

# 18.100C: Fixed point algorithms for computing n-th roots

Tony Kim

November 4, 2007

## 1 Introduction

In this assignment I demonstrate some simple theorems regarding the fixed points of continuous, real-valued functions on  $\mathfrak{R}$ . We then consider numerical algorithms of computing n-th roots as an application of the results.

We examine the iterates of the function  $f$  defined by

$$f(t) = \frac{1}{2} \cdot \left(t + \frac{\alpha}{t}\right)$$

as a means to calculate the square root of  $\alpha > 1$ .<sup>1</sup> After examining the validity of this procedure, we attempt to generalize  $f$  to find n-th roots. However, we conclude that for  $n \geq 4$ , the theorems derived in this paper cannot be applied, and convergence cannot be guaranteed.

### 1.1 Definitions

Let  $f$  be a continuous real function on  $(-\infty, \infty)$ . We call  $x$  a *fixed point* of  $f$  if  $f(x) = x$ . Furthermore, given some initial number  $x_0$ , we define the sequence  $\{x_i\}$  by the relation  $x_{i+1} = f(x_i)$ . The n-th term of this sequence is then called the n-th *iterate* of  $f$ .

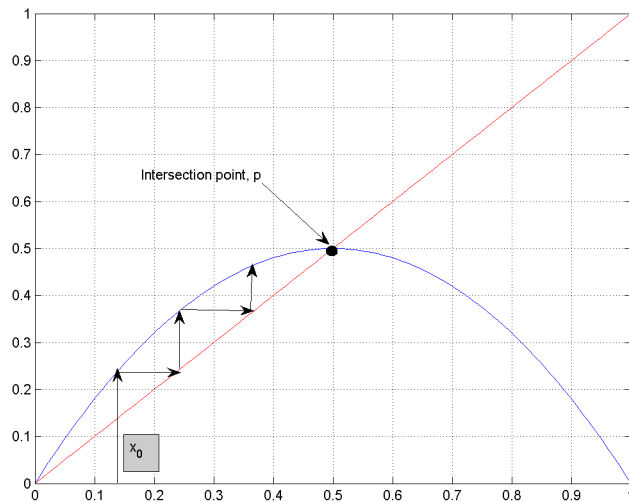
### 1.2 An intuitive picture of iteration

We deal with real-valued functions on  $\mathfrak{R}$ . Hence, it is useful to consider the graph of  $f$  in order to visualize the function's fixed points and iterates. In particular, when superimposed with the identity map, many of the results that we discuss in an analytic setting are quite evident even at a glance.

For example, consider the "logistic map"  $f(t) = 2t(1 - t)$  over the interval  $[0, 1]$ . (See Figure (1.2).) Since the fixed point is an  $x$  such that  $x = f(x)$ , it is clear that the abscissa of the intersection point  $p$  is a fixed point of  $f$ .

---

<sup>1</sup>It can easily be shown that the function  $f$  presented here is a special case of the Newton-Raphson (NR) method of computing roots of functions. In discussing this topic, I desperately hoped to *not* mention the NR algorithm since its application (to real functions, at least) is a topic that everyone is familiar with. But it just seems that I can't get away from this guy!



This construction is also very helpful to visualize the “dynamics” of the iterates. Referring to the figure, choose some seed point  $x_0$  in the domain. Move vertically to obtain an intersection with the graph of  $f$ . Then, move horizontally to intersect the identity map. From this point, move vertically to  $f(t)$  and so on. It is clear that the coordinates of the successive intersections with  $f$  give the sequence  $\{x_i\}$ .

Some facts are readily apparent from this graphical perspective:

1. Let  $a, b$  be two distinct elements from an interval in the domain of  $f$ . Let  $a < b$  and assume that they are not fixed points. If  $f(a) > a$  and  $f(b) < b$ , then there is a fixed point  $c \in (a, b)$ . The same conclusion holds if  $f(a) < a$  and  $f(b) > b$ .
2. Evidently, certain fixed points are attractive while others are not. (Try the graphical iteration near the origin and near  $p$  in the above figure!) In this paper, we will consider the conditions that *guarantee*  $x_n \rightarrow x$  where  $x$  is a fixed point of  $f$ .<sup>2</sup>

We investigate these notions more precisely in the following section.

