CHAOS IN NEWTONIAN ITERATIONS: SEARCHING FOR ZEROS WHICH ARE NOT THERE*

Łukasz Skowronek[a†], Paweł F. Góra[a,b‡]

[a]M. Smoluchowski Institute of Physics, Jagellonian University
Reymonta 4, 30-059 Kraków, Poland

[b]Mark Kac Complex Systems Research Center, Jagellonian University
Reymonta 4, 30-059 Kraków, Poland

(Received February 28, 2007)

We show analytically that Newtonian iterations, when applied to a polynomial equation, have a positive topological entropy. In a specific example of an attempt to "find" the real solutions of the equation $x^2 + 1 = 0$, we show that the Newton method is chaotic. We analytically find the invariant density and show how this problem relates to that of a piecewise linear map.

PACS numbers: 05.45.–a

1. Introduction

Suppose we want to numerically solve a nonlinear equation $f(z) = 0$, where $z \in \mathbb{R}$ or $z \in \mathbb{C}$. If only calculating the derivative of f is possible, perhaps the famous Newton (or Newton–Raphson) method [1] is the first method that comes to mind. Starting from some z_0, this method uses the following iterations:

$$z_{n+1} = z_n - \frac{f(z_n)}{f'(z_n)} = \eta(z_n)\,. \tag{1}$$

If this iteration converges, it usually does so very fast. Because of that, and because of its simplicity, the Newton method is one of the most frequently used numerical methods ever. Yet it is widely known that even if $f(z)$ is a low-order polynomial, boundaries between basins of attractions of different

* Presented at the XIX Marian Smoluchowski Symposium on Statistical Physics, Kraków, Poland, May 14–17, 2006.

† cwirus@poczta.onet.pl

‡ gora@if.uj.edu.pl

roots can be very complicated. Indeed, many pieces of popular software use this particular property of the Newtonian iterations to produce aesthetically pleasing fractals.

The Newton method fails if it hits a zero of the derivative, or when the denominator in Eq. (1) vanishes, unless the zero of the derivative corresponds to a multiple root. It is less commonly known that the Newton method also fails to find a root if it forms a multicycle. Many textbooks that teach the Newton method do not even mention this fact and those that do make the impression that multicycles in the Newton method are a rare and unimportant peculiarity. In reality, however, constructing examples that display multicycles is quite easy. A reader may verify, for instance, that the points $\{0, 1\}$ form a stable 2-cycle, shown in Fig. 1, in the Newton method applied to the polynomial $x^3 - 2x + 2$. Several questions then arise: How many such multicycles are there? Is it possible to encounter them in practical numerical applications of the Newton method? Is there a connection between the multicycles and, say, basins of attraction?

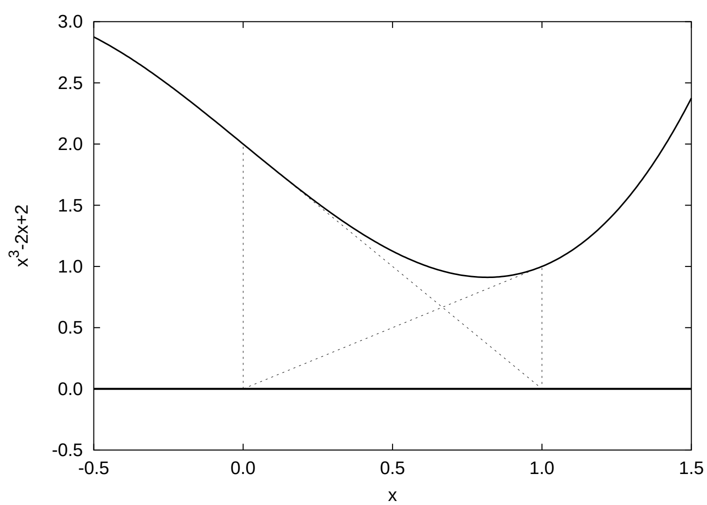

Fig. 1. A 2-cycle generated by the Newton method applied to the equation $x^3 - 2x + 2 = 0$.

We may regard consecutive iterations of the Newton method as a dynamical system. Multicycles generated by the Newtonian iterations are the periodic orbits of this dynamics. Since the seminal works of Cvitanovic and co-workers [2], it is known that the unstable periodic orbits (UPOs) carry important information about the dynamical system that generates them. Yet to extract this information, one needs to know at least how many periodic orbits there are.

We shall discuss these questions in the case of $f(z)$ being a polynomial.

This paper is organized as follows: In Section 2 we show that the Newtonian iterations applied to a polynomial equation can generate infinitely many multicycles and we calculate the topological entropy associated with them. In Section 3 we discuss a particularly simple case of the polynomial z^2+1. We show that dynamics resulting from the Newton method, when restricted to the real axis, is chaotic and equivalent to the dynamics generated by a piecewise linear map. We explicitly calculate the invariant density of the Newtonian dynamics in this case. Since the number of chaotic systems whose invariant densities are known analytically is fairly limited, this result is very interesting, at least from a pedagogical point of view. In Section 4 we partially generalize these results to the case of polynomials $z^{2m}+1$, $m \geqslant 2$. Conclusions are given in Section 5. The Appendix contains a proof of the theorem that we use to calculate the topological entropy.

