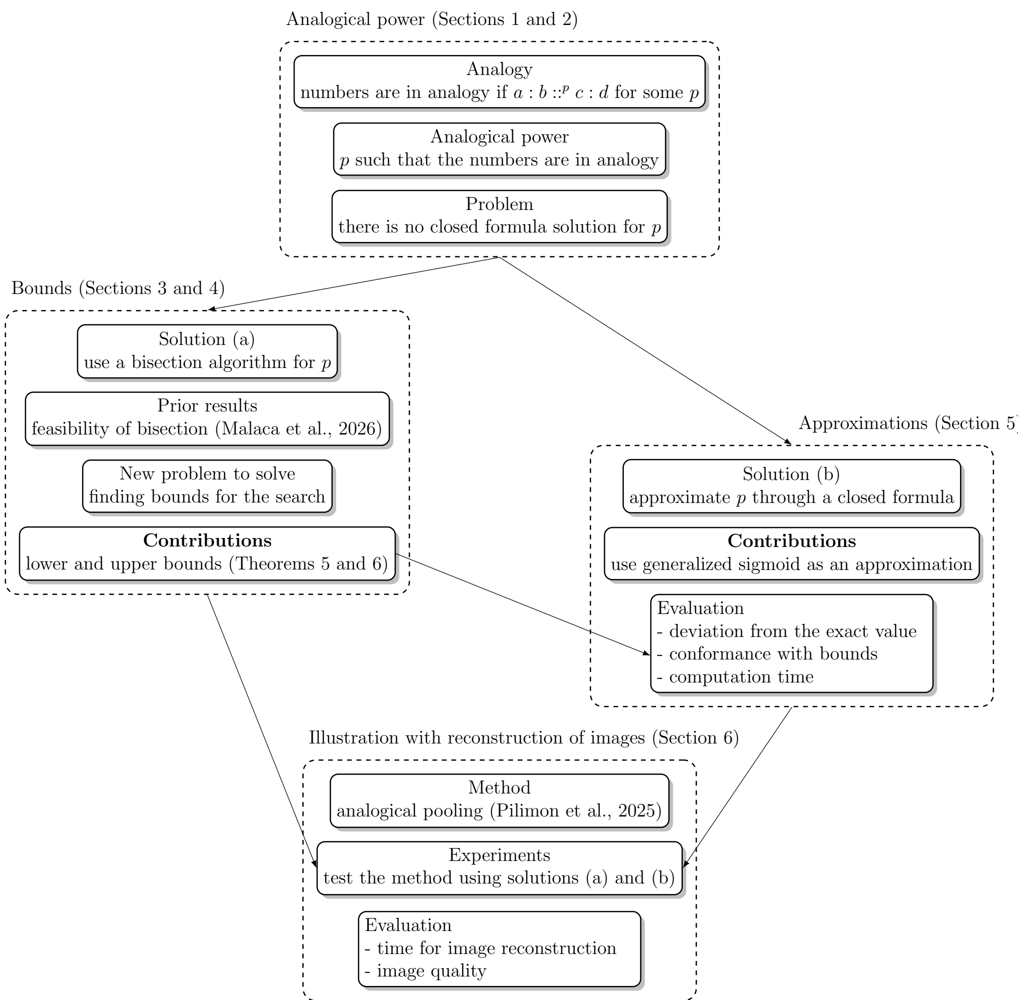
Graphical Abstract

Bounds and approximations of analogical powers

Francisco Malaca, Yves Lepage, Miguel Couceiro



Highlights

Bounds and approximations of analogical powers

Francisco Malaca, Yves Lepage, Miguel Couceiro

- **Reminder** on numerical analogy, analogical power, and justification of the work: no analytical expression exists for analogical powers that can be solved directly.
- **Contribution:** Theoretical results on bounds for analogical powers: theorems and proofs for all lower and higher bounds that completely cover all possible cases.
- **Contribution:** Proposal of several approximations of analogical powers: comparison in terms of execution time and framing by the theoretical bounds; application to the task of image reconstruction.

Bounds and approximations of analogical powers

Francisco Malaca[a], Yves Lepage[b], Miguel Couceiro[a]

[a] *INESC-ID, IST, Universidade de Lisboa, 9 Rua Alves Redol, 1000-029, Lisbon, Portugal*
[b] *Waseda University, IPS, 2-7 Hibikino, Wakamatsu-ku, Kitakyushu-shi, 808-0135, Fukuoka-ken, Japan*

Abstract

In this paper, we address fundamental computational challenges within the framework of numerical analogies. In particular, we consider the computation of analogical powers, which somewhat quantify the analogical relationship between the four numbers. Prior work established the feasibility of using a bisection algorithm to search for this value, but the efficiency of this search is entirely dependent on having tight bounds.

The latter motivates this paper's first objective, namely, to precisely define the necessary search space. For that, we prove lower and upper bounds for the analogical power of a given numerical analogy that reliably minimize the search space for the bisection algorithm that ensures the exact computation of analogical powers with guaranteed termination. To avoid computation burdens, we also propose a closed-form approximation of analogical powers based on a generalized sigmoid function. Both the bounds and the approximation methods are rigorously assessed in terms of their conformance, deviation, and computational time complexity.

The practical utility of our results is demonstrated through an application to image processing. We integrate both the exact (bound-enabled) and approximate computational methods into the process of "analogical pooling" for image reduction and reconstruction tasks, using the MNIST dataset as a benchmark. The experimental results provide a comparative analysis of the trade-offs between computation time and resulting image quality.

This work significantly advances the feasibility and applicability of analogical reasoning in real-world applications and complex high-dimensional downstream tasks. It provides theoretical results immediately usable in these real-world applications.

Keywords: Analogy, numerical analogy, generalized means, analogical

powers, image reconstruction

1. Introduction

Research on analogy in artificial intelligence (AI) originated from works in cognitive sciences. Back in the era of symbolic AI, one of the most influential works was the book by Hofstadter and the Fluid Analogies Research Group (1994) in which a chapter, Chapter 5, written in collaboration with Mitchell, proposed computational approaches to solve analogy puzzles. An example is: "Suppose that abc is changed into abd, how would you change ijk in the same way?" This can be written abc : abd :: ijk : x. The word *analogy* itself applied to this kind of pattern comes from ancient Greece, where it first referred to an equality of *ratios* between numbers, like, say, "2 is to 3 as