Newton's versus Halley's Method: A Dynamical Systems Approach

Gareth E. Roberts* Jeremy Horgan-Kobelski†

October 8, 2003

Abstract

We compare the iterative root-finding methods of Newton and Halley applied to cubic polynomials in the complex plane. Of specific interest are those "bad" polynomials for which a given numerical method contains an attracting cycle distinct from the roots. This implies the existence of an open set of initial guesses whose iterates do not converge to one of the roots (ie. the numerical method fails). Searching for a set of bad parameter values leads to Mandelbrot-like sets and interesting figures in the parameter plane. We provide some analytic and geometric arguments to explain the contrasting parameter plane pictures. In particular, we show that there exists a sequence of parameter values λ_n for which the corresponding numerical method has a superattracting n cycle. The λ_n lie at the centers of a converging sequence of Mandelbrot-like sets.

1 Introduction

The numerical root-finding methods of Newton and Halley are investigated from a dynamical systems perspective. Our primary focus is to determine when a given method fails, that is, when does there exist an open set of initial seeds which fails to converge to a solution under iteration? Such a phenomenon occurs whenever an attracting cycle exists other than the roots of the equation. In the case of quadratic polynomials, both methods succeed perfectly well, with the only failure occurring along the perpendicular bisector of the segment joining the two roots. However, when applied to cubics, the methods become considerably more complicated and failure can occur on an open set of initial guesses.

By studying each method applied to the family of cubic complex polynomials

$$p_\lambda(z) = (z-1)(z+1)(z-\lambda),$$

we numerically find the presence of Mandelbrot-like sets in the λ-parameter plane corresponding to polynomials for which the given method has an extraneous attracting cycle. Exploiting the symmetry in this family, we offer some straight-forward analytical arguments for the existence and location of these Mandelbrot-like sets. In particular, we prove that each numerical method contains parameter values λ_n for which there exists superattracting cycles of period $n \geq 2$. Each of these parameter values lies at the center of the main cardioid of a Mandelbrot-like set M_n. As $n \to \infty$, the sets M_n coalesce along the imaginary axis, shrinking in size as they converge towards $\lambda = \sqrt{3}\,i$ (Newton) or $\lambda = i$ (Halley). We also contrast the two methods from a geometric perspective, focusing on qualitative ideas to explain the parameter plane structure in each case.

The discovery of Mandelbrot-like sets for Newton's method began with the computer experiments of Curry, Garnett and Sullivan [5]. By studying Newton's method applied to a particular class of

*Department of Mathematics and Computer Science, 1 College St., College of the Holy Cross, Worcester, MA 01610. Email: groberts@radius.holycross.edu.

†Department of Applied Mathematics, University of Colorado, Boulder

complex cubic polynomials, the authors surprisingly found the presence of Mandelbrot-like sets in the parameter plane. The appearance of these sets was explained by Douady and Hubbard [7] in their crucial development of the theory of polynomial-like mappings. In certain neighborhoods, Newton's method (although a rational map of degree 3 for the case of cubic polynomials) is conjugate to a polynomial map of degree 2 (in the sense that points have two local pre-images). This explains why Julia sets for Newton's method have figures resembling certain Julia sets from the quadratic family $z \mapsto z^2 + c$. Moreover, Douady and Hubbard prove that branched covers of the Mandelbrot set will appear in the parameter plane. Blanchard gives a nice exposition concerning these works in [3].

A typical geometric feature of both numerical methods is the **nearest root principal**, where initial seeds iterate to the closest root. This principal precisely summarizes the dynamics when each method is applied to a quadratic polynomial. The method succeeds when the initial guess is closest to one of the roots and fails when the initial guess is equidistant between roots. Both methods leave the perpendicular bisector between the two roots invariant, as if unable to choose between them. Although the nearest root principal can break down in the case of cubic polynomials, it is still useful for predicting where problems in the numerical method may arise. For example, in the case of Newton's method, problems can occur when the triangle formed by the roots is isosceles. In fact, the *only* place where Newton's method applied to cubics fails on an open set, is at or very near the isosceles case. It is these cases where an attracting cycle exists other than those provided by the roots.

It is interesting to note that Schröder's iteration formulas, another well-known numerical method, do not obey the nearest root principal when applied to quadratic polynomials. Our primary reason for choosing Halley's method instead of other higher-order methods is that it shares this nice geometric property with Newton's method.

Acknowledgments: Much of the research for this paper was supported by NSF-VIGRE grant DMS-9810751. The authors would like to thank James Curry for suggesting the topic and for many interesting discussions regarding this subject. We would also like to thank Bruce Peckham for insightful comments regarding the presentation of this work.

2 Newton's and Halley's Method

2.1 Preliminaries

Let $p(z)$ be a polynomial from $\bar{C}$ to $\bar{C}$. Throughout the paper we will assume that $p(z)$ has distinct roots. Both Newton's and Halley's root-finding methods applied to $p(z)$ are equivalent to iterating a rational function on $\bar{C}$. The *Julia set* for a rational map is the closure of the set of repelling periodic points [2]. This is an invariant and perfect set on which the map is chaotic. Any neighborhood of a point in the Julia set is mapped under iteration to cover all of the extended complex plane except at most two points. The complement of the Julia set, the *Fatou set*, is where the tame dynamics occurs (eg. attracting cycles and their basins of attraction).

The roots of $p(z)$ and their basins of attraction are in the Fatou set. These basins represent places where the numerical method works. There are two ways in which the numerical method may fail. First, any initial seed z_0 chosen in the Julia set will never converge to a root of $p(z)$. However, any neighborhood of such a z_0 contains points which converge to any of the roots, so a small perturbation of z_0 will lead to convergence. A more dangerous phenomenon occurs when an attracting cycle exists other than the roots. This cycle is necessarily in the Fatou set and more importantly, its basin of attraction is an entire region in $\bar{C}$ for which initial seeds never converge to a root. In this case, a small perturbation of a failing initial seed may *not* lead to convergence to a root.

Newton's method for finding the roots of $p(z)$ is equivalent to iterating the rational map

$$N_p(z) = z - \frac{p(z)}{p'(z)}. \tag{1}$$

The second-order method is based on approximating the polynomial $p(z)$ by a linear approximation. If $p(z)$ is degree d and has distinct roots, then $N_p(z)$ is a rational map of degree d. It is clear that roots of $p(z)$ correspond to fixed points of $N_p(z)$. In fact, the only fixed point of Newton's method other than a root is the point at infinity. A short computation yields

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

which shows that the roots are superattracting. It is important to note that inflection points of $p(z)$ are critical points of $N_p(z)$.

Halley's method, sometimes referred to as Bailey's method or Lambert's method (see [8]) is identical to K_3 of the König iteration family [9]. Applied to $p(z)$, this method is equivalent to iterating the rational map

$$H_p(z) = z - \frac{p(z)}{p'(z) - \frac{p(z)\,p''(z)}{2p'(z)}}. \tag{2}$$

Halley's method is third-order, based on approximating $p(z)$ by a quadratic function as opposed to a linear one. This accounts for the improved rate of convergence. It can be derived using an osculating hyperbola and is therefore often referred to as the method of tangent hyperbolas [1, 11, 13]. If $p(z)$ is degree d and has distinct roots, then $H_p(z)$ is a rational map of degree $2d-1$. As in the case of Newton's method, the roots of $p(z)$ are fixed points of $H_p(z)$, although other fixed points exist as well. Since we are assuming that the roots of $p(z)$ are distinct, the critical points of $p(z)$ are also fixed points under Halley's method (even if they are repeated roots of $p'(z)$).

The derivative of Halley's method is

$$H_p'(z) = -\frac{p(z)^2\,S[p](z)}{2(p'(z) - \frac{p(z)\,p''(z)}{2p'(z)})^2} \tag{3}$$

where $S[p](z)$ is the Schwarzian derivative of $p(z)$

$$S[p](z) = \frac{p'''(z)}{p'(z)} - \frac{3}{2}\left(\frac{p''(z)}{p'(z)}\right)^2$$

(see [6]). From expression (3) we see that the roots are superattracting, but of one degree higher order than for Newton's method. If ξ is a nondegenerate critical point of $p(z)$, then $H'(\xi) = 3$. If ξ is a degenerate critical point of $p(z)$, then it is a pole for $H'(z)$. Therefore, the additional fixed points of Halley's method at the critical points of $p(z)$ are repelling. It can also be shown that the point at infinity is a repelling fixed point for each method and that Halley's method is more effective at repelling very large initial seeds towards roots.

The critical points of $H_p(z)$ which don't correspond to roots of $p(z)$ are places where the Schwarzian derivative of $p(z)$ vanishes. We will call z a **free critical point** if it is an inflection point of $p(z)$ in the case of Newton's method or a place where the Schwarzian derivative $S[p](z)$ vanishes in the case of Halley's method. In order to find polynomials containing extraneous attracting cycles, we follow the orbit of the free critical point(s). Any such cycle must attract the orbit of a free critical point, since the other critical points (those corresponding to the roots) are fixed. This is based on the important theorem of Fatou and Julia [2]:

Theorem 2.1 *Every attracting cycle of a rational map attracts at least one critical point.*

Before undertaking this analysis, we mention some important facts which hold true for both numerical methods. These facts will be used to reduce the dimension of the parameter space for cubics and to deduce much of the symmetry found in the figures.

