ADDENDUM TO “ON RECURRANCES CONVERGING TO THE WRONG LIMIT
IN FINITE PRECISION AND SOME NEW EXAMPLES”∗

SIEGFRIED M. RUMP†

Abstract. In a recent paper [Electron. Trans. Numer. Anal, 52 (2020), pp. 358–369], we analyzed Muller’s famous recurrence, where, for particular initial values, the iteration over real numbers converges to a repellent fixed point, whereas finite precision arithmetic produces a different result, the attracting fixed point. We gave necessary and sufficient conditions for such recurrences to produce only nonzero iterates. In the above-mentioned paper, an example was given where only finitely many terms of the recurrence over \( \mathbb{R} \) are well defined, but floating-point evaluation indicates convergence to the attracting fixed point. The input data of that example, however, are not representable in binary floating-point, and the question was posed whether such examples exist with binary representable data. This note answers that question in the affirmative.

Key words. recurrences, rounding errors, IEEE-754, exactly representable data, bfloat, half precision (binary16), single precision (binary32), double precision (binary64)

AMS subject classifications. 65G50, 11B37

1. Main result. In 1989, Muller [3] presented the recurrence

\[
x_0 := 11/2, \quad x_1 := 61/11, \quad \text{and} \quad x_{n+1} := 111 - (1130 - 3000/x_{n-1})/x_n.
\]

The limit of the recurrence over the field of real numbers is 6, whereas in double precision it converges to 100. Subsequently, similar examples were given by Kahan [2], together with some analysis, and again also by Muller [4].

In [5] these recurrences were analyzed stating a necessary and sufficient criterion for such a sequence being well defined, i.e., no zero iterate is encountered. More precisely, let

\[
x_{n+1} := a + (b + c/x_{n-1})/x_n, \quad \text{with} \quad a, b, c \in \mathbb{R},
\]

for given initial values \((x_0, x_1) \in \mathbb{R}^2\). Setting \(y_{n+1} := x_n y_n\), for \(0 \leq n \in \mathbb{N}\) and \(y_0 := 1\), defines the characteristic polynomial

\[
\chi(y) = y^3 - ay^2 - by - c =: (y - \alpha)(y - \beta)(y - \gamma)
\]

as in [5, Equation (2.3)]. We restrict our attention to recurrences satisfying

\[
|\alpha| > |\beta| > |\gamma| > 0 \quad \text{and} \quad \alpha, \beta, \gamma \in \mathbb{R}.
\]

**Lemma 1.1** ([5, Lemma 2.1]). Let \(x_0, x_1 \in \mathbb{R}\) be given, and let the recurrence (1.1) with the characteristic polynomial (1.2) satisfy (1.3). Then (1.1) is well defined and \(x_1 \to \beta\) if and only if

\[
\begin{align*}
x_0 &\neq \gamma \\
x_1 &\neq \beta + \gamma - \beta \gamma/x_0 \\
x_0 &\neq \gamma - \gamma^n(\beta - \gamma)/\beta^n - \gamma^n \quad \text{for all } n \geq 1.
\end{align*}
\]
By this lemma, the recurrence \((x_i)\) is well defined and converges to \(\beta\) for \((x_0, x_1)\) on the hyperbola \(H\) defined by \(x_1 = \beta + \gamma - \beta \gamma / x_0\) except for infinitely many discrete points. Moreover, it was shown in [5] that in every \(\varepsilon\)-neighborhood of the initial values \((x_0, x_1)\) with a well-defined recurrence converging to \(\beta\), there exists a pair of initial values with not well-defined recurrence.

In [5] we presented the recurrence

\[
x_0 := \frac{109225}{43691}, \quad x_1 := \frac{10923}{4369} k, \quad \text{and} \quad x_{n+1} := 56.5 + \left(160 - \frac{737.5}{x_{n-1}}\right) / x_n.
\]

Over \(\mathbb{R}\), this produces \(x_{16} = 0\), but when evaluated in half, single, or double precision, the floating-point iteration is well defined and becomes stationary at the attracting fixed point \(\alpha = 59\); see [5, Table 2.1].

