Elusive Zeros Under Newton's Method

Trevor M. O'Brien

May 31, 2005

Abstract

Given that Newton's method is an iterative process, it is natural to study it as a discrete dynamical system. Of particular interest are the various open sets of initial seeds that fail to converge to a root under Newton's method. Thus we examine certain "bad" polynomials that contain extraneous, attracting periodic cycles.

In particular, we chose to examine Newton's method applied to a particular family of fourth degree polynomials that rely on only one parameter value. These are polynomials of the form:

$$p_\lambda(z) = (z+1)(z-1)(z-\lambda)(z-\bar{\lambda})$$

where $\lambda \in \mathbb{C}$. We extensively analyze the parameter plane for this family of polynomials in cases where λ is both real and purely imaginary. More specifically, we have developed and implemented computer programs to locate λ values for which Newton's method fails on an open, often relatively large, set of initial conditions. In doing so, we have discovered some rather surprising dynamical figures in the λ-parameter plane, including Mandelbrot-like sets, tricorns, and swallowtails. Through symmetry and by restricting to the imaginary axis, we have uncovered certain analytic and numerical evidence that aids in explaining the existence of such figures.

Acknowledgments

First and foremost, I would like to thank my thesis advisor, Professor Gareth Roberts, without whom this thesis would not exist. Professor Roberts is responsible for suggesting this problem, and I owe him a great deal for offering so much of his time and efforts in providing me with valuable insight and advice. I would also like to thank Professor Damiano for taking the time out of his busy schedule to read through and scrutinize this thesis. His help and support are greatly appreciated. I would also like to thank the entire Mathematics and Computer Science Department at Holy Cross for providing an enjoyable and supportive atmosphere at all times. I cannot imagine working in a better environment. Lastly, I would like to thank the Council on Undergraduate Research (CUR), as this project was partially supported by a CUR Undergraduate Summer Research Fellowship in Science and Mathematics.

Contents

1 Introduction to Newton's Method

Newton's method for approximating the zeros of a function is one of the most well known and widely used numerical algorithms. Given that Newton's method is an iterative process, it is natural that we study it as a discrete dynamical system. When dealing with a smooth function of one real variable, the implementation of this algorithm is a fairly straightforward application of the derivative.

Suppose we are trying to find the roots of a smooth function $p(x)$. For the purposes of this work, p will usually be considered a polynomial. Given an initial seed x_0, we construct the tangent line approximation for $p(x)$, and record the point at which this line crosses the x-axis. We label this point x_1, and begin the process again with x_1 as our new initial seed.

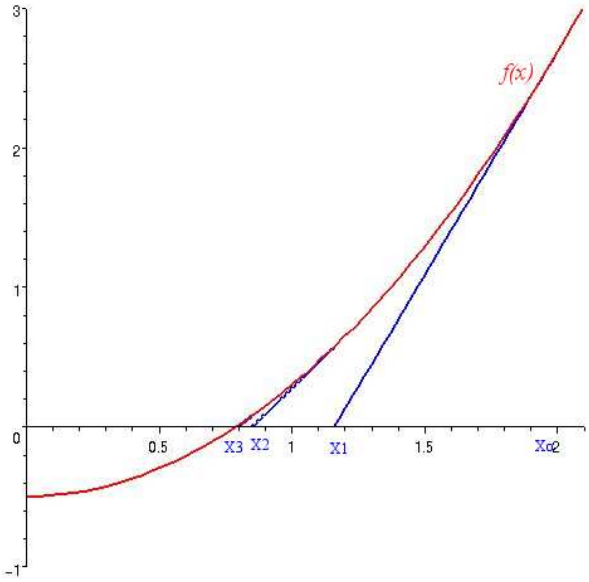

Figure 1.1: A simple example of Newton's method applied to $f(x)$.

This process can be continued until a sufficient approximation for the root of $p(x)$ is found. It is defined iteratively as follows

$$x_{i+1} = x_i - \frac{p(x_i)}{p'(x_i)} .$$

The success of this method, however, is problematic. Given the construction of the tangent line approximation, if the initial guess x_0 is a critical point of $p(x)$, Newton's method fails to converge. For example, if $p(x) = x^2 - 1$, and $x_0 = 0$, then x_1 is undefined, and we are unable to continue Newton's method further. Thus, Newton's method failed to converge to either of the roots of p, 1 and -1.

This does not appear to bode well for the reliability of Newton's method. However, we are always guaranteed that Newton's method will converge on some neighborhood of the roots in question. The size of this neighborhood relies on certain factors about the function

to which we are applying Newton's method and its domain. Recent work has shown not only that Newton's method will always converge, but that we can actually locate and iterate extremely small sets of initial guesses in order to find all roots under Newton's method. More information on this can be found in [8].

Nevertheless, while we are guaranteed that there exists a neighborhood about each root on which Newton's method converges and is successful, the main focus in this work will actually be on the *failure* of this method.

Beyond mappings of one real variable, one can apply Newton's method to functions of complex variables. This can be accomplished by iterating the Newton Map

$$N_p(z) = z - \frac{p(z)}{p'(z)}$$

If p is a polynomial, then N_p is a rational map over the Riemann Sphere, $\overline{\mathbb{C}}$. As with Newton's method applied to polynomials of one real variable, we are not assured of the convergence of N_p. For example, consider the polynomial $p(z) = z^4 - 6z^2 - 11$. In this case, for the initial guess of $z_0 = 1$, the solutions of N_p oscillate between 1 and -1. It turns out that 1 and -1 lie in a particular region of the complex plane that consists of other initial guesses that oscillate in the same manner, without converging to a root. Furthermore, we also see that $N_p'(\pm 1) = 0$. In this example it is relatively easy for one to work through the calculations to see that 1 and -1 oscillate back and forth between each other under iteration of N_p. Locating *all* initial seeds that fail to converge to a root, however, is not nearly as obvious.

The main theme of this paper will be to locate and examine the open sets of initial seeds on which Newton's method does not converge when applied to a particular family of fourth degree polynomials. These are polynomials of the form:

$$p_\lambda(z) = (z+1)(z-1)(z-\lambda)(z-\overline{\lambda}).$$

In studying the λ-parameter plane for this family of polynomials, we uncover some noteworthy dynamical behavior. With the aid of computer software, we numerically locate Mandelbrot-like sets, tricorns, and other dynamical figures that have previously been discovered in the case of general cubic maps. Making use of the abundance of symmetry that exists within our chosen family of polynomials, we are able to significantly reduce the parameter plane and better explain the existence of the aforementioned figures. We make note of the strong correlation between our Newton map and a general cubic map, given that both have two free critical points. Then, in accordance with Milnor's work on cubics, found in [11], we find the same type of dynamical behavior in our study of the λ-parameter plane. While a direct conjugacy or correlation between our cubic-*like* map and the general cubic case is lacking, we have strong numerical evidence that suggests that such a correlation does exist. Furthermore, using Milnor's classification of dynamical components, we offer a conjecture as to where we should expect certain figures to exist in the parameter plane.

2 Introduction to Complex Dynamical Systems

In order to best understand Newton's Method as a dynamical system of one complex variable, we will first review some standard facts and results from complex dynamics. Historically it should be noted that the field of complex dynamics was pioneered by the work of G. Julia and P. Fatou in the 1920's. For a more extensive history of Fatou and Julia, and the entire field of complex dynamics, see [1, 10].

Consider some analytic function $f : \overline{\mathbb{C}} \to \overline{\mathbb{C}}$. As with dynamical systems of real variables, the main focus of complex dynamical systems is understanding and explaining the behavior of points under iteration of f. In more basic terms, we are interested in following and describing where various points *go* under iteration of f. In order to do so, we must first familiarize ourselves with some notation.

Given f as above, $f^2(z) = f(f(z))$ denotes the second iterate of f. Following in this fashion, $f^3(z)$ is the third iterate of f, and, in general, we write the n-fold composition of f with itself as $f^n(z)$. Given some $z_0 \in \overline{\mathbb{C}}$ we define the **orbit** of z_0 to be the sequence of points: $\{z_0, z_1 = f(z_0), z_2 = f(z_1) = f^2(z_0), \ldots, z_n = f^n(z_0), \ldots\} = \{f^n(z_0)\}_{n=1}^{\infty}$. We refer to z_0 as the **initial seed** of the orbit.

A point z is said to be a **periodic point** if $f^n(\mathrm{z}) = \mathrm{z}$ for some $n \in \mathbb{N}$, and the smallest such n is known as the **period** of z. A periodic point of period 1 is known as a **fixed point**.

A periodic point z with period n is said to be **attracting** if $|(f^n)'(z)| < 1$ and **repelling** if $|(f^n)'(z)| > 1$. The dynamical behavior near these attracting and repelling points is exactly as one what would expect from their respective titles. Points nearby an attracting periodic point are *attracted* to that periodic point under iteration. Similarly, points nearby a repelling periodic points are *repelled* away from it under iteration. Furthermore, a periodic point z with period n is said to be **indifferent** or **neutral** if $|(f^n)'(z)| = 1$. These indifferent or neutral points may be weakly attracting, repelling, or neither. It should be noted that the same holds for the entire orbit of z, so we may refer to entire orbits as attracting or repelling. Lastly, there is a special type of attracting periodic points that will be of particular interest to us later in this paper. A periodic point z satisfying $(f^n)'(z) = 0$ is called **superattracting**, a title that corresponds to the rate at which nearby points converge.

Definition 2.1 Let α be an attracting periodic point of period n for the analytic function $f(z)$. Then the set

$$\{z : f^{mn}(z) \to \alpha \text{ as } m \to \infty,\ m \in \mathbb{N}\}$$

is defined to be the **basin of attraction** of α. The connected component of this set that contains α is referred to as the **immediate attractive basin**, denoted $B(\alpha)$.

Definition 2.2 Let U be an open subset of $\mathbb{C}$, and let $g_k : U \to \mathbb{C}$ be a family of complex analytic functions. The family $\{g_k\}$ is said to be **normal on U** if every sequence of functions selected from $\{g_k\}$ has a subsequence that converges uniformly on every compact subset of U, either to a bounded analytic function or to ∞. The family $\{g_k\}$ is **normal at the point w** of U if there is some open subset V of U containing w such that $\{g_k\}$ is a normal family on V.

This notion of a normal family leads us to the following definition, which is vital to the study of complex dynamics.

Definition 2.3 Given an analytic function $f : \overline{\mathbb{C}} \to \overline{\mathbb{C}}$, the **Julia set** of f, denoted J_f, is the set of all points $z \in \overline{\mathbb{C}}$ such that the family $\{f_k\}$ is *not* normal at z. The **Fatou set** is the complement of the Julia set.

For the purposes of this work, we will assume that we are dealing with the Julia set of a rational function of degree at least two. It is not necessarily the case that all of the properties and characteristics of the Julia set that we mention will hold in general. It is straight-forward to see that repelling periodic points must lie in the Julia set. Fatou and Julia proved even more, showing that for rational maps, the Julia set of f is precisely the closure of the repelling fixed points of f [6, 9],

$$J_f = \{\overline{\text{repelling periodic points of f}}\}$$

We may conclude from the previous definitions, that given any neighborhood about some point $z \in J_f$, there is at least one subsequence of $\{f_k\}$ that does not converge uniformly on some compact subset of $U \in \overline{\mathbb{C}}$. In fact, given the following theorem, we can actually say quite a bit more.

Theorem 2.4 (Montel's Theorem) *Let $\{g_k\}$ be a family of complex analytic functions on an open domain* U. *If $\{g_k\}$ is not a normal family, then for all $w \in \overline{\mathbb{C}}$, with at most two exceptions, $g_k(z) = w$ for some $z \in U$ and some k.* (See [10].)

From Montel's Theorem, we see that any neighborhood about a point in the Julia Set is smeared over the entire complex plane under iteration of the given function (with at most two exceptions). This property is known as the **supersensitivity** of J_f.

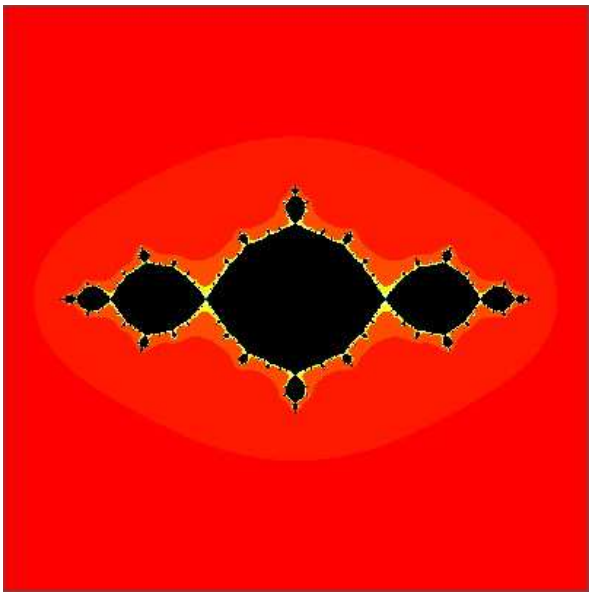

Figure 2.1: The Julia set for $z^2 - 1$.

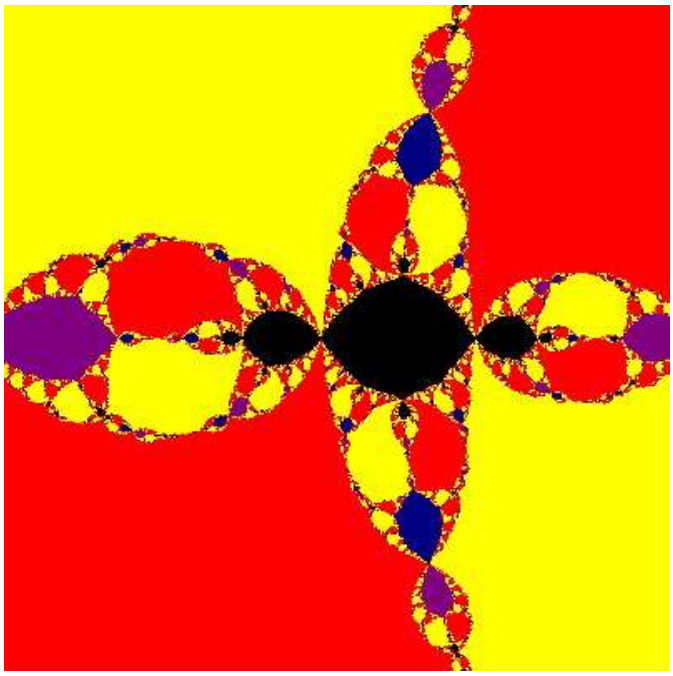

Figure 2.2: A sneak preview of a component of the Julia set for the family of polynomials we will be working with later. Note the similarity to Figure 2.1.

The dynamical behavior on the Fatou set is often considered to be somewhat boring, as it contains all of the attractive periodic cycles and their basins of attraction, which behave in an orderly fashion. On the Julia set, however, the dynamical behavior is far from orderly.

Definition 2.5 (Devaney) The function f: E $\to$ E is **chaotic** if
a) the periodic points of f are dense in D,
b) f is topologically transitive, and
c) f exhibits sensitive dependence on initial conditions. (See [4].)

To best understand this definition, the terms *topologically transitive* and *sensitive dependence on initial conditions* must be explained more thoroughly.

Saying that f is topologically transitive is a more precise way of saying that it is "well-mixed." That is to say, under iteration of f, points in D are smeared over all of D in an equally distributed manner. A more rigorous definition is as follows.

Definition 2.6 f is **topologically transitive on f** if for any two points $x, y \in D$, and any $\epsilon > 0$, there is some $z \in D$ and some $n \in \mathbb{N}$ such that $|z - x| < \epsilon$ and $|f^n(z) - y| < \epsilon$.

Now, when we say f exhibits sensitive dependence on initial conditions, we simply mean that under iteration of f, points that begin nearby one another do not remain nearby.

Definition 2.7 The function $f : D \to D$ exhibits **sensitive dependence on initial conditions** if there exists a $\delta > 0$ such that for any x $\in$ D and any $\epsilon > 0$, there is a $y \in$ D and some $n \in \mathbb{N}$ such that $|x - y| < \epsilon$ and $|f^n(x) - f^n(y)| > \delta$. For more on the criteria for chaotic maps, see [4] and [7].

Proposition 2.8 *i. The Julia Set of a rational function is always closed and non-empty.*
ii. J_f is invariant under f, and f^{-1}. That is to say, J_f is forward and backward invariant: $J_f = f(J) = f^{-1}(J)$.
iii. J_f is a perfect set (closed with no isolated points), and is therefore uncountable.
iv. Dynamical behavior on J_f is chaotic.
v. J_f is supersensitive. (See [1].)

Since we are applying N to a family of polynomials, N_p will always be a rational map. For this reason, throughout the rest of this work we shall restrict our attention to rational maps.

Theorem 2.9 *Every immediate attractive basin of a rational map contains a critical point of that rational map.* (See [10].)

Example 2.10 Let's consider the squaring map, $Q_0 : \overline{\mathbb{C}} \to \overline{\mathbb{C}}$, $Q_0(z) = z^2$.

