Spring 6-11-2022

Analyzing Domain of Convergence for Broyden's Method

Michael Bonthron
DePaul University, MBONTHRO@depaul.edu

Analyzing Domain of Convergence for Broyden's Method

Thesis by

Michael R. Bonthron

In Partial Fulfillment of the Requirements

For the Degree of Applied Mathematics

Master of Science

DePaul University

Chicago, Illinois

ABSTRACT

Analyzing Domain of Convergence for Broyden's Method

Broyden's method is a quasi-Newton iterative method used to find roots of non-linear systems of equations. Research has shown and improved the rate of convergence for special cases and specific applications of the method. However, there is limited literature regarding the well-posedness of the method. In practice, a numerical method must reliably converge to the appropriate root. This paper will discuss the domain of attraction for the roots of a system found by using Broyden's method. A method of approximating the radius of convergence of a root will be described which considers the largest disk centered at the root such that all values within the disk converge to the root. Literature on Broyden's method has conflicting claims about the initial approximation of the Jacobian. Plots will demonstrate the effect of the initial guess of the Jacobian for the iterative scheme. In this paper, the importance of using a finite difference approximation for the initial guess of the Jacobian will be shown through examples of 2×2 systems of equations.

ACKNOWLEDGEMENTS

I would first like to acknowledge DePaul University's mathematics department. DePaul University and its enthusiastic professors made this work possible.

I would like to thank my advisor Dr. Ilie Ugarcovici who, from the first time I set foot on campus, encouraged research and taking a deeper look at the material presented in class. I am appreciative of such a supportive network at DePaul that allowed the flexibility in the program to take research credits and challenging courses.

I would also like to thank Professor Kevin Meade (Illinios Institute of Technology) who met weekly throughout the pandemic and held me responsible for my research. Professor Meade has also been an invaluable sounding board and has helped shape me into the researcher that I am.

TABLE OF CONTENTS

Chapter 1

Background

This section will provide background on related research and applications of Broyden's method for solving systems of non-linear equations.

1.1 Motivation

Solving or finding roots of systems of equations are required in many engineering and physics problems. In engineering, assumptions are made to mathematically represent a physical system. Governing equations such as Bernoulli's law or conservation of mass, among others, are then used to formulate a system of n equations and m unknowns. Solutions to the system can determine design parameters, factors of safety, or optimize designs. Solutions provide answers to engineering questions, but methods of finding the solutions are in the domain of mathematics.

It is common practice to use computational methods to find solutions to systems when analytical solutions are not readily available. Analytical solutions cannot be reliably used to solve every system of equations. Depending on the size, assumptions, and complexity of the model, analytical solutions can be unreasonable or impractical to formulate. Even in the one-dimensional case, roots of the equation cannot always be found analytically. It is well known that polynomials of degree 5 or greater have no closed form solution [1].

For one dimensional systems, Newton's method is a familiar iterative scheme which can find the equation's roots. For higher dimension systems, Newton's method can be modified to apply to systems of equations, but not without its limitations. Newton's method for systems of equations requires an expression of the Jacobian of the system. An explicit calculation requires the partial derivative of each function with respect to each variable, demanding n^2 partial derivatives for an $n \times n$ system [2] [3] [4]. Alternatively, approximating the Jacobian with a so-called "Secant method" is not always a practical alternative, still requiring $2n^2$ evaluations of the function [2]. So called quasi-newton methods have been developed as alternatives which seek to resolve the problems caused by the Jacobian.

One popular quasi-newton method is Broyden's method, an iterative scheme that approximates the root of the function and the Jacobian of the function [1]. Each iteration, the approximation of the Jacobian is used to improve the approximation of the root. The Jacobian approximation is improved using the evaluation of the function. Broyden's method locally converges q-superlinearly, slower than Newton's method, but without requiring an explicit formulation of the Jacobian [4]. For this reason, Broyden's method is a popular alternative to Newton's method.

1.2 Applications of Broyden's Method

Broyden's method has been widely applied and studied. Applications include finding solutions to the Copenhagen problem, a specific case of the three body problem [5]. The method has been studied regarding special cases of mixed linear and non-linear systems of equations [6] or studying the effect of combining Broyden's method with other well-known iterative schemes [7]. The improvement of the method is a relevant research area to the study of numerical methods and optimization.

Numerical methods can be evaluated by a number of criteria. Accuracy, computation time, and computer storage required were originally used to argue the power of Broyden's method [2]. This paper uses a fourth merit to assess an algorithm: the reliability of convergence. There are many papers interested in improving the rate of convergence of the method [6] [7] [8], however there is limited literature regarding the reliability of the convergence. Broyden's method, like many other iterative methods, requires an initial guess "close" to a root. The understanding of what is meant by "close" is vague and without a formal definition [4]. In practice, the location of roots will be unknown, making impossible to know if an initial guess is "close" to the root. Ideally, a method will converge despite poorly chosen initial guesses. Broyden's method requires two initial guesses: an initial guess of the root and an initial guess for the Jacobian. How important is it that these initial guesses are close to the root? Are both parameters equally important for convergence? This paper explores both questions using numerical examples.

1.3 Initial Guess of the Jacobian

Unlike Newton's method, Broyden's method requires two inputs: an initial approximation of the root and an initial approximation of the Jacobian. The initial guess of the root is straightforward and should be the closest approximation to the root using the available information of the system. What to use as the initial approximation of the Jacobian is not as clear. In the original paper, Broyden used a forward finite difference approximation at the initial guess of the root to approximate the Jacobian [2]. Forward finite difference approximation was also used in [9]. Towaiq initialized the Jacobian with a central finite difference approximation and showed examples of faster convergence [10]. Dennis generally references using a finite difference method, but does not specify any method in particular [4]. Kalvouridis explicitly calculates the Jacobian and uses the exact formulation for the initial guess of the Jacobian in Broyden's method [5] . Lastly, Sauer states that if no knowledge is available about the Jacobian, the identity matrix can be used as the initial guess for the Jacobian [3].

These are contradicting claims about how to initialize the method. Currently, to the best of the author's knowledge, there is no discussion regarding the effect or the importance of the initial approximation of the Jacobian. This paper will consider four numerical methods of varying computational complexity to investigate the effect of the initial approximation of the Jacobian.

1.4 Domain of Attraction

Broyden's method requires that an initial guess of the root be sufficiently close to the root to ensure superlinear convergence. Ideally, a well posed system will converge, even for guesses not initially close to the root. "How close must an initial guess be to guarantee convergence?" and "What is the set of points of initial guesses that converge to a root?" are important questions to consider when evaluating a numerical method.

The domain of attraction for different numerical methods for a system of single non-linear oscillators have been studied, but so far are limited in scope to specific applications [9]. In physics, [5] studied the domain of attraction for a specific case of the three body problem. The paper by Kalvouridis explored how many points within a specified region successfully converged using Broyden's method compared to the number of points which converged using Newton's method.

Both [5] and [9] used plots to show where initial guesses converged. In both papers, the plots exhibited a chaotic nature. Cheney considered the domain of attraction for Newton's method [11]. The plots presented using Newton's method demonstrated a fractal pattern of symmetry. This paper investigates the domain of attraction for Broyden's method.

Graphically, this paper will show similar plots to [5] and [9] demonstrating the convergence behavior for different initial guesses. This paper will introduce two methods of approximating the radius of convergence by considering the largest disk centered around a root, such that all values within the disk or on the edge of the disk converge to the root at the center. These graphical and numerical representations of the domain of attraction will be used to compare numerical methods.

1.5 Paper Structure

So far, this paper has introduced the topic of Broyden's method for solving systems of non-linear equations. This paper will work through the original motivation of Broyden's method by considering the limitations of Newton's method. In chapter 3 this paper will discuss the domain of attraction. This paper will introduce a method of graphically presenting the root to which an initial guess converges. Graphs will be presented which demonstrate a range of structures of both chaotic and non-chaotic behavior. These graphs will show the beauty and complexity of this numerical method.

This paper will present two methods of numerically approximating the region around a root which will converge to the root for a given method. The numerical approximation and the graphical representation of the radius of convergence and domain of attraction will be used to demonstrate the strength of a numerical method. This paper will present a number of examples for a range of varying equation complexity. In each example, graphical representation of the domain of attraction will be presented along with the numerical approximation of the radius of convergence. Finally, this paper will conclude by summarizing the findings and emphasizing the importance of approximating the Jacobian with a finite difference method when initializing the method.

Chapter 2

Broyden's Method Theory, Derivation, and Application

This section will provide background on Broyden's method for solving systems of non-linear equations. First, Newton's method of solving systems of equations will be described as well as the limitations of the model. These disadvantages will motivate the derivation of Broyden's method. Examples of Broyden's method will show the intermediate calculations performed.

2.1 Newton's Method for Systems of Non-Linear Equations

For systems of one variable, Newton's method for finding roots of a function $f(x)$ is widely known to be:

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)} \tag{2.1}$$

At a simple root ($f'(r) \neq 0$), the method has quadratic convergence while at a root of higher multiplicity, the method only converges linearly based on the multiplicity of the root [3] [4]. For both cases, the convergence requires an initial guess "close" to the root. Global convergence is not guaranteed. The quadratic convergence for simple roots makes the method an attractive option for finding roots in one dimension.

The method can be expanded to consider systems of equations with higher dimensional analogs of each component. Consider the vector value function:

$$F(\boldsymbol{x}) = (f_1(\boldsymbol{x}), f_2(\boldsymbol{x})..., f_n(\boldsymbol{x})) \tag{2.2}$$

A root of $F(\boldsymbol{x})$ is a value for the vector $\boldsymbol{x}$ such that:

$$\begin{aligned} f_1(x_1, x_2, ..., x_n) &= 0 \\ f_2(x_1, x_2, ..., x_n) &= 0 \\ &\vdots \\ f_n(x_1, x_2, ..., x_n) &= 0 \end{aligned} \tag{2.3}$$

Although it might be impossible to know mathematically if a system has a solution, there is an intuitive expectation that a solution will exist for those that represent physical systems [2].

The analog of the derivative, the Jacobian, is needed to apply Newton's method to a system of equations. The Jacobian of a system of equations is defined as:

$$DF(\boldsymbol{x}) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_n} \\ \vdots & \vdots & & \vdots \\ \frac{\partial f_n}{\partial x_1} & \frac{\partial f_n}{\partial x_2} & \cdots & \frac{\partial f_n}{\partial x_n} \end{bmatrix} \tag{2.4}$$

The multivariate Newton's method uses the Jacobian as the multidimensional analog to the derivative. To apply Newton's multivariate method, the Jacobian must be non-singular and must be Lipschitz continuous in a neighborhood around the root [1]. Newton's multivariate method is expressed as:

$$\boldsymbol{x}_{k+1} = \boldsymbol{x}_k - DF(\boldsymbol{x}_k)^{-1} F(\boldsymbol{x}_k) \tag{2.5}$$

Newton's method for solving systems of equations also convergences q-quadratically [4]. Although the convergence of the method is appealing, the method is not without its drawbacks, primarily with regards to the Jacobian. A considerable disadvantage of Newton's method is the need to explicitly calculate the Jacobian. For an $n \times n$ system of equations,

n^2 partial derivatives are required to formulate Newton's method. Finite difference approximations could be used to approximate the Jacobian at each iteration, but if the function is computationally expensive to evaluate, this is not a practical solution [2]. In addition to calculating the Jacobian, Newton's method requires the inverse of the Jacobian. Inverting matrices is another computationally expensive task requiring n^3 operations for an $n \times n$ matrix [12]. Both the requirement for the Jacobian and the need for an inverse are solved with methods developed by Broyden.

2.2 Broyden's Method I

Broyden's method eliminates the need to explicitly calculate the Jacobian by approximating the Jacobian each iteration. The method requires no additional evaluations of the function than what would be required with Newton's method.

An approximation of the Jacobian on the k^{th} iteration must satisfy the equation:

$$A_k \delta_k = \Delta_k \tag{2.6}$$

where:

$$\begin{aligned} \delta_k &= \boldsymbol{x}_k - \boldsymbol{x}_{k-1} \\ \Delta_k &= F(\boldsymbol{x}_k) - F(\boldsymbol{x}_{k-1}) \end{aligned} \tag{2.7}$$

Equation 2.6 does not define a single matrix, but rather a class of matrices. Additional constraints are needed to define how to update the Jacobian approximation. In Broyden's method I, the additional constraint requires that the approximation to the Jacobian only be updated in the direction of δ_k. In any direction perpendicular to δ_k, the approximation to the Jacobian must remain the same between iterations. This is to say:

$$A_k \omega = A_{k-1} \omega \tag{2.8}$$

for all ω satisfying:

$$\delta_k^T \omega = 0 \tag{2.9}$$

One method of updating the Jacobian that satisfies both conditions is:

$$A_k = A_{k-1} + \frac{(\Delta_i - A_{k-1}\delta_i)\delta_k^T}{\delta_k^T \delta_k} \tag{2.10}$$

Broyden's method I uses the approximation of the Jacobian each iteration to update the root. In addition, each iteration the Jacobian is updated in the direction δ_k to include the new information about the function that has been found. Expressed as an algorithm, Broyden's method I is:

Broyden's Method I

x_0 := initial guess of root

A_0 := initial guess of Jacobian

for k = 1,2,3... **do**

$$\boldsymbol{x}_{k+1} = \boldsymbol{x}_k - A_k^{-1} F(\boldsymbol{x}_k)$$
$$A_{k+1} = A_k + \frac{(\Delta_k - A_k \delta_k)\delta_k^T}{\delta_k^T \delta_k}$$

The advantage to Broyden's method I is that there is no need to explicitly calculate the Jacobian for the method. The Jacobian is not approximated using a finite difference method, but by using the evaluations of the function as the method progresses.

Broyden's method I still requires the inverse of the approximation of the Jacobian. Since inverting a matrix can be computationally expensive, rather than approximating the Jacobian each iteration, it is possible for the method to be modified to approximate the inverse of the Jacobian.

2.3 Broyden's Method II

Broyden's method II is a modification to Broyden's method I which approximates the inverse to the Jacobian rather than approximating the Jacobian itself each iteration. The derivation is explained in [3].

$$A_k^{-1} = A_{k-1}^{-1} + \frac{(\delta_k - A_{k-1}^{-1}\Delta_i)\delta_i^T A_{k-1}^{-1}}{\delta_k^T A_{k-1}^{-1}\Delta_k} \tag{2.11}$$

If the approximation to the inverse of the Jacobian is represented as B_k then the iterative scheme can rely on only considering the approximation to the inverse, B_k by:

Broyden's Method II

x_0 := initial guess of root

B_0 := initial guess of Inverse of Jacobian

for k = 1,2,3... **do**

$$\boldsymbol{x}_{k+1} = \boldsymbol{x}_k - B_k F(\boldsymbol{x}_k)$$
$$B_{k+1} = B_k + \frac{(\delta_k - B_k\Delta_i)\delta_i^T B_k}{\delta_k^T B_k \Delta_k}$$

Broyden's method II improves computation time when the inverse of the Jacobian is expensive to calculate. Unless there is a need for the Jacobian during each iteration, Broyden's method II is preferred over Broyden's method I since it removes the need for an inversion of a matrix. This paper uses Broyden's method I over Broyden's method II since the inversion of a 2×2 matrix is inexpensive, and the surrounding literature used the first method. Additional research considering the application of both methods will be discussed in the future work section of this paper.

Both of Broyden's methods share the same rate of convergence and same challenges of global convergence. Broyden's method converges q-superlinearly, as shown in [4]. Although the rate of convergence is slower than in Newton's method, in practice, the computation time for Broyden's method can be faster if the Jacobian is computationally complex to evaluate, or the function is sufficiently expensive to evaluate.

2.4 Termination Criteria for Broyden's Method

Broyden's method, like all numerical methods, requires termination criteria to end the scheme. Since Broyden's method is used to find roots for systems of equations, the norm of the function is a satisfactory way to measure the error of the system. Unless otherwise specified, this paper will use the termination criteria of six decimal points of accuracy:

$$|F(\boldsymbol{x})| \leq 0.5 \times 10^{-6} \tag{2.12}$$

In addition, there must practically be some upper limit on the number of iterations allowed before the program exits. Unless otherwise stated, this paper will use 100 for the number of iterations before terminating. Increasing the number of iterations will allow for more initial guesses to converge, although 100 iterations serves as a practical limitation. If a system is unable to achieve the termination criteria of 2.12 within 100 iterations, the initial guess will be considered as non-converging.

Since Broyden's method I requires an inverse of a matrix, the case of a singular matrix must be discussed. In the event the approximation to the Jacobian is singular, the method will terminate and will be denoted accordingly, but considered non-converging. Alternative ways of handling singular matrices could include re-approximating the Jacobian with a finite difference method. This paper will not consider any reapproximation methods and will mark cases where a singular matrix occurs.

2.5 Example Application

Broyden's method can be applied to solve the general problem of finding the roots to a system of non-linear equations. As discussed in the background section of this paper, there are applications for this method in engineering, physics, computing, economics and any other situation where a problem can be formulated as a system of equations whose roots are of interest. This section will demonstrate the application of Broyden's method to two examples. In each example the intermediate calculations for each iteration will be shown.