2. Polynomials and multicycles

Let $P(z)$ be a complex polynomial of order n:

$$P(z) = \sum_{s=0}^{n} a_s z^s \,. \tag{2}$$

Suppose the polynomial Eq. (2) has $n_{\rm d}$ distinct roots[1]. We solve the polynomial equation

$$P(z) = 0 \,, \tag{3}$$

numerically by Newtonian iterations. This procedure defines a function

$$\eta(z) = z - \frac{P(z)}{P'(z)} \,. \tag{4}$$

Consecutive iterates of Eq. (4) satisfy

$$\eta^k(z) = \eta^{k-1}(z) - \frac{P\left(\eta^{k-1}(z)\right)}{P'\left(\eta^{k-1}(z)\right)} \,, \quad k = 1, 2, \ldots \,, \tag{5}$$

with $\eta^0(z) \equiv z$. The following Theorem gives a more detailed characteristics of the iterates of the function Eq. (4). It is pivotal in our subsequent discussion:

[1] Distinct roots of a polynomial must not be confused with roots of muliplicity one. For example, the polynomial $x^2(x-1)^2(x-2)$, of order five, has *three* distinct roots.

Theorem 1. $\forall\, k \geqslant 1$

$$\eta^k(z) = z - \frac{P(z)}{P'(z)}\,\frac{A_k(z)}{B_k(z)}\,, \tag{6}$$

where A_k, B_k are polynomials of order $n_{\mathrm{d}}^k - n_{\mathrm{d}}$, where n_{d} is the number of *distinct* roots of the polynomial $P(z)$.

An elementary, but rather lengthy and technical, proof of this Theorem is given in the Appendix.

Theorem 1 is very useful when we consider multicycles generated by the function Eq. (4). Specifically, all points belonging to a k-cycle satisfy

$$\eta^k(z) = z\,. \tag{7}$$

Using Eq. (6) we get that either $P(z) = 0$, which means that z is a fixed point of the function Eq. (4), or that

$$A_k(z) = 0\,. \tag{8}$$

This equation has $n_{\mathrm{d}}^k - n_{\mathrm{d}}$ solutions on the complex plane. If k is prime, then there are $N_k = (n_{\mathrm{d}}^k - n_{\mathrm{d}})/k$ different k-cycles. If k is not prime, we need to subtract "spurious" k-cycles; for example, two rounds over a 2-cycle can be mistakenly taken for a 4-cycle. Eventually, for the number of k-cycles that are generated by the Newton method applied to a polynomial with n_{d} distinct zeros, we obtain

$$N_k = \frac{1}{k}\Big(n_{\mathrm{d}}^k - n_{\mathrm{d}} - \sum_{j\in D_k} j\,N_j\Big), \tag{9}$$

where D_k is a set of all proper divisors of k (if k is prime, $D_k = \emptyset$). These multicycles are periodic orbits of the dynamics generated by successive iterations of the function Eq. (4). A limited number of the multicycles may be stable, but as in any dynamical system, an overwhelming majority of them are unstable and form the unstable periodic orbits (UPOs) of the dynamics.

The number of multicycles grows very rapidly with k; for example, if $n_{\mathrm{d}} = 5$, there are approximately 4.5×10^{10} 17-cycles. To quantify this observation, we can calculate the topological entropy [3] of the dynamics generated by $\eta(x)$ by appropriately counting the UPOs [4]:

$$S = \lim_{k\to\infty} \frac{1}{k} \ln N_k = \ln n_{\mathrm{d}}\,. \tag{10}$$

$S > 0$ for all $n_{\mathrm{d}} \geqslant 2$. It is well known that if a map acting on an interval has a positive topological entropy, it is chaotic [5]. In a more general setting, if the topological entropy is positive, we can expect *some* form of chaotic behavior to show up in the dynamics.

We can see that the Newton method, when applied to a polynomial equation, leads to an extremely rich structure of multicycles. But where are they? If any of them is stable, as in our example in the Introduction, it can show up in a practical application of the Newtonian iterations. The unstable majority live on fractal boundaries between basins of attraction. It is the dynamics on these boundaries that is the chaotic behavior whose presence is indicated by a positive value of the topological entropy. It is also interesting to see that if a polynomial has multiple roots, it has less multicycles than a polynomial of the same order and with all roots distinct: It is the number of *distinct* roots, rather than the degree of the polynomial or the number of all the roots, that determines the value of the topological entropy.

3. The equation $x^2 + 1 = 0$

In general, multicycles generated by the Newton method live on complicated geometric objects somewhere on the complex plane. There is, however, one nontrivial example where the locale of the multicycles can be pinpointed quite accurately. Consider the equation

$$z^2 + 1 = 0 \,. \tag{11}$$

It generates the following Newtonian dynamics:

$$\eta(z) = \frac{1}{2}\left(z - \frac{1}{z}\right) . \tag{12}$$

It can be proved that if we start with any z_0 such that $\mathrm{Im}(z_0) > 0$, the Newtonian iterations converge to $z_\infty = +i$. Similarly, if we start with $\mathrm{Im}(z_0) < 0$, the Newtonian iterations converge to $z_\infty = -i$. Therefore, all the multicycles of the function Eq. (12) and all the points that eventually end up on them must lie on the real axis.