## 2 Fixed point theorems

We take  $f$  to be a real, continuous function defined on  $(-\infty, \infty)$ .

---

<sup>2</sup>Incidentally, the logistic map is a well-known example of a function whose iterates are known to exhibit very complicated dynamics, such as the phenomena of period doubling and chaos. In this paper, I keep things modest and discuss only the conditions that guarantee  $x_n \rightarrow x$ , although it may not be the most general condition

**2.1** Suppose  $a, b \in I$  where  $I$  is an interval in the domain of  $f$ , and  $a < b$ . If  $f(a) > a$  and  $f(b) < b$ , then there is a point  $c \in (a, b)$  that is a fixed point of  $f$ . The conclusion also holds when  $f(a) < a$  and  $f(b) > b$ .

**Proof.** We will demonstrate the theorem for the case  $f(a) > a$  and  $f(b) < b$ . The other case proceeds similarly.

Consider the function  $g(x) = f(x) - x$ . By the premise, we have:

$$\begin{aligned}g(a) &= f(a) - a > 0 \\g(b) &= f(b) - b < 0\end{aligned}$$

Furthermore,  $g$  is continuous since  $f$  and the identity map are continuous. So the intermediate value theorem applies and there is a  $c \in (a, b)$  such that  $g(c) = 0$ . Or, in other words:

$$\begin{aligned}g(c) &= f(c) - c = 0 \\f(c) &= c\end{aligned}$$

So  $c$  is a fixed point.

For the remainder of this section, assume that  $f$  is differentiable.

**2.2** Suppose  $f'(t) \neq 1$  for every real  $t$ . Then  $f$  has at most one fixed point.

**Proof.** Suppose not. We assume that  $f$  has more than one fixed point. Choose any two fixed points; call them  $x$  and  $y$  such that  $x < y$ . The function  $f$  satisfies the requirements of the mean value theorem on  $[x, y]$ , so there is a point  $c \in (x, y)$  such that:

$$f'(c) = \frac{f(y) - f(x)}{y - x} = \frac{y - x}{y - x} = 1 \tag{1}$$

contradicting the property that  $f'(t) \neq 1, \forall t$ .

This result is easily generalized to the statement that a function  $f$  can have at most  $(n + 1)$  fixed points if it possesses  $n$  distinct points at which  $f'(t) = 1$ .

It is also clear that we can restrict the domain of  $f$  to an interval  $I \subset \mathbb{R}$ . If  $f'(t) \neq 1$  for all  $t \in I$ , then there can be at most one fixed point of  $f$  in  $I$ .

**2.3** If there is a positive constant  $A < 1$  such that  $|f'(t)| \leq A$  for all real  $t$ , then a fixed point  $x$  of  $f$  exists.

**Proof.** By Theorem 2.1, if  $f$  does not have a fixed point, it must be the case that  $f(t) > t$  for every  $t$  or  $f(t) < t$  for every  $t$ . We show that neither of these scenarios are consistent with the requirement that  $|f'(t)| \leq A < 1$  for all  $t$ .

We focus our attention on the case  $f(t) > t, \forall t$ . The alternate case is proven analogously.

For any  $x > 0$ , we have by the mean value theorem,

$$\begin{aligned} f'(c) &= \frac{f(x) - f(0)}{x - 0} = \frac{f(x) - f(0)}{x} > \frac{x - f(0)}{x} \\ &> 1 - \frac{f(0)}{x} \end{aligned}$$

for some  $c \in (0, x)$ . Note that  $f(0)$  is, by assumption, some nonnegative fixed value.

