $\widetilde{f} = \frac{g(x)}{h(x)}$ mod $p = \frac{\widetilde{g}(x)}{\widetilde{h}(x)}$. If p is a prime of good reduction, then $\widetilde{f^n(x)} \equiv \widetilde{f^n}(\widetilde{x})$ mod p.*

An important consequence of this theorem is that for all rational periodic points α of the rational function f, if p is a prime of good reduction, then $\widetilde{\alpha}$ is a periodic point for $\widetilde{f}$. However, our main objective of exploring the reduction of rational functions modulo primes is to use the dynamics of $\widetilde{f}$ to help us determine the periodic points of f. Thus, it is worthwhile to explore what happens to the least periods of periodic points upon reduction. Since we need good reduction in order to maintain this good structure of periodic points modulo primes, we would like to be able to find these primes of good reduction. In particular, if there are infinitely many primes of bad reduction, then there is no guarantee that we will be able to find a prime of good reduction. Fortunately, we have the following result.

7.1 Finitely Many Primes of Bad Reduction

Given a function $f(x) = \frac{g(x)}{h(x)}$ we know that a prime p of bad reduction occurs only when the resultant of the polynomials $g(x)$ and $h(x)$ is congruent to 0 modulo p. Since the resultant R is finite and an integer (since we can define $f(x)$ such that $g(x), h(x) \in \mathbb{Z}[x]$), we know that R has a unique, finite, prime factorization. Therefore, there are finitely many primes which divide R, and thus there are finitely many primes p such that $R \equiv 0(p)$.

Now that we have the reassurance that there are only finitely many primes of bad reduction that we have to weed through, we can refocus on finding periodic points using primes of good reduction.

8 Finding periodic points

The key to finding periodic points of a function is looking at the interplay between rational periodic points and periodic points modulo primes. One final important piece that we must define is the multiplier of a periodic point.

Definition 13. *Given a periodic point z with minimal period n of rational function $f(x)$, the multiplier of z is*

$$\lambda_z = (f^{\circ n})'(z).$$

To see this computation, let's look at the rational function $f(x) = x^2 - \frac{21}{16}$.

Example 8.1. Observe that:

$$f(\frac{1}{4}) = \frac{-5}{4}$$

with

$$f(\frac{-5}{4}) = \frac{1}{4}$$

Thus, $\frac{1}{4}$ is a periodic point with period 2. Now, the multiplier

$$\lambda_{\frac{1}{4}} = (f^{\circ 2})'(\frac{1}{4})$$

Since

$$(f^{\circ 2})(x) = x^4 - \frac{21}{8}x^2 + \frac{105}{256}$$

we have

$$(f^{\circ 2})'(x) = 4x^3 - \frac{21}{4}x$$

Finally, we have the multiplier

$$\lambda_{\frac{1}{4}} = (f^{\circ 2})'(\frac{1}{4}) = \frac{-5}{4}$$

Now that we have an understanding of how to compute a multiplier for a rational function and its periodic points, it begs the question: what exactly is the multiplier and why does it matter? Well, if we define the following terms

$f(x)$ is a rational function $\frac{g(x)}{h(x)}$ of degree at least 2.

z is a periodic point of f.

n is the minimal period of z.

p is a prime of good reduction for f.

$\widetilde{z}$ and $\widetilde{f}$ are the reductions of z and f modulo p.

m is the minimal period of $\widetilde{z}$.

r is the multiplicative order of $\widetilde{\lambda_z}$ modulo p.

Then we have a couple of very useful results for finding rational periodic points of f.

Theorem 11. *Defining everything as above, we have that*

$$n = m \text{ or } n = mrp^e$$

for some nonnegative integer e.

To illustrate this, let's revisit our previous example of $f(x) = x^2 - \frac{21}{16}$.

Example 8.2. To find a prime of good reduction, let's compute the resultant of

$$f(x) = x^2 - \frac{21}{16} = \frac{16x^2 - 21}{16}$$

$$Res(f) = det \begin{pmatrix} 16 & 0 & -21 & 0 \\ 0 & 16 & 0 & -21 \\ 0 & 0 & 16 & 0 \\ 0 & 0 & 0 & 16 \end{pmatrix} = 65536 = 2^{16}$$

Thus, the only prime of bad reduction is 2 and for our example, we shall choose 3 as our prime of good reduction.

