Cadernos do IME - Série Matemática *DOI: 10.12957/cadmat.2020.52532*
***e-ISSN:**2236-2797 (online)* *Recebido em 07 Jul. 2020 | Aceito em 31 Ago. 2020*

A NOTE ON THE CONVERGENCE OF ITERATED INFINITE EXPONENTIALS OF REAL NUMBERS

SOBRE A CONVERGÊNCIA DE EXPONENCIAIS INFINITAS DE NÚMEROS REAIS

ANDRÉ LUIZ CORDEIRO DOS SANTOS[a]
ALEXANDRE DE SOUZA SOARES[b] PATRÍCIA NUNES DA SILVA[c]

Resumo

Nesta nota, discutimos um erro capcioso na solução de um problema que envolve exponenciação infinita. Analisamos quando uma sequência infinita de exponenciais iteradas de um número real converge. A questão é motivada pelo fato de que a exponenciação infinita geralmente aparece no contexto em que problemas desafiadores são apresentados a estudantes de Matemática sem muita preocupação com a convergência. Para alcançar uma compreensão mais profunda deste tema, discutimos um teorema provado originalmente por Knoebel [1] porém com diferentes enfoques e métodos. Diferentemente da prova de Knoebel, a presente nota utiliza apenas funções de uma variável.

Palavras-chave: Tetração, Sequências, Convergência.

Abstract

In this note, we deal with a tricky mistake on solving a problem involving infinite exponentiation. We analyze the question of when an infinite sequence of iterated exponentials of a real number converges. Our motivation is that infinite exponentiation often appears in the context of 'challenge' problems presented to students of Mathematics without much regard for convergence. Looking for a deeper understanding of it, a theorem proved initially by Knoebel in [1] is established, however, with different focus and methods. In contrast to Knoebel's proof, the present note makes use only of one-variable functions.

Keywords: Tetration, sequences, convergence.

MSC2010: 40A05, 26Axx

[a]Centro Federal de Educação Tecnológica Celso Suckow da Fonseca, Rio de Janeiro, Brasil; ORCID: https://orcid.org/0000-0001-6381-0527 **E-mail:** andre.santos@cefet-rj.br

[b]Centro Federal de Educação Tecnológica Celso Suckow da Fonseca, Rio de Janeiro, Brasil; ORCID: https://orcid.org/0000-0002-2679-9602 **E-mail:** alsoares@gmail.com

[c]Universidade do Estado do Rio de Janeiro, Rio de Janeiro, Brasil; ORCID: https://orcid.org/0000-0002-1852-7746 **E-mail:** nunes@ime.uerj.br

1. Introduction

Commonly, the equation

$$x^{x^{x^{\cdot^{\cdot^{\cdot}}}}} = 2 \tag{1}$$

is presented as a challenge for students and enthusiasts of Mathematics. Forgoing formality, an argument often used to address the problem is to rewrite (1) as

$$p^q = 2,$$

where $p = x$ and $q = x^{x^{x^{\cdot^{\cdot^{\cdot}}}}}$. This is simply $p^2 = 2$, seeing as $q = 2$ due to (1). Thus we have $x = \sqrt{2}$. This gives rise to the beautiful (albeit not rigorous) identity

$$\sqrt{2}^{\sqrt{2}^{\sqrt{2}^{\cdot^{\cdot^{\cdot}}}}} = 2 \tag{2}$$

Still disregarding the precise meaning of equation (1) and of identity (2), it is possible to repeat the argument in order to solve the general case

$$x^{x^{x^{\cdot^{\cdot^{\cdot}}}}} = L, \tag{3}$$

in which L is a positive real number. Analogously, it follows that $x = L^{1/L}$ is the solution of (3). A brief consideration of this general solution shows that if $L = 4$, then again $x = \sqrt{2}$ is the solution of equation (3), from whence it follows that

$$\sqrt{2}^{\sqrt{2}^{\sqrt{2}^{\cdot^{\cdot^{\cdot}}}}} = 4. \tag{4}$$

This shows that there is something not quite right, considering that identities (2) and (4) contradict each other. This should raise an alarm, especially to beginners in Mathematics, that a careful analysis of the previous argument is desired, or, which amounts to the same, that one cannot afford to neglect mathematical rigor. This should not, nevertheless, discourage the pursuit of intuitive ideas, seeing that as much as the above manipulation may be informal, it is responsible for revealing the existence of a problem that begs further analysis.