The input data \(x_0\) and \(x_1\) are not representable in binary format in any precision, and it was asked in [5, p. 364] whether there are similar examples with all data representable in some binary format. To answer that in the affirmative, we use the following lemma.

**Lemma 1.2.** For given \(a, b, c \in \mathbb{C}, c \neq 0\), let \(\beta\) and \(\gamma\) be any roots of \(x^3 - ax^2 - bx - c = 0\). Let \(n \in \mathbb{N}\) with \(n \geq 3\) be given, and assume that \(\beta^j \neq \gamma^j\), for \(j \in \{1, \ldots, n\}\). Then,

\[
x_0 = \gamma - \gamma^n (\beta - \gamma) / \beta^n - \gamma^n, \quad x_1 = \beta + \gamma - \beta \gamma / x_0,
\]

and \(x_{k+1} := a + (b + c / x_k) / x_k\), for \(k \geq 1\), imply

\[
x_k = \frac{\beta \gamma (\beta^{n-k-1} - \gamma^{n-k-1})}{\beta^n - \gamma^n}, \quad \text{for} \ 0 \leq k \leq n - 1.
\]

**Remark 1.3.** Note that \(\beta \gamma \neq 0\) because \(c \neq 0\), and that (1.4) implies \(x_0 x_1 \neq 0\) and \(x_{n-1} = 0\).

**Proof of Lemma 1.2.** A computation shows that (1.4) is true for \(k = 0\), and similarly, the assumption \(x_1 = \beta + \gamma - \beta \gamma / x_0\) implies (1.4) for \(k = 1\). Abbreviate \(\delta_j := \beta^j - \gamma^j\), and note that \(\delta_j \neq 0\) for \(j \in \{1, \ldots, n\}\). We have to prove that \(x_k = \frac{\beta \gamma \delta_{n-k-1}}{\delta_{n-k}}\). The definition of the recurrence implies

\[
x_{k+1} = a + \left( b + \frac{c \delta_{n-k+1}}{\beta \gamma \delta_{n-k}} \right) \frac{\delta_{n-k}}{\beta \gamma \delta_{n-k-1}}
\]

\[
= a \beta^2 \gamma^2 \delta_{n-k-1} + b \beta \gamma \delta_{n-k} + c \delta_{n-k+1}
\]

\[
= \frac{\beta^{n-k+1} (a \gamma^2 + b \gamma + c) - \gamma^{n-k+1} (a \beta^2 + b \beta + c)}{\beta^2 \gamma^2 \delta_{n-k-1}}
\]

\[
= \frac{\beta^{n-k+1} \gamma^3 - \gamma^{n-k+1} \beta^3}{\beta^2 \gamma^2 \delta_{n-k-1}} = \frac{\beta \gamma \delta_{n-k-2}}{\delta_{n-k-1}},
\]

and this proves the result. \(\Box\)
Let \( x_{n+1} = a + (b + c/x_{n-1})/x_n \) for given \( a, b, c, x_0, x_1 \in \mathbb{R} \). Then, for \( \varphi \in \mathbb{R} \), the recurrence

\[
X_{n+1} := A + (B + C/X_{n-1})/X_n
\]

with

\[
(1.5) \quad A := \varphi a, \quad B := \varphi^2 b, \quad C := \varphi^3 c, \quad X_0 := \varphi x_0, \quad X_1 := \varphi x_1
\]

satisfies \( X_k = \varphi x_k \) for \( k \geq 0 \). Hence, a recurrence with rational \( a, b, c, x_0, x_1 \) can be transformed into a similar one with integer quantities. Using Lemma 1.2, a desired example with integer data may be constructed as follows:

- Choose some integer \( n \geq 2 \).
- Choose \( p, q \in \mathbb{Q}, q \neq 0 \), and denote the roots of \( x^2 + px + q \) by \( \beta, \gamma \).
- Make sure that \( \beta^j \neq \gamma^j \) for \( j \in \{1, \ldots, n\} \).
- Choose \( \alpha \in \mathbb{Q} \) with \( |\alpha| > \max(|\beta|, |\gamma|) \).
- Let \( x^3 - ax^2 - bx - c = (x - \alpha)(x^2 + px + q) \).
- Define \( x_{n-1} := 0 \) and \( x_{n-2} := \beta^j \gamma = -q/p \).
- Compute \( x_0, x_1 \) recursively by \( x_{k-1} = c(x_k x_{k+1} - ax_k - b)^{-1} \).

