Bounding the Degree of Belyi Polynomials

Jose Rodriguez

November 12, 2011

Abstract

Belyi’s theorem states that a Riemann surface $X$, as an algebraic curve, is defined over $\mathbb{Q}$ if and only if there exists a holomorphic function $B$ taking $X$ to $\mathbb{P}^1 \mathbb{C}$ with at most three critical values $\{0, 1, \infty\}$. By restricting to the case where $X = \mathbb{P}^1 \mathbb{C}$ and our holomorphic functions are Belyi polynomials, we define a Belyi height of an algebraic number, $H(\lambda)$, to be the minimal degree of Belyi polynomials with $B(\lambda) \in \{0, 1\}$. Using the combinatorics of Newton polygons, we prove for non-zero $\lambda$ with non-zero $p$-adic valuation, the Belyi height of $\lambda$ is greater than or equal to $p$. We also give examples of algebraic numbers which show our bounds are sharp.

1 Introduction

In this paper we fix an algebraic closure of $p$-adic numbers and denote it as $\mathbb{Q}_p$. We denote an embedded algebraic closure of the rational numbers in $\mathbb{Q}_p$ as $\overline{\mathbb{Q}}_p$. A polynomial $B(x) \in \mathbb{Q}_p[x]$ is said to have a critical point at $x_i$ if its derivative $B'(x)$ vanishes at $x_i$. We say $B(x)$ has a critical value of $B(x_i)$ when $x_i$ is a critical point. A polynomial is said to be a general Belyi polynomial if its critical values are contained in $\{0, 1\}$. Since composing a general Belyi polynomial with any linear factor $(\gamma x - \alpha)$ yields another general Belyi polynomial, we normalize our set of polynomials by requiring $B(0), B(1) \in \{0, 1\}$.

Definition 1. A polynomial $B(x) \in \mathbb{Q}_p[x]$ is said to be a normalized Belyi polynomial or Belyi polynomial if $B(0), B(1) \in \{0, 1\}$ and $\{B(x_i) : B'(x_i) = 0\} \subset \{0, 1\}$.

Equivalently we note that $B(x)$ is a Belyi polynomial if $B(0), B(1) \in \{0, 1\}$, and

$$B'(x) | B(x)(1 - B(x)).$$

We call these the two Belyi conditions. With these conditions, a Belyi polynomial composed with a linear factor $(\gamma x - \alpha)$ is a Belyi polynomial if and only if $B(\gamma), B(\gamma - \alpha) \in \{0, 1\}$. For a fixed Belyi polynomial there exist finitely many linear factors we may compose with and yield a Belyi polynomial. This finiteness condition is essential to define our Belyi height with the property that there exist finitely many Belyi polynomials of a given degree.

Example 1. The simplest examples of Belyi polynomials are $f(x) = x^n, f(x) = 1 - x$, and

$$B_{a,b}(x) = b^{a-b}(b-a)^{-a} x^a (1-x)^{b-a}, \text{ where } a,b \in \mathbb{N}, \text{ and } (b-a) \geq 0.$$

The Belyi polynomial $B_{a,b}(x)$ maps $\{\frac{a}{b}, 0, 1\}$ to $\{0, 1\}$. When we compose $B_{a,b}(x)$ with certain polynomials $C(x)$ the result, $B_{a,b}(C(x))$, as fewer critical values than $C(x)$. Specifically, when $C(x)$ satisfies the first Belyi condition and has a critical value of $\frac{a}{b}$, composing with $B_{a,b}$ reduces the number of critical values.
Example 2. The Chebyshev polynomials of the first kind, $T_n(x)$, $n \geq 1$,

$$T_0(x) = 1, \quad T_1(x) = x, \quad T_{n+1}(x) = 2x \cdot T_n(x) - T_{n-1}(x)$$

have critical values contained in $\{-1, 1\}$ and $T_n(1), T_n(-1) \in \{-1, 1\}$. Therefore $\frac{1}{2}(T_n(x) + 1)$ are general Belyi polynomials and $\frac{1}{2}(T_n(2x - 1)) + 1$ are Belyi polynomials.

This example is studied in detail in [11] where the normalization of Belyi polynomials is done with respect to $\{-1, 1\}$ instead of $\{0, 1\}$.

Example 3. The composition of any two Belyi polynomials is a Belyi polynomial.

