# An iteration for the k-th root with cubic convergence

The other day I encountered a formula for f such that the iteration $y := f.k.a.y$ converged cubically to $y = \sqrt[k]{a}$. Cubic convergence means that in the transition from one estimate to the next, the number of correct leading digits is roughly tripled: a relative error $\delta$ leads to a relative error of the order of $\delta^3$ for the next approximation. This note is not about the merits of this iteration or the lack thereof, but on how it could have been derived. When I saw the formula, it was patently obvious that $\sqrt[k]{a}$ was a fixpoint of $f.k.a$, the cubic convergence was not obvious at all, and the formula was, as far as I was concerned, a Big Rabbit. The purpose of this note is to do something about these latter issues.

* * *

Let $\alpha$ denote our target value for $y$, i.e. $\alpha = \sqrt[k]{a}$. Then, as remarked, $\alpha$ has to be a fixpoint of $f.k.a$. To eliminate the unknown $\alpha$ from that requirement, we observe for any $f, k, a, y$ and $\alpha$:

$\alpha$ is fixpoint of $f.k.a$

$\equiv$ { definition of fixpoint }

$\alpha = f.k.a.\alpha$

$\Leftarrow$ { Leibniz Principle }

$y = \alpha \ \wedge \ y = f.k.a.y$

$\equiv$ { because we don't consider negative or complex values, $y = \alpha \equiv y^k = a$ }

$y^k = a \ \wedge \ y = f.k.a.y$ .

When $y^k = a$, the above says that

$$y := f.k.a.y$$

should leave $y$ unchanged, or – in terms of familiar operators plus and times – should increase $y$ by $0$ or should multiply $y$ by $1$. For positive $y$ these alternatives are equivalent, we choose the multiplicative version, more precisely we rewrite

$$f.k.a.y = y \cdot g.k.a.y^k$$

where $g$ has the property

$$a = b \Rightarrow g.k.a.b = 1 \quad .$$

In this last requirement the constant $1$ was introduced because it is the neutral element of the multiplication. For the same reason, the requirement is trivially met by choosing

$$g.k.a.b = \frac{a}{b} \quad \text{or} \quad g.k.a.b = \frac{b}{a} \quad ,$$

but neither gives a converging iteration: for an approximation $y = \alpha \cdot (1+\delta)$ they lead to next approximations

$$y = \alpha \cdot (1 - (k-1) \cdot \delta + \ldots) \quad \text{and} \quad y = \alpha \cdot (1 + (k+1) \cdot \delta + \ldots)$$

which are not actually improvements. (With either choice, $y = \alpha$ would be a fixpoint, but the iteration would be unstable.)

<u>Note</u> "+..." is short for "plus higher powers of $\delta$", which itself is short for something else. (End of Note.)

The question is whether we can reach our goal by suitably "averaging" the last two choices for $g$ considered. To get enough freedom I propose to "average" denominators and numerators separately, and to consider

$$g.k.a.b = \frac{p \cdot a + q \cdot b}{r \cdot a + s \cdot b}$$

for suitable $p, q, r, s$ . Since these are 4 homogenous parameters, they give us only 3 degrees of freedom, but that is enough to aim for cubic convergence.

More precisely, we are now considering the iteration

$$y := y \cdot \frac{p \cdot a + q \cdot y^k}{r \cdot a + s \cdot y^k} \quad .$$

To study its convergence we substitute $y := \alpha \cdot (1+\delta)$ in the right-hand side and remove all canceling factors $\alpha$. One ends up studying

$$c_0 + c_1 \cdot \delta + c_2 \cdot \delta^2 + \ldots$$

when these are the first terms of the Taylor expansion in $\delta$ of

$$\frac{(1+\delta) \cdot (p + q \cdot (1+\delta)^k)}{r + s \cdot (1+\delta)^k}$$

---

The requirements $c_0 = 1$ , $c_1 = 0$ , $c_2 = 0$ lead in turn to

(0) $\quad p + q = r + s \neq 0$

(1) $\quad p + q + qk - sk = 0 \qquad$ (using (0))

(2) $\quad qk(k+1) - sk(k-1) = 0 \qquad$ (using (0)) .

With $k \geq 1$ , (2) gives $q = k-1$ , $s = k+1$ , (1) then gives $p = k+1$ , and (2) $r = k-1$. The iteration with cubic convergence is

$$y := y \cdot \frac{(k+1) \cdot a + (k-1) \cdot y^k}{(k-1) \cdot a + (k+1) \cdot y^k}$$

* * *

Looking for other things, I encountered this iteration at the end of a letter to my parents of August 1951. I had added that I was delighted by this discovery, but that was all: no hint of a proof of the cubic convergence and no indication of how (or why) I had derived this formula. So I set myself the task of finding a way in which that could be done. I am sure that at the time I did it differently.

Austin, 29 January 1999

prof. dr Edsger W. Dijkstra
Department of Computer Sciences
The University of Texas at Austin
Austin, TX 78712-1188
USA