Obviously all data are rational, and by using (1.5) we may produce integer data. By construction, the recurrence (1.1) with the initial values \( x_0, x_1 \) produces \( x_{n-1} = 0 \) over \( \mathbb{R} \). If in some finite precision format, one of the \( x_k \) for \( 2 \leq k \leq n-2 \) is not representable, then likely the floating-point approximation of \( x_{n-1} \) will be nonzero, and the recurrence will converge to the attracting fixed point \( \alpha \).

**Lemma 1.4.** For given \( a, b, c \in \mathbb{R} \) assume that the roots \( \alpha, \beta, \gamma \) of \( x^3 - ax^2 - bx - c = 0 \) satisfy \( |\alpha| > |\beta| > |\gamma| > 0 \). For given \( x_0 \in \mathbb{R}, x_0 \neq \gamma \), let \( x_1 := \beta + \gamma - \beta \gamma/x_0 \), and assume that \( x_0 x_1 \neq 0 \). Finally, assume that

\[
x_0 = \gamma - \frac{\gamma^n(\beta - \gamma)}{\beta^n - \gamma^n}
\]

for some integer \( n \geq 2 \). Then in every \( \varepsilon \)-neighborhood of \((x_0, x_1)\) there exist \((x'_0, x'_1)\) and \((x''_0, x''_1)\) for which the recurrence \( x_{k+1} := a + (b + c/x_{k-1})/x_k \) is well defined for all \( k \) such that for the initial values \((x'_0, x'_1)\) it converges to the repelling fixed point \( \beta \), whereas for the initial values \((x''_0, x''_1)\) it converges to the attracting fixed point \( \alpha \).

**Proof.** By [5, Lemma 2.1], for each pair of initial values \((x_0, x_1)\) on the hyperbola \( x_1 := \beta + \gamma - \beta \gamma/x_0 \), the recurrence converges to the repelling fixed point \( \beta \), provided it is well defined, i.e., \( x_0 \neq \gamma - \frac{\gamma^n(\beta - \gamma)}{\beta^n - \gamma^n} \) for all \( n \in \mathbb{N} \). Thus, the set of exceptional pairs \((x_0, x_1)\) for which the recurrence is not well defined is countable, implying the existence of initial values \((x'_0, x'_1)\) with the desired property. The existence of a pair \((x''_0, x''_1)\) follows by [5, Corollary 2.4].

Based on the previous considerations it is not difficult to construct examples with the desired property, for instance,

\[
x_{n+1} := 6496 - (4205 \cdot 2^{10} + 609725 \cdot 2^{15}/x_{n-1})/x_n \quad \text{for} \quad x_0 := -1305, \; x_1 := -1440.
\]

The roots of the characteristic polynomial are

\[
\alpha = 4640 \quad \text{and} \quad \beta, \gamma = 928 \pm 928\sqrt{6} \approx [-1345.13, 3201.13].
\]
Results for $x_{n+1} := 6496 - (4205 \cdot 2^{10} + 609725 \cdot 2^{15}/x_{n-1})/x_n$ with the initial values $x_0 := -1305$, $x_1 := -1440$.