First off, we note that any complex number z can be written in the form $z = re^{i\theta}$, where $|z| = r$. Then, $z^2 = r^2 e^{2i\theta}$. It is plain to see that under the squaring map, the origin is a fixed point. Similarly, since $Q_0(\infty) = \infty$, we see that ∞ is also a fixed point under Q_0. Moreover, it is plain to see that points inside the unit circle, where $r < 1$, spiral in counterclockwise to the origin under iteration of Q_0, while points outside the unit circle, where $r > 1$, spiral out to ∞. Thus we see that $J_{Q_0} = \{z : |z| = 1\}$, the unit circle in $\overline{\mathbb{C}}$. □

It should be noted that the terminology for this mapping, Q_0, is significant in the sense that it represents a particular parameter value for functions of the form $Q_c(z) = z^2 + c$. While this parameter value merely translates the mapping of z^2 by some complex number c, the behavior that results from varying c is extremely interesting, and certainly non-trivial. The family Q_c has been studied extensively by Douady and Hubbard, who were able to prove a great deal about its dynamical behavior, which has become fundamental in the study of complex dynamics. For our purposes, we will only discuss the family Q_c briefly. A more thorough analysis of this family of quadratics can be found in [5].

With respect to the family of functions $Q_c = z^2 + c$, we have only the critical point $z = 0$. By our previous theorem, if there exists an attracting periodic point w, then $0 \in B(w)$. Therefore, the orbit of 0 is bounded, as it cannot leave its respective attractive basin, $A(w)$.

Remark 2.11 J_f is either connected, or totally disconnected (a Cantor set or fractal dust). □

This leads to another important definition.

Definition 2.12 The **Mandelbrot Set**, M, is the set of parameters c for which the Julia set of Q_c is connected.

$$M = \{\text{c} \in \overline{\mathbb{C}} : J_{Q_c} \text{ is connected}\}.$$

See Figure 2.3.

Theorem 2.13

$$\begin{aligned} M &= \{c \in \overline{\mathbb{C}} : \{Q_c^k(0)\}_{k \geq 1} \textit{ is bounded}\} \\ &= \{c \in \overline{\mathbb{C}} : Q_c^k(0) \nrightarrow \infty \textit{ as } k \to \infty\} \end{aligned}$$

This is sometimes referred to as the **Fundamental Dichotomy** *of the Mandelbrot Set.* (See [4].)

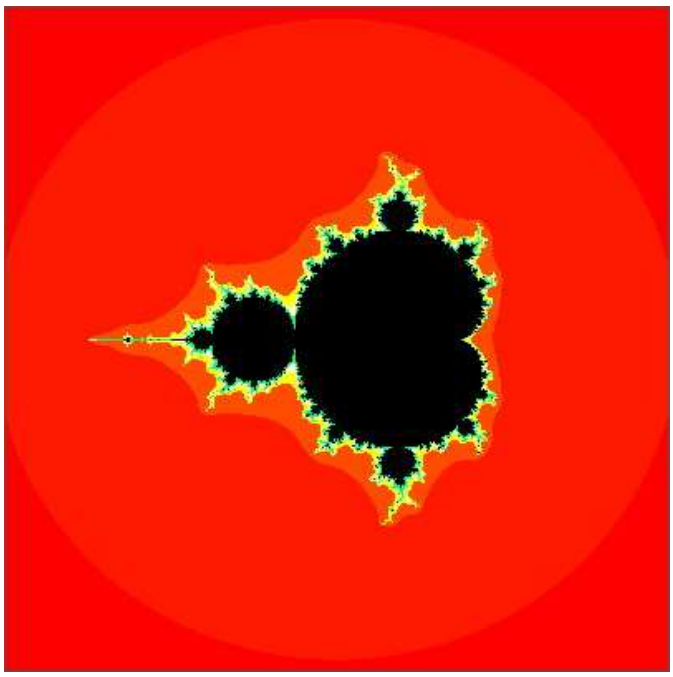

Figure 2.3: The c-parameter plane for the family $Q_c = z^2 + c$. The black region represents parameter values for which the orbit of the critical point $z = 0$ is bounded. This is known as the Mandelbrot set.

There is one final definition that must be given which will prove useful later in reducing the space on which we must study our given dynamical system.

Definition 2.14 Let $f : A \to A$ and $g : B \to B$ be two maps. Then f is **topologically conjugate** to g if there is a homeomorphism $\tau : A \to B$ such that $\tau \circ f = g \circ \tau$. In this case, τ is called a **topological conjugacy**.

It is important to note that topological conjugacy preserves dynamical behavior. That is to say, given a topological conjugacy τ, fixed points, periodic points and attractive basins of f are sent under τ to fixed points, periodic points and attractive basins of g.

3 Newton's Method as a Complex Dynamical System

Following our brief introduction to complex dynamics, we will now shift our attention back to Newton's method. Given a polynomial $p(z)$, we define the *Newton map* to be

$$N_p(z) = z - \frac{p(z)}{p'(z)}.$$

Proposition 3.1 *i. If $p(z)$ has n distinct roots, then N_p is a degree n rational map.*
ii. If $q(z) = p(az+b)$, then N_p is conjugate to N_q by the map $\tau(z) = az+b$.
iii. The roots of $p(z)$ are fixed points of N_p.
iv. The point at ∞ is a fixed point, and since $N_p'(\infty) = \frac{d}{d-1}$, it is repelling.
v. The derivative of N_p is

$$N_p'(z) = \frac{p(z)p''(z)}{(p'(z))^2}$$

Thus, simple roots of $p(z)$ are superattracting fixed points of N_p.
vi. Multiple roots are attracting fixed points, though they are not superattracting. (See [2].)

For a root α with multiplicity m, $N_p' = \frac{m-1}{m}$, which suggests slow convergence for roots with high multiplicity.

It is also important to note that inflection points of $p(z)$ are critical points of $N_p(z)$. We refer to an inflection point of $p(z)$ as a *free critical point*, a point where the second derivative $p''(z)$ vanishes. Since the only other critical points of N_p are the roots of $p(z)$, which are fixed points of N_p, we will attempt to uncover the dynamical behavior of N_p by examining its free critical points.

It should be noted that much work has been done involving the study of Newton's method applied to quadratic and cubic polynomials. In both cases, Newton's method generally follows the **Nearest Root Principle**, which is to say that initial seeds will typically converge to the root to which they are nearest. This fails to hold true in certain cubic cases, though it is still useful in pointing us in the direction of where to look for sets of bad initial seeds. In the quadratic case, however, it turns out to be perfectly accurate in classifying the dynamical behavior of Newton's method.

Given a quadratic polynomial with two distinct roots, r_1 and r_2, the only region on which Newton's method fails is the perpendicular bisector of the line segment joining r_1 to r_2, which is left invariant. This makes sense intuitively, since points along this line are equidistant from both r_1 and r_2. Thus, Newton's method seems to be confused as to which root it should converge to. More importantly, the perpendicular bisector is full of repelling periodic cycles, exactly as the unit circle was under the squaring map. Thus, Newton's method applied to the quadratic map is chaotic along this perpendicular bisector. In other words, the perpendicular bisector of the line segment joining r_1 to r_2 is actually the Julia set for N_p.

The only other type of quadratic polynomial left to discuss are those with one root of multiplicity two, i.e. the squaring map. It turns out that there is a distinct relationship between quadratic mappings with two distinct roots and the squaring map, Q_0. This relationship can be seen through the topological conjugacy in the following example.

Example 3.2 Suppose $p(z)$ is a quadratic polynomial with distinct roots r_1 and r_2. Then $N_p(z)$ is topologically conjugate to the squaring map, $Q_0(z) = z^2$, under the conjugacy

$$h(z) = \frac{z-r_1}{z-r_2}.$$

The perpendicular bisector of the line segment joining r_1 to r_2 is the set $\mathrm{B} = \{z : |z-r_1| = |z-r_2|\}$. Then, we see that for any $z_0 \in \mathrm{B}$, $|h(z_0)| = 1$. Thus, under h, B is mapped to the

unit circle in $\overline{\mathbb{C}}$.

Since Julia sets are preserved under topological conjugacy, this confirms that B, the set of points on the perpendicular bisector of the line segment joining r_1 to r_2 is the Julia set of N_p, since we know that the unit circle is the Julia set for Q_0. More specifically, the dynamics on the perpendicular bisector are conjugate to angle doubling on the unit circle, which we know is a chaotic map. □

Then, following the Nearest Root Principle, points on the r_1 side of the Julia set converge to r_1, and points on the opposing side converge to r_2. From a dynamical systems point of view, this is about all we can say, and we care about, with respect to Newton's method applied to quadratic polynomials.

As stated earlier, the Nearest Root Principle is also helpful in exploring the dynamics of Newton's method applied to cubic polynomials. Naturally, as was the case when dealing with quadratic polynomials, sets of bad initial seeds typically lie along areas of the complex plane where the roots of our polynomial are equidistant. In the case of Newton's method applied to a cubic polynomial with three distinct roots, this would occur when the convex hull of the three roots forms an isosceles triangle. In fact, cases where this triangular hull is isosceles or very close to isosceles are the *only* places where Newton's method applied to cubics fails on an open set. In these cases, extraneous periodic cycles exist and cause Newton's method to fail. This was first discovered in 1983 by Curry, Garnett and Sullivan in their ground-breaking paper "On the iteration of a rational function: Computer experiments with Newton's method" [3]. A more detailed analysis of Newton's method applied to cubic polynomials can be found in [12].

4 A Symmetric Family of Fourth Degree Polynomials

Consider the family of fourth degree polynomials defined by

$$P_\lambda(z) = (z+1)(z-1)(z-\lambda)(z-\overline{\lambda})$$

where $\lambda \in \mathbb{C}$ is a complex parameter. It is plain to see that the roots of P_λ are 1, -1, λ and $\overline{\lambda}$. By expanding the equation for P_λ, we find that

$$\begin{aligned} P_\lambda(z) &= (z^2-1)(z^2 - 2\mathrm{Re}(\lambda)z + |\lambda|^2) \\ &= z^4 - 2\mathrm{Re}(\lambda)z^3 + (|\lambda|^2 - 1)z^2 + 2\mathrm{Re}(\lambda)z - |\lambda|^2. \end{aligned}$$

Thus, we observe that P_λ is a polynomial with real coefficients. As an immediate corollary, it can be seen that the real axis is invariant under P_λ. That is to say, if we restrict the domain of P_λ to $\mathbb{R}$, the image of P_λ will also be contained in $\mathbb{R}$. Furthermore, given the definition of P_λ, it is clear that $P_\lambda = P_{\bar{\lambda}}$, since the roots or P_λ are left unchanged if λ is replaced by $\overline{\lambda}$.

For the remainder of this work, we will denote N_λ as the complex map obtained by applying Newton's method to p_λ. From the facts just mentioned, we see that N_λ leaves the real axis invariant for any $\lambda \in \mathbb{C}$ and that $N_\lambda = N_{\bar{\lambda}}$. This gives a symmetry about the real axis to our figures in the λ-parameter plane.

Another nice property of this family of polynomials is that since P_λ is a polynomial with real coefficients,

$$\overline{P_\lambda(z)} = P_\lambda(\overline{z}).$$

This follows easily from the fact that $\overline{ab} = \overline{a}\overline{b}$. Since each derivative of P_λ is also a polynomial of real coefficients, the general property holds that

$$\overline{P_\lambda^{(k)}(z)} = P_\lambda^{(k)}(\overline{z}).$$

From our expanded form of P_λ, a simple calculation yields its first and second derivatives

$$\begin{aligned} P_\lambda' &= 4z^3 - 6\,\mathrm{Re}(\lambda)z^2 + 2(|\lambda|^2 - 1)z + 2\,\mathrm{Re}(\lambda) \\ P_\lambda'' &= 12z^2 - 12\,\mathrm{Re}(\lambda)z + 2(|\lambda|^2 - 1) \\ &= 12(z^2 - \mathrm{Re}(\lambda)z + \frac{|\lambda|^2 - 1}{6}). \end{aligned}$$

We may then calculate the free critical points of N_λ, where $P_\lambda''(z)$ vanishes. This gives

$$c_\pm = \frac{\mathrm{Re}(\lambda) \pm \sqrt{(\mathrm{Re}(\lambda))^2 - \frac{2}{3}(|\lambda|^2 - 1)}}{2}. \tag{4.1}$$

Since these free critical points will be the primary focus of our study, it will be useful to qualify their existence in the complex plane as thoroughly as possible. To do so, we shall consider the discriminant in equation (4.1).

Suppose $\lambda = a + bi$, and let $\delta = (Re(\lambda))^2 - \frac{2}{3}(|\lambda|^2 - 1)$. Then, by substitution

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

As the sign of δ varies, we see a fundamental change in the nature of the free critical points of N_λ.

1. $\delta > 0 \implies$ Two real free critical points, symmetric about $\frac{\mathrm{Re}(\lambda)}{2}$.
2. $\delta = 0 \implies$ Double free critical point at $\frac{\mathrm{Re}(\lambda)}{2}$.
3. $\delta < 0 \implies$ Two free critical points that form complex conjugate pairs, with real part $\frac{\mathrm{Re}(\lambda)}{2}$.

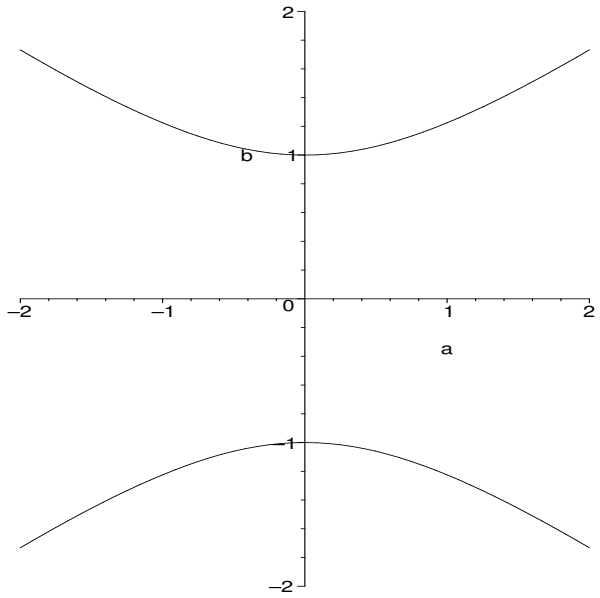

Figure 4.1: The hyperbola that divides the complex plane with respect to the nature of the free critical points of N_λ. Within the curves, our free critical points are real. Outside of the curves they form a complex conjugate pair.

In order to better understand where this dichotomy takes place in the complex plane, we solve for $\delta = 0$. This results in the hyperbola defined by $b^2 - \frac{1}{2}a^2 = 1$. This can be seen in Figure 4.1.

The region between the defining curves of the hyperbola corresponds to our first case where $\delta > 0$. The actual curves are where the dichotomy occurs, which is where $\delta = 0$. Thus, as we saw earlier, within these regions of the λ-parameter plane we have real free critical points whose orbits are unrelated. However, the outermost regions correspond to the third case, where $\delta < 0$. This is where the free critical points are complex conjugates of each other. As we shall see, the orbits of these critical points under N_λ are also complex conjugates of each other.

Now that we have uncovered the necessary information about P_λ and its free critical points, we shift our attention toward Newton's method applied to this family of fourth degree polynomials. Using the general definition of Newton's function applied to P_λ, we have

$$\begin{aligned} N_\lambda(z) &= z - \frac{P_\lambda(z)}{P'_\lambda(z)} \\ &= z - \frac{z^4 - 2\,\mathrm{Re}(\lambda)z^3 + (|\lambda|^2 - 1)z^2 + 2\,\mathrm{Re}(\lambda)z - |\lambda|^2}{4z^3 - 6\,\mathrm{Re}(\lambda)z^2 + 2(|\lambda|^2 - 1)z + 2\,\mathrm{Re}(\lambda)} \\ &= \frac{3z^4 - 4\,\mathrm{Re}(\lambda)z^3 + (|\lambda|^2 - 1)z^2 + |\lambda|^2}{4z^3 - 6\,\mathrm{Re}(\lambda)z^2 + 2(|\lambda|^2 - 1)z + 2\,\mathrm{Re}(\lambda)} \end{aligned}$$

While this rational map may appear imposing, it possesses certain symmetric properties

under conjugacy that significantly reduce the size of the λ-parameter plane we need to consider. It should also be noted that since P_λ has real coefficients, so does N_λ.

As described earlier, the parameter plane for Newton's method applied to P_λ exhibits symmetry about the real axis, which can be seen in Figure 4.2. This image was created by following the orbit of the free critical point c_+ under iteration of N_λ. If c_+ converges to one of the four roots, $\lambda, \overline{\lambda}, 1$, or -1, for some value of λ, then that λ value is colored red, yellow, blue or purple, respectively. If c_+ does not converge to within 10^{-6} of any root after 100 iterations, the parameter value is colored black. Figures 4.3, 4.4 and 4.5 (computed with the same color scheme) exhibit some of the interesting behavior found in the parameter plane that we will examine more closely in the coming sections.

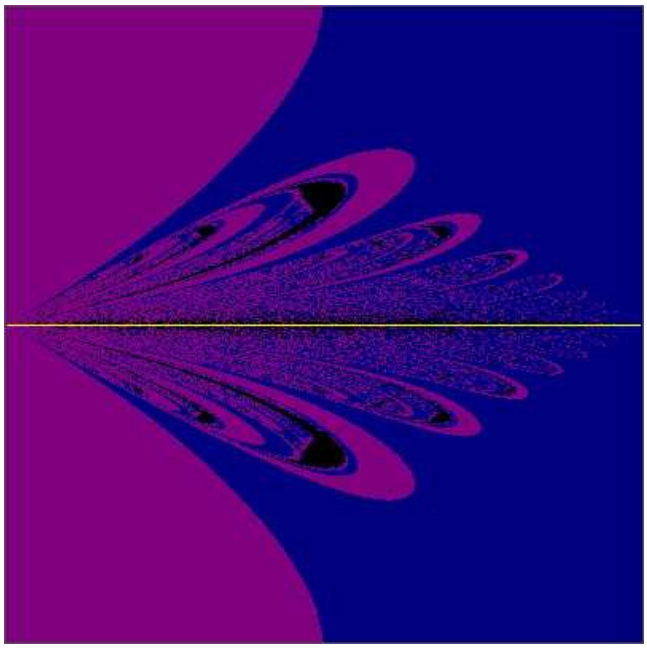

Figure 4.2: The parameter plane for N_λ. We make note of the symmetry about the real axis. This image was produced by following the orbit of the free critical point c_+. The window is $[-1,1] \times [-i,i]$.