This example is a simple application of the chain rule and gives the set of Belyi polynomials a monoid structure under composition with identity, $x$. This structure has been used to study the absolute Galois group in number theory [13], [4] and dynamical systems [10].

Belyi polynomials belong to the larger set of Belyi functions. A Belyi function $f$ maps a Riemann surface $X$ to the Riemann sphere $P^1\mathbb{C}$ with critical values contained in $\{0, 1, \infty\}$. Grothendieck was drawn into this subject because of Belyi’s theorem [3], which states a Riemann surface $X$ is defined over $\overline{\mathbb{Q}}$ if and only if there exist a Belyi function mapping $X$ to $P^1\mathbb{C}$. This marked the beginning of his program on dessin d’enfants [11], which is directly related to Belyi functions due to the well-known categorical equivalence between the two.

In the case where $X = P^1\mathbb{C}$ we normalize Belyi functions by requiring the set $\{0, 1, \infty\}$ be mapped to $\{0, 1, \infty\}$. As a corollary [1] of the Riemann Existence Theorem [9] there exist finitely many normalized Belyi functions that map $P^1\mathbb{C}$ to $P^1\mathbb{C}$ of degree at most $n$, where degree is the cardinality of the pre-image of a point in $P^1\mathbb{C} \setminus \{0, 1, \infty\}$. This means there are finitely many normalized Belyi polynomials of a given degree, hence finitely many algebraic numbers mapped to zero or one by normalized Belyi polynomials of degree $d$. The question we address in this paper is the following: for fixed $\lambda \in \overline{\mathbb{Q}}$, what is the minimal degree of normalized Belyi polynomials that map $\lambda$ to zero or one? We call this minimum the Belyi height of a number and denote it as $\mathcal{H}(\lambda)$. In [7], an upper bound of $\mathcal{H}(\lambda)$ is given, in addition to bounds for the case when $X$ is an elliptic curve. In this paper we will prove a sharp lower bound on the degree. Our results follow directly from [8] and [2]. As in Beckman’s paper our result says bad reduction implies wild ramification. What this paper contributes is a proof which uses elementary combinatorial techniques and Newton polygons. We will prove that Belyi polynomials with degree less than $p$ and $B(0) = 0$ have Newton polynomials with respect to $p$ (for the remainder of the paper all Newton polynomials will be with respect to $p$) contained in the Newton polygon of $B(x) - 1$ [Theorem 1]. We then prove the Newton polygon of $B(x) - 1$ is contained in a single line segment [Theorem 2]. Using a classical lemma [Lemma 2] relating the Newton polygon of a polynomial to the $p$-adic valuation of its roots we prove:

Main Result (Theorem 3). The Belyi height of $\lambda$, $\mathcal{H}(\lambda)$, is greater than or equal to $p$ for $\lambda \neq 0$ in $\overline{\mathbb{Q}}$ with non-zero $p$-adic valuation.

We remark that it is nontrivial to show that such a height is well defined, that is, for all algebraic numbers over $\mathbb{Q}$ there exists a Belyi polynomial, which maps it to either zero or one. Given $\lambda \in \overline{\mathbb{Q}}$ Belyi provided a way to construct [11] a Belyi function, which maps $\{0, 1, \lambda, \infty\}$ to $\{0, 1, \infty\}$ by first constructing a polynomial $g_\lambda(x) \in \mathbb{Q}[x]$ having rational critical values, $g_\lambda(\lambda) \in \mathbb{Q}$, and $\{0, 1\}$ mapped to $\{0, 1\}$. We compose $g_\lambda(x)$ with a linear factor $l_1(x)$, preserving the number of critical values, so that $l_1 \circ g_\lambda(x)$ has a rational critical value $\frac{\lambda}{n}$ between zero and one. We compose $l_1(x) \circ g_\lambda(x)$ with $B_{a_1, b_1}(x)$ so $B_{a_1, b_1} \circ l_1 \circ g_\lambda(x)$ has fewer critical values than $g_\lambda(x)$ as mentioned in Example 1. Repeating this finitely many times yields a Belyi polynomial $B_{a_k, b_k} \circ l_k \circ \cdots \circ B_{a_1, b_1} \circ l_1 \circ g_\lambda(x)$ that maps $\lambda$ to a rational number $\frac{a_{k+1}}{b_{k+1}}$. We do a final iteration so that $\lambda$ is mapped to zero or one. While this algorithm gives us a way of constructing Belyi polynomials it does not provide us a way of constructing all of them.
2 Newton Polygon Factorization

