Continued Powers and a Sufficient Condition for Their Convergence

DIXON J. JONES
5112 Fairchild Ave.
Fairbanks, AK 99709-4523

Background Continued powers are the conceptual ancestors and historical descendants of *continued square roots*, which are expressions of the form

$$\lim_{k\to\infty} x_0+\sqrt{x_1+\sqrt{x_2+\sqrt{\cdots+\sqrt{x_k}}}}\ .$$

Examples of continued square roots have been around for quite a while; Pólya and Szegö in 1916 considered their convergence properties [3], and relevant problems date back to the late 19th century. A. Herschfeld's 1935 paper on continued square roots and similar expressions (he called them "infinite radicals" [1]) contained many interesting observations, among them two in particular. First, Herschfeld proved that a continued square root of real nonnegative terms x_n converges if, and only if, $(x_n)^{2^{-n}}$ is bounded. (This was independently discovered some 50 years later by Sizer, who also coined the term "continued square root" [4]). Second, Herschfeld noted in passing that the general form

$$\lim_{k\to\infty} x_0+\Big(x_1+\Big(x_2+\big(\cdots+(x_k)^p\cdots\big)^p\Big)^p\Big)^p, \tag{1}$$

includes not only continued square roots ($p=1/2$), but infinite series and continued fractions as well ($p=1$ and $p=-1$, respectively). Herschfeld's investigation of expression (1) extended his continued square roots work to arbitrary roots; here is a somewhat restricted version of his generalization.

HERSCHFELD'S CONVERGENCE THEOREM. *For real nonnegative terms x_n and real $p, 0<p<1$, the expression* (1) *converges if, and only if, $(x_n)^{p^n}$ is bounded.*

Definitions, notation, and the main result The terms "continued fraction" and "continued root" are accepted, in spite of a certain slipperiness about the adjective "continued." Allowing for some syntactic guilt by association, therefore, we call expression (1) a *continued (pth) power*, especially when $p>1$. In imitation of the continued root notation, we adopt the alternate notation ${}^p(x)$ for exponentiation and write continued powers as

$$\lim_{k\to\infty} x_0+{}^p\big(x_1+{}^p\big(x_2+{}^p\big(\cdots+{}^p(x_k)\cdots\big)\big)\big),$$

or sometimes more informally as

$$x_0+{}^p\big(x_1+{}^p\big(x_2+{}^p(\cdots)\big)\big).$$

We abbreviate a continued power with the notation $C_{k=0}^{\infty}(p,x_k)$, and use C_0^{∞} as a nickname when p and the sequence $\{x_n\}$ of terms are understood. The finite expression

$$x_0+{}^p\big(x_1+{}^p\big(x_2+{}^p\big(\cdots+{}^p(x_n)\cdots\big)\big)\big)$$

we call the *nth approximant* of the continued power, shortened likewise to $C_{k=0}^{n}(p, x_k)$ and C_0^n. Just as an infinite series is rigorously defined as the limit of its partial sums, it is important to think of a continued power as the limit of its sequence of approximants, especially in view of certain difficulties encountered in generating an approximant from its predecessor.

Can continued powers converge for positive values of p not covered by Herschfeld's Theorem? For $p = 1$ the question has been, shall we say, extensively answered in the affirmative. What can we say for $p = 2$, 100, or even $10^{10^{10}}$? Some results are known for the case $p > 1$, including a ratio test strikingly reminiscent of d'Alembert's test for the convergence of series **[2]**. Here we prove a condition, analogous to Herschfeld's Theorem, that is sufficient for the convergence of a continued power when $p > 1$.

THEOREM. *For real $p > 1$, the continued pth power with real nonnegative terms x_n converges if $(x_n/R)^{p^n}$ is bounded, where $R = (p-1)/p^{p/(p-1)}$.*

Examples and amplifications To get the feel of the theorem and its elements, a few examples are helpful. Some properties of continued powers and their approximants, assumed or referred to casually in this section, receive more careful scrutiny in the proof of the theorem.