Lemma 4.1 $\overline{N_\lambda(z)} = N_\lambda(\overline{z})$.

Proof This is a straightforward calculation. We make use of the fact given earlier, that

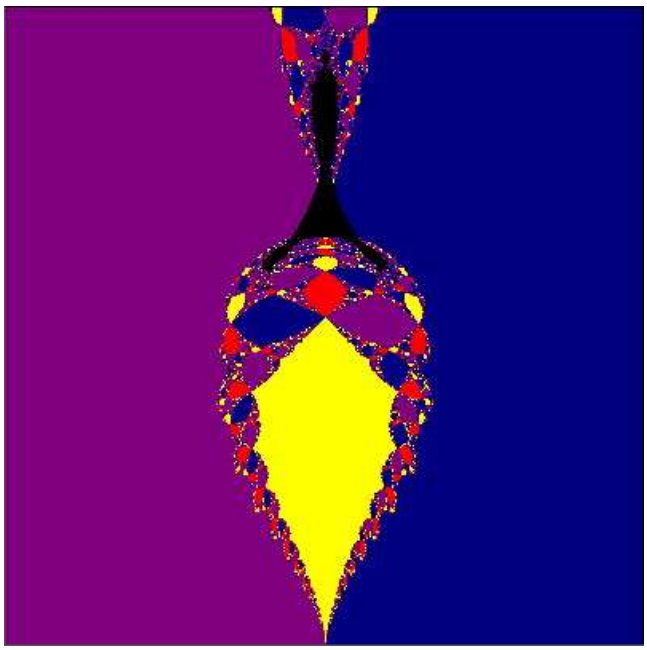

Figure 4.3: The "beetle" and "tricorn" located in the λ-parameter plane along the imaginary axis.

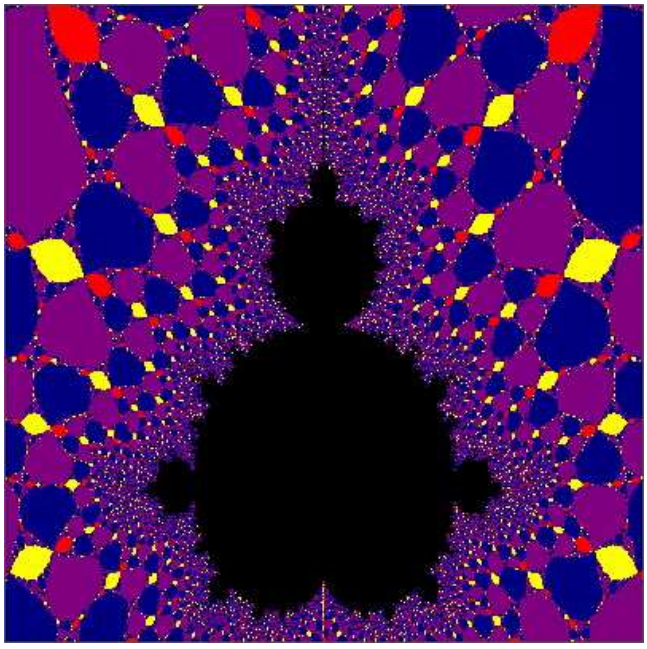

Figure 4.4: A Mandelbrot-like set located along the imaginary axis.

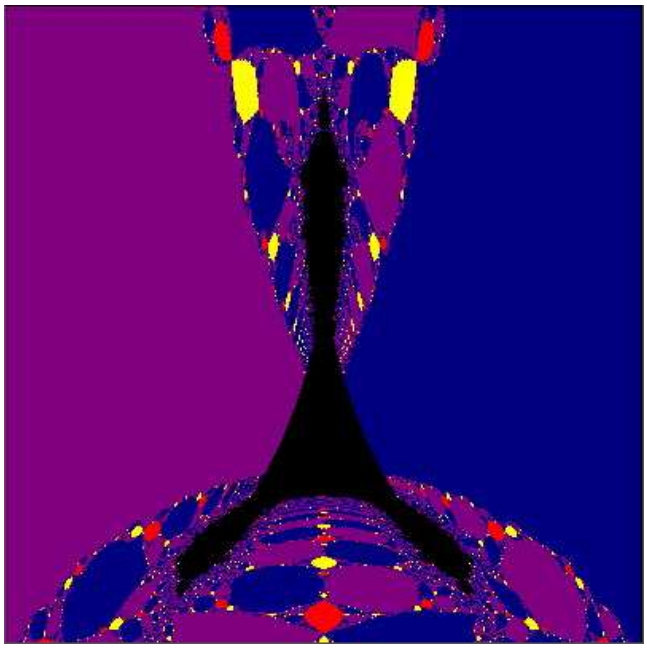

Figure 4.5: Another tricorn located along the imaginary axis.

$\overline{P_\lambda(z)} = P_\lambda(\overline{z})$, and $\overline{P'_\lambda(z)} = P'_\lambda(\overline{z})$.

$$\begin{aligned}
\overline{N_\lambda(z)} &= \overline{z - \frac{P_\lambda(z)}{P'_\lambda(z)}} \\
&= \overline{z} - \frac{\overline{P_\lambda(z)}}{\overline{P'_\lambda(z)}} \\
&= \overline{z} - \frac{P_\lambda(\overline{z})}{P'_\lambda(\overline{z})} \\
&= N_\lambda(\overline{z}).
\end{aligned}$$

□

Lemma 4.2 *N_λ is topologically conjugate to $N_{-\lambda}$ via the conjugacy $h(z) = -z$.*

Proof

$$\begin{aligned}
N_{-\lambda} \circ h &= N_{-\lambda}(-z) \\
&= \frac{3(-z)^4 - 4\,\mathrm{Re}(-\lambda)(-z)^3 + (|-\lambda|^2 - 1)(-z)^2 + |-\lambda|^2}{4(-z)^3 - 6\,\mathrm{Re}(-\lambda)(-z)^2 + 2(|-\lambda|^2 - 1)(-z) + 2\,\mathrm{Re}(-\lambda)} \\
&= \frac{3z^4 - 4\,\mathrm{Re}(\lambda)z^3 + (|\lambda|^2 - 1)z^2 + |\lambda|^2}{-4z^3 + 6\,\mathrm{Re}(\lambda)z^2 - 2(|\lambda|^2 - 1)z - 2\,\mathrm{Re}(\lambda)} \\
&= -\frac{3z^4 - 4\,\mathrm{Re}(\lambda)z^3 + (|\lambda|^2 - 1)z^2 + |\lambda|^2}{4z^3 - 6\,\mathrm{Re}(\lambda)z^2 + 2(|\lambda|^2 - 1)z + 2\,\mathrm{Re}(\lambda)} \\
&= -N_\lambda(z) \\
&= h \circ N_\lambda.
\end{aligned}$$

□

Thus, given the symmetries of N_λ under conjugacy, it suffices to study the region $R = \{\lambda \in \mathbb{C} : \mathrm{Re}(\lambda) \geq 0, \mathrm{Im}(\lambda) \geq 0, \lambda \neq 1\}$. For instance, given Lemma 4.1, we see that the first quadrant of the λ-parameter plane, R, is conjugate to the fourth quadrant. Then, by the result of Lemma 4.2, we see that the second quadrant is conjugate to the fourth, and the third is conjugate to R. So, by applying these two Lemmas, we are able to analyze the entire λ-parameter plane by simply studying the region R.

While this is the region that we will be concerned with throughout the remainder of this paper, the following lemma shows that in the case where λ is purely imaginary, we can, in fact, reduce our parameter space further.

Lemma 4.3 *Suppose $\lambda = \beta i$. Then $N_{\beta i}$ is conjugate to $N_{-\frac{1}{\beta}i}$ via the conjugacy $h(z) = \frac{z}{\lambda} = \frac{z}{\beta i}$.*

Proof This is simply an affine conjugacy that maps roots to roots. Thus, this result follows from the proof of Lemma 3.1. To emphasize this result, however, we offer an alternative proof below.

We check directly that $N_{-\frac{1}{\beta}i} \circ h = h \circ N_{\beta i}$. For $\lambda = \beta i$, we have

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

$$\text{Therefore, } N_{-\frac{1}{\beta}i}(z) = \frac{3z^4 + \left(\frac{1}{\beta^2} - 1\right) z^2 + \frac{1}{\beta^2}}{4z^3 + 2\left(\frac{1}{\beta^2} - 1\right) z}$$

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

$$\text{and } N_{-\frac{1}{\beta}i}\left(\frac{z}{\beta i}\right) = \frac{\frac{3}{\beta^2} z^4 + \left(1 - \frac{1}{\beta^2}\right) z^2 + 1}{\frac{4}{\beta} z^3 i + 2(1-\beta^2)\left(\frac{z}{\beta i}\right)}$$

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

$$= \frac{1}{\beta i} \cdot N_{\beta i}(z)$$

$$= h \circ N_{\beta i} \, .$$

□

Remark 4.4 The imaginary axis is the only region of the complex plane on which this conjugacy holds.

To see this, suppose the conjugacy in Lemma 4.3 held for general λ values not necessarily on the imaginary axis. Further, let's consider the polar coordinate representation of λ. Since any complex number z can be written in the form $re^{i\theta}$, where r is the modulus of z and θ is its angle from the real axis, we may make the following substitutions.

$$\lambda = re^{i\theta}, \quad \frac{1}{\lambda} = \frac{1}{r} e^{-i\theta},$$

In order for our conjugacy to hold, we must have

$$\frac{1}{\lambda} N_\lambda(z) = N_{\frac{1}{\lambda}}(\frac{z}{\lambda}).$$

Inserting our polar representation for λ, we have

$$\frac{1}{\lambda} N_\lambda(z) = \frac{1}{\lambda}\left(\frac{3z^4 - 4r\cos\theta z^3 + (r^2-1)z^2 + r^2}{4z^3 - 6r\cos\theta z^2 + 2(r^2-1)z + 2r\cos\theta}\right).$$

Then,

$$N_{\frac{1}{\lambda}}(\frac{z}{\lambda}) = \frac{3(\frac{z}{\lambda})^4 - \frac{4}{r}\cos\theta(\frac{z}{\lambda})^3 + (\frac{1}{r^2} - 1)(\frac{z}{\lambda})^2 + \frac{1}{r^2}}{4(\frac{z}{\lambda})^3 - \frac{6}{r}\cos\theta(\frac{z}{\lambda})^2 + 2(\frac{1}{r^2} - 1)(\frac{z}{\lambda}) + \frac{2}{r}\cos\theta}$$

$$= \frac{3z^4 - 4\cos\theta e^{i\theta} z^3 + (1-r^2)z^2 e^{2i\theta} + r^2 e^{4i\theta}}{\lambda\left(4z^3 - 6\cos\theta e^{i\theta} z^2 + 2(1-r^2)z e^{2i\theta} + 2r^2\cos\theta e^{4i\theta}\right)} \, .$$

Thus, we see that for $\frac{1}{\lambda}N_\lambda(z) = N_{\frac{1}{\lambda}}(\frac{z}{\lambda})$, we must have $\theta = \frac{\pi}{2}$. Thus, it is clear that the conjugacy between N_λ and $N_{\frac{1}{\lambda}}$ only holds when λ is on the imaginary axis. □

This makes sense, given that we are dealing with an affine conjugacy. Under an affine conjugacy, roots are always mapped to other roots. For instance, on the imaginary axis, we see that $\frac{\lambda}{\overline{\lambda}} = -1$. However, when we move λ off the imaginary axis, this is no longer the case, as we see that $\overline{\lambda}$ is not mapped to any other root.

Even though our use of this conjugacy is limited to the imaginary axis, it will prove to be useful not only in the case where λ is purely imaginary, but also in the more general case where λ has both real and imaginary parts. We will be able to use the result of Lemma 4.3 to predict and explain the existence of various dynamical components in the λ-parameter plane.

5 Restriction to the Real Axis

We begin our study of N_λ by restricting to λ values on the real axis. Let us first take a look at how this effects our family of polynomials, P_λ.

Since $\mathrm{Im}(\lambda) = 0$, it is plain to see that $\lambda = \overline{\lambda}$. Thus, our family of functions now has a double root at λ, and P_λ is defined by

$$P_\lambda(z) = (z^2 - 1)(z - \lambda)^2.$$

Furthermore,

$$\begin{aligned} P'_\lambda(z) &= 2(z^2 - 1)(z - \lambda) + 2z(z - \lambda)^2 \\ &= 2(z - \lambda)(2z^2 - \lambda z - 1). \end{aligned}$$

As mentioned earlier, since N_λ is a rational map with real coefficients, the real axis is invariant under N_λ. Then, since all free critical points for $\lambda \in \mathbb{R}$ are real, we can study N_λ as a real-valued dynamical system when we restrict our domain to the real axis.

Given the symmetric nature of N_λ discussed in the previous section, we need only study $\lambda \geq 0$. Since we are dealing with a real-valued system in this case, we will use the standard convention and replace our general complex variable z with x, where $x \in \mathbb{R}$. Thus, our new Newton's function is

$$N_\lambda(x) = \frac{3x^4 - 4\lambda x^3 + (\lambda^2 - 1)x^2 + \lambda^2}{4x^3 - 6\lambda x^2 + 2(\lambda^2 - 1)x + 2\lambda}.$$

At first glance this does not appear to be much nicer to work with than our general Newton's function. Upon inspection, however, we can greatly simplify this map by factoring out the term $(x - \lambda)$ from both the numerator and denominator. We are then left with

$$N_\lambda(x) = \frac{3x^3 - \lambda x^2 - x - \lambda}{2(2x^2 - \lambda x - 1)}.$$

Along with the reduction of the Newton function, we can also redefine our free critical points for $\lambda \in \mathbb{R}$. We now have

$$c_{\pm} = \frac{\lambda \pm \sqrt{\frac{1}{3}\lambda^2 + \frac{2}{3}}}{2}.$$

Furthermore, now that the denominator of N_λ has been reduced to a quadratic polynomial, we are able to solve for the poles of this rational map.

$$p_{\pm} = \frac{\lambda \pm \sqrt{\lambda^2 + 8}}{4}$$

The derivative of our Newton function has also become significantly reduced. Recall, for any polynomial p, $N_p' = \frac{(p)(p'')}{(p')^2}$. Then, since we can write

$$P_\lambda''(z) = 12(z - c_-)(z - c_+),$$

we see that

$$N_\lambda' = \frac{12(z^2-1)(z-\lambda)^2(z-c_-)(z-c_+)}{(2(z-\lambda)(2z^2-\lambda z-1))^2}$$

$$= \frac{3(z^2-1)(z-c_-)(z-c_+)}{(2z^2-\lambda z-1)^2}$$

At this point, we would like to begin examining N_λ as a real-valued dynamical system. Due to the results of Lemma 4.1 and Lemma 4.2, we are only concerned with λ values such that $\lambda \geq 0$. Nevertheless, along the real axis, we would like to be able to explain how our system changes qualitatively as we vary λ. More specifically, we would like to know what happens to our free critical points, $c_\pm$ as λ changes. By studying the relationships between our roots, free critical points, and poles, we are able to predict λ values where we might expect bifurcations to occur. As it turns out, there are three distinct cases that we must consider.

Case 1: $0 \leq \lambda < 1$

Lemma 5.1 *For* $0 \leq \lambda < 1$,

$$-\frac{\sqrt{2}}{2} < p_- < c_- < 0 \leq \lambda < c_+ < p_+ < 1.$$

Proof We are given $0 \leq \lambda < 1$. We then see that $\lambda < c_+$ if and only if

$$\lambda < \frac{\lambda + \sqrt{\frac{\lambda^2}{3} + \frac{2}{3}}}{2}$$

$$\iff 2\lambda < \lambda + \sqrt{\frac{\lambda^2}{3} + \frac{2}{3}}$$

$$\iff \lambda < \sqrt{\frac{\lambda^2}{3} + \frac{2}{3}}\,.$$

Then, after squaring both sides and equating terms, we have

$$\frac{2}{3}\lambda^2 < \frac{2}{3}$$

$$\iff \lambda^2 < 1.$$

Thus, since we had assumed $\lambda < 1$, we see that $\lambda < c_+$ is valid.