Problems that involve iterated *infinite exponentiation* (also called infinite tetration) similar to the one presented in equation (1) are not new. A rich and detailed reference is [1], which presents interesting historical accounts on the evolution of the subject besides the treatment of this problem. Our approach to the problem is based on the analysis of certain one-variable functions and the properties of real sequences. This is a remarkable distinction that sets the present note apart from the aforementioned [1], which makes substantial use of analysis of functions on several variables.

The central motivation for this note is to characterize the values of $L > 0$ for which the equation (3) has a solution. In this direction, the first step is to give a precise definition of what it means for an x to be a solution of (3). We do this in Section 2, where we also state our main theorem. In Section 3, we pursue the steps necessary to establish the main theorem.

2. Formalization of the problem

For each $x > 0$, consider the sequence $\{a_n(x)\}_{n\in\mathbb{N}}$ recursively defined by

$$a_n(x) = \begin{cases} x, & (n = 1) \\ x^{a_{n-1}(x)}, & (n \geq 2) \end{cases} \tag{5}$$

We shall let the notation $x^{x^{x^{\cdot^{\cdot^{\cdot}}}}}$ represent the limit of the sequence $\{a_n(x)\}_{n\in\mathbb{N}}$ when it exists; thus, $x > 0$ will be a *solution* of equation (3) when the sequence $\{a_n(x)\}_{n\in\mathbb{N}}$ converges to L.

Our main theorem is the following.

Theorem 2.1. *The equation*

$$x^{x^{x^{\cdot^{\cdot^{\cdot}}}}} = L$$

has a solution (*in the sense defined above*) *if and only if* $e^{-1} \leq L \leq e$. *In this case,* $x = L^{1/L}$.

The latter claim of the theorem is straightforward to prove.

Proposition 2.2. *If* $x^{x^{x^{\cdot^{\cdot^{\cdot}}}}} = L$ *has a solution* x*, then* $x = L^{1/L}$.

Proof. Assume that the sequence $\{a_n(x)\}_{n\in\mathbb{N}}$ converges to a limit L, then take the limit on both sides of the recurrence relation (5); since exponentiation is continuous, the right-hand side becomes x^L, and solving for x yields the claim. □

To establish the existence part of Theorem 2.1, one needs to investigate among the values of x which ones generate a convergent sequence, and what is the corresponding limit. We shall carry out this investigation for different values of x in what follows, and then summarize our findings to complete the theorem.

3. Analysis of the sequence $\{a_n(x)\}_{n\in\mathbb{N}}$.

We already established in Proposition 2.2 that in case the sequence $\{a_n(x)\}_{n\in\mathbb{N}}$ converges to L, then $x = L^{1/L}$, which is regarded as the solution of (3). The structure of this solution motivates the following lemma.

Lemma 3.1. *The function* $\varphi(y) = y^{1/y} \quad (y > 0)$ *possesses a global maximum at* $y = e$. *Moreover, the restriction of this function to the interval* $(0, e]$ *is one-to-one and onto the interval* $(0, e^{1/e}]$.

Proof. The global maximum follows immediately from the analysis of the signal of the derivative of the function φ:

$$\varphi'(y) = (\varphi(y)/y^2)(1 - \ln y).$$

The signal analysis reveals that φ is strictly increasing on the interval $(0, e]$ and strictly decreasing on the interval $[e, +\infty)$. Furthermore, the function is clearly one-to-one on $(0, e]$, since it is strictly monotonous there; the last claim of the lemma then follows observing that $(0, e^{1/e}] \subseteq (0, e]$. □

Since we know the form x needs to have when $\{a_n(x)\}_{n\in\mathbb{N}}$ converges, the previous lemma allows us to rule out some values of x for which the sequence cannot converge. This is the content of the next proposition.

Proposition 3.2. *If $\{a_n(x)\}_{n\in\mathbb{N}}$ converges to $L > 0$, then $x \leq e^{1/e}$. In particular, the sequence fails to converge for $x > e^{1/e}$.*

Proof. Proposition 2.2 tells us that $x = L^{1/L}$; by the first part of Lemma 3.1, the function φ has a global maximum at $y = e$, so $\varphi(L) \leq \varphi(e)$ for positive L. It follows that $\varphi(L) = L^{1/L} \leq e^{1/e} = \varphi(e)$, so the claim holds. □

Now we would like to know for which values of $x \leq e^{1/e}$ the sequence $\{a_n(x)\}_{n\in\mathbb{R}}$ is actually convergent. The next lemma establishes an upper bound for the sequences generated by these values of x.

