Big-O Notation and
Problem A2 of the 2001 Putnam Exam

John McCuan

August 28, 2022

Original Statement: For every positive real number x, let

$$g(x) = \lim_{r\to 0}[(x+1)^{r+1} - x^{r+1}]^{1/r}.$$

Find

$$\lim_{x\to\infty} \frac{g(x)}{x}.$$

The second posted solution of this problem starts out by replacing the "base part" of the limiting quantity in the definition of the function g as follows:

$$(x+1)^{r+1} - x^{r+1} = (r+1)x^r + O\left(x^{r-1}\right). \tag{1}$$

Let me start by explaining what is meant by this. Given real valued functions p and q defined in an open interval (a,b) with $t_0 \in (a,b)$, one says

$p(t) = O(q(t))$ **as** t **limits to** t_0

if the quotient $p(t)/q(t)$ remains bounded as t tends to t_0. More precisely, there is some $M > 0$ and some $\epsilon > 0$ such that

$$\left|\frac{p(t)}{q(t)}\right| \le M \qquad \text{for all } t \text{ satisfying } 0 < |t - t_0| < \epsilon.$$

Notice the positivity of $|t - t_0|$ rules out the possibility $t = t_0$, so the condition says nothing about the values of p and q at the point $t = t_0$. Furthermore, it is assumed there is some punctured neighborhood $0 < |t - t_0| < \epsilon$ of t_0 on which $q(t) \neq 0$.

This may perhaps be considered the "basic" definition of big-O notation.

The first thing you may note is that specification of the variable to which the big-O notation applies, e.g., something like **as** t **limits to** t_0, is missing from the posted solution. Furthermore, since there are two variables in play (x and r) one needs to know which one the author has in mind. The following suggestion of the use of Taylor's theorem gives at least a little bit of context, and indeed it turns out (I'm pretty sure) the variable referenced is x. Thus, I think (1) means

$$(x+1)^{r+1} - x^{r+1} - (r+1)x^r = O\left(x^{r-1}\right) \qquad \text{as } x \to \infty. \tag{2}$$

But now there's a problem with the basic definition because it requires a finite limit $t_0 \in (a, b) \subset \mathbb{R}$. Here is the relevant modifiction to suit the desired application:

Definition 1 *Given real valued functions p and q defined in an open interval $(a, \infty) \subset \mathbb{R}$, one says*

$p(t) = O(q(t))$ **as** t **limits to** ∞

if the quotient $p(t)/q(t)$ remains bounded as t tends to ∞. More precisely, there is some $M > 0$ and some $N > 0$ such that

$$\left|\frac{p(t)}{q(t)}\right| \le M \qquad \textit{for all } t > N.$$

Here it is assume $q(t) \neq 0$ for $t > N$ as well.

Translating (2) into what it actually means then: The author first claims there is some $M > 0$ and some $N > 0$ such that

$$\left|\frac{(x+1)^{r+1} - x^{r+1} - (r+1)x^r}{x^{r-1}}\right| \le M \qquad \text{for all } x > N, \tag{3}$$

Before we address why the author wrote this, much less the validity of the claim, let's check to see if this is plausible. First note that $(x+1)^{r+1} - x^{r+1} > 0$ for $x > 0$. The same applies to the denominator $x^{r-1} > 0$ for $x > 0$. It might not be entirely clear that the full numerator $(x+1)^{r+1} - x^{r+1} - (r+1)x^r$ is positive, but that turns out to be true too. Because of this, we can take away the absolute values and write the main inequality of the claim as

$$(x+1)^{r+1} - x^{r+1} - (r+1)x^r \le Mx^{r-1}.$$

Here is probably what the author is thinking: *There is a generalized binomial expansion (series) for the first term on the left side.* This series looks like

$$(x+1)^{r+1} \sim x^{r+1} + (r+1)x^r + \frac{r(r+1)}{2}x^{r-1} + \frac{(r-1)r(r+1)}{3!}x^{r-2} + \cdots. \tag{4}$$