Similarly, $c_+ < p_+$ if and only if

$$\frac{\lambda + \sqrt{\frac{\lambda^2}{3} + \frac{2}{3}}}{2} < \frac{\lambda + \sqrt{\lambda^2 + 8}}{4}$$

$$\iff \lambda + 2\sqrt{\frac{\lambda^2}{3} + \frac{2}{3}} < \sqrt{\lambda^2 + 8}\,.$$

After squaring both sides and simplifying, we have

$$\frac{4}{3}\lambda^2 + 4\lambda\sqrt{\frac{\lambda^2}{3} + \frac{2}{3}} + \frac{8}{3} < 8$$

$$\iff \lambda^2 + 3\lambda\sqrt{\frac{\lambda^2}{3} + \frac{2}{3}} < 4$$

$$\iff 3\lambda\sqrt{\frac{\lambda^2}{3} + \frac{2}{3}} < 4 - \lambda^2\,.$$

Squaring again yields

$$3\lambda^4 + 6\lambda^2 < \lambda^4 - 8\lambda^2 + 16$$

$$\iff \lambda^4 + 7\lambda^2 - 8 < 0$$

$$\iff (\lambda^2 + 8)(\lambda^2 - 1) < 0$$

$$\implies \lambda^2 - 1 < 0 \text{ since } \lambda^2 + 8 \text{ is positive } \forall \lambda \in \mathbb{C}.$$

Again, given our hypothesis, we see that this series of inequalities holds true. We will now show that $p_+ < 1$. This is true if and only if

$$\frac{\lambda + \sqrt{\lambda^2 + 8}}{4} < 1$$

$$\iff \sqrt{\lambda^2 + 8} < 4 - \lambda\,.$$

Squaring both sides gives us

$$\lambda^2 + 8 < 16 - 8\lambda + \lambda^2$$

$$\iff -8 < -8\lambda$$

$$\iff \lambda < 1.$$

At this point, we have shown the nonnegative portion of our inequality to be true. We will now show $c_- < 0$, which holds true if and only if

$$\frac{\lambda - \sqrt{\frac{\lambda^2}{3} + \frac{2}{3}}}{2} < 0$$

$$\iff \lambda < \sqrt{\frac{\lambda^2}{3} + \frac{2}{3}}\,.$$

After squaring, we have

$$\lambda^2 < \frac{\lambda^2}{3} + \frac{2}{3}$$

$$\iff 0 < -2\lambda^2 + 2$$

which is satisfied by $\lambda < 1$.

Next, $p_- < c_-$ if and only if

$$\frac{\lambda - \sqrt{\lambda^2 + 8}}{4} < \frac{\lambda - \sqrt{\frac{\lambda^2}{3} + \frac{2}{3}}}{2}$$

$$\iff -\sqrt{\lambda^2 + 8} < \lambda - 2\sqrt{\frac{\lambda^2}{3} + \frac{2}{3}}$$

$$\iff 2\sqrt{\frac{\lambda^2}{3} + \frac{2}{3}} < \lambda + \sqrt{\lambda^2 + 8}\,.$$

After squaring and some simplification, we have

$$\frac{4}{3}\lambda^2 + \frac{8}{3} < 2\lambda^2 + 2\lambda\sqrt{\lambda^2 + 8} + 8$$

$$\iff 0 < 2\lambda^2 + 6\lambda\sqrt{\lambda^2 + 8} + 16$$

which is valid since $\lambda > 0$.

Lastly, we must show that $\frac{-\sqrt{2}}{2} < p_-$, which holds if and only if

$$\frac{-\sqrt{2}}{2} < \frac{\lambda - \sqrt{\lambda^2 + 8}}{4}$$

$$\iff -2\sqrt{2} < \lambda - \sqrt{\lambda^2 + 8}$$

$$\iff \sqrt{\lambda^2 + 8} < \lambda + 2\sqrt{2}\,.$$

Squaring both sides gives us

$$\begin{aligned} & \lambda^2 + 8 < \lambda^2 + 4\lambda\sqrt{2} + 8 \\ \iff & 4\lambda\sqrt{2} > 0 \\ \iff & \lambda > 0. \end{aligned}$$

□

Lemma 5.2 *Suppose* $0 \leq \lambda < 1$. *Then for all* $x \in [c_-, c_+]$, ${N_\lambda}^k(x) \to \lambda$ *as* $k \to \infty$.

Proof We begin by showing that $N_\lambda(c_-) > c_-$.

$$N_\lambda(c_-) = \frac{3(c_-)^3 - \lambda(c_-)^2 - (c_-) - \lambda}{4(c_-) - 2\lambda(c_-) - 2}.$$

$$\begin{aligned} N_\lambda(c_-) - c_- &= \frac{3(c_-)^3 - \lambda(c_-)^2 - (c_-) - \lambda - 4(c_-)^3 + 2\lambda(c_-)^2 + 2(c_-)}{4(c_-) - 2\lambda(c_-) - 2} \\ &= \frac{-(c_-)^3 + \lambda(c_-)^2 + (c_-) - \lambda}{4(c_-) - 2\lambda(c_-) - 2}. \end{aligned}$$

To determine the sign of $N_\lambda(c_-) - c_-$, we first calculate the signs of the numerator and denominator, respectively. We begin with the denominator, $4(c_-) - 2\lambda(c_-) - 2$. Since we showed in Lemma 5.1 that $c_- < 0$, and since $\lambda > 0$, we see that $4(c_-)$ and $-2\lambda(c_-)$ are positive terms. In fact,

$$\begin{aligned} 4(c_-) - 2\lambda(c_-) &= \frac{4(\lambda^2 - 2\lambda\sqrt{\frac{\lambda^2}{3} + \frac{2}{3}} + \frac{\lambda^2}{3} + \frac{2}{3})}{4} - 2\lambda\left(\frac{\lambda - \sqrt{\frac{\lambda^2}{3} + \frac{2}{3}}}{2}\right) \\ &= \frac{1}{3}\lambda^2 - \lambda\sqrt{\frac{\lambda^2}{3} + \frac{2}{3}} + \frac{2}{3} \\ &< \frac{2}{3}. \end{aligned}$$

Thus, the denominator in this case is negative.

Let us now consider the numerator of $N_\lambda(c_-)$.

$$\begin{aligned}
-(c_-)^3 + \lambda(c_-)^2 + (c_-) - \lambda &= (c_-)^2(\lambda - (c_-)) + (c_-) - \lambda \\
&= (c_-)^2 \left(\lambda - \left(\frac{\lambda - \sqrt{\frac{\lambda^2}{3} + \frac{2}{3}}}{2} \right) \right) + \frac{\lambda - \sqrt{\frac{\lambda^2}{3} + \frac{2}{3}}}{2} - \lambda \\
&= (c_-)^2 \left(\frac{\lambda}{2} + \frac{\sqrt{\frac{\lambda^2}{3} + \frac{2}{3}}}{2} \right) - \left(\frac{\lambda}{2} + \frac{\sqrt{\frac{\lambda^2}{3} + \frac{2}{3}}}{2} \right) \\
&= (c_+)((c_-)^2 - 1).
\end{aligned}$$

As we saw in Lemma 5.1, $-1 < c_- < 0$, which implies that $(c_-)^2 - 1 < 0$. Further, since the same lemma tells us that $c_+ > 0$, we see that the numerator of $N_\lambda(c_-) - c_-$ is negative in this case. Further, since we have already shown the denominator to be negative, we see that

$$N_\lambda(c_-) - c_- > 0,$$

which implies that

$$N_\lambda(c_-) > c_-$$

for $0 < \lambda < 1$.

A similar series of calculations shows that $N_\lambda(c_+) < c_+$. Further, since we know that the only fixed point in the interval (c_-, c_+) is λ, it must be the case that $N_\lambda(x) > x$ for all $x \in [c_-, \lambda)$, and $N_\lambda(x) < x$ for all $x \in (\lambda, c_+]$. This can be seen in Figure 5.1.

Now that we know the relation between N_λ and the diagonal on $[c_-, c_+]$, we must consider the slope of N_λ on this interval. Let $x \in [c_-, c_+]$, and consider $N'_\lambda(x)$. In our definition of N'_λ, it is plain to see that the denominator is always positive. Thus, to determine the sign of $N'_\lambda(x)$, we simply check each term in the numerator. Since $c_+ < 1$, we see that $(x^2 - 1) < 0$. Further, since x is located between c_- and c_+, it is clear that the product of the remaining terms, $(x - c_-)$ and $(x - c_+)$ must be negative as well (or zero). Therefore, $N'_\lambda(x) \geq 0$. So, we see that N_λ is non-decreasing on $[c_-, c_+]$.

Case 1A: $c_- \leq x < \lambda$.

As shown above, for $x \in [c_-, \lambda)$, $N_\lambda(x)$ is decreasing and $N_\lambda(x) > x$. Thus, we see that ${N_\lambda}^k(x)$ forms an increasing sequence that is bounded above by the fixed point λ. We would like to show that ${N_\lambda}^k(x)$ must converge to λ.

Assume

$$\lim_{k \to \infty} {N_\lambda}^k(x) = \lambda^* \text{ where } \lambda^* < \lambda.$$

Then, since N_λ is continuous, we have

$$\lim_{x \to \lambda^*} N_\lambda = N_\lambda(\lambda^*).$$

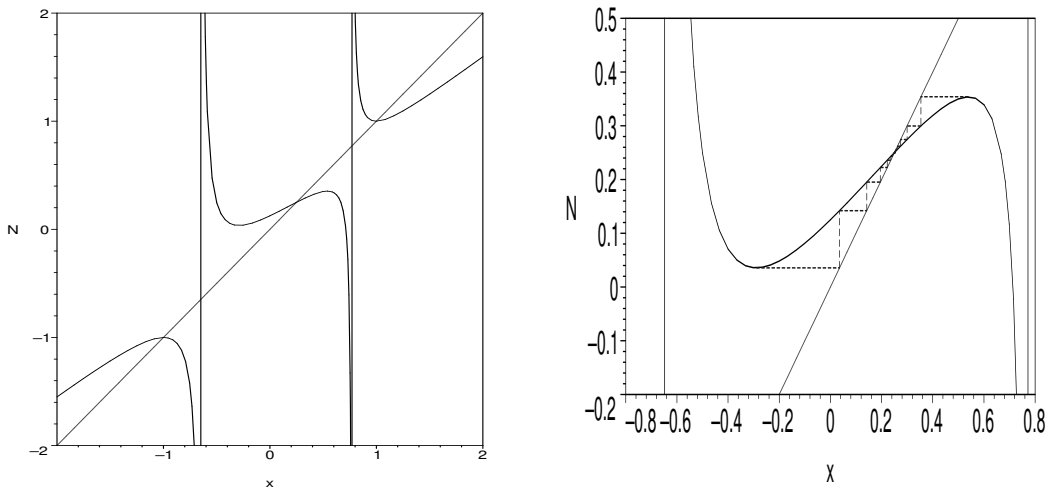

Figure 5.1: To the left is the general picture of N_λ for $0 \leq \lambda < 1$. To the right is an orbit diagram showing the dynamical behavior of the free critical points c_+ and c_-. From this picture we see that both free critical points converge to the root λ under iteration of N_λ.

By substitution

$$\begin{aligned} N_\lambda(\lambda^*) &= N_\lambda(\lim_{k\to\infty} {N_\lambda}^k(x)) \\ &= \lim_{k\to\infty} {N_\lambda}^{k+1}(x) \\ &= \lambda^* \end{aligned}$$

This is a contradiction, since λ^* is not a fixed point of N_λ (λ is the only fixed point between c_- and c_+.) Therefore, in this case, ${N_\lambda}^k(x)$ converges to λ.

Case 1B: $\lambda < x \leq c_+$.

In this case, we have shown that $N_\lambda(x) < x$ for all $x \in (\lambda, c_+]$, and that $N_\lambda(x)$ is decreasing for all such x. Thus, we see that ${N_\lambda}^k(x)$ forms a decreasing sequence that is bounded below by λ. A similar argument to that above shows that all x in this interval converge to λ under iteration of N_λ. □

At $\lambda = 1$, we find our first bifurcation point. Here, we see that $c_+ = 1 = \lambda$. Thus, we now have a triple root at $\lambda = 1$. So dynamically, c_+ has become a fixed point, while c_- continues to converge to λ, which happens to be 1 in this case. For $\lambda = 1$,

$$N_\lambda(x) = \frac{3x^2 + 2x + 1}{4x + 2},$$

from which similar arguments as above show that $c_- \to \lambda$. Moving further along the real axis, as λ becomes greater than 1, we see that c_- converges to 1, which is now less than λ. As we shall see in Case 2, we have a bit of a reversal of roles with respect to our important values.

Case 2: $1 < \lambda < 5$

Lemma 5.3 *For* $1 < \lambda < 5$,

$$-\frac{\sqrt{2}}{2} < p_- < 0 < c_- < 1 < p_+ < c_+ < \lambda.$$

Proof As mentioned previously, in this case we see that λ and 1 swap positions from what we saw in Case 1. Furthermore, as a result of this, $c_+ = \frac{\lambda+\sqrt{\frac{\lambda^2}{2}+\frac{2}{3}}}{2}$ jumps to the other side of the positive pole, p_+. Showing that this is the case follows from a straightforward series of calculations similar to that found in the proof of Lemma 5.1. A generic picture of this case can be seen in Figure 5.2. □

Lemma 5.4 *Suppose* $1 < \lambda < 5$. *Then* ${N_\lambda}^k(c_-) \to 1$, *and* ${N_\lambda}^k(c_+) \to \lambda$ *as* $k \to \infty$.

Proof Again, the proof here is very similar to that of Lemma 5.4 found in Case 1. While the positioning of our important values has changed somewhat, no radically different dynamical behavior has been introduced. As discussed earlier, while c_- converged to λ in Case 1, at the bifurcation $\lambda = 1$, λ and 1 essentially swap roles with respect to the behavior of c_-. Furthermore, now that c_+ is greater than p_+, its dynamical behavior is very straightforward given the nature of N_λ in this region. Since λ is the only fixed point greater than p_+, we see that c_+ converges to λ. The details of this proof are similar to those found in the proof of Lemma 5.4. The orbit diagram in Figure 5.2 is an example of this behavior. □

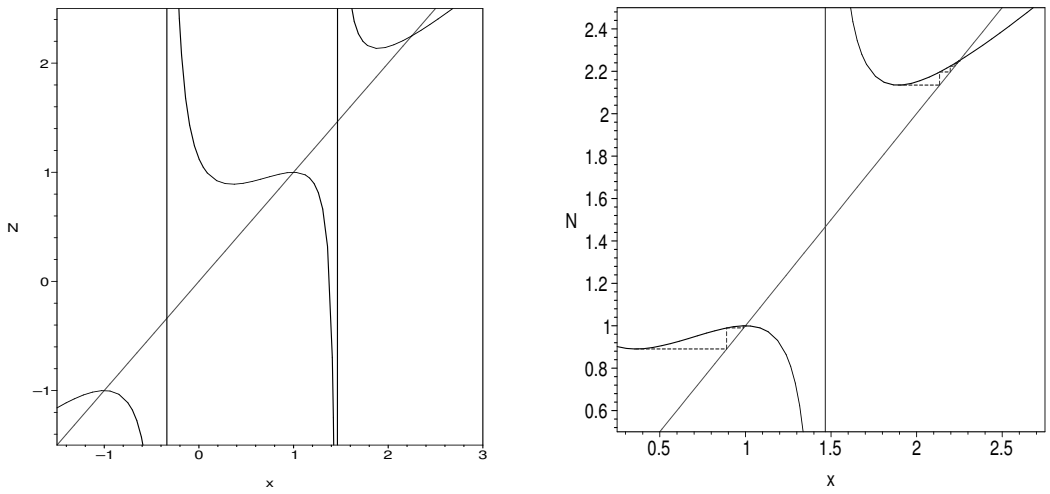

Figure 5.2: To the left is the general picture of N_λ for $1 < \lambda < 5$. To the right is an orbit diagram which shows the dynamic behavior of c_+ and c_-. This verifies that $c_- \to 1$ and $c_+ \to \lambda$ under iteration of N_λ.

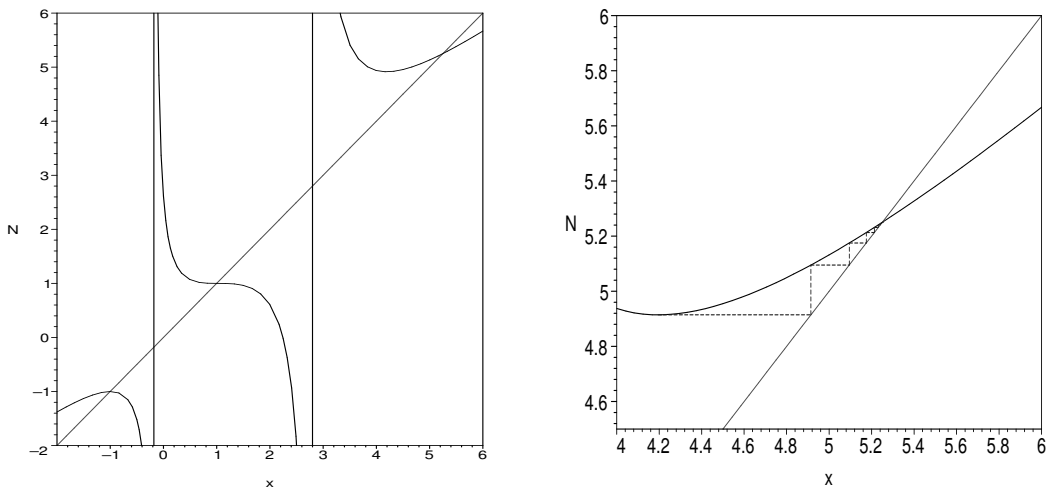

Figure 5.3: To the left is the general graph of N_λ for $\lambda > 5$. To the right is a web diagram showing the orbit of c_+ as it converges to the root λ.