Lemma 2.2 *Let $A(z) = \alpha z + \beta$ with $\alpha \neq 0$ be an affine map of the complex plane and let $q(z) = p(A(z))$. Then N_p (H_p) is topologically conjugate to N_q (H_q, respectively.) Specifically, $A \circ N_q \circ A^{-1} = N_p$ and $A \circ H_q \circ A^{-1} = H_p$.*

Proof: This is a straight-forward calculation. We have

$$A \circ N_q = \alpha \left(z - \frac{p(A(z))}{\alpha\, p'(A(z))} \right) + \beta = A(z) - \frac{p(A(z))}{p'(A(z))} = N_p \circ A.$$

A similar calculation works for Halley's method. □

Lemma 2.3 *Let $\eta(z) = \bar{z}$ be the map which sends z to its complex conjugate $\bar{z}$. Suppose that $p(z) = \prod(z - r_i)$ and $q(z) = \prod(z - \overline{r_i})$. Then N_p (H_p) is topologically conjugate to N_q (H_q, respectively) via the conjugacy η.*

Proof: This is also a straight-forward calculation. We make use of the fact that $\overline{p(z)} = q(\bar{z})$. A similar result holds for any of the derivatives, $\overline{p^{(j)}(z)} = q^{(j)}(\bar{z})$, where $p^{(j)}(z)$ represents the j-th derivative of the polynomial p. We have

$$\eta \circ N_p = \overline{z - \frac{p(z)}{p'(z)}} = \bar{z} - \frac{\overline{p(z)}}{\overline{p'(z)}} = \bar{z} - \frac{q(\bar{z})}{q'(\bar{z})} = N_q \circ \eta.$$

A similar calculation works for Halley's method. □

2.2 Newton's and Halley's method applied to quadratics

As discussed in the introduction, the dynamics of Newton's or Halley's method applied to a quadratic polynomial with distinct roots is easy to describe. Initial seeds converge to the closest root. The dynamics on the line of points equidistant from the roots are chaotic, conjugate to angle doubling (tripling) for Newton's (Halley's, respectively) method. The result on Newton's method dates back to Schröder's work in 1871 [12] and Cayley's paper in 1879 [4]. The result on Halley's method can be found in [14] where it is shown that the König iteration functions K_m applied to $p(z) = z^2 - 1$ are conjugate to $z \mapsto z^m$. (Newton's method is equivalent to K_2 and Halley's method is equivalent to K_3.) Since it can be proven that K_m is conjugate under affine transformations, the result holds for the general quadratic with distinct roots.

Theorem 2.4 *Suppose that $p(z)$ is a quadratic polynomial with distinct roots. Newton's (Halley's) method for $p(z)$ is topologically conjugate to $z \mapsto z^2$ ($z \mapsto z^3$, respectively). Moreover, the Julia set in each case is all points lying on the perpendicular bisector of the segment joining the two roots.*

Proof: We extend the proof for Newton's method given in [3] to apply to Halley's method. Let the two distinct roots of $p(z)$ be r_1 and r_2. The conjugacy for both methods is the map

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

Note that h maps r_1 to 0, r_2 to ∞, ∞ to 1 and sends the perpendicular bisector l defined by $\{z : |z - r_1| = |z - r_2|\}$ to the unit circle.

One way to prove the theorem is to directly calculate that

$$h \circ N_p \circ h^{-1}(z) = z^2 \quad \text{and} \quad h \circ H_p \circ h^{-1}(z) = z^3.$$

A more elegant and generalizable approach is to consider the map $S = h \circ H_p \circ h^{-1}$ which is rational of degree 3 because H_p is degree 3. We compute that $S(0) = 0, S'(0) = 0, S''(0) = 0, S(\infty) = \infty, S'(\infty) = 0, S''(\infty) = 0$ and $S(1) = 1$. This makes use of the fact that H_p has superattracting fixed points of order 2 at r_1 and r_2. Therefore, S is a rational map of degree 3 which has superattracting fixed points of order 2 at 0 and ∞, and fixes $z = 1$. It follows that $S(z) = z^3$. A similar argument works for Newton's method. Since the Julia set for the maps $z \mapsto z^2$ and $z \mapsto z^3$ is the unit circle, and since the Julia set is preserved under conjugacy, it follows that the perpendicular bisector l is the Julia set for each method. □

Remark:
The previous argument is easy to generalize. Suppose that $p(z)$ is a quadratic polynomial with distinct roots and that Φ is a rational iterative method of order k. If Φ applied to $p(z)$ is a degree k rational map which fixes ∞, then Φ_p is conjugate to $z \mapsto z^k$.

3 Newton's Method Applied to Cubics

3.1 Geometry and reducing the parameter space

Let $p(z) = (z-a)(z-b)(z-c)$ be a complex cubic polynomial. The inflection point of p (the free critical point for N_p) is $\rho = (a+b+c)/3$. Assuming that a, b and c are non-collinear, ρ lies at the centroid of the triangle formed by the roots, that is, the point at which all three medians of the triangle intersect. Based on the side-lengths of the triangle, it is possible to determine which root is closest to ρ. For example, if $|a-c| \leq |a-b|$, then any point z on the median emanating from a satisfies $|z-c| \leq |z-b|$. In particular, this means that $|\rho - c| \leq |\rho - b|$. It follows that ρ is *closest* to the vertex (or vertices if the triangle is isosceles) between the two shortest sides.

Does ρ converge to its closest root? We explore this question by examining a particular set of polynomials whose roots form a triangle with longest side between -1 and 1. A similar family is studied in [3]. The advantage of our family is its symmetry about the imaginary axis, a very useful feature for investigating the important isosceles case.

Given any triangle in the complex plane, there is an affine map which takes the longest side to the real interval $[-1, 1]$ and the remaining vertex into the upper half-plane. Throughout the paper, our parameter region will be given by

$$\Lambda = \{\lambda \in \mathbf{C} : |\lambda - 1| \leq 2, |\lambda + 1| \leq 2, \mathrm{Im}(\lambda) \geq 0, \lambda \neq \pm 1\}.$$

Also, let p_λ denote the family of cubic polynomials

$$p_\lambda(z) = (z-1)(z+1)(z-\lambda)$$

where the third root λ is treated as a complex parameter. By Lemma 2.2, we can explore Newton's method on all cubic polynomials with distinct roots, by studying it applied to the one-parameter family $\{p_\lambda : \lambda \in \Lambda\}$. For any cubic polynomial $q(z)$ with distinct roots, there is a parameter value $\lambda \in \Lambda$ such that N_q is topologically conjugate to N_{p_λ}.

Note that a further reduction of the parameter space is possible using Lemma 2.3. Consider $\zeta(z) = -\bar{z}$ as the composition of complex conjugation and an affine transformation. Given a parameter value λ, Newton's method applied to p_λ is topologically conjugate to Newton's method applied to the polynomial $p_{-\bar{\lambda}}(z) = (z+1)(z-1)(z+\bar{\lambda})$ with ζ serving as the conjugacy. This implies a symmetry about the imaginary axis. Although we could restrict Λ to exclude the second quadrant, we wish to preserve the symmetry in our figures by defining it as above.

On the boundary of Λ the triangle formed by the roots is isosceles. This boundary is super-imposed onto the parameter plane for Newton's method applied to p_λ in Figure 1. The picture is created by

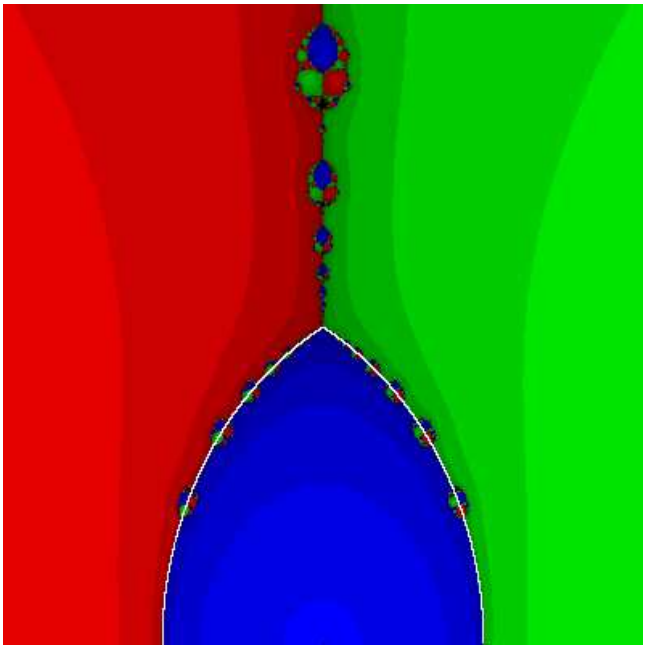

Figure 1: The parameter plane for Newton's method colored according to the convergence of the free critical point $\rho = \lambda/3$. The boundary of the region Λ is shown in white. The plot range for the figure is $[-2, 2] \times [0, 3.5]i$.

following the orbit of the free critical point $\rho = \lambda/3$ under N_{p_λ}. If ρ converges to a root $\lambda, 1$ or -1, then the parameter value λ is colored blue, green or red, respectively. Since N_{p_λ} is conjugate to $N_{p_{-\bar{\lambda}}}$, green and red colors are symmetric with respect to the imaginary axis, as indicated in the figure. Blue regions also display this vertical symmetry. If ρ does not converge to within 10^{-6} of any root after 60 iterations, the parameter value is colored black.