Then the reduction of $\lambda_{\frac{1}{4}}$ mod 3 is

$$\lambda_{\frac{1}{4}} = f'(\frac{1}{4}) = \frac{1}{2} \equiv -1$$

This makes the multiplicative order, r to be equal to 2. Note that at $p = 3$, since $\frac{1}{4} \equiv \frac{-5}{4}$ mod 3, the cycle of $\frac{1}{4}$ is simply 1. Thus

$$n = 2 = 2 \cdot 1 \cdot 3^0 = mrp^0$$

Lastly, a corollary to this result is as follows:

Theorem 12. *Let p and q be distinct primes of good reduction for a rational function $f(x)$. Then, given a rational periodic point z of $f(x)$ with minimal period n, we know that*

$$n \leq (p^2 - p)(q^2 - q).$$

8.1 Using our results to find rational periodic points

So, how do these two theorems allow us to find rational periodic points of a generic rational function $f(x)$? Well, following this algorithm will yield all rational periodic points.

1. Compute the resultant of $f(x)$ and use it to find two distinct primes, p and q, of good reduction. Recall that a prime is a prime of good reduction if and only if it does not divide the resultant of $f(x)$.
2. Compute the orbits of the integers modulo p and q under $f(x)$. Then, find the lengths of all the periods. This will give us $\{m_{pi} \mid i \text{ is a periodic point modulo p}\}$ and $\{m_{qi} \mid i \text{ is a periodic point modulo q}\}$, the sets of potential lengths of periods modulo p and q.
3. For both p and q, use theorems 11 and 12 to get lists of potential minimal period lengths for a rational periodic point z.
4. Find the intersection of these lists
5. For each n on this list, write $f^{\circ n}(x)$ as $\frac{g(x)}{h(x)}$. Then, create the polynomial $k(x) = g(x) - xh(x)$. All periodic points of minimal period n will be roots of this polynomial. So, if there is a rational root of this polynomial, add it to the list of rational periodic points.
6. After the previous step, we now have a list of all rational periodic points of $f(x)$.

As an example to show this algorithm in action, we turn to the rational function $f(x) = x^2 + 3x - 1$.

Example 8.3. Let $f(x) = x^2 + 3x - 1$.

Then its resultant is

$$Res(f) = det \begin{pmatrix} 1 & 3 & -1 & 0 \\ 0 & 1 & 3 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = 0$$

As $f(x)$ has no primes of bad reduction, we can choose $p = 2$ and $q = 3$ as our two distinct primes of good reduction.

For $p = 2$

$$f(0) = -1 \equiv 1 \, mod \, 2$$

$$f(1) = 3 \equiv 1 \, mod \, 2$$

Thus, 1 is a point of period 1 modulo $p = 2$, making the multiplier

$$\lambda_1 = f'(1) = 5 \equiv 1 \, mod \, 2$$

which makes the multiplicative order

$$r = 1$$

This makes our list of potential minimal period lengths to be

$$\{1 \cdot 1 \cdot 2^0, 1 \cdot 1 \cdot 2^1 ...\} = \{1, 2, 4, 8, ...\}$$

Now we move on to our second prime of good reduction $q = 3$.

$$f(0) = -1 \equiv 2 \, mod \, 3$$

$$f(1) = 3 \equiv 0 \, mod \, 3$$

$$f(2) = 9 \equiv 0 \, mod \, 3$$

Thus, 0 is a point of period 2 modulo 3. This makes its multiplier to be

$$\lambda_0 = (f^{\circ 2})'(0) = 0$$

making the multiplicative order

$$r = 0$$

Thus, the only possible period is

$$\{2\}$$

with the intersection of these 2 lists being

$$\{2\}$$

Now, we can create the polynomial $k(x)$ whose roots will be the rational periodic points of minimal period 2.

$$f^{\circ 2}(x) = x^4 + 6x^3 + 10x^2 + 3x - 3 = \frac{x^4 + 6x^3 + 10x^2 + 3x - 3}{1}$$

Then

$$k(x) = x^4 + 6x^3 + 10x^2 + 3x - 3 - x(1) = x^4 + 6x^3 + 10x^2 + 2x - 3$$

which can be factored as

$$x^4 + 6x^3 + 10x^2 + 2x - 3 = (x+1)(x+3)(x^2 + 2x - 1)$$

which has rational roots

$$x = -1, -3$$

Thus, the function $f(x) = x^2 + 3x - 1$ has rational periodic points -1 and -3 with minimal periods of 2.

9 Acknowledgements

We would like to extend our dearest thanks to our comps advisor Rafe Jones for his direction, advice, and moral support. We would also like to thank the air conditioning unit under the CMC-Boliou patio for serenading us on many an autumn afternoon.

10 Bibliography

[1] Benjamin Hutz, *An Experimental Introduction to Number Theory*, Providence, Rhode Island: American Mathematical Society, 2018.