We begin this section with an introduction to $p$-adic numbers, Newton polygons, and convex sets to state Lemma 1 which allows us to classify the roots of a polynomial using these objects (see [6], [12], [5] for a thorough introduction). The $p$-adic metric on $\mathbb{Q}$ is defined as:

$$|| \cdot ||_p : \mathbb{Q} \rightarrow \mathbb{R} \quad \xrightarrow{\text{p}} p^{-k}$$

where $p \nmid ab \neq 0$ and $|0|_p \equiv 0$. The completion of $\mathbb{Q}$ under this metric will be denoted as $\mathbb{Q}_p$. The algebraic closure of $\mathbb{Q}_p$ is denoted as $\overline{\mathbb{Q}}_p$ and has a $p$-adic absolute value. Thus it makes sense to talk of the $p$-adic absolute value of any algebraic number over $\mathbb{Q}$. Frequently, it will be easier to state results using the $p$-adic valuation

$$\nu : \overline{\mathbb{Q}}_p \rightarrow \mathbb{R} \cup \{\infty\}$$

where $\nu_p(0) \equiv \infty$. The $p$-adic valuation has properties induced by the $p$-adic metric:

1. $\nu_p(\lambda \lambda_2) = \nu_p(\lambda_1) + \nu_p(\lambda_2)$
2. $\nu_p(\lambda) = \infty$ if and only if $\lambda = 0$
3. $\nu_p(\lambda_1 + \lambda_2) \geq \min\{\nu_p(\lambda_1), \nu_p(\lambda_2)\}$.

The last property is induced because the $p$-adic metric is non-Archimedian meaning

$$|a + b|_p \leq \max\{|a|_p, |b|_p\}.$$

Define the valuation ring with respect to $p$ as $\mathcal{O}_p = \{\lambda \in \overline{\mathbb{Q}}_p : \nu_p(\lambda) \geq 0\}$, the elements of the field $\overline{\mathbb{Q}}_p$ with non-negative valuation. This ring has the maximal ideal $m_p = \{\lambda \in \overline{\mathbb{Q}}_p : \nu_p(\lambda) > 0\}$, the elements of $\overline{\mathbb{Q}}_p$ with positive valuation. We denote the reduction map as $\pi : \mathcal{O}_p[x] \rightarrow \mathbb{F}[x]$ where $\mathbb{F}$ is the field $\mathcal{O}_p/m_p$.

The convex hull of a set of points is the intersection of all convex sets containing the points. When we find the convex hull of finitely many points $\{(x_0, y_0), \ldots, (x_n, y_n)\} \subset \mathbb{R}^2$, the result is a point, line segment, or convex polygon described algebraically as

$$\{\left(\sum_{i=0}^{n} c_i x_i, \sum_{i=0}^{n} c_i y_i\right) \in \mathbb{R}^2 : \sum_{i=0}^{n} c_i = 1\}$$

where $c_i \geq 0$ for all $i$. Given a polynomial $f(x) = a_0 + a_1 x + \cdots + a_n x^n$ over $\mathbb{Q}$, then $\text{Conv}_p(f)$ denotes the convex hull of

$$\{(i, \nu_p(a_i)) \in \mathbb{R}^2 : a_i \neq 0\}.$$

Our notation will be that $[v_i, v_j]$ denotes the line segment connecting the points $v_i$ and $v_j$. By convention $[v_i, v_i]$ denotes the point $v_i$. When $\text{Conv}_p(f)$ is a polygon we label a subset of the polygon’s vertices counter-clockwise from the left-most, $v_0$, ending at the right-most, $v_m$. The lower boundary of $\text{Conv}_p(f)$ is the union of the $m$ line segments connecting $v_{i-1}$ to $v_i$, denoted as

$$\bigcup_{i=1}^{m} [v_{i-1}, v_i].$$

When $\text{Conv}_p(f)$ is a line segment or point, the lower boundary of $\text{Conv}_p(f)$ is $[v_0, v_1]$ or $[v_0, v_0]$ respectively.

Definition 2. The Newton polygon of a polynomial $f(x) \in \mathbb{Q}[x]$ with respect to $p$, is the lower
The Newton polygon of a polynomial is a single vertex precisely when \( f(x) \) is a monomial. When \( f(0) = 0 \) the Newton polygons of \( f(x) \) and \( f(x) - 1 \) are closely related.