Note that Figure 1 confirms our geometric intuition. Frequently, the free critical point finds the root closest to it. For example, most of the interior of Λ is colored blue since λ is the closest root to the free critical point ρ. To the left of the boundary of Λ, -1 becomes the closest root to ρ and thus most of this region is colored red. Separating these regions is the crucial isosceles case, where ρ has a choice between two or more roots.

The isosceles case occurs for three specific regimes, the left circular boundary of Λ ($|\lambda - 1| = 2$), the right circular boundary of Λ ($|\lambda + 1| = 2$) and the imaginary axis ($\lambda = \beta\, i, \beta \geq \sqrt{3}$). In each case, the congruent legs are longer than the base. The intersection of the three cases (when $\lambda = \sqrt{3}\, i$) represents the equilateral triangle configuration. It is easy to check that the three configurations can be mapped to each other via an affine transformation. This fact is crucial as it allows us to study those parameter values on the boundary of Λ by restricting to those λ-values on the imaginary axis with $\mathrm{Im}(\lambda) \geq \sqrt{3}$.

A more detailed picture of the parameter plane for N_{p_λ} is shown in Figure 2. Here we are only interested in whether ρ converges to a root or not. The lighter shaded values correspond to faster convergence. The black regions represent values where the free critical point does not converge to one of the roots. In most of these cases (the interior of the black regions), the free critical point finds a different attracting cycle other than the roots. For any such parameter value λ, N_{p_λ} fails on an open set of initial guesses. As in the papers [3, 5] and [7], we see the existence of Mandelbrot-like sets in the parameter plane (see Figure 3 for enlargements.) We investigate the periods of the main cardioids of these Mandelbrot-like sets in the next section.

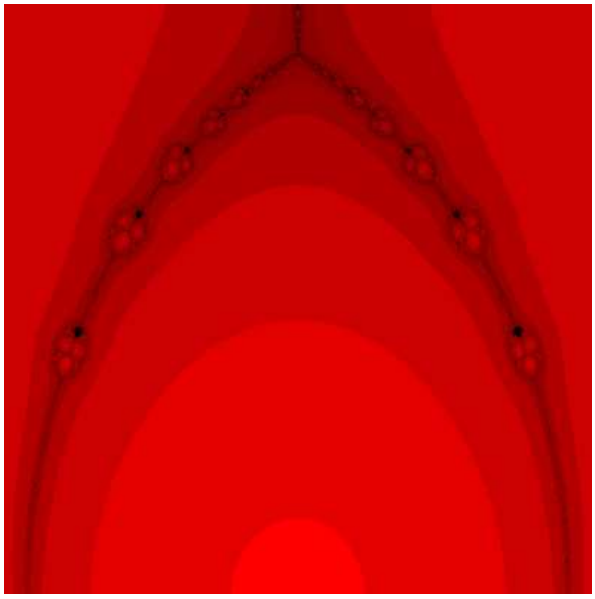

Figure 2: The parameter plane for Newton's method applied to p_λ (plot range $[-1.1, 1.1] \times [0, 1.9]i$).

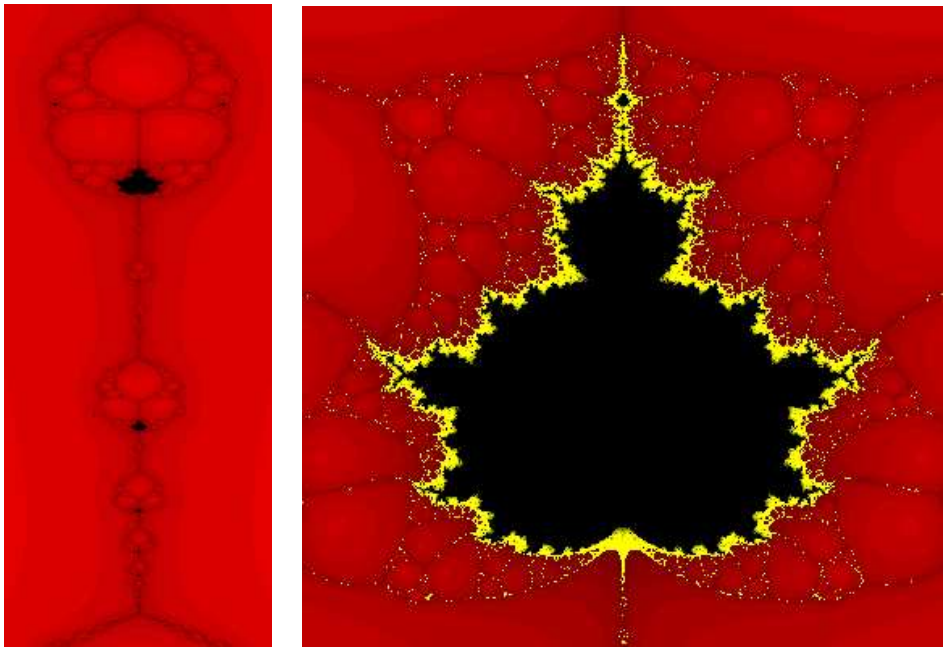

Figure 3: To the left is an enlargement around the imaginary axis of the parameter plane for N_{p_λ} (plot range $[-.55, .55] \times [1.5, 5.7]i$). To the right is a magnification of the period 2 Mandelbrot-like set (plot range $[-0.13, 0.13] \times [4.41, 4.69]i$). Yellow shading represents values where the critical point took between 25 and 60 iterations to converge to a root.

3.2 Restricting to λ-values on the imaginary axis

We now restrict our study to the case when $\lambda = \beta\, i, \beta \in \mathbf{R}^+$, where the roots form an isosceles triangle. A similar approach, although less rigorous, was undertaken in [15]. We provide some evidence for the existence of Mandelbrot-like sets in the parameter plane by showing that superattracting periodic cycles are born in the regime $\beta > \sqrt{3}$. By affine conjugacy, the same events take place on the left and right circular boundaries of Λ. The geometry of Newton's method leaves the axis of symmetry of the isosceles triangle invariant. This mimics the behavior in the quadratic case. The orbit does not know which of the two outer roots to choose from and consequently remains equidistant from each. The difference here, of course, is that a third root exists on the axis as a point of possible convergence.

Setting $\lambda = \beta\, i$, we have $p_\lambda(z) = (z^2 - 1)(z - \beta\, i)$ and $p'_\lambda(z) = 3z^2 - 1 - 2\beta\, iz$. It follows that $p_\lambda(y\, i)$ is pure imaginary and $p'_\lambda(y\, i)$ is real. Consequently, $\mathrm{Re}(N_{p_\lambda}(y\, i)) = 0$ so that the imaginary axis is invariant. Since the free critical point $\rho = \lambda/3 = \beta\, i/3$ is also pure imaginary, we can restrict our study to a rational map of one variable. This map is determined by computing the imaginary part of $N_{p_{\beta i}}(y\, i)$ and is given by

$$R_\beta(y) = \frac{2y^3 - \beta y^2 + \beta}{3y^2 - 2\beta y + 1}$$

where β is our new real parameter and y is a real variable. Note that $R_\beta(\beta) = \beta$ as to be expected since $\beta\, i$ is fixed under Newton's method as one of the roots. Also, we have

$$R'_\beta(y) = \frac{2(y - \beta)(3y - \beta)(y^2 + 1)}{(3y^2 - 2\beta y + 1)^2}$$

which confirms that β is a superattracting fixed point and that $\beta/3$ is the free critical point.

Lemma 3.1 *For $0 \le \beta < \sqrt{3}$, all orbits under R_β converge to β.*

Proof: First, we notice that the poles of R_β, $y_\pm = (\beta \pm \sqrt{\beta^2 - 3})/3$, are real only for $\beta \ge \sqrt{3}$. At $\beta = \sqrt{3}$, corresponding to the equilateral triangle case, there is a unique pole at the free critical point $\rho = \sqrt{3}/3$. Thus the denominator of R_β is strictly positive in the regime $0 \le \beta < \sqrt{3}$. From this it is straight-forward to compute that $R_\beta(y) > y$ when $y < \beta$ and $R_\beta(y) < y$ when $y > \beta$ (see Figure 4).

