Approximate Squaring

J. C. Lagarias and *N. J. A. Sloane*
Information Sciences Research Center
AT&T Shannon Lab
Florham Park, NJ 07932–0971

Email addresses: jcl@research.att.com, njas@research.att.com

Aug. 16, 2003; revised Sept. 8, 2003

Abstract

We study the "approximate squaring" map $f(x) := x\lceil x \rceil$ and its behavior when iterated. We conjecture that if f is repeatedly applied to a rational number $r = l/d > 1$ then eventually an integer will be reached. We prove this when $d = 2$, and provide evidence that it is true in general by giving an upper bound on the density of the "exceptional set" of numbers which fail to reach an integer. We give similar results for a p-adic analogue of f, when the exceptional set is nonempty, and for iterating the "approximate multiplication" map $f_r(x) := r\lceil x \rceil$, where r is a fixed rational number.

AMS 2000 Classification: Primary 26A18; Secondary: 11B83, 11K31, 11Y99

1. Introduction

In this paper we study the "approximate squaring" map $f : \mathbb{Q} \to \mathbb{Q}$ given by

$$f(x) := x\lceil x \rceil \ , \tag{1}$$

and consider its behavior when iterated. Although there is an extensive literature on iterated maps (see for example [Collet and Eckmann 1980], [Beardon 1991], [Lagarias 1992]), including the study of various first order recurrences involving the ceiling function ([Eisele and Hadeler 1990], [Graham and Yan 1999]), the approximate squaring map seems not to have been treated before, and has some interesting features.

The function f behaves qualitatively like iterating the rational function $R(x) = x^2$. Indeed, all points $|x| \le 1$ have a bounded orbit under $f(x)$, while all points $|x| > 1$ have unbounded orbits and diverge to ∞, just as they do when $R(x)$ is iterated. However, $f(x)$ has the additional feature that it is discontinuous at integer points. It follows that the n-th iterate $f^{(n)}$ is discontinuous at a certain set of rational points, namely, those points x where $f^{(n)}(x)$ is an integer.

It is therefore natural to ask: if we start with a rational number r and iterate f, will we always eventually reach an integer? This question is the subject of our paper.

Numerical experiments suggest that the answer to our question is "Yes", although it may take many steps, and consequently involve some very large numbers.

For example, starting at $r = \frac{3}{2}$, $f(r) = \frac{3}{2} \cdot 2 = 3$, reaching an integer in one step; and starting at $r = \frac{8}{7}$ we get $f(r) = \frac{16}{7}$, $f^{(2)}(r) = \frac{48}{7}$, $f^{(3)}(r) = 48$, taking three steps. On the other hand, starting at $r = \frac{6}{5}$, we find

$$f(r) = \frac{12}{5},\ f^{(2)}(r) = \frac{36}{5},\ f^{(3)}(r) = \frac{288}{5},\ f^{(4)}(r) = \frac{16704}{5},\ f^{(5)}(r) = \frac{55808064}{5},$$

$$f^{(6)}(r) = \frac{622908012647232}{5},\ f^{(7)}(r) = \frac{776028784440252019977703040704}{5},\ \ldots,$$

and we do not reach an integer until $f^{(18)}(r)$, which is a number with 57735 digits.

We note that for any rational starting point r, since $\lceil x \rceil$ is an integer, the denominators d_j of the iterates $f^{(j)}(r)$ must form a nonincreasing sequence with d_{j+1} dividing d_j. For $0 < r \le 1$, $\lceil r \rceil = 1$ and $f(r) = r$, so there the denominator is fixed. For $-1 < r \le 0$, $f(r) = 0$, and for $r \le -1$, $f(r) \ge 1$. So it is sufficient to restrict our attention to the case of rationals $r > 1$.

We make the following conjecture:

Conjecture 1. *For each rational $r \in \mathbb{Q}$ with $r > 1$, there is an integer $m \ge 0$ such that $f^{(j)}(r)$ is an integer for all $j \ge m$.*

We establish the conjecture in the special case when the denominator is 2, where a complete analysis is possible. This is done in Section 2.

In Section 3 we consider the case of rational starting values r with a fixed denominator $d \ge 3$. We show that at most a sparse subset of the rationals $\{r = \frac{l}{d} : d < l \le x\}$ can fail to become integers under iteration, in the sense that the cardinality of this subset is bounded above by $C(d, \epsilon)x^{1-\alpha_d+\epsilon}$ for a certain positive constant α_d and any positive ϵ, where $C(d, \epsilon)$ is a positive constant depending only on d and ϵ. Showing that this "exceptional set" of starting values which fail to reach integers is in fact empty (or even finite) appears to be a difficult problem, for reasons indicated below.

In Section 4 we consider a p-adic analogue of the approximate squaring map. In this case we show that there is a nonempty exceptional set of elements in $\frac{1}{p^k}\mathbb{Z}_p$ which under iteration never "escape" to the smaller invariant set $\frac{1}{p^{k-1}}\mathbb{Z}_p$. This set has Hausdorff dimension exactly $1 - \alpha_{p^k}$, where α_{p^k} is the same constant that appeared in Section 3. The existence of this exceptional set is a reason why it may be a difficult problem to obtain better upper bounds on the cardinality of the exceptional set in Section 3.

In Section 5 we study similar questions concerning the "approximate multiplication" map

$$f_r(x) := r\lceil x \rceil\ , \tag{2}$$

where r is a fixed rational number. For $r = \frac{1}{b}$, this map is a special case of the map $x \mapsto a + \lceil \frac{x}{b} \rceil$, where $a, b \in \mathbb{Z}$, $b \ge 2$, studied by P. Eisele and R. P. Hadeler [Eisele and Hadeler 1990]. Recently, J. S. Tanton [Tanton 2002], together with Charles Adler, formulated a game-theoretic problem "Survivor", and noted that its analysis leads to the study of the sequence of rational numbers

$$a_0 = r,\ a_n = r\lceil a_{n-1} \rceil \text{ for } n \ge 1\ ,$$

for $r > 1$, which is the trajectory of r under the map $f_r(x)$. He raised the question, "Must some a_n be an integer?", and conjectured that the answer is "Yes". This question differs from the case of the approximate squaring map in that the denominators of successive iterates, though bounded by the denominator of r, may increase or decrease. We note that the long-term dynamics of iterating this map differs according to whether $|r| < 1$, $|r| = 1$ or $|r| > 1$, with the case $r > 1$ being most analogous to the approximate squaring map.

The approximate multiplication maps have some resemblance to the map occurring in the $3x+1$ problem. Setting $r = \frac{l}{d}$, we observe that $f_r(x)$ maps the domain $\frac{1}{d}\mathbb{Z}$ into itself, and on this domain is conjugate to the map $g_r : \mathbb{Z} \to \mathbb{Z}$ given by

$$g_r(n) = \begin{cases} \frac{1}{d}n & \text{if } n \equiv 0 \pmod{d}\,, \\ \frac{1}{d}(n + l(d-b)) & \text{if } n \equiv b \pmod{d}, 1 \le b \le d-1\,, \end{cases}$$

(see (40) of Section 5). In terms of the conjugated map the question we consider becomes whether for most starting values some iterate of g_r is an integer divisible by d. For $r = \frac{3}{2}$ the conjugated map is

$$g_{3/2}(n) = \begin{cases} \frac{3}{2}n & \text{if } n \equiv 0 \pmod 2, \\ \frac{3}{2}n + \frac{3}{2} & \text{if } n \equiv 1 \pmod 2. \end{cases}$$

This is similar in form to the $3x+1$ function

$$T(n) = \begin{cases} \frac{1}{2}n & \text{if } n \equiv 0 \pmod 2, \\ \frac{3}{2}n + \frac{1}{2} & \text{if } n \equiv 1 \pmod 2. \end{cases}$$

as given in [Lagarias 1985] and [Wirsching 1998], although the long-term dynamics of $g_{3/2}$ and T are different.

We formulate a conjecture for approximate multiplication maps analogous to the one above for the approximate squaring map. We define the exceptional set $E(r)$ for the map f_r to be

$$E(r) := \{n : \ n \in \mathbb{Z} \ \text{ and } \ \text{no} \ \text{ iterate } \ f_r^{(j)}(n) \in \mathbb{Z} \text{ for } j \geq 1\} \ . \tag{3}$$

Then we have:

Conjecture 2. *For each nonintegral rational $r \in \mathbb{Q}$ with $|r| > 1$, the exceptional set $E(r)$ for the approximate multiplication map f_r is finite.*

The "expanding map" hypothesis $|r| > 1$ is necessary in the statement of this conjecture, for the conclusion fails for $r = \frac{1}{d}$ with $d \geq 3$, as remarked below. In parallel to the results for the approximate squaring map we prove Conjecture 2 for rational r having denominator 2; it remains open for all rationals with $|r| > 1$ having denominator $d \geq 3$.

In Section 5 we exhibit an analogy between the problem of showing that $E(r)$ is finite for a given nonintegral $r > 1$ and the problem of showing that there exist no Mahler Z-numbers, a notorious problem connected with powers of $\frac{3}{2}$ ([Mahler 1968], [Flatto, 1991]; also [Choquet 1980], [Lagarias 1985]). This suggests that Conjecture 2 may be difficult.