Lemma 3.3. *If $0 < x \leq e^{1/e}$ we have $0 < a_n(x) \leq e$ for all $n \geq 1$. In particular, the sequence $\{a_n(x)\}_{n\in\mathbb{N}}$ is bounded.*

Proof. Let $0 < x \leq e^{1/e}$. Clearly $a_1(x) = x \leq e$. Suppose by induction that for some $k \geq 1$ we have $a_k(x) \leq e$. It follows from the recurrence relation (5) that $a_{k+1}(x) = x^{a_k(x)} \leq (e^{1/e})^{a_k(x)} \leq e$. The induction principle, then, implies that $a_n(x) \leq e$ for all $n \geq 1$. □

Remark 3.4. It follows from Proposition 3.2 and Lemma 3.3 that if $\{a_n(x)\}_{n\in\mathbb{N}}$ converges to $L > 0$, then $L \leq e$. This implies that if $L > e$, then Equation (3) has no solution. In particular, we see that the informal procedure used to obtain (4) is a fallacy (this is also a consequence of Theorem 2.1, but we do not need the full power of the theorem here).

Since the previous lemma informed us that $\{a_n(x)\}_{n\in\mathbb{N}}$ is bounded for $0 \leq x \leq e^{1/e}$, we can deduce convergence if we know that the sequence is monotonic. Although this is not true of every x on the interval, as we will see later, it is obviously true for $x \geq 1$.

Proposition 3.5. *If $1 \leq x \leq e^{1/e}$, then the sequence $\{a_n(x)\}_{n\in\mathbb{N}}$ converges.*

Proof. In this case, the sequence $\{a_n(x)\}_{n\in\mathbb{N}}$ is non-decreasing and, therefore, converges since it is bounded according to Lemma 3.3. □

So far, we have deduced some facts about $\{a_n(x)\}_{n\in\mathbb{N}}$ with relative ease. However, for $0 < x < 1$, discussing the convergence of the sequence will be more involved and shall require an understanding of the behavior of the subsequences of even and odd indexes.

From (5), we obtain

$$a_{2k-1}(x) = \begin{cases} x, & (k = 1) \\ x^{\left(x^{a_{2k-3}(x)}\right)}, & (k \geq 2) \end{cases}$$

and

$$a_{2k}(x) = \begin{cases} x^x, & (k = 1) \\ x^{\left(x^{a_{2k-2}(x)}\right)}, & (k \geq 2) \end{cases}$$

The function $\delta_x(y) = x^{(x^y)}$ $(y \in \mathbb{R})$ will be useful in the study of the convergence of the subsequences.

Lemma 3.6. *For every* $0 < x < 1$*, the function* $\delta_x(y) = x^{(x^y)}$ $(y \in \mathbb{R})$ *is strictly increasing.*

Proof. This may be verified by direct differentiation or by considering that $\delta_x(y) = \nu_x(\nu_x(y))$, where $\nu_x(y) = x^y$ $(y \in \mathbb{R})$, that is, δ_x is a composition of two strictly decreasing functions. □

Lemma 3.7. *If* $0 < x < 1$*, then the subsequence of even indexes* $\{a_{2k}(x)\}_{k\in\mathbb{N}}$ *is strictly decreasing, while the subsequence of odd indexes* $\{a_{2k-1}(x)\}_{k\in\mathbb{N}}$ *is strictly increasing.*

Proof. Lemma 3.6 yields that $a_1(x) = \delta_x(0) < \delta_x(x) = a_3(x)$. Assume by induction that for some $k \geq 1$, we have proved that $a_{2k-1}(x) < a_{2k+1}(x)$. Then again, by Lemma 3.6, it follows that

$$a_{2k+1}(x) = \delta_x(a_{2k-1}(x)) < \delta_x(a_{2k+1}(x)) = a_{2k+3}(x).$$

Thus it follows by induction that the subsequence of odd indexes is strictly increasing.

It is straightforward to conclude by an analogous argument using the function δ_x that the subsequence of even indexes is strictly decreasing. □

The function in the next definition will also be useful in the study of the convergence of the subsequences of even and odd indexes.

Definition 3.8. For each $x > 0$, define

$$H_x(y) = y - x^{(x^y)} \quad (y \in \mathbb{R}).$$

The importance of the function H_x is due to the fact stated in the following lemma.

Lemma 3.9. *If the subsequences of even and odd indexes of* $\{a_n(x)\}_{n\in\mathbb{N}}$ *are both convergent, their respective limits* (*which may be distinct*) *are zeroes of the function* H_x*.*

Proof. Let $\{a_{2k}(x)\}_{k\in\mathbb{N}}$ converge to P and $\{a_{2k-1}(x)\}_{k\in\mathbb{N}}$ converge to Q. Taking the limit on (5) for even and odd n, it follows that

$$P = x^Q \qquad \text{and} \qquad Q = x^P. \tag{6}$$

From the equations on (6), one verifies that P and Q satisfy

$$H_x(P) = H_x(Q) = 0.$$ □

Since a sequence converges if and only if the subsequences of even and odd indexes have the same limit, we are interested in knowing when these limits are equal; because the limits are zeroes of H_x, it is useful to obtain a condition that prevents this function from having multiple zeroes. The next lemmas accomplish this.