Do you see the pattern of the terms? It may also help to compare to the standard binomial expansion

$$(a+b)^n = a^n + na^{n-1}b + \frac{n(n-1)}{2}a^{n-2}b^2 + \cdots + b^n = \sum_{j=0}^{n} \left(\begin{array}{c} n \\ j \end{array} \right) a^{n-j}b^j.$$

This standard binomial expansion is easy to prove by induction. The expansion series of $(a+b)^\alpha$ where $a, b, \alpha > 0$ was discovered by Isaac Newton, and if I'm not mistaken the resulting series is known as **Newton's binomail expansion**. Formally,

$$(a+b)^\rho \sim \sum_{j=0}^{\infty} \frac{\rho(\rho-1)\cdots(\rho-j-1)}{j!}a^{\alpha-j}b^j. \tag{5}$$

The series actually converges, so "$\sim$" in (5) can be replaced by "=," but Newton almost certainly didn't prove this. He was too busy trying to find a clever way to put fig jam in cookies, and also Weierstrass didn't formulate what it would actually mean to prove convergence for another couple hundred years (after Newton's clever discoveries). In any case, we really don't need any of these binomial expansion details to work this problem. They can be grouped together under the heading of "heuristics," but they are fun to know about, and there is no doubt that if you do know about Newton's expansion, then the quantity on the left in (2) can be written as

$$(x+1)^{r+1} - x^{r+1} - (r+1)x^r = \frac{r(r+1)}{2}x^{r-1} + \frac{(r-1)r(r+1)}{3!}x^{r-2} + \cdots$$

where each of the terms on the right has a factor of x^{r-1}. That is,

$$\begin{aligned}(x+1)^{r+1} - x^{r+1} - (r+1)x^r &= \left(\frac{r(r+1)}{2} + \frac{(r-1)r(r+1)}{3!}x^{-1} + \cdots \right) x^{r-1} \\ &\le \left(\frac{r(r+1)}{2} + \frac{(r-1)r(r+1)}{3!} + \cdots \right) x^{r-1}.\end{aligned}$$

In particular, if you believe the quantity on the left is positive (which we have not yet established) and the binomial series converges, then you can take

$$M_1 = \frac{r(r+1)}{2} + \frac{(r-1)r(r+1)}{3!} + \cdots,$$

and the claim (3) certainly holds.

Exercise 1 *Assuming Newton's binomial expansion converges (to the binomial expression Newton said it represented), what is*

$$M_1 = \frac{r(r+1)}{2} + \frac{(r-1)r(r+1)}{3!} + \cdots?$$

At this point, I hope we understand (3) of the posted solution. The thing is the author immediately claims something much stronger. Here is the real claim needed for this solution:

Lemma 1 $f(x) = (x+1)^{r+1} - x^{r+1} - (r+1)x^r$ *satisfies*

$$|f(x)| \le (r+1)rx^{r-1} \qquad \text{for } x \ge 1 \text{ and } r \le 1. \tag{6}$$

This may be viewed as a statement $f(x) = O(x^{r-1})$ with an explicit bound of $M = (r+1)r$ depending on r and with range of applicability $r \le 1 \le x$. But really it's better viewed as simply a statement about the function $\phi(x,r) = (x+1)^{r+1} - x^{r+1} - (r+1)x^r$ as a function of two variables. It is suggested that we use "Taylor's theorem with remainder." Before we explore what that might mean, let's see if it is reasonable to even imagine this is correct. In Figure 1 I have plotted $\phi(x,r)$ as a function of x for three values of r with $0 < r < 1$. This suggests the function ϕ is positive and

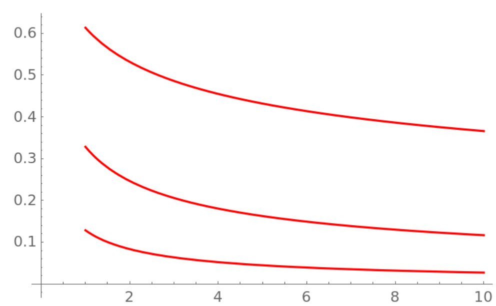

Figure 1: Plots of ϕ as a function of x.

