Chaos in Relaxed Newton's Method: The Quadratic Case

L. Billings, J. H. Curry, V. Robins
Department of Applied Mathematics
University of Colorado at Boulder

August 19, 1999

Abstract

In [8], Gilbert Strang explored the chaotic behavior of Newton's method applied to $x^2 + 1$ when the iteration starts from only real initial conditions. We embed Strang's problem in the one-parameter family of iteration methods known as relaxed Newton's method. By extending this analysis to include a discussion of invariant measures, probability densities, and Lyapunov exponents, we prove that in the quadratic case, relaxed Newton's method exhibits chaotic dynamics as well.

1 Introduction

This article makes a contribution to an old problem first considered by Sir Arthur Cayley in 1879, the convergence of Newton's method for quadratic polynomials. Gilbert Strang later used the divergence of Newton's method to illustrate the idea of chaos in [8], since most calculus textbooks cover this familiar root finding algorithm. Newton's method is defined here for any function $f(x)$

$$x_{n+1} = x_n - f(x_n)/f'(x_n). \tag{1}$$

Generating x_{n+1} from x_n is called an iteration step. Repeating the iteration process and listing each x_{n+1} forms a sequence, or orbit, generated by an initial condition, which is the x_n used in the first step. The sequence may or may not converge to a root of the function $f(x)$ due to the choice of the initial condition or the function $f(x)$ itself.

In the search of alternative algorithms, variations of Newton's method were developed. One such example is relaxed Newton's method, defined by

$$x_{n+1} = N_\beta(x_n) = x_n - \beta\left(\frac{f(x_n)}{f'(x_n)}\right). \tag{2}$$

Notice that Newton's method is embedded in this family at the parameter value $\beta = 1$. For simple roots, relaxed Newton's method exhibits linear convergence—not quite as good as the quadratic convergence of Newton's method. The advantage is that a full Newton step ($\beta = 1$) may not be suitable to ensure that convergence is monotone. (For a classical treatment of the convergence properties of Newton's method applied to quadratic polynomials see Appendix F of [6] or see [4].) Further, this method can be used to improve the convergence of Newton's method when applied to a polynomial with multiple roots, assuming that the multiplicity of the root is known. (See [2] for details.)

If $f(x)$ is a quadratic with a complex conjugate pair of roots, there are no real roots for relaxed Newton's method to find. Almost every initial condition produces an orbit that never converges or repeats an iteration step. Our approach is to view relaxed Newton's method as a dynamical system on the real line in this situation and prove the long term behavior of its orbits is in fact chaotic. We also extend this analysis to include a discussion of how to find invariant measures, probability densities, and Lyapunov exponents analytically. But first, we revisit the classical problem discussed by Gilbert Strang in [8] and prove that Newton's method applied to $f(x) = x^2 + 1$ has chaotic orbits.

1.1 The case $\beta = 1$: Newton's method

For $\beta = 1$, the map N_1 is standard Newton's method applied to the quadratic $f(x) = x^2 + 1$

$$N_1(x) = x - \frac{x^2+1}{2x} = \frac{1}{2}\left(x - \frac{1}{x}\right). \tag{3}$$

N_1 has fixed points at $\pm i$. However, since x and β are real numbers, relaxed Newton's method fails to converge to the roots of the given quadratic. When viewed as a dynamical system, we can ask additional questions about the iteration, for example: Is the sequence that the iteration produces chaotic? And can we establish this analytically?

In [8], it is established that $N_1(x)$ is topologically conjugate to the piecewise linear map $g(x) = 2x(\text{mod } 1)$. That is:

Definition: Two maps, f and g, are *conjugate* if they are related by a homeomorphism h, that is a continuous one-to-one change of coordinates, such that $h \circ f \circ h^{-1} = g$.

In the case of N_1 this means, $h \circ N_1 \circ h^{-1} = 2x(\text{mod } 1)$. Where

$$h(x) = \frac{1}{2} + \frac{1}{\pi}\tan^{-1}(x). \tag{4}$$

The associated conjugacy diagram is

$$\begin{tikzcd} \mathbb{R} \arrow[r, "N_1"] \arrow[d, "h"'] & \mathbb{R} \arrow[d, "h"] \\ {[0,1]} \arrow[r, "g"] & {[0,1]} \end{tikzcd} \; . \tag{5}$$

Conjugacy means that this diagram commutes. See Figure 1 for graphs of $N_1(x)$ and $g(x)$.

