Newton's Method on a System of Nonlinear Equations

Nicolle Eagan, University at Buffalo
George Hauser, Brown University
Research Advisor: Dr. Timothy Flaherty, Carnegie Mellon University

Abstract

Newton's method is an algorithm for finding the roots of differentiable functions, that uses iterated local linearization of a function to approximate its roots. Newton's method also extends to systems of n differentiable functions in n variables. In this paper, we examine the dynamics of Newton's method on system of two bivariate polynomials. We explore the generalization of Newton's method to systems of two bivariate polynomials, as well as techniques of computer visualization for the corresponding dynamics. In particular, we investigate whether the attracting cycles that arise in the dynamics of Newton's Method on certain cubic polynomials of one complex variable also arise in the case of bivariate quadratics.

1 Introduction

Complex dynamics is defined as the study of dynamical systems in the context of iterated function systems in the complex plane. It is a broad field that can be examined through many different perspectives, including Newton's method. Newton's method, applied to a polynomial equation, allows us to approximate its roots through iteration. Newton's method is effective for finding roots of polynomials because the roots happen to be fixed points of Newton's method, so when a root is passed through Newton's method, it will still return the exact same value. We can see that the points found through iteration of Newton's method correspond to distinct components of the Fatou set. From this, we can determine a function's basins of attraction- a set of starting points in which each point in the set iterates to a particular root. Since this can be easily done for one polynomial, we will consider a system of two nonlinear equations. We will see that by iterating Newton's method on the inverse of the Jacobian matrix for the system, we can calculate the distance for each root and create an image which displays the basins of attraction for the system. We will see that the quadratic systems behave quite like the one variable case, while other systems show interesting results. We will also see that the quadratic systems behave quite like the one-variable case, in that no attracting cycles will be found;

however in other systems, attracting cycles may exist due to the fact that there exists cases where the second derivative maps to noninvertible matrices.

2 Complex Dynamics and Newton's Method

2.1 Newton's Method

As we have said, Newton's method is an iterative algorithm for finding the roots of a differentiable function. But before we define Newton's method precisely, let us make a few normalizing assumptions. In this paper, we will consider Newton's method applied specifically to polynomials either real or complex. The advantage to working with polynomials specifically is that they are well behaved, in that they are infinitely differentiable and have at most finitely many roots and critical points. The advantage to working over $\mathbb{C}$ is that every non constant polynomial over $\mathbb{C}$ has at least one complex root, by the fundamental theorem of algebra.

So let us now define Newton's method. Let $f : \mathbb{C} \to \mathbb{C}$ be polynomial with a root α, so that $f(\alpha) = 0$. Let z_0 be a point chosen roughly to approximate α. Given that f is differentiable at z_0 we may construct an even better better approximation z_1 of α as follows. First we locally linearization of f at z_0. This gives us an affine function $L_{z_0} : \mathbb{C} \to \mathbb{C}$ whose graph is tangent to f at z_0, given by

$$L_{z_0}(z) = f'(z_0)(z - z_0) + f(z_0)$$

Assuming that $f'(z_0) \neq 0$, we may solve the linear equation $L_{z_0}(z) = 0$ and denote the solution z_1:

$$z_1 = z_0 - \frac{f(z_0)}{f'(z_0)}$$

(We note that this assumption is valid for all but finitely many z_0 since f has at most finitely many critical points). Iterating this process, we obtain a sequence of points $(z_n)_n$ that converges, as we will see, to α, provided that z_0 is chosen sufficiently close to α to begin with.

Here is a minor, though theoretically preferable, abstraction of the above outlined procedure:

Given a polynomial f, define the *Newton's function* of f, denoted $N = N_f$, given by

$$N(z) = N_f(z) = z - \frac{f(z)}{f'(z)}$$

Here we generally consider the domain and range of N to be the Riemann sphere (i.e. $\mathbb{C} \cup \infty$) so that N is defined even at singularities of f.

Now, from this perspective, Newton's method can be seen as the iteration of Newton's function on a starting point z_0 that is sufficiently close to α.

2.2 Attracting Fixed Points and Basins of Attraction

We make the following useful observation:

Observation The roots of a polynomial f are precisely the fixed points of N_f.