decreasing in x. We will verify this later. In Figure 2 I have plotted each of the curves from Figure 1 along with the proposed upper bound $m(x,r) = (r+1)rx^{r-1}$. We know m is decreasing in x and limits to $m = 0$ as $x \to \infty$. The plots in Figure 2 suggest this stronger version of the author's claim:

Lemma 2 $f(x) = (x+1)^{r+1} - x^{r+1} - (r+1)x^r$ *satisfies*

$$0 < (x+1)^{r+1} - x^{r+1} - (r+1)x^r < (r+1)rx^{r-1} \qquad \text{for } x \ge 1 \text{ and } r \le 1. \tag{7}$$

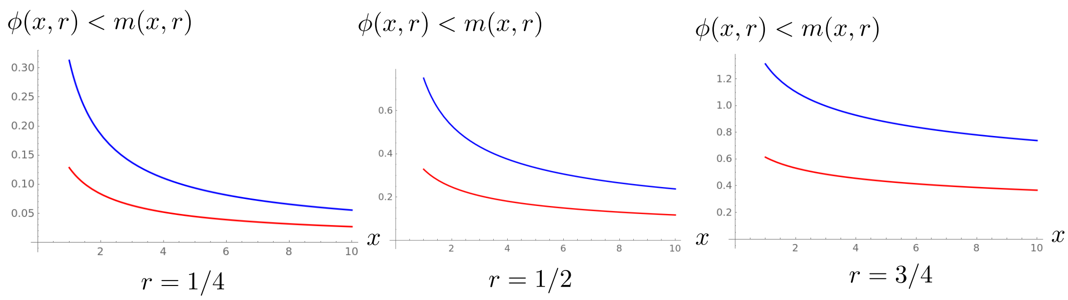

Figure 2: Plots of ϕ as a function of x compared to $m(x, r) = (r + 1)rx^{r-1}$.

This is, in fact, correct, and we will prove it. We will not prove

$$\frac{\partial \phi}{\partial x} < 0,$$

but this is presumably true too.

Proof of (7): It seems to me the easiest way to prove (7) is to use the **mean value theorem** which says a continuously differentiable real valued function ψ defined on a closed interval $[a, b] \subset \mathbb{R}$ with $a < b$ satisfies, for some $t_* \in (a, b)$

$$\frac{\psi(b) - \psi(a)}{b - a} = \psi'(t_*).$$

Taking $\psi(t) = t^{r+1}$, $a = x$ and $b = x + 1$, the mean value theorem gives us a point t_* with $x < t_* < x + 1$ for which

$$(x + 1)^{r+1} - x^{r+1} = (r + 1)t_*^r. \tag{8}$$

Therefore,

$$(x + 1)^{r+1} - x^{r+1} - (r + 1)x^r = (r + 1)t_*^r - (r + 1)x^r = (r + 1)(t_*^r - x^r).$$

Applying the mean value theorem a second time with $\psi(t) = t^r$, $a = x$ and $b = t_*$ we get a second point t_{**} for which $x < t_{**} < t_*$ and

$$t_*^r - x_r = rt_{**}^{r-1}(t_* - x).$$

From this we see

$$(x + 1)^{r+1} - x^{r+1} - (r + 1)x^r = r(r + 1)t_{**}^{r-1}(t_* - x).$$

On the one hand, the quantity on the right is clearly positive, so

$$0 < (x+1)^{r+1} - x^{r+1} - (r+1)x^r.$$

On the other hand, the function $\psi(t) = t^{r-1}$ is strictly decreasing (in t), at least for $r < 1$ and $x < t_{**} < t_* < x+1$. This means

$$t_{**}^{r-1}(t_* - x) < x^{r-1}$$

and

$$0 < (x+1)^{r+1} - x^{r+1} - (r+1)x^r < r(r+1)x^{r-1}$$

as desired. $\square$

Exercise 2 *What happens in the case $r = 1$?*

Exercise 3 *Can you show*

$$\frac{\partial \phi}{\partial x} < 0 \qquad \text{for } x \geq 0 \text{ and } 0 < r < 1?$$

What happens when $r = 1$?