Lemma 3.10. *The function* $\psi(y) = ye^y$ $(y \in \mathbb{R})$ *admits a unique global minimum at* $y = -1$*.*

Proof. The global minimum follows immediately from the analysis of the signal of the derivative of ψ: $\psi'(y) = e^y(1 + y)$. □

Lemma 3.11. *If* $e^{-e} \leq x \leq 1$*, then* H_x *is one-to-one.*

Proof. If $x = 1$ the function H_x is trivially one-to-one, so assume $x < 1$. Denoting by H'_x the derivative of the function H_x, we have

$$H'_x(y) = 1 - (\ln x)^2 (x^y) \left(x^{(x^y)}\right).$$

Now use the change of variable $z = x^y \ln x$ to write H'_x in terms of z as

$$h_x(z) = 1 - (\ln x) z e^z,$$

that is, $h_x(z) = H'_x(y)$. Since $0 < x < 1$, we have $\ln x < 0$; this and Lemma 3.10 together yield

$$h_x(z) \geq 1 + e^{-1}(\ln x), \tag{7}$$

and equality will only hold at the global minimum $z = -1$. If $\ln x > -e$, the right-hand side of (7) is positive, thus H_x is strictly increasing when $x > e^{-e}$; for $x = e^{-e}$ the right-hand side of (7) is 0, but since equality only holds for a single value of z we have that h_x is positive save for a single point, and so H_x is yet strictly increasing. Thus the function H_x is one to one for every $e^{-e} \leq x \leq 1$. □

The intended use of the previous lemma is as follows: if we know that the subsequences of even and odd indexes of $\{a_n(x)\}_{n\in\mathbb{N}}$ are both convergent, then Lemma 3.9 tells us that their limits are zeroes of H_x; Lemma 3.11 further assures us that if $e^{-e} \leq x \leq 1$, H_x cannot have multiple zeroes, and thus $\{a_n(x)\}_{n\in\mathbb{N}}$ converges. Of course, we still need to prove that the even and odd subsequences actually converge; we will do this momentarily.

Note also that the injectivity of H_x provided by Lemma 3.11 does not impose a necessary condition for convergence of $\{a_n(x)\}_{n\in\mathbb{N}}$. Indeed, it is possible to prove that if $x > 1$, the function H_x is not one-to-one, so it may very well have multiple zeroes. However, the zeroes corresponding to the limits of the even and odd subsequences need not be distinct (and in fact we know that they cannot be distinct if $1 < x \leq e^{1/e}$, since we already know that $\{a_n(x)\}_{n\in\mathbb{N}}$ converges for these values of x).

We are now in position to extend the interval of convergence of the sequence $\{a_n(x)\}_{n\mathbb{N}}$ beyond what we have obtained thus far.

Proposition 3.12. *The sequence $\{a_n(x)\}_{n\in\mathbb{N}}$ converges for $e^{-e} \leq x < 1$.*

Proof. Lemmas 3.3 and 3.7 imply that there exist real numbers P and Q such that $\{a_{2k}(x)\}_{k\in\mathbb{N}}$ converges to P and $\{a_{2k-1}(x)\}_{k\in\mathbb{N}}$ converges to Q. Lemma 3.9 yields that $H_x(P) = H_x(Q) = 0$. Since by Lemma 3.11, the function H_x is one-to-one, we must have $P = Q$. □

Now it is left to discuss the case of $0 < x < e^{-e}$. We shall establish that the sequence $\{a_n(x)\}_{n\in\mathbb{N}}$ diverges for these values of x. We will do this by proving that the subsequences $\{a_{2k}(x)\}_{k\in\mathbb{N}}$ and $\{a_{2k-1}(x)\}_{k\in\mathbb{N}}$ possess distinct limits. A little more work is needed before reaching this goal.