We shall, therefore, discuss properties of the map Eq. (12) restricted to the real axis:

$$\eta(x) = \frac{1}{2}\left(x - \frac{1}{x}\right), \quad x \in \mathbb{R} \,. \tag{13}$$

Formally speaking, this map results from an attempt to find the real zeros of the equation $x^2 + 1 = 0$. There are no such zeros, but that does not necessarily mean that properties of the map Eq. (13) are not interesting. Fig. 2 shows a typical trajectory generated by this map.

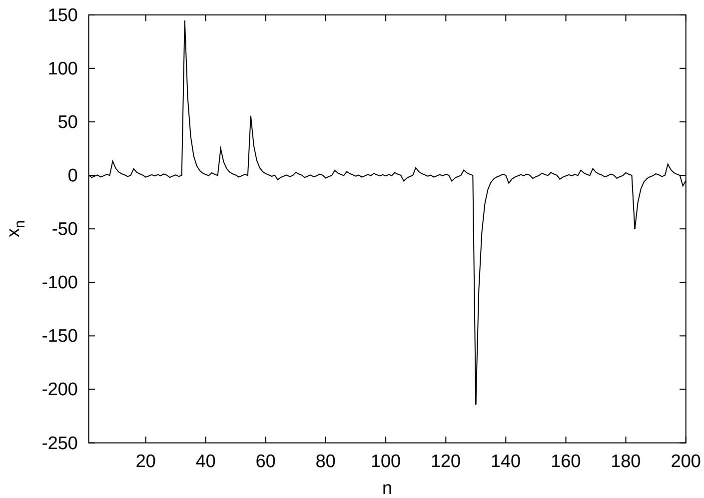

Fig. 2. A typical trajectory generated by the map Eq. (13). The trajectory spends most of its time in the vicinity of $x = 0$, but from time to time it makes large excursions, and then slowly relaxes towards zero. If one waits long enough, arbitrarily large excursions can be observed.

3.1. Properties of the map Eq. (13)

We shall list the most important properties of the map (Eq. 13):

- This function is singular in $x = 0$ and it increases monotonically for both $x < 0$ and $x > 0$. It does not have a fixed point.

- Each point has two pre-images. The pre-images of a point x satisfy

$$x^{\pm} = x \pm \sqrt{x^2 + 1}\,. \tag{14}$$

- There are countably many multicycles (UPOs). For each multicycle, there are countably many points that eventually fall onto it.

- $x = 0$ corresponds to the escape to infinity. There are countably many points that are eventually mapped into $x = 0$ and escape to infinity.

- The union of all the multicycles, points that lead to them, and points that eventually escape to infinity, is also countable and dense in $\mathbb{R}$. This leaves us with uncountably many points that neither escape, nor fall on a multicycle, but form a "truly chaotic" orbit of the dynamical system Eq. (13).

Simulations suggest that the map Eq. (13) has an invariant density that does not depend on the starting point, provided this point does not escape to infinity. This eliminates only countably many points, and in practice only three of them: $x \in \{-1, 0, 1\}$ (the pre-images of ± 1 are irrational and inaccessible in numerical simulations). Indeed, to compute the invariant density, denoted here by $\rho(x)$, we use the Frobenius–Perron equation [6]:

$$\rho(x) = \int_{-\infty}^{\infty} \delta\left(\eta(x) - y\right) \rho(y)\, dy\,. \tag{15}$$

Using Eq. (14), this equation takes the explicit form

$$\begin{aligned} \sqrt{x^2+1}\,\rho(x) \; &= \; \left(x + \sqrt{x^2+1}\right) \rho\left(x + \sqrt{x^2+1}\right) \\ &\quad - \left(x - \sqrt{x^2+1}\right) \rho\left(x - \sqrt{x^2+1}\right). \end{aligned} \tag{16}$$

It can be verified by a direct substitution that the Lorentzian distribution

$$\rho(x) = \frac{1}{\pi(1+x^2)}\,, \tag{17}$$

solves the Frobenius–Perron equation Eq. (16). Thus the map Eq. (13) joins the elite club of maps whose invariant densities are known explicitly. Slowly decaying tails of the invariant density Eq. (17) explain the large deviations from zero made by the chaotic trajectory.

3.2. The stereographic projection

It is interesting to see that the stereographic projection converts the function Eq. (13) into a piecewise linear map on the unit interval. If we substitute

$$x = \tan \pi\phi\,, \tag{18}$$

we obtain

$$\eta(x) = \cot 2\pi\phi = \tan \pi \left(2\phi + \frac{1}{2}\right). \tag{19}$$

Given the periodicity of the trigonometric function, we can see that the map Eq. (13) is equivalent to

$$\zeta(\phi) = 2\phi + \frac{1}{2} \bmod 1\,. \tag{20}$$

Properties of the map Eq. (20) are easy to find and they nicely illustrate the multicycles.