Our analysis in Section 5 applies more generally to the family $\mathcal{P}_r$ of maps $h_r : \mathbb{Z} \to \mathbb{Z}$ having the form

$$h_r(n) = \frac{1}{d}(ln + l_b) \text{ when } n \equiv b \pmod d \ , \tag{4}$$

where the integers l_b satisfy $l_b \equiv -lb \pmod d$. We obtain for all functions $h_r \in \mathcal{P}_r$ an explicit upper bound on the cardinality of the exceptional set $E(h_r, x)$ consisting of all integers $|n| \leq x$ that do not have some iterate that is a multiple of d. We show that for all rationals r,

$$\#E(h_r, x) \leq 4dx^{\beta_d},$$

with $\beta_d = \frac{\log(d-1)}{\log d}$.

We also show that this upper bound is of the correct order of magnitude (to within a multiplicative constant) for certain values of r lying in $0 < r < 1$. This is the case for the function g_r for $r = \frac{1}{d}$ with $d \geq 3$. This result implies that the conclusion of Conjecture 2 does not hold for these values of r.

The final section gives some numerical results related to these questions.

Notation: $\lceil \ \rceil$ denotes the ceiling function and $\{\!\{ \ \}\!\}$ the fractional part. For a prime p, $| \ |_p$ denotes the p-adic valuation. (If $r \in \mathbb{Q}$, $r = p^a \frac{b}{c}$ with $a, b, c \in \mathbb{Z}, c \neq 0, \gcd(p, b) = \gcd(p, c) = 1$, then $|r|_p = a$.) $\mathbb{Q}_p$ and $\mathbb{Z}_p$ denote the p-adic rationals and integers, respectively. For integers r, s, i, $r|s$ means r divides s, and $r^i||s$ means r^i divides s but r^{i+1} does not.

2. Denominator 2

In this section we investigate the case when the starting value r has denominator 2. Here we are able to give a complete analysis. The following table shows what happens for the first few values of r. It gives the initial term, the number of steps to reach an integer, and the integer that is reached.

start :	$\frac{3}{2}$	$\frac{5}{2}$	$\frac{7}{2}$	$\frac{9}{2}$	$\frac{11}{2}$	$\frac{13}{2}$	$\frac{15}{2}$	$\frac{17}{2}$	$\frac{19}{2}$	$\cdots$
steps :	1	2	1	3	1	2	1	4	1	$\cdots$
reaches :	3	60	14	268065	33	2093	60	1204154941925628	95	$\cdots$

The number of steps appears to match sequence A001511 in [Sloane 1995–2003] (and the numbers reached now form sequence A081853 in that database). Indeed, we have:

Theorem 1. *Let $r = \frac{2l+1}{2}$, with $l \geq 1$. Then $f^{(m)}(r)$ reaches an integer for the first time when $m = |l|_2 + 1$.*

Proof. Note that if $x \in \mathbb{Q}$ has denominator 2 and is not an integer then $\lceil x \rceil = x + \frac{1}{2}$.

We use induction on $|l|_2 = v$. If $v = 0$ then l is odd, $\lceil r \rceil = r + \frac{1}{2} = l + 1$ is even, and $r\lceil r \rceil$ has become an integer in one step, as claimed.

Suppose $v \geq 1$, and

$$l = 2^v + l_{v+1}2^{v+1} + l_{v+2}2^{v+2} + \cdots$$

is the binary expansion of l, where each $l_i = 0$ or 1. Then

$$\begin{aligned}
r\lceil r \rceil &= \left(l + \frac{1}{2}\right)(l+1) \\
&= \frac{1}{2} + \frac{l}{2} + l + l^2 \\
&= \frac{1}{2} + 2^{v-1} + (l_{v+1}+1)2^v + (l_{v+1}+l_{v+2})2^{v+1} + \cdots + 2^{2v} + \cdots \\
&= \frac{2l'+1}{2}
\end{aligned}$$

where

$$l' = 2^{v-1} + (l_{v+1}+1)2^v + (l_{v+1}+l_{v+2})2^{v+1} + \cdots + 2^{2v} + \cdots .$$

and $|l'|_2 = v - 1$. By the induction hypothesis, this will reach an integer in $v - 1$ steps, so we are done. ∎

Remark. The numbers $2l+1$ for which $|l|_2 = v$ are precisely the numbers that are congruent to $2^{v+1} + 1 \pmod{2^{v+2}}$. For example, if $v = 0$, $2l+1 \in \{3, 7, 11, 15, \ldots\}$, of the form $3 \pmod 4$; if $v = 1$, $2l+1 \in \{5, 13, 21, 29, \ldots\}$, of the form $5 \pmod 8$; and so on.

Corollary 1. *Let $r = \frac{2l+1}{2}, l \geq 1, |l|_2 = v$. Then the first integer value taken by $f^{(m)}(r)$ is*

$$\frac{1}{2}\,\theta^{(v+1)}(2l+1) ,$$

where $\theta(y) = y(y+1)/2$.

Proof. This is now a straightforward calculation, again using the fact that if $x \in \mathbb{Q} \setminus \mathbb{Z}$ has denominator 2 then $\lceil x \rceil = x + \frac{1}{2}$. ∎

For example, if $v = 0$, and $r = (4k+3)/2 = y/2$ (say), then in one step we reach the integer $\frac{1}{2}\theta(y) = y(y+1)/4$. If $v = 1$, and $r = (8k+5)/2 = y/2$, then in two steps we reach the integer

$$\frac{1}{2}\theta(\theta(y)) = \frac{y(y+1)(y^2+y+2)}{16} ;$$

if $v = 2$, and $r = (16k+9)/2 = y/2$, then in three steps we reach the integer

$$\frac{1}{2}\theta(\theta(\theta(y))) = \frac{y\,(y+1)\,\left(y^2+y+2\right)\left(y^2-y+2\right)\left(y^2+3\,y+4\right)}{256} ,$$

and so on.

3. Denominator d

We now analyze the case of rationals with a general denominator d, obtaining less complete results. The next theorem shows that most rationals will eventually reach an integer. More precisely, it gives an upper bound on the number of such rationals below x that never reach an integer. Given an integer $d \geq 2$, and a bound $x \geq 1$ we study the "exceptional set"

$$\mathcal{M}_d(x) \;:=\; \{l \;:\; 1 \leq l \leq x,\; f^{(m)}(\frac{l}{d}) \notin \mathbb{Z} \text{ for each } m \geq 1\} , \tag{5}$$

and let $M_d(x) = |\mathcal{M}_d(x)|$. The finite set $[1, d-1] := \{1, 2, \ldots, d-1\}$ is contained in $\mathcal{M}_d(x)$, and Conjecture 1 asserts that $\mathcal{M}_d(x) = [1, d-1]$.

Theorem 2. *For each integer $d \geq 2$, there is a positive exponent α_d such that for each $\varepsilon > 0$ and all $x > 1$,*

$$M_d(x) \leq C(d, \varepsilon)\, x^{1-\alpha_d+\varepsilon} \tag{6}$$

for a positive constant $C(d, \varepsilon)$, with α_d given by

$$\alpha_d \;=\; \min_{d'|d,\, d'>1} \log_{d'}\left(\frac{d'}{\phi(d')}\right) , \tag{7}$$

where ϕ is the Euler totient function. In fact

$$\alpha_d = \min_{p^j||d} \frac{\log\left(1+\frac{1}{p-1}\right)}{j\log p} . \tag{8}$$

Note: It follows immediately from (7) that $0 < \alpha_d \leq 1$ and $\alpha_d = 1$ only for $d = 2$.

Proof. We need only consider numbers l satisfying $d < l \leq x$. For the numbers $0 < l \leq d$ contribute $d = O(1)$ (as $x \to \infty$) to $M_d(x)$.

We write $\frac{l}{d} = \frac{l_0}{d_0}$ with $\gcd(l_0, d_0) = 1$, and set

$$f^{(j)}\left(\frac{l}{d}\right) = \frac{l_j}{d_j} , \; j \geq 1 ,$$

where $\gcd(l_j, d_j) = 1$ and $d_j | d$ with $d_j \geq 1$. The pair (l, d) determines the sequence $d_0, d_1, d_2, \ldots$.

We call a value l "bad" at size x if there is an iterate $m \geq 0$ such that

$$\prod_{j=0}^{m-1} d_j \leq x < \prod_{j=0}^{m} d_j \ , \tag{9}$$

with $d_m > 1$, and we let $N_d(x)$ denote the number of bad l at size x. We will obtain an upper bound for $N_d(x)$. Every l counted in $M_d(x)$ has $d_j \geq 2$ for all $j \geq 0$, hence certainly such l are bad for all sizes, and so in particular

$$M_d(x) \leq N_d(x) \ .$$

We establish the theorem by showing the stronger result that

$$N_d(x) \leq C(d, \varepsilon)\, x^{1-\alpha_d+\varepsilon} \ . \tag{10}$$

Suppose d has s prime factors (counted with multiplicity), where $s \leq \log_2 d$. Let l be bad at size x. Then $d_j \geq 2$ for $j = 0, 1, \ldots, m$ (cf. (9)). Define r_j by

$$r_j = \frac{d_{j-1}}{d_j}$$

for $j = 1, 2, \ldots, m$, and call j a *break-point* if $r_j > 1$. Suppose there are t break-points. At each break-point a nontrivial factor is removed from the denominator, and since d has s prime factors and $d_m > 1$, it follows that $t \leq s - 1$. Let the break-points be $j_1, j_2, \ldots, j_t$. We call the sequence of pairs

$$\mathcal{Y} := \{(j_1, d_{j_1}), (j_2, d_{j_2}), \ldots, (j_t, d_{j_t})\} \tag{11}$$

the *chain* associated with (l, d, x).