Since  $1 - A > 0$ , we can always find an  $x$  such that:

$$x > \frac{f(0)}{1 - A} \tag{2}$$

so that for some  $c \in (0, x)$

$$f'(c) > 1 - \frac{f(0)}{x} > A \tag{3}$$

Hence, if  $f(t) > t$  for every  $t$ , then there cannot be a bound  $A$  for  $f'$  such that  $A < 1$ . We get a similar contradiction if we assume that  $f(t) < t$  for every  $t$ .

It should be clear that this result remains valid if  $f$  is defined over some other connected, *unbounded* subset of  $\mathfrak{R}$ . For example,  $(0, \infty)$  is just as valid.

### 2.3.1 Example

To demonstrate the requirement that  $A$  needs to be strictly less than 1 in order to guarantee a fixed point by the previous theorem, we consider the function:

$$f(t) = t + \frac{1}{1 + e^t} \tag{4}$$

Its derivative is:

$$f'(t) = 1 - \frac{e^t}{(1 + e^t)^2} \tag{5}$$

Since  $e^t$  is positive for  $t \in \mathfrak{R}$ , we easily see that  $|f'(t)| < 1$  for all  $t$ . However, by the same property of the exponential function, it is also clear that  $f$  cannot have any fixed points since the equation

$$\begin{aligned} t &= f(t) \\ &= t + \frac{1}{1 + e^t} \\ 0 &= \frac{1}{1 + e^t} \end{aligned}$$

cannot have a valid solution for finite  $t$ . Hence, we see explicitly that the bound must be strictly less than 1 for the conclusion of Theorem 2.3 to hold.

**2.4 Let  $I$  be an interval in  $\mathfrak{R}$ , such that  $f(I) \subset I$ . Let  $0 < A < 1$  be such that  $|f'(t)| \leq A$  for all  $t \in E$ . Suppose there is a fixed point  $x$  in  $I$ . For any arbitrary real number  $x_0 \in I$ , define the sequence  $\{x_n\}$  by  $x_{n+1} = f(x_n)$ . Then  $\lim x_n = x$**

**Proof.** By Theorem 2.2 we know that the fixed point  $x$  is unique in  $E$ .

Put  $x_2 = f(x_1)$ . If  $x_2 = x_1$ , then  $x_2 = x$  and hence  $x_n = x$  for  $n \geq 2$ . We are done.

Suppose instead  $x_2 \neq x_1$ . Define  $D = d(x_2, x_1)$ .

By the mean value theorem, there is a  $c \in (x_1, x_2)$  (or  $(x_2, x_1)$  depending on the order of  $x_1$  and  $x_2$ ). This is immaterial to the proof) such that:

$$\frac{f(x_2) - f(x_1)}{x_2 - x_1} = f'(c) \leq A \quad (6)$$

so that:

$$\begin{aligned} f(x_2) - f(x_1) &\leq A \cdot (x_2 - x_1) \\ |x_3 - x_2| &\leq AD \end{aligned}$$

By the property that  $f(I) \subset I$ , we can continue to apply this to the  $n$ -th iterate, which yields:  $|x_{n+2} - x_{n+1}| \leq A|x_{n+1} - x_n|$ . This in turn implies:

$$|x_{n+1} - x_n| \leq A^{n-1} \cdot D \quad (7)$$

This claim is readily verified by induction. The  $n = 2$  case has already been demonstrated. For the  $(n+1)$ -th result, we have:

$$\begin{aligned} |x_{n+2} - x_{n+1}| &\leq A|x_{n+1} - x_n| \\ &\leq AA^{n-1}D \\ &\leq A^n D \end{aligned}$$

So the result is proved.