When $\lambda = 5$, we have a bifurcation similar to what we saw for $\lambda = 1$. In this case, we have $c_- = 1$, though our Newton map is not significantly reduced. Furthermore, the dynamical behavior of the free critical points does not change, although the relative location of our key values does change somewhat.

Case 3: $\lambda > 5$

Lemma 5.5 *For* $\lambda > 5$,

$$-\frac{\sqrt{2}}{2} < p_- < 0 < 1 < c_- < p_+ < c_+ < \lambda.$$

Proof This proof is very similar to that of Lemma 5.1 from Case 1. The only difference here is that c_- has become greater than 1. The rest of the inequality is left unchanged from Case 2 (see Figure 5.3). □

Lemma 5.6 *Suppose* $\lambda > 5$. *Then* ${N_\lambda}^k(c_-) \to 1$, *and* ${N_\lambda}^k(c_+) \to \lambda$ *as* $k \to \infty$.

Proof This proof is very similar to that of Lemma 5.2 from Case 1. The behavior of c_+ and c_- can be seen from the orbit diagram in Figure 5.3. □

After examining these three cases, it is plain to see that when $\lambda \in \mathbb{R}$, there is no interesting dynamical behavior. For each $\lambda \in \mathbb{R}$, $N_\lambda(c_+)$ and $N_\lambda(c_+)$ converge to one of the three roots of P_λ monotonically. Moreover, by Theorem 2.9, there are no extraneous periodic cycles in $\mathbb{C}$ since the orbits of the free critical points are accounted for. Thus, we see that in the parameter plane, the real axis is dynamically straight-forward. As we shall see in the following section, the same cannot be said of the imaginary axis.

6 Restriction to the Imaginary Axis

Let us now suppose that $\lambda = \beta$i. Here, λ lies on the imaginary axis, so $\mathrm{Re}(\lambda) = 0$. We will denote N_β for the map corresponding to Newton's method applied to $P_{\beta i}$. Recall from Lemma 4.3 that we need only study β values for $0 < \beta \leq 1$. Furthermore, it should be noted that when λ lies along the imaginary axis, the complex hull of the four roots of P_λ forms a rhombus, which one might expect to lead to interesting dynamical behavior. Similar to our work in the real case, we would like to study N_β as a real dynamical system. To do this, however, we must first verify that the imaginary axis is, in fact, invariant under N_β.

We begin by simplifying our polynomial $P_{\beta i}$,

$$\begin{aligned} P_{\beta i}(z) &= (z^2 - 1)(z^2 + \beta^2) \\ &= z^4 + (\beta^2 - 1)z^2 - \beta^2 \end{aligned}$$

As we did in the case of λ real, we will now compute our key dynamical values. We begin with the poles of N_β, which are critical points of $P_{\beta i}$. Since $P'_{\beta i}(z) = 4z^3 + 2(\beta^2 - 1)z$, we see that zero will always be a pole for our Newton function in this case. Solving the remaining quadratic term, we find the other two poles

$$\begin{aligned} p_\pm &= \pm\frac{\sqrt{(-4)(4)(2\beta^2 - 2)}}{8} \\ &= \pm\frac{4\sqrt{2 - 2\beta^2}}{2} \\ &= \pm\sqrt{\frac{1 - \beta^2}{2}}\,. \end{aligned}$$

Now that we have located the poles of N_β, we must do the same for its free critical points, where

$$P''_{\beta i}(z) = 12z^2 + 2(\beta^2 - 1)$$

vanishes. In a similar calculation to that above, we solve for $P''_{\beta i}$ to obtain our free critical points

$$\begin{aligned} c_\pm &= \pm\frac{\sqrt{96 - 96\beta^2}}{24} \\ &= \pm\frac{4\sqrt{6 - 6\beta^2}}{24} \\ &= \pm\sqrt{\frac{1 - \beta^2}{6}}\,. \end{aligned}$$

A rather nice relationship between $c_\pm$ and $p_\pm$ emerges from their respective definitions. We see that our poles and free critical points are actually scalar multiples of each other, with $c_\pm = \frac{1}{\sqrt{3}}(p_\pm)$, respectively.

Once again, we have a simplified version of our Newton function.

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

Similar to the Newton map in the real case, N_β has real coefficients, and thus leaves the real axis invariant. This is vital to our study of the imaginary axis due to the fact that as we saw in Section 4, for $0 \leq \beta \leq 1$, the free critical points of N_β are real. Thus it is important to confirm that the real axis is invariant under N_β so that we may proceed to study it as a real-valued dynamical system. Moreover, it is also the case that the imaginary axis is invariant under N_β.

Let z = bi. Then,

$$\begin{aligned} N_\beta(z) = N_\beta(bi) &= \frac{3(bi)^4 + (\beta^2 - 1)(bi)^2 + \beta^2}{4(bi)^3 + 2(\beta^2 - 1)(bi)} \\ &= \frac{3b^4 + (\beta^2 - 1)(-b^2) + \beta^2}{i(4b^3 + 2(\beta^2 - 1)b)} \\ &= \frac{-3b^4 + (\beta^2 - 1)b^2 - \beta^2}{4b^3 + 2(\beta^2 - 1)b} i. \end{aligned}$$

It is then plain to see that as we iterate purely imaginary values, we do not stray from the imaginary axis under N_β. While we have already confirmed that we need only study β values between 0 and 1, this fact is useful in the sense that for $\beta > 1$, our free critical points are purely imaginary. Therefore, as it was important to ensure the invariance of the real axis under N_β, it is also important to confirm that N_β leaves the imaginary axis invariant, because otherwise we would not be able to iterate our simplified Newton map on the entire axis. Nevertheless, given our conjugacy earlier in Lemma 4.3, this confirms what we expected. In the proof of Lemma 4.3, we saw that N_β was conjugate to $N_{-\frac{1}{\beta}}$ via an inversion and rotation by $\frac{\pi}{2}$. Thus, under this rotation, we see that our free critical points on the real axis are scaled and rotated to the imaginary axis.

We are now able to study N_β as a real-valued dynamical system. From this point on, we will use the standard convention and replace our complex variable z with the real variable x. We begin with some basic properties of N_β, and its dynamical behavior. For instance, given the following, we see that N_β is an odd function:

$$\begin{aligned} N_\beta(-x) &= \frac{3(-x)^4 + (\beta^2 - 1)(-x)^2 + \beta^2}{4(-x)^3 + 2(\beta^2 - 1)(-x)} \\ &= -\frac{3x^4 + (\beta^2 - 1)x^2 + \beta^2}{4x^3 + 2(\beta^2 - 1)x} \\ &= -N_\beta(x). \end{aligned}$$

Lemma 6.1 $-\frac{1}{\sqrt{2}} < p_- < c_- \leq 0 \leq c_+ < p_+ < \frac{1}{\sqrt{2}}$.

Proof This is a fairly straightforward series of calculations. First off, since $0 < \beta \leq 1$, we see that $1 - \beta^2 \geq 0$. Thus, given the equations for $c_\pm$ we have $c_- \leq 0 \leq c_+$. Further, as mentioned above, $c_\pm = \frac{1}{\sqrt{3}}(p_\pm)$. It is then plain to see that $p_- < c_-$ and $c_+ < p_+$. Lastly, since $\beta > 0$, it is clear that $p_- > -\frac{1}{\sqrt{2}}$ and $p_+ < \frac{1}{\sqrt{2}}$, which completes our inequality (see Figure 6.1). □

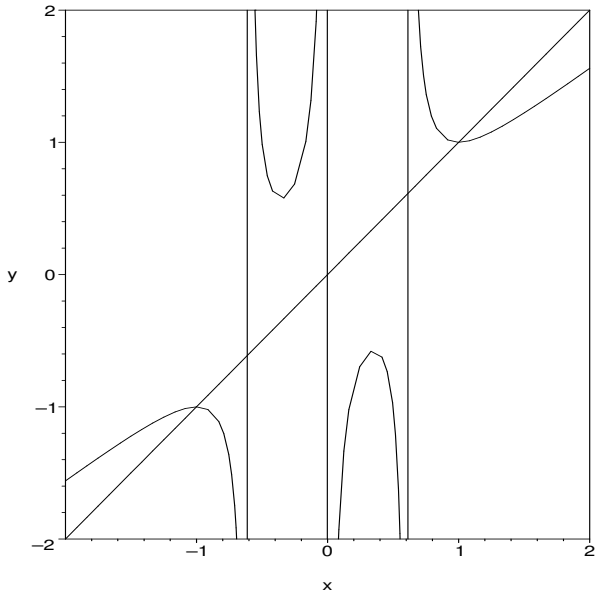

Figure 6.1: A general graph of N_β for $0 \leq \beta \leq 1$.

Lemma 6.2 *For $0 < \beta \leq 1$, $N_\beta(c_+)$ is decreasing with respect to β, while $N_\beta(c_-)$ is increasing with respect to β.*

Proof Recall, $c_+ = \sqrt{\frac{1-\beta^2}{6}}$. We then begin by simplifying $N_\beta(c_+)$.

$$N_\beta(c_+) = \frac{3(\frac{1-\beta^2}{6})^2 + (\beta^2 - 1)(\frac{1-\beta^2}{6}) + \beta^2}{\sqrt{\frac{1-\beta^2}{6}}(4(\frac{1-\beta^2}{6}) + 2(\beta^2 - 1))}$$

$$= \frac{-\frac{1}{12}(1-\beta^2)^2 + \beta^2}{\sqrt{\frac{1-\beta^2}{6}}(-\frac{4}{3}(1-\beta^2))}$$

$$= \frac{-\frac{1}{12}(1-\beta^2)^2 + \beta^2}{-\frac{4}{3\sqrt{6}}(1-\beta^2)^{\frac{3}{2}}}.$$

For simplicity, we let $\gamma = 1 - \beta^2$. Taking the derivative of $N_\beta(c_+)$ with respect to β and simplifying the numerator gives

$$\begin{aligned}
\text{Num}\left[\frac{d}{d\beta}\left(\frac{-\frac{1}{12}(\gamma)^2+\beta^2}{-\frac{4}{3\sqrt{6}}(\gamma)^{\frac{3}{2}}}\right)\right] &= -\frac{4}{3\sqrt{6}}(\gamma)^{\frac{3}{2}}(\frac{1}{3}(\gamma)\beta+2\beta) - (-\frac{1}{12}(\gamma)^2+\beta^2)(\frac{4\beta}{\sqrt{6}}(\gamma)^{\frac{1}{2}}) \\
&= \frac{-4\beta}{9\sqrt{6}}(\gamma)^{\frac{5}{2}} + \frac{\beta}{3\sqrt{6}}(\gamma)^{\frac{5}{2}} - \frac{8\beta}{3\sqrt{6}}(\gamma)^{\frac{3}{2}} - \frac{4\beta^3}{\sqrt{6}}(\gamma)^{\frac{1}{2}} \\
&= \frac{-\beta}{9\sqrt{6}} - \frac{8\beta}{3\sqrt{6}}(\gamma)^{\frac{3}{2}} - \frac{4\beta^3}{\sqrt{6}}(\gamma)^{\frac{1}{2}}.
\end{aligned}$$

Then, since β is positive, we see that each of these terms is negative. Thus, we have shown that $N_\beta(c_+)$ is decreasing in β for $0 < \beta \leq 1$.

Moreover, since we have shown that N_β is an odd function, and that $c_- = -c_+$, we see that $N_\beta(c_-) = -N_\beta(c_+)$. Therefore, differentiating $N_\beta(c_-)$ with respect to β would result in multiplying $\frac{d}{d\beta}\left(N_\beta(c_+)\right)$ by -1. Thus, since we have shown $N_\beta(c_+)$ to be decreasing in β, it follows that $N_\beta(c_-)$ is increasing in β for $0 \leq \beta \leq 1$. □

Lemma 6.3 *For $1/\sqrt{3} \leq \beta \leq 1$, $N_\beta(c_+) \leq -1$ and $N_\beta(c_-) \geq 1$. Therefore, c_+ converges to -1 and c_- converges to 1 under iteration of N_β.*

Proof As we saw in the proof of Lemma 6.2, $N_\beta(c_+) = \frac{-\frac{1}{12}(1-\beta^2)^2+\beta^2}{-\frac{4}{3\sqrt{6}}(1-\beta^2)^{\frac{3}{2}}}$. Thus for $\beta = 1/\sqrt{3}$, we have

$$\begin{aligned}
N_{\frac{1}{\sqrt{3}}}(c_+) &= \frac{\left(-\frac{1}{12}\right)\left(\frac{4}{9}\right)+\frac{1}{3}}{\left(-\frac{4}{3\sqrt{6}}\right)\left(\frac{2}{3}\right)^{\frac{3}{2}}} \\
&= \frac{\frac{8}{27}}{-\frac{8}{9\sqrt{6}}\cdot\sqrt{\frac{2}{3}}} \\
&= -1.
\end{aligned}$$

Using Lemma 6.2, since $N_\beta(c_+)$ decreases with respect to β, it follows that for $1/\sqrt{3} \leq \beta \leq 1$, $N_\beta(c_+) \leq -1$. Given that N_β is odd, this same argument shows that $N_\beta(c_-) \geq 1$ for the given range of β values.

From this point, the proof that c_+ and c_- converge to -1 and 1 respectively is very similar to that of Lemma 5.2 in the real case. □

Given our general picture of N_β and the fact that $N_\beta(c_+)$ increases in β for $0 < \beta \leq 1$, it appears that for some β value we will have $N_\beta(c_+) = N_\beta(c_-) = 0$. We compute this value

for future inspection:

$$
\begin{aligned}
0 &= \frac{3(c_+)^4 + (\beta^2-1)(c_+)^2 + \beta^2}{4(c_+)^3 + 2(\beta^2-1)(c_+)} \implies \\
0 &= 3(c_+)^4 + (\beta^2-1)(c_+)^2 + \beta^2 \\
&= 3(\frac{(1-\beta^2)^2}{36}) + (\beta^2-1)(\frac{1-\beta^2}{6}) + \beta^2 \\
&= \frac{1}{12}(1-\beta^2)^2 - \frac{1}{6}(1-\beta^2)^2 + \beta^2 \\
&= -\frac{1}{12}(1-\beta^2)^2 + \beta^2 \\
&= -\frac{1}{12}\beta^4 + \frac{7}{6}\beta^2 - \frac{1}{12} \\
&= -\frac{1}{12}\left(\beta^4 - 14\beta^2 + 1\right)
\end{aligned}
$$

Then, letting $A = \beta^2$, the quartic in parentheses vanishes iff

$$
\begin{aligned}
A &= \frac{14 \pm \sqrt{192}}{2} \\
&= 7 \pm 4\sqrt{3}\,.
\end{aligned}
$$

Then, since $(2-\sqrt{3})^2 = 7-4\sqrt{3}$, we see that $\beta = 2-\sqrt{3} \approx 0.2679491924$ is the only solution to $N_\beta(c_+) = N_\beta(c_-) = 0$ for $0 \le \beta \le 1$. As it turns out, this value will prove to be a very valuable landmark in our study of the λ-parameter plane.

While the orbits of c_- and c_+ converged quickly to roots in the case where λ was real, it is readily apparent that this is not the case when $\lambda = \beta i$.