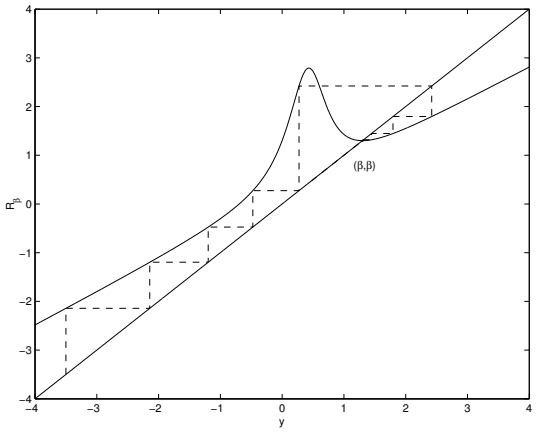

Figure 4: A typical graph of R_β for $0 \le \beta < \sqrt{3}$. A sample web diagram (dashed) shows the convergence to β.

Next, the formula for $R'_\beta(y)$ reveals that $R'_\beta(y) > 0$ for $y < \beta/3$ and for $y > \beta$ and that $R'_\beta(y) < 0$ whenever $\beta/3 < y < \beta$. It follows that R_β maps the interval $[\beta, \infty)$ into itself. Because $R_\beta(y) < y$ when

$y > \beta$, the sequence $R_\beta^n(y_0)$ is strictly decreasing and bounded below by β for any $y_0 > \beta$. Since β is the only fixed point of the map, the decreasing sequence $R_\beta^n(y_0)$ must converge to β as $n \to \infty$.

For initial seeds $y_0 < \beta$, calculation shows that

$$R_\beta(\frac{\beta}{3}) = \frac{\beta(\beta^2 - 27)}{9(\beta^2 - 3)} \tag{4}$$

which is larger than β for $0 \leq \beta < \sqrt{3}$. Because $R_\beta(y) > y$ when $y < \beta$ and since β is the only fixed point, it follows that any point $y_0 < \beta$ eventually iterates into the region $y \geq \beta$, from which it must then converge to β. □

From Lemma 3.1 we can conclude three interesting facts for the regime $0 \leq \beta < \sqrt{3}$. First, initial seeds on the negative imaginary axis far from the root $\beta\, i$ still converge to $\beta\, i$ even though the other two roots at -1 and 1 are closer. Second, unlike the quadratic case, it is *not* the case that the point equidistant from all three roots is in the Julia set of N_{p_λ}. For the regime $0 \leq \beta < \sqrt{3}$, the point equidistant from $-1, 1$ and $\beta\, i$ lies on the imaginary axis and consequently, its orbit converges to $\beta\, i$, placing it in the Fatou set. Third, no interesting dynamical behavior occurs in the parameter plane for this regime because the free critical point always converges to $\beta\, i$. This is indicated in the parameter plane picture in Figure 1 as the imaginary axis is colored blue for $\beta < \sqrt{3}$. As we shall see, the same result occurs for Halley's method but on the interval $\beta \in [0, 1)$. For Newton's method, the interesting behavior (eg. the birth of extraneous attracting cycles) occurs for $\beta > \sqrt{3}$ and consequently, around the left and right circular boundaries of Λ.

Proposition 3.2 *For each integer $n \geq 2$, there exists a parameter value $\beta_n < 3\sqrt{3}$ such that the free critical point $\rho_n = \beta_n/3$ is on a superattracting n cycle. Moreover, the sequence β_n can be chosen so that it is strictly decreasing, converging to $\sqrt{3}$.*

Proof: Figure 5 shows the graph for R_β for $\beta = 2.39638545$. Note the difference from the previous case. For $\beta > \sqrt{3}$, R_β has two poles y_- and y_+ symmetric about the free critical point $\beta/3$. As before, the formula for $R'_\beta(y)$ reveals that where the derivative exists, it is positive for $y < \beta/3$ and $y > \beta$, and is negative whenever $\beta/3 < y < \beta$. However, for $\sqrt{3} < \beta < 3\sqrt{3}$, expression (4) shows that $R_\beta(\beta/3) = R_\beta(\rho) < 0$. Consequently, the graph of R_β takes on negative values between the two poles when $\sqrt{3} < \beta < 3\sqrt{3}$. Note that $0 < y_- < y_+ < \beta$. It is straight-forward to see that $R_\beta(y) > y$ on the intervals $(-\infty, y_-)$ and (y_+, β) while $R_\beta(y) < y$ on (y_-, y_+) and (β, ∞). Thus the generic graph for $R_\beta(y)$ when $\sqrt{3} < \beta < 3\sqrt{3}$ is as shown in Figure 5.

To show the existence of a superattracting n cycle, we compute

$$\frac{d}{d\,\beta}\,(R_\beta(\beta/3)) \;=\; \frac{(\beta^2+9)^2}{9(\beta^2-3)^2} \;>\; 0.$$

Taken with the fact that $\lim_{\beta \to \sqrt{3}^+} R_\beta(\beta/3) = -\infty$, it follows that the image of the critical point ρ is decreasing towards $-\infty$ as β decreases towards $\sqrt{3}$. This is the key fact behind the proposition.

Let n be some period with $n \geq 2$. Consider $g_n(\beta) = R_\beta^n(\beta/3) - \beta/3$ as a function of the parameter β. A root of $g_n(\beta)$ gives a period n cycle for R_β and this cycle is superattracting because it contains the critical point ρ in its orbit. We claim that there exists a value ξ_n with $\sqrt{3} < \xi_n < 3\sqrt{3}$ such that $g_n(\beta)$ is a continuous function in β on the open interval $(\sqrt{3}, \xi_n)$ and that

$$\lim_{\beta \to \sqrt{3}^+} g_n(\beta) = -\infty \quad \text{and} \quad \lim_{\beta \to \xi_n^-} g_n(\beta) = \infty. \tag{5}$$

The intermediate value theorem then gives the existence of a root for $g_n(\beta)$ in the interval $(\sqrt{3}, \xi_n)$.

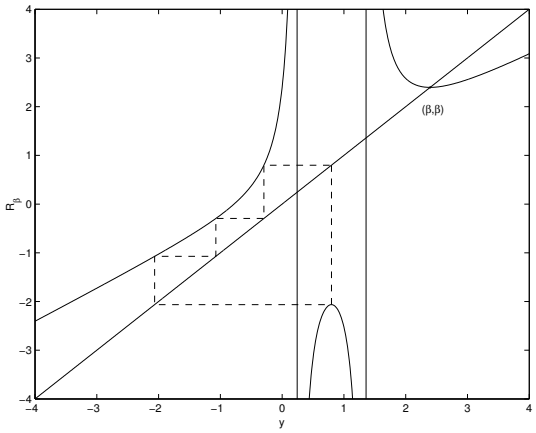

Figure 5: The graph of R_β for $\beta = 2.39638545$ and the web diagram (dashed) demonstrating the superattracting period 4 cycle.

The first limit in (5) follows from $\lim_{\beta\to\sqrt{3}+} R_\beta(\beta/3) = -\infty$. Polynomial division then yields

$$R_\beta(y) = \frac{2}{3}y + \frac{1}{9}\beta + \frac{y(\frac{2}{9}\beta^2 - \frac{2}{3}) + \frac{8}{9}\beta}{3y^2 - 2\beta y + 1} \tag{6}$$

which implies that for y values large in magnitude, $R_\beta(y)$ can be approximated by the linear function $2y/3 + \beta/9$. Thus, as β decreases towards $\sqrt{3}$, the orbit of the critical point ρ takes more and more iterates to return back near itself. As β increases away from $\sqrt{3}$, the image of ρ increases away from $-\infty$ and consequently, $g_n(\beta)$ is also increasing. The first place where $g_n(\beta)$ may lose continuity is if $g_{n-1}(\beta) = y_-$, the left most pole. Since the image of the critical point ρ is strictly increasing and since $R^k_{3\sqrt{3}}(\sqrt{3}) = \beta$ for any $k \geq 2$, there will be a value $\xi_n < 3\sqrt{3}$ where $g_{n-1}(\xi_n)$ lands on the left-most pole. This verifies the second limit in (5) and proves the existence of the superattracting n cycle.

The last statement of the proposition follows from the fact that the image of the critical point ρ is strictly increasing in β and from equation (6). Alternatively, the larger the period of the superattracting n cycle, the more negative the image of the critical point ρ and the closer β must be to $\sqrt{3}$. □

Table 1 shows the values β_n which yield a superattracting n cycle for Newton's method applied to $p_{\beta_n i}$. Each parameter value lies at the center of the main cardioid of a period n Mandelbrot-like set shown on the left in Figure 3. As $n \to \infty$, these sets coalesce along the imaginary axis, converging to $\lambda = \sqrt{3}\,i$.

Table 1: The (approximate) parameter values $\lambda_n = \beta_n\, i$ for which Newton's method applied to p_{λ_n} has a superattracting n cycle.

n	β_n	n	β_n	n	β_n
2	4.50016263	6	1.98209864	10	1.77722708
3	2.93806914	7	1.89245943	20	1.73281876
4	2.39638545	8	1.83640072	25	1.7321519067
5	2.13108922	9	1.80052219	50	1.7320508116

It is important to note that the values given in Table 1 are not unique. For instance, there is a superattracting period 4 cycle at $\beta \approx 4.58784584$ which arises from period-doubling off of the main

cardioid about $\beta_2 \approx 4.50016263$. There are also smaller rogue Mandelbrot-like sets containing superattracting cycles such as a period 5 cycle at $\beta \approx 3.8891390787$ and a period 6 cycle at $\beta \approx 3.667823184$. These parameter values correspond to places where the orbit of the free critical point passes close to itself once before returning to complete the cycle.