2.5.1 Unconstrained Optimization

This section will show how Broyden's method can be used to solve unconstrained optimization problems. Consider the objective function of two variables:

$$f(x, y) = x^2 + 2xy - 4y^2 + 4x - 6y + 4 \tag{2.13}$$

To minimize the function, critical points are first located such that:

$$\begin{aligned} \frac{\partial f}{\partial x} &= 0 \\ \frac{\partial f}{\partial y} &= 0 \end{aligned} \tag{2.14}$$

The partial derivatives are found to be:

$$\begin{aligned} f_x(x, y) &= 2x + 2y + 4 \\ f_y(x, y) &= 2x - 8y - 6 \end{aligned} \tag{2.15}$$

In this case, critical points can be found using linear algebra, although it is easy to imagine examples where this is not the case. It is well known that the minimum of the function will occur at a critical point such that:

$$\begin{aligned} f_1(x, y) &= 2x + 2y + 4 = 0 \\ f_2(x, y) &= 2x - 8y - 6 = 0 \end{aligned} \tag{2.16}$$

Broyden's method can be used to find values of (x, y) such that equation 2.16 is satisfied. Beginning with the initial guess of the root as $(2, 2)$ and with an initial guess of the Jacobian as the 2×2 identity matrix, we can follow the calculations performed by Broyden's method with each iteration:

After only 4 iterations, the method converges to a critical point at $(-1, -1)$. Plotting the root approximation after each iteration results in Figure 2.1. It is interesting to note that the first, second, and third approximation of the root all have a greater error than the initial

k	x_k	A_k	$\|F(x_k)\|$
0	$(2,2)$	$\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}$	21.63
1	$(-10,20)$	$\begin{bmatrix}0.38 & 0.92\\ 4.76 & -6.15\end{bmatrix}$	187.54
2	$(-6.45,-7.47)$	$\begin{bmatrix}0.27 & 1.77\\ 4.95 & -7.618\end{bmatrix}$	47.36
3	$(3.55,4.04)$	$\begin{bmatrix}1.10 & 2.75\\ 3.54 & -9.29\end{bmatrix}$	39.51
4	$(-1,-1)$	-	0

Table 2.1: Intermediate calculations used in Broyden's method for equation 2.16

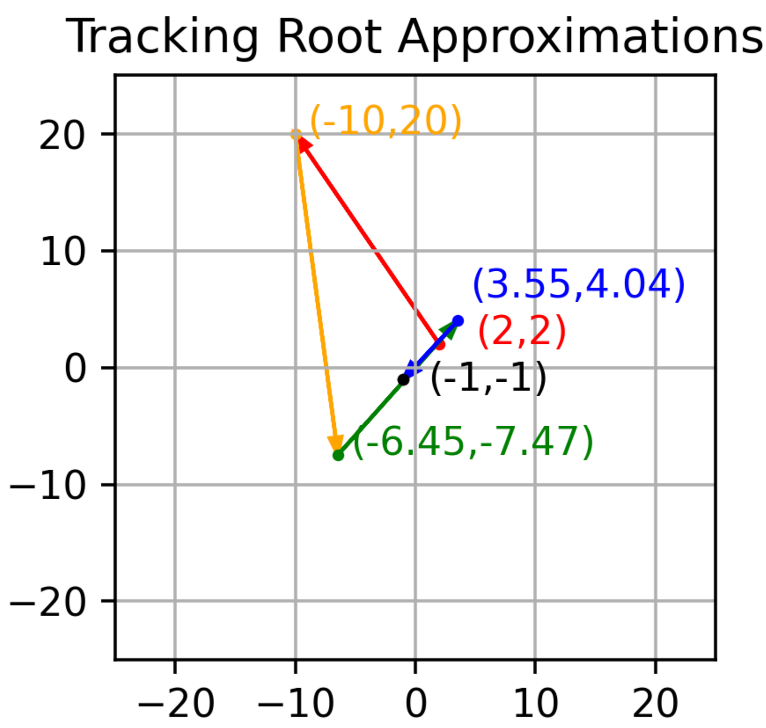

Figure 2.1: Plotting δ_i for each iteration from table 2.1

guess. The reason that the approximation does not necessarily improve with each iteration is due to the initial approximation of the Jacobian. For the first iteration, the Jacobian was arbitrarily approximated using the identity matrix. Each iteration, the approximation to the Jacobian is only updated in the direction δ_k. However, just 4 iterations and updates to the approximation to the Jacobian are necessary to reduce the error to six decimal points of accuracy. It is interesting to note that the final approximation of the Jacobian is not near the true value of the Jacobian at the root. Although it is often the case that the Jacobian approximation does converge to the true Jacobian, there is no guarantee this is the case [4].

In practice, the critical point will need to be analyzed to check if it is a local maximum, minimum, or saddle point. This method does not prove that $(-1, -1)$ is the only critical point of the function. For this example, and with a linear set of equations for the partial derivatives of the objective function, it works out that $(-1, -1)$ is the only critical point and is a global minimum, however this is not necessarily the case. If a system has multiple roots, Broyden's method could converge to different roots based on the initial guesses for the root and the Jacobian. Without analytical insight, there is no way of knowing for certain that all the roots of a particular system have been found. In the following sections this paper will further discuss the nature of converging to different roots based on the nature of the initial guesses of the method.

2.5.2 Inverse Kinematics

This section will show how Broyden's method can be applied to an inverse kinematics question. Consider a simple 2-dimensional robot arm that has two degrees of motion: being able to rotate around its base and at a joint located halfway along its arm 2.2. The location of the end of the robot arm can be expressed by:

$$\begin{aligned} x &= l_1 \cos(\theta) + l_2 \cos(\psi) \\ y &= l_1 \sin(\theta) + l_2 \sin(\psi) \end{aligned} \tag{2.17}$$

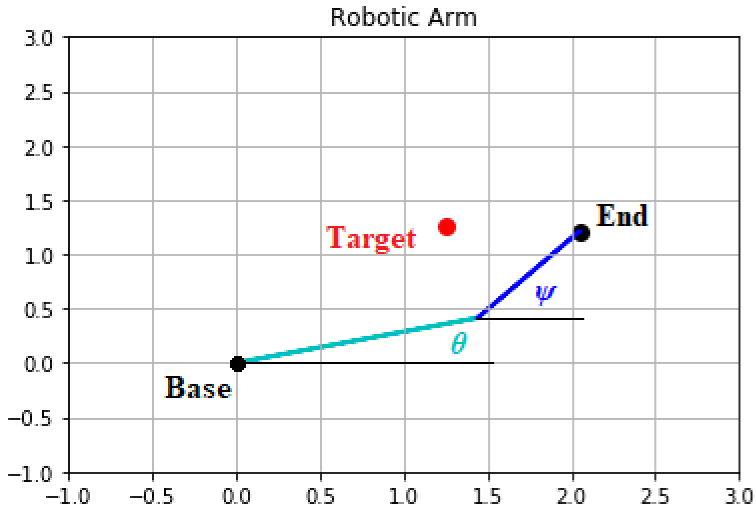

Figure 2.2: Diagram of the robotic arm described by the equation 2.17

where: θ is the angle between the base and the first portion of the arm, ψ is the angle between the horizontal and the second portion of the arm, l_1 is the length of the first portion of the arm, and l_2 is the length of the second portion of the arm.

Given a target position (x_{tar}, y_{tar}), a parameter of interest is what values must θ and ψ be in order to position the end of the robot arm at the target location. This can be expressed as finding the roots of the system of equations:

$$\begin{aligned} f_1(\theta, \psi) &= x_{tar} - l_1 \cos(\theta) - l_2 \cos(\psi) = 0 \\ f_2(\theta, \psi) &= y_{tar} - l_1 \sin(\theta) - l_2 \sin(\psi) = 0 \end{aligned} \tag{2.18}$$

Consider the target position $(1.25, 1.25)$ with arm length $l_1 = 1.5$ and $l_2 = 1$. Using Broyden's method to find a root of 2.18 with an initial guess of $(0, 0)$ for the root and using the 2×2 identity matrix for the initial guess of the Jacobian results in the following iterations.

The nature of the *sine* and *cosine* terms increases the difficulty for the method to converge. The error of the approximation does not steadily decrease over the first few iterations. Once again, the approximation to the Jacobian requires several iterations before it becomes a "good" approximation. After 5 iterations, the Jacobian becomes reasonable enough that the error of the method decreases for all subsequent iterations.

k	x_k	A_k	$\lvert F(x_k)\rvert$
0	$(0, 0)$	$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$	1.767
1	$(1.25, -1.25)$	$\begin{bmatrix} 1.18 & -0.18 \\ 0.31 & 0.68 \end{bmatrix}$	0.902
2	$(0.72, -2.13)$	$\begin{bmatrix} 0.85 & -0.73 \\ -0.23 & -0.22 \end{bmatrix}$	1.285
3	$(2.50, 0.84)$	$\begin{bmatrix} 1.12 & -0.29 \\ -0.29 & -0.32 \end{bmatrix}$	1.836
4	$(0.96, 1.05)$	$\begin{bmatrix} 1.18 & -0.30 \\ 0.24 & 0.39 \end{bmatrix}$	0.854
5	$(0.42, -1.42)$	$\begin{bmatrix} 1.21 & -0.20 \\ 0.10 & -1.02 \end{bmatrix}$	1.644
6	$(0.90, 0.21)$	$\begin{bmatrix} 1.10 & -0.56 \\ 0.08 & -1.10 \end{bmatrix}$	0.664
7	$(1.45, 0.12)$	$\begin{bmatrix} 1.24 & -0.58 \\ -0.56 & -0.99 \end{bmatrix}$	0.370
8	$(1.26 - 0.13)$	$\begin{bmatrix} 1.59 & -0.09 \\ -0.48 & -0.87 \end{bmatrix}$	0.196
9	$(1.37, -0.24)$	$\begin{bmatrix} 1.56 & -0.07 \\ -0.37 & -0.98 \end{bmatrix}$	0.025
10	$(1.38, -0.22)$	-	0.004

Table 2.2: Intermediate calculations used in Broyden's method for equation 2.18

k	x_k	A_k	$\|F(x_k)\|$
0	$(1.5, 0)$	$\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}$	0.767
1	$(1.50, 0.76)$	$\begin{bmatrix}1 & -0.24\\0 & 1.41\end{bmatrix}$	0.367
2	$(1.63, 0.54)$	$\begin{bmatrix}0.57 & 0.48\\0 & 1.41\end{bmatrix}$	0.218
3	$(2.01, 0.54)$	$\begin{bmatrix}-0.95 & 0.48\\0 & 1.41\end{bmatrix}$	0.582
4	$(1.40, 0.54)$	-	0.003

Table 2.3: Intermediate calculations used in Broyden's method for equation 2.19

Of course, this example could be reformulated such that it wouldn't require any *sine* or *cosine* terms. As an alternate formulation, rather than trying to set the angles of each section of the arm, the position of the joint can be solved using constraints from the distance formula between the target point and the base:

$$\begin{aligned} f_1(x, y) &= l_1 - \sqrt{x^2 + y^2} = 0 \\ f_2(x, y) &= l_2 - \sqrt{(x_{tar} - x)^2 + (y_{tar} - y)^2} = 0 \end{aligned} \tag{2.19}$$

where (x, y) represents the position of the joint of the robot arm. It can be shown that this new formulation converges to a root much faster in table 2.3.

Many applications use Broyden's method to find roots of systems of equations. This section demonstrated two examples and showed the calculations performed by the method. An optimization type question was solved by considering the partial derivatives of the objective function. Critical points were found by solving the system of equations when the partial derivatives were zero. The second example in this section demonstrated an inverse kinematics type question. By reformulating the problem to remove trigonometric functions, it was shown that the method demonstrated superior convergence.

This section arbitrarily selected initial guesses and used the identity matrix for the initial guess of the Jacobian which were sufficient to satisfy converging criteria of two decimal points within 10 iterations. However, it is not the case that all initial guesses will converge to a root. If a system has multiple roots, it is also interesting to consider which root is found for a given initial guess. The next section will graphically consider the domain of attraction for a system with multiple roots.

Chapter 3

Domain of Attraction

This section will consider the domain of attraction of a root and its importance in numerical methods. After defining the domain of attraction, this section will describe how the domain of attraction can be visualized using plots for systems of two equations and two variables. This section will show both chaotic and non chaotic behavior based on the functions used in the method. This section will conclude by demonstrating the effect the initial guess of the root and the initial guess of the Jacobian has on the domain of attraction.

3.1 Definition

The domain of attraction, $\mathscr{D}$ for a root r, is the region of initial guesses that will converge to a specified root for a given method. Formally, the domain of attraction is the set of all points $\boldsymbol{x}_0$ such that the method, beginning at $\boldsymbol{x}_0$, will converge to that zero [11].

$$\mathscr{D}(r) = \{\boldsymbol{x} \in \mathbb{R}^n | \lim_{i \to \infty} B_i(\boldsymbol{x}) = \boldsymbol{r}\} \tag{3.1}$$

Where $B_i(\boldsymbol{x})$ represents i iterations of Broyden's method.

If a system of equations has only one root, the domain of attraction is all values which converge to the root. If a system of equations has multiple roots, then it is interesting to note where each initial guess converges.

Ideally, the domain of attraction for a root will have three features: located near the root, large, and predictable near the root. Broyden's method requires the initial guess be close to the root, but a method that is able to converge for initial guesses with 1×10^{-1} of the true root is much preferred over a method that requires the initial guesses to be within 1×10^{-6} of the root. Ideally, there is a large region surrounding the root that will converge to the root so the initial guess does not require a high level of accuracy. A predictable domain of attraction is analogous to a well-posed system. Ideally, values that converge to the same root will be located nearby one another. If initial guesses near one another converge to different roots, then the system is not well posed. A small change in the initial guess leading to a significant change in the output of the method is not desirable.

3.1.1 System with One Root

Consider the following system of equations:

$$\begin{aligned} f_1(x,y) &= y - x^3 = 0 \\ f_2(x,y) &= y - \tfrac{1}{2}x^2 + \tfrac{1}{2} = 0 \end{aligned} \tag{3.2}$$

There is only one real root for this system located at $(1,1)$. The domain of attraction for the system is all values which converge to the root. Practically, we cannot run the method for infinite iterations, so some cutoff value must be declared. After 100 iterations, if the termination criteria for the method has not been reached, it will be treated that the method did not converge for the initial guess. Plotting all the points which converge to the root $(1,1)$ using Broyden's method I results in Figure 3.1.

3.1.2 System with Two Roots

Consider the following system of equations 3.3:

$$\begin{aligned} f_1(x,y) &= y - x^2 = 0 \\ f_2(x,y) &= y - 1 = 0 \end{aligned} \tag{3.3}$$

By substitution, we find that the roots of the equation are $(1,1)$ and $(-1,1)$. Consider how the initial guess of the root affects which of the two roots the method converges to. For

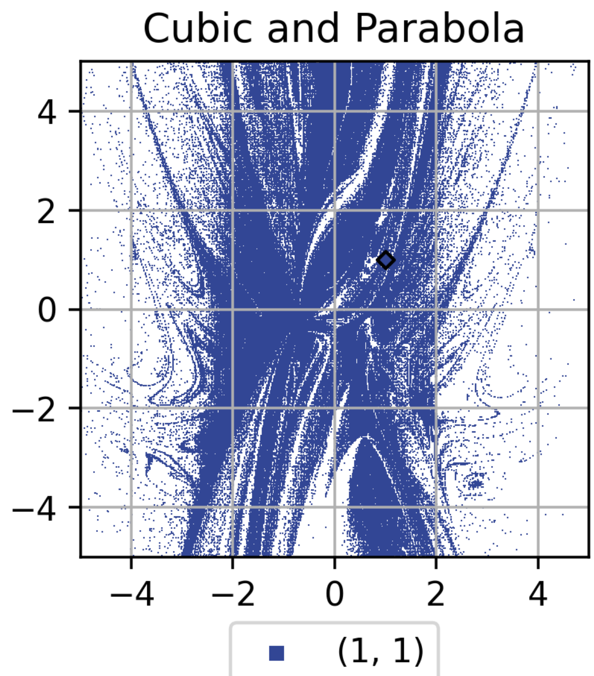

Figure 3.1: Domain of attraction for equation 3.2

an array of initial guesses evenly spaced out in the Cartesian plane, we can track which root is found, and color all the initial guesses which find the root $(1, 1)$ blue and all the initial guesses which find the root $(-1, 1)$ red. Values which do not converge after 30 iterations are left uncolored, and initial guesses which encounter a singular matrix during the method are colored black. Using Broyden's method with the initial approximation of the Jacobian being the identity matrix, the initial guess of the root (x_0, y_0), and coloring each point as described results in Figure 3.2.