Since $h(x)$ is a homeomorphism between $N_1(x)$ and $g(x)$, there is a one-to-one correspondence between the dynamics of Newton's method and a map of the interval. So, for example, periodic

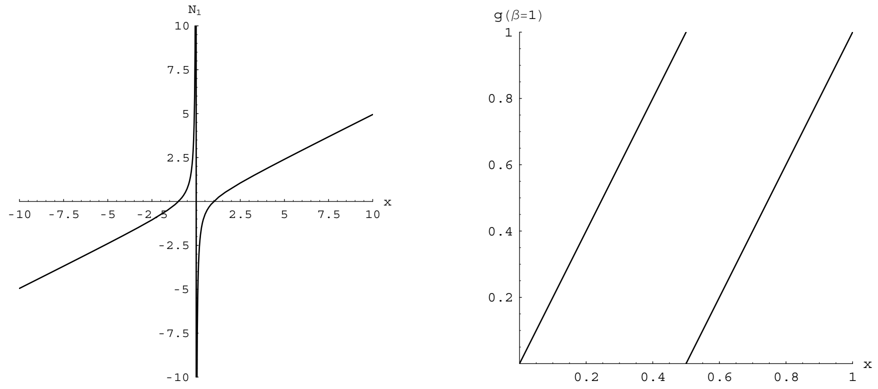

Figure 1: The left graph is N_1 and the right graph is $2x(\text{mod } 1)$.

orbits for Newton's method are transformed into similar orbits for $g(x)$. As we shall see this correspondence proves to be convenient. It allows us to establish results in the best of all possible phase spaces and sometimes with only a modest amount of work.

We use the property that a conjugacy preserves chaos.

Definition: A function $f(x)$ is *chaotic* on a set if (Robinson [7]):

(a) $f(x)$ is transitive on that set.

(b) $f(x)$ has sensitive dependence on initial conditions.

Example: $g(x) = 2x(\text{mod } 1)$ is chaotic on $[0, 1]$:

(a) $g(x)$ is transitive on $[0, 1]$.
The definition of transitive is equivalent to the condition that given any two open sets U and V in $[0, 1]$, there is an integer n such that $g^n(U) \cap V \neq \emptyset$. This follows from the fact that g is expanding. Since any open set will be stretched in length by a factor of 2 at each iteration, it will eventually cover the whole set, therefore intersecting V.

(b) $g(x)$ has sensitive dependence on initial conditions. Sensitive dependence on initial conditions is implied by a positive Lyapunov exponent. The Lyapunov exponent measures the exponential rate at which neighboring orbits are moving apart. It is determined by averaging the natural logarithm of the derivative evaluated along an orbit.

$$\begin{aligned}
\Lambda &= \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \ln |g'(x_i)| \\
&= \int_0^1 \ln |g'| \, d\lambda \\
&= \ln 2.
\end{aligned}$$

The second expression follows from the fact that Lebesgue measure, λ, is an invariant, ergodic measure for g and so we can apply the Birkhoff Ergodic theorem that time averages are the same as space averages. Clearly, $\ln 2 > 0$ and so g has sensitive dependence on initial conditions.

Finally, the properties of transitivity and sensitive dependence are preserved under conjugacy. In fact, Lyapunov exponents are identical for topologically conjugate maps. Since the map $2x(\text{mod } 1)$ is chaotic, $N_1(x)$ must also be chaotic.

1.2 Outline

By applying the same conjugacy h to the entire family of relaxed Newton's method maps, N_β, we obtain a family, g_β, of maps of the unit interval. The approach outlined in the above example will be followed below for the families N_β and g_β. We establish that the members of these families exhibit chaotic dynamics by showing that they are transitive on their respective phase spaces and that they have sensitive dependence on initial conditions. A consequence of this study is that we determine invariant, ergodic measures for both families of maps and establish an analytic formula for the Lyapunov exponents of relaxed Newton's method. These results are based on elementary mathematics.

In section 2 we establish our main results.