From here we show that  $\{x_n\}$  is Cauchy. Pick some integer  $N$ . For any  $m \geq n \geq N$ , the distance between the  $n$ -th and  $m$ -th iterate is:

$$\begin{aligned} |x_m - x_n| &= |x_m - x_{m-1} + x_{m-1} - \dots - x_{n+1} + x_{n+1} - x_n| \\ &\leq |x_m - x_{m-1}| + |x_{m-1} - x_{m-2}| + \dots + |x_{n+1} - x_n| \\ &\leq A^{m-2}D + A^{m-3}D + \dots + A^{n-1}D \\ &\leq A^{n-1}D \cdot (A^{m-n-1} + A^{m-n-2} + \dots + 1) \\ &\leq A^{n-1}D \cdot \frac{1 - A^{m-n}}{1 - A} < \frac{DA^n}{1 - A} \leq \frac{DA^N}{1 - A} \\ |x_m - x_n| &< \frac{DA^N}{1 - A} \end{aligned}$$

So, for any  $\epsilon > 0$ , we pick an  $N$  such that  $\frac{DA^N}{1-A} < \epsilon$ . This is always possible since  $0 < A < 1$ .

Then, for any  $n, m \geq N$ ,  $d(x_m, x_n) < \frac{DA^N}{1-A} < \epsilon$ . Hence,  $\{x_n\}$  is Cauchy. In addition, since  $x_n$  is a sequence in  $\mathfrak{R}$ , we conclude that it must converge to some number  $x^*$ .

Finally, we wish to show that if  $x_n \rightarrow x^*$ , then that  $x^* = x$ .

By the continuity of  $f$ :

$$f(x^*) = f(\lim x_n) = \lim f(x_n) = \lim x_{n+1} = x^* \quad (8)$$

so we find that  $x^*$  is a fixed point of  $f$ . Since  $x$  is unique, we must have  $x^* = x$  as desired.

Note, in particular, that the limit of the sequence of iterates, if it exists, must be a fixed point of the function.

### 3 Applications

In this section we give examples of using the theorems we have derived. The first is a rather unmotivated, purely mathematical problem of iterations on a particular  $f$ .

#### 3.1 Example 1

The function  $f(t) = \frac{t^3+1}{3}$  has three fixed points, labeled  $\alpha, \beta, \gamma$  such that  $-2 < \alpha < -1, 0 < \beta < 1, 1 < \gamma < 2$ . Using the theorems, we will show that:

1. If  $x_1 < \alpha$ , then  $x_n \rightarrow -\infty$  as  $n \rightarrow \infty$ .
2. If  $\alpha < x_1 < \gamma$ , then  $x_n \rightarrow \beta$  as  $n \rightarrow \infty$ .
3. If  $\gamma < x_1$ , then  $x_n \rightarrow +\infty$  as  $n \rightarrow \infty$ .

#### Proof.

First we establish the existence of the three fixed points. Evaluating  $f$  at various points gives:

$x$	$f(x)$	Order
-2	$\frac{-7}{3}$	>
-1	0	<
0	$\frac{1}{3}$	<
1	$\frac{2}{3}$	>
2	$\frac{3}{3}$	<

We see that the relative ordering of  $x$  and  $f(x)$  swaps during the intervals  $[-2, -1]$ ,  $[0, 1]$  and  $[1, 2]$ . Then, by Theorem 2.1, there exist three fixed points  $\alpha, \beta, \gamma$  in the respective intervals. Furthermore, it is clear that the equation  $1 = f'(x) = x^2$  has only two solutions. Hence, Theorem 2.2 shows that  $f$  has at most three fixed points.

We wish to show that for  $x_1 < \alpha$ ,  $x_n \rightarrow -\infty$ . By Theorem 2.1, we require that  $f(t) < t$  for every  $t \in (-\infty, \alpha)$ . This is necessary since otherwise we must have a fixed point in  $(-\infty, \alpha)$  in addition to the three accounted for above. By definition, this inequality shows that  $x_{n+1} < x_n, \forall x_n \in (-\infty, \alpha)$ . So,  $\{x_n\}$  is a monotonically decreasing sequence.