Ideally, the regions colored red and blue in Figure 3.2 would be uniformly distributed around each root, however this is clearly not the case. The root $(-1, 1)$, indicated by the red diamond, has a nicely defined region around the root whose initial guesses all converge to the root. The root $(1, 1)$ does not have a nicely defined region. Looking closer at the range $(0, 2)$ appears there are four 'arms' of red that extend into the region near the root. These red regions protruding towards the blue root, although close to the root $(1, 1)$, converge to the other root. This is not a desired feature of the system. For an initial guess 'close' to a

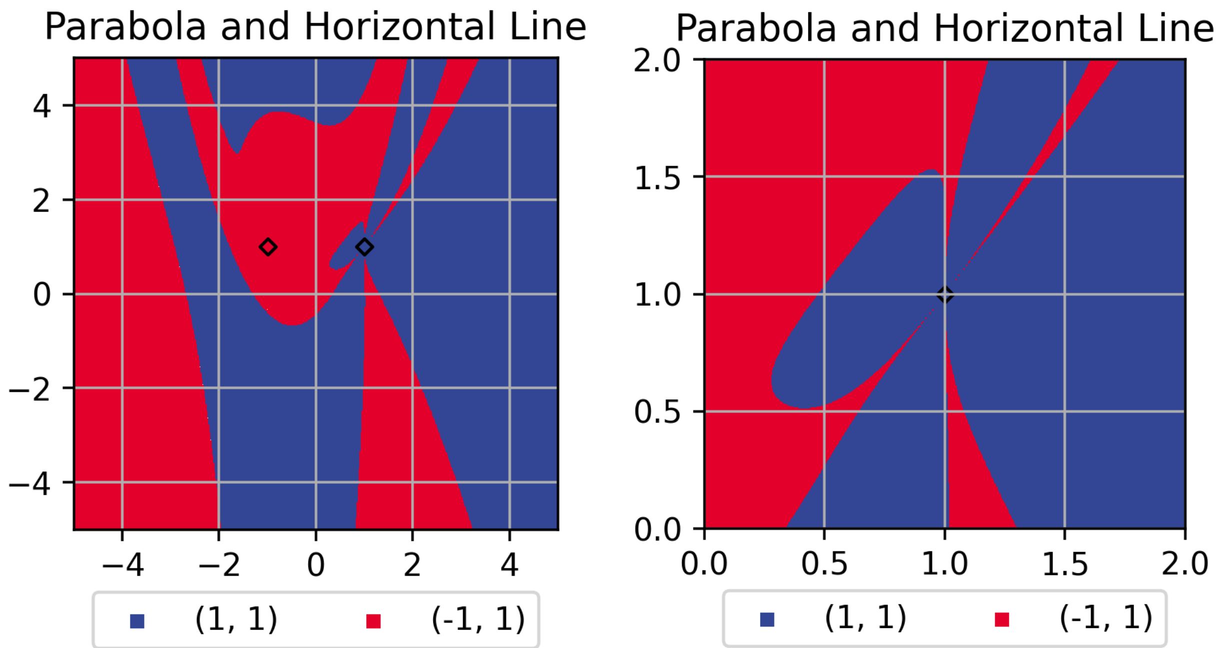

Figure 3.2: Parabola and horizontal line domain of attraction from equation 3.3.

root, ideally, the nearby root is found.

An additional parameter of interest is not just which root is found, but the rate at which the root is found. The number of iterations needed to achieve six decimal points accuracy of a root can be tracked by shading each initial guess based on how many iterations were needed. If we color the roots such that initial guesses converging in fewer than 7 iterations are darker and roots converging in greater than 13 iterations are colored lighter, the following plot can be produced (Figure 3.3).

Coloring the roots in this manner reveals additional structures of the graph. It would be expected that the boundaries between regions gradually shift to include the lighter colors, representing a weaker convergence of the method. In some regions, such as the lower left corner of the plot, this is seen. There is a gradual shift from initial guesses converging from one root to the other. However, there are also regions with sharp distinct boundaries separating the regions. Near the root $(1, 1)$, for example, the red 'arms' protruding from the root quickly transition from red to blue without a large region of the lighter color.

Throughout the remainder of this paper this coloring scheme will be included to represent

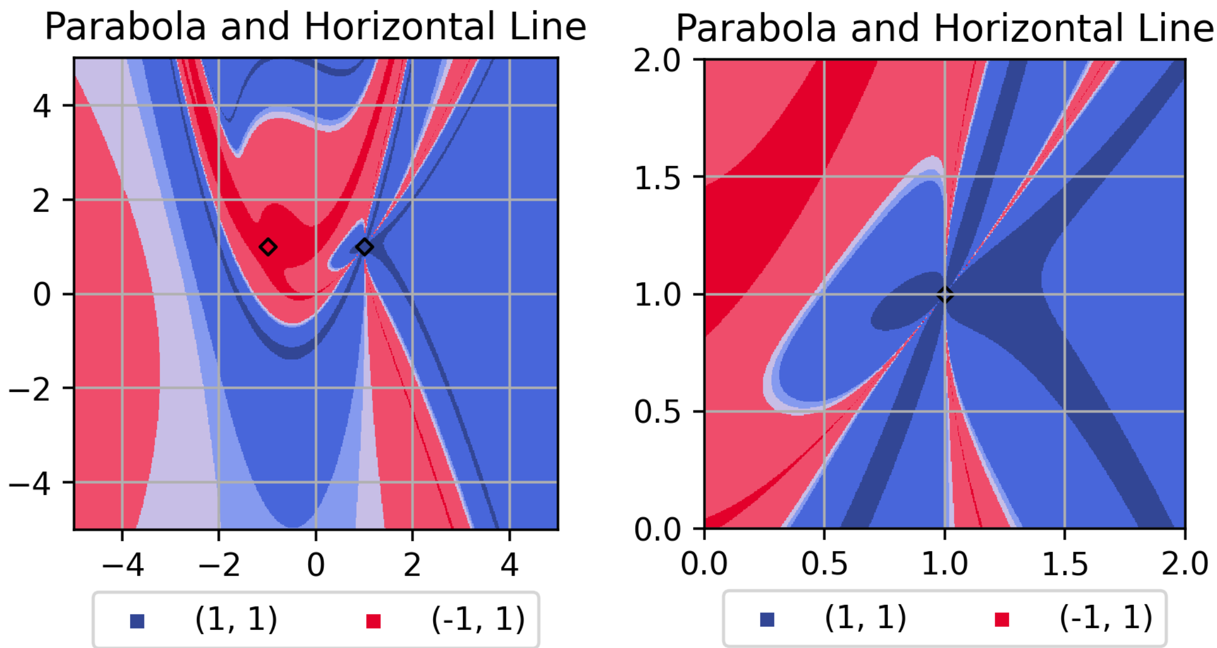

Figure 3.3: Parabola and horizontal line domain of attraction from equation 3.3, colored so initial guesses which converge quicker are darker colors. Initial values which converge in fewer than 7 iterations are colored darker and initial values which converge in more than 13 iterations are colored lighter.

the rate of convergence for each initial guess.

3.2 Systems with Multiple Roots

Previously, systems with only two roots were considered for convenience in introducing the concepts. However, the domain of attraction extends to systems with any number of roots. Consider the following system of equations with three roots:

$$\begin{aligned} f_1(x, y) &= x^2 + y^2 - 1 = 0 \\ f_1(x, y) &= y - x^2 + 1 = 0 \end{aligned} \tag{3.4}$$

The roots of equation 3.4 can be interpreted as the intersections between the unit circle and the parabola $y = x^2 - 1$. The roots of the function are $(1, 0)$, $(-1, 0)$, and $(0, -1)$. If we follow the same process discussed previously and color each initial guess based on which

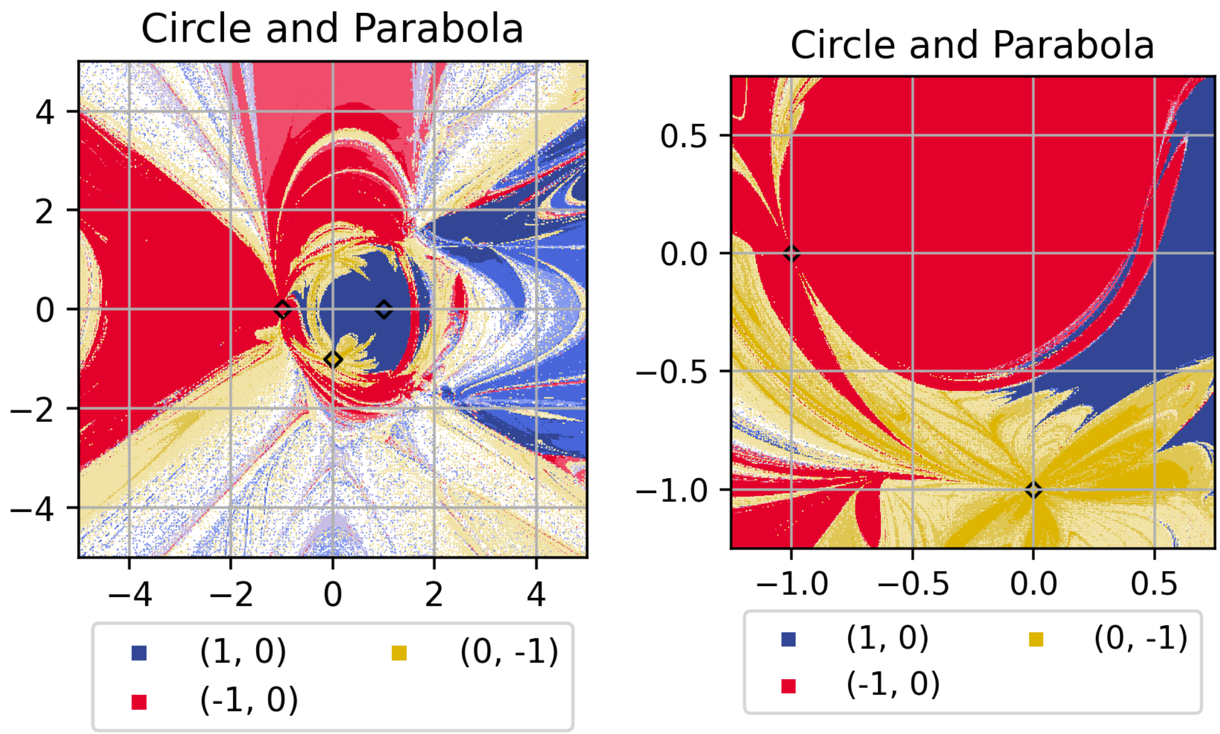

Figure 3.4: Finding roots of 3.4 and colored such that initial guesses converging in fewer than 20 iterations are darker, initial guesses converging in more than 30 iterations are lighter and initial guesses that did not converge after 100 iterations are white.

root is found and the number of iterations required to obtain six decimal points of accuracy, the following plot can be generated (Figure 3.4).

The equations from (3.4) show much more of the plot is colored white, representing initial guesses that did not converge to a root after 100 iterations. In addition to large regions which did not converge, the regions around roots are not colored uniformly. The root at $(1, 0)$ seems 'fairly' well-posed, as most surrounding initial values will converge to that root. The roots at $(-1, 0)$ are moderately well-posed, with a dark red region surrounding the root. Although there are yellow 'arms' that reach near the root, most initial guesses near the root appear to converge to the desired root. Lastly, the root at $(0, -1)$ is the most poorly conditioned. Although there is a region surrounding the root that converges, it is not colored darkly and represents initial guesses that take more iterations to converge.

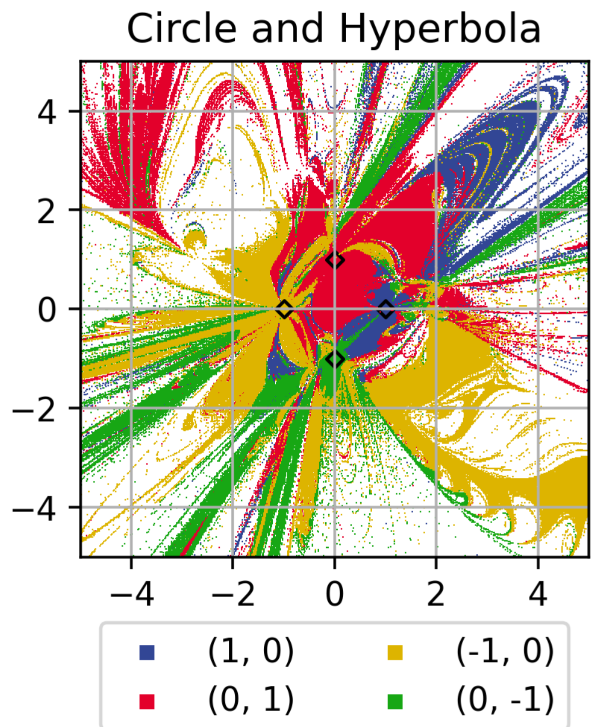

Figure 3.5: Finding roots of 3.5 and colored such that initial guesses converging in fewer than 20 iterations are darker, initial guesses converging in more than 30 iterations are lighter and initial guesses that did not converge after 100 iterations are white.

Similarly, a system with four roots can be considered.

$$\begin{aligned} f_1(x, y) &= x^2 + y^2 - 1 = 0 \\ f_2(x, y) &= x^2 + y^2 - 3xy - 1 = 0 \end{aligned} \tag{3.5}$$

For this system, the roots can be interpreted as the intersections between the unit circle and a hyperbola. This system has four roots at $(1, 0)$, $(0, 1)$, (-1.0), and $(0, -1)$. Once again, plotting the roots and coloring each initial guess based on which root is found and the number of iterations required to arrive at six decimal places of accuracy can be created (Figure 3.5 and Figure 3.6).

The same game can be played to evaluate the strength of each root by considering the region surrounding each root. Each root appears to be moderately well-posed. This is to say the region surrounding each root is colored the same as the root. Each root has ‘arms’

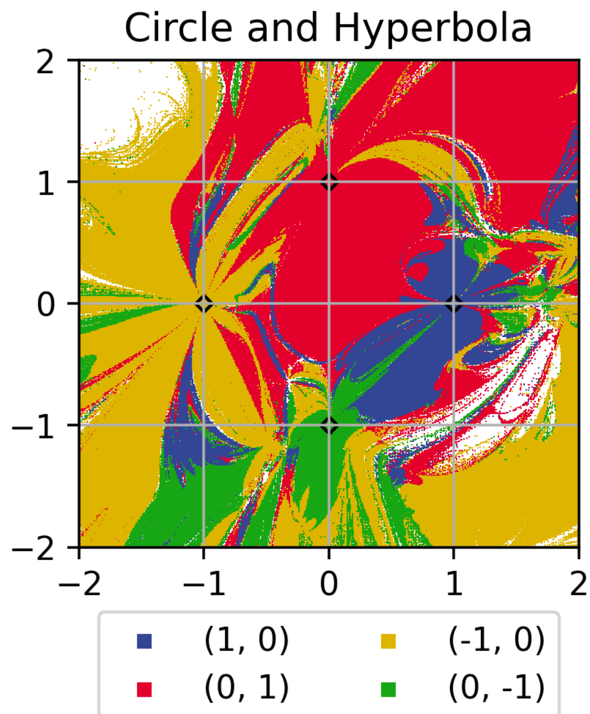

Figure 3.6: Closer look at Figure 3.5.

of other colors approaching the root. With an initial guess sufficiently close to a root, it is likely that the method will converge to the nearest root, but depending on the distance from the root, it is impossible to be certain.

It is important to note that for systems with multiple roots, the graphs are not always chaotic with poorly defined domains of attraction. Consider the intersection points between a circle and a hypocycloid with the following equations 3.6.

$$f_1(x, y) = x^2 + y^2 - 1 = 0$$
$$f_2(x, y) = x^{\frac{2}{3}} + y^{\frac{2}{3}} - 1 = 0 \tag{3.6}$$

This system similarly has four roots at $(1, 0)$, $(0, 1)$, (-1.0), and $(0, -1)$. However, considering the same method of coloring initial guesses based on the root found and the iterations required results in a much less chaotic image (Figure 3.7).

Examining each of the roots reveals an incredibly well-defined and predictable domain of

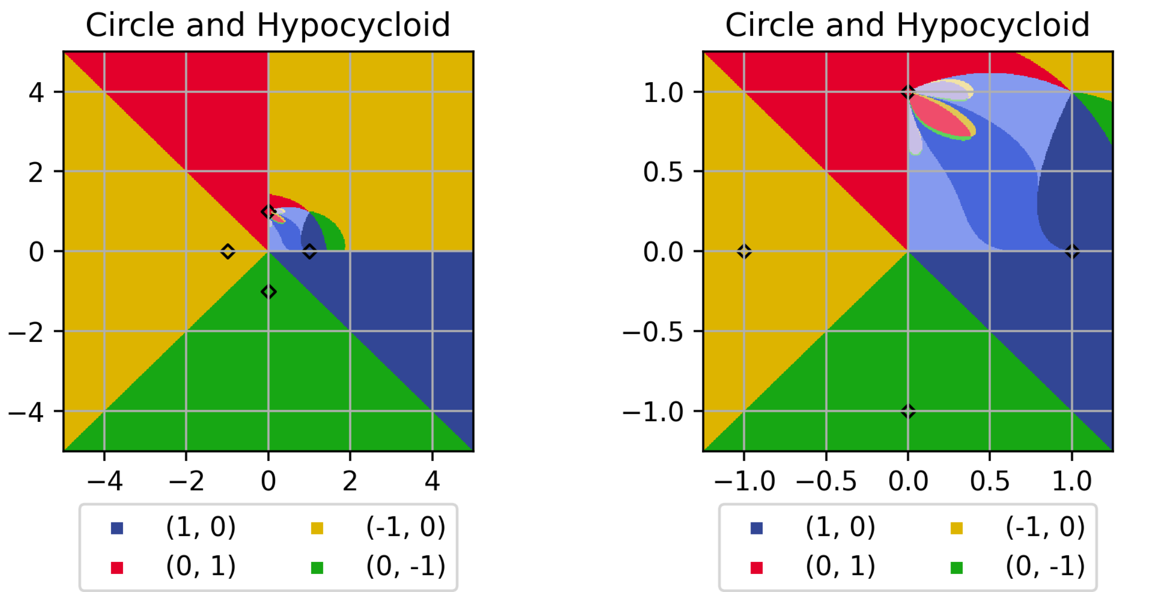

Figure 3.7: Finding roots of 3.5 and colored such that initial guesses converging in fewer than 2 iterations are darker, initial guesses converging in more than 3 iterations are lighter.

attraction. The graph is nicely partitioned into clearly defined regions. The root of $(0, 1)$ is the least well conditioned with 'petals' of the blue region approaching the root. Compared to the examples seen previously, the intersection of a unit circle with a unit hypocycloid with four petals produces a well-defined boundary for the domain of attraction for each root.