What about using Taylor's formula with remainder? Here's one version:

Theorem 1 *If $\psi \in C^k(a,b)$, i.e., ψ is continuously differentiable k times, and $a < t_0 < b$, then for each $t \in (a,b)$ there is some t_* such that*

$$\psi(t) = \psi(t_0) + \psi'(t_0)(t-t_0) + \frac{\psi''(t_0)}{2}(t-t_0)^2 + \cdots + \frac{\psi^{(k-1)}(t_0)}{(k-1)!}(t-t_0)^{k-1} + \frac{\psi^{(k)}(t_*)}{k!}(t-t_0)^k.$$

The last term

$$R_k(t,t_0) = \frac{\psi^{(k)}(t_*)}{k!}(t-t_0)^k$$

is called the **remainder** and there are different formulas for it, but this is one of the easier ones. A good feature of this formula is that it gives an approximation for $f(t)$ by the polynomial

$$\sum_{j=0}^{k-1} \frac{\psi^{(j)}(t_0)}{j!}(t-t_0)^j$$

of order $k-1$, with an estimate that holds, in a certain sense, for all values of t in the interval (a,b) of definition. If one has uniform estimates for the k-th derivative $\psi^{(k)}(t)$

(from above and below) on the interval (a, b), then this is really true. One thing one must take into account, however, is that the point t_* appearing in the remainder formula changes when t changes. More precisely, $t_* = t_*(t, t_0)$ is a function of t and t_0, and of course k as well. One does know also that t_* is between t and t_0.

Note that the zero order expansion with remainder formula of order one is essentially the same thing as the mean value theorem:

$$\psi(t) = \psi(t_0) + \psi'(t_*)(t - t_0).$$

Therefore, we should be able to follow our proof of Lemma 2 above and replace the use of the mean value theorem with the use of Taylor's formula with remainder. It's a little bit of a weird application, however. We take the function $\psi(t) = t^{r+1}$, which is easy enough, but then we take an expansion (the zero order expansion) centered at $t_0 = x$ to get

$$t^{r+1} = x^{r+1} + (r + 1)t_*^r(t - x).$$

Then we evaluate at $t = x + 1$:

$$(x + 1)^{r+1} = x^{r+1} + (r + 1)t_*^r(1),$$

for some t_* with $x < t_* < x + 1$. This puts us right in the position we were in with (8). So it all works just fine, but if we hadn't used the mean value theorem first, it might have been difficult to know Taylor's formula needed to be centered at x and applied at $t = x + 1$.

Exercise 4 *Complete the second step/second application of the mean value theorem in the proof of Lemma 2 using Taylor's formula with remainder again in place of the mean value theorem.*

With this exercise behind us, we can probably see what the author had in mind. Take the first order Taylor formula for $\psi(t) = t^{r+1}$ with center of expansion at $x \geq 1$. If we do this, we get

$$t^{r+1} = x^{r+1} + (r + 1)x^r(t - x) + \frac{r(r + 1)t_*^{r-1}}{2}(t - x)^2.$$

Evaluating at $t = x + 1$ this becomes

$$f(x) = \phi(x, r) = (x + 1)^{r+1} - x^{r+1} - (r + 1)x^r = \frac{r(r + 1)t_*^{r-1}}{2}$$

for some t_* with $x < t_* < x+1$. If $x \geq 1$, then $t_* < x+1 \leq 2x$, so

$$0 < (x+1)^{r+1} - x^{r+1} - (r+1)x^r < r(r+1)x^{r-1}$$

as claimed. $\square$

Exercise 5 *Can you get the same estimate using the second order Taylor formula (and the third order remainder term)?*

With Lemma 2 in hand, the author of the second solution seems to want to apply some version of the Moore-Osgood theorem:

Theorem 2 *Let $h = h(x, r)$ be a real valued function of two variables on on open rectangle $U = (a, x_0) \times (r_0, b)$ with $p_0 = (x_0, r_0)$ the bottom right corner of $\overline{U}$. If*

$$\lim_{r \searrow r_0} h(x, r) = h_1(x) \quad \text{exists and is real valued for each } x \in (a, x_0), \tag{9}$$

$$\lim_{x \nearrow x_0} h(x, r) = h_2(r) \quad \text{exists and is real valued for each } r \in (r_0, b), \tag{10}$$

and the convergence to the function h_1 in (9) is uniform for $a < x < x_0$, then the three limits

$$\lim_{U \ni (x,r) \to (x_0, r_0)} h(x, r), \qquad \lim_{x \nearrow x_0} \lim_{r \searrow r_0} h(x, r), \qquad \text{and} \qquad \lim_{r \searrow r_0} \lim_{x \nearrow x_0} h(x, r)$$

all exist and are all equal (to one another).