4 Halley's Method Applied to Cubics

In this section, we explore the dynamics of Halley's method applied to cubic polynomials with distinct roots. Due to Lemma 2.2, the problem is reduced to studying the same region Λ in the parameter plane as for Newton's method. One important difference for this case is that there are two free critical points rather than one. This complicates the parameter plane picture considerably. A second important difference is that the presence of attracting cycles other than roots does not occur on the boundary of Λ as it did for Newton's method. Although Halley's method can fail in the isosceles case, there are other configurations which can cause difficulty as well.

4.1 Parameter plane

As before, let $p_\lambda(z) = (z-1)(z+1)(z-\lambda)$ denote the one-parameter family of cubic polynomials whose roots we search for under Halley's method. The Schwarzian derivative of p_λ vanishes at the roots of the quadratic $6z^2 - 4\lambda z + 1 + \lambda^2$ yielding two free critical points

$$\rho_\pm \;=\; \frac{1}{6}\left(2\lambda \;\pm\; \sqrt{-2\lambda^2 - 6}\right).$$

Here we take the principle branch of the square root (real part non-negative). Note that the free critical points for Halley's method are symmetric around the free critical point for Newton's method $\lambda/3$. We also have that $\mathrm{Re}(\rho_+) \geq \mathrm{Re}(\rho_-)$ with equality holding only when λ is real or for $\lambda = \beta\, i$ with $|\beta| \leq \sqrt{3}$. Specific places of interest include

- $\lambda \in \mathbf{R} \implies \overline{\rho_+} = \rho_-$,
- $\lambda = \beta\, i$ with $|\beta| \leq \sqrt{3} \implies \rho_\pm$ are pure imaginary,
- $\lambda = \sqrt{3}\, i \implies \rho_+ = \rho_- = \sqrt{3}\, i/3$ and
- $\lambda = \beta\, i$ with $|\beta| > \sqrt{3} \implies \rho_\pm$ are truly complex and symmetric about the imaginary axis.

Figure 6 captures the behavior of the orbit of the free critical points $\rho_\pm$ under H_{p_λ} for varying parameter values λ. Since there are two critical orbits to follow and three possible roots for convergence, there are nine possible convergence schemes. However, one of these cases never occurs, perhaps due to the fact that Halley's method often respects the nearest root principal. The case which never occurs is when the orbits of the critical points ρ_- and ρ_+ cross over each other to converge to 1 and -1, respectively. Thus, we require 8 colors to distinguish the different regions in parameter space, as indicated in Table 2.

Note that the behavior along the imaginary axis between i and $\sqrt{3}\, i$ mimics that of Newton's method along the imaginary axis from $\sqrt{3}\, i$ to $3\sqrt{3}\, i$ revealing the importance of the isosceles case for Halley's method. However, there is also a new branch traveling from i to 1 (and its symmetric sister from i to -1) which increases the complexity of the parameter plane picture. This extra branch arises from the addition of a free critical point, as the following qualitative argument indicates.

Consider the convergence of the two free critical points as a function of λ and look for bifurcations. Beginning in the blue region in Figure 6, both critical points converge to the root λ. Draw an arc moving right to left into the red region, where both critical points converge to the root -1. Somewhere

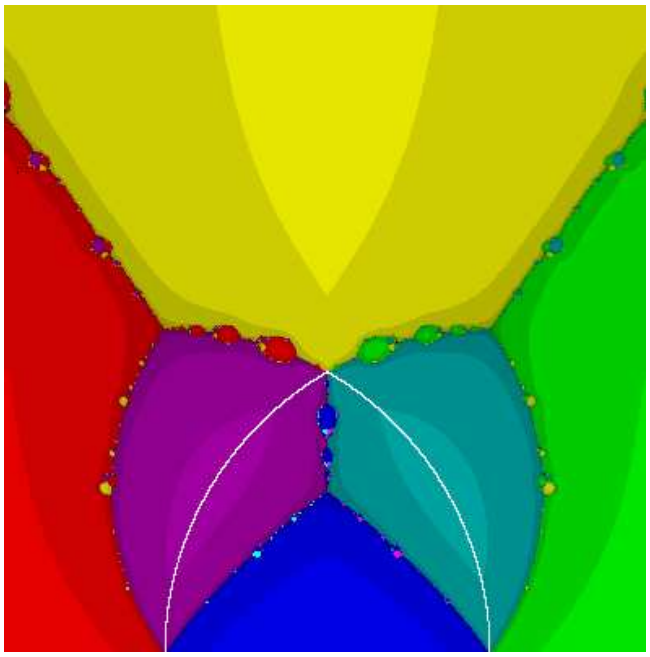

Figure 6: The parameter plane for Halley's method colored according to the convergence of the two free critical points $\rho_\pm$ as indicated in Table 2. The boundary of the region Λ, shown in white, is superimposed on the figure.

Table 2: The coloring scheme for the parameter plane in Figure 6. For example, λ is colored purple whenever the critical point ρ_- converges to the root -1 and the critical point ρ_+ converges to the root λ.

color	$(\rho_-, \rho_+) \to$	color	$(\rho_-, \rho_+) \to$	color	$(\rho_-, \rho_+) \to$
red	$(-1,-1)$	yellow	$(-1,1)$		
blue	(λ,λ)	purple	$(-1,\lambda)$	pink	$(\lambda,-1)$
green	$(1,1)$	hazel	$(\lambda,1)$	light blue	$(1,\lambda)$

in between we should expect two bifurcations. The first occurs as we enter the purple region, where the critical point ρ_- (farthest to the left) switches to converge to -1, while the other critical point's convergence remains unchanged. The second occurs as we enter the red region where the critical point ρ_+ switches to converge to -1 while the other critical point's convergence remains unchanged. By symmetry, a similar phenomenon happens moving from the blue to green region.

Note that along the real axis, both critical points converge to the same root (green, red or blue colors only). If $\lambda = \bar{\lambda}$, then Lemma 2.3 implies that $\overline{H_{p_\lambda}(z)} = H_{p_\lambda}(\bar{z})$. Since $\overline{\rho_+} = \rho_-$, if ρ_+ converges to a root r, then ρ_- converges to $\bar{r} = r$. Hence both critical points converge to the same root.

Just as in the case of Newton's method, there is symmetry in Figure 6 about the imaginary axis. The key fact here is that the map $\zeta(z) = -\bar{z}$ sends the free critical points $\rho_\pm$ for Halley's method applied to p_λ to the free critical points $\rho_\mp$ for Halley's method applied to $p_{-\bar{\lambda}}$. In other words, if ρ_+ converges to the root r under H_{p_λ}, then $-\overline{\rho_+}$, which is the new critical point ρ_- for $H_{p_{-\bar{\lambda}}}$, converges to the root $-\bar{r}$ under Halley's method applied to $p_{-\bar{\lambda}}$. Thus, if (ρ_-, ρ_+) converge to the roots $(-1, \lambda)$ under H_{p_λ}, then the new free critical points (ρ_-, ρ_+) converge to $(-\bar{\lambda}, 1)$ under $H_{p_{-\bar{\lambda}}}$. In terms of color, this means that red and green are symmetric with respect to the imaginary axis, as are the pairs purple and hazel,

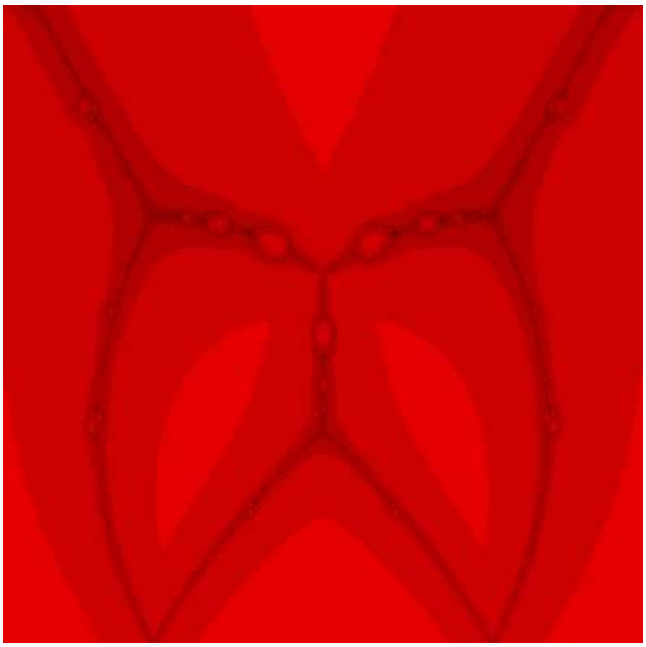

Figure 7: The parameter plane for Halley's method applied to p_λ (plot range $[-1.9, 1.9] \times [0, 3]i$).

and pink and light blue. Both the blue and yellow regions are symmetric with respect to themselves. These symmetries are clearly indicated in Figure 6.