3.3 Change of Variable

Consider the intersection of a unit circle and a horizontal line. The system of equations can be represented by the functions 3.7.

$$\begin{aligned} f_1(x, y) &= x^2 + y^2 - 1 = 0 \\ f_2(x, y) &= y = 0 \end{aligned} \tag{3.7}$$

The same method of producing a plot based on the root the method converges to can be used. However, we can also consider the effect of the initial guess of the Jacobian for the method. Previous examples have used the 2×2 identity matrix as the initial guess for the Jacobian, but consider the permutation of the identity matrix 3.8.

$$\hat{I} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \tag{3.8}$$

For the system described by 3.7, we can apply Broyden's method with the initial guess of the Jacobian being the identity matrix (Figure 3.8) and with the permutation of the identity matrix (Figure 3.9).

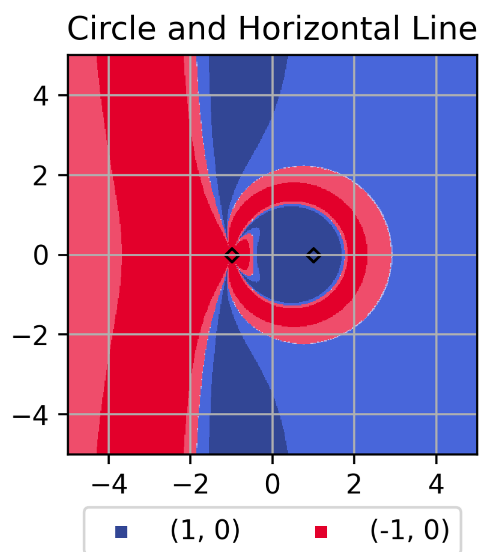

Figure 3.8: Equation 3.7 with $A_0 = I$.

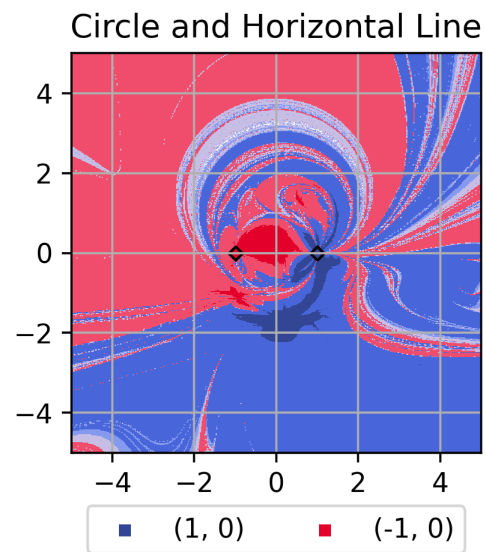

Figure 3.9: Equation 3.7 with $A_0 = \hat{I}$.

The two plots vary drastically. The plot which uses the identity matrix as the initial guess has well-defined and predictable boundaries between the domains of attraction. Using the permutation of the identity matrix results in a far more chaotic graph.

If, instead of considering a circle intersected by a horizontal line, a circle is intersected with a vertical line represented by the equations 3.9.

$$\begin{aligned} f_1(x, y) &= x^2 + y^2 - 1 = 0 \\ f_2(x, y) &= x = 0 \end{aligned} \tag{3.9}$$

The same pair of plots can be produced considering the initial guess of the Jacobian as the identity matrix (Figure 3.10) and the permutation of the identity matrix (Figure 3.11).

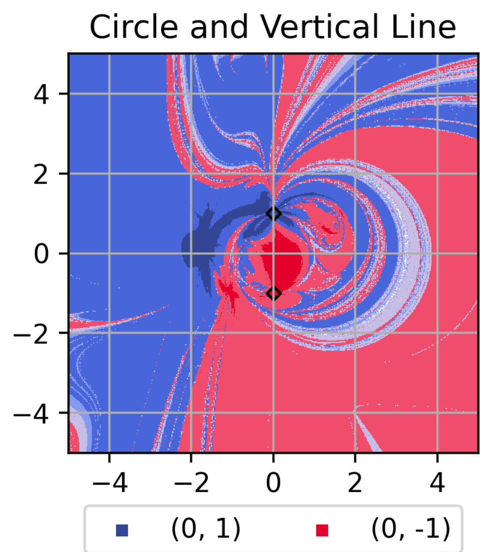

Figure 3.10: Equation 3.9 with $A_0 = I$.

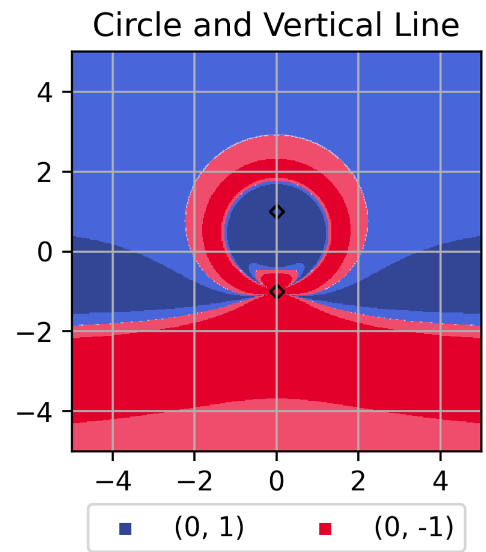

Figure 3.11: Equation 3.9 with $A_0 = \hat{I}$.

Comparing the four plots reveals the method is consistent with a change of variable. Figure 3.8 and Figure 3.11 are rotationally symmetric. If a change of variable was performed to 3.7 such that $x = y'$ and $y = x'$, then for an equivalent system, the initial guess of the Jacobian will change from $I = \hat{I}$. It is expected that the two plots before and after the change of basis should have rotational symmetry. This symmetry can be seen by comparing Figure 3.8 and Figure 3.11. Similarly, the method of a circle and horizontal line which used $\hat{I}$ as the initial guess of the Jacobian will have rotational symmetry to the circle and vertical line system that used the identity matrix as the initial guess.

The set of initial guesses that converge to a root can be described as the domain of attraction of the root. In addition to rate of convergence and computational complexity, quantifying the domain of attraction of a method can describe the strength of a method. Ideally, guesses sufficiently close to a root will converge to the root. Creating plots of 2×2 systems of equations and coloring initial guesses based on the root which is converged to and based on the rate of convergence gives insight to the behavior of the methods. This section showed systems of equations with well posed domains of attraction as well as systems of equations with chaotic domains of attraction. This section concluded with an example showing that changing the initial guess of the Jacobian from the identity matrix to the permutation of the identity matrix can greatly affect the domain of attraction. In the next section, discussion of how to quantifiably approximating the domain of attraction will be presented.

Chapter 4

Approximating the Radius of Convergence

This section will introduce two methods of representing the radius of convergence of a root. The radius of convergence is a region surrounding a root such that all values in the region converge to the root. There is no closed-form solution or explicit formula for defining the domain of attraction of an arbitrary system of equations. As shown in the examples, there are many cases in which describing the entire domain of attraction is practically impossible. This paper proposes a method of representing and approximating the radius of convergence to quantifiably compare the well-posedness of a system.

This section will consider the two dimensional case of a system with only two variables, although higher dimensional analogs of the methods exist. The first method presented will find the circle at the root with the largest radius such that all evaluated values on the circumference of the circle converge to the root at the center. The second method will use a Monte Carlo simulation to approximate a disk whose points converge to the root a specified percentage of the time. Both methods utilize sampling to approximate the entire region, but serve as a way to represent the behavior of values close to each root.

The motivation of both methods is to find the circle or disk centered at a root with the largest radius such that all the values within the disk converge to the root centered at the disk. The region represents an upper bound to the absolute error between a root and a guess needed to be confident the initial guess will converge to the desired root.

$$|\boldsymbol{x_0} - \boldsymbol{r}| \leq \delta \tag{4.1}$$

Where: $\boldsymbol{x_0}$ is the initial guess and $\boldsymbol{r}$ is the root of interest. The goal is to find the value of δ such that for values of $\boldsymbol{x_0}$ for which the inequality is true, will converge to the root r.

Of course, there may be values outside of the disk that will converge to the root due to the chaotic nature of the method. This does not prove that all initial guesses within the disk will converge to the root at the center, but creates a reasonable expectation that the initial guesses will converge to said root. In addition, a disk might not always be the most appropriate contour of describing the domain of attraction. As it has been shown in Figure 3.7, some domains of attraction could more precisely be described with inequalities representing broad regions that converge to the root. It is unlikely the domain of attraction will be perfectly described by a disk and it is likely impossible to show all values within a disk will converge to the specified root. However, this method will provide a starting point to compare the different domain of attraction plots.

4.1 Circle Bisection Method

This section will discuss one method of approximating the domain of attraction of a root using Broyden's method evaluating two circles centered at the root. The method requires five inputs: the root to search around, an initial lower bound radius (δ_{low}), an initial upper bound radius (δ_{high}), the linear density that points should be evaluated along the circumference of the circle, (λ), and some termination criteria. To begin, two circles are required, both centered at the root. One circle, radius denoted δ_{low}, is to be defined such that all values evaluated on the circumference converge to the root at the center. The other circle, radius denoted δ_{high}, is to be defined such that there are values along the circumference of the root that do not converge to the root 4.1.

Each iteration will consider a circle with the average of the two radii.

$$\delta_{new} = \frac{\delta_{low} + \delta_{high}}{2} \tag{4.2}$$

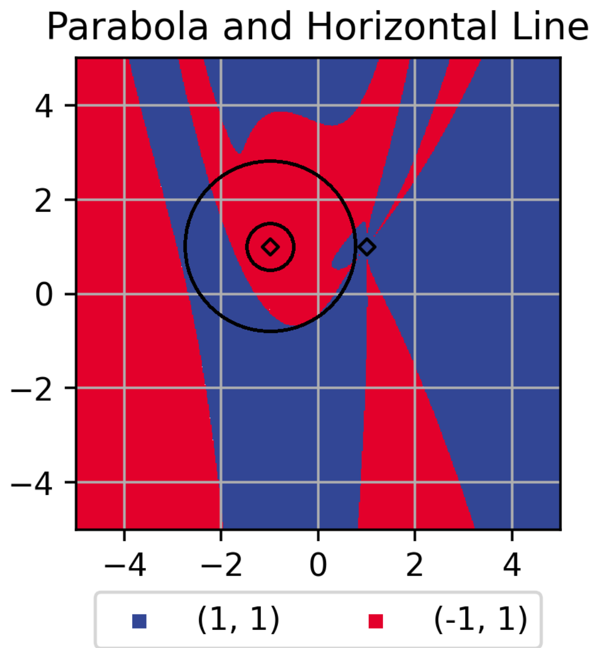

Figure 4.1: Circle bisection method where the larger circle represents the initial guess for a radius whose values on the circumference do not all converge to the root at the center and where the smaller circle represents the initial guess for a radius where all values on the circumference converge to the root at the center

If all the points evaluated on the circle's circumference converge, then δ_{new} replaces δ_{low} and the process repeats. If any points evaluated on the circle's circumference do not converge, then δ_{new} replaces δ_{high} and the process is repeated. Once the difference in δ_{low} and δ_{high} is less than the termination criteria, the method is stopped and the circle with the largest radius that is found whose values evaluated along the circumference all converge to the specified root will be used as an approximation of the domain of attraction.

The number of points evaluated along the circumference is calculated with a user defined value of linear density, λ. The linear density parameter λ has units of points per linear distance. A circle with radius δ will be evaluated according to the equation 4.3.

$$\text{num} = \text{ceil}(2\pi\delta\lambda) \tag{4.3}$$

This way, regardless of the circle radii being evaluated by the method, the same density of

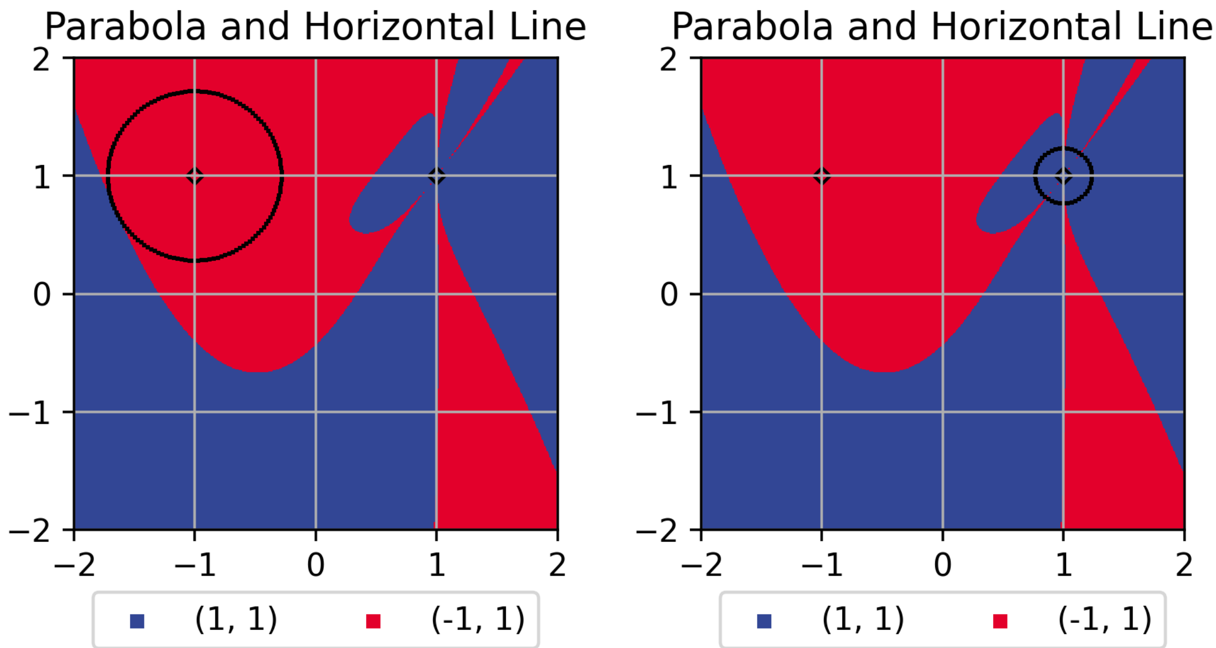

Figure 4.2: Applying the method of bisecting circles to 3.3. For the root $(-1, 1)$ in red, $\delta = 0.71$ while for the root $(1, 1)$ in blue $\delta = 0.23$. For both methods $\lambda = 30$ and δ is found to two decimal points of accuracy

points are evaluated.

Consider the example of the parabola and horizontal line governed by 3.3. For each root, use the method of bisecting circles to approximate the domain of attraction.

The values of δ for each root match what is seen on the images (Figure 4.2). The red root representing (-1.1) had a δ value calculated to be 0.71 while the blue root representing $(1, 1)$ had a δ value of only 0.23. From the graph alone, it was easy to see that the red root had a stronger domain of attraction, however, the method of bisecting circles was able to introduce a parameter to measure the domain.

To improve the method either the density of points evaluated λ can be increased to a higher density. However, there is a trade-off in precision with computational runtime. Any method which evaluates more points will take longer to complete.

4.2 Monte Carlo Disk

This section will describe how a Monte Carlo simulation can be used to approximate the domain of attraction. Imagine a system where it is important that 99% of trials the method will converge to the desired root as long as the initial guess is within some distance to the true root. This method will find the radius δ such that a specified percentage, α, of points on the disk converge to the desired root. Rather than only considering points on the circumference of a circle, this method will randomly select points on the disk to evaluate.

The method begins with two disks. The smaller disk of radius δ_{low} is to underestimate the region. More than α of the points in the disk of radius δ_{low} are to converge to the root at the center of the disk. The larger disk of radius δ_{high} is to overestimate the region. Less than α of the points in the disk are to converge to the root. In each iteration, equation 4.2 is used to determine the average of the two radii. Since it is unreasonable to evaluate all the points in the disk, a random selection process is used to sample a small number of them. The quantity of points evaluated on the disk is determined by an area density parameter, ρ_A, with the unit points per unit area. The number of points to be selected can be calculated by

$$\text{num} = \text{ceil}(\pi\delta^2\rho_A) \tag{4.4}$$

To determine the coordinates for each point evaluated on the disk of radius δ for the root centered at (x_0, y_0) in two dimensions, the following expression is used:

$$\begin{aligned} x &= x_0 + \delta\omega_1 cos(2\pi\omega_2) \\ y &= y_0 + \delta\omega_3 sin(2\pi\omega_4) \end{aligned} \tag{4.5}$$

Where ω_n is a randomly generated variable between 0 and 1.