**Lemma 3.** Suppose \( f(0) = 0 \) and \( \text{New}_p(f(x)) = [v_0, v_1] \cup \ldots \cup [v_{m-1}, v_m] \).

1. \( \text{New}_p(f(x) - 1) = [v_{-1}, v_j] \cup [v_j, v_{j+1}] \cup \ldots \cup [v_{m-1}, v_m] \) for some \( j, 0 \leq j \leq m \), and \( v_{-1} \) denotes the origin.
2. Let \( s_i \) denote the slope of \([v_{i-1}, v_i]\) and \( s_0 \) denote the slope of \([v_{-1}, v_j]\). Then
   \[
   s_1 < \ldots < s_j < s_0 \leq s_{j+1} < \ldots < s_m.
   \]
3. If the degree of \( f(x) \) is less than \( p \), then \( \text{New}_p(f(x)) = \text{New}_p(x \cdot f'(x)) \).

**Proof.** This follows directly from properties of convex sets and the definition of Newton polygon.

When \( f(0) = 0 \) and \( v_j \) denotes the left-most point of \( \text{New}_p(f(x)) \cap \text{New}_p(f(x) - 1) \), combinatorially, the first two parts of the lemma say: the points in \( \text{New}_p(f(x)) \cup \text{New}_p(f(x) - 1) \) to right of \( v_j \) are in \( \text{New}_p(f(x)) \cap \text{New}_p(f(x) - 1) \); the \( \text{New}_p(f(x) - 1) \) has only one segment, \([v_{-1}, v_j]\), to the left of \( v_j \); and the slope of \([v_{-1}, v_j]\) is bounded by the slopes of line segments of \( \text{New}_p(f(x)) \). The third part combinatorial means that the \( \text{New}_p(f(x)) \) is \( \text{New}_p(f(x)) \) but shifted to left one unit.

We will prove Theorem 1 and Theorem 2 by taking full advantage of the following classical lemma:

**Lemma 4.** Let \( f(x) \) be a polynomial over \( \overline{\mathbb{Q}} \) such that \( \text{New}_p(f(x)) = [v_0, v_1] \cup \ldots \cup [v_{m-1}, v_m] \). Let \( s_i \) equal the slope of \([v_{i-1}, v_i]\), and \( d_i \) equal the length of the projection of \([v_{i-1}, v_i]\) to the \( x \)-axis. Then the polynomial \( f(x) \) may be written as

\[
 f(x) = a_n x^{d_0} f_1(x) \cdots f_m(x)
\]

where \( f_i(x) \) is monic with \( d_i \) roots of valuation \( -s_i \), counting multiplicity.

**Proof.** We refer to [12], p.74.

We call this factorization of \( f(x) \) its **Newton polygon factorization** with respect to \( p \).

**Example 4.** Setting \( p \) equal to five, the Newton polygons of three polynomials, \( h_1(x), h_2(x), h_3(x) \), are shown in bold. The thin line segment is the left-most line segment of \( \text{New}_p(h_1(x) - 1) \).

The left is an example where \( \text{New}(f(x)) \not\subset \text{New}(f(x) - 1) \). The center is an example where \( \text{New}(f(x) - 1) \) is not contained in a line segment. The right is an example of a Newton polygon of a Belyi polynomial.
\[ h_1 = 5^4 \cdot x \left( x - \frac{1}{3} \right)^2 (x - 5), \quad h_2 = 5^7 x^2 \left( x - \frac{1}{3} \right) \left( x - \frac{1}{17} \right), \quad h_3 = \frac{5^5}{3^4} x^4 (1 - x), \]

\[ h_1 = 5^4 x^4 - 3^3 \cdot 5^3 x^3 + 51 \cdot 5^2 x^2 - 5^2 x, \quad h_2 = 5^7 x^4 - 6 \cdot 5^4 x^3 + 5^2 x^2, \quad h_3 = -\frac{5^5}{3^4} x^5 + \frac{5^5}{3^4} x^4 \]

### 3 Newton Polygons of Belyi Polynomials

We prove in the case where \( B(x) \) is a Belyi polynomial of degree less than \( p \) with zero as a root that \( \text{New}_p(B(x)) \subset \text{New}_p(B(x) - 1) \). We then prove \( \text{New}_p(B(x) - 1) \) is contained in a single line segment. Using these two results we are able to give a lower bound on the Belyi height.