- The map Eq. (20) is chaotic with the Lyapunov exponent $\lambda = \ln 2$. This map has a flat invariant density, $\rho(\phi) = 1$, $\phi \in [0, 1]$.
- $\phi = 1/2$ is the fixed point of the map Eq. (20). The fixed point corresponds to an escape to infinity in the language of x given by Eq. (18).
- If we represent any $\phi \in [0, 1]$ by its binary expansion, the map Eq. (20) acts on it by *(i)* the binary shift, *(ii)* rejecting the highest bit, and *(iii)* flipping the remaining highest bit. For example,

$$0.1100101\ldots_2 \to 0.000101\ldots_2 \,, \tag{21}$$

  where the subscript indicates a binary expansion. Thus, all points with finite binary expansions eventually converge to $0.1_2 = 1/2$, *i.e.*, escape to infinity, points with periodic (starting form some point) binary expansions belong to multicycles or settle on one after a finite number of steps, and points with nonperiodic, infinite binary expansions lie on the chaotic trajectory. Thus the rational numbers either escape to infinity under the action of Eq. (20), or end up on multicycles. As the transformation Eq. (18) maps rational numbers from the unit interval to irrational numbers in $\boldsymbol{R}$, encountering a true multicycle of Eq. (13) during a computer simulation is impossible. However, the fact that computer calculations are performed with a finite precision, leads to the conclusion that in a computer experiment, the map Eq. (13) behaves as if it were periodic, with a fairly large period, almost regardless of the starting point. This is a well known fact about pseudo-random sequences.
- We can count the k-cycles by counting nontrivially different binary sequences of the length k. For example, there are 2^k possibilities of distributing $\{0, 1\}$ among k sites. Two constant sequences must be excluded, and as all cyclic permutations are equivalent, we arrive, if k is prime, on $(2^k - 2)/k$. (If k is not prime, the result of $(2^k - 2)/k$ is fractional.) This result, and its generalizations to non-prime k's, are the same as those given by Eq. (9) with $n_{\mathrm{d}} = 2$.

4. The equation $x^{2m} + 1 = 0$

The equation

$$z^{2m} + 1 = 0 \,, \quad z \in \mathbb{C} \,, \quad m \geqslant 2 \,, \tag{22}$$

is a natural generalization of the equation Eq. (11). However, the multicycles generated by the Newtonian iterations resulting from the equation Eq. (22) are not restricted to the real axis. Indeed, in case of Eq. (22) the Newton

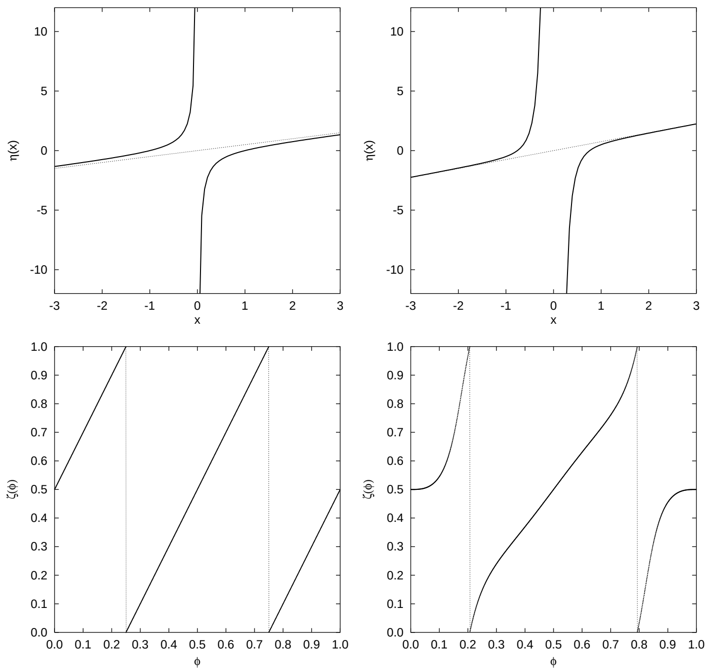

Fig. 3. The map Eq. (13) (top left) and the map Eq. (20) (bottom left), compared to the map Eq. (23) with $m = 2$ (top right) and the corresponding map on the unit interval, obtained numerically (bottom right).

method generates identical dynamics on each line $e^{i\pi l}$ on the complex plane, with $l = 0, 1, \dots\ m-1$, but there are also fractal boundaries between the basins of attraction that do not lie on these lines and, apparently, a majority of the multicycles are located there. If the Newtonian iterations are restricted to the real axis, they are the iterates of the function

$$\eta(x) = \left(1 - \frac{1}{2m}\right) x - \frac{1}{2m\, x^{2m-1}}\,, \quad x \in \mathbb{R}\,. \tag{23}$$

This function has *qualitatively* the same properties as the function Eq. (13): it is singular at $x = 0$, grows monotonically from $-\infty$ to ∞ in both domains $x < 0$ and $x > 0$, and each point has two pre-images. Therefore, the number of its multicycles (again, on the real axis) must be the same as that of Eq. (13). If we use the substitution Eq. (18), the function Eq. (23) does not convert to a piecewise linear map. Fig. 3 compares the maps Eq. (13) and Eq. (20) with the map Eq. (23) with $m = 2$ and the counterpart of $\zeta(\phi)$, obtained numerically from Eq. (23) with Eq. (18).