We calculate an upper bound on the number $\#(\mathcal{Y})$ of distinct chains. There are $\binom{m}{t}$ choices for the break-points, and (9) yields the bound

$$m \leq \log_2 x \ ,$$

so there are no more than

$$\sum_{t=0}^{s-1} \binom{\log_2 x}{t} \leq s (\log_2 x)^{s-1}$$

break-point patterns. Since in each chain a nontrivial divisor r_i of d_{i-1} is chosen at each step, we conclude that the number of distinct chains is at most $d! s (\log_2 x)^{s-1}$. For d fixed and large x this gives

$$\#(\mathcal{Y}) \leq C_1(d, \varepsilon) x^{\varepsilon}. \tag{12}$$

Therefore it suffices to prove an upper bound of the form (10) for the number $N_d(\mathcal{Y}, x)$ of bad l at size x having a given chain $\mathcal{Y}$.

Suppose now that l, d, x, m are as above, and that l has chain (11). We claim that these l fall into a certain set of arithmetic progressions modulo $d_0 d_1 \ldots d_{m-1}$, and in fact there are exactly

$$\phi(d_1)\phi(d_2) \cdots \phi(d_m) \tag{13}$$

such arithmetic progressions.

To see this, we write $\frac{l_k}{d_k}$ (for $k = 0, 1, \ldots$) in a mixed-radix expansion where the radices depend on k:

$$\frac{l_k}{d_k} = \frac{a_{-1}(k)}{d_k} + a_0(k) + \sum_{j=1}^{\infty} a_j(k) \prod_{l=0}^{j-1} d_{k+l} \ , \tag{14}$$

in which the "digits" $a_j(k)$ satisfy

$$0 < a_{-1}(k) < d_k \, , \text{ and } 0 \le a_j(k) < d_{j+k} \text{ for each } j \ge 0 \ .$$

Here we set $d_{m+j} := d_m$ for all $j \ge 1$. The sum on the right-hand side of (14) is actually a finite sum. By definition,

$$\begin{aligned} \frac{l_{k+1}}{d_{k+1}} = f\left(\frac{l_k}{d_k}\right) &= \left(\frac{a_{-1}(k)}{d_k} + a_0(k) + \sum_{j=1}^{\infty} a_j(k) d_k d_{k+1} \cdots d_{k+j-1}\right) \\ &\times \left(1 + a_0(k) + \sum_{j=1}^{\infty} a_j(k) d_k d_{k+1} \cdots d_{k+j-1}\right) \ . \end{aligned}$$

We use induction on k to establish four properties of this mixed-radix expansion:

$$r_{k+1} = \gcd(a_0(k) + 1, d_k) \ , \tag{15}$$

$$a_{-1}(k+1) \equiv a_{-1}(k) \frac{a_0(k) + 1}{r_{k+1}} \ (\mathrm{mod}\ d_{k+1}) \ , \tag{16}$$

$$\gcd(a_{-1}(k+1), d_{k+1}) = 1 \ , \tag{17}$$

and, for $0 \le j \le m$,

$$a_j(k+1) \equiv a_{-1}(k) a_{j+1}(k) + G(a_{-1}(k), a_0(k), a_1(k), \ldots, a_j(k)) \ (\mathrm{mod}\ d_{j+k+1}) \ , \tag{18}$$

in which the function $G(a_{-1}(k), a_0(k), a_1(k), \ldots, a_j(k))$ includes all the necessary information about "carries" in the multiple-radix expansion.

The base case $k = 0$ is checked directly. Since d_0 divides all terms in the sum on the right-hand side of (14), the right-hand side of (15) (when $k = 0$) has a single term $\frac{a_{-1}(0)(a_0(0)+1)}{d_0}$ having a denominator, and this term equals $\frac{a_{-1}(1)}{d_1} (\mathrm{mod}\ 1)$. Since $\gcd(a_{-1}(0), d_0) = 1$ by hypothesis, we must have

$$\gcd(a_0(0) + 1, d_0) = \frac{d_0}{d_1} = r_1 \ ,$$

which is (15) for $k = 0$. The term with a denominator in (15) is then

$$\frac{a_{-1}(0) \frac{a_0(0)+1}{r_1}}{d_1}$$

hence

$$a_{-1}(1) \equiv a_{-1}(0) \left(\frac{a_0(0) + 1}{r_1}\right) \ (\mathrm{mod}\ d_1) \ , \tag{19}$$

which is (16). Now $\gcd(a_{-1}(1), d_1)$ divides $\gcd(a_{-1}(0), d_1) \gcd(\frac{a_0(0)+1}{r_1}, d_1)$, both terms of which are 1, so (17) follows. Finally, to establish (18) when $k = 0$, we drop the terms involving $d_1 d_2 \cdots d_{j+1}$ from (15) and observe that there is a term

$$a_{-1}(0) a_{j+1}(0) d_1 d_2 \cdots d_j \ , \tag{20}$$

while all the other terms containing any $a_{j+l}(0)$ for $l \geq 1$ are divisible by $d_1 d_2 \cdots d_{j+1}$. (Note that $d_1 d_2 \cdots d_{j+1}$ divides $d_0 d_1 d_2 \cdots d_j$.) This establishes that a congruence of the form (18) holds for the digit $a_j(1)$ in the expansion (15) for $\frac{l_1}{d_1}$, completing the proof of the base case.

The induction step for general k follows using exactly the same reasoning.

Next, we claim that (18) implies that

$$a_0(k+1) \equiv a_{k+1}(0) \prod_{l=0}^{k} a_{-1}(l) + \widetilde{G}(a_{-1}(0), a_0(0), a_1(0), \ldots, a_k(0)) \pmod{d_{k+1}} . \qquad (21)$$

for some function $\widetilde{G}$ depending on the indicated variables.

To prove this, we again use induction on k. For $k = 0$ the assertion is that

$$a_0(1) \equiv a_1(0) a_{-1}(0) + \widetilde{G}(a_{-1}(0), a_0(0)) \pmod{d_1} . \qquad (22)$$

which is (18) with $k = j = 0$. For $k = 1$ we wish to show

$$a_0(2) \equiv a_2(0) a_{-1}(0) a_{-1}(1) + \widetilde{G}(a_{-1}(0), a_0(0), a_1(0)) \pmod{d_2} . \qquad (23)$$

Setting $k = 1, j = 0$ and $k = 0, j = 1$ in (18) we obtain

$$a_0(2) \equiv a_{-1}(1) a_1(1) + \widetilde{G}(a_{-1}(1), a_0(1)) \pmod{d_2}$$

and

$$a_1(1) \equiv a_{-1}(0) a_2(0) + \widetilde{G}(a_{-1}(0), a_0(0), a_1(0)) \pmod{d_2} ,$$

hence

$$a_0(2) \equiv a_2(0) a_{-1}(0) a_{-1}(1) + a_{-1}(1) \widetilde{G}(a_{-1}(0), a_0(0), a_1(0)) + \widetilde{G}(a_{-1}(1), a_0(1)) \pmod{d_2} .$$

However, from (19) and the fact that $0 < a_{-1}(1) < d_1$, $a_{-1}(1)$ is uniquely determined by $a_{-1}(0)$ and $a_0(0)$. Also the induction hypothesis allows us to use (22) to eliminate $a_0(1)$. Equation (23) follows. The case of general k follows in the same way; we leave the details to the reader. The important point about (21) is that the dependence on $a_{k+1}(0)$ is linear, even though the dependence on the other initial terms $a_{-1}(0), a_0(0), \ldots, a_k(0)$ is nonlinear.

We have already seen that

$$\gcd(a_{-1}(0), d_1) = 1 .$$

From (17) and (16) ,

$$\gcd(a_{-1}(2), d_2) = 1 ,$$

$$a_{-1}(2) \equiv a_{-1}(1) \frac{a_0(1)+1}{r_2} \pmod{d_2} .$$

Therefore $\gcd(a_{-1}(1), d_2) = 1$ and so

$$\gcd(a_{-1}(0) a_{-1}(1), d_2) = 1 .$$

Continuing in this way we obtain

$$\gcd(\prod_{l=0}^{k} a_{-1}(l), d_{k+1}) = 1 , \qquad (24)$$

for $k = 0, 1, \ldots, m$.