<table>
<thead>
<tr>
<th>n</th>
<th>single</th>
<th>double</th>
<th>over $\mathbb{R}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>-1305.00000000000000000000</td>
<td>-1305.00000000000000000000</td>
<td>-1305</td>
</tr>
<tr>
<td>1</td>
<td>-1440.00000000000000000000</td>
<td>-1440.00000000000000000000</td>
<td>-1440</td>
</tr>
<tr>
<td>2</td>
<td>-1145.679199218750000000</td>
<td>-1145.67901234567943900000</td>
<td>-92800/81</td>
</tr>
<tr>
<td>3</td>
<td>-1855.999023437500000000</td>
<td>-1855.99999999999818100000</td>
<td>-1856</td>
</tr>
<tr>
<td>4</td>
<td>-580.002441406250000000</td>
<td>-580.00000000000027285000</td>
<td>-580</td>
</tr>
<tr>
<td>5</td>
<td>-4639.963867187500000000</td>
<td>-4639.99999999672582000000</td>
<td>-4640</td>
</tr>
<tr>
<td>6</td>
<td>-0.01953125000000000000</td>
<td>-0.0000000000109139000000</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>4780.799804687500000000</td>
<td>3680.00000000000000000000</td>
<td>497456029492482816.00000000</td>
</tr>
<tr>
<td>8</td>
<td>213975808.00000000000000</td>
<td>497456029492482816.00000000</td>
<td>497456029492482816.00000000</td>
</tr>
<tr>
<td>9</td>
<td>6495.960449218750000000</td>
<td>6495.99999999979991100000</td>
<td>6495.99999999999818100000</td>
</tr>
<tr>
<td>10</td>
<td>5833.124511718750000000</td>
<td>5833.14285714284869750000</td>
<td>5833.14285714284869750000</td>
</tr>
<tr>
<td>46</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>47</td>
<td>4640.000097656250000000</td>
<td>4640.00097735409963210000</td>
<td>4640.00097735409963210000</td>
</tr>
<tr>
<td>48</td>
<td>4640.000048828125000000</td>
<td>4640.00067427444628270000</td>
<td>4640.00067427444628270000</td>
</tr>
<tr>
<td>49</td>
<td>4640.000048828125000000</td>
<td>4640.00046518048930010000</td>
<td>4640.00046518048930010000</td>
</tr>
<tr>
<td>50</td>
<td>4640.000000000000000000</td>
<td>4640.00032092703349920000</td>
<td>4640.00032092703349920000</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>102</td>
<td>4640.000000000000000000</td>
<td>4640.00000000000000000000</td>
<td>4640.00000000000000000000</td>
</tr>
<tr>
<td>103</td>
<td>4640.000000000000000000</td>
<td>4640.00000000000000000000</td>
<td>4640.00000000000000000000</td>
</tr>
<tr>
<td>104</td>
<td>4640.000000000000000000</td>
<td>4640.00000000000000000000</td>
<td>4640.00000000000000000000</td>
</tr>
</tbody>
</table>

The data $x_0$, $x_1$, a, b, c are exactly representable in 20-bit binary format. The left two columns of Table 1.1 display the result in IEEE-754 [1] single (binary32) and double (binary64) precision.

As can be seen, both in single and double precision, the recurrence is defined and converges to the attracting fixed point $\alpha = 4640$. However, at the 8-th iterate, it becomes visible that something happened during the iteration. The second example was constructed by Paul Zimmermann [7] from INRIA using Sage [6]:

$$x_{n+1} := -256 + (131072/x_{n-1})/x_n$$

for $x_0 := 3$, $x_1 := 170$.

The roots of the characteristic polynomial are approximately $-253.97$, $-23.76$, and 21.72, and the data $x_0$, $x_1$, $a$, $b$, $c$ are representable in 7 bits. The results of the floating-point iteration in bfloat (8 bits), half (11 bits), single and double precision are displayed in the left four columns of Table 1.2.

In all used floating-point formats, the recurrence converges to the floating-point number nearest to the attracting fixed point $\alpha$. In bfloat, half, and single precision, the floating-point iteration camouflages the true behavior of the recurrence—yet another example of the smoothing effect of rounding operations.

Acknowledgment. Many thanks to Paul Zimmermann for fruitful discussions, really many nice examples, and very careful reading, resulting in several suggestions to improve the note. Also many thanks to Florian Bünger, who read the manuscript very carefully and gave several valuable comments.
ADDENDUM TO RECURRENCES CONVERGING TO THE WRONG LIMIT

Table 1.2

Results for $x_{n+1} := -256 + (131072/x_{n-1})/x_n$ with $x_0 := 3, x_1 := 170$.