The points generated by equation 4.5 are used as initial guesses in Broyden's method. The total percentage of points evaluated that converge to the desired root is calculated and compared against α. If the percentage of points that converge to the root is greater than

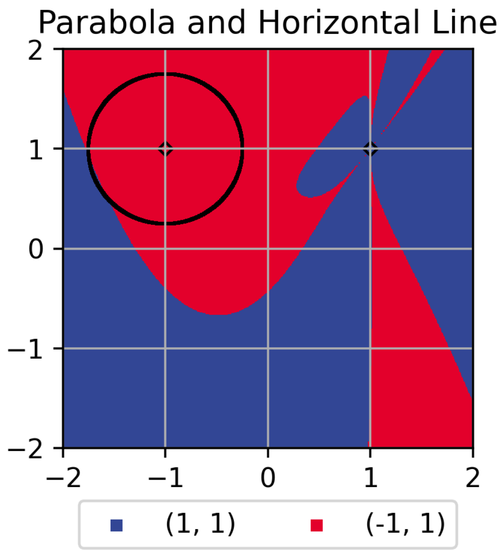

Figure 4.3: $\delta = 0.76$, $\alpha = 99\%$

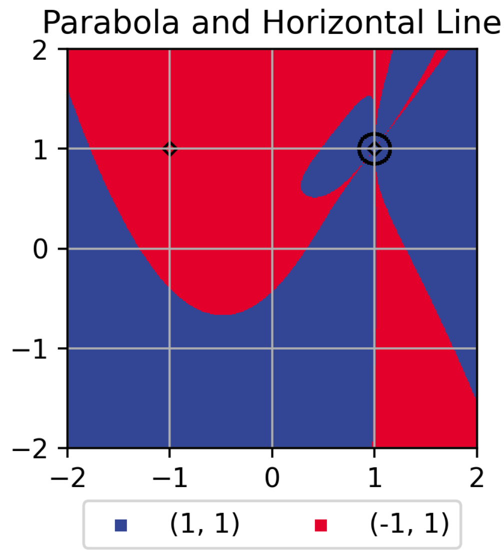

Figure 4.4: $\delta = 0.15$, $\alpha = 99\%$.

α, the lower estimate of the radius is replaced by δ_1. If fewer initial guesses converge to the desired root, the upper estimate of the radius is replaced by δ_1. The process repeats until a specified termination parameter is reached.

Consider the example from before of the parabola and the horizontal line. Two values of α can be compared representing 99% and 95% of the values converging to the desired root. For each trial the area density ρ_A was set at 500 points per unit area. Figure 4.7 shows the result of each method.

As expected, the region produced by this method admits some amount of initial guesses that will converge to the opposite root. For a tighter tolerance, a higher α value can be specified at the cost of a smaller region. Both values of α showed that the root at $(-1, 1)$ was better conditioned than the root at $(1, 1)$. Although the values of δ did not match that of the bisecting circle methods, that is not to be expected since the two methods are searching for different criteria. It is reassuring to note that for an α value of 99%, the values for δ were approximately the same.

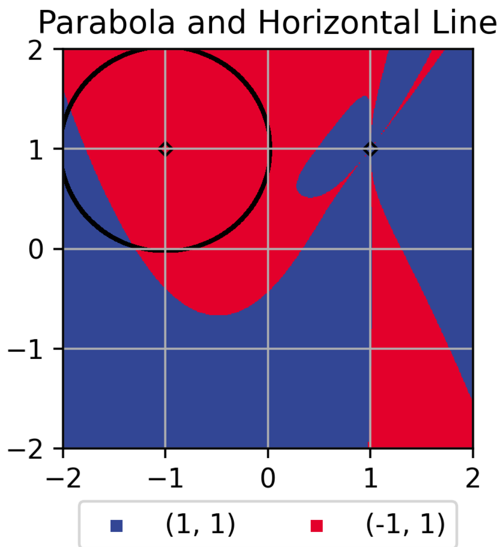

Figure 4.5: $\delta = 1.03$, $\alpha = 95\%$.

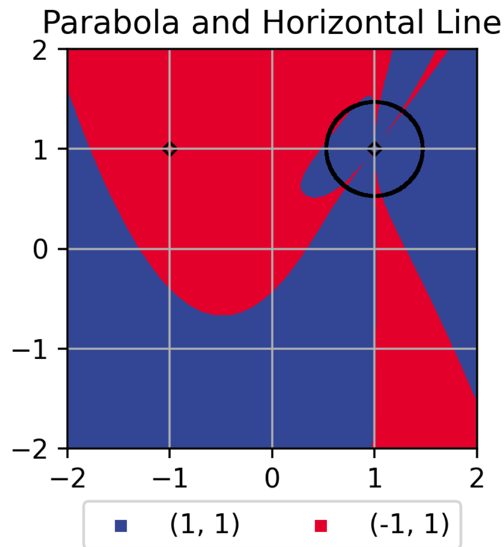

Figure 4.6: $\delta = 0.48$, $\alpha = 95\%$

Figure 4.7: Applying Monte Carlo Disk method to 3.3 with $\rho_A = 500$

4.3 Discussion

Both methods presented in this section can be used to represent the domain of attraction of a root. Since the domain of attraction is often a complicated boundary, for most cases it is unreasonable to try to develop a closed-form representation of the region. These methods seek to represent the domain of attraction with the largest circle or disk where the points within the circle or on the disk will converge to the desired root to some level of accuracy. Of course, neither method guarantees that all points within the region will converge to the correct root. Using the bisecting circle method, it is possible for points to 'slip through the cracks' on the border of points considered. If the linear density of points evaluated is not high enough, the method could miss points that will converge to the other root. With the Monte Carlo method, the specification of α determines the ratio of points that will converge to the desired root. The two methods can be used in conjunction with one another to validate the results.

The remainder of this paper will primarily consider the circle bisection method since there is not a need to develop a random list of points in the method. Therefore, the results for the circle bisection method are more repeatable and easily verified. However, it is easy to consider a case with the circle bisection method where a region of points that do not converge to the proper root are 'skipped over.' To check for such regions, after applying the circle bisection method to a system, a Monte Carlo simulation can be used to approximate the percentage of roots that converge to the desired root. A Monte Carlo simulation after approximating the radius of convergence supports the strength of the radius found with the bisecting circle method. Of course, this still does not prove that all values within the region will converge, but supports a percentage of expected values to converge.

Other methods could be used to approximate the domain of attraction. The number of 'pixels' that are connected to a root which converge to the root could be measured, approximating the area of convergence. However, it is likely this area could have long 'arms' that would not properly demonstrate the true nature of the domain of attraction. Other methods could include the largest shapes trying to fill a region. To extend this method to three dimensions, the domain of attraction will be represented by the largest sphere whose internal values converge to the root of interest.

This section proposed two methods of numerically approximating the domain of attraction for Broyden's method. The methods can be used in tandem to support an approximation of the radius of convergence for a root. Both methods serve to quantify how much absolute error an initial guess can have from a root while still being confident the initial guess will converge to the desired root. In the next section, this method of approximating a domain of attraction numerically will be used to demonstrate the importance of a 'good' initial guess of the Jacobian.

Chapter 5

Effect of Initial Guess of the Jacobian on the Radius of Convergence

This section will demonstrate the importance of the initial guess of the Jacobian for Broyden's method. Previous sections have graphically shown how a method converges to different roots based on the initial guess of the root, but only minor consideration has gone into the selection of the Jacobian. As previously stated, work on this method has produced varying claims about the initial guess of the Jacobian. In this section it will be shown both graphically and quantitatively that the radius of convergence of a root improves with an approximation of the Jacobian used in place of the identity matrix to initialize the method. Four methods of varying computational complexity will be compared:

1. Newton's method for Solving Non-Linear Systems of Equations
2. Broyden's method with $A_0 = J(x_0)$
3. Broyden's method with $A_0 \approx J(x_0)$
4. Broyden's method with $A_0 = I$

For the first method, the Jacobian must be explicitly calculated and evaluated with each iteration. For the last method, no knowledge of the Jacobian is required. The trade-off between needing to explicitly calculate the Jacobian and the computational improvement will be seen. Ultimately this section will propose that the initial Jacobian should be approximated using a finite difference method.

5.1 Newton's Method Domain of Attraction

This section will include examples of the domain of attraction for Newton's method for solving systems of non-linear equations. For the remainder of this paper, Newton's method will be used as the "control" group while evaluating numerical methods. It has been shown that Newton's method is "q-quadratically" convergent under standard assumptions and for initial guesses sufficiently close to the root in question [4]. However, as has been discussed in this paper, Newton's method comes with the trade-off of needing to explicitly calculate the Jacobian of a system. The use of Newton's method as a baseline comparison will demonstrate a best-case scenario for a method. If the domain of attraction of Newton's method is ill conditioned, it is unreasonable to expect the quasi-newton's method to be well conditioned.

Consider the example of the parabola and horizontal line from equation 3.3.

$$\begin{aligned} f_1(x,y) &= y - x^2 = 0 \\ f_2(x,y) &= y - 1 = 0 \end{aligned} \tag{5.1}$$

Whose roots are $(1,1)$ and $(-1,1)$. The Jacobian of this system can easily be explicitly calculated to be:

$$J(x,y) = \begin{bmatrix} 2x & 1 \\ 0 & 1 \end{bmatrix} \tag{5.2}$$

Using the same method as before, we can track which root each initial guess converged to. If we color all initial guesses that converge to the root $(1,1)$ as blue and all the initial guesses that converge to $(-1,1)$ as red, the result is Figure 5.1.

Compared to the domain of attraction seen for Broyden's method (Figure 3.2), with Newton's method the domain of attraction is clearly defined.

Consider instead a system with four roots, such as the intersection of the circle and

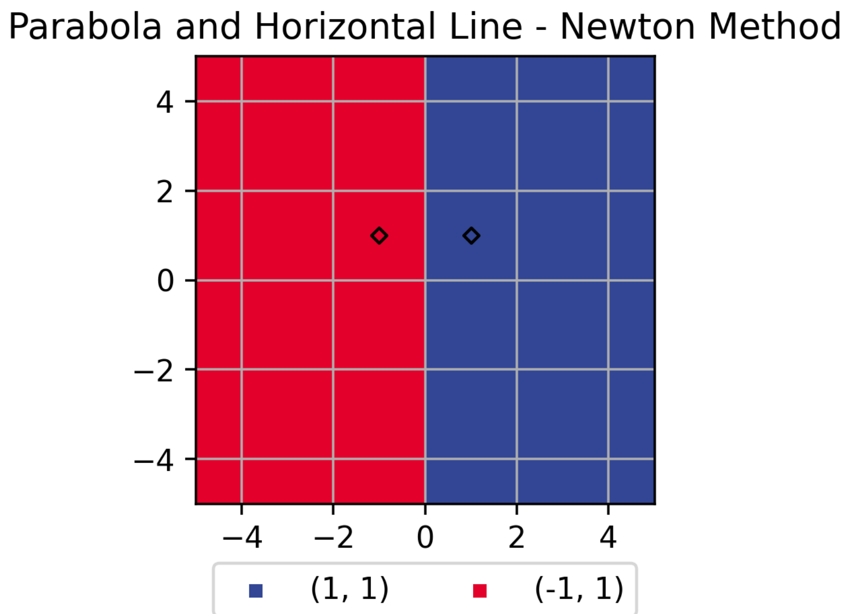

Figure 5.1: Newton's method applied to equation 3.3. The domain of attraction for each root is clearly defined based on the sign of the x value of the initial guess

hyperbola given by equation 3.5.

$$\begin{aligned} f_1(x,y) &= x^2 + y^2 - 1 = 0 \\ f_2(x,y) &= x^2 + y^2 - 3xy - 1 = 0 \end{aligned} \tag{5.3}$$

whose roots are $(1,0), (0,1), (-1,0)$ and $(0,-1)$. The Jacobian of this system can be calculated to be:

$$J(x,y) = \begin{bmatrix} 2x & 2y \\ 2x - 3y & 2y - 3x \end{bmatrix} \tag{5.4}$$

The plot which tracks which root each initial guess converges to is given by Figure 5.2. Once again, the improvement of the domain of attraction by applying Newton's method is drastic. Compared to that which was seen in Figure 3.5, the domain of attraction can be clearly defined. The method converges to the closest root.

Compared to Broyden's method, so far the domains of attraction for Newton's method have been much more well conditioned. So far for these polynomial equations, the domain

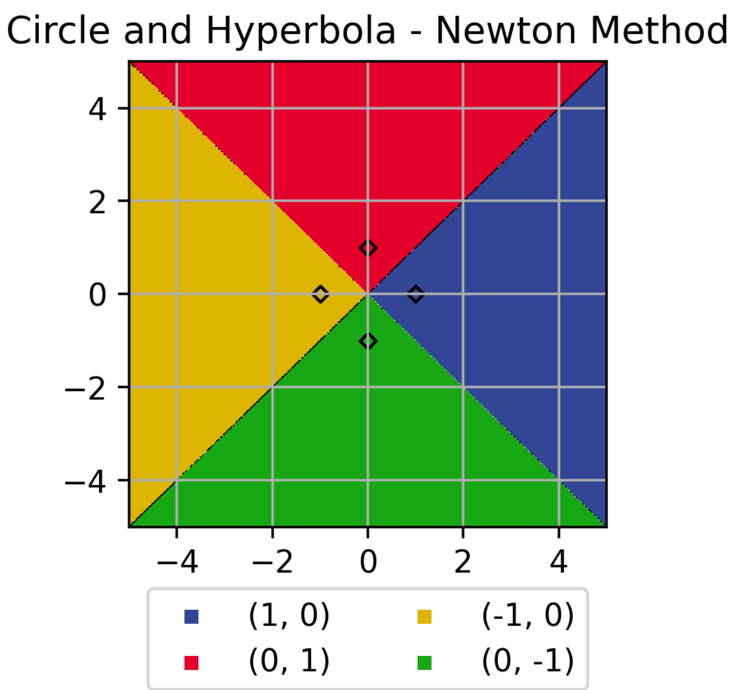

Figure 5.2: Newton's method applied to equation 3.5. The domain of attraction for each root is clearly defined based on which root is closest.

of attraction for each root could not be improved significantly. For this reason, Newton's method will serve as the standard to which other methods will be compared against. The following sections will include examples of systems of equations with increasing degrees of complexity. For each example, the domain of attraction will be shown graphically and will be numerically approximated to compare the effect each method has on the convergence behavior of the method.

5.2 Testing Method

This section compares four iterative methods of solving systems of non-linear equations. In addition to Newton's method and Broyden's method using the identity matrix as the initial guess of the Jacobian, this section will also consider Broyden's method with different choices for the Jacobian initial guess: an explicit calculation of the Jacobian for the initial guess x_0 and an approximation of the Jacobian using central difference theory.

Consider Broyden's method where the Jacobian was explicitly calculated for only the first guess, x_0. Rather than evaluating the Jacobian for each iteration, the Jacobian is calculated for the initial guess only. All subsequent iterations used the approximation of the Jacobian derived from Broyden's method. An explicit evaluation of the Jacobian at the initial guess represents the best initial guess of a Jacobian that a value can have. Broyden's method, which uses the explicitly calculated value of the Jacobian as the initial guess, provided an upper bound of improvement that can be made to the system by choosing better guesses for the Jacobian. In contrast to the method which uses the identity matrix as the initial guess for the Jacobian, this method requires more initial computation, but produces a better initial guess. This method will be referred to as explicitly calculating the Jacobian, or $A_0 = J(x_0)$.

In the middle between the arbitrary selection of the identity matrix for the initial guess of the Jacobian and explicitly calculating the Jacobian for the initial guess, exists approximating the Jacobian using a central difference approximation. For an $n \times n$ matrix, $2n^2$ evaluations of functions are needed to approximate the Jacobian by:

$$J(x_1, x_2, ..., x_n) \approx \begin{bmatrix} \frac{f_1(x_1+h,x_2,...x_n)-f_1(x_1-h,x_2,...x_n)}{2h} & \cdots & \frac{f_1(x_2,x_2,...x_n+h)-f_1(x_2,x_2,...x_n-h)}{2h} \\ \frac{f_2(x_1+h,x_2,...x_n)-f_2(x_1-h,x_2,...x_n)}{2h} & \cdots & \frac{f_2(x_2,x_2,...x_n+h)-f_2(x_2,x_2,...x_n-h)}{2h} \\ \vdots & \vdots & \vdots \\ \frac{f_n(x_1+h,x_2,...x_n)-f_n(x_1-h,x_2,...x_n)}{2h} & \cdots & \frac{f_n(x_2,x_2,...x_n+h)-f_n(x_2,x_2,...x_n-h)}{2h} \end{bmatrix} \tag{5.5}$$

Approximating the Jacobian with a central difference formulation will serve as an intermediate initial guess for the Jacobian. For Jacobians with low degree terms (linear or quadratic), choosing small values for h will approximate the Jacobian effectively. Higher degree terms or trigonometric or transcendental functions will be less well approximated with the central difference methods, but will still have insight into the behavior of the Jacobian over the method which only uses the identity matrix as the initial guess. This paper chose the standard value of $h = 0.1$ to approximate the Jacobian. This method will be referred to as approximating the Jacobian, or $A_0 \approx J(x_0)$.

5.3 Polynomial Examples

This section will compare each of the four methods to systems of non-linear equations, limiting the examples to polynomials.