Proof. Suppose α is a fixed point of N. Then $N(\alpha) = \alpha - \frac{f(\alpha)}{f'(\alpha)} = \alpha$ only if $f(\alpha) = 0$.

In the other direction, suppose α is a root of f of multiplicity m. Then $f(z) = (z-\alpha)^m q(z)$ where q is a polynomial such that $q(\alpha) \neq 0$. We compute

$$\begin{aligned} N(z) &= z - \frac{(z-\alpha)^m q(z)}{m(z-\alpha)^{m-1}q(z) + (z-\alpha)^m q'(z)} \\ &= z - \frac{(z-\alpha)q(z)}{mq(z) + (z-\alpha)q'(z)} \end{aligned}$$

Now, since $q(\alpha) \neq 0$, it follows that $N(\alpha) = \alpha$. □

In addition, roots of f are actually what are known as *attracting* fixed points of N.

Definition A fixed point α of N is said to be *attracting* if $|N'(\alpha)| < 1$. In the special case that $N'(\alpha) = 0$, α is said to be *super-attracting*.

To see that roots of f are attracting fixed points of N, we compute that

$$N'(z) = \frac{f(z)f''(z)}{[f'(z)]^2}$$

Once again using the factorization $f(z) = (z-\alpha)^m q(z)$, one checks that $N'(\alpha) = 1 - 1/m$, where m is the multiplicity of α as a root of f. Therefore roots of f are generally attracting fixed points of N, and moreover single roots are super-attracting fixed points of N.

The technical use of the word "attracting" here is motivated by the local dynamics of N near its attracting fixed points. Imprecisely put, N "pulls" points close to an attracting fixed point α closer to α. Here is the more rigorous version.

Proposition 2.1. *Let α be an attracting fixed point of N. Then the sequence $(N, N \circ N, N^{\circ 3}, \dots)$ of iterates of N converges uniformly to α on some neighborhood of α.*

Proof. The proof of this proposition follows from considering the Taylor expansion of N about α. □

This proposition justifies the previously made claim that Newton's method applied to a starting point sufficiently close to a root of f will necessarily converge to that root.

Moreover, this proposition also implies that any starting point that eventually iterates into such a neighborhood of a root will converge to that root. This motivates the following definition

The *basin of attraction* of an attracting fixed point of N is the set of points that converge to that fixed point under iteration of N.

We note that the basin of attraction itself is an open set.

Basins of attractions of attracting fixed points of N can be visualized with computers in the following way. Create a grid of points representing a points in the complex plane. Iterate Newton's method on each of these points. Then color points that converge to different attracting fixed points of N different colors.

The following figure illustrates this by visualizing the basins of attraction of a cubic polynomial whose roots form the vertices of an equilateral triangle in the complex plane.

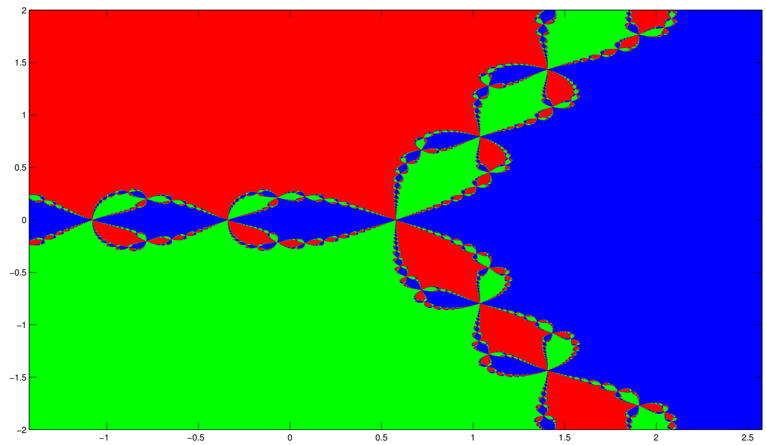

Figure 1: *Newton basins for* $f(z) = (z^2 + 1)(z - \sqrt{3})$ *in* $\mathbb{C}$.

In this image, red represents the basin for the root i, green represents the basin for the root $-i$, and blue represents the basin for the root $\sqrt{3}$. We see three main regions emerge, separated by "bead-shaped" lines. Each main region is a different color, meaning each root has a main region of iteration.