I. Perhaps the simplest kind of example is the continued square ($p = 2$) having nonnegative, constant terms:

$$C_{n=0}^{\infty}(2, c) = c +^2 \left(c +^2 \left(c +^2 (\cdots)\right)\right), c \geq 0.$$

For $p = 2$ we have $R = 1/4$. The theorem guarantees convergence if $(x_n/R)^{2^n} = (4c)^{2^n}$ is bounded, so any c between 0 and 1/4, inclusive, will work. Unlike most continued powers, the limit of this continued square is easily found. If it converges (*i.e.*, if $0 \leq c \leq 1/4$), we may write

$$C_{n=0}^{\infty}(2, c) = c +^2 (C_{n=0}^{\infty}(2, c)),$$

and solve the equation as a quadratic to obtain $C_{n=0}^{\infty}(2, c) = \frac{1}{2} \pm \frac{1}{2}\sqrt{1-4c}$. (Note the corroborative evidence under the radical for $0 \leq c \leq 1/4$.) The choice of "plus" or "minus" is resolved as follows: When c is at its maximum of 1/4, the continued square has a limit of 1/2. We expect the limit for smaller values of c to be less than 1/2, hence

$$C_{n=0}^{\infty}(2, c) = \tfrac{1}{2} \pm \tfrac{1}{2}\sqrt{1-4c}\,.$$

An immediate corollary to the theorem can be gleaned from this example: *For $p > 1$, a continued pth power of nonnegative constant terms c converges if $0 \leq c \leq R$.* In fact it can be shown that the continued power *diverges* if $c > R$. More about necessary conditions for convergence can be found in **[2]**. Now, it happens that as p grows, R approaches 1. Hence the "interval of convergence" for a continued power of constants grows larger as p increases.

II. Moving up the scale of complexity, the continued square

$$C_{n=0}^{\infty}(2, 2^{2^{-n}-2}) = 1/2 +^2 \left(\frac{1}{\sqrt{2^3}} +^2 \left(\frac{1}{\sqrt[4]{2^7}} +^2 \left(\frac{1}{\sqrt[8]{2^{15}}} +^2 (\cdots)\right)\right)\right)$$

converges, since

$$(x_n/R)^{2^n} = (2^{2^{-n}})^{2^n} = 2$$

for all n. Or *did* we move up the scale of complexity? Because this example converges, we may multiply by $2/2$ and distribute the denominator to the right:

$$C_{n=0}^{\infty}(2,2^{2^{-n}-2}) = 2/2\left[1/2 +^2\left(\frac{1}{\sqrt{2^3}} +^2\left(\frac{1}{\sqrt[4]{2^7}} +^2\left(\frac{1}{\sqrt[8]{2^{15}}} +^2(\cdots)\right)\right)\right)\right]$$

$$= 2\left[1/4 +^2\left(1/4 +^2\left(1/4 +^2\left(1/4 +^2(\cdots)\right)\right)\right)\right].$$

It's just a constant multiple of Example I for $c = 1/4$. The value of this continued square is $2(1/2) = 1$.

III. For an example that really *is* different, and shows the limitations of a convergence condition that is sufficient but not necessary, consider

$$C_{n=0}^{\infty}(2,(4^n+1)/4^{n+1}) = 1/2 +^2\left(5/16 +^2\left(17/64 +^2(\cdots)\right)\right).$$

This fails the test of the theorem: $(x_n/R)^{2^n} = (1 + 1/4^n)^{2^n}$ is unbounded as $n \to \infty$. Yet this continued square does converge. Each term after the zeroth is less than the corresponding term in Example II. (Consider it an exercise to show why.) We can compare corresponding approximants in Examples II and III to see that III's approximants are increasing and yet smaller than II's, and thereby conclude that because II converges, so does III.

Proof of the theorem The gist of the proof is that the approximants of a continued power are nondecreasing, and are bounded by the limit of a convergent iterated map. The only trick needed is a careful justification for "reversing" the associativity of an approximant when it suits our purpose. (For proving the theorem's sufficiency we would be technically correct in assuming $R = (p-1)/p^{p/(p-1)}$. An expression so exotic, however, is probably better appreciated when unearthed, rather than unveiled. Therefore, abandoning technicality for pedagogy (and a little drama), we assume only that $R > 0$, and derive its value in terms of p as a consequence of the mechanics of the proof.)

Let us begin the proof by reading the condition of the theorem more precisely. Suppose there is a real $M > 0$ and an integer $N \geq 0$ such that $(x_n/R)^{p^n} < M$ for $n \geq N$. Using the equivalent expression $x_n < RM^{p^{-n}}$, we begin construction of a continued power's nth approximant, proceeding right-to-left as the associativity of the form suggests, and freely using the convention ${}^p(x) = x^p$:

$$\begin{aligned}
{}^p(x_n) &< R^p M^{p^{-n+1}}\\
x_{n-1} +^p(x_n) &< RM^{p^{-n+1}} + R^p M^{p^{-n+1}}\\
&= RM^{p^{-n+1}}(1 + R^{p-1})\\
{}^p(x_{n-1} +^p(x_n)) &< R^p M^{p^{-n+2}}\{{}^p(1 + R^{p-1})\}\\
x_{n-2} +^p(x_{n-1} +^p(x_n)) &< RM^{p^{-n+2}} + R^p M^{p^{-n+2}}\{{}^p(1 + R^{p-1})\}\\
&= RM^{p^{-n+2}}(1 + R^{p-1}\cdot{}^p(1 + R^{p-1})).
\end{aligned}$$