Figure 7 shows the parameter plane for Halley's method without the color coding for convergence. Here, λ is colored black if one of the two free critical points does not converge to a root. Bifurcations occur along three "wishbones", one inside Λ and two located in the exterior of Λ, on either side of the imaginary axis. These wishbones are conjugate under an affine map. For example, the map

$$S_\lambda(z) = \frac{-2z + \lambda + 1}{1 - \lambda}$$

is an affine conjugacy between H_{p_λ} for λ-values on the main wishbone in Λ and H_{p_λ} for λ-values on the right exterior wishbone. This map conjugates the "stem" of the main wishbone on the imaginary axis to the stem of the right exterior wishbone on the circle $|z - 1| = 2$. Geometrically, this conjugacy is visualized by rotating the main wishbone inside Λ counterclockwise to line up with the outer right wishbone. The value i at the center of the main wishbone is identified with the value $1 + 2i$ at the center of the right wishbone. Similar arguments work for the left exterior wishbone using the map $T_\lambda(z) = (-2z + \lambda - 1)/(1 + \lambda)$.

As with Newton's method, we see the existence of Mandelbrot-like sets in the parameter plane for Halley's method (see Figure 8). However, the sets for Halley's method are considerably smaller than the corresponding ones found in Figure 3. For example, the period 2 Mandelbrot-like set shown for Newton's method is roughly 0.25 units of width while the corresponding set for Halley's method is only about 0.02 units of width. This is probably due to the higher rate of convergence for Halley's method.

In Figure 9, we show the dynamical plane for Halley's method applied to p_λ with $\lambda = 1.34853\,i$. This parameter value was taken from the circular bulb just above the main cardioid shown to the right in Figure 8. As before, a point z is colored blue (respectively, red or green) if the orbit of z converges to λ (respectively, -1 or 1). If z fails to converge to within 10^{-6} of any root after 60 iterations, it is colored black. In this example the critical point ρ_+ converges to the root λ while the critical point ρ_- limits on an attracting period 4 cycle. Part of the filled in Julia set for H_{p_λ} is shown on the right. This

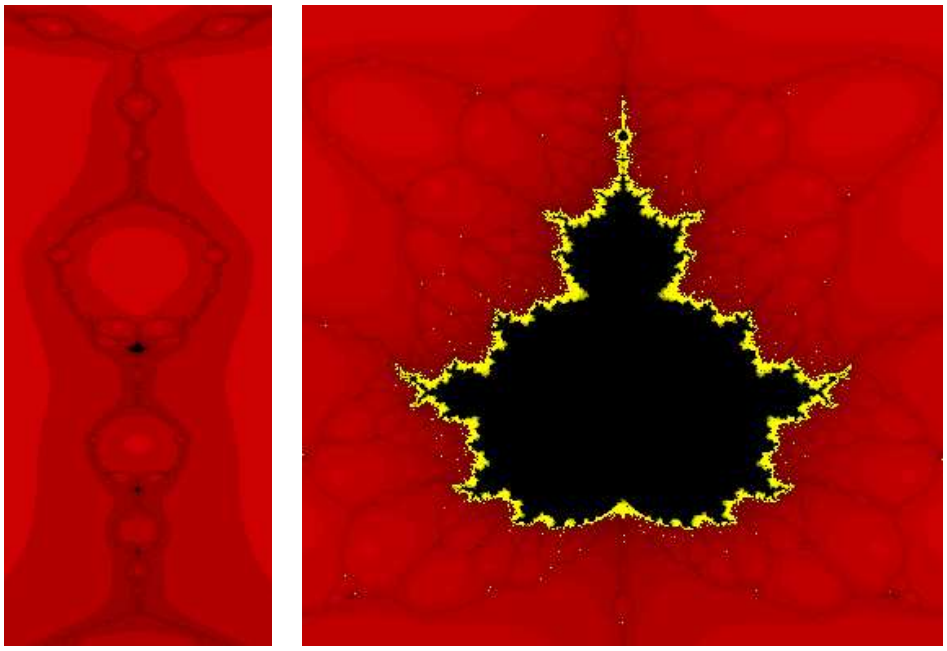

Figure 8: To the left is an enlargement around the imaginary axis of Figure 7 (plot range $[-0.1, 0.1] \times [0.95, 1.8]i$.) To the right is a magnification of the period 2 Mandelbrot-like set in the parameter plane for Halley's method (plot range $[-0.01, 0.01] \times [1.335, 1.357]i$.)

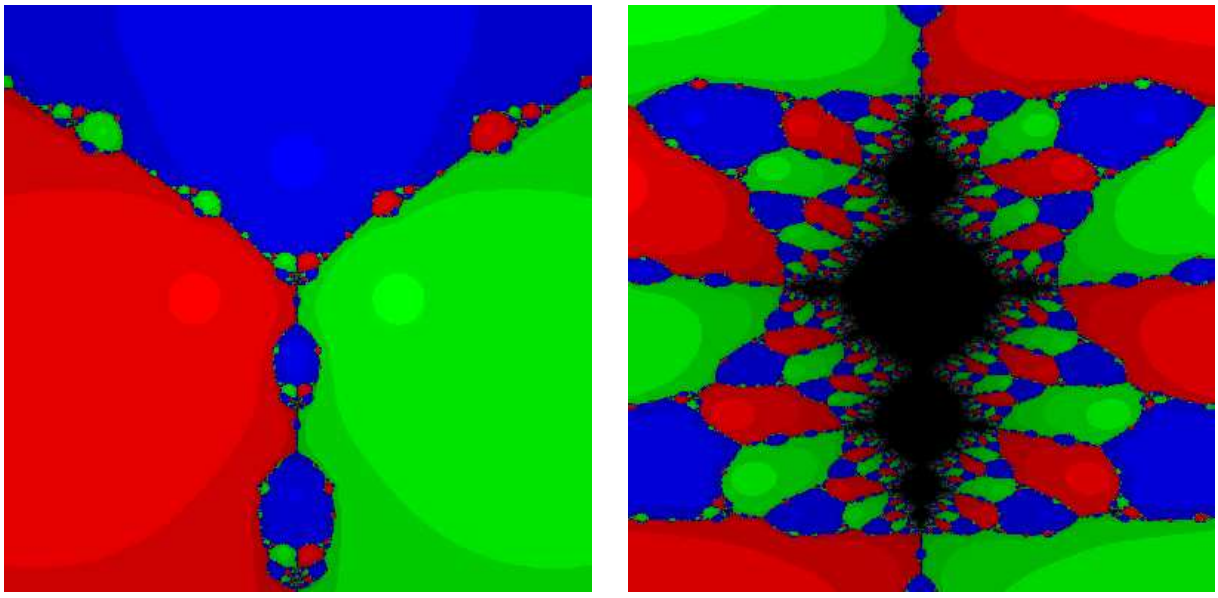

Figure 9: To the left is the dynamical plane for Halley's method applied to p_λ for $\lambda = 1.34853\,i$ (plot range $[-3, 3] \times [-3, 3]i$). To the right is an enlargement about the free critical point ρ_- (plot range $[-0.065, 0.065] \times [0.1, 0.282]i$).

set bears a striking resemblance to the filled in Julia set for the quadratic map $z \mapsto z^2 - 1$, where the critical point $z = 0$ is on a super-attracting period 2 cycle. For the family $z \mapsto z^2 + c$, $c = -1$ is a parameter value in the circular period 2 bulb directly to the left of the main cardioid of the Mandelbrot set. This is the analogous bulb to the one in Figure 8 from which λ was chosen.

4.2 Restricting to λ-values on the imaginary axis

We now discuss the behavior of Halley's method applied to p_λ for pure imaginary λ-values, obtaining similar results to those we found for Newton's method in Section 3.2. The only difference here is that the parameter values where superattracting cycles occur are in the regime $(1, \sqrt{3})$ and the parameter values found decrease towards $\beta = 1$ as opposed to $\beta = \sqrt{3}$. This puts the interesting dynamical behavior in the interior of Λ rather than the exterior, as was the case in Newton's method.

By choosing $\lambda = \beta\, i, \beta \in \mathbf{R}^+$, Halley's method applied to p_λ leaves the imaginary axis invariant. This is easy to see by examining equation (2), as $p_\lambda(y\, i)$ is pure imaginary, $p'_\lambda(y\, i)$ is real and $p''_\lambda(y\, i)$ is pure imaginary. Moreover, for the regime $0 \leq \beta < \sqrt{3}$, both free critical points $\rho_\pm$ are pure imaginary. Thus, as with Newton's method, we can reduce Halley's method to studying a one-parameter family of rational maps.