**Theorem 1.** If \( B(x) \in \overline{\mathbb{Q}}[x] \) is a Belyi polynomial of degree less than \( p \) such that \( B(0) = 0 \), then \( \text{New}_p(B(x)) \subset \text{New}_p(B(x) - 1) \).

**Proof.** Let \( B(x) = \sum a_k x^k \). If \( B(x) \) is a monomial the result is trivial so we consider the case where \( \text{New}_p(B(x)) = [v_0, v_1] \cup \ldots \cup [v_{m-1}, v_m] \), \( m > 0 \). Using the same notation as Lemma 3, we may assume \( \text{New}_p(B(x) - 1) = [v_{-1}, v_j] \cup [v_j, v_{j+1}] \cup \ldots \cup [v_{m-1}, v_m] \) and

\[ s_1 < \ldots < s_j < s_0 \leq s_{j+1} < \ldots < s_m. \]

Our goal is to show \( v_j = v_0 \) and the result follows. Lemma 4 allows us to write

\[ B(x) = a_n x^{d_0} f_1(x) \cdots f_m(x), \]

\[ B(x) - 1 = a_n g_0(x) g_{j+1}(x) \cdots g_m(x) \]

where every root of the monic polynomials \( f_i \) and \( g_i \), \( i \neq 0 \), has valuation \( -s_i \). In addition \( \deg(g_i) = \deg(f_i) \) when \( i > j \), while each root of \( g_0 \) has valuation \( -s_0 \).

Since \( \deg(B(x)) < p \) and \( B(0) = 0 \) then \( \nu_p(a_i) = \nu_p(i \cdot a_i) \) for every \( a_i \neq 0 \). Therefore \( \text{Conv}_p(B(x)) = \text{Conv}_p(x B'(x)) \) and \( \text{New}_p(B(x)) = \text{New}_p(x B'(x)) \). Hence \( B'(x) \) may be written as

\[ B'(x) = \deg(B) a_n x^{d_0-1} h_1(x) \cdots h_m(x) \]

where \( \deg(h_i) = \deg(f_i) \) and \( h_i \) is monic with roots of valuation \( -s_i \). By the Belyi conditions,

\[ \deg(B) a_n x^{d_0-1} h_1(x) \cdots h_m(x) | a_n x^{d_0} f_1(x) \cdots f_m(x) \cdot a_n g_0(x) g_{j+1}(x) \cdots g_m(x). \]

Therefore \( h_i(x) | f_j(x) \) when \( i \leq j \). Since the degrees of the monic polynomials are also equal it follows \( f_1 = h_1 \) when \( j \geq 1 \). Taking the derivative of \( B(x) \) and substituting \( f_1(x) \) for \( h_1(x) \) we have

\[ \deg(B) a_n x^{d_0-1} f_1(x) h_2(x) \cdots h_m(x) = f'_1(x)(a_n x^{d_0} f_2(x) \cdots f_m(x)) + f_1(x)(a_n x^{d_0} f_2(x) \cdots f_m(x))' \]

and so

\[ f_1(x) | f'_1(x)(a_n x^{d_0} f_2(x) \cdots f_m(x)). \]

Because \( f_1(x) \) and \( f_i(x) \) share no common roots when \( i \neq 1 \), \( f_1(x) \) is not divisible by \( f_i(x) \), yielding a contradiction when \( j \geq 1 \). Hence \( v_0 = v_j \).

Next we show that if \( B(x) \) is a Belyi polynomial such that \( B(0) = 0 \), then \( \text{New}(B(x) - 1) \) must be a line segment, and in preparation prove two lemmas.

**Lemma 5.** Suppose \( f(x) \) is a nonzero polynomial over an algebraically closed field of characteristic zero. If \( f'(x) \) divides \( f(x)^2 \) and \( f(0) = 0 \) then \( f(x) = a_n x^d \).
Proof. Suppose \( f(x) = a_n \prod_{i=1}^{m} (x - \alpha_i)^{d_i} \) where \( \alpha_i \) are distinct. Then

\[
f'(x) = a_n \prod_{i=1}^{m} (x - \alpha_i)^{d_i - 1} g(x), \quad \text{where } g(x) = \sum_{i=1}^{m} d_i (x - \alpha_1) ... (x - \alpha_i) ... (x - \alpha_m)
\]

and \((x - \alpha_i)\) denotes omitting a term. Note that \( \deg(g(x)) = m - 1 \) and the coefficient of the leading term is \( \deg(f(x)) \). For each root \( \alpha_j \) of \( f(x) \), \( g(\alpha_j) \neq 0 \). But \( g(x) \) also divides \( f(x)^2 \) so \( g(x) \) must have degree zero and \( m = 1 \). 