Example 6.4 We will show that for a particular β value, $N_\beta(c_-) = c_+$ and $N_\beta(c_+) = c_-$. Since N_β is an odd function, and since $c_- = -c+$, we need only show that one of the previous equations holds true. Without loss of generality, we show that $N_\beta(c_+) = c_-$.

$$
\begin{aligned}
& N_\beta(c_+) = c_- \\
\implies\ & N_\beta(c_+) - c_- = 0 \\
\implies\ & N_\beta(c_+) + c_+ = 0.
\end{aligned}
$$

To simplify the following series of calculations, we let $A = \beta^2$. Thus, we have $c_\pm = \pm\sqrt{\frac{1-A}{6}}$. So,

$$
N_\beta(c_+) + c_+ = \frac{3(\sqrt{\frac{1-A}{6}})^4 + (A-1)(\sqrt{\frac{1-A}{6}})^2 + A}{4(\sqrt{\frac{1-A}{6}})^3 + 2(A-1)\sqrt{\frac{1-A}{6}}}.
$$

Setting this equal to zero yields

$$\begin{aligned}
0 &= 3(\frac{(1-A)^2}{36} - (\frac{(1-A)^2}{6}) + A + \frac{1-A}{6}(4(\frac{1-A}{6}) + 2(A-1)) \\
&= \frac{1}{12}(1-A)^2 - \frac{1}{6}(1-A)^2 + A + \frac{1}{9}(1-A)^2 - \frac{1}{3}(1-A)^2 \\
&= \frac{-11}{36}(1-A)^2 + A \\
&= \frac{-11}{36} + \frac{22}{36}A - \frac{11}{36}A^2 + A \\
&= \frac{-11}{36}A^2 + \frac{58}{36}A - \frac{11}{36}
\end{aligned}$$

Thus,

$$\begin{aligned}
A &= \frac{\frac{-58}{36} \pm \sqrt{\frac{2880}{1296}}}{\frac{-11}{18}} \\
&= \frac{29 \pm \sqrt{720}}{11}
\end{aligned}$$

Then, since we are only concerned with β values between 0 and 1, and $\beta^2 = A = \frac{29 \pm \sqrt{720}}{11}$, we have

$$\begin{aligned}
\beta^2 &= \frac{29 - 12\sqrt{5}}{11} \\
\implies \beta &= \sqrt{\frac{29 - 12\sqrt{5}}{11}}\,.
\end{aligned}$$

Therefore, taking the positive root, we have found that for $\beta \approx 0.4438656912$, the orbits of c_- and c_+ fail to converge to a root under Newton's method, since they lie on a superattracting cycle of period 2. This can be seen in Figure 6.2. □

Since it is virtually impossible to deal with higher iterations of N_β analytically, a computer program was written to help us uncover periodic points numerically. The full code for the program is given in the Appendix. The program works as follows. We are given some period n. Then, since we are assured that no periodic behavior occurs on the imaginary axis for $1/\sqrt{3} < \beta < 1$, the initial β value used by the program is 0.58, which is slightly greater than our landmark value $1/\sqrt{3}$. Then, we utilize the bisection method to find solutions to the equation $N_\beta^n(c_+) - c_+ = 0$. These solutions are superattracting periodic points of period n. By making small changes in β values (10^{-6}), we are able to straddle the x-axis and then iterate the bisection method 100 times to hone in on a solution. After these iterations are completed, β is decreased by 10^{-7}, and the process repeats itself. This is done until β reaches

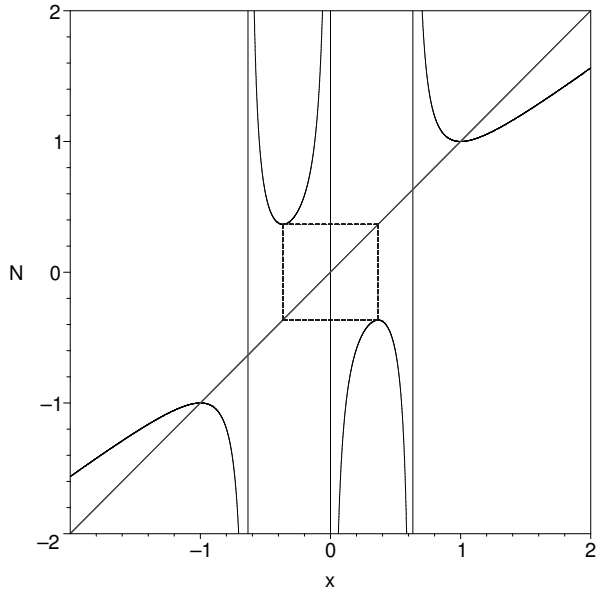

Figure 6.2: The orbit diagram for N_β with $\beta \approx 0.4438656912$.

n	β	n	β	n	β	n	β
2	0.4438657165	4	0.2158225775	5	0.2275660932	5	0.1125293225
2	0.3835689425	4	0.2113012969	5	0.2249682546	5	0.0917167962
3	0.2291103601	4	0.1134351641	5	0.1846443415	5	0.0570865125
3	0.1341462433	4	0.0616595671	5	0.1577119529	5	0.0298646167
4	0.3642913699	5	0.2299712598	5	0.1301919222		
4	0.3363839984	5	0.2296915054	5	0.1289675832		

Table 6.1: A table of all periodic points up to period 5 on the imaginary axis. These are approximations found using a bisection method program.

0. It should also be noted that error checking devices are also put into place to ensure we do not mistake poles for periodic points.

Using this bisection technique, our program gives a fairly exhaustive list of periodic points along the imaginary axis for $\beta \in [0, 1]$. It should be noted, however, that these are approximate values, subject to some round-off error. Nevertheless, according to this program, we find that our first periodic value, $\beta \approx 0.4438657165$, is confirmed by this program. Table 6.1 gives a comprehensive listing of the data that was found using this program.

Using this numerical tool, we observed some noteworthy behavior. For instance, a period doubling cascade to chaos is readily apparent from our data, and can be seen in Table 6.2. While the periodic points of period 256 are too numerous to list in their entirety, (letting the program run exhaustively on only period 16 yields 2,525 periodic points), the evidence exhibited in Table 6.2 certainly substantiates the claim to a period doubling cascade to chaos.

n	β
2	0.3835689425
4	0.3642913699
8	0.3601377606
16	0.3592396379
32	0.3590471745
64	0.359005928
128	0.3589970469

Table 6.2: A table exhibiting the period doubling behavior found on the imaginary axis.

To visualize the dynamical behavior of N_β, we created another computer program to plot the bifurcation diagram as a function of β. The vertical axis captures the long-term behavior of both critical points. As can be seen in Figures 6.3 and 6.4, our period doubling claim is further supported. Furthermore, our landmark β value of $\beta = 2 - \sqrt{3} \approx 0.268$ appears to be a divider for much of the interesting dynamical behavior. For β-values slightly smaller than $2 - \sqrt{3}$, the image of both critical points is close to the pole at zero and consequently, further iteration leads to convergence to either 1 or -1. As long as the image of the critical points is trapped close to the pole at zero, the long-term behavior will be convergence to roots. This explains the gap in the center of Figure 6.3.

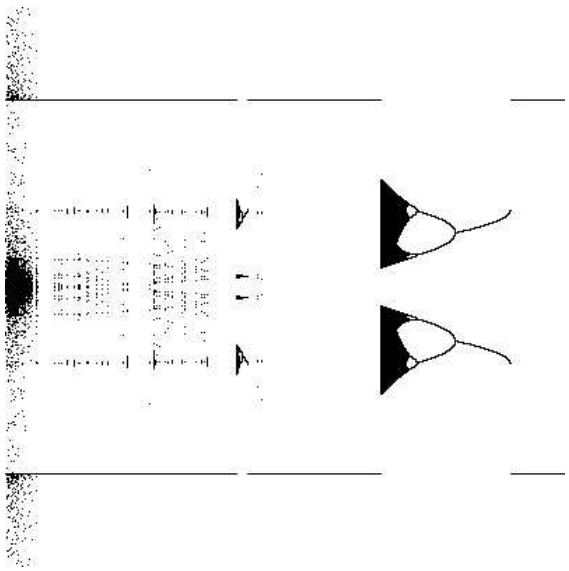

Figure 6.3: The bifurcation diagram for N_β showing the asymptotic behavior of both free critical points under iteration as a function of β. The horizontal line segments at the top and bottom of the figure are 1 and -1 representing β-values where the free critical points converge to those roots. The significant gap in the center of the figure begins at $\beta = 2 - \sqrt{3}$.

After compiling this extensive amount of data, some observations about the behavior

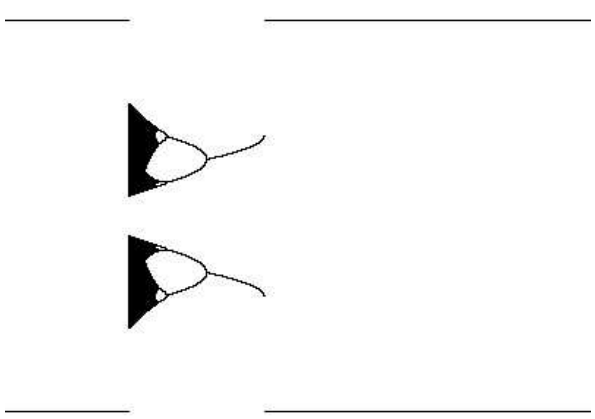

Figure 6.4: A zoom of the bifurcation diagram from Figure 6.3 for $0.268 \approx 2 - \sqrt{3} < \beta < 1/\sqrt{3} \approx 0.577$.

of these periodic points is warranted. For instance, when we computed the value $\beta \approx 0.4438656912$, we saw that c_+ and c_- existed on the same periodic orbit. However, as we will see, this is not always the case. Consider the only other period 2 value, $\beta \approx 0.3836689425$. In this case we see that c_+ and c_- lie on disjoint periodic orbits. See Figure 7.2 for a web diagram exhibiting this behavior. After realizing that two distinct types of behavior were possible for superattracting critical points, a second program was written that not only produced the periodic β values, but also labeled them according to the type of behavior exhibited by their periodic orbits. For reasons we will see in the next section, we refer to β-values for which c_+ and c_- lie on the same orbit as *bitransitive*, while values for which the orbits of the free critical points do not coincide are referred to as *disjoint*.

As we can see, β-values for which c_+ and c_- lie on the same orbit occur relatively rarely in contrast with the disjoint case. Note also that there do not exist β-values yielding a bitransitive case with odd period. In fact, this will never be the case.

Lemma 6.5 *For β-values corresponding to superattracting periodic points with odd period, c_+ and c_- do not lie on the same orbit.*

Proof Without loss of generality, suppose $N_\beta^n(c_+) = c_+$ for some primitive odd period n. Further, suppose that c_- and c_+ lie on the same periodic orbit. This implies that $N_\beta^k(c_+) = c_-$ for some $k < n$. However, since N_β is an odd function and $c_+ = -c_-$, this also implies that $N_\beta^k(c_-) = c_+$. Then, by a simple substitution we have

$$\begin{aligned} N_\beta^k(c_+) &= N_\beta^k\left(N_\beta^k(c_-)\right) \\ &= N_\beta^{2k}(c_-). \end{aligned}$$

n	β	Type	n	β	Type
2	0.4438657165	Bitransitive	5	0.2296915054	Disjoint
2	0.3835689425	Disjoint	5	0.2275660932	Disjoint
3	0.2291103601	Disjoint	5	0.2249682546	Disjoint
3	0.1341462433	Disjoint	5	0.1846443415	Disjoint
4	0.3642913699	Disjoint	5	0.1577119529	Disjoint
4	0.3363839984	Disjoint	5	0.1301919222	Disjoint
4	0.2158225775	Bitransitive	5	0.1289675832	Disjoint
4	0.2113012969	Disjoint	5	0.1125293225	Disjoint
4	0.1134351641	Disjoint	5	0.0917167962	Disjoint
4	0.0616595671	Disjoint	5	0.0570865125	Disjoint
5	0.2299712598	Disjoint	5	0.0298646167	Disjoint

Table 6.3: The table of periodic β values listed along with their type of dynamical behavior.

However, we also have that $N_\beta^k(c_+) = c_-$. Therefore, we see that c_- is periodic with *even* period $2k$. Again, given the symmetry of N_β, this also implies that c_+ is periodic with even period $2k$. Since $k < n$ and n is the least period of the orbit, it follows that $2k$ must equal n. But this contradicts the fact that n is odd. Thus, we see that for β-values with odd period superattracting orbits, c_+ and c_- cannot lie on the same orbit. □

Given Lemma 6.5, we focus our attention on periodic values of β with even period. It is apparent through our collection of data that some type of bifurcation exists between the two period 2 values, $\beta = 0.4438657165$ and $\beta = 0.3835689425$. The first of these points exists as the first special case we discovered where the two free critical points lie on the same orbit. However, as we proceed down the imaginary axis to the second period 2 point, this behavior is lost. In order to get a better understanding of why this occurs, we look to the second iterate of N_β. As mentioned earlier, dealing with even the second iterate of N_β analytically is very difficult, so for this reason we turn to Maple to get a graphical understanding of the second iterate.

We notice two distinct types of bifurcations occurring at our bitransitive and disjoint periodic values. For instance, at the first bitransitive value, $\beta = 0.4438656192$, we notice that the portion of the graph that breaks through the graph is parabolic in nature, which lead to a saddle-node bifurcation. However, as we decrease β slightly, the graph of the second iterate of N_β begins to alter in a qualitative way. The graph goes from having a parabolic-like shape to a fourth degree-like "w" shape. When we look at the second iterate for $\beta = 0.3835689425$, our second period 2 value, we notice the resulting behavior of a pitchfork bifurcation (see Figure 6.5). This behavior seems to generalize well to all of the periodic values produced by our program.

After examining the graphs of the higher iterates of N_β via Maple, it seems as though each bitransitive value for β is a periodic value that results from a saddle node bifurcation, while our disjoint cases are the result of pitchfork bifurcations. For example, this exact behavior can be seen once again in the case of our bitransitive period 4 value, $\beta = 0.2158225775$ and

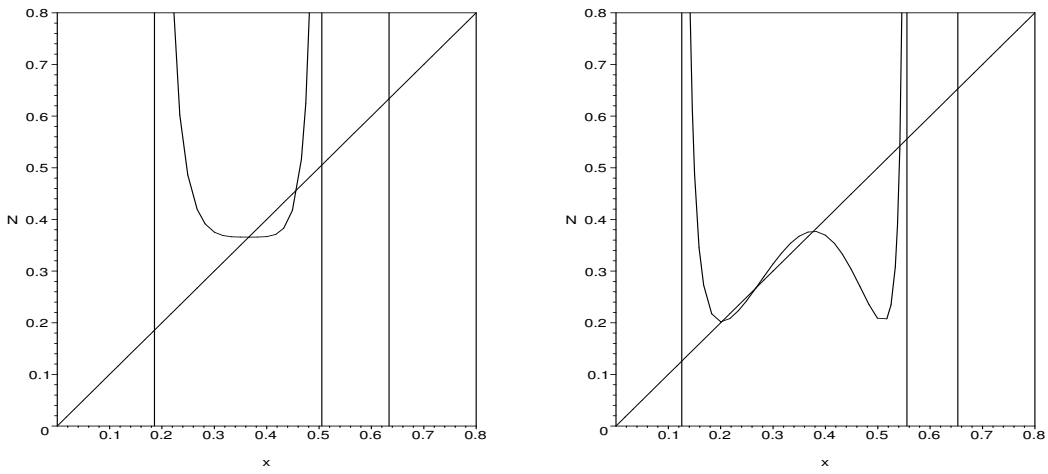

Figure 6.5: Graphs of the second iterate of N_β. To the left we have the bitransitive case and to the right we have the disjoint case for period 2 values of β.

its closest disjoint neighbor, $\beta = 0.2113012969$. This can be seen in Figure 6.6.

Although we are only dealing with a cross section of the complex plane in this case, the peculiar behavior that we have observed of N_β proves to be very insightful for the general case. As we shall see, the type of superattracting periodic orbit that occurs on the imaginary axis provides us with valuable information about what we should expect, and what we see, when we move λ off of the imaginary axis and into the first quadrant of the parameter plane.

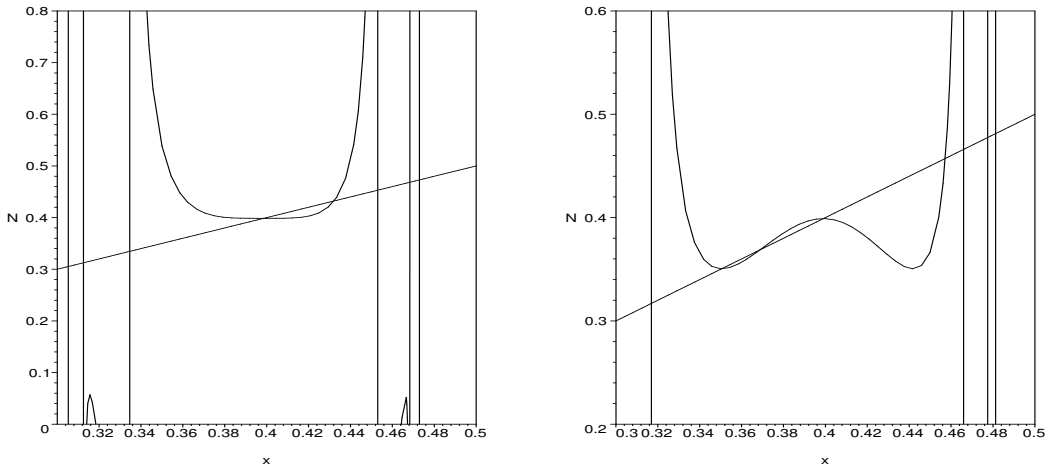

Figure 6.6: Graphs of the fourth iterate of N_β. To the left ($\beta \approx 0.2158225775$) we have the bitransitive case and to the right ($\beta \approx 0.2113012969$) we have the disjoint case for period 4 orbits.

7 The General Case: Moving off Both Axes

Unfortunately, once we move off both the real and imaginary axes, we are no longer able to work with the simplified versions of our Newton function as we had been doing in the previous sections. Recall from Section 4 that the definition of the Newton function in the general case is

$$N_\lambda(z) = \frac{3z^4 - 4Re(\lambda)z^3 + (|\lambda|^2 - 1)z^2 + |\lambda|^2}{4z^3 - 6Re(\lambda)z^2 + 2(|\lambda|^2 - 1)z + 2Re(\lambda)}.$$

Similarly, we must now deal with the general definition of the free critical points as well

$$c_\pm = \frac{\mathrm{Re}(\lambda) \pm \sqrt{(\mathrm{Re}\,\lambda))^2 - \frac{2}{3}(|\lambda|^2 - 1)}}{2}.$$

It is important to note, however, that we have only two free critical points. The λ-parameter plane for N_λ is shown in Figure 7.1. Here we follow the orbits of both critical points under iteration of N_λ for varying λ. The deeper red colors indicate faster convergence to a root, while the lighter colors indicate a slower convergence. Black represents λ-values where one or both critical points fails to converge to a root after some specified number of iterations. The same color scheme is used for all remaining figures in the parameter plane.