5.3.1 Circle and Horizontal Line

Consider the system of equations from 3.7.

$$\begin{aligned} f_1(x, y) &= x^2 + y^2 - 1 = 0 \\ f_2(x, y) &= y = 0 \end{aligned} \tag{5.6}$$

Each of the methods described at the beginning of this chapter will be applied to the system to both solve for both roots and compare the behavior of the domain of attraction for each method. Considering the root at $(1, 0)$ and plotting the four methods reveals a significant improvement in the domain of attraction.

With the introduction of an informed initial guess of the Jacobian, the plots for the domain of attraction greatly improved. The radius of convergence increased from $\delta = 0.772$ to $\delta = 0.998$, representing an improvement of 29%, but the more impressive improvement came from considering the plots. The same domain of attraction for Newton's method (Figure 5.6) is now seen in both Broyden's methods that use an approximation of the Jacobian to initialize the method (Figures 5.4, 5.5).

It should not come as a surprise that the method which approximated the Jacboian for the initial guess of A_0 performed exactly as well as the method which used the explicit Jacobian for the initial guess of A_0. The Jacobian of this system is:

$$\begin{bmatrix} 2x & 2y \\ 0 & 1 \end{bmatrix} \tag{5.7}$$

A central difference approximation solves exactly for the linear components of the Jacobian.

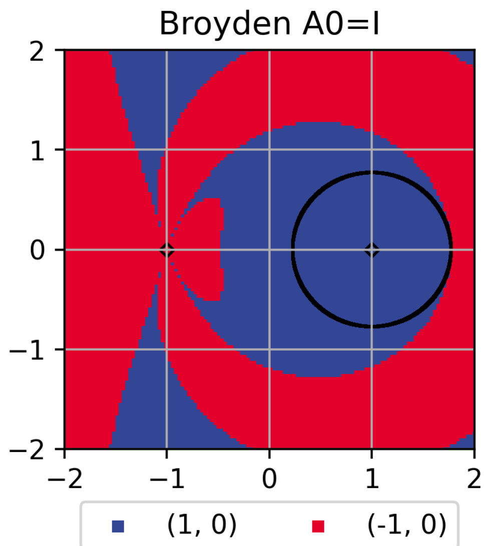

Figure 5.3: $\delta = 0.772$

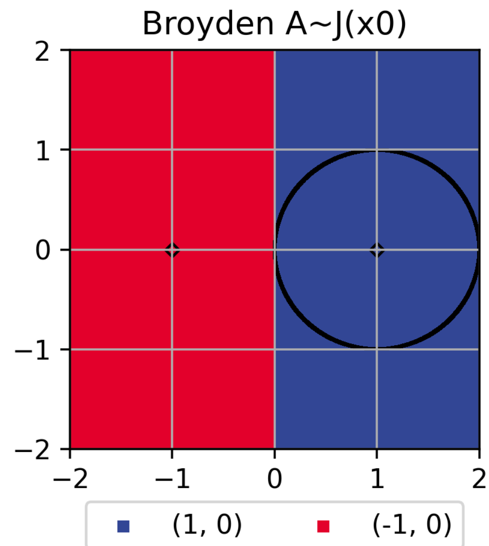

Figure 5.4: $\delta = 0.998$

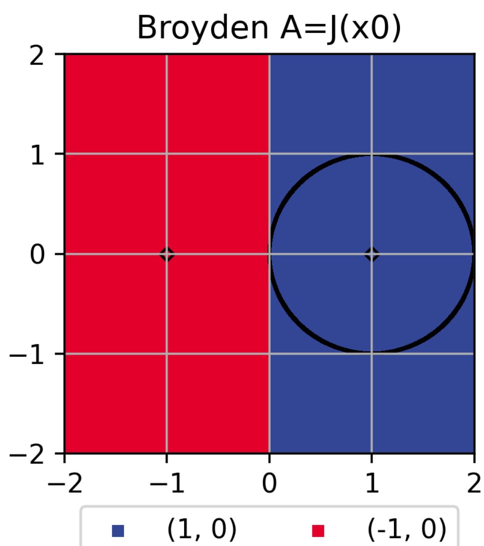

Figure 5.5: $\delta = 0.998$

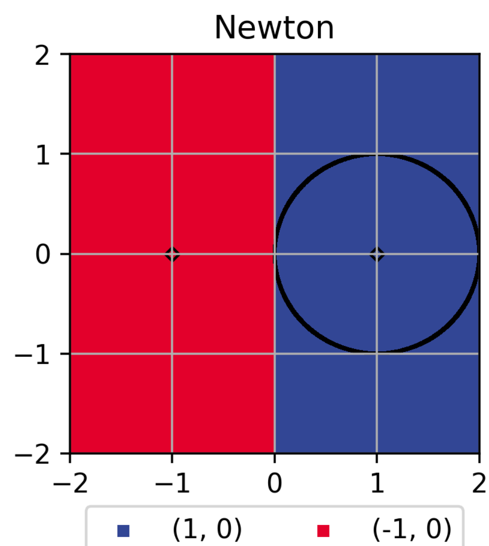

Figure 5.6: $\delta = 0.998$

Figure 5.7: Applying the four numerical methods to equation 3.7 with $\lambda = 1000$ to the root $(1, 0)$.

It will only be when the Jacobian contains higher degree terms that the central difference method will produce a worse approximation of the Jacobian.

Considering the other root of the system, the same plots can be produced to approximate the radius of convergence.

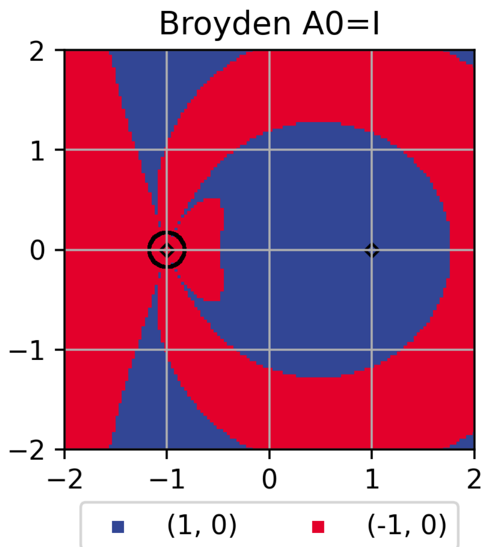

Figure 5.8: $\delta = 0.172$

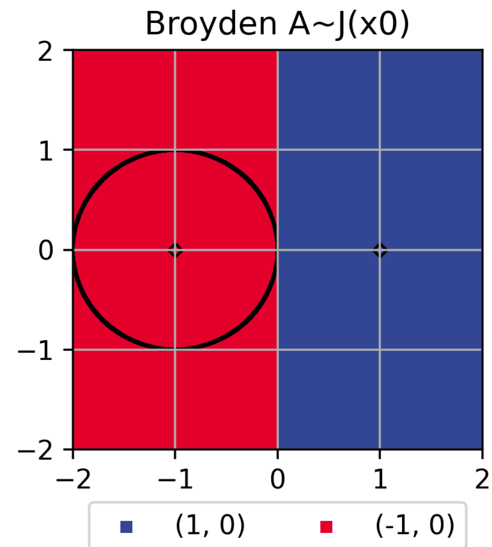

Figure 5.9: $\delta = 0.998$

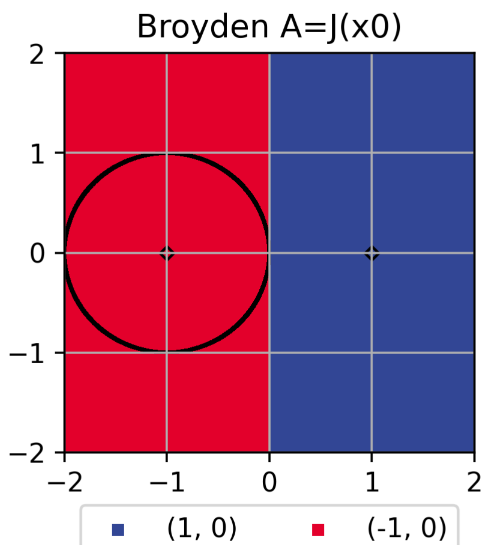

Figure 5.10: $\delta = 0.998$

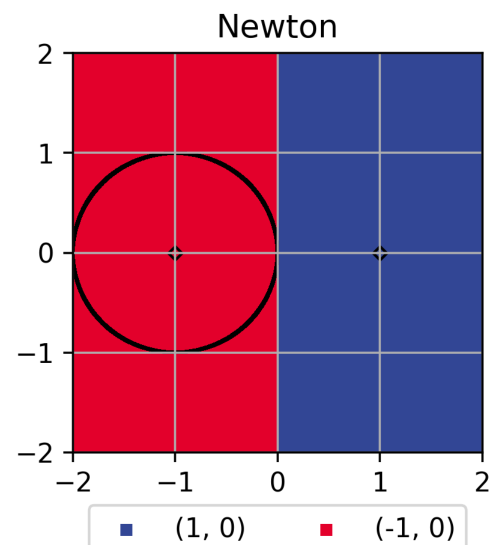

Figure 5.11: $\delta = 0.998$

Figure 5.12: Applying the four numerical methods to equation 3.7 with $\lambda = 1000$ to the root $(-1, 0)$.

A 480% improvement was made to the root at $(-1, 0)$ by using an informed guess for the

Jacobian. When using the identity as the initial guess, the radius of convergence is $\delta = 0.172$, however with an informed first guess, the radius increased to $\delta = 0.998$. In addition, the informed guess revealed a new line of symmetry of the system. Both roots have an equal area domain of attraction, symmetric across the y-axis.

5.3.2 Circle and Vertical Line

Considering the system of the circle and vertical line from equation 3.8, A chaotic behavior for this system was seen when the identity matrix was used for the initial guess of the Jacobian (Figure 3.10). The chaotic behavior can be removed with an informed selection for the initial guess of the Jacobian.

For both roots, significant improvements were seen with an approximation of the Jacobian. For the positive root, the radius of convergence increased from 0.102 to 0.998. For the negative root, the radius of convergence increased from 0.079 to 0.998. The improvement to the system is apparent when considering the plots of the domains of attraction. What began as a chaotic system was quickly improved with an addition of 8 evaluations of the function to approximate the Jacobian with the central difference theory. Since the Jacobian of the system only included linear terms, the central difference approximation exactly solved for the Jacobian. Explicitly calculating the Jacobian, or applying Newton's method in place of Broyden's method, would not provide any further improvement to the domain of attraction.

5.3.3 Circle and Cubic System

Consider a system of polynomials of a higher degree:

$$\begin{aligned} f_1(x, y) &= x^3 + y^3 - 2 = 0 \\ f_2(x, y) &= x^2 + y^2 - 2 = 0 \end{aligned} \tag{5.8}$$

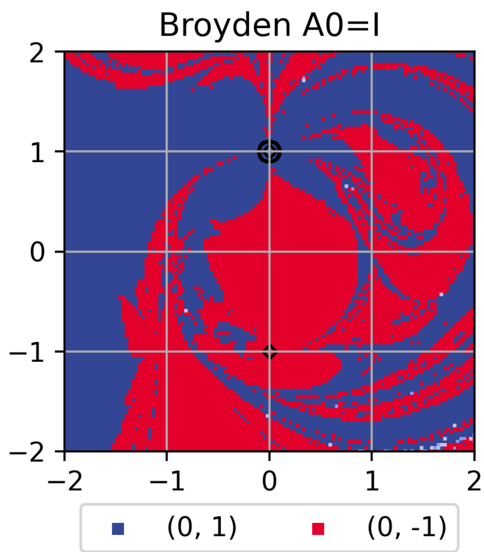

Figure 5.13: $\delta = 0.102$

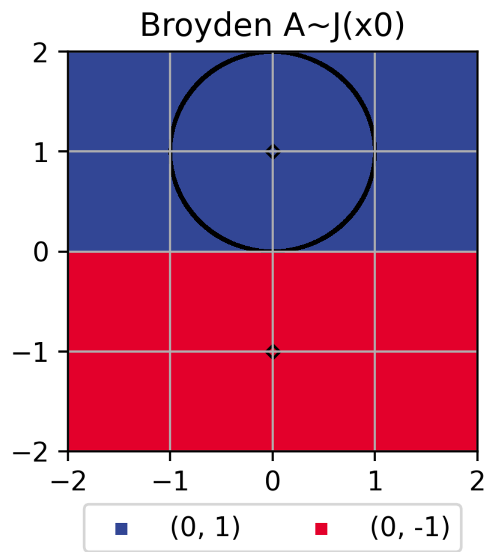

Figure 5.14: $\delta = 0.998$

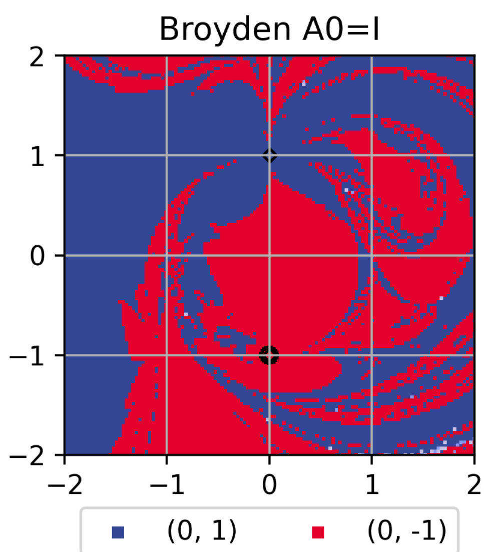

Figure 5.15: $\delta = 0.079$

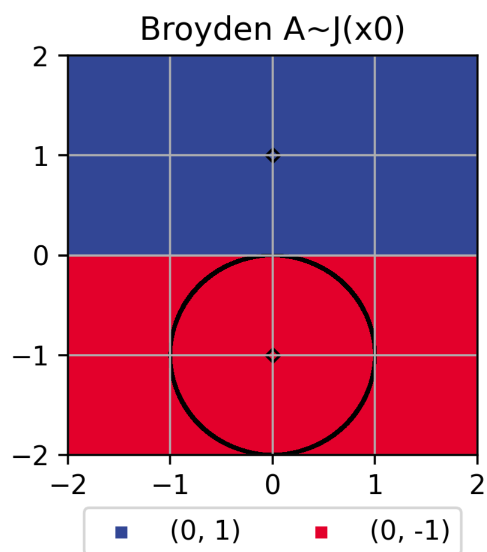

Figure 5.16: $\delta = 0.998$

Figure 5.17: Applying the four numerical methods to equation 3.8 with $\lambda = 1000$.

whose roots are $(1, 1)$,$(1.29, -0.56)$ and $(-0.56, 1.29)$. In this case the Jacobian is calculated to be:

$$J(x, y) = \begin{bmatrix} 3x^2 & 3y^2 \\ 2x & 2y \end{bmatrix} \tag{5.9}$$

Consider the effect of each method on the radius of convergence for the root $(1, 1)$. In

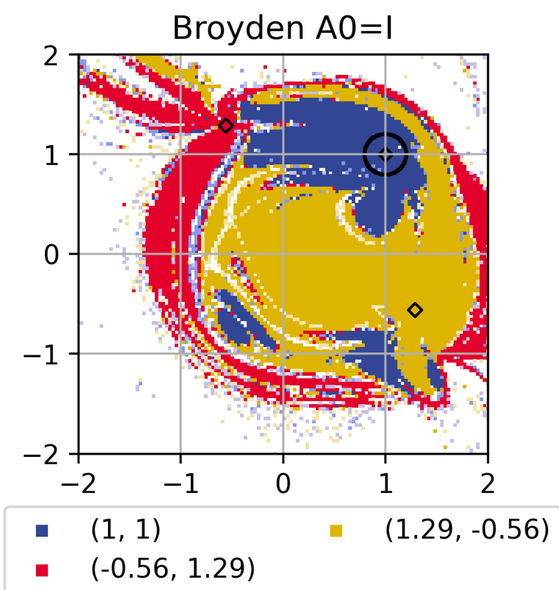

Figure 5.18: $\delta = 0.203$

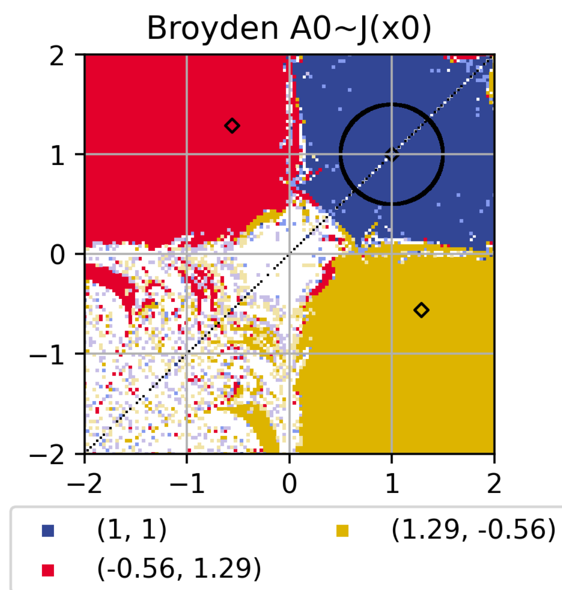

Figure 5.19: $\delta = 0.499$ and 99.97% converge tested via Monte Carlo Simulation

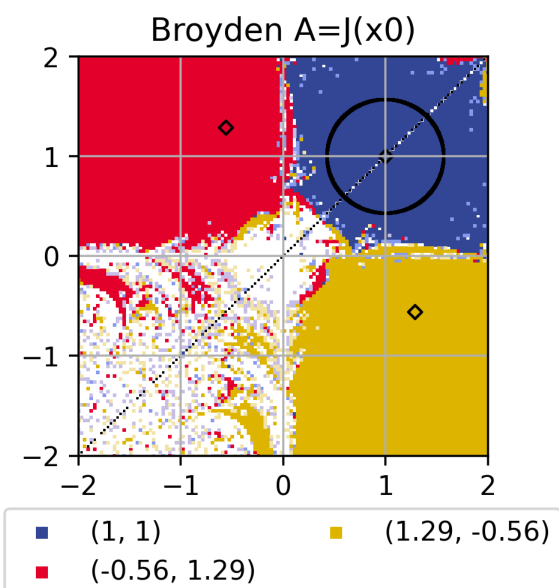

Figure 5.20: $\delta = 0.569$ and 99.88% converge tested via Monte Carlo Simulation

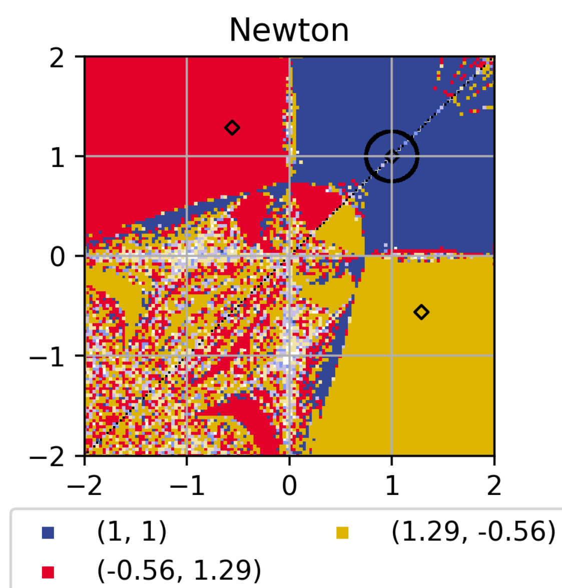

Figure 5.21: $\delta = 0.250$ and 99.94% converge tested via Monte Carlo Simulation

Figure 5.22: Applying the four numerical methods to equation 5.8 with $\lambda = 1000$.

this case, Newton's method does not result in the largest radius of convergence. Rather, the largest domain of attraction is seen in Broyden's method with either an approximation or an explicitly calculated initial guess of the Jacobian. Although Newton's method still outperformed Broyden's method with an initial guess of the identity matrix for the Jacobian, its radius of convergence was only 23% larger despite the additional computational complexity required for Newton's method. In the cases of Broyden's method with either educated initial guess of the Jacobian, the radius of convergence was 0.569, a 127% improvement over Newton's method. This result is interesting since with an educated initial guess in Broyden's method, the method is able to outperform Newton's method in this specific example.

5.4 Trigonometric Functions

This section will consider systems of equations which incorporate trigonometric functions. With fewer linear functions, we expect the system to not behave as nicely. Trigonometric functions are common in physics and engineering systems, and cannot be ignored due to computational complexity.

5.4.1 Cosine and Parabola

Consider the system of equations:

$$\begin{aligned} f_1(x,y) &= y - \cos(\pi x) \\ f_2(x,y) &= y - x^2 + \tfrac{1}{4} \end{aligned} \tag{5.10}$$

Whose roots are $(0.5, 0)$ and $(-0.5, 0)$. The Jacobian of the system is found to be:

$$J(x,y) = \begin{bmatrix} \pi \sin(\pi x) & 1 \\ -2x & 1 \end{bmatrix} \tag{5.11}$$

Considering each of the four numerical methods in this paper results in the following plots (Figure 5.27). The approximate radius of convergence for each method is listed below its plot. In this case, Broyden's method with the identity matrix displays a chaotic nature (Figure 5.23).

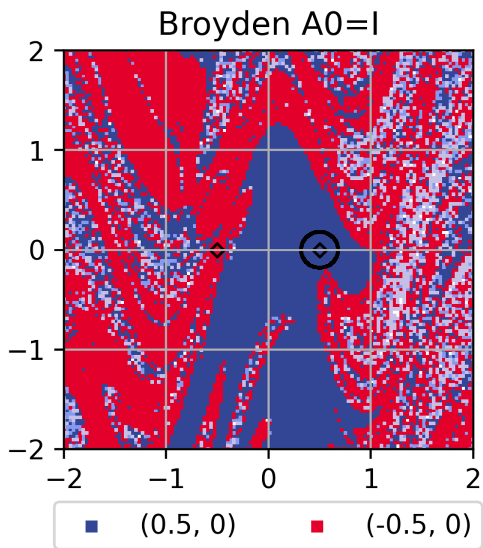

Figure 5.23: $\delta = 0.180$

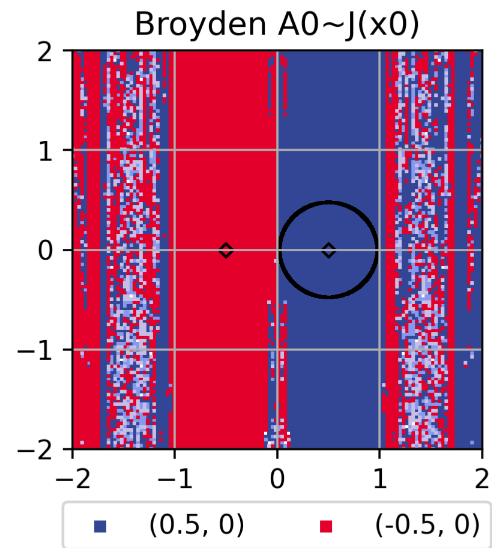

Figure 5.24: $h = 0.1, \delta = 0.476$

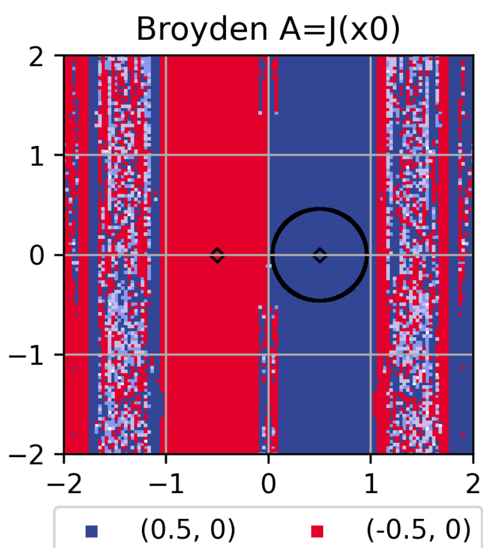

Figure 5.25: $\delta = 0.460$

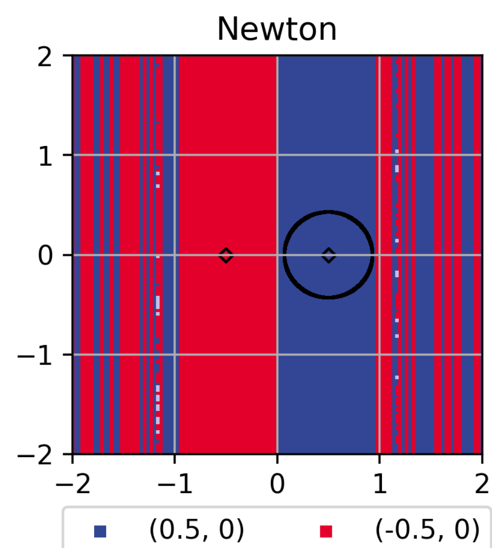

Figure 5.26: $\delta = 0.429$

Figure 5.27: Applying the four numerical methods to equation 5.10 with $\lambda = 1000$.

The root at $(0.5, 0)$ has a more clearly defined region around it for the first method. Improving the initial guess of the Jacobian to consider an approximation of the Jacobian with $h = 0.1$ quickly reduces some of the chaotic behavior of the graph 5.24. Using an explicit calculation for the Jacobian rather than a numerical example does not significantly reduce the 'noise' in the plot, nor is there an improvement in the radius of convergence 5.25. Finally, considering the computationally expensive Newton's method reduces the chaotic nature to clean vertical bands which converge to one root or another 5.26.

5.4.2 Sine, Cosine, and Circle

Consider the system of equations:

$$\begin{aligned} f_1(x,y) &= y - \cos(\pi x)\sin(\pi x) = 0 \\ f_2(x,y) &= x^2 + y^2 - 1 = 0 \end{aligned} \tag{5.12}$$

whose roots are $(-1, 0)$ and $(1, 0)$. The Jacobian of the system is calculated to be:

$$J(x,y) = \begin{bmatrix} -\pi\cos^2(\pi x) + \pi\sin^2(\pi x) & 1 \\ 2x & 2y \end{bmatrix} \tag{5.13}$$

Producing the plots considering the domain of attraction for each method results in 5.32. Broyden's method using the identity matrix as the initial guess produces an incredibly chaotic plot 5.28. Once again, with a more highly non-linear system, a more chaotic behavior is found. The radius of convergence for the root at $(1, 0)$ was found to be only 0.117. Requiring an error of less than 0.117 for an initial guess to be confident in the convergence of the method is impractical.

Improving the initial guess of the Jacobian to a central difference approximation with step size $h = 0.1$ improves the plot considerably 5.29. Although there is still chaotic behavior in a large plot of the graph, the region around each root has been improved. The radius of convergence of this method doubled to 0.226 with the central difference approximation. An explicit calculation moderately improved the radius to 0.234 5.30. As expected, Newton's method introduced a level of symmetry to the graph, which could provide its own usefulness

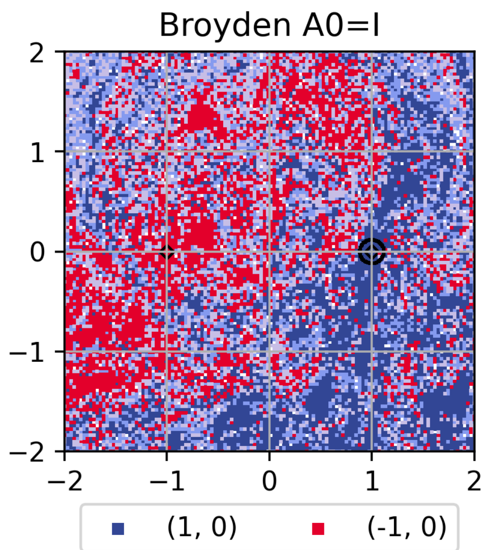

Figure 5.28: $\delta = 0.117$

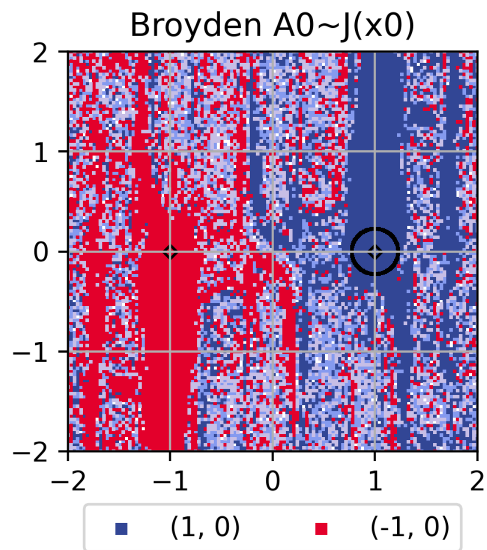

Figure 5.29: $h = 0.1\delta = 0.227$

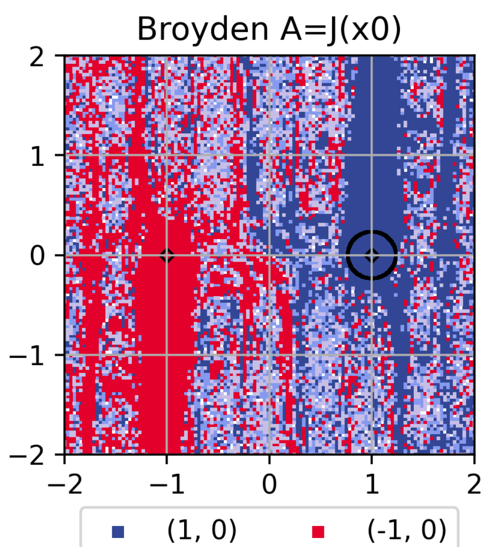

Figure 5.30: $\delta = 0.234$

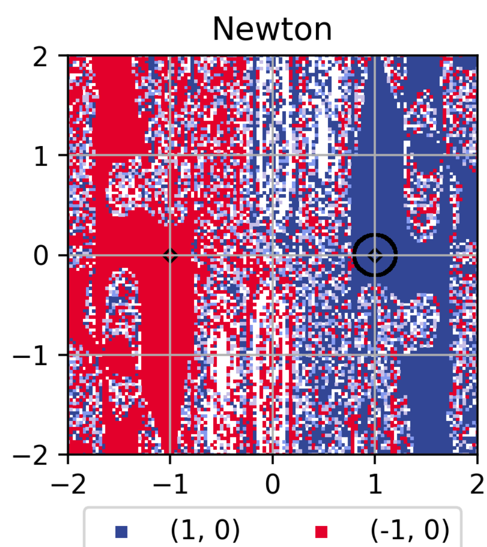

Figure 5.31: $\delta = 0.203$

Figure 5.32: Applying the four numerical methods to the trigonometric equation 5.12 with $\lambda = 1000$.

when looking for multiple roots, but still had a smaller radius of convergence compared to both Broyden's methods with only 0.203 5.31.

5.5 Transcendental Functions and Global Minimization

This section will consider systems of equations with transcendental terms.

5.5.1 Exponential Function

Consider the system of equations

$$\begin{aligned} f_1(x,y) &= y - e^x + \tfrac{e}{2} \\ f_2(x,y) &= x - e^y + \tfrac{e}{2} \end{aligned} \tag{5.14}$$

whose roots are $(0.743, 0.743)$ and $(-0.986, -0.986)$.

The Jacobian of the system is:

$$J(x,y) = \begin{bmatrix} -e^x & 1 \\ 1 & -e^y \end{bmatrix} \tag{5.15}$$

Considering the domain of attraction for each root reveals a relatively well behaving plot. In the first case of Broyden's method using the identity matrix as the initial guess, the radius of convergence was only 0.271 for the blue root in the first quadrant 5.33. Significant improvement was made with an educated initial guess for the Jacobian with either a numerical approximation or explicit calculation 5.34 5.35. Ultimately, Newton's method had the highest radius of convergence at 1.037, but did not have significant improvement over the two educated guesses of Broyden's method.

5.5.2 Optimization Rosenbrock Function

As previously discussed, one application of Broyden's method is with global minimization. Consider the objective commonly used to test performance of optimization algorithms, Rosenbrock's function:

$$f(x,y) = 100(y - x^2)^2 + (x-1)^2 \tag{5.16}$$

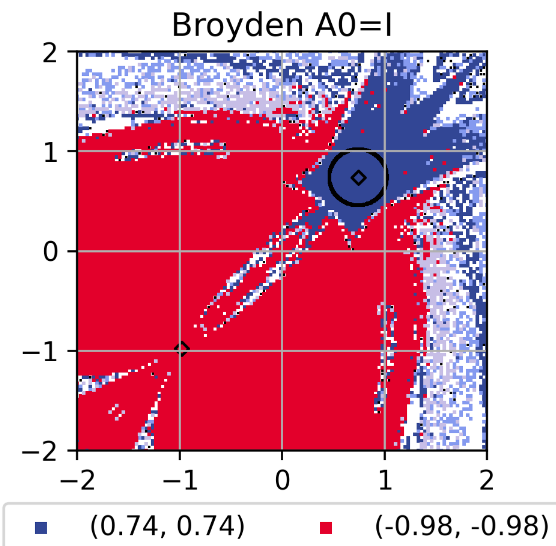

Figure 5.33: $\delta = 0.281$

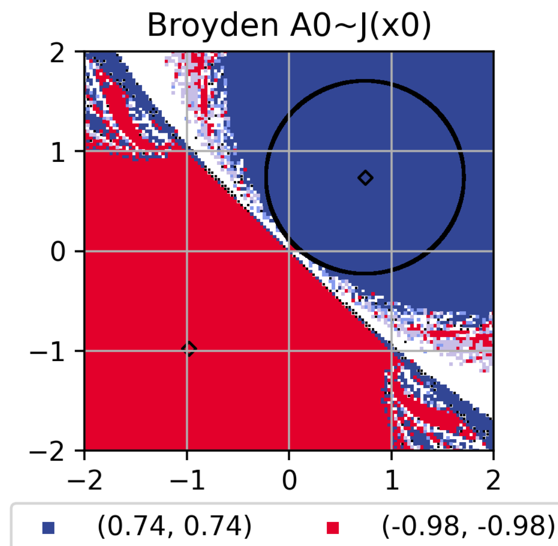

Figure 5.34: $\delta = 0.967$

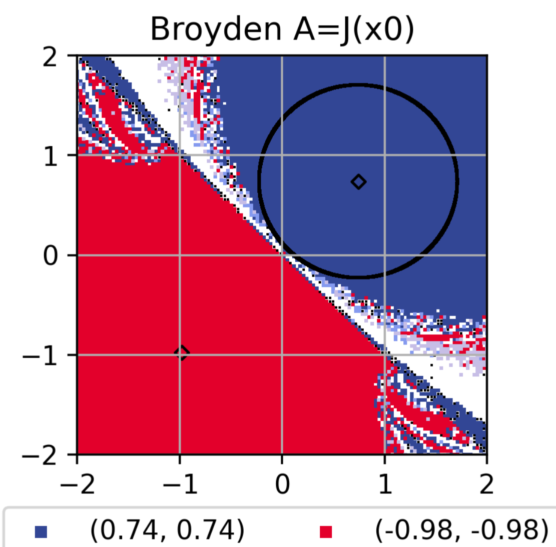

Figure 5.35: $\delta = 0.967$

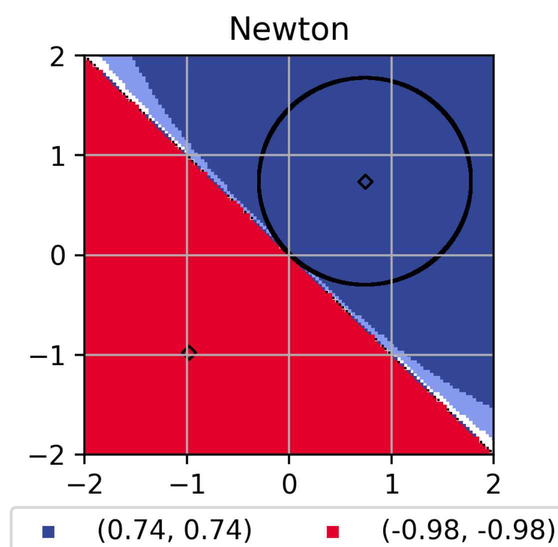

Figure 5.36: $\delta = 1.037$

Figure 5.37: Applying the four numerical methods to the transcendental equation 5.14 with $\lambda = 1000$.

A minimum occurs when the partial derivatives with respect to each variable are zero:

$$\begin{aligned} \frac{\partial f}{\partial x} &= 0 \\ \frac{\partial f}{\partial y} &= 0 \end{aligned} \tag{5.17}$$

In the terms of Broyden's method, the system of equations to find the roots of is:

$$\begin{aligned} f_1(x,y) &= -400xy + 400x^3 + 2x - 1 = 0 \\ f_2(x,y) &= 200(y - x^2) = 0 \end{aligned} \tag{5.18}$$

The root of the function is $(1,1)$. Since this system only has one root, and thus the only local minimum is also the global minimum, the method is concerned with not so much which root is found, but the reliability of the method to converge.

To apply Newton's method, the explicit Jacobian is needed. The Jacobian of the system is:

$$J(x,y) = \begin{bmatrix} -400y + 1200x^2 + 2 & -400x \\ -400x & 200y \end{bmatrix} \tag{5.19}$$

In this example, Newton's method 5.41 and Broyden's method using the identity matrix 5.38 as the initial approximation of the Jacobian performed poorly. Terminating after 100 iterations, few points resulted convergence. Explicitly calculating the Jacobian at the initial guess improved the method significantly, but still showed poor performance. The best performing method approximated the Jacobian using the central difference approximation.

This section presented numerous examples of systems of equations. This section compared four numerical methods of increasing computational complexity:

1. Broyden's method using the identity matrix as the initial approximation to the Jacobian

2. Broyden's method approximating the initial guess of the Jacobian with central difference approximation

3. Broyden's method with the explicitly calculated Jacobian as the initial approximation

4. Newton's method for systems of equations

It was shown that the effect of the initial approximation for the Jacobian greatly depends on the system. For the cases presented in this paper, the radius of convergence was improved by

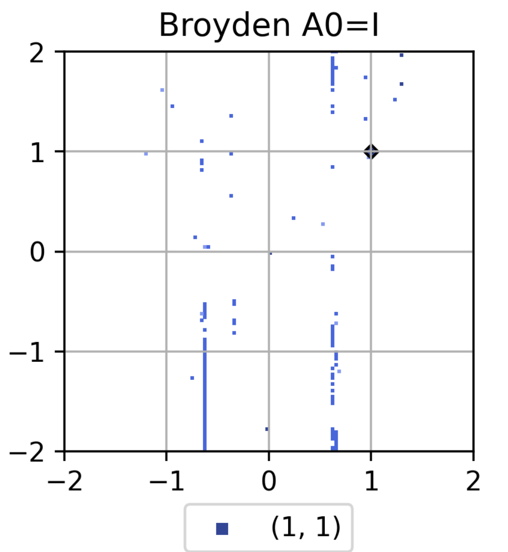

Figure 5.38: $\delta = 0.004$ and $\alpha = 29.37\%$

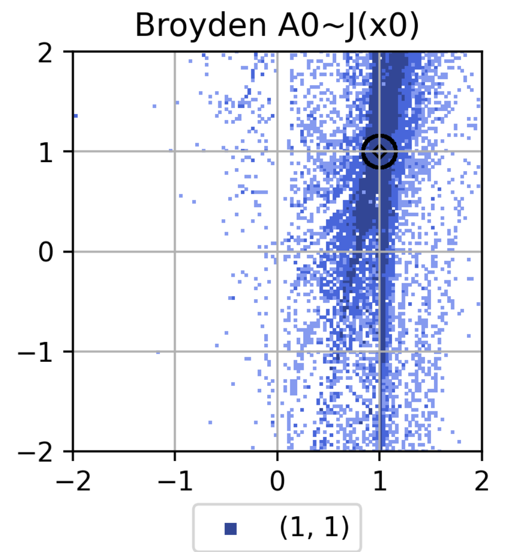

Figure 5.39: $\delta = 0.160$ and $\alpha = 97.06\%$

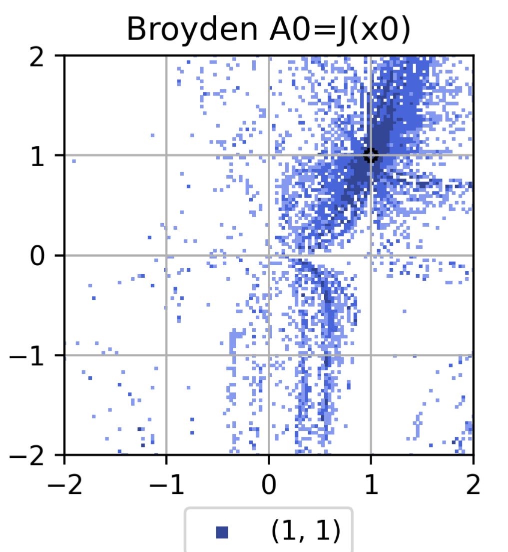

Figure 5.40: $\delta = 0.058$ and $\alpha = 99.08\%$

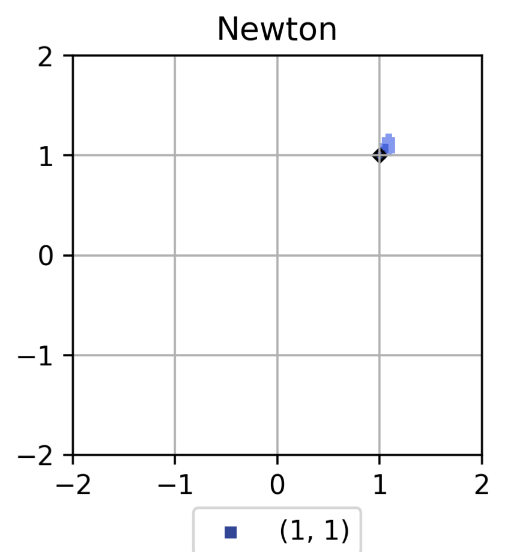

Figure 5.41: $\delta = 0.0274$ and $\alpha = 45.30\%$

Figure 5.42: Applying the four numerical methods to equation 5.18 with $\lambda = 1000$.

using a central difference approximation for the initial guess of the Jacobian. By plotting the domain of attraction for each root and using the previously described method approximating the radius of convergence, this paper supports the importance of an informed initial guess of the Jacobian and the importance of considering the reliability of convergence when evaluating a numerical method. The next section will summarize and discuss future work.

Chapter 6

Concluding Remarks

6.1 Future Work

This section will conclude the paper by discussing future directions this work could take and providing closing remarks.

6.1.1 Method I & Method II

This work studied the domain of attraction for Broyden's method I, but similar work can be applied to Broyden's method II. Each iteration, Broyden's method II approximates the inverse of the Jacobian rather than the Jacobian. Because of this, Broyden's method II will not encounter the problem of needing to invert a singular matrix. In the system considering the intersection of the circle and cubic system (equation 5.8), comparing the domain of attraction of each root between Broyden's method I and Broyden's method II highlights the subtle differences between the methods 6.3.

Although the general shape and structure of the graphs are similar, they are not identical. One future direction of work could be performing a similar analysis of the effect of the initial guess of the Jacobian on Broyden's method II. It would be particularly insightful to see the impact of the two methods on systems with singular Jacobian matrices.

6.1.2 Higher Dimension Systems

This paper considered only systems of two equations and two variables. Two dimensional examples were chosen for their convenience to plot solutions and interpret the roots. Future

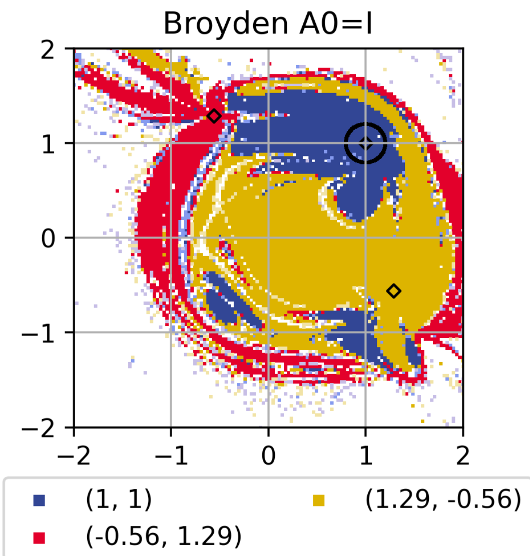

Figure 6.1: Broyden's method I

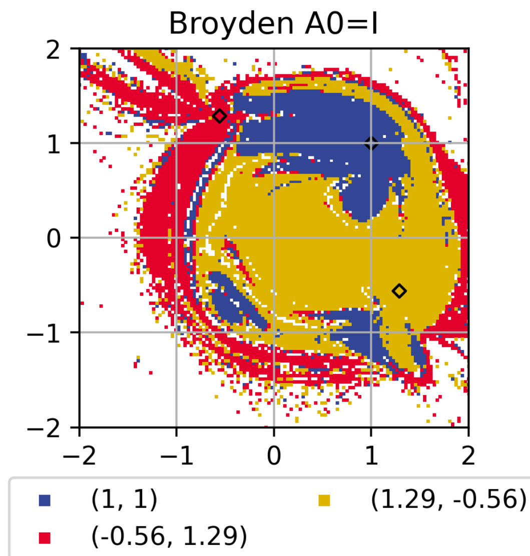

Figure 6.2: Broyden's method II

Figure 6.3: Comparing domain of attraction of equation 5.8 for Broyden's method I and Broyden's method II.

work should include investigating the behavior of higher dimensional cases. Higher dimensions will require more advanced graphing methods to represent the basin of attraction, and the consideration of a hyper-sphere rather than a circle to approximate the radius of convergence. In higher dimensional systems it will be interesting to see if the initial guess of the Jacobian is more important than in lower dimensional cases.

This work presented cases of systems with one root, two roots or multiple roots. The number of roots were all determined before analysis. Interesting questions to explore in future research include determining teh number of roots of a system and determinign if an understanding of the domain of attraction can be used in finding all solutions. If a root is found, but the root does not have interpretation back to the physical system (negative mass for example), then the method should be repeated with different initial guesses until a satisfactory root is found. At what point can the search for alternative roots be called off? Future work can use the understanding of the domain of attraction to provide initial guesses for a system to search for different roots.

6.1.3 Behavior of Systems

Systems presented in this paper exhibited a variety of behaviors based on the system of equations. There is limited research involving system behavior or commonalities between well behaved and poorly behaving systems. Future work can include examining why systems behave one way or another. Understanding what causes a system to behave chaotically or not could be useful in predicting if a method will converge. Recognizing a system as chaotic without needing to explicitly calculate the root and measure the basin of attraction could save computation time and create more reliable numerical methods.

In this paper, *a priori* knowledge of the roots was used for each system, though this is unlikely in practice. Future work could continue to develop methods to approximate the radius of convergence without prior knowledge of the root and without needing to explicitly evaluate each value.

6.2 Closing Remarks

This paper provided background on Newton's method for systems of equations. Although able to locally converge quickly, a drawback of the method was needing to explicitly calculate the Jacobian. To avoid needing to explicitly calculate the Jacobian, Broyden's method was introduced. Broyden's method used an iterative scheme to approximate the root and approximate the Jacobian at each iteration.

This paper then showed two examples of applications of Broyden's method. Broyden's method was used to optimize an objective function using partial derivatives. Next, an example of a simple inverse kinematics question showed how the formulation of the system affects the convergence of the method.

After demonstrating examples of the applications of Broyden's method, the effect of the initial guesses was considered. The domain of attraction of a root was defined to be the set of all initial guesses which converge to the root. Plots were made for systems with multiple roots to show how the initial guess affects the root the method converges to. For some systems, the plots produced well-defined boundaries between regions which converged to different roots. Other plots exhibited chaotic behavior, with no well-defined regions separating the roots each initial guess converged to. This motivated further examination into what could be done to improve the basin of attraction of the method.

Two methods of numerically approximating the radius of convergence were developed. The first method considered finding the circle of the largest radius centered on the root such that all values evaluated on the circumference of the circle converge to the root at the center. A parameter of linear density, λ, was introduced to specify the points per unit length to evaluate along the circumference of the circle. Higher values of λ increased confidence in knowing that all values of the circumference would converge but would require evaluating more points in the method. The second method used a Monte Carlo simulation with a specified α value. The method approximated the largest circle centered at a root such that α of the points within the circle would converge to the root. Practically, it was easy to understand the importance of being 99%, 99.9% or 99.99...% confident in a method converging to the desired root. A parameter ρ_A was introduced which specified the area density of points to evaluate. Since there is a random nature to Monte Carlo methods, the results required sufficiently large values of ρ_A for consistent results. This paper favored the first method of to ensure repeatability in testing.

This paper concluded by demonstrating the improvement made to the basin of attraction by using an educated initial guess for the Jacobian. This paper compared four methods of solving systems of non-linear equations with decreasing computational complexity:

1. Newton's method
2. Broyden's method with the Jacobian explicitly calculated for the initial guess
3. Broyden's method with the Jacobian initially approximated using the central difference theorem
4. Broyden's method with the initial guess of the Jacobian the identity matrix.

This paper showed improved convergence properties when an informed first guess for the Jacobian was used. Examples were shown which exhibited significant improvement in the basin of attraction diagram by using a Jacobian approximation. In some cases, Broyden's method with an initial approximation of the Jacobian was able to have a larger radius of convergence than Newton's method. Although for an $n \times n$ system, approximating the Jacobian with central difference requires $2n^2$ function evaluations, the improvement made to the method is significant. This paper demonstrates the importance of approximating the Jacobian when initializing Broyden's method.

REFERENCES

[1] M. Raydan, *Nonlinear Equations and Newton's Method in Princeton Companion to Applied Mathematics*, 1st ed. Princeton, New Jersey: Princeton University Press, 2015.

[2] C. Broyden, "A class of methods for solving nonlinear simultaneous equations," *Mathematics of Computation*, vol. 19, no. 92, pp. 577–593, Oct. 1965.

[3] T. Sauer, *Numerical Analysis*, 1st ed. Boston, Massachusetts: Pearson Education, 2006.

[4] J. Dennis and R. B. Schnabel, *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*, 1st ed. Englewood Cliffs, New Jersey: Prentice-Hall, 1983.

[5] T. Kalvouridis and M. Gousidou-Koutita, "Basins of attraction in the copenhagen problem where the primaries are magnetic dipoles," *Applied Mathematics*.

[6] F. Mannel, "On the order of convergence of broyden's method," *Calcolo*, vol. 58, no. 4, pp. 47–21.

[7] R. Byrd, D. Liu, and J. Nocedal, "On the behavior of broyden's class of quasi-newton methods," *Siam Journal on Optimization*, vol. 2, no. 4, pp. 533–557.

[8] S. Lui and S. Nataj, "Superlinear convergence of broyden's method and bfgs algorithm using kantorovich-type assumptions," *Journal of Computational and Applied Mathematics*, vol. 385, p. 113204–21.

[9] M. Conner, P. Donescu, and L. Virgin, "On the global convergence characteristics of numerically evaluated jacobian matrices," *Nonlinear Dynamics*, vol. 10, pp. 165–174.

[10] M. Al-Towaiq and Y. A. Hour, "Two improved classes of broyden's method for solving nonlinear systems of equations," *Journal of Mathematics and Computer Science*.

[11] W. Cheney and D. Kincaid, *Numerical Mathematics and Computing*, 5th ed. Belmont, California: Brooks/Cole-Thomson Learning, 2004.

[12] G. Strang, *Linear Algebra and its Applications*, 3rd ed. Orlando, Florida: Harcourt Brace Jovanovich, 1988.