The same proof holds in characteristic \( p \) if every \( d_i \) is not divisible by \( p \) and \( p \nmid \deg(f(x)) \), giving us:

**Corollary 6.** Suppose \( f(x) \) is a nonzero polynomial over an algebraically closed field of arbitrary characteristic. If \( f'(x) \) divides \( f(x)^2 \) and \( f(0) = 0 \) then \( f(x) = a_n x^d \) or \( \deg(f(x)) \geq p \).

**Lemma 7.** Given \( f(x) \in \mathbb{Q} \) of degree \( n \), \( f(0) = 0 \), \( \text{New}_p(f(x)) = [v_0, v_1] \cup ... \cup [v_{m-1}, v_m] \), and \( m > 0 \), then there exists \( \gamma \) such that \( R(x) = \frac{1}{a_n \gamma^n} f(\gamma x) \) with \( \text{New}_p(R(x)) = [w_0, w_1] \cup ... \cup [w_{m-1}, w_m] \) has \([w_{m-1}, w_m]\) contained in the \( x \)-axis.

Proof. The polynomial \( f(x) \) has Newton factorization \( f = a_n f_1(x)...f_m(x) \) with the roots of \( f_m(x) \) of least valuation. Let \( \gamma \) be a root of \( f_m(x) \). For \( R(x) := \frac{1}{a_n \gamma^n} f(\gamma x) \) the roots are of the form \( \frac{\gamma}{\gamma_i} \) where \( \gamma_i \) is a root of \( f_i(x) \). Therefore the valuation of a root of \( R(x) \) equals \( \nu_p(\gamma_i) - \nu_p(\gamma) \geq 0 \). It follows the slopes of \([w_{i-1}, w_i]\) of \( \text{New}_p(R(x)) \) are less than zero if \( i \neq m \) and equal to zero when \( i = m \). Since \( R(x) \) is monic and \([w_{m-1}, w_m]\) has slope zero then \( w_{m-1} \) and \( w_m \) are in the \( x \)-axis. 

**Theorem 2.** If \( B(x) \) is a Belyi polynomial of degree less than \( p \) with \( B(0) = 0 \), then \( \text{New}_p(B - 1) \) is a line segment.

Proof. If \( B(x) \) is a monomial the result is trivial. Now suppose \( B(x) = a_1 x + a_2 x^2 + ... + a_n x^n \) with Newton factorization \( a_n x^d_1 f_1(x)...f_m(x) \). By Lemma 3 using the already defined notation from Theorem 1 we see for \( m \geq 1 \)

\[
\text{New}_p(B(x)) = [v_0, v_1] \cup ... \cup [v_{m-1}, v_m],
\]

\[
\text{New}(B(x) - 1) = [v_1, v_2] \cup ... \cup [v_{m-1}, v_m].
\]

By Theorem 1, \( j = 0 \) so the slopes satisfy

\[
s_0 \leq s_1 < s_2 < ... < s_m.
\]

To prove the theorem we must show that \( m = 1 \) and \( s_0 = s_1 \). Let \( \gamma \) be a root of \( B(x) \) with least valuation. This is a root of \( f_m(x) \) and \( \nu_p(\gamma) = -s_m \). Let \( R(x) := \frac{1}{a_n \gamma^n} B(\gamma x) \). Then the Newton factorization of \( R(x) \) is

\[
R(x) = \frac{(\gamma x)^{d_0}}{\gamma^{d_0}} f_1(\gamma x) \frac{f_m(\gamma x)}{\gamma^{d_m}}
\]

By Lemma 7

\[
\text{New}_p(R(x)) = [w_0, w_1] \cup ... \cup [w_{m-1}, w_m]
\]

has \([w_{m-1}, w_m]\) contained in the \( x \)-axis. So \( R(x) \) has \( d_m \) roots of valuation zero. Since the slope of each \([w_{i-1}, w_i]\), is non-positive it follows \( \text{New}_p(R(x)) \) is contained in the upper half plane, hence \( R(x) \) is in \( \mathcal{O}_p[x] \) as is each of its factors.