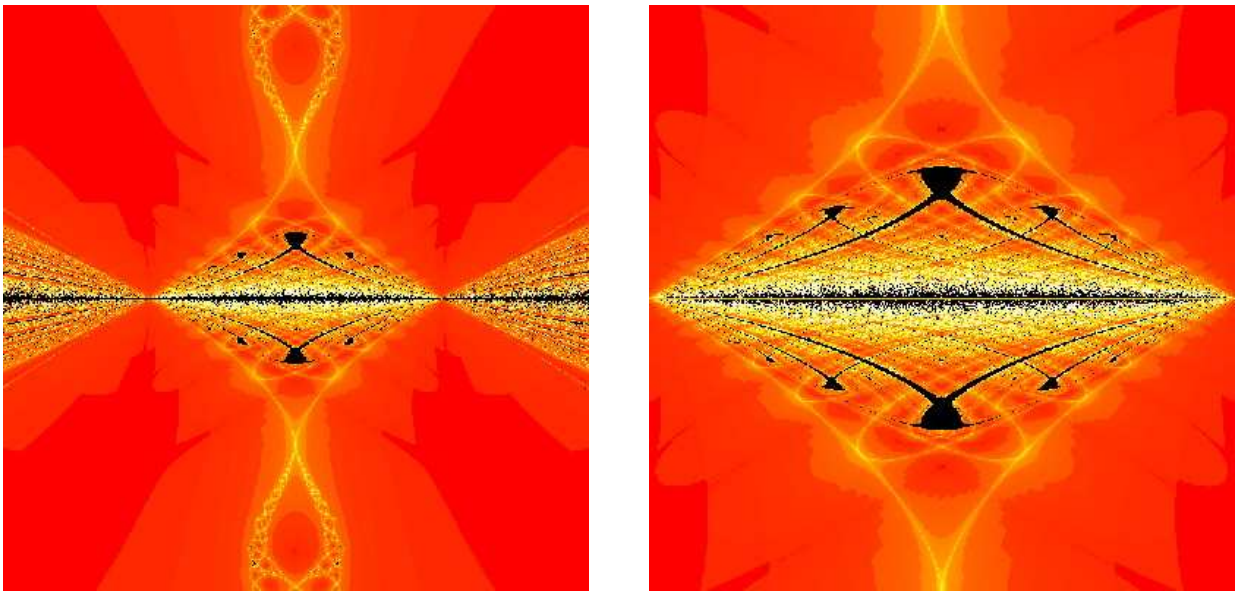

Figure 7.1: The λ-parameter plane for N_λ. On the left, the window is $[-2, 2] \times [-2i, 2i]$ and on the right the window is $[-1, 1] \times [-i, i]$.

We find a striking similarity between the dynamical behavior of N_λ and that of a general complex cubic polynomial. This is not entirely surprising, since a general cubic map has two free critical points, which is exactly what we have for N_λ.

In [11], Milnor explains in detail the different types of dynamical behavior found in the general cubic case. According to Milnor, there are four distinct classifications of hyperbolic

components in the cubic case. These four cases are referred to as *adjacent critical points*, *bitransitive*, *capture*, and *disjoint periodic sinks*. As suggested by our discussion in Section 6, only two of these cases are relevant for our study of N_λ: the bitransitive and disjoint periodic sinks cases (see Figure 7.2).

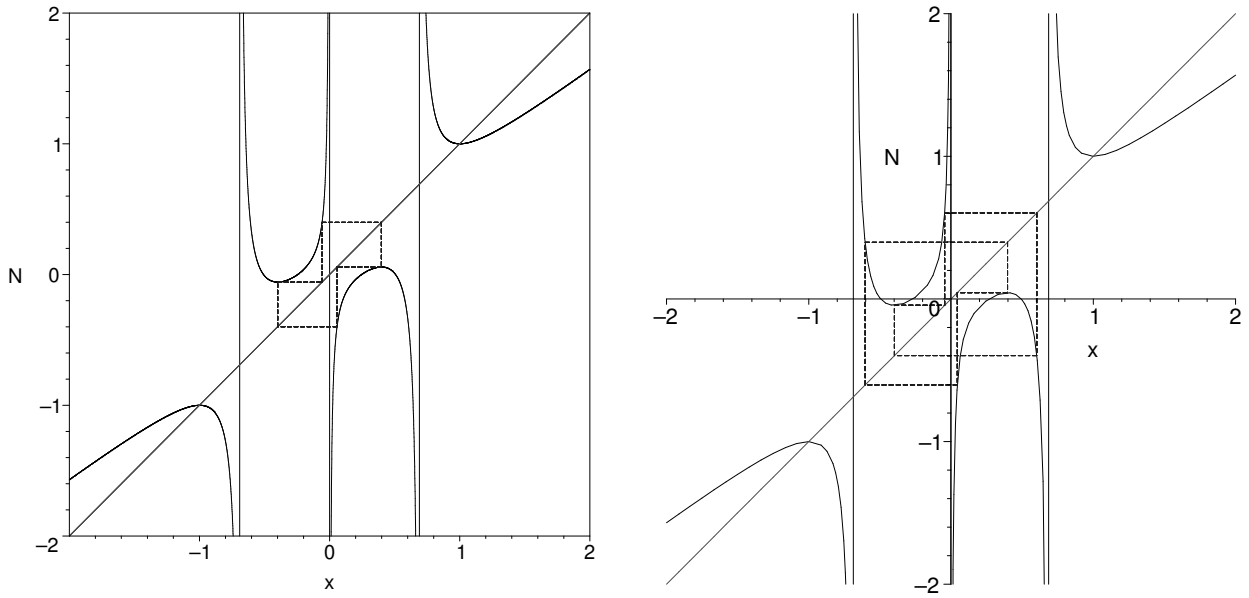

Figure 7.2: An example of bitransitive behavior (left) for a period 4 value of β and of disjoint periodic sink behavior (right) for a period 3 value of β.

Definition 7.1 Let U be the entire attractive basin for N_λ. We find the following two types of dynamical behavior in the λ-parameter plane.
Bitransitive The two free critical points belong to different components U_0 and U_1 of U, but there exist integers p, $q > 0$ such that $f^p(U_0) = U_1$ and $f^q(U_1) = U_0$. We assume that p and q are primitive, so that both U_0 and U_1 have period $p + q$.
Disjoint Periodic Sinks: The two free critical points belong to different components U_0 and U_1, where no forward image of U_1 is equal to U_0. In this case, each of the two components U_0 and U_1 must be periodic, though their periods p and q may be different.

Milnor goes on to explain which types of dynamical figures we should expect from each of these cases. In the Bitransitive case, there are two possibilities, either the swallow configuration (indicated in Figure 7.3) or a three pointed configuration which Milnor terms the "tricorn" (see Figure 7.4). The tricorn actually contains three embedded copies of the Mandelbrot set, where the cusp of each has been stretched out over a triangular region, joining them in this peculiar fashion. In the case of Disjoint Periodic Sinks, we are told to expect either a product configuration (see Figure 7.5) or an actual copy of the Mandelbrot set itself (see Figure 7.6).

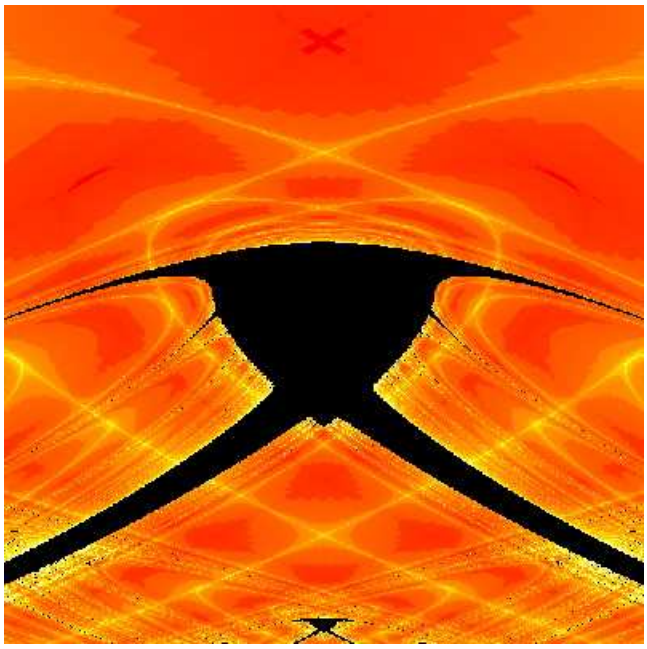

Figure 7.3: A swallow configuration in the parameter plane for N_λ centered at our bitransitive value $\lambda \approx 0.443865$i.

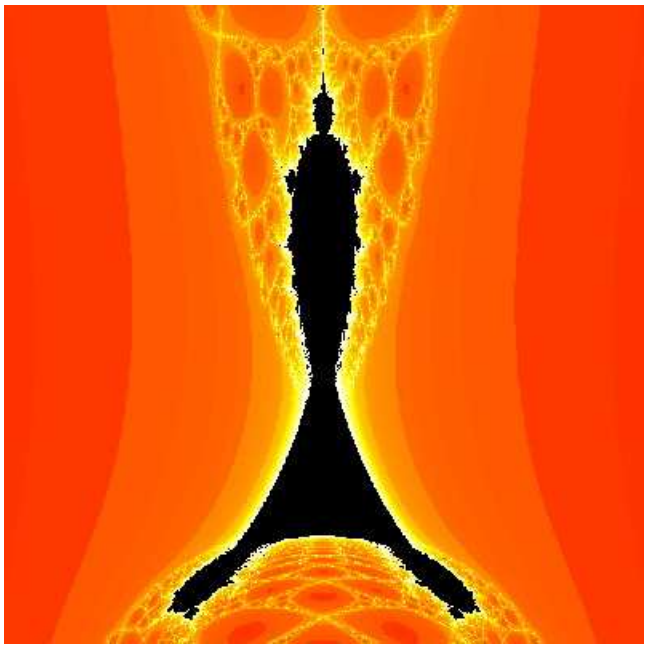

Figure 7.4: A tricorn in the parameter plane for N_λ centered at $1/0.443865\,i$, the inversion of the bitransitive value from Figure 7.3.

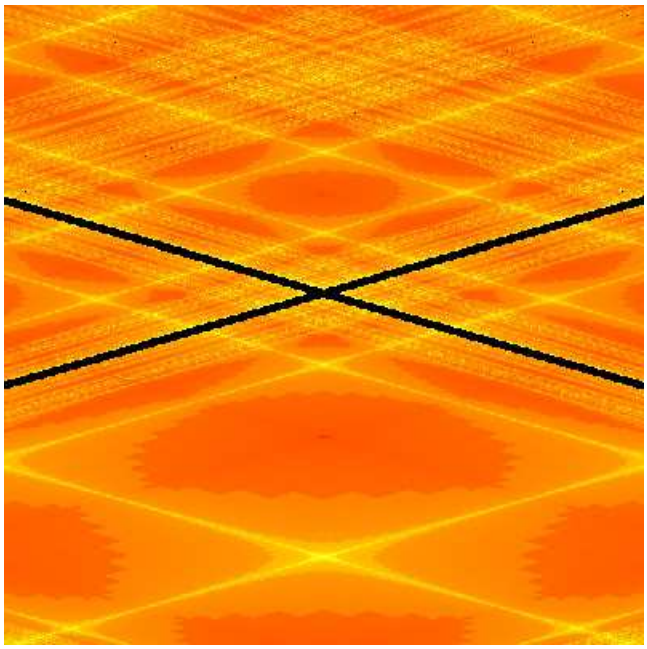

Figure 7.5: A zoom of the parameter plane near a disjoint periodic value $\lambda \approx 0.2291\text{i}$, exhibiting a product-like configuration.

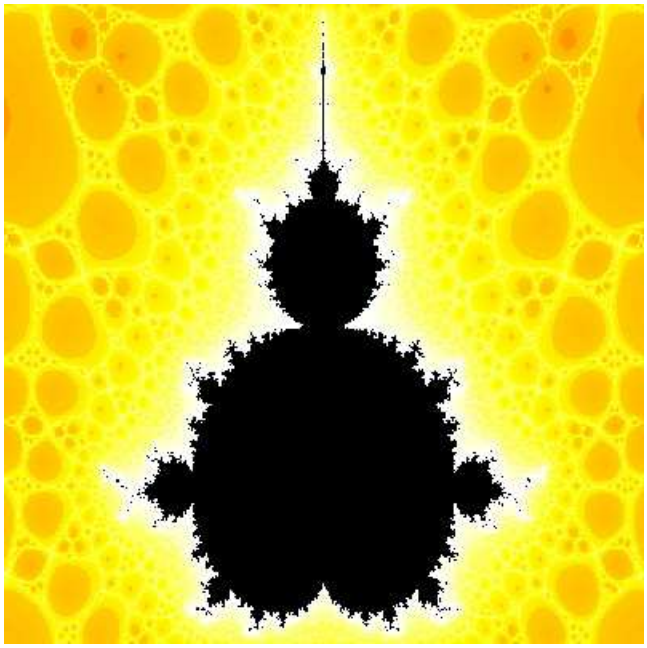

Figure 7.6: A Mandelbrot-like set in the parameter plane centered at $1/0.2291\,i$, the inversion of the disjoint periodic sink value found in Figure 7.5.

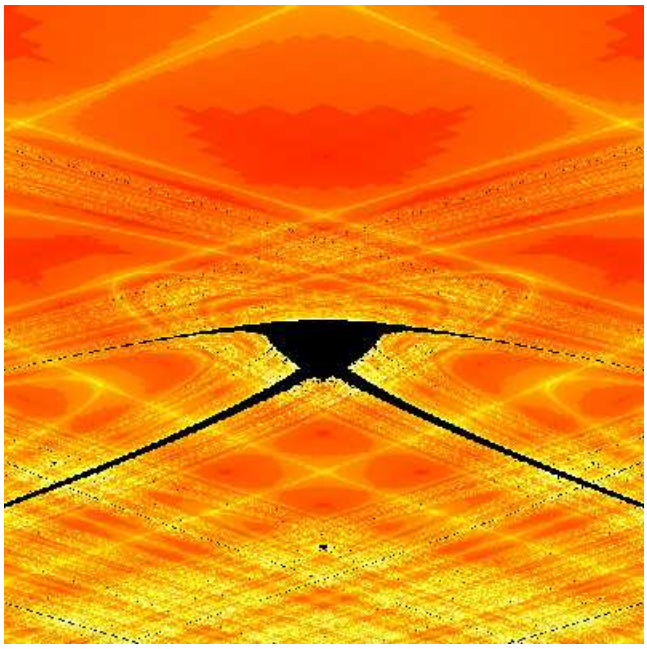

Figure 7.7: The swallow which contains our second bitransitive value, $\lambda \approx 0.2158\text{i}$.

Now, while we are not assured analytically where this behavior occurs off of the real and imaginary axes, we do have bitransitive and disjoint periodic sink values of β from our work on the imaginary axis. Thus, we can attempt to locate these points in our more detailed pictures of the complex planes, in hopes of explaining the general dynamical behavior of N_λ. Furthermore, it should be noted that Milnor's results cannot be directly correlated to our work on Newton's method simply due to the fact that we are dealing with a *cubic-like* map and not an actual cubic polynomial.

Nevertheless, if we recall the hyperbola from Section 4 that helped us to define the nature of our free critical points, we find some interesting parallels between our study and Milnor's. For instance, all of our bitransitive values that occur between the curves of the hyperbola, where the free critical points are real, result in a swallow configuration in the parameter plane (Figure 7.7). When these bitransitive values are inverted, and the free critical points become complex, the result is a tricorn (Figure 7.8). Similarly, for our disjoint periodic sink values, whenever the free critical points are real, the result is a product configuration. Upon inversion, making our free critical points complex, we find Mandelbrot-like sets. Thus, it seems that the nature of our free critical points is the determining factor behind which of the dynamical components we should expect in the λ-parameter plane. However, there are some exceptions to this. For instance, there are disjoint periodic sink values that are very close to bitransitive values and are contained inside of a swallow configuration in the parameter plane. After studying this behavior extensively via computer experiments, we developed the following conjecture:

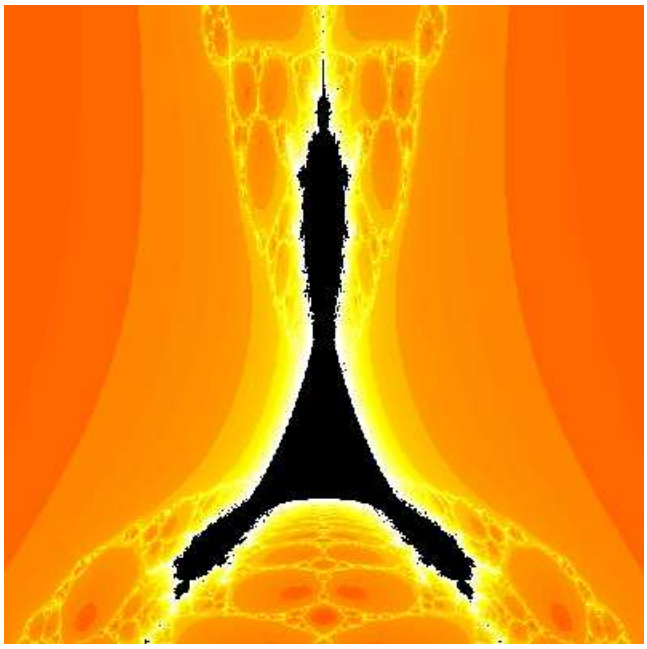

Figure 7.8: Another tricorn, this one containing the inversion of the bitransitive value in Figure 7.7.

Conjecture: Each bitransitive λ-value where $\mathrm{Re}(\lambda) = 0$ and $0 \leq \mathrm{Im}(\lambda) \leq 1$ is the center of a swallow configuration in the parameter plane.

There is certainly a good deal more to be said about the λ-parameter plane, though we believe that the key to understanding exactly what is going on relies on a better comprehension of the general cubic case and how that information can be related to our Newton map N_λ.