We now turn to the invariant density. In the asymptotic regime, $x \gg 1$, pre-images of a point x satisfy

$$\eta^{-1}(x) \;\simeq\; x^{+} = \frac{2m}{2m-1}x\,, \tag{24a}$$

$$\eta^{-1}(x) \;\simeq\; x^{-} = -\frac{1}{(2mx)^{1/(2m-1)}}\,. \tag{24b}$$

Moreover, if an invariant density exists and is continuous, it must satisfy

$$\rho(x^{-}) \simeq \rho(0)\,. \tag{24c}$$

Thus, if $x \gg 1$, the Frobenius–Perron equation Eq. (15) takes the approximate form

$$x^{\frac{2m}{2m-1}}\left[\rho(x) - \frac{2m}{2m-1}\,\rho\left(\frac{2mx}{2m-1}\right)\right] = \frac{\rho(0)}{(2m-1)(2m)^{1/(2m-1)}}\,. \tag{25}$$

Because of the symmetry of the function Eq. (23), the invariant density, if it exists, must satisfy $\rho(x) = \rho(-x)$. The right-hand side of Eq. (25) is constant. Thus, the left-hand side can be constant only if

$$\rho(x) \sim \mathrm{const}\cdot|x|^{-2m/(2m-1)}\,, \quad |x| \to \infty\,. \tag{26}$$

Fig. 4 shows the invariant density, found numerically, corresponding to the function Eq. (23) with $m = 2$. The tails agree perfectly with the theoretical prediction Eq. (26). Simulations show that for all $m \geqslant 2$, the invariant density is bimodal, and that the valley between the peaks gets deeper as m increases.

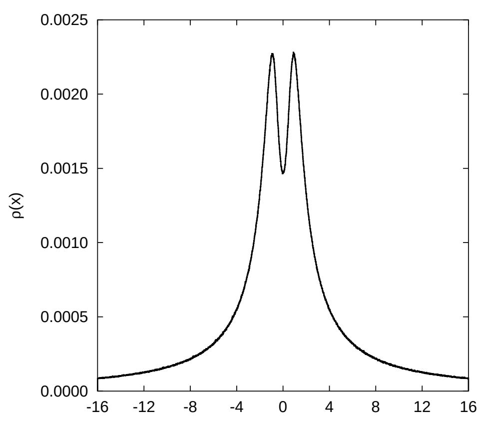

Fig. 4. The invariant density corresponding to the function Eq. (23) with $m = 2$, obtained numerically. The tails behave like $\sim |x|^{-4/3}$, in a full agreement with our theoretical predictions.

5. Conclusions

We have shown that the Newton method, when applied to a polynomial equation that has more than one distinct solution, generates infinitely many multicycles, and we have calculated the topological entropy associated with them. We have also shown that the map Eq. (13) is chaotic, equivalent to a piecewise linear map on the unit interval, and we have explicitly found its invariant density. This result is interesting by itself as the number of maps with analytically known invariant densities is very limited. Because the invariant density in question coincides with the Lorentzian probability distribution, one could use the function Eq. (13) as a basis of a pseudo-random generator of the latter.

From a practical point of view, hitting an unstable multicycle is virtually impossible, although it is well known that starting the Newtonian iterations in a close vicinity of a fractal boundary between the basins of attractions, where the unstable multicycles live, can slow down the convergence significantly. In some cases, stable multicycles can also appear and they can prevent the Newton method from converging to a root; the polynomial presented in Fig. 1 provides one such example. It would be interesting to see what are the criteria for the existence of stable multicycles. In practice, however, if a multicycle appears to form, the Newtonian iterations should be interrupted by a couple of steps taken with a different method, and with the damped Newton method in particular. This usually breaks the multicycle that has started to form, but can lead to new multicycles in their own right. Discussing this point goes beyond the scope of the present paper. We should finally mention that the Newton method is *not* the best method for a numerical search for zeros of a polynomial. There are other algorithms, better tailored to polynomials. From those the Laguerre method [1, 7], coupled with the deflation of the polynomial, is now regarded as the method of choice. Despite many arguments in favor of the latter method, the Newton method is probably most commonly used for the task, and this is why a study of its UPOs structure is, in our opinion, important.

To end on a lighter note, we have shown that a numerical search for zeros of a polynomial with the Newton method can be quite chaotic, in particular when the zeros are just not there.

Ł.S. thanks Mr Wojciech Brzezicki for a helpful discussion. This work was supported in part by the Marie Curie Actions Transfer of Knowledge project COCOS (contract MTKD-CT-2004-517186).