We compute the imaginary part of $H_{p_{\beta\, i}}(y\, i)$ for $0 \leq \beta < \sqrt{3}$ and obtain the rational map

$$S_\beta(y) \;=\; \frac{3y^5 - 3\beta y^4 + (\beta^2 - 1)y^3 + 6\beta y^2 - 3\beta^2 y + \beta}{6y^4 - 8\beta y^3 + 3(1 + \beta^2)y^2 + 1 - \beta^2}$$

where β is our new real parameter and y is a real variable. We are interested in studying the orbit of two real free critical points

$$\rho_\pm \;=\; \frac{1}{6}\left(2\beta \;\pm\; \sqrt{6 - 2\beta^2}\right)$$

as β varies. As expected, $S_\beta(\beta) = \beta$ because $\beta\, i$ is fixed under Halley's method. We compute

$$S'_\beta(y) \;=\; \frac{3(y - \beta)^2(y^2 + 1)^2(6y^2 - 4\beta y + \beta^2 - 1)}{(6y^4 - 8\beta y^3 + 3(1 + \beta^2)y^2 + 1 - \beta^2)^2} \tag{7}$$

which confirms that $\beta\, i$ is a superattracting fixed point of order two and that $\rho_\pm$ are the two free critical points. Finally, solving $S_\beta(y) = y$ yields the quintic equation

$$(y - \beta)(y^2 + 1)(3y^2 - 2\beta y + 1) = 0. \tag{8}$$

The roots of the quadratic term $3y^2 - 2\beta y + 1$ in the above expression represent the two extra fixed points for Halley's method located at the critical points of the polynomial p_λ. However, these roots are complex since $3y^2 - 2\beta y + 1$ is strictly positive for $0 \leq \beta < \sqrt{3}$. Thus, $y = \beta$ is the only fixed point of S_β when $0 \leq \beta < \sqrt{3}$. For $\beta = \sqrt{3}$, (equilateral case) $\rho_+ = \rho_- = \sqrt{3}/3$ is a double root of $3y^2 - 2\beta y + 1$ because $\sqrt{3}/3$ is a degenerate critical point for p_λ.

Lemma 4.1 *For $0 \leq \beta < 1$, all orbits under S_β converge to β.*

Proof: First, note that the denominator of S_β can be factored into

$$6y^2\left((y - \frac{2}{3}\beta)^2 + \frac{1}{2} + \frac{1}{18}\beta^2\right) + 1 - \beta^2$$

which is always positive for $0 \leq \beta < 1$. Thus, S_β has no poles for $0 \leq \beta < 1$. When $\beta = 1$, S_β has a unique pole of order 2 at the free critical point $\rho_- = 0$. We found the same traits with R_β in Newton's

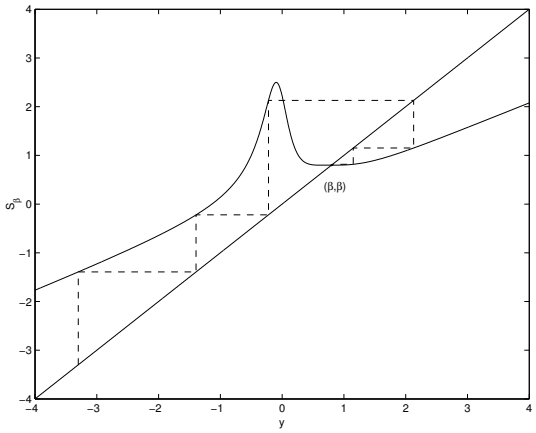

Figure 10: A typical graph of S_β for $0 \le \beta < 1$ and the web diagram (dashed) demonstrating convergence to $y = \beta$.

method at the parameter value $\beta = \sqrt{3}$. Using equation (8), it is easy to compute that $S_\beta(y) > y$ when $y < \beta$ and $S_\beta(y) < y$ when $y > \beta$ (see Figure 10).

Next, the formula for $S'_\beta(y)$ reveals that the derivative can only change sign at the two free critical points $\rho_\pm$. It follows that $S'_\beta(y) \ge 0$ for $y \le \rho_-$ and $y \ge \rho_+$, and that $S'_\beta(y) < 0$ whenever $\rho_- < y < \rho_+$. There are two cases determined by whether the critical point ρ_+ is larger or smaller than β. Note that $\rho_+ = \beta$ when $\beta = \sqrt{3}/3$.

Case 1: $\sqrt{3}/3 \le \beta < 1$

We do the easier case first, where the argument is the same as in the case of Newton's method. In this regime, we have $\rho_+ \le \beta$ and therefore $S_\beta(y) < y$ and $S'_\beta(y) > 0$ for $y \ge \beta$. It follows that S_β maps the interval $[\beta, \infty)$ into itself. Since β is the only fixed point of the map, the decreasing sequence $S^n_\beta(y_0)$ must converge to β for $y_0 \in [\beta, \infty)$. The interval $[\rho_+, \beta]$ is also invariant for this case and similar arguments give that the increasing sequence $S^n_\beta(y_0)$ must converge to β for initial seeds $y_0 \in [\rho_+, \beta]$. Finally, the orbit of any initial seed $y_0 < \rho_+$ will increase until reaching the region $[\rho_+, \infty)$ from which it must then converge to β.

Case 2: $0 \le \beta < \sqrt{3}/3$

In this case, we have $\rho_+ > \beta$ which means that the interval $[\beta, \infty)$ is no longer invariant since $S_\beta(\rho_+) < \beta$. We must estimate the derivative on a sufficiently large neighborhood of β to obtain global convergence. Let $\mu_\beta = S_\beta(\rho_+) < \beta$ be the image of the larger critical point. We claim that S_β is a contraction on the interval $[\mu_\beta, \rho_+]$ containing the attracting fixed point β. Since the orbit of every initial seed eventually must fall into this interval, it follows that β is globally attracting.

We first find an upper-bound for $|S'_\beta(y)|$ on the interval $[\beta, \rho_+]$. The denominator of $S'\beta(y)$ is an increasing function in y and therefore attains its minimum at $y = \beta$ on $[\beta, \rho_+]$. The quadratic term $6y^2 - 4\beta y + \beta^2 - 1$ is negative for $\beta \le y < \rho_+$ so this term is largest in absolute value at $y = \beta$. This yields the estimate

$$|S'_\beta(y)| \;\le\; \frac{3\,d^2((\beta+d)^2+1)^2(1-3\beta^2)}{(\beta^2+1)^4}$$

where $d \equiv d(\beta) = -2\beta/3 + \sqrt{6-2\beta^2}/6$ is the distance between β and ρ_+. Using the rather crude estimate $d \le \sqrt{6}/6$ and the fact that the denominator is always greater than or equal to 1, we obtain

$$|S'_\beta(y)| \;\le\; \frac{1}{2}(\beta^2 + \frac{\sqrt{6}}{3}\beta + \frac{7}{6})^2(1-3\beta^2).$$

This function is concave down for $0 \le \beta < \sqrt{3}/3$ and has a maximum of approximately 0.834572 at $\beta \approx 0.249055$. This shows that for any β such that $0 \le \beta < \sqrt{3}/3$, the magnitude of the derivative is less than 0.85 on the interval $y \in [\beta, \rho_+]$.

Next, we find an upper-bound for $|S'_\beta(y)|$ on the interval $[\mu_\beta, \beta]$. First, a tedious but straight-forward calculation shows that

$$\mu_\beta - \rho_- \;=\; \frac{(3-\beta^2)(19\sqrt{6-2\beta^2}(3-\beta^2)+40\beta(\beta^2+9))}{24(5(\beta^2-3)^2+\beta(\beta^2+9)\sqrt{6-2\beta^2})} \;>\; 0$$

which means that the image of ρ_+ is between ρ_- and β. Knowing this, the quadratic term $6y^2 - 4\beta y + \beta^2 - 1$ will be negative and thus can be bounded in magnitude using the value at its vertex $y = \beta/3$, yielding $|\beta^2/3 - 1|$. We then estimate the magnitude of the derivative as

$$|S'_\beta(y)| \;\le\; \frac{3(\beta-\mu_\beta)^2(\beta^2+1)^2(1-\beta^2/3)}{(1-\beta^2)^2}$$

on the interval $[\mu_\beta, \beta]$. Further calculation shows that

$$\beta - \mu_\beta \;=\; \frac{(1-3\beta^2)^4}{4\sqrt{6-2\beta^2}(\beta^2+5)(\beta^4+18\beta^2+1)+8\beta(11\beta^6-47\beta^4+113\beta^2+43)}. \tag{9}$$

The numerator in expression (9) is decreasing in β while the denominator is positive and an increasing function of β. Therefore, the expression in (9) is bounded above by its value at $\beta = 0$, yielding $\beta - \mu_\beta \le 1/(20\sqrt{6})$. Using this estimate, we obtain

$$|S'_\beta(y)| \;\le\; \frac{3(1/2400)(4/3)^2(1)}{(2/3)^2} \;=\; 1/200$$

so that the magnitude of the derivative on $[\mu_\beta, \beta]$ is quite small. This is not surprising because μ_β and β are very close and β is a double root of $S'_\beta(y)$. This shows that S_β is a contraction on the interval $[\mu_\beta, \rho_+]$. □

Proposition 4.2 *For each integer $n \ge 2$, there exists a parameter value $\beta_n < \sqrt{3}$ such that the free critical point ρ_- is on a superattracting n cycle. Moreover, the sequence β_n can be chosen as strictly decreasing, converging to* 1.