2.3 Julia and Fatou Sets

In the above image, we note that almost every point iterates to one of the three roots. However, we notice points that do not iterate to any root are located on the boundary of the basins. This is known as the *Julia set* [3]. The Julia set is the set on which the dynamics of N are chaotic. One was to see this is to consider a point on the boundary between two basins of attraction. We observe that any neighborhood of that point contains points in multiple basins of attraction. Thus, under iteration of N, this neighborhood gets "torn apart" because different parts of the neighborhood get pulled to different roots.

The *Fatou set* is the complement of the Julia set; it is the set on which the dynamics of N are "well-behaved". More rigorously, it is a set with regular behavior in which the sequence of iterates is a normal family in some neighborhood of a set of points in $\mathbb{C}$. Thus, the sequence of iterates is equicontinuous in the Fatou set. The Fatou set also contains basins of attraction of attracting fixed points. We now arrive at the following question: is the Fatou set entirely basins of attraction for attracting fixed points? Let us consider the basins of another cubic polynomial.

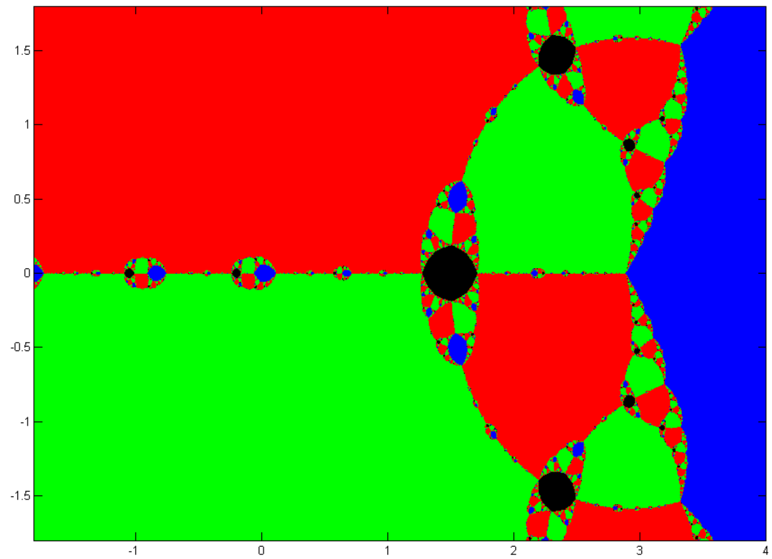

Figure 2: *Newton basins for* $f(z) = (z^2 + 1)(z - 4.5)$ *in* $\mathbb{C}$.

We note that red represents the basin for the root i, green represents the basin for the root $-i$, and blue represents the basin for the root 4.5. We now observe the roughly circular black regions in the image. These black regions, which have positive area, do not iterate to any of the roots. If we pick the starting point $z_0 = 1.4997$, one calculates that $x_1 = N(x_0) \approx -0.1957$ and $N(x_1) \approx 1.4997 = x_0$. Thus, what is known as a *two-cycle* emerges.

Moreover, it is in fact the case the the black regions, in some sense, represent the basin of attraction for this two cycle, in the sense that points in these regions converge under iteration of N to the two cycle.

We make this precise as follows. Denote the two-cycle by $\{x_1, x_2\}$, so that $N(x_1) = x_2$ and $N(x_2) = x_1$. Also denote $N^{\circ 2} = N \circ N$. Observe that $N^{\circ 2}(x_i) = x_i$ for both $i = 1, 2$. From this perspective, both elements of the two-cycle are fixed points of $N^{\circ 2}$. Let us check the derivative of $N^{\circ 2}$ at both x_1 and x_2. Using the chain rule, we see that

$$(N \circ N)'(x_1) = N'(N(x_1))N'(x_1) = N'(x_2)N'(x_1)$$

Likewise,

$$(N \circ N)'(x_2) = N'(x_1)N'(x_2)$$

Plugging in our values, we see:

$$N'(x_1) = 5.998025658, \quad N'(x_2) = 0.0004937078289$$

$$N'(x_1)N'(x_2) = 0.002961272225 < 1$$

Therefore, both x_1 and x_2 are attracting fixed points of $N^{\circ 2}$, and this is precisely what we mean when we refer the two-cycle $\{x_1, x_2\}$ as an *attracting* two-cycle of N.