Then (21) shows that the congruence class of $a_0(k+1) \pmod{d_{k+1}}$ is uniquely determined by the congruence class of $a_{k+1}(0) \pmod{d_{k+1}}$, once $a_{-1}(0), a_0(0), ..., a_k(0)$ are specified.

In particular there are exactly $\phi(d_{k+2})$ congruence classes of $a_{k+1}(0) \pmod{d_{k+1}}$ that give

$$\gcd(a_0(k+1)+1, d_{k+2}) \;=\; 1 \;,$$

or in other words which give

$$\gcd(a_0(k+1)+1, d_{k+1}) \;=\; r_{k+2} \;.$$

At each iteration we impose one such condition, and thus (13) follows.

Now (9) shows that each arithmetic progression $(\bmod\ d_0 d_1 \cdots d_{m-1})$ contains at most d_m elements below x. From (13), $N_d(\mathcal{Y}, x)$, the number of bad elements l at size x with chain $\mathcal{Y}$ in (11) satisfies

$$N_d(\mathcal{Y}, x) \le d_m \phi(d_1)\phi(d_2)\cdots\phi(d_m) \;.$$

Now $x \ge d_0 d_1 \cdots d_{m-1} \ge d_1 d_2 \cdots d_{m-1}$, hence

$$\frac{N_d(\mathcal{Y}, x)}{x} \le d_m \prod_{j=1}^{m} \frac{\phi(d_j)}{d_j} \;,$$

and therefore

$$\begin{aligned} \frac{N_d(\mathcal{Y}, x)}{x} \quad &\le \quad d_m\, e^{\sum_{j=1}^{m} \log \frac{\phi(d_j)}{d_j}} \\ &\le \quad d\, x^{\frac{\sum_{j=0}^{m} \log \frac{\phi(d_j)}{d_j}}{\log x}} \;. \end{aligned} \tag{25}$$

Now $x \le d_0 d_1 \cdots d_m$, so that

$$\begin{aligned} \frac{\sum_{j=0}^{m} \left|\log \frac{\phi(d_j)}{d_j}\right|}{\log x} \quad &\ge \quad \frac{\sum_{j=0}^{m} \left|\log \frac{\phi(d_j)}{d_j}\right|}{\sum_{j=0}^{m} \log d_j} \\ &\ge \quad \min_{0\le j\le m} \frac{\left|\log \frac{\phi(d_j)}{d_j}\right|}{\log d_j} \\ &\ge \quad \min_{0\le j\le m} \left(\log_{d_j} \left(\frac{d_j}{\phi(d_j)} \right) \right) \\ &\ge \quad \alpha_d \;, \end{aligned}$$

where we used the definition (7). Substituting this in (25) yields

$$\frac{N_d(\mathcal{Y}, x)}{x} \le d\, x^{-\alpha_d} \;,$$

since $\frac{\phi(d_j)}{d_j} < 1$ gives $\log \frac{\phi(d_j)}{d_j} < 0$, and so

$$N_d(\mathcal{Y}, x) \le d\, x^{1-\alpha_d} \;,$$

as required. Combined with (12), this gives (10), hence (6).

Finally, we establish the equivalence of (7) and (8). Since ϕ is multiplicative, for a general $d' > 1$ we have

$$\begin{aligned} \log_{d'}\left(\frac{d'}{\phi(d')}\right) &= \frac{\sum_{p^j\|d'} \log\left(\frac{p^j}{\phi(p^j)}\right)}{\sum_{p^j\|d} \log p^j} \\ &\geq \min_{p^j\|d'} \left\{ \log_{p^j}\left(\frac{p^j}{\phi(p^j)}\right) \right\} . \end{aligned} \tag{26}$$

Thus the minimum in (7) is attained when d' is a prime power. Now

$$\log_{p^j}\left(\frac{p^j}{\phi(p^j)}\right) = \frac{\log\left(\frac{1}{1-\frac{1}{p}}\right)}{\log\ p^j} = \frac{\log\left(1+\frac{1}{p-1}\right)}{j\log p} \tag{27}$$

is minimized by making j as large as possible, so we obtain the formula (8). ■

Remark. If l, d, x, m are such that

$$d_1 = d_2 = \cdots = d_m = d'$$

where d' is the value that minimizes (7), so that $t = 1$ and the chain is simply $\mathcal{Y} = \{(1, d')\}$, we have

$$N_d(\mathcal{Y}, x) \geq C_d' x^{1-\alpha_d} ,$$

for some positive constant C_d'. It follows that the upper bound in (10) has essentially the best possible exponent.

4. p-Adic iteration

We now consider an approximate squaring map defined on the p-adic numbers analogous to the approximate squaring map on $\mathbb{Q}$, and study the question of whether some iterate will eventually become a p-adic integer. We show that now there is always a nonempty exceptional set of p-adic numbers which never become p-adic integers.

Let p be a prime ≥ 2 and let $\mathbb{Q}_p$ denote the p-adic numbers, with typical element $\alpha = \sum_{j=-k}^{\infty} a_j p^j$, where $k \in \mathbb{Z}$ and the a_j satisfy $0 \leq a_j \leq p-1$. The p-adic integral part of α is given by the function $F_p : \mathbb{Q}_p \to \mathbb{Q}_p$, where

$$F_p(\alpha) := \sum_{j=\max\{0,k\}}^{\infty} a_j p^j ,$$

while the p-adic fractional part (or "principal part") of α is

$$P_p(\alpha) := \sum_{j<0} a_j p^j ,$$

which is a finite sum (possibly empty); thus $\alpha = P_p(\alpha) + F_p(\alpha)$. We investigate the function

$$f_p(\alpha) := \alpha(F_p(\alpha) + 1) .$$

which is a p-adic analogue of the approximate squaring map defined in (1). If we regard the rationals $\mathbb{Q}$ as embedded in $\mathbb{Q}_p$, then for nonintegral r in the subring $\mathbb{Z}[\frac{1}{p}] \subseteq \mathbb{Q} \subseteq \mathbb{Q}_p$, we have $f_p(r) = r\lceil r \rceil$, so the iterates there agree with the approximate squaring map.

For each $k \ge 0$ the set $\frac{1}{p^k}\mathbb{Z}_p$ is invariant under the action of $f_p(x)$, and

$$\mathbb{Z} \subseteq \frac{1}{p}\mathbb{Z} \subseteq \frac{1}{p^2}\mathbb{Z} \subseteq \cdots .$$

We define the exceptional set

$$\Omega_k(p) := \left\{ \alpha \in \frac{1}{p^k}\mathbb{Z}_p \ : \ f^{(m)}(\alpha) \notin \frac{1}{p^{k-1}}\mathbb{Z}_p \ \text{for each} \ m \ge 1 \right\} , \tag{28}$$

for $k \ge 1$. The set $\Omega_k(p)$ is an analogue of the exceptional sets $\mathcal{M}_d(x)$ studied in the last section, corresponding to the denominator $d = p^k$. (Note that $\Omega_k(p) \cap \frac{1}{p^k}\mathbb{Z}_{>0}$ is contained in $\frac{1}{p^k}\mathcal{M}_{p^k}(\infty)$). We will show that these sets are nonempty, determine their Hausdorff dimension s, and get upper and lower bounds on their Hausdorff s-dimensional measure.

The s-dimensional p-adic Hausdorff measure $\mathcal{H}_p^s(\Omega)$ of a closed set Ω in $\mathbb{Q}_p$ is defined by the general prescription in [Falconer 1990, Chapter 2] or [Federer 1969, Section §2.10]. Here $0 < s \le 1$. The *diameter* of a measurable set $S \subseteq \mathbb{Q}_p$ is

$$|S| := \sup\{|\alpha - \beta|_p \ : \ \alpha, \beta \in S\} ,$$

and its *p-adic measure* $\mu_p(S)$ is Haar measure with the normalization $\mu_p(\mathbb{Z}_p) = 1$. A p-adic ball

$$B(\alpha; p^l) := \{\beta \in \mathbb{Q}_p \ : \ |\beta - \alpha|_p \le p^l\}$$

is both closed and open, and has the property that its diameter equals its measure:

$$|B(\alpha; p^l)| = \mu(B(\alpha; p^l)) = \frac{1}{p^l} .$$

For each $\delta > 0$ we define

$$\mathcal{H}^s(\Omega, \delta) := \inf \left\{ \sum_{j=1}^{\infty} |I_j|^s \ : \ \Omega \subseteq \bigcup_{j=1}^{\infty} I_j , \ |I_j| \le \delta \text{ for all } j \right\} , \tag{29}$$

and[1]

$$\mathcal{H}_p^s(\Omega) := \lim_{\delta \to 0} \mathcal{H}^s(\Omega, \delta) . \tag{30}$$

The *p-adic Hausdorff dimension* of Ω is the unique value s_0 such that $\mathcal{H}_p^s(\Omega) = \infty$ for $s < s_0$ and $\mathcal{H}_p^s(\Omega) = 0$ for $s > s_0$. The value of $\mathcal{H}_p^{s_0}(\Omega)$ may be zero, finite or infinite.