Proof: The argument is similar to that for Newton's method in Proposition 3.2. The graph of S_β displays the same properties as R_β did for Newton's method. There are two poles y_- and y_+ which contain the free critical point ρ_- between them and as $\beta \to 1^+$, the image of ρ_- approaches $-\infty$. See Figure 11 for a sample graph demonstrating a superattracting period 3 cycle.

Denote the denominator of S_β as $f_\beta(y) = 6y^4 - 8\beta y^3 + 3(1+\beta^2)y^2 + 1 - \beta^2$. In the regime $\beta > 1$, f_β has exactly two roots. This follows from $f'_\beta(y) = 24y((y-\beta/2)^2 + 1/4)$ which means $f_\beta(y)$ is decreasing for $y < 0$ and increasing for $y > 0$. Since $f_\beta(0) = 1 - \beta^2 < 0$, we know that S_β has exactly two poles $y_\pm$ when $\beta > 1$ and we have $y_- < 0 < y_+$. In the case $\beta = 1$, S_β has a single pole at $y = 0$ of order 2.

Next we claim that for $1 < \beta < \sqrt{3}$, we have

$$y_- \;<\; 0 \;<\; \rho_- \;<\; y_+ \;<\; \beta/3 \;<\; \rho_+ \;<\; \beta. \tag{10}$$

That $\rho_- > 0$ and $\rho_+ < \beta$ follow by direct calculation. To show $y_+ < \beta/3$ we compute that $f_\beta(\beta/3) = (1 - \beta^2/3)^2 > 0$ which means that the roots of f_β are less than $\beta/3$. Next we compute that

$$f_\beta(\rho_-) \;=\; \frac{1}{27}\left(5(3-\beta^2)^2 - \beta\sqrt{6-2\beta^2}\,(\beta^2+9)\right).$$

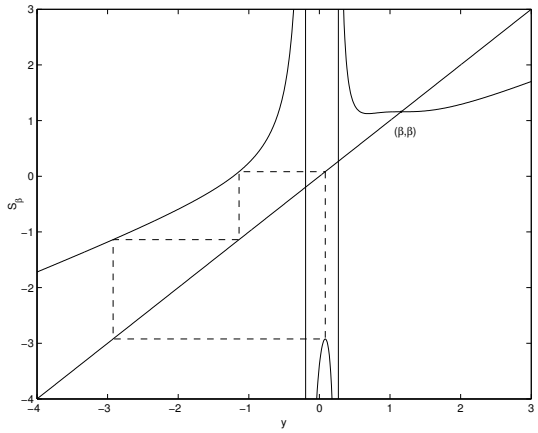

Figure 11: The graph of S_β for $\beta = 1.15838303$ and the web diagram demonstrating a superattracting period 3 cycle containing the free critical point ρ_-.

We see that $f_1(\rho_-) = f_{\sqrt{3}}(\rho_-) = 0$. In other words, ρ_- is a pole of S_β at the extremes $\beta = 1$ and $\beta = \sqrt{3}$. Solving the inequality $f_\beta(\rho_-) < 0$ leads to the expression

$$(3 - \beta^2)(\beta^2 - 1)((\beta^2 - 3)^2 + 16) > 0$$

which is satisfied for the regime $1 < \beta < \sqrt{3}$. This shows that $\rho_- < y_+$ and verifies expression (10).

Using expression (7), we see that S_β is decreasing on the intervals (ρ_-, y_+) and (y_+, ρ_+) and increasing on the intervals $(-\infty, y_-), (y_-, \rho_-), (\rho_+, \beta)$ and (β, ∞). We compute that

$$S_\beta(\rho_-) \;=\; \frac{-8\beta(51 - \beta^2)(3 - \beta^2) + \sqrt{6 - 2\beta^2}\,(7\beta^4 + 78\beta^2 - 9)}{-648 f_\beta(\rho_-)}. \tag{11}$$

The denominator of this expression is positive for $1 < \beta < \sqrt{3}$ and vanishes when $\beta = 1$. Moreover, solving for when the numerator is negative yields the inequality

$$(3 - \beta^2)(\beta^8 - 28\beta^6 + 1222\beta^4 - 3100\beta^2 + 1) < 0$$

which, using Mathematica, is satisfied for $1 \le \beta \le 1.5$. This shows that the image of the critical point ρ_- is negative for $\beta \in (1, 1.5)$, and that

$$\lim_{\beta \to 1^+} S_\beta(\rho_-) \;=\; -\infty.$$

Furthermore, using expression (8) and the results about $f_\beta(y)$, we have that for $\beta \in (1, 1.5)$, $S_\beta(y) > y$ for y in the intervals $(-\infty, y_-)$ and (y_+, β), and $S_\beta(y) < y$ for the intervals (y_-, y_+) and (β, ∞).

The above arguments show that the qualitative properties of $S_\beta(y)$ are as indicated in Figure 11. They also show that it is the orbit of ρ_- which is interesting to follow. The larger free critical point ρ_+ is always attracted to the root at $y = \beta$.

In Figure 12, we graph the image of the free critical point ρ_- as a function of β, for $1 < \beta < \sqrt{3}$. It is clear from the graph that the image of ρ_- decreases as β decreases. Moreover, polynomial division yields

$$S_\beta(y) \;=\; \frac{y}{2} + \frac{\beta}{6} + \frac{O(y^3)}{O(y^4)}$$

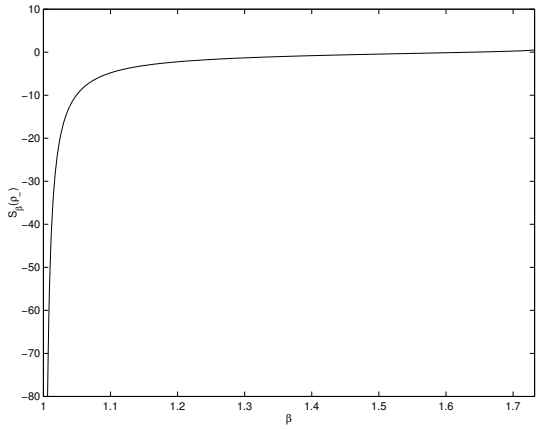

Figure 12: The graph of $S_\beta(\rho_-)$ (the image of the smaller free critical point) as a function of β.

which means that for y large in magnitude, S_β can be approximated by a linear function. The theorem now follows using the same arguments as with Newton's method in Proposition 3.2, where $g_n(\beta)$ is replaced with the function $S_\beta^n(\rho_-) - \rho_-$ and the open interval of continuity is $(1, \xi_n)$ with $\xi_n < \sqrt{3}$. As with Newton's method, to obtain a large period n, the image of ρ_- needs to be very negative which forces β to be very close to 1. □

Table 3 shows the parameter values $\lambda_n = \beta_n\, i$ which yield a superattracting period n cycle for Halley's method applied to p_{λ_n}. These parameter values correspond to polynomials with open sets of initial guesses whose orbits converge to the attracting n cycle rather than to any of the roots of the polynomial. Each parameter value lies at the center of the main cardioid of a period n Mandelbrot-like set and as $n \to \infty$, these sets converge to $\lambda = i$ (see Figure 8).

Table 3: The (approximate) parameter values $\lambda_n = \beta_n\, i$ for which Halley's method applied to p_{λ_n} has a superattracting n cycle.

n	β_n	n	β_n	n	β_n
2	1.34223206	6	1.01865616	10	1.0011575040
3	1.15838303	7	1.00929155	20	1.0000011298
4	1.07647075	8	1.0046367455	25	1.0000000353
5	1.03761109	9	1.0023161274	30	1.0000000011

References

[1] G. Alefeld, "On the convergence of Halley's method," Amer. Math. Monthly **88**, no. 7, 530-536 (1981).

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

[3] 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).

[4] A. Cayley, "Applications of the Newton-Fourier method to an imaginary root of an equation," Quarterly Journal of Pure and Applied Mathematics **16**, 179-185 (1879).

[5] 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).

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

[7] 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).

[8] F. B. Hildebrand, *Introduction to Numerical Analysis* 2nd edition, (McGraw-Hill, New York, 1974).

[9] A. S. Householder, *Principles of Numerical Analysis*, (McGraw-Hill, New York, 1953).

[10] K. Kneisl, "Julia sets for the super-Newton method, Cauchy's method, and Halley's method," Chaos **11**, no. 2, 359-370 (2001).

[11] T. R. Scavo and J. B. Thoo, "On the geometry of Halley's Method," Amer. Math. Monthly **102**, no. 5, 417-426 (1995).

[12] E. Schröder, "Ueber iterirte Funktionen, Mathematische Annalen, **3**, 296-322 (1871).

[13] J. F. Traub, *Iterative Methods for the Solution of Equations* (Chelsea Publishing Company, New York, 1982).

[14] E. Vrscay and W. Gilbert, "Extraneous Fixed Points, Basin Boundaries and Chaotic Dynamics for Schröder and König Rational Iteration Functions," Numer. Math. **52**, 1-16 (1998).

[15] J. Walsh, "The Dynamics of Newton's Method for Cubic Polynomials," The College Mathematics Journal **26**, no. 1, 22-28 (1995).