This is about the most general form of the Moore-Osgood theorem I have been able to find, and it is not a particularly easy theorem to prove either. There are several things to note at this point. The first one is that the author of the second solution is wanting to apply a version of the Moore-Osgood theorem stated above on an unbounded region in the plane, namely for $(x, r) \in (0, \infty) \times (0, 1)$. Thus, the "bottom right corner" is not a point in the plane. In many instances, this kind of generalization is not too difficult, but I'm not entirely sure it doesn't cause some difficulty in this case.

Conjecture 1 *Let $h = h(x, r)$ be a real valued function of two variables on on open strip $U = (a, \infty) \times (r_0, b)$. If*

$$\lim_{r \searrow r_0} h(x, r) = h_1(x) \quad \text{exists and is real valued for each } x \in (a, \infty), \tag{11}$$

$$\lim_{x \nearrow \infty} h(x, r) = h_2(r) \quad \text{exists and is real valued for each } r \in (r_0, b), \tag{12}$$

and the convergence to the function h_1 in (11) is uniform for $a < x < \infty$, then the limits

$$\lim_{x \nearrow \infty} \lim_{r \searrow r_0} h(x,r), \qquad \text{and} \qquad \lim_{r \searrow r_0} \lim_{x \nearrow \infty} h(x,r)$$

both exist and are all equal to each another.

The second thing to note is that even if this result is correct, the author of the second solution has not verified the hypotheses. In order to apply the conjecture we would have to go all the way back to the original statement of the problem and take

$$h(x,r) = \frac{1}{x}\left[(x+1)^{r+1} - x^{r+1}\right]^{1/r}$$

which we can write as

$$h(x,r) = \left[r + 1 + \frac{\phi(x,r)}{x^r}\right]^{1/r}$$

where

$$\phi(x,r) = (x+1)^{r+1} - x^{r+1} - (r+1)x^r$$

is defined as above. In view of Lemma 2, the limit as $x \nearrow \infty$ for r fixed exists:

$$\lim_{x \nearrow \infty} h(x,r) = (r+1)^{1/r}.$$

Also, the limit as $r \searrow 0$ for x fixed exists:

$$\lim_{r \searrow 0} h(x,r) = \frac{g(x)}{x}. \tag{13}$$

The limit g is more or less assumed to exist in the original statement of the problem, but this is also shown in the first solution. But it is now required that we show the convergence of (13) is uniform in x, that is,

$$\lim_{r \searrow 0} \left[r + 1 + \frac{\phi(x,r)}{x^r}\right]^{1/r} = \frac{g(x)}{x}$$

with the convergence being uniform in x. We have that

$$\lim_{r \searrow 0} \left[r + 1 + \frac{\phi(x,r)}{x^r}\right] = 1$$

with the convergence being uniform in x, but this does not seem to me to be the same thing, and the power $1/r$ seems to quite obviously make a pretty big difference.

Thus, my final evaluation is that the author of the second solution has made many correct assertions, but has not given adequate justification for changing the order of the limits.

Finally, there is no doubt that if one changes the order of the limits one gets

$$\lim_{r \searrow 0} \lim_{x \nearrow \infty} h(x, r) = \lim_{r \searrow 0} (r + 1)^{1/r} = e$$

which is the correct answer.

Exercise 6 *Can you prove the Moore-Osgood theorem?*

Exercise 7 *Can you prove the generalized Moore-Osgood theorem stated as Conjecture 1 above?*

Exercise 8 *Can you show the convergence in (13) is uniform in x?*