Lemma 3.13. *The following hold.*

(a) *If $s \geq e^{-1}$, then $x^{(x^s)} > e^{-1}$ for all $x \neq e^{-1/s}$.*

(b) *If $0 < x < e^{-e}$, then e^{-1} is not a zero of the function H_x.*

Proof. Let $s > 0$ and consider the function $\rho_s(x) = x^{(x^s)}$ $(x > 0)$. An analysis of the signal of the derivative of ρ_s,

$$\rho_s'(x) = x^{s-1}\rho_s(x)(s \ln x + 1),$$

shows that the function has a unique global minimum at $x = e^{-1/s}$. Thus if $s \geq e^{-1}$ and $x \neq e^{-1/s}$, we have

$$x^{(x^s)} > e^{-1/(se)} \geq e^{-1}.$$

For Item (b), set $s = e^{-1}$ and notice that $e^{-1/s} = e^{-e}$, so by Item (a) we have $x^{(x^s)} > e^{-1}$ so long as $x \neq e^{-e}$, which yields $H_x(e^{-1}) < 0$; in particular, H_x is negative if $0 < x < e^{-e}$. □

Lemma 3.14. *For each $0 < x < e^{-e}$, we have $a_{2p-1}(x) < e^{-1} < a_{2q}(x)$ for every $p, q \in \mathbb{N}$.*

Proof. We first prove by induction that $a_{2k}(x) > e^{-1}$ for every $k \in \mathbb{N}$. Item (a) of Lemma 3.13 with $s = 1$ implies that for each $0 < x < e^{-1}$ we have $x^x > e^{-1}$, hence $a_2(x) > e^{-1}$, which is the base case of the induction. Now assume that for some $k \geq 1$, we have $a_{2k}(x) > e^{-1}$. We want to again apply Item (a) of Lemma 3.13, this time with $s = a_{2k}(x)$; since $s > e^{-1}$ and $x < e^{-e}$, it follows that $x \neq e^{-1/s}$, thus we conclude that

$$a_{2k+2}(x) = x^{\left(x^{a_{2k}(x)}\right)} > e^{-1},$$

so the induction is complete for the subsequence of even indexes.

Finally, note that $a_1(x) < e^{-1}$ and, for all $k \geq 1$, it is straightforward that

$$a_{2k+1}(x) = x^{a_{2k}(x)} < x^{e^{-1}} < (e^{-e})^{e^{-1}} = e^{-1}.$$ □

The following Proposition summarizes the latter results.

Proposition 3.15. *For each $0 < x < e^{-e}$, the sequence $\{a_n(x)\}_{n\in\mathbb{N}}$ diverges.*

Proof. It follows from Lemmas 3.3 and 3.7 that the subsequences $\{a_{2k}(x)\}_{k\in\mathbb{N}}$ and $\{a_{k-1}(x)\}_{k\in\mathbb{N}}$ converge to real numbers R and S, respectively. From Lemma 3.14 we know that $S \leq e^{-1} \leq R$. In addition, Lemma 3.9 and Item (b) of Lemma 3.13 assure that $S < e^{-1} < R$. Thus the sequence $\{a_n(x)\}_{n\in\mathbb{N}}$ diverges. □

Our work is almost done. We now have a complete understanding of which values of x generate a convergent sequence $\{a_n(x)\}_{n\in\mathbb{N}}$. Propositions 3.2, 3.5, 3.12, and 3.15 together imply the following:

Proposition 3.16. *The sequence $\{a_n(x)\}_{n\in\mathbb{N}}$ converges if and only if $e^{-e} \leq x \leq e^{1/e}$.*

All we have left in order to establish Theorem 2.1 is to relate the values of x from the last Proposition with the corresponding limits of $\{a_n(x)\}_{n\in\mathbb{N}}$.

Proof of Theorem 2.1. We know from Proposition 3.16 that $\{a_n(x)\}_{n\in\mathbb{N}}$ converges if and only if $e^{-e} \leq x \leq e^{1/e}$. For each x on this interval, let $\lambda(x) = \lim a_n(x)$. Then, the values L for which Equation (3) has a solution are precisely the numbers on the image of the interval $[e^{-e}, e^{1/e}]$ by the function λ. By Proposition 2.2, the function λ satisfies $x = \lambda(x)^{1/\lambda(x)}$; thus $\varphi(\lambda(x)) = x$, where φ is the function in Lemma 3.1, that is, φ is a left inverse for the function λ. By Remark 3.4 and Lemma 3.7, we have $\lambda(x) \in (0, e]$. But, Lemma 3.1 also tells us that φ is one-to-one and onto the interval $(0, e^{1/e}]$; since this interval contains the domain of λ and $\varphi(e^{-1}) = e^{-e}$, it folows that the image of λ is precisely the interval $[e^{-1}, e]$, establishing the theorem. □

Acknowledgments

The authors would like to thank FAPERJ and CNPq for their funding of this research.

References

[1] KNOEBEL, R. A. Exponentials Reiterated. Amer. Math. Monthly. 88(4):235–52, 1981.