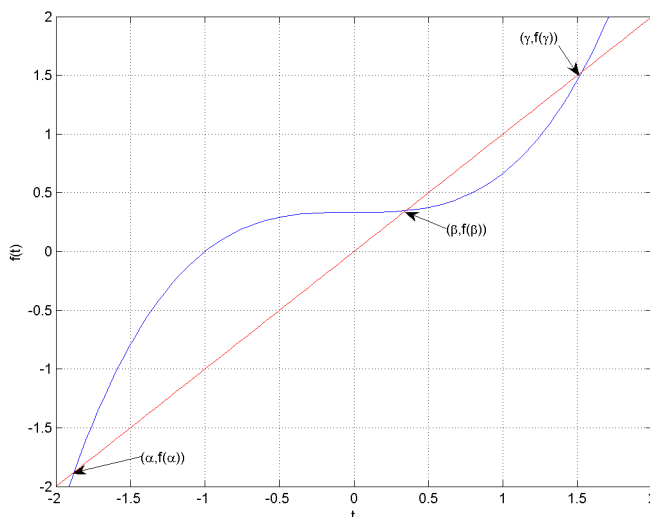
Suppose that  $\{x_n\}$  is bounded. Since we know the sequence is monotonic, it must then converge. However, Theorem 2.4 shows that the limit of the iterates, if it exists, must be a fixed point of the function. Of course,

this cannot be, because no fixed points of  $f$  live in the interval  $(-\infty, \alpha)$ . Instead, we have  $x_n \rightarrow -\infty$  as  $n \rightarrow \infty$ .

The proof that  $x_1 > \gamma$  implies  $x_n \rightarrow \infty$  is identical. This is all very obvious when one considers the graph of  $f$  and the iterations using the graphical method described in Section 1.

However, of some interest is the proof that if  $\alpha < x_1 < \gamma$ , then  $x_n \rightarrow \beta$ . In particular, it is instructive in illustrating how to apply the convergence criterion (Theorem 2.4).

We wish to identify an interval  $I$  containing  $\beta$  such that  $f(I) \subset I$  and for which  $f'(t) \leq A < 1$  for all  $t \in I$ . Since  $f'(t) = t^2$ , it is obvious that the latter condition forces us to look for  $I$  in the segment  $(-1, 1)$ . From Figure (3.1), it is evident that  $I = [-.5, .5]$  satisfies these conditions.



So, Theorem 2.4 guarantees that any sequence generated by  $x_0 \in I$  will converge to the fixed point  $\beta$ . It remains to be shown that any point in  $(\alpha, \gamma)$  will eventually evolve into  $I$ . This is also easy to demonstrate.

Suppose  $x_0 \in (\alpha, \beta)$ . Then, as before, we require that  $x_{n+1} > x_n$ , so the sequence is monotonically increasing. If this sequence never enters  $I$ , then it is bound above by  $-0.5$  and must converge somewhere within  $(\alpha, -0.5)$ . But this is impossible since there are no fixed points within that segment. We argue similarly for initial conditions in  $(\beta, \gamma)$ .<sup>3</sup>

This simple example illustrates how our theorems may be applied to demonstrate the necessary convergence of iterates. The strategy is:

1. Identify an interval  $I$  containing the fixed point that satisfies the premises of Theorem 2.4.
2. Identify a larger domain  $E$  for from which any sequence began by  $x_0 \in E$  eventually evolves into  $I$ .

---

<sup>3</sup>As of this writing, I recognize that the above argument does *not* treat the possibility of oscillations between  $(\alpha, -0.5)$  and  $(0.5, \gamma)$ . For the case considered here, this is not possible since  $f$  is monotonic, indicating that an iterate of some element in  $(\alpha, \beta)$  cannot land in  $(\beta, \gamma)$ . I'll slide this under the table for the time being, but I would like to make this point cleaner in a future draft.

### 3.2 Example 2: Algorithm for square roots

With this background, we now discuss the function  $f$  for computing square roots. In particular, we demonstrate, using our theoretical results, that the procedure is *correct*.

The function  $f(t) = \frac{1}{2} \cdot (t + \frac{\alpha}{t})$  with  $\alpha > 1$  is proposed as a function whose iterates can be used to compute the square root of  $\alpha$ .