With n fixed, and making the substitution $r = R^{p-1}$, an induction proof on the index i $(i \leq n - N)$ shows that

$$\begin{aligned}
&x_{n-i} +^p(x_{n-i+1} +^p(\cdots +^p(x_n)\cdots))\\
&\quad < RM^{p^{-n+i}}(1 + r\cdot{}^p(1 + r\cdot{}^p(\cdots + r\cdot{}^p(1+r)\cdots))),
\end{aligned}$$

where r appears i times on the right. When $i = n - N$, the left side becomes an expression for the nth approximant, truncated at x_N:

$$C_{k=N}^{n}(p, x_k) < RM^{p^{-N}}(1 + r^{\cdot p}(1 + r^{\cdot p}(\cdots + r^{\cdot p}(1 + r)\cdots))), \qquad (2)$$

where there are now $(n - N)$ constants r on the right.

The proof focuses now on the behavior of inequality (2) as n grows without bound. We consider each side in turn, beginning with C_N^n. Although we regard a continued power as the limit of the sequence of approximants C_0^n, it suffices to consider just the sequence of truncated approximants C_N^n, because the identity

$$C_0^n = x_0 +^p (x_1 +^p (\cdots +^p (x_{N-1} +^p (C_N^n))\cdots))$$

assures us that C_0^n converges if C_N^n converges as $n \to \infty$, the former being at most finitely larger than the latter. Now, bearing in mind that $p > 1$ and $x_n \geq 0$, the inequality $x_{n-1} \leq x_{n-1} +^p (x_n)$ immediately becomes

$$C_{n-1}^{n-1} \leq C_{n-1}^{n}, \qquad n \geq 1.$$

From this auspicious beginning we can build truncated approximants on each side, term-by-term, layer-by-layer, from right to left, to obtain inequalities of the form

$$C_i^{n-1} \leq C_i^n, \qquad 0 \leq i \leq n-1.$$

These inequalities show that a sequence of truncated approximants having the same initial term is nondecreasing. In particular, as regards (2), the truncated approximants C_N^n are nondecreasing as n increases.

We now turn our attention to the right-hand side of (2), and of course what interests us here is not the constant quantity $RM^{p^{-N}}$, but the long, continued-power-like expression in which r is repeated $(n - N)$ times. One approach to unraveling such a sequence of nested operations is through the theory of *functional iteration*, the application of a function f to an appropriate input and the successive application of f to the resulting output. The *kth iterate* of f at a (meaning $f \circ f \circ \ldots f \circ f(a)$ with f repeated k times) is written $f^{(k)}(a)$, with $f^{(0)}(a) = a$ and $f^{(1)}(a) = f(a)$ by definition; the iterates obey the forward recursion formula $f^{(k)} = f(f^{(k-1)})$. The form of (2) suggests that a function $f(x) = 1 + rx^p$ might receive an initial input of, say $x_0 = 1$, then by iteration of f a total of $(n - N)$ times the entire expression might be generated:

$$f^{(1)}(1) = 1 + r,$$
$$f^{(2)}(1) = 1 + r^{\cdot p}(1 + r),$$

and so on. The behavior of this iterated function can be analyzed pretty easily; if it converges, we'll use its limit as an upper bound for inequality (2).

This is the point in the proof where associativity must be considered very carefully. Both sides of expression (2) were constructed from right to left. The left-hand side, the truncated approximant C_N^n, is possessed of an unambiguous order of operations, which can be posed in the form of a backward recursion formula:

$$C_j^n = x_j +^p (C_{j+1}^n), \qquad N \leq j \leq n-1.$$

The right-hand side of (2) is possessed of an *ambiguous* order of operations, meaning that forward and backward recursion formulas both apply. To see this, define

$$S_j^n = 1 + r^{\cdot p}(\cdots + r^{\cdot p}(1 + r)\cdots))$$

so that $(n-j)$ counts the number of times r appears, whereby for instance inequality (2) becomes

$$C_N^n \leq RM^{p^{-N}} S_N^n.$$

Then by construction, a backward recursion holds:

$$S_j^n = 1 + r^{\cdot p}\left(S_{j+1}^n\right), \qquad N \leq j \leq n-1.$$

But also, by luck, providence, and/or the absence of indices on the rs, we have the forward recursion

$$S_j^n = 1 + r^{\cdot p}\left(S_j^{n-1}\right), \qquad N \leq j \leq n-1. \tag{3}$$