Appendix A

Proof. We give a proof of Theorem 1 from Section 2 in this Appendix. First observe that if the polynomial $P(z)$ has multiple roots, there exists a polynomial $Q(z)$ such that $P(z) = Q(z)R(z)$ and $P'(z) = Q(z)S(z)$, where $\deg R(z) = n_{\mathrm{d}}$, $\deg S(z) = n_{\mathrm{d}} - 1$, n_{d} is the number of *distinct* roots of the polynomial $P(z)$ and the fraction $R(z)/S(z)$ cannot be further canceled. (If all roots of $P(z)$ are distinct, $n_{\mathrm{d}} = n$ and $Q(z) \equiv 1$.) Therefore, (*cf.* Eq. (6))

$$\eta^k(z) = z - \frac{R(z)}{S(z)}\,\frac{A_k(z)}{B_k(z)}\,. \tag{A.1}$$

To prove the theorem, we need to show that

I. The polynomials $A_k(z)$, $B_k(z)$, generated according to Eq. (5), have degrees $n_{\mathrm{d}}^k - n_{\mathrm{d}}$, and

II. The polynomials $R(z)$, $S(z)$, $A_k(z)$, $B_k(z)$ do not have common roots, from which it follows that the fraction in Eq. (A.1) cannot be further canceled.

Both parts of the proof proceed by induction. Note that $A_1(z) \equiv B_1(z) \equiv 1$, so the theorem holds for $k=1$. For the sake of the notation, let us assume that

$$R(z) = \sum_{j=0}^{n_{\mathrm{d}}} r_j z^j\,. \tag{A.2}$$

Part I

Suppose the theorem holds for some $k \geqslant 1$ and calculate

$$\begin{aligned}
R\big(z^{(k)}\big) &= R\left(z - \frac{R(z)}{S(z)}\,\frac{A_k(z)}{B_k(z)}\right)\\
&= r_0 + \sum_{s=1}^{n_{\mathrm{d}}} r_s \left(z - \frac{R(z)}{S(z)}\,\frac{A_k(z)}{B_k(z)}\right)^s\\
&= r_0 + \sum_{s=1}^{n_{\mathrm{d}}} r_s \sum_{l=0}^{s} \binom{s}{l} z^{s-l}(-1)^l \left(\frac{R(z)}{S(z)}\,\frac{A_k(z)}{B_k(z)}\right)^l\\
&= \sum_{s=0}^{n_{\mathrm{d}}} r_s z^s + \sum_{s=1}^{n_{\mathrm{d}}} r_s \sum_{l=1}^{s} \binom{s}{l} z^{s-l}(-1)^l \left(\frac{A_k(z)}{S(z)B_k(z)}\right)^l (R(z))^l\\
&= R(z)\left[1 + \sum_{s=1}^{n_{\mathrm{d}}} r_s \sum_{l=1}^{s} \binom{s}{l} z^{s-l}(-1)^l \left(\frac{A_k(z)}{S(z)B_k(z)}\right)^l (R(z))^{l-1}\right]
\end{aligned}$$

(note that $l - 1 \geqslant 0$)

$$= \frac{R(z)}{(S(z)B_k(z))^{n_\mathrm{d}}}\left[(S(z)B_k(z))^{n_\mathrm{d}} + \sum_{s=1}^{n_\mathrm{d}} r_s (S(z)B_k(z))^{n_\mathrm{d}-s}\right.$$

$$\left.\times \sum_{l=1}^{s}\binom{s}{l}(-1)^l (zS(z)B_k(z))^{s-l}(A_k(z))^l (R(z))^{l-1}\right]. \qquad (A.3)$$

Similarly

$$S\left(z^{(k)}\right) = S\left(z - \frac{R(z)}{S(z)}\frac{A_k(z)}{B_k(z)}\right) = \frac{1}{(S(z)B_k(z))^{n_\mathrm{d}-1}}$$

$$\times \sum_{s=0}^{n_\mathrm{d}-1}(s+1)r_{s+1}(SB_k)^{n_\mathrm{d}-1-s}(z\,SB_k - RA_k)^s\,. \qquad (A.4)$$

Therefore

$$\frac{R\left(z^{(k)}\right)}{S\left(z^{(k)}\right)} = \frac{R(z)}{S(z)}\,\frac{\mathcal{N}(z)}{B_k(z)\mathcal{D}(z)}\,, \qquad (A.5)$$

where $\mathcal{N}(z)$, $\mathcal{D}(z)$ are certain polynomials whose forms are implied in Eqs. (A.3) and Eq. (A.4). What are the orders of these polynomials? $\deg S = n_\mathrm{d} - 1$ and, by assumption, $\deg A_k = \deg B_k = n_\mathrm{d}^k - n_\mathrm{d}$. We thus have

$$\begin{aligned}
\deg\,(SB_k)^{n_\mathrm{d}-s} &= (n_\mathrm{d}-s)(n_\mathrm{d}-1+n_\mathrm{d}^k-n_\mathrm{d}) \\
&= n_\mathrm{d}^{k+1} - n_\mathrm{d}^k s - n_\mathrm{d} + s\,, && (A.6)\\
\deg\,(z\,SB_k)^{s-l}(A_k)^l R^{l-1} &= (s-l)(1+n_\mathrm{d}-1+n_\mathrm{d}^k-n_\mathrm{d}) + l(n_\mathrm{d}^k-n_\mathrm{d}) + (l-1)n_\mathrm{d} \\
&= n_\mathrm{d}^k s - n_\mathrm{d}\,. && (A.7)
\end{aligned}$$