To verify this within the framework of fixed point analysis, we first verify that  $\alpha$  is a fixed point of  $f$ :

$$\begin{aligned} f(\sqrt{\alpha}) &= \frac{1}{2} \cdot (\sqrt{\alpha} + \frac{\alpha}{\sqrt{\alpha}}) \\ &= \frac{1}{2} \cdot (\sqrt{\alpha} + \sqrt{\alpha}) \\ &= \sqrt{\alpha} \end{aligned}$$

so indeed the algorithm can be treated as a fixed point iteration problem.

The derivative of  $f$  is:

$$f'(t) = \frac{1}{2} \cdot (1 - \frac{\alpha}{t^2})$$

Conveniently,  $f'$  is bounded above by  $1/2$ ; and  $f'(\sqrt{\alpha}) = 0$ . With a little more work, we find that  $f(\sqrt{\alpha}) = \sqrt{\alpha}$  is the minimum of  $f$ . This implies by Theorem 2.2 that there can be only one fixed point. Since we have identified  $x = \sqrt{\alpha}$  as a fixed point, there are no others to consider.

We proceed by identifying an  $I \subset (0, \infty)$  such that  $|f'(t)| \leq A < 1$  for all  $t \in I$ . Since we have seen that  $f'$  is bounded above by  $1/2$ , we simply need to establish a lower bound. The relevant inequality is:

$$\begin{aligned} -1 &< f'(t) \\ &< \frac{1}{2} \cdot (1 - \frac{\alpha}{t^2}) \\ -2 &< 1 - \frac{\alpha}{t^2} \\ \frac{\alpha}{t^2} &< 3 \\ \sqrt{\frac{\alpha}{3}} &< t \end{aligned}$$

So, we find that in the segment  $(\sqrt{\frac{\alpha}{3}}, \infty)$ ,  $f' < 1$ . But since we have seen that the bound for  $f'$  must be strictly less than one, we pick a more restricted subset that still contains the fixed point. Arbitrarily, we can choose  $I = [\sqrt{\frac{\alpha}{2}}, \infty)$ .

It was pointed out earlier that  $f(\sqrt{\alpha}) = \sqrt{\alpha}$  represents the minimum of  $f$ . It then follows that  $f(I) \subset I$ . We have now satisfied the requirements of Theorem 2.4. So a sequence began with  $x_1 \in I$  will converge to  $\alpha$ .

As a matter of fact, even if  $x_1 \in (0, \sqrt{\frac{\alpha}{2}})$ , convergence is guaranteed. This is so because  $(x_2 = f(x_1)) > \sqrt{\alpha}$ , and hence  $x_2 \in I$ . So, we see that for *any*  $x_1 \in (0, \infty)$ , the resulting sequence is guaranteed to converge to  $x = \sqrt{\alpha}$ .

Hence, we have *verified* this procedure for computing the square roots of  $\alpha$ .

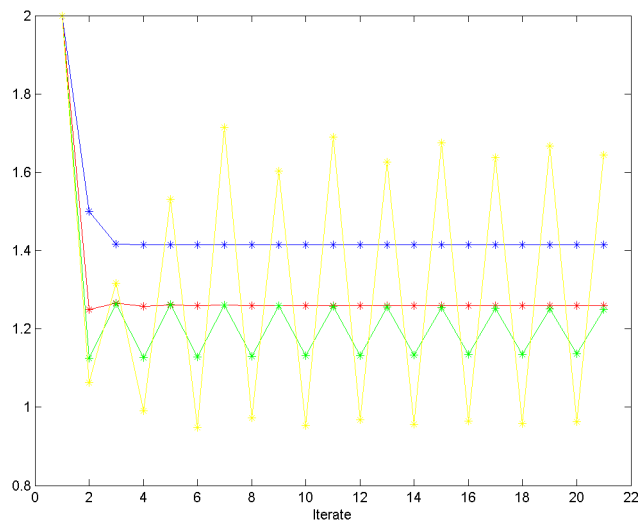
## 4 Generalization to n-th roots

We now wish to see if the function  $f$  can be generalized to allow for calculations of the  $n$ -th root of  $\alpha$ . To treat as a fixed point problem, this requires  $x = \sqrt[n]{\alpha}$  to be a fixed point of  $f$ . The most straightforward approach to achieve this is to let:

$$f(t) = \frac{1}{2} \cdot \left( t + \frac{\alpha}{t^{n-1}} \right)$$

### 4.1 Numerical investigation

As a means of exploration, we fix  $\alpha = 2$  and  $x_0 = 2$ , and attempt the procedure numerically for various  $n \geq 2$ . Consider the results below:



In Figure (4.1), it is seen that the iterates converge for  $n = 2$  and  $n = 3$ , but not for larger values of  $n$ . Similar results are obtained even if one chooses different  $\alpha$  and  $x_0$ .

The goal of this section then is to see if we can understand this behavior using the apparatus we have developed. It turns out that the divergence of the sequences for  $n \geq 4$  can be predicted.

As before, we see if it is possible to identify a subset  $I$  such that  $f(I) \subset I$  and  $f' \leq A < 1$  in  $I$ . The derivative  $f'$  is:

$$f'(t) = \frac{1}{2} \cdot \left( 1 - \frac{(n-1)\alpha}{t^n} \right) \tag{9}$$

and, important to our discussion, we find that,

$$\begin{aligned} f'(x = \sqrt[n]{\alpha}) &= \frac{1}{2} \cdot \left( 1 - \frac{(n-1)\alpha}{\sqrt[n]{\alpha^n}} \right) \\ &= \frac{1}{2} \cdot (1 - (n-1)) \end{aligned}$$

$$= 1 - \frac{n}{2}$$

Now, recall that we want to identify an interval  $I$  containing the fixed point  $x = \sqrt[n]{\alpha}$  such that the derivative is always bounded by one. However, this is *impossible* for the case  $n \geq 4$  since in such a case  $|f'(x)| \geq 1$ . Hence, for  $n \geq 4$  Theorem 2.4 does not apply and we can no longer guarantee convergence.

## 4.2 How to beat the system

We have shown so far that the naive generalization of  $f(t)$  does not allow us to compute the fourth-roots of numbers. However, it is easy to see that the original method of computing square roots via  $f(t) = \frac{1}{2}(t + \frac{\alpha}{t})$  can be extended to compute fourth roots. This is so because if we are interested in the fourth root of  $\alpha$ , we can first compute  $\sqrt{\alpha}$  and then repeat the procedure with  $\alpha' = \sqrt{\alpha}$ . Since we have shown that  $n = 2, 3$  cases can be computed directly, we can calculate  $N$ -th roots by function composition, where  $N = 2^n 3^m$  with  $n, m$  positive integers. The question of fixed points of composition of  $f$ 's is interesting and I hope to investigate it in future editions of this assignment.

## 5 Conclusion

In this paper I have derived some simple theorems regarding fixed points of continuous real functions and have applied them to establish the validity of an algorithm to calculate square roots. In an attempt to generalize the procedure to  $n$ -th roots, we found that the procedure is no longer valid for  $n \geq 4$ .

As mentioned before, in future edits investigations of the fixed points of composition of  $f$ 's is interesting, as well as a closer examination at what seems to be a closed cycle about the fixed point for  $n = 4$ . (See Figure (4.1).)

Furthermore, another idea that emerged, but haven't been examined, is the notion of stability. Theorem 2.4 suggests that if one can obtain any interval containing the fixed point that maps to itself, as well as having a derivative bound by 1, then the fixed point is *stable* in the sense that perturbations away from the fixed point will die down in future iterations. The development of the orbit for  $n = 4$  about the fixed point  $x$  must be a result of the loss of stability that results when  $f'(x) \geq 1$ .

**Acknowledgement.** In writing this assignment, I did not consult any other work than Rudin. The main bulk of the paper was suggested by Problems 22, 23, 24 in Chapter 5 of Rudin. The notation (i.e.  $x_n$  being the “ $n$ -th iterate”) and background information regarding iteration dynamics were remnants from a previous course I had taken, 18.353: Nonlinear Dynamics, in the Fall of 2006.