A closed set $\Omega \subset \mathbb{Z}_p$ is called *weakly self-similar* (mod p^k) *with branching ratio* b, where b is an integer ≥ 2, if the following "equal branching" property holds for $l = 1, 2, \ldots$. Let $W_l(\Omega)$ denote the set of initial sequences of digits of length lk in Ω, i.e.

$$W_l(\Omega) = \{\beta = \sum_{j=0}^{lk-1} a_j p^j \ : \ \text{there exists some } \alpha \in \Omega \text{ whose "initial part" } p^{lk} P_p(p^{-lk}\alpha) = \beta\} . \tag{31}$$

Then each sequence in $W_l(\Omega)$ should extend to exactly b sequences in $W_{l+1}(\Omega)$. That is, if the digits of an element $\beta \in \Omega$ are grouped in blocks of size k, once the first l blocks of digits are specified, there are exactly b allowable choices for the next block of digits.

[1] The limit (which may be ∞) exists since $\mathcal{H}^s(\Omega, \delta_1) \ge \mathcal{H}^s(\Omega, \delta_2)$ if $\delta_1 \le \delta_2$.

Theorem 3. *For each $k \geq 1$ the set $p^k\Omega_k(p)$ is weakly self-similar (mod p^k) with branching ratio $b = \phi(p^k) = p^k - p^{k-1}$.*

Proof. The set $W_l(\Omega_k(p))$ specifies the conditions under which the first l iterates $f(\alpha)$, $f^{(2)}(\alpha)$, $\ldots$, $f^{(l)}(\alpha) \notin \frac{1}{p^{k-1}}\mathbb{Z}_p$. Each such condition is a congruence (mod p^{lk}), which has exactly $\phi(p^k) = p^k - p^{k-1}$ solutions for the next digit (compare (18) in the proof of Theorem 2, taking each $d_j = p^k$). Finally the definition of $\Omega_k(p)$ implies it is a closed set. Thus $p^k\Omega_k(p)$ is weakly self-similar. ∎

We can determine the Hausdorff dimension of weakly self-similar sets in $\mathbb{Z}_p$, together with upper and lower bounds for the Hausdorff measure at this dimension.

Theorem 4. *Let $\Omega \subseteq \mathbb{Z}_p$ be a weakly self-similar set (mod p^k) with branching ratio b satisfying $2 \leq b < p^k$. Then Ω is a compact set and has Hausdorff dimension $s(\Omega)$ given by*

$$\dim_H(\Omega) = \frac{\log b}{\log p^k} \ .$$

Its $s(\Omega)$-dimensional Hausdorff measure satisfies

$$\left(\frac{b}{p^k}\right)^{1-\frac{1}{k}} \ \leq \ \mathcal{H}_p^{s(\Omega)}(\Omega) \ \leq \ 1 \ . \tag{32}$$

Proof. The compactness of Ω is established similarly to Theorem 3. We define $W_l(\Omega)$ as in (31). This set has cardinality b^l by hypothesis, and

$$\Omega = \bigcap_{l=1}^{\infty} \widetilde{W}_l(\Omega) \ ,$$

where $\widetilde{W}_l(\Omega)$ is the compact set

$$\widetilde{W}_l(\Omega) = \left\{\tilde{\beta} \in \mathbb{Z}_p \ : \ \tilde{\beta} \equiv \beta \ (\mathrm{mod}\ p^{lk}) \text{ for some } \beta \in W_l(\Omega)\right\} \ .$$

To determine the Hausdorff dimension it suffices to establish the inequalities (32), since the fact that the Hausdorff measure is positive and finite determines the Hausdorff dimension. Set $s = (\log b)/(\log p^k)$, so that $p^{ks} = b$. Also $0 < s < 1$.

For the upper bound in (32) we consider the sets $\widetilde{W}_l(\Omega)$. Now b^l balls of diameter p^{lk} cover $\widetilde{W}_l(\Omega)$, hence cover Ω. Thus for $\delta = p^{-lk}$ this covering gives

$$\mathcal{H}^s(\Omega, p^{-lk}) \leq \frac{b^l}{p^{lks}} = \frac{b^l}{b^l} = 1 \ ,$$

which implies

$$\mathcal{H}_p^s(\Omega)) \leq 1 \, .$$

The lower bound argument is similar in spirit to that used for Cantor sets in [Falconer 1990, pp. 31–32]. By the compactness of Ω we need only prove that the lower bound holds for finite coverings. The non-archimedean property of the valuation $|\ |_p$ means that each I_j has diameter p^m for some m, and hence we can enlarge I_j to a ball $B(\alpha; p^m) \supseteq I_j$ without changing its diameter. But $B(\alpha; p^m)$ gives an open cover, so it has a finite subcover:

$$\Omega \subseteq \bigcup_{j=1}^{m} B(\alpha_j; p^{m_j}) \ ,$$

and

$$\sum_{j=1}^{m} |B(\alpha_j; p^{m_j})|^s \;=\; \sum_{j=1}^{m} p^{m_j s} \; .$$

We want to show

$$\sum_{j=1}^{m} |B(\alpha_j; p^{m_j})|^s \;\geq\; \left(\frac{b}{p^k}\right)^{\frac{k-1}{k}} \; . \tag{33}$$

We first replace these balls with balls of diameter p^{-kl_j} for integers l_j. Write

$$m_j = -kl_j + k_j \; , \;\; 0 \leq k_j \leq k-1 \; .$$

Then we can cover $B(\alpha_j; p^{m_j})$ with p^{k_j} balls $B(\alpha_{j'}; p^{-kl_j})$. We claim that

$$\sum_{j=1}^{m} |B(\alpha_j; p^{m_j})|^s \;\geq\; \left(\frac{b}{p^k}\right)^{\frac{k-1}{k}} \sum_{j'} |B(\alpha_{j'}; p^{-kl_j})|^s \; . \tag{34}$$

This will follow if we show that

$$\left(p^{-kl_j+k_j}\right)^s \geq \left(\frac{b}{p^k}\right)^{\frac{k-1}{k}} p^{k_j} \left(p^{-kl_j}\right)^s \; ,$$

holds for each j. This in turn is equivalent to showing

$$p^{k_j(s-1)} \geq \left(\frac{b}{p^k}\right)^{\frac{k-1}{k}} \; .$$

Since $p^{ks} = b$ we have

$$p^{k_j(s-1)} = \left(\frac{b}{p^k}\right)^{\frac{k_j}{k}} \geq \left(\frac{b}{p^k}\right)^{\frac{k-1}{k}} \; ,$$

which proves (34). Thus (33) will follow from showing

$$\sum_{j=1}^{m} |B(\alpha_j; p^{-kl_j})|^s \;\geq\; 1 \tag{35}$$

for any set of such balls that covers Ω. We may suppose $l_1 \leq l_2 \leq \cdots \leq l_m$. By the weak self-similarity of Ω there are b^{l_m} principal parts to cover with balls of diameter p^{-kl_m} in $\widetilde{W}_m(\Omega)$. Weak self-similarity also says that each ball of radius p^{-kl_j} covers either none or else exactly $b^{l_m - l_j}$ such principal parts in $\widetilde{W}_m(\Omega)$. Since the balls cover Ω, we must have

$$\sum_{j=1}^{m} b^{l_m - l_j} \geq b^{l_m} \; .$$

Dividing by b^{l_m} yields

$$\sum_{j=1}^{m} b^{-l_j} \geq 1 \; . \tag{36}$$

Using this bound we obtain

$$\begin{aligned}\sum_{j=1}^{m} |B(\alpha_j; p^{-kl_j})|^s &= \sum_{j=1}^{m} p^{-kl_j s} \\ &= \sum_{j=1}^{m} b^{-l_j} \\ &\geq 1 ,\end{aligned}$$

which is (35). Thus (33) holds. ∎

Corollary 2. *The set $\Omega_k(p)$ has Hausdorff dimension*

$$\dim_H(\Omega_k(p)) = s(p^k) \;:=\; 1 \;-\; \frac{\log\left(1+\frac{1}{p-1}\right)}{k\log p} \tag{37}$$

for $k \geq 1$. Furthermore, its $s(p^k)$-dimensional Hausdorff measure satisfies

$$\left(1-\frac{1}{p}\right)^{1-\frac{1}{k}} (p^k - p^{k-1}) \;\leq\; \mathcal{H}_p^{s(p^k)}(\Omega_k(p)) \;\leq\; p^k - p^{k-1} \;. \tag{38}$$

Proof. This follows by applying Theorems 3 and 4 to the set $p^k\Omega_k(p)$ and using the fact that

$$\mathcal{H}_p^{s(p^k)}(\Omega_k(p)) = b\mathcal{H}_p^{s(p^k)}(p^k\Omega_k(p)) \;,$$

since $(p^k)^s = b$. Using the branching ratio $b = \phi(p^k) = p^k - p^{k-1}$, we have

$$\frac{\log b}{\log p^k} = \frac{\log p^k + \log(1-1/p)}{\log p^k}$$

which gives (37). ∎

Remark. For the branching ratio $b = \phi(p^k) = p^k - p^{k-1}$, one can show that equality may occur on either side of the Hausdorff measure bounds in (32) (or (38)). For the lower bound, take the $p^k - p^{k-1}$ allowed digit sets in each layer to be $\sum_{j=lk}^{k(l+1)-1} a_j\, p^j$ with the restriction that $a_{lk} \neq 0$. Then we can cover Ω with $(p^k - p^{k-1})^{l+1}(p-1)$ balls of radius $p^{-(lk+1)}$, and get

$$\begin{aligned}\mathcal{H}_p^s(\Omega_k(p)) &\leq (p^k - p^{k-1})(p^k - p^{k-1})^l p^{-lks}[(p-1)p^{-s}] \\ &\leq (p^k - p^{k-1})\left[\left(1-\frac{1}{p}\right)\left(1-\frac{1}{p}\right)^{-\frac{1}{k}}\right] .\end{aligned}$$

Since this coincides with the lower bound in (38), $\mathcal{H}_p^s(\Omega_k(p))$ must equal this bound.