2 Chaotic Dynamics: The conjugacy does all the work

We first show N_β satisfies both transitivity and sensitive dependence on initial conditions by applying the same conjugacy transformation used to make Newton's method conjugate to the map $2x(\text{mod } 1)$. The result is a family of maps denoted by g_β. The topological conjugacy between N_β and g_β is represented by the diagram below.

$$\begin{tikzcd}
\mathbb{R} \arrow[r, "N_\beta"] \arrow[d, "h"'] & \mathbb{R} \arrow[d, "h"] \\
{[0,1]} \arrow[r, "g_\beta"] & {[0,1]}
\end{tikzcd} \quad . \tag{6}$$

The composite function $g_\beta = h \circ N_\beta \circ h^{-1}$ simplifies to

$$g_\beta = \frac{1}{2} + \frac{1}{\pi} \tan^{-1}\left(\frac{1 - \beta + \cos(2\pi x)}{\sin(2\pi x)}\right). \tag{7}$$

The g_β maps are piecewise monotone increasing functions on the interval $[0,1]$ as illustrated by the two examples in Figure 2. In the analysis that follows we exploit this conjugacy and make use of both N_β and g_β.

2.1 Transitivity

To demonstrate transitivity we will construct the proof for g_β.

Proposition The map $g_\beta : [0,1] \to [0,1]$ is expanding for $0 < \beta < 2$.

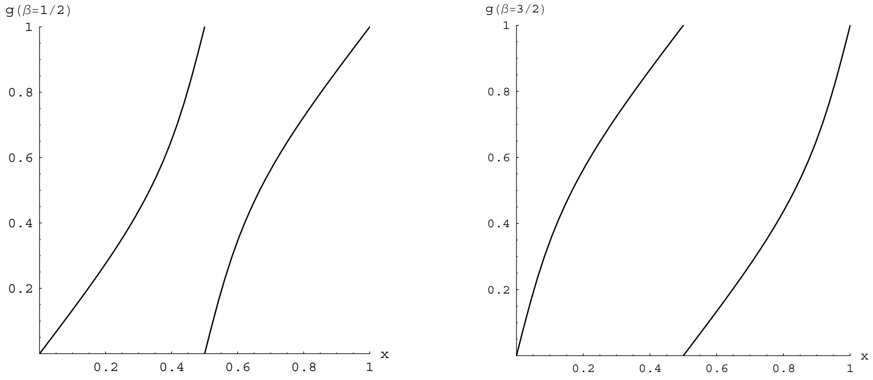

Figure 2: The left graph is $g_{1/2}$ and the right graph is $g_{3/2}$.

Proof. The derivative of g_β is

$$g'_\beta(x) \quad = \quad \frac{2(1+(1-\beta)\cos(2\pi x))}{1+(1-\beta)^2+2(1-\beta)\cos(2\pi x)}. \tag{8}$$

First observe that the denominator is positive:

$$1+(1-\beta)^2+2(1-\beta)\cos(2\pi x) \quad = \quad (1-\beta+\cos(2\pi x))^2+(sin(2\pi x))^2 \ > \ 0. \tag{9}$$

We now want to show that $g'_\beta > 1$ for $0 < \beta < 2$, a fact that follows from the chain of inequalities below:

$$\begin{aligned} 1+(1-\beta)^2 \quad &< \quad 2 \\ 1+(1-\beta)^2+2(1-\beta)\cos(2\pi x) \quad &< \quad 2+2(1-\beta)\cos(2\pi x) \\ \Longrightarrow \quad g'_\beta(x) \quad &= \quad \frac{2(1+(1-\beta)\cos(2\pi x))}{1+(1-\beta)^2-2(1-\beta)\cos(2\pi x)} \ > \ 1. \end{aligned} \tag{10}$$

It now follows that any open set will be stretched in length after each iteration, hence it will eventually cover the whole set. Therefore, for any two open sets U and V in $[0,1]$, there is an integer n such that $g^n_\beta(U) \cap V \neq \emptyset$, that is, $g_\beta(x)$ is transitive on $[0,1]$. As a consequence of the conjugacy, N_β is also transitive.