Formula (3) is evidently our link to functional iteration. With $f(x) = 1 + rx^p$, interpret (3) as $S_j^n = f(S_j^{n-1})$ for $N \leq j \leq n-1$. Employing $S_N^{N+1} = 1 + r = f(1)$ as a basis for induction, we have in short order

$$S_N^n = f^{(n-N)}(1),$$

which validates our treatment of $1 + r^{\cdot p}(\cdots + r^{\cdot p}(1+r)\cdots))$ as an iterated function.

(If it seems that the preceding discourse makes a mountain of technicality out of a transparently obvious molehill, be cautioned: Already there are published accounts of continued square roots that are confused and even wrong about the associativity of infinitely nested expressions. The attempt at conciseness here follows the example set by Herschfeld, who carefully distinguished between "left infinite radicals" and "right infinite radicals"—having forward and backward recursion formulas, respectively—and found their convergence properties completely dissimilar.)

To summarize the proof thus far: The inequality (2) having been constructed for each $n \geq N$ in right-to-left fashion, we accept that its right-hand side can be interpreted as $(n-N)$ iterations of the function $f(x) = 1 + rx^p$ from the initial input $x_0 = 1$, all multiplied by the constant $RM^{p^{-N}}$. We continue now by examining the function f and its iterates more closely. Inasmuch as our entire discourse is concerned only with nonnegative real numbers; and whereas the exponent p is strictly greater than 1; and because $r = R^{p-1}$ is strictly positive, it follows by elementary calculus that the graph of $f(x) = 1 + rx^p$ must resemble one of the curves shown in FIGURE 1.

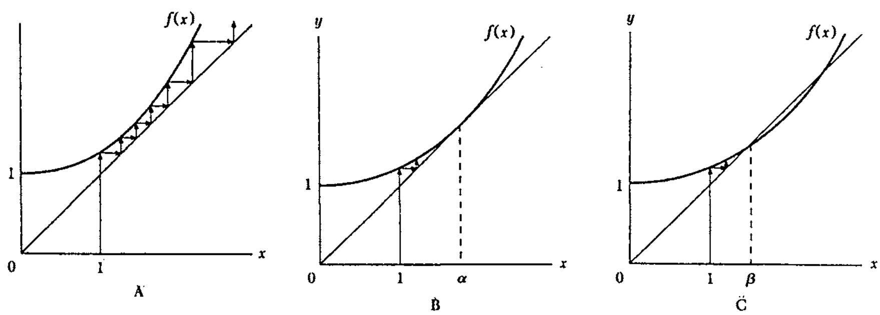

FIGURE 1

Beyond illustrating the possible graphs of f, FIGURE 1 shows in each case the iteration of f from $x_0 = 1$ in what has come to be called a "graphical analysis." From the point $(1, f(1))$, the horizontal skip to the line $y = x$ and the vertical bounce back up to $y = f(x)$ represent, respectively, the transfer of $f(1)$ from the range of f to its domain, followed by the application of f to the domain element $f(1)$ to produce $f^{(2)}(1)$. The process then continues from the point $(f(1), f^{(2)}(1))$.

With its foot anchored firmly at $(0,1)$, f stretches farther to the right as the scaling factor r decreases, and from FIGURE 1A it is clear that the iterates of f will race off to $+\infty$ until r reaches some critical value forcing f to touch the line $y=x$. When r attains this value and f is tangent to $y=x$, as in FIGURE 1B, the point of tangency (α, α) becomes the limit of the sequence of iterates. And as r drops below its critical value, f intersects $y=x$ in two points (FIGURE 1C), of which (β, β)—the left-most—is the limit point of the sequence of iterates. Note that α and β lie to the right of 1 on the x-axis because f increases from 1 on the y-axis.