As in Theorem 1, \( B(x) - 1 \) has a Newton factorization

\[
B(x) - 1 = a_n g_0(x) g_1(x) ... g_m(x)
\]
where \( \text{deg } g_i = d_i \) and \( g_i \) has roots of valuation \(-s_i\). So

\[
R(x) - \frac{1}{a_n \gamma^n} = \frac{1}{a_n \gamma^n} (B(\gamma x) - 1) = \frac{g_0(\gamma x)}{\gamma^{d_0}} g_1(\gamma x) \cdots \frac{g_m(\gamma x)}{\gamma^{d_m}} \in \mathcal{O}_p[x]
\]

and \( R(x) - \frac{1}{a_n \gamma^n} \) also has \( d_i \) roots of valuation \((s_i - s_m)\).

Since \( R(x) - \frac{1}{a_n \gamma^n} \) is monic, the product of its roots is \( \frac{(-1)^{n+1}}{a_n \gamma^n} \in \mathcal{O}_p \), and

\[
0 \leq \nu_p\left(\frac{(-1)^{n+1}}{a_n \gamma^n}\right) = \nu_p\left(\frac{-1}{a_n \gamma^n}\right) = -d_0(s_0 - s_m) - d_1(s_1 - s_m) - \cdots - d_m(s_m - s_m).
\]

With this, we see \( \nu_p\left(\frac{-1}{a_n \gamma^n}\right) = 0 \) if and only if \((s_i - s_m) = 0 \). So in the case where \( \nu_p\left(\frac{-1}{a_n \gamma^n}\right) = 0 \) it follows \( m \) is necessarily one and \( s_0 = s_1 \).

We conclude the proof by using the reduction map and Corollary 6 to show that the remaining case where \( \nu_p\left(\frac{-1}{a_n \gamma^n}\right) > 0 \) leads to a contradiction. Since \( \deg R(x) < p \) and \( R(0) = 0 \) then \( \text{New}_p(R(x)) = \text{New}_p(x \cdot R'(x)) \). So \( R'(x) \) also has leading coefficient and \( d_m \) roots of valuation zero. In particular \( R(x), R(x) - \frac{1}{a_n \gamma^n}, R'(x) \), and each of their factors are in \( \mathcal{O}_p[x] \) and \( \pi(R(x)), \pi(R(x) - \frac{1}{a_n \gamma^n}), \pi(R'(x)) \) are nonzero. The Belyi condition \( B(x) \mid B(x)B(x-1) \) imply \( R'(x) \mid R(x)R(x) - \frac{1}{a_n \gamma^n} \). So \( \pi(R'(x)) \mid \pi(R(x)\pi(R(x) - \frac{1}{a_n \gamma^n}) \). But when \( \nu_p\left(\frac{-1}{a_n \gamma^n}\right) > 0 \), \( \pi(R(x)) = \pi(R(x) - \frac{1}{a_n \gamma^n}) \). We can then apply Corollary 7 which says \( \pi(R(x)) \) has no nonzero roots. But this contradicts the fact that \( R(x) \) has \( d_m \) roots of valuation zero.

Theorem 3. The Belyi height of \( \lambda, \mathcal{H}(\lambda) \), is greater than or equal to \( p \) for \( \lambda \neq 0 \) in \( \overline{\mathbb{Q}} \) with non-zero \( p \)-adic valuation.

Proof. If \( B(0) = 1 \), then consider the Belyi polynomial \( 1 - B(x) \), so without loss of generality we may assume \( B(0) = 0 \). If \( \text{deg}(B(x)) < p \), then by Theorem 2 \( \text{New}(B(x)) \) and \( \text{New}(B(x) - 1) \) imply \( R'(x) \mid R(x)R(x) - \frac{1}{a_n \gamma^n} \). So \( \pi(R'(x)) \mid \pi(R(x))\pi(R(x) - \frac{1}{a_n \gamma^n}) \). But when \( \nu_p\left(\frac{-1}{a_n \gamma^n}\right) > 0 \), \( \pi(R(x)) = \pi(R(x) - \frac{1}{a_n \gamma^n}) \). We can then apply Corollary 7 which says \( \pi(R(x)) \) has no nonzero roots. But this contradicts the fact that \( R(x) \) has \( d_m \) roots of valuation zero.