For the upper bound, choose the $p^k - p^{k-1}$ allowed digits in each layer to be $\sum_{j=lk}^{k(l+1)-1} a_j\, p^j$ with the restriction that $a_{k(l+1)-1} \neq 0$. Then each residue class (mod $p^{-k(l+1)+r}$) covers exactly $(p-1)p^{r-1}$ classes of Ω (mod p^{-lk}), and we can do no better than the upper bound.

It would be interesting to know how many elements $r = \frac{l}{p^k}$ have no iterate $f_p^n(r) \in \frac{1}{p^{k-1}}\mathbb{Z}_p$. We know that this set contains the $p^k - p^{k-1}$ elements $1 \leq l \leq p^k - 1$ with $(l,p) = 1$. We conjecture that these are the only such elements.

5. Approximate multiplication maps

We can use similar methods to study iteration of the approximate multiplication map $f_r : \mathbb{Q} \to \mathbb{Q}$ given by

$$f_r(x) = r\lceil x \rceil \ , \tag{39}$$

where r is a fixed rational number, say $r = \frac{l}{d}$ with $\gcd(l, d) = 1$. In this case, since x enters into the iteration only as the integer $\lceil x \rceil$, we may restrict attention to initial values $x \in \mathbb{Z}$. We consider the case when the denominator $d > 1$, and study the question of whether some iterate $f_r^{(j)}(x)$ will be an integer for some $j \geq 1$. Note that all iterates lie in $\frac{1}{d}\mathbb{Z}$.

Unlike the case of approximate squaring, the iterates do not remain integral once they become integral. However, the truth of Conjecture 2 would imply that infinitely many members of a sequence of iterations $\{f_r^{(j)}(n) \ : \ j \geq 1\}$ will be integers, provided $|r| > 1$.

It is convenient to rescale the map to eliminate the denominators d, by conjugating f_r by the dilation $\Phi_d(x) = d\,x$. The result is the map $g_r : \mathbb{Z} \to \mathbb{Z}$ given by $g_r(x) := \Phi_d \circ f_r \circ \Phi_d^{-1}(x)$. Thus

$$g_r(x) = d\,f_r(\frac{x}{d}) = l\lceil \frac{x}{d} \rceil \ , \tag{40}$$

with

$$g_r^{(j)}(x) = d\,f_r^{(j)}(\frac{x}{d})$$

for $j = 1, 2, \ldots$. We have

$$g_r(n) = \frac{1}{d}(ln + l_b) \text{ when } n \equiv b \ (\mathrm{mod}\ d) \ ,$$

where $l_0 = 0$ and

$$l_b = l(d - b) \ \text{ for } \ 1 \leq b \leq d - 1 \ .$$

For example, when $r = \frac{3}{2}$, we have

$$g_{3/2}(n) = \begin{cases} \frac{3}{2}n & \text{if } \ n \equiv 0 \ (\mathrm{mod}\ 2)\,, \\ \frac{3}{2}n + \frac{3}{2} & \text{if } \ n \equiv 1 \ (\mathrm{mod}\ 2)\,. \end{cases}$$

Our question then becomes: when does the sequence of iterations $\{g_r^{(j)}(n) \ : \ j \geq 1\}$ contain a term which is divisible by d?

The map g_r belongs a general class of functions which we will denote by $\mathcal{P}_r$, consisting of all "periodically linear" functions $h_r : \mathbb{Z} \to \mathbb{Z}$ of the form

$$h_r(n) = \frac{1}{d}(ln + l_b) \text{ when } n \equiv b \ (\mathrm{mod}\ d) \ , \tag{41}$$

where $r = l/d$ is rational and the integers $\{l_b : \ 0 \leq b \leq d - 1\}$ satisfy the conditions

$$l_b \equiv -lb \ (\mathrm{mod}\ d) \tag{42}$$

needed to give an integer-valued map. Thus, although the notation does not reflect this, h_r is defined by specifying $r = l/d$ and constants $l_0, l_1, \ldots, l_{d-1}$ satisfying (42). We note that any map h_r is "self-similar" in the sense that its linear part $\frac{l}{d}n$ is independent of the residue class. Various other classes of periodically linear functions have been studied in connection with the $3x+1$ problem – see §3.2 of [Lagarias 1985].

Another interesting map in this class is

$$\tilde{g}_r(x) = \lceil rx \rceil \ . \tag{43}$$

This map has $l_0 = 0$ and

$$l_b = d - (lb \mod d) \text{ for } 1 \le b \le d-1 \ .$$

The function $\tilde{g}_{3/2}(x)$ appears in Mahler's study of Z-numbers [Mahler 1968], as explained below. We have

$$\tilde{g}_{3/2}(n) = \begin{cases} \frac{3}{2}n & \text{if } n \equiv 0 \pmod 2, \\ \frac{3}{2}n + \frac{1}{2} & \text{if } n \equiv 1 \pmod 2. \end{cases}$$

Our methods apply generally to the question of whether a particular function $h_r \in \mathcal{P}_r$ has an iterate that is divisible by d (the denominator of r). The behavior of the map h_r depends on whether $|r| > 1$, the "expanding map" case; $r = \pm 1$, the "indifferent map" case; or $|r| \le 1$, the "contracting map" case. Our motivation comes from the expanding map case, but the results proved below apply to all cases.

We formulate a general conjecture concerning functions in this class that are expanding maps, of which Conjecture 2 is a special case.

Conjecture 3. *Let $r = \frac{l}{d}$ with $\gcd(l, d) = 1$ and $|r| > 1$. Let $h_r : \mathbb{Z} \to \mathbb{Z}$ be a function in the class $\mathcal{P}_r$. Then for each integer n, with at most a finite number of exceptions, there is some iterate $j \ge 1$ such that $h_r^{(j)}(n) \equiv 0 \pmod d$.*

The "expanding" condition on r is necessary, for the conjecture fails for certain r with $0 < r < 1$, as shown in Theorem 7.

Mahler's study of Z-numbers [Mahler 1968] led to questions similar in spirit to Conjecture 3. A *Z-number* is a positive real number ξ with the property that

$$0 \ \le \ \{\!\{ (\frac{3}{2})^n \xi \}\!\} \ \le \ \frac{1}{2} \text{ for all } \ n \ge 1 \ ,$$

where $\{\!\{x\}\!\} = x - \lfloor x \rfloor$ denotes the fractional part of x. Mahler conjectured that Z-numbers do not exist, and showed that a necessary and sufficient condition for their non-existence is that for each $n \ge 1$ there exists some $j \ge 1$ (depending on n) such that

$$\tilde{g}_{3/2}^{(j)}(n) \equiv 3 \pmod 4. \tag{44}$$

Mahler's conjecture remains open. Mahler obtained a nontrivial upper bound on the number of Z-numbers smaller than x, and [Flatto 1991] improved the upper bound to $O(x^{0.59})$ for $x \to \infty$.

In comparison, Conjecture 3 asserts for $r = \frac{3}{2}$ that for each $n \in \mathbb{Z}$ there exists some $j \ge 1$ (depending on n) with

$$\tilde{g}_{3/2}^{(j)}(n) \equiv 0 \pmod 2. \tag{45}$$

This special case of Conjecture 3 is true, as a consequence of the following theorem. There is exactly one exceptional integer, -1, whose iterates never satisfy (45).

More generally, for the case of rational numbers r with denominator $d = 2$, Conjecture 3 is provable for all functions in the class $\mathcal{P}_r$ in a fashion analogous to that used for the approximate squaring map in Section 2.

Theorem 5. *Let h_r be a function in the class $\mathcal{P}_r$, for fixed $r = \frac{2t+1}{2}$ where t is an integer. Then for each $n \in \mathbb{Z}$, with at most two exceptions, there exists some iterate $k \ge 1$ with*

$$h_r^{(k)}(n) \equiv 0 \pmod 2. \tag{46}$$

Proof. We have

$$h_r(n) = \begin{cases} r\,n + l_0 & \text{if } \; n \equiv 0 \pmod{2}, \\ r\,n + \frac{1}{2} + l_1 & \text{if } \; n \equiv 1 \pmod{2}. \end{cases}$$

for some integers l_0, l_1.