Thus the term under the sum over s in Eq. (A.3) is of the order $n_\mathrm{d}^{k+1} - n_\mathrm{d}^k s - n_\mathrm{d} + s + n_\mathrm{d}^k s = n_\mathrm{d}^{k+1} - 2n + s$, which takes a maximal value of $n_\mathrm{d}^{k+1} - n_\mathrm{d}$ for $s = n_\mathrm{d}$. Therefore

$$\deg \mathcal{N}(z) = n_\mathrm{d}^{k+1} - n_\mathrm{d}\,. \qquad (A.8)$$

Similarly, because $\deg z\,S = \deg R = n_\mathrm{d}$,

$$\begin{aligned}
\deg\,(SB_k)^{n_\mathrm{d}-s-1}(z\,SB_k - RA_k)^s &= (n_\mathrm{d}-s-1)(n_\mathrm{d}-1+n_\mathrm{d}^k-n_\mathrm{d}) \\
&\quad + s(n_\mathrm{d}+n_\mathrm{d}^k-n_\mathrm{d}) \\
&= n_\mathrm{d}^{k+1} - n_\mathrm{d}^k - n_\mathrm{d} + s + 1\,, && (A.9)
\end{aligned}$$

which takes a maximal value of $n_\mathrm{d}^{k+1} - n_\mathrm{d}^k$ for $s = n_\mathrm{d} - 1$. Therefore

$$\deg \mathcal{D}(z) = n_\mathrm{d}^{k+1} - n_\mathrm{d}^k\,. \qquad (A.10)$$

We finally have

$$\begin{aligned}\eta^{k+1}(z) &= \eta^k(z) - \frac{R\left(\eta^k(z)\right)}{S\left(\eta^k(z)\right)} = z - \frac{R(z)}{S(z)}\,\frac{A_k(z)}{B_k(z)} - \frac{R(z)}{S(z)}\,\frac{\mathcal{N}(z)}{B_k(z)\mathcal{D}(z)} \\ &= z - \frac{R(z)}{S(z)}\,\frac{\mathcal{D}(z)A_k(z)+\mathcal{N}(z)}{B_k(z)\mathcal{D}(z)} = z - \frac{R(z)}{S(z)}\,\frac{A_{k+1}(z)}{B_{k+1}(z)}\,. \qquad \text{(A.11)}\end{aligned}$$

The last statement defines polynomials $A_{k+1}(z)$, $B_{k+1}(z)$. Using Eq. (A.8) and Eq. (A.10), it is now easy to verify that $\deg A_{k+1} = \deg B_{k+1} = n_{\mathrm{d}}^{k+1} - n_{\mathrm{d}}$, provided the last fraction in Eq. (A.11) cannot be canceled.

By direct calculations, it is easy to show that

$$\begin{aligned}A_{k+1}(z) &= A_k(z)B_k^{n_{\mathrm{d}}-1}(z)S^{n_{\mathrm{d}}-1}(z)\,S\left(z - \frac{R(z)}{S(z)}\,\frac{A_k(z)}{B_k(z)}\right) \\ &\quad + B_k^{n_{\mathrm{d}}}(z)S^{n_{\mathrm{d}}}(z)R^{n_{\mathrm{d}}-1}(z)\,R\left(z - \frac{R(z)}{S(z)}\,\frac{A_k(z)}{B_k(z)}\right), \qquad \text{(A.12a)} \\ B_{k+1}(z) &= B_k^{n_{\mathrm{d}}}(z)S^{n_{\mathrm{d}}-1}(z)\,S\left(z - \frac{R(z)}{S(z)}\,\frac{A_k(z)}{B_k(z)}\right). \qquad \text{(A.12b)}\end{aligned}$$

Note that we have already shown that $A_{k+1}(z)$, $B_{k+1}(z)$ *are* polynomials.

Part II

To complete the proof, we still need to show that the polynomials $R(z)$, $S(z)$, $A_k(z)$, $B_k(z)$ do not have common roots. The polynomials $R(z)$, $S(z)$ do not have common roots by construction. Suppose the polynomials $R(z)$, $S(z)$, $A_k(z)$, $B_k(z)$ do not have common roots for some $k \geqslant 1$.

$R(z) = 0$

Let $R(z) = 0$. Then $S(z) \neq 0$ and $B_k(z) \neq 0$, and because of Eq. (A.12b), $B_{k+1}(z) = (B_k(z)S(z))^{n_{\mathrm{d}}} \neq 0$. Furthermore, if $R(z) = 0$, then, by virtue of Eq. (A.12a), $A_{k+1}(z) = A_k(z)(B_k(z)S(z))^{n_{\mathrm{d}}-1}S(z) \neq 0$, because $A_k(z) \neq 0$, either. In other words, $R(z)$ does not have common roots with the other polynomials.