With this theorem, we know for every Belyi polynomial with rational number \( \frac{a}{b} \) in lowest terms as a root will have degree greater than or equal to every prime \( p \) that divides \( ab \). The well-known Belyi polynomial from Example 1 \( B_{1,1}(x) \) has as its critical points \( \{\frac{1}{x}, 0, 1\} \). Therefore \( B_{1,1}(\frac{1}{x}) \) is a normalized Belyi polynomial mapping \( p \) to zero showing our bound is sharp. However, in general, it is not true \( \mathcal{H}(a) \geq a \), for \( a \in \mathbb{Z} \), as the following example shows.

Example 5. If we consider the Belyi polynomial \( B(x) = -\frac{1}{4}(x-1)^2(x-4) \), then \( B(4) = 0 \) and \( \mathcal{H}(4) \leq 3 \). By Theorem 3 \( \mathcal{H}(4) \geq 2 \). A direct calculation by solving a quadratic shows that \( \mathcal{H}(4) \neq 2 \), so it follows \( \mathcal{H}(4) = 3 \).

We end with a few open questions. First, how can one express \( \mathcal{H} : \overline{\mathbb{Q}} \to \mathbb{R}^+ \) in a closed form? By Example 1 we know this is not a simple function such as \( \mathcal{H}(a) = a \) when we restrict \( \mathcal{H} \) to the natural numbers. Second, when is \( \mathcal{H}(ab) \geq \mathcal{H}(a) + \mathcal{H}(b) \)? By Theorem 3 we know \( \mathcal{H}(pq) \geq \max\{\mathcal{H}(p), \mathcal{H}(q)\} \) for primes \( p \) and \( q \). Third, we ask for fixed \( h \in \mathbb{R}^+ \) how many distinct \( \lambda \) satisfy the inequality \( \mathcal{H}(\lambda) \leq h \) in \( \mathbb{R}^+ \)? In addition, can we adjust the definition of Belyi height so that the number of such \( \lambda \) grows on the order of a polynomial as we vary \( h \). Finally, does there exist a unique Belyi polynomial of degree equal to \( \mathcal{H}(\lambda) \) with \( \lambda \) as one of its roots? If not can we classify such polynomials, and do they have the same Newton polygon?

I would like to especially thank Eric Katz for supervising this research and the Ronald E. McNair Postbaccalaureate Achievement Program for funding this project.
References

[1] Ingrid Bauer, Fabrizio Catanese, and Fritz Grunewald. Chebycheff and Belyi polynomials, dessins d’enfants, Beauville surfaces and group theory. *Mediterr. J. Math.*, 3(2):121–146, 2006.

[2] Sybilla Beckmann. Ramified primes in the field of moduli of branched coverings of curves. *J. Algebra*, 125(1):236–255, 1989.

[3] G. V. Bely˘ı. On extensions of the maximal cyclotomic field having a given classical Galois group. *J. Reine Angew. Math.*, 341:147–156, 1983.

[4] Jordan S. Ellenberg. Galois invariants of dessins d’enfants. In *Arithmetic fundamental groups and noncommutative algebra* (Berkeley, CA, 1999), volume 70 of *Proc. Sympos. Pure Math.*, pages 27–42. Amer. Math. Soc., Providence, RI, 2002.

[5] Antonio J. Engler and Alexander Prestel. *Valued fields*. Springer Monographs in Mathematics. Springer-Verlag, Berlin, 2005.

[6] Fernando Q. Gouvêa. *p-adic numbers: An introduction*. Universitext. Springer-Verlag, Berlin, second edition, 1997.

[7] Lily S. Khadjavi. An effective version of Belyi’s theorem. *J. Number Theory*, 96(1):22–47, 2002.

[8] Zapponi Leaonardo. On the degree of a belyi number field. *Arxiv*, 2008.

[9] Rick Miranda. *Algebraic curves and Riemann surfaces*, volume 5 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 1995.

[10] Kevin M. Pilgrim. Dessins d’enfants and Hubbard trees. *Ann. Sci. École Norm. Sup. (4)*, 33(5):671–693, 2000.

[11] Leila Schneps, editor. *The Grothendieck theory of dessins d’enfants*, volume 200 of *London Mathematical Society Lecture Note Series*. Cambridge University Press, Cambridge, 1994. Papers from the Conference on Dessins d’Enfant held in Luminy, April 19–24, 1993.

[12] Edwin Weiss. *Algebraic number theory*. Dover Publications Inc., Mineola, NY, 1998. Reprint of the 1963 original.

[13] Melanie Matchett Wood. Belyi-extending maps and the Galois action on dessins d’enfants. *Publ. Res. Inst. Math. Sci.*, 42(3):721–737, 2006.