Therefore, these are *attracting* fixed pints of $N^{\circ 2}$. We now note the following theorem about cycles:

Theorem 1. *An attracting cycle of N attracts at least one critical point of N [4].*

This follows from the fact that all basins of attraction contain at least one critical point of the function. Thus, $N'(x) = 0$ implies $f(x) = 0$ or $f''(x) = 0$. So, if there is an attracting cycle that is not trivial (i.e. not a fixed point or a one-cycle), then it attracts an inflection point of f, or a root of f''. However, if f has degree d, then f'' has degree $d-2$, hence there can be at most $d-2$ non-trivial attracting cycles.

3 Newton's Method in Several Variables

Newton's method can be generalized for finding zeros of systems of n functions in n variables. For our purposes, we consider systems of two bivariate polynomials.

The idea of Newton's method in two variables is the same as in one variable. Just like the one variable case, we choose a starting point, locally linearize, solve the system of linear equations, and repeat.

First, we formalize the notion of a system of equations as a single, vector-valued function

$$f = (f^1, f^2) : \mathbb{R}^2 \to \mathbb{R}^2$$

Like the one variable case, we find the local linearization of f at a vector $v_0 = (x_0, y_0) \in \mathbb{R}^2$ is given by $L_{v_0(v)} = J_{v_0}(v - v_0) + f(v_0)$. Here, J_{v_0} represents the Jacobian matrix of f at v_0, which is a multivariable generalization of the derivative:

$$J_{(x_0,y_0)} = \begin{pmatrix} f^1_x & f^1_y \\ f^2_x & f^2_y \end{pmatrix}\Bigg|_{(x_0,y_0)}$$

where f^1_x is the partial derivative of f^1 with respect to x, and likewise for the other entries of the matrix. Assuming J_{v_0} is invertible, we can solve the linear system $L_{v_0}(v) = 0$ for v. We calculate:

$$v_1 = v_0 - J^{-1}_{v_0} f(v_0)$$

As in the one variable case, this motivates definition of the *Newton function* of f: $N = N_f : \mathbb{R}^2 \to \mathbb{R}^2$, given by:

$$N(v) = v - J^{-1}_v f(v)$$

Thus, Newton's method in two variables can be seen as iteration of N. Essentially, start at a vector v_0, then recursively define $v_{n+1} = N(v_n)$ until within the desired accuracy of a root.

4 Results

4.1 Circle and Ellipse System of Equations

Our first system will be of a circle and ellipse, where the function f is given by:

$$\begin{cases} f^1(x,y) = x^2 + y^2 - 1 \\ f^2(x,y) = \frac{x^2}{a^2} + \frac{y^2}{b^2} - 1 \end{cases}$$

where $0 < b < 1 \leq a$. The Newton function for f is:

$$N(x,y) = (x,y) - \frac{a^2b^2}{2xy(a^2-b^2)} \begin{pmatrix} \frac{y}{b^2} & -y \\ -\frac{x}{a^2} & x \end{pmatrix} \begin{pmatrix} x^2+y^2-1 \\ \frac{x^2}{a^2} + \frac{y^2}{b^2} - 1 \end{pmatrix}$$

We then see the following basins of attraction:

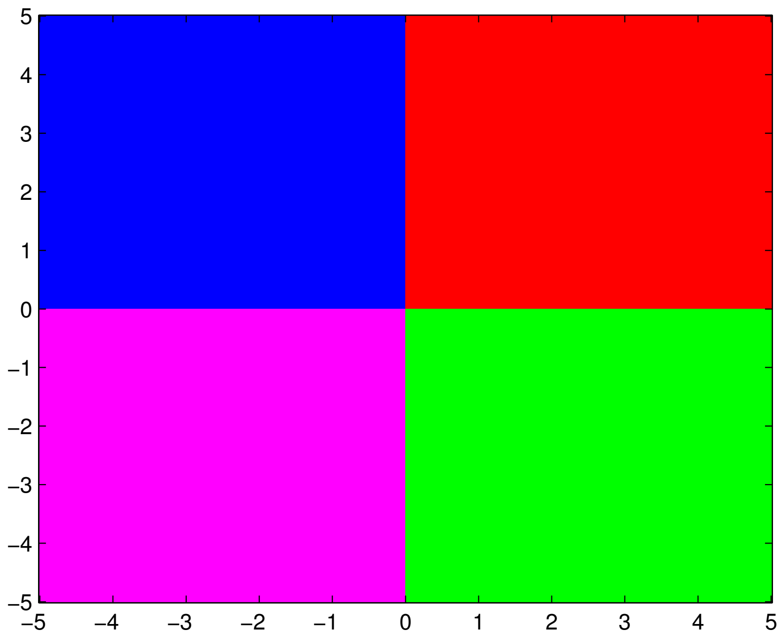

Figure 3: *Newton basins for circle and ellipse with $a = 2$ and $b = 1/2$.*

We notice that the dynamics of the system are quite simple. There are four roots; one root per quadrant. This makes sense with what the image is showing, because each quadrant is a different color, meaning each quadrant iterates to a different root.

4.2 Quadratic System of Equations

We know consider another system of quadratic equations, namely

$$\begin{cases} f^1(x,y) = x^2 - y \\ f^2(x,y) = y^2 - x \end{cases}$$

Newton's function is

$$N(x,y) = (x,y) - \frac{1}{4xy-1}\begin{pmatrix} 2y & 1 \\ 1 & 2x \end{pmatrix}\begin{pmatrix} x^2 - y \\ y^2 - x \end{pmatrix}$$

The following is the image produced:

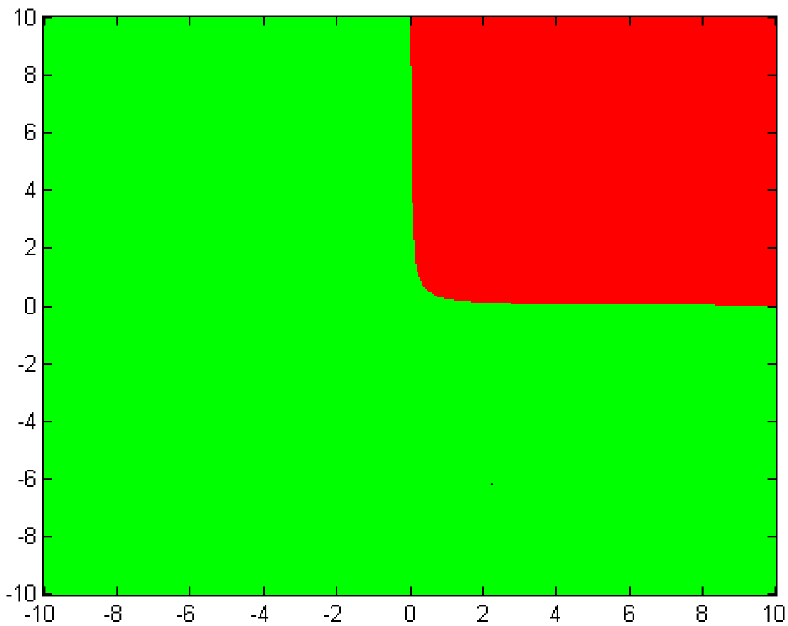

Figure 4: *Newton basins for* $\{x^2 - y, y^2 - x\}$

Seeing theses basins of attraction, let's examine the roots of the system, which are $(0,0), (1,1), (-\frac{1}{2} + \frac{\sqrt{3}}{2}i, -\frac{1}{2} - \frac{\sqrt{3}}{2}i)$, and$(-\frac{1}{2} - \frac{\sqrt{3}}{2}i, -\frac{1}{2} + \frac{\sqrt{3}}{2}i)$. We observe two of these roots have complex coordinates. Thus, we can consider our function f as a function $f : \mathbb{C}^2 \to \mathbb{C}^2$ ($\mathbb{C}^2$ represents the set of two-tuples of complex numbers).