We claim that the set of integers n satisfying

$$h_r^{(j)}(n) \not\equiv 0 \pmod{2} \text{ for } 1 \le j \le k \tag{47}$$

consist of the integers in exactly two arithmetic progressions $b \pmod{2^{k+1}}$, one consisting of even integers and one of odd integers. We prove the claim by induction on $k \ge 1$. For the base case, let $b \equiv a_0 \pmod{2}$ with $a_0 = 0$ or 1 fixed, and consider the arithmetic progression $n = b + 2m$, with $m \in \mathbb{Z}$. Then

$$h_r(n) = h_r(b) + (2t+1)m,$$

and the condition $h_r(n) \equiv 0 \pmod{2}$ then restricts m to lie in a single congruence class $m \equiv a_1(b) \pmod{2}$. We conclude that exactly two congruence classes $b \equiv a_0 + 2a_1 \pmod{4}$ satisfy (47) for $k = 1$, one consisting of even integers and one of odd integers, completing the base case.

For the induction step, supposing (47) true for k, let $b \pmod{2^{k+1}}$ run over the two allowed congruence classes for the given k. Consider the arithmetic progression $n = b + 2^{k+1}m$, with $m \in \mathbb{Z}$. Then we have

$$h_r^{(j)}(n) \equiv h_r^{(j)}(b) \pmod{2} \;\; \text{for} \;\; 1 \le j \le k\;,$$

and

$$h_r^{(k+1)}(n) = h_r^{k+1}(b) + (2t+1)^{k+1}m\;.$$

The condition that

$$h_r^{(k+1)}(n) \equiv 0 \pmod{2}$$

is equivalent to $m \equiv a_{k+1}(b) \pmod{2}$, which excludes the congruence class

$$b' \equiv b + a_{k+1}(b)2^{k+1} \pmod{2^{k+2}}\;.$$

Thus two congruence classes $\pmod{2^{k+2}}$ remain which satisfy (47) for $1 \le j \le k+1$. Since each of the previous classes $b \pmod{2^{k+1}}$ contributed one of these classes, one contains even integers and the other contains odd integers. This completes the induction step.

Denote these two classes by $b_0(k+1) \pmod{2^{k+2}}$ and $b_1(k+1) \pmod{2^{k+2}}$, respectively. It follows that (46) holds except for integers in the sets

$$\bigcup_{k=1}^{\infty} \{n \equiv b_0(k) \pmod{2^{k+1}}\}.$$

and

$$\bigcup_{k=1}^{\infty} \{n \equiv b_1(k) \pmod{2^{k+1}}\}.$$

Each of these sets contains at most one element, so there are at most two exceptional elements. ∎

Remarks. (1) One can find examples of functions $h_r \in \mathcal{P}_r$ with denominator $d = 2$ such that there are zero, one or two elements in the exceptional set. For example, for $r = \frac{3}{2}$ the function $\tilde{g}_{3/2}$ has the single exceptional element $n = -1$. The map

$$h_{3/2}(n) = \begin{cases} \frac{3}{2}n + \frac{1}{2} & \text{if } n \equiv 1 \pmod{2}, \\ \frac{3}{2}n - 1 & \text{if } n \equiv 0 \pmod{2} \end{cases} \tag{48}$$

has two exceptional points, 0 and -1, all the iterates of which are odd.

(2) The analysis applies to the cases $r = \pm\frac{1}{2}$ where the map h_r is contracting.

Concerning Conjecture 3 for denominators $d \geq 3$, we can bound the number of exceptions below x for a general function in the class $\mathcal{P}_r$ in a fashion similar to that for the approximate squaring map studied in Section 3. Given a rational number $r = \frac{l}{d}$ with $\gcd(l, d) = 1$ and a function h_r in the class $\mathcal{P}_r$, we define the exceptional set by

$$E(h_r; x) := \{n \in \mathbb{Z} : \ |n| \leq x, \ h_r^{(k)}(n) \not\equiv 0 \pmod{d} \text{ for all } k \geq 1\}, \tag{49}$$

and let

$$N(h_r; x) := \#(E(h_r; x)).$$

(The exceptional set in (3) is $E(r) = E(g_r; \infty)$.) The following result holds for all rational r, including those with $|r| \leq 1$.

Theorem 6. *Let $r = \frac{l}{d}$ be a rational number with $\gcd(l, d) = 1$, and suppose that $d \geq 2$. There is a constant $0 \leq \beta_d < 1$ depending only on d such that for every function $h_r \in \mathcal{P}_r$ we have*

$$N(h_r; x) \leq 4dx^{\beta_d}. \tag{50}$$

The precise value of the constant is

$$\beta_d = \frac{\log(d-1)}{\log d} = 1 - \log_d\left(1 + \frac{1}{d-1}\right). \tag{51}$$

Proof. We will prove a stronger result. Set

$$E^*(h_r; x) := \{n \in \mathbb{Z} : \ |n| \leq x, \ h_r^{(k)}(n) \not\equiv 0 \pmod{d} \text{ for } 1 \leq k \leq \log_d x\} \tag{52}$$

and let $N^*(h_r; x) = \#(E^*(h_r; x))$. Certainly $N^*(h_r; x) \geq N(h_r; x)$, so it suffices to establish

$$N^*(h_r; x) \leq 4dx^{\beta_d}, \tag{53}$$

where β_d is given in (51).

The argument we will use to establish (53) is simpler than that for the approximate squaring map because we can use radix expansions to the fixed base d. We suppose $d^k \leq x < d^{k+1}$. If $n \equiv b \pmod{d}$ then

$$h_r(n) = \frac{l}{d}n + \frac{l_b}{d}, \tag{54}$$

where $0 \leq l_b \leq d - 1$ with $l_b \equiv -lb \pmod{d}$.

We claim that the elements $n \in \mathbb{Z}$ such that

$$h_r^{(j)}(n) \not\equiv 0 \pmod{d} \text{ for } 1 \leq j \leq k \tag{55}$$

consist of a certain set of $d(d-1)^k$ residue classes (mod d^{k+1}). We proceed by induction on $k \ge 1$. For the base case $k = 1$, given $b \pmod d$ the elements of the arithmetic progression $n = b + dm$ with $m \in \mathbb{Z}$ have

$$h_r(n) = h_r(b) + lm \ .$$

Since $\gcd(l, m) = 1$, as $m \in \mathbb{Z}$ varies these numbers cycle through every residue class $\pmod d$. In particular there is one class $m \equiv a_1(b) \pmod d$, say, that gives $g_r(n) \equiv 0 \pmod d$. This arithmetic progression $b + a_1(b)d \pmod{d^2}$ is ruled out and all elements of the remaining $d-1$ arithmetic progressions $\pmod{d^2}$ satisfy (55) with $k = 1$. This completes the base case.

For the induction step, suppose (55) holds for k, and there are $d(d-1)^k$ allowed residue classes $b \pmod{d^{k+1}}$. For each of these residue classes consider the arithmetic progression $n = b + d^{k+1}m$ with $m \in \mathbb{Z}$. Using (54) repeatedly, we have

$$h_r^{k+1}(n) = h_r^{(k+1)}(b) + l^{k+1}m \ .$$

Since $\gcd(l, m) = 1$ this progression cycles through all residue classes $\pmod d$, and the condition $h_r^{k+1}(n) \equiv 0 \pmod d$ rules out one residue class $b + a_{k+1}(b)d^{k+1} \pmod{d^{k+2}}$, say. Thus imposing the additional condition

$$h_r^{k+1}(n) \not\equiv 0 \pmod d$$

leaves $d(d-1)^{k+1}$ allowed residue classes $\pmod{d^{k+2}}$ whose elements satisfy (55) for $k+1$. This completes the induction step.

Applying (55), with k replaced by $k-1$, we see that the elements in $E^*(h_r; x)$ are necessarily contained in a set of $d(d-1)^{k-1}$ residue classes $\pmod{d^k}$, since $\log_d x > k-1$. It follows that

$$N^*(h_r; x) \le d(d-1)^{k-1} \cdot 2\left\lceil \frac{x}{d^k} \right\rceil \le 2d^2(d-1)^{k-1} \le 4d(d-1)^k,$$

since $d \ge 2$. Now $x \ge d^k$, so $k \le \frac{\log x}{\log d}$ and we have

$$\begin{aligned} N^*(h_r; x) &\le 4de^{k\log(d-1)} \\ &\le 4dx^{\frac{\log(d-1)}{\log d}}, \end{aligned}$$

as asserted. ∎

Theorem 6 is nearly best possible for certain values of r in the interval $0 < r < 1$, where the map is a contracting map. That is, for suitable maps in the class $\mathcal{P}_r$, the upper bound (50) of Theorem 6 is within a multiplicative constant of the best possible upper bound.

Theorem 7. *Let $r = \frac{1}{d}$ with $d \ge 3$. Then, for the conjugated approximate multiplication map g_r of (40), the exceptional set $E(g_r; x)$ has cardinality*

$$N(g_r; x) \ge \frac{1}{d} x^{\beta_d} \ , \ \textit{for all } x \ge d \ , \tag{56}$$

with $\beta_d = \frac{\log(d-1)}{\log d}$. In particular, the full exceptional set $E(g_r; \infty)$ is infinite.