<table>
<thead>
<tr>
<th>n</th>
<th>bfloat</th>
<th>half</th>
<th>single</th>
<th>double</th>
<th>over R</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>3.00</td>
<td>3.00</td>
<td>3.000000000000</td>
<td>3.000000000000000000000000000000</td>
<td>3</td>
</tr>
<tr>
<td>1</td>
<td>170.00</td>
<td>170.00</td>
<td>170.0000000000000000000000000000</td>
<td>170.0000000000000000000000000000</td>
<td>170</td>
</tr>
<tr>
<td>2</td>
<td>2.00</td>
<td>1.00</td>
<td>1.00393676758</td>
<td>1.0039215686274474</td>
<td>-256/255</td>
</tr>
<tr>
<td>3</td>
<td>130.00</td>
<td>515.00</td>
<td>511.98840320312</td>
<td>512.000000000000027285</td>
<td>512</td>
</tr>
<tr>
<td>4</td>
<td>248.00</td>
<td>-1.50</td>
<td>512.00000000000004547</td>
<td>-1.000000000000004547</td>
<td>-1</td>
</tr>
<tr>
<td>5</td>
<td>-252.00</td>
<td>-425.50</td>
<td>512.0000000000000000000000000000</td>
<td>-511.9999999999822204</td>
<td>-512</td>
</tr>
<tr>
<td>6</td>
<td>-258.00</td>
<td>-50.75</td>
<td>512.0000000000000000000000000000</td>
<td>0.243438720702</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>-254.00</td>
<td>-249.875</td>
<td>-1306.57568359375</td>
<td>-4.45019 \cdot 10^{12}</td>
<td>-254.00</td>
</tr>
<tr>
<td>8</td>
<td>-254.00</td>
<td>-245.625</td>
<td>512.0000000000000000000000000000</td>
<td>512.0000000000000000000000000000</td>
<td>512</td>
</tr>
<tr>
<td>9</td>
<td>-254.00</td>
<td>-253.875</td>
<td>512.0000000000000000000000000000</td>
<td>512.0000000000000000000000000000</td>
<td>512</td>
</tr>
<tr>
<td>10</td>
<td>-254.00</td>
<td>-254.00</td>
<td>512.0000000000000000000000000000</td>
<td>512.0000000000000000000000000000</td>
<td>512</td>
</tr>
<tr>
<td>11</td>
<td>-254.00</td>
<td>-254.00</td>
<td>512.0000000000000000000000000000</td>
<td>512.0000000000000000000000000000</td>
<td>512</td>
</tr>
<tr>
<td>12</td>
<td>-254.00</td>
<td>-254.00</td>
<td>512.0000000000000000000000000000</td>
<td>512.0000000000000000000000000000</td>
<td>512</td>
</tr>
<tr>
<td>13</td>
<td>-254.00</td>
<td>-254.00</td>
<td>512.0000000000000000000000000000</td>
<td>512.0000000000000000000000000000</td>
<td>512</td>
</tr>
<tr>
<td>14</td>
<td>-254.00</td>
<td>-254.00</td>
<td>512.0000000000000000000000000000</td>
<td>512.0000000000000000000000000000</td>
<td>512</td>
</tr>
<tr>
<td>15</td>
<td>-254.00</td>
<td>-254.00</td>
<td>512.0000000000000000000000000000</td>
<td>512.0000000000000000000000000000</td>
<td>512</td>
</tr>
<tr>
<td>16</td>
<td>-254.00</td>
<td>-254.00</td>
<td>512.0000000000000000000000000000</td>
<td>512.0000000000000000000000000000</td>
<td>512</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>23</td>
<td>-254.00</td>
<td>-254.00</td>
<td>512.0000000000000000000000000000</td>
<td>512.0000000000000000000000000000</td>
<td>512</td>
</tr>
<tr>
<td>24</td>
<td>-254.00</td>
<td>-254.00</td>
<td>512.0000000000000000000000000000</td>
<td>512.0000000000000000000000000000</td>
<td>512</td>
</tr>
<tr>
<td>25</td>
<td>-254.00</td>
<td>-254.00</td>
<td>512.0000000000000000000000000000</td>
<td>512.0000000000000000000000000000</td>
<td>512</td>
</tr>
</tbody>
</table>

REFERENCES