8 Extending Work on this Family Further

Throughout this work, our main concern has been examining the λ-parameter plane and trying to categorize the various types of behavior that we find there. Of course, while we were able to explain the existence of certain figures in correlation with Milnor's work on cubics, there is still much of the λ-parameter plane that is left unexplained. It should be noted however, that we were able to define the curves that bound the diamond shaped region we find about the origin in the parameter plane (see Figure 7.1). These curves are the result of λ-values for which both p' and p'' vanish. That is, λ-values for which the free critical points of N_λ happen to coincide with one of its poles. A contour plot of these values then reveals the four curves that bound the region of what appears to be chaotic dynamical behavior about the origin.

We then suspect that the self-replicating "leaves" that can be seen approaching the real axis in Figure 7.1 are preimages of the outermost bounding curves. Unfortunately, taking

preimages of our Newton map is an extremely difficult task. While our suspicions can be substantiated by numerical evidence, we are unable to define these curves analytically to ensure that they are in fact preimages of our defined outermost curves.

Switching gears all together, an alternative approach to this family of polynomials that has not yet been mentioned has to do with the dynamical plane. While we explained the interesting behavior that occurs in the parameter plane at some length, it should be noted that we have also observed some noteworthy phenomena in the dynamical plane. For instance, as mentioned in Section 3, the Nearest Root Principle does not hold with respect to our family of polynomials.

Furthermore, in studying the dynamical plane, we are able to get a better sense of exactly how successful (or unsuccessful) Newton's method can become in this particular instance. While only studying the dynamical plane briefly, we were able to draw a few parallels from our work with the parameter plane.

For example, the important value of $\lambda \approx 0.4438656912i$, turns out to be quite useful again in the dynamical plane, as it appears to be in a neighborhood of λ values for which Newton's method fails on a fairly substantial portion of the complex plane. This can be seen in Figure 8.1. This image was created by iterating each point in the complex plane under N_λ where $\lambda = 0.4438656912i$. If the initial seed converges to one of the four roots, $\lambda, \overline{\lambda}, 1$, or -1, then it is colored yellow, red, blue or purple, respectively. If the initial seed does not converge to within 10^{-6} of any root after 100 iterations, it is colored black. This same scheme is used for the remaining figures in the dynamical plane.

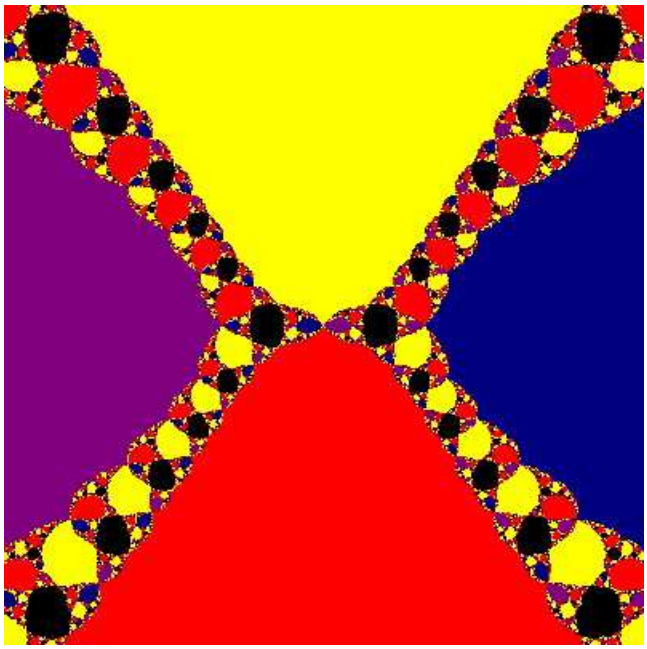

Figure 8.1: The dynamical plane for $\lambda \approx 0.4438656912$i, a bitransitive value.

However, once we cross our landmark value of $\lambda = (2 - \sqrt{3})$i, we see that the size of the attractive basin for extraneous attractive cycles decreases significantly. In Figure 8.2,

which depicts the Julia set for a disjoint period 3 value of λ, it is difficult to detect any black regions at all. Naturally, it would be worthwhile to try and understand this phenomenon more rigorously.

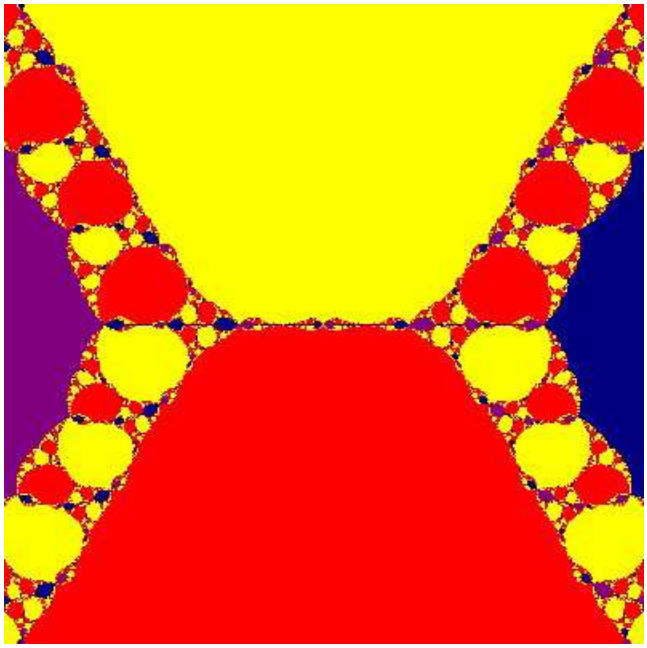

Figure 8.2: The dynamical plane for $\lambda \approx 0.2291\text{i}$, a disjoint periodic value of odd period.

Through our brief look at the dynamical plane, however, we were able to confirm some of our beliefs having to do with the figures we found in the parameter plane. For instance, recall the definition of bitransitive behavior from Section 7. In this case, the open sets U_1 and U_2, each containing one of the free critical points, become interchanged after some number of iterations of N_λ. However, in the case of disjoint periodic sink behavior, something prevents this from happening. The nature of these neighborhoods U_1 and U_2 about the free critical points changes in some qualitative way.

As it turns out, after viewing the dynamical plane for our known bitransitive and disjoint periodic sink values of λ, some conclusions can be drawn. In the bitransitive case, each component in the attractive basin is a quasi-circular, connected set. This makes sense since each component is mapped onto another component at some periodic interval. Examples of this can be seen in Figures 8.1 and 8.3.

In the case of disjoint periodic sink behavior, the components in the attractive basin appear to have been pinched off and separated, most likely as the result of a pole being introduced (see Figures 8.4 and 8.5). This limits the way in which we are able to choose U_1 and U_2 about the free critical points of N_λ. As a result, forward iterates of U_1 and U_2 never intersect, resulting in the disjoint behavior discussed previously in Section 6.

While this confirms our theories on bitransitive and disjoint periodic sink behavior, the more interesting question still remains. That is, while we are familiar with some of the dynamical components that we find in the λ-parameter plane, are we assured that these are

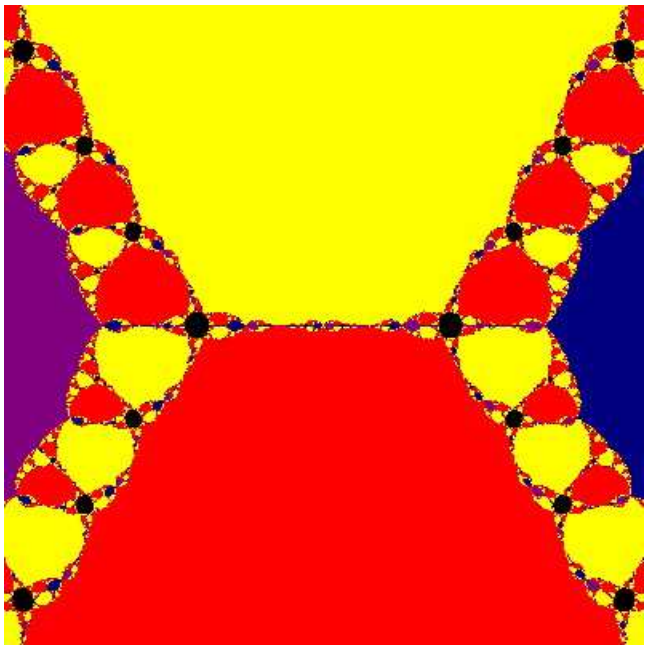

Figure 8.3: The dynamical plane for $\lambda \approx 0.2158i$, a bitransitive value.

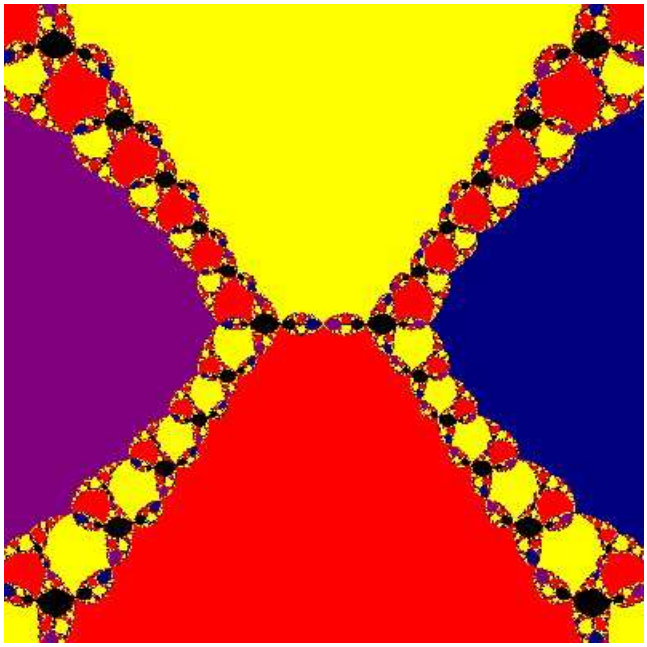

Figure 8.4: The dynamical plane for $\lambda \approx 0.3835i$, a disjoint periodic value.

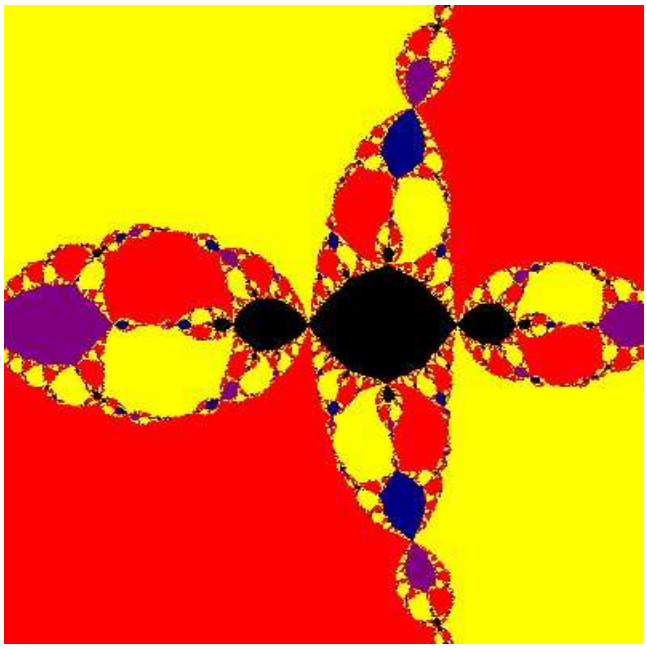

Figure 8.5: The dynamical plane for $\lambda \approx 0.2113$i, a disjoint periodic value.

the only types of components that can arise in the case of Newton's method applied to a fourth degree polynomial?

9 Appendix

```
/*****************************************************************
 *
 *  Author:  Trevor O'Brien
 *  Date:    June 25, 2004
 *  Purpose: Given a polynomial of the form $(x^2 - 1)(x^2 + \beta^2)$,
 *           this program will determine which critical points are periodic
 *           points of the corresponding Newton's function.  The user
 *           may input the period they are looking for, and an upper bound
 *           for the initial guess.  A bisection method is used to narrow
 *           in on the zero of the function $N^k(c) - c$.  Such zeros are
 *           periodic points of $N(x)$ with period $k$.
 *
 *****************************************************************/

\#include <iostream>
\#include <math.h>
\#include <iomanip>
```

```
float Newton(float beta, int period);
void  GetInfo(int\& period);
void  BisectionMethod(int period, float beta, float\& betaDifference,
      float\& midpoint);
void  DisplayResults(float betaDifference, float midpoint, int\& count, int period);
bool  Bitransitive(float beta, int period);

int main()
{
   float beta, midpoint, betaDifference;
   int period, count;
   beta = 0.58;
   count = 0;

   GetInfo(period);

   while(beta > 0){
      BisectionMethod(period, beta, betaDifference, midpoint);
      DisplayResults(betaDifference, midpoint, count, period);
      beta = midpoint - .000001;
   }

   cout << endl << setw(50) << "Goodbye for now!" << endl;
   return 0;
}

void BisectionMethod(int period, float beta, float\& betaDifference,
     float\& midpoint)
{
   float newBeta = beta;
   float value = Newton(beta, period);

   do{
      newBeta = newBeta - 0.000001;
   }while((Newton(newBeta, period)*value) > 0);

   midpoint = (beta + newBeta)/2;

   for(int i = 0; i < 100; i++){

      if(Newton(midpoint, period)*Newton(beta, period) < 0){
 newBeta = midpoint;
 midpoint = (beta + midpoint)/2;
```

```
 betaDifference = Newton(midpoint, period) - Newton(beta, period);
      }else{
 beta = midpoint;
 midpoint = (newBeta + midpoint)/2;
 betaDifference = Newton(midpoint, period) - Newton(newBeta, period);
      }
   }
}

void GetInfo(int\& period){
   cout << endl << "Please enter period:  ";
   cin  >> period;
   cout << endl;
   cout << "And the results..." << endl << endl;
   cout << setw(35) << "The periodic points of period " << period
<< " are:" << endl << endl;
}

float Newton(float beta, int period)
{
   float estimate, x, criticalPt;

    criticalPt = sqrt((1 - pow(beta, 2))/6);
    x = criticalPt;

    for(int j = 0; j < period; j++){

      estimate = (3*pow(x, 4) + (pow(beta, 2) - 1)*pow(x, 2) + pow(beta,2))/
 (4*pow(x, 3) + 2*(pow(beta, 2) - 1)*x);

      x = estimate;
   }

   return x - criticalPt;
}

void DisplayResults(float betaDifference, float midpoint, int\& count, int period)
{
   if(fabs(betaDifference) < 1){
      count++;
      cout << count << ". ";
      cout << setw(21) << setprecision(10) << midpoint;
      if(Bitransitive(midpoint, period)){
 cout << " BITRANSITIVE";
```

```
      }
      cout << endl;
   }

}

bool Bitransitive(float beta, int period){

   float estimate, criticalPt, x;
   bool bitransitive = false;

   criticalPt = sqrt((1 - pow(beta, 2))/6);
   x = criticalPt;

   for(int j = 0; j < period; j++){

      estimate = (3*pow(x, 4) + (pow(beta, 2) - 1)*pow(x, 2) + pow(beta,2))/
 (4*pow(x, 3) + 2*(pow(beta, 2) - 1)*x);

      x = estimate;
      if(fabs(x + criticalPt) < .000001){
 bitransitive = true;
      }
   }

   return bitransitive;
}
```

References

[1] P. Blanchard, "Complex analytic dynamics on the Riemann sphere," Bull. Amer. Math. Soc. (New Series) **11**, 85-141 (1981).

[2] P. Blanchard, "The dynamics of Newton's method," *Complex Dynamical Systems* (Cincinnati, OH, 1994), Proc. Sympos. Appl. Math., 49, AMS, Providence, RI, 139-154 (1994).

[3] J. H. Curry, L. Garnett and D. Sullivan, "On the iteration of a rational function: Computer experiments with Newton's method," Communications in Mathematical Physics **91**, 267-277 (1983).

[4] R. L. Devaney, *An Introduction to Chaotic Dynamical Systems,* 2nd edition, (Addison-Wesley Publishing Company, Inc., Redwood City, 1989).

[5] A. Douady and J. H. Hubbard, "On the dynamics of polynomial-like mappings," Annales Scientifiques de L'Ecole Normal Superieure, 4^{e} serie, t. 18, 287-343 (1985).

[6] P. Fatou, "Sur les équations fonctionelles," Bull. Soc. Math. France **47**, 161-271 (1919); **48**, 33-94 and 208-314 (1920).

[7] R. A. Holmgren, *A First Course in Discrete Dynamical Systems,* 2nd edition, (Springer-Verlag New York, Inc., New York, 2000).

[8] J. Hubbard, D. Schleicher and S. Sutherland, "How to find all roots of complex polynomials by Newton's method," Inventiones Mathematicae **146**, No. 1, 1-33 (2001).

[9] G. Julia, "Memoire sur l'itération des fonctions rationelles," J. Math. pures et app. **8**, 47-245 (1918).

[10] J. Milnor, *Dynamics in One Complex Variable: Introductory Lectures,* 2nd edition, (Friedrick Vieweg and Sohn, Braunschweig, 2000).

[11] J. Milnor, "Remarks on iterated cubic maps," Experimental Mathematics **1**, No. 1, 5-24 (1992).

[12] G. E. Roberts and J. Horgan-Kobelski, "Newton's versus Halley's method: A dynamical systems approach," International Journal of Bifurcation and Chaos **14**, No. 10, 3459-3475 (2004).

[13] S. Sutherland, "Finding roots of complex polynomials with Newton's method", Doctoral Thesis, Boston University (1989).