Proof. Note that for $r = \frac{1}{d}$ the functions g_r and $\tilde{g}_r$ coincide. We claim that, for each $k \geq 1$, the subset Σ_k of $[1, d^k]$ given by

$$\{n : 1 \leq n \leq d^k,\ n = a_0 + a_1 d + \ldots + a_{k-1} d^{k-1} \text{ with all } a_i \not\equiv -1 \pmod{d} \ \text{ for } i \geq 1\}$$

is contained in $E(g_r; d^k)$. (The cardinality of Σ_k is $d(d-1)^{k-1}$.) We prove this by induction on $k \geq 1$.

For the base case $k = 1$ we have

$$\Sigma_1 = \{1, 2, ..., d-1\} \subseteq E(g_r; d)$$

because each $g_r(n) = 1 \not\equiv 0 \pmod{d}$ and 1 is a fixed point of g_r. Next,

$$g_r(\Sigma_k) \subset \Sigma_{k-1} ,$$

because

$$g_r(n) = (a_1 + 1) + a_2 d + a_3 d^2 + \cdots + a_{k-1} d^{k-2} \in \Sigma_{k-1} ,$$

and $1 \leq a_1 + 1 \leq d - 1$ by hypothesis. Thus $\Sigma_k \subset E(g_r; d^k)$, which completes the induction step.

For $d^k \leq x < d^{k+1}$, with $k \geq 1$, we have

$$N(g_r; x) \geq N(g_r; d^k) \geq \#\Sigma_k = d(d-1)^{k-1} \geq \frac{1}{d} x^{\beta_d} , \tag{57}$$

as asserted. ∎

Since the exceptional set $E(r)$ of the approximate multiplication map f_r has $E(r) = E(g_r; \infty)$ and $E(g_r, x) \subset E(g_r; \infty)$, Theorem 7 shows that the conclusion of Conjecture 2 does not hold for these values of r.

It seems plausible that for all values $-1 < r < 1$ (except $r = 0$) there is some function in the class $\mathcal{P}_r$ for which the conclusion of Conjecture 3 does not hold. However we do not attempt to construct such functions here.

6. Numerical results

The simplest case where we do not know if the approximate squaring map f of (1) will always reach an integer is when the starting value $r = l/d$ has denominator $d = 3$. We wish to determine $\theta(r)$ (say), the smallest value of $k \geq 0$ for which $f^{(k)}(r)$ is an integer.

Testing any particular value of r is complicated by the fact–already illustrated in Section 1– that the iterates grow so rapidly. This difficulty can be overcome by writing the k-th iterate $l_k/d_k := f^{(k)}(r)$ in "base d":

$$\frac{l_k}{d_k} = \sum_{j=-1}^{\infty} a_j(k) d^k , \tag{58}$$

where the "digits" $a_j(k)$ satisfy $0 \leq a_j(k) < d$ (compare (14)), but storing only the terms in (58) with $j \leq M$. That is, we work mod d^{M+1}. As long as $\theta(r) \leq M - 1$, we get the correct answer by finding the smallest k for which $a_{-1}(k) = 0$. If this has not happened by the time k reaches M we increase M and repeat.

For denominator 3 the value $M = 25$ is sufficient to show that $\theta(l/3)$ is finite for $3 \leq l \leq 2000$. The following table shows what happens for the first few values. It gives the initial term, the number of steps to reach an integer, and the integer that is reached.

start :	$\frac{3}{3}$	$\frac{4}{3}$	$\frac{5}{3}$	$\frac{6}{3}$	$\frac{7}{3}$	$\frac{8}{3}$	$\frac{9}{3}$	$\frac{10}{3}$	$\frac{11}{3}$	$\cdots$
θ :	0	2	6	0	1	1	0	5	2	$\cdots$
reaches :	1	8	1484710602474311520	2	7	8	3	1484710602474311520	220	$\cdots$

(These are sequences A072340 and A085276 in [Sloane 1995–2003].) In the range $l \le 2000$ large values of $\theta(l/3)$ are scarce. The first few record values are $\theta(l/3) = 0, 2, 6, 22, 23$, reached at $l = 3, 4, 5, 28, 1783$ respectively.

Starting values $r = \frac{d+1}{d}$ take longer to converge–we discussed the cases $d \le 8$ in Section 1. The initial values of $\theta((d+1)/d)$ can be found in sequence A073524 in [Sloane 1995–2003]. The first few record values are $0, 1, 2, 3, 18, 26, 56, 79, 200, 225, 388, 1444$, reached at $d = 1, 2, 3, 4, 5, 11, 19, 31, 37, 67,$ 149, 199 respectively (sequences A073529, A073528). R. G. Wilson, v. [Wilson 2002] has checked that $\theta((d+1)/d)$ is finite for $d \le 500$.

It is amusing to note that the record value 1444 has the following interpretation: starting with $\frac{200}{199}$ and repeatedly approximately squaring, the first integer reached is

$$200^{2^{1444}},$$

a number with about 10^{435} digits.

The approximate multiplication map $f_r(x)$ of (39) is easier to compute since it grows more slowly. Let $\theta_r(n)$ denote the smallest value of $k \ge 1$ for which $f_r^{(k)}(n)$ is an integer. We give just one example. This table shows what happens when $f_{4/3}(n)$ is iterated with starting value n:

n:	0	1	2	3	4	5	6	7	8	9	10	11	12	$\cdots$
$\theta_{4/3}(n)$:	1	3	2	1	2	9	1	8	3	1	7	2	1	$\cdots$
reaches:	0	4	4	4	8	84	8	84	20	12	84	20	16	$\cdots$

(sequences A085068 and A085071). Large values of $\theta_{4/3}(n)$ are again scarce. The first few record values are $\theta_{4/3}(n) = 1, 3, 9, 15, 17, 18, 24, 27, 28, 30, 40$, reached at $0, 1, 5, 161, 1772, 3097, 3473, 23084,$ 38752, 335165, 491729 respectively (sequences A085328 and A085330). We thank J. Earls [Earls 2003] for computing the last six terms in these two sequences.

All the evidence supports the conjectures made here; it would be nice to know more.

Acknowledgements

We thank R. Bacher, B. Cloitre, J. Earls, T. D. Noe and R. G. Wilson, v., for their help in contributing and extending several sequences in [Sloane 1995–2003] during the early stages of this investigation.

References

[Beardon 1991] A. F. Beardon, *Iteration of Rational Functions*, Springer-Verlag, NY, 1991.

[Choquet 1980] G. Choquet, Construction effective de suites $(k(3/2)^n)$. Étude des mesures (3/2)-stables, *C. R. Acad. Sci. Paris*, Sér. A-B **291** (1980), A69–A74.

[Collet and Eckmann 1980] P. Collet and J.-P. Eckmann, *Iterated Maps on the Interval as Dynamical Systems*, Birkhäuser: Boston 1980.

[Earls 2003] J. Earls, Comment on sequence A085328 in [Sloane 1995–2003], August 14, 2003.

[Eisele and Hadeler 1990] P. Eisele and K. P. Hadeler, Game of cards, dynamical systems, and a characterization of the floor and ceiling functions, *Amer. Math. Monthly*, **97** (1990), 466–477.

[Falconer 1990] K. Falconer, *Fractal Geometry: Mathematical Foundations and Applications*, John Wiley and Sons, Chichester, 1990.

[**Federer 1969**] H. Federer, *Geometric Measure Theory*, Springer-Verlag, NY, 1969.

[**Flatto 1991**] L. Flatto, Z-numbers and β-transformations, in P. Walters, ed., *Symbolic Dynamics and its Applications (New Haven, CT, 1991)*, Contemp. Math., **135**, Amer. Math. Soc., Providence, RI, 1992, pp. 181–201.

[**Graham and Yan 1999**] R. L. Graham and C. H. Yan, On the limit of a recurrence relation, J. Difference Eqn. Appl. **5** (1999), 71–95.

[**Lagarias 1985**] J. C. Lagarias, The 3x+1 problem and its generalizations, *Amer. Math. Monthly*, **92** (1985), 3–23.

[**Lagarias 1992**] J. C. Lagarias, Number theory and dynamical systems, in: *The Unreasonable Effectiveness of Number Theory (S. A. Burr, Ed.),* Proc. Symp. Applied Math., No. 46 (1992), 35–72.

[**Mahler 1968**] K. Mahler, An unsolved problem on the powers of 3/2, *J. Australian Math. Soc.*, **8** (1968), 313–321.

[**Sloane 1995–2003**] N. J. A. Sloane, *The On-Line Encyclopedia of Integer Sequences*, published electronically at www.research.att.com/∼njas/sequences/.

[**Tanton 2002**] J. S. Tanton, *A Collection of Research Problems*, published electronically at www.themathcircle.org/dozen%20problems.htm.

[**Wilson 2002**] R. G. Wilson, v., Comment on sequence A073524 in [Sloane 1995–2003], September 11, 2002.

[**Wirsching 1998**] G. J. Wirsching, *The Dynamical System Generated by the* $3n + 1$ *Function*, Lecture Notes in Math., Vo1. 1681, Springer-Verlag: New York 1998.