However, basins of attraction in $\mathbb{C}^2$ are hard to visualize, since $\mathbb{C}^2$ is four-dimensional over $\mathbb{R}$. So, to visualize basins of attraction in $\mathbb{C}^2$, we take two-dimensional slices of $\mathbb{C}^2$. Three linearly independent vectors, v, w, and u in $\mathbb{C}^2$, define a 2-D slice:

$$\{av + bw + u : a, b \in \mathbb{R}\}$$

The previously shown basins of attraction in two variables were both the $\mathbb{R}^2$ slice, generated by the vectors $v = (1,0)$, $w = (0,1)$, and $u = (0,0)$. We note that dynamics are generally more interesting on slices that include three or more roots. Using this technique, we can know visualize the following images:

We see very interesting dynamics appear in all of these images. We see "feather-like swirls" appearing in close-ups just off of the origin. These basins

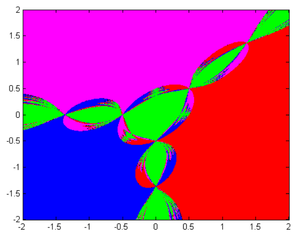

Figure 5: *Newton basins for* $\{x^2 - y, y^2 - x\}$*. Slice generated by* $v = (1,1)$*,* $w = (\zeta_3, \zeta_3^2)$*, and* $u = (0,0)$

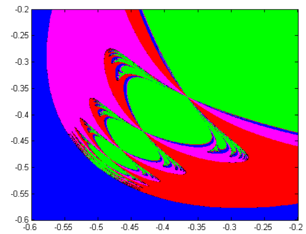

Figure 6: *Newton basins for* $\{x^2 - y, y^2 - x\}$*. Slice generated by* $v = (1,1)$*,* $w = (\zeta_3, \zeta_3^2)$*, and* $u = (0,0)$

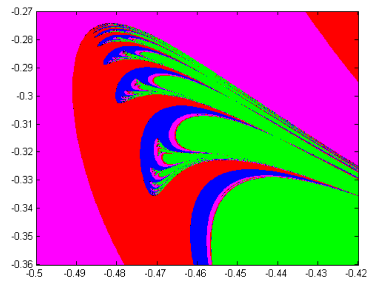

Figure 7: *Newton basins for* $\{x^2 - y, y^2 - x\}$*. Slice generated by* $v = (1,1)$*,* $w = (\zeta_3, \zeta_3^2)$*, and* $u = (0,0)$

of attraction are quite intricate and are a very interesting result.

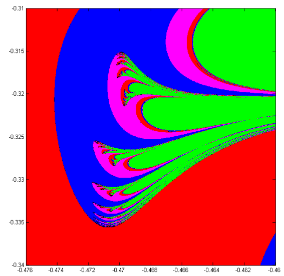

Figure 8: *Newton basins for $\{x^2 - y, y^2 - x\}$. Slice generated by $v = (1,1)$, $w = (\zeta_3, \zeta_3^2)$, and $u = (0,0)$*

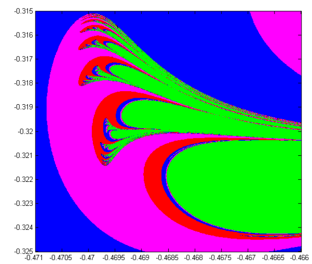

Figure 9: *Newton basins for $\{x^2 - y, y^2 - x\}$. Slice generated by $v = (1,1)$, $w = (\zeta_3, \zeta_3^2)$, and $u = (0,0)$*

4.3 Another Quadratic System of Equations

Our results thus far have not exhibited any attracting cycles. Thus, we now consider a system of two polynomials in two variables that ought to have attracting cycles:

$$\begin{cases} (x^2+1)(y-c) \\ (y^2+1)(x-c) \end{cases}$$

for $c = 9/2$. The roots for this system are $(c,c), (i,-i), (i,i), (-i,i), (-i,-i)$. We also calculate that the Newton function is as follows:

$$N(x,y) = \begin{pmatrix} x & y \end{pmatrix} - \frac{1}{det} \begin{pmatrix} x^2+1 & -2y(x-c) \\ -2x(y-c) & y^2+1 \end{pmatrix} \begin{pmatrix} (y^2+1)(x-c) \\ (x^2+1)(y-c) \end{pmatrix}$$

where $det = -3x^2y^2 + x^2 + y^2 + 1 + 4xy(xc + yc - c^2)$ The following basins of attraction appear:

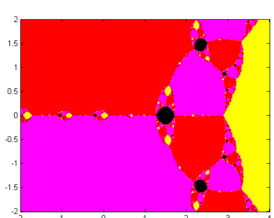

Figure 10: *Basins of attraction for $\{(x^2+1)(y-4.5), (y^2+1)(x-c)\}$, y=x slice.*

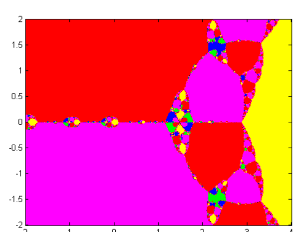

Figure 11: *Basins of attraction for $\{(x^2+1)(y-4.5), (y^2+1)(x-c)\}$, y=x slice translated.*

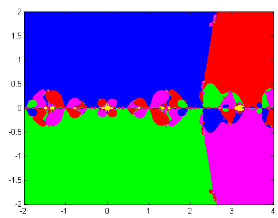

Figure 12: *Basins of attraction for* $\{(x^2+1)(y-4.5), (y^2+1)(x-c)\}$, *y=-x slice.*

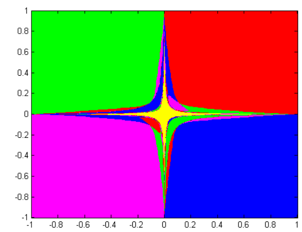

Figure 13: *Basins of attraction for* $\{(x^2+1)(y-4.5), (y^2+1)(x-c)\}$, *imaginary slice.*

The result of *Figure 10* is extremely interesting. It looks strikingly similar to *Figure 2*, meaning that an attracting two-cycle may in fact exist. To see if this is true, we analyze *Figure 11.* We see that *Figure 10* and *Figure 11* are almost the same picture. However, in *Figure 11*, the black regions are now colored in by the other roots that were previously missing from *Figure 10.* Thus, this is not truly an attracting two-cycle, because the black regions are not consistently showing up throughout. The fact that no two-cycle existed is a very interesting result. Even though we constructed the system so that we would see a two-cycle emerge, it still did not happen. The last two figures, *Figure 12* and *Figure 13*, exhibit very intriguing dynamics. *Figure 12* shows very nice elaborations, while *Figure 13* shows a star-looking design.

Our conjecture as to why this behavior occurs is that although the two-cycle does exist, it is "attracting" only in one complex direction, and it is repelling in the transverse complex direction.

In order to test this, we examined the Jacobian matrix of $N \circ N$ near each point on the two-cycle. We found that it had two distinct eigenvalues, one with complex modulus less than one, and one with complex modulus greater than one. The corresponding eigenvectors of the first were in the direction given by the "complex line" $y = x$, and the eigenvectors of the second were in the direction of $y = -x$, which supports our graphical findings.

This would be valuable to explore further in the future. Specifically, it would be valuable somehow to make precise the notion of an attracting fixed point and an attracting cycle in the two-variable case. This notion does not so easily carry over from the single variable case, because the derivative of Newton's function is no longer just a number, but rather a linear map represented in coordinates by the Jacobian matrix. The fact that such a derivative could have multiple distinct eigenvalues complicates the matter significantly.

5 Conclusion

Newton's method not only works with one variable functions, but also with multivariable nonlinear systems of equations to find the system's basins of attraction. Through our investigations, we saw Newton's method reveals interesting dynamical behavior in several variables, but we were unable to find any true attracting cycles. There were several behaviors we observed that we would love to research further in the future, such as actually finding a true attracting cycle, if one exists. It would also be interesting to explore the dynamics of more complicated nonlinear systems to see if any more interesting appear, and to also see if any characteristics of previously explored systems hold.

References

[1] Blanchard, P., "The Dynamics of Newton's Method" *Proceedings of Symposia in Applied Mathematics*, Vol. 49, pp. 139-155, 1994.

[2] Hubbard, J., Schleicher, D., Sutherland, S., *How to Really Find Roots of Polynomials by Newton's Method*, December 17, 1998.

[3] Alexander, D., Devaney, R., *A Century of Complex Dynamics*, April 1, 2014.

[4] Milnor, J., *Dynamics in One Complex Variable, Third Edition*. Princeton University Press, Princeton, New Jersey, 2006.