Which graph optimally represents the function $f(x)=1+rx^p$ as we require it for our proof? We'd prefer that the iterates of f converge—which automatically rules out FIGURE 1A—and we'd like their limit to be the largest possible, which among the remaining choices puts FIGURE 1C out of the running. FIGURE 1B therefore gets the nod as the graph of f that best meets our needs. Given that f is tangent to $y=x$ at (α, α), clearly α is the largest number that bounds the iterates of f. Therefore α is the largest number that bounds the right side of inequality (2). We can be a bit more specific: At (α, α), we have $f'(\alpha)=1$, which yields

$$\begin{aligned}\alpha &= (rp)^{1/(1-p)}\\ &= p^{1/(1-p)}/R. \end{aligned} \tag{4}$$

Equation (2) thereby becomes well-disciplined:

$$\begin{aligned} C_{k=N}^{n}(p, x_k) &< RM^{p^{-N}}\left(1+r^{\cdot p}\left(1+r^{\cdot p}\left(\cdots+r^{\cdot p}(1+r)\cdots\right)\right)\right)\\ &< RM^{p^{-N}}\left(p^{1/(1-p)}/R\right)\\ &= M^{p^{-N}}p^{1/(1-p)}. \end{aligned}$$

The last quantity above is composed entirely of known constants, and the truncated approximant C_N^n is therefore bounded for arbitrary $n \geq N$. Being increasing and bounded, the sequence of truncated approximants is convergent, and thus the continued power is convergent.

The only loose end to be taken care of is verifying the value of R given in the theorem. At (α, α) it is true that $f(\alpha)=\alpha$, so by substitution for α from equation (4) we get

$$1+r(rp)^{p/(1-p)}=(rp)^{1/(1-p)},$$

which can be solved for r as follows:

$$\begin{aligned} 1+r^{1/(1-p)}p^{p/(1-p)} &= r^{1/(1-p)}p^{1/(1-p)}\\ r^{1/(1-p)}p^{1/(1-p)}(1-1/p) &= 1\\ r &= (1/p)[(p-1)/p]^{p-1}\\ &= (p-1)^{(p-1)}/p^p. \end{aligned}$$

Since $r=R^{p-1}$, we conclude that $R=(p-1)/p^{p/(p-1)}$, and the proof is complete.

REFERENCES

1. Aaron Herschfeld, On infinite radicals, *Amer. Math. Monthly* 42 (1935), 419–429.
2. Dixon J. Jones, Continued powers and roots, *Fibonnaci Quarterly* 29 (1991), 37–46.
3. G. Pólya and G. Szegö, *Problems and Theorems in Analysis*, Vol. 1, Springer-Verlag, New York, 1972.
4. Walter S. Sizer, Continued roots, this MAGAZINE 59 (1986), 23–27.

NOTE: An error in Example III was noted in a Letter to the Editor by J. Nichols-Barrer, **69** (3) June 1996, p. 238. Following is my correction, **69** (4) Oct. 1996, p. 316.

Allendoerfer Awards, in particular, for their generous recognition.

As mathematicians one of our great pleasures is working in collaboration to solve a challenging problem. Sharing our results and, whenever possible, explaining how they were discovered is an equal joy. Writing "Permutations and Combination Locks" gave me the opportunity to enjoy each of these pleasures. In addition, Dan Velleman and I tried to provide our students with an accessible model of mathematical research which we hope will encourage them to undertake their own investigations. I look forward to those investigations and the opportunity to share them with my colleagues in the MAA.

Letters to the Editor

Dear Editor:

I thank Josh Nichols-Barrer (Letters to the Editor, June 1996, p. 238) for bringing to light an error in my article *Continued powers and a sufficient condition for their convergence* (this MAGAZINE, December 1995, pp. 387–392). He points out that since it does not in fact violate my convergence condition for continued squares, my Example III doesn't show that the condition for general powers $p > 1$ is not necessary.

As my penance for publicly transgressing first-year calculus, I offer the following replacement for the lightly-conceived and ill-fated Example III. Consider the continued square

$$S = b+{}^2(0+{}^2(b+{}^2(0+{}^2(b+{}^2(0+{}^2(\ldots))))))$$

with $b = 3/(4^{4/3})$. This fails the convergence test for a continued square. With $p = 2$, we have $R = (p-1)/p^{p/(p-1)} = 1/4$, $x_n = b$ for n even and 0 for n odd, and

$$\left(\frac{x_n}{R}\right)^{p^n} = \begin{cases} [3/(4^{1/3})]^{2^n} & n \text{ even;} \\ 0 & n \text{ odd.} \end{cases}$$

The dominant subsequence of even terms results in an unbounded expression, and the test fails.

However, S is equivalent to the continued fourth power

$$b + {}^4(b + {}^4(b + {}^4(\ldots))),$$

which converges by the boundedness test: with $p = 4$, one has $R = 3/(4^{4/3})$ and $(x_n/R)^{p^n} = 1^{4^n} = 1$. The continued square S therefore converges, but since it fails the continued squares convergence test, the test remains sufficient but not necessary.

Dixon J. Jones
5112 Fairchild Avenue
Fairbanks, Alaska 99709–4523