2.2 Sensitive dependence on initial conditions

One way to establish sensitive dependence on initial conditions is to prove that neighboring orbits diverge exponentially on a local scale. This is a consequence of showing that the map has positive Lyapunov exponent, a quantity also preserved by the conjugacy. Recall that the Lyapunov exponent is determined by appealing to the formula

$$\begin{aligned} \Lambda \quad &= \quad \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \ln\left|N'_\beta(x_i)\right| \\ &= \quad \int_{-\infty}^{\infty} \ln\left|N'_\beta\right| d\rho. \end{aligned} \tag{11}$$

The first formula is the one implemented numerically and the second holds by Birkhoff's ergodic theorem if there is an invariant, ergodic measure, ρ, for N_β. This is the formula we use to obtain analytical results.

2.2.1 The invariant measure

As discussed in [1], the "natural", invariant, ergodic measure associated with the dynamics allows us to calculate quantities that are sampled and averaged over orbits. Its existence guarantees that for almost all initial conditions we will obtain the same average. We need to determine this measure in order to calculate the Lyapunov exponents.

Definition: A measure is *invariant* for a map if the measure of any Borel set is the same as the measure of its preimage.

The function g_β must have an invariant measure by the following theorem:

Theorem 1 *(Lasota and Yorke [5]) Assume that the map g on $[0,1]$ is piecewise smooth and piecewise expanding. Then g has an absolutely continuous invariant measure μ.*

Since the invariant measure, μ, is absolutely continuous with respect to Lebesgue measure, it has a density function, $p(x)$. Given any measurable set $S \subset [0,1]$ we have

$$\mu(S) = \int_S p(x)\ dx \tag{12}$$

Therefore, the density function can be used to find the invariant measure. In fact, the density function of g_β is $p(x) = 1$, making the invariant measure simply Lebesgue measure. Since the preimage of g_β cannot be found explicitly, the proof that this measure is invariant under g_β is not trivial. With this intuition in the g_β phase space, we return to the N_β phase space, suspecting that the conjugacy has contributed h' as the density function.

First, we prove that the measure, defined by

$$\tilde{\mu}(S) = \int_S \frac{1}{\pi\left(x^2+1\right)}\ dx, \tag{13}$$

is invariant. Name the two preimages of the function N_β as z_1 and z_2. Therefore, the preimage of any definite interval (excluding $x = 0$) on the real line $[x_1, x_2]$ is the two intervals $[z_1(x_1), z_1(x_2))]$ and $[z_2(x_1), z_2(x_2)]$. The measure of the sum of the preimages is equal to the measure of set, or equivalently:

$$\begin{aligned}
\int_{z_1(x_1)}^{z_1(x_2)} \tfrac{1}{\pi(x^2+1)}\ dx + \int_{z_2(x_1)}^{z_2(x_2)} \tfrac{1}{\pi(x^2+1)}\ dx &= \tan^{-1}\left(z_1\left(x_2\right)\right) - \tan^{-1}\left(z_1\left(x_1\right)\right) \\
&\quad + \tan^{-1}\left(z_2\left(x_2\right)\right) - \tan^{-1}\left(z_2\left(x_1\right)\right) \\
&= \tan^{-1}\left(\tfrac{z_1(x_2)+z_2(x_2)}{1-z_1(x_2)z_2(x_2)}\right) - \tan^{-1}\left(\tfrac{z_1(x_1)+z_2(x_1)}{1-z_1(x_1)z_2(x_1)}\right) \\
&= \tan^{-1}\left(x_2\right) - \tan^{-1}\left(x_1\right) \\
&= \int_{x_1}^{x_2} \tfrac{1}{\pi(x^2+1)}\ dx.
\end{aligned} \tag{14}$$

The middle step follows from the formulae for the sum and product of the roots of a quadratic.

Another way to prove h' is the density for an invariant measure of $N_\beta(x)$ is to show it is a fixed point of the Frobenius-Perron operator.

Given a map τ, the operator P_τ acts on $\mathcal{L}^1$ functions and is defined as follows.

$$\int_A P_\tau f d\lambda = \int_{\tau^{-1}(A)} f d\lambda, \tag{15}$$

where $d\lambda$ is Lebesgue measure and A is any measurable compact set in the domain. The goal is to have $P_\tau f = f$, implying that f is the density for a measure that is invariant under τ.