$S(z) = 0$

If $S(z) = 0$, then z is a zero of the derivative of the original polynomial $P(z)$ that is not a multiple root of the latter. In this case the Newton iterations diverge and we cannot use the expressions Eq. (A.12). We need to use the definitions implied in Eq. (A.11) instead; they are still valid by the argument of continuity. We have $B_{k+1}(z) = B_k(z)\mathcal{D}(z)$, and if $S(z) = 0$, this reduces to $B_{k+1}(z) = s_{n_\mathrm{d}-1}(R(z)A_k(z))^{n_\mathrm{d}-1} \neq 0$, where $s_{n_\mathrm{d}-1}$ is the highest-order coefficient in the polynomial $S(z)$, *cf.* Eq. (A.4) above. For the remaining polynomial we have $A_{k+1}(z) = A_k(z)\mathcal{D}(z) + \mathcal{N}(z)$ which, for $S(z) = 0$, reduces to $A_{k+1}(z) = (-1)^{n_\mathrm{d}} R^{n_\mathrm{d}-1}(z) A_k^{n_\mathrm{d}}(z)(r_{n_\mathrm{d}} - s_{n_\mathrm{d}-1})$, where r_{n_d} is the highest coefficient in $R(z)$. Thus if $S(z) = 0$, $A_{k+1}(z)$ could vanish only if $r_{n_\mathrm{d}} = s_{n_\mathrm{d}-1}$, but this is impossible, given the fact that $R(z)Q(z) = P(z)$ and $S(z)Q(z) = P'(z)$ for a certain polynomial $Q(z)$. In other words, $S(z)$ does not have common roots with the other polynomials.

$A_{k+1}(z)$ and $B_{k+1}(z)$ do not have common roots

It remains to be shown that $A_{k+1}(z)$ and $B_{k+1}(z)$ do not have common roots. Let us assume that $B_{k+1}(z) = 0$. Because $B_{k+1}(z) = B_k(z)\mathcal{D}(z)$, all roots of $B_k(z)$ are also roots of $B_{k+1}(z)$. However, if $B_k(z) = 0$, we can directly repeat the argument from the preceding paragraph to show that in this case $A_{k+1}(z) = (-1)^{n_\mathrm{d}} R^{n_\mathrm{d}-1}(z) A_k^{n_\mathrm{d}}(z)(r_{n_\mathrm{d}} - s_{n_\mathrm{d}-1}) \neq 0$.

Because we have already shown that $S(z)$ does not have common roots with $B_{k+1}(z)$, $B_k(z) \neq 0$ and $S\left(z - \frac{R(z)}{S(z)}\frac{A_k(z)}{B_k(z)}\right) = 0$ is the only remaining possibility for $B_{k+1}(z)$ to vanish, *cf.* Eq. (A.12b). If this is the case, then form Eq. (A.12a) we have

$$A_{k+1}(z) = B_k^{n_\mathrm{d}}(z) S^{n_\mathrm{d}}(z) R^{n_\mathrm{d}-1}(z)\, R\left(z - \frac{R(z)}{S(z)}\frac{A_k(z)}{B_k(z)}\right). \qquad \text{(A.13)}$$

By assumption, $B_k(z) \neq 0$, and because we have already shown that neither $R(z)$, nor $S(z)$ can have common roots with $B_{k+1}(z)$, $R(z) \neq 0$ and $S(z) \neq 0$. Finally, because, by construction, the polynomials $R(z)$ and $S(z)$ do not have common roots, if $S\left(z - \frac{R(z)}{S(z)}\frac{A_k(z)}{B_k(z)}\right) = 0$, it follows that $R\left(z - \frac{R(z)}{S(z)}\frac{A_k(z)}{B_k(z)}\right) \neq 0$, and therefore, $A_{k+1}(z) \neq 0$.

We have thus shown that $A_{k+1}(z)$ and $B_{k+1}(z)$ do not have common roots. This completes the proof. □

REFERENCES

[1] W.H. Press, B.R. Flannery, S.A. Teukolsky, W.T. Vetterling, *Numerical Recipes in Fortran. The Art of Scientific Computing*, 2nd ed. Cambridge University Press, 1993.

[2] D. Auerbach, R. Cvitanovic, J-R. Eckmann, G. Gunarante, *Rhys. Rev. Lett.* **23**, 2387 (1987); R. Cvitanovic, *Phys. Rev. Lett.* **24**, 2729 (1988); R. Cvitanovic, *Physica D* **83**, 109 (1995).

[3] R. Walters, *An Introduction to Ergodic Theory*, Springer, 1992; E. Ott, *Chaos in Dynamical Systems*, 2nd ed. Cambridge University Press, 2003.

[4] C. Greborgi, E. Ott, J.A. Yorke, *Phys. Rev.* **A37**, 1711 (1988).

[5] N. Franzowa, J. Smitel, *Proc. Am. Math. Soc.* **112**, 1083 (1991).

[6] C. Beck, F. Schlögl, *Thermodynamics of Chaotic Systems*, Cambridge University Press, 1993.

[7] A. Ralston, *A First Course in Numerical Analysis*, McGraw-Hill, 1965.