If τ is a piecewise monotone function, the action of the operator is simply (see [3] for a derivation)

$$P_\tau f(x) = \sum_{z \in \{\tau^{-1}(x)\}} \frac{f(z)}{|\tau'(z)|}. \tag{16}$$

Setting $f(x) = h'(x) = 1/(\pi(x^2+1))$ and $\tau = N_\beta$ it follows that

$$\begin{aligned} P_{N_\beta} h'(x) &= \frac{h'(z_1(x))}{|N'_\beta(z_1(x))|} + \frac{h'(z_2(x))}{|N'_\beta(z_2(x))|} \\ &= \frac{1}{\pi\,(x^2+1)} \;=\; h'(x). \end{aligned} \tag{17}$$

We can now write the invariant, ergodic measure for N_β as

$$\rho(A) = \int_A d\rho = \int_A \frac{dx}{\pi(x^2+1)}. \tag{18}$$

By showing h' is the density for N_β, we have also confirmed that Lebesgue measure on $[0,1]$ is ergodic and invariant for the g_β family of maps.

2.2.2 The Lyapunov exponent

Now that we have found the absolutely continuous invariant measure for the family N_β, we can use it to find the Lyapunov exponent, and therefore quantify the average rate of expansion or contraction for an interval under iteration. Formula (11) becomes

$$\begin{aligned} \Lambda_\beta &= \int_{-\infty}^{\infty} \ln \left|N'_\beta(x)\right| \frac{dx}{\pi(x^2+1)} \\ &= \ln\left(1+\sqrt{\beta(2-\beta)}\right). \end{aligned} \tag{19}$$

For a graph of Λ_β, see Figure 3. The integral is calculated using contour integration. We remark that this is also the Lyapunov exponent for g_β, however, the integral in that case is more difficult to compute.

Because the Lyapunov exponent is positive for $0 < \beta < 2$, N_β has sensitive dependence on initial conditions for these parameters. Therefore both requirements for chaotic dynamics are fulfilled and our main result is proved.

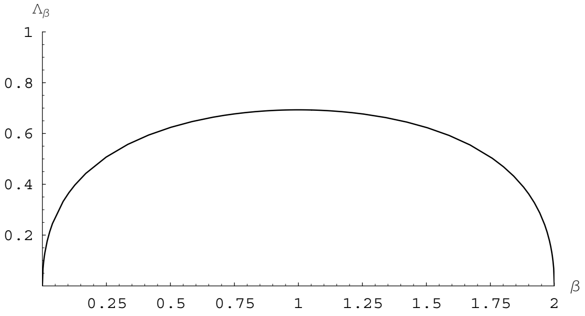

Figure 3: The Lyapunov exponent of N_β.

3 Conclusions

The results we have presented are in some sense a complement to a problem first considered by Sir Arthur Cayley in 1879, who also was the first to establish the convergence of Newton's method for quadratic polynomials. We have provided a complete analysis of the dynamics on the separatrix of the two roots. Here, we show that relaxed Newton's method applied to a quadratic with complex roots generates chaotic dynamics. Our results are accessible via elementary means, but less elementary extensions are possible using the technology of one dimensional dynamics and the Frobenius-Perron operator.

References

[1] K. Alligood, T. Sauer, and J. Yorke, *Chaos*, Springer-Verlag, 1996.

[2] K. Atkinson, *An Introduction to Numerical Analysis*, Wiley, 1978.

[3] A. Boyarsky and P. Góra, *Laws of Chaos: Invariant Measures and Dynamical Systems in One Dimension*, Birkhäuser, 1997.

[4] R. Holmgren, *A First Course in Discrete Dynamical Systems*, Springer Verlag, 1996.

[5] A. Lasota and J. Yorke, *On the existence of invariant measures for piecewise monotonic transformations*, Trans. Amer. Math. Soc., 186 (1973), pp. 481–488.

[6] A. M. Ostrowski, *Solutions of Equations in Euclidean and Banach Spaces*, Academic Press, 1973.

[7] C. Robinson, *Dynamical Systems*, CRC Press, Inc., 1995.

[8] G. Strang, *A chaotic search for i*, The Coll. Math. J., 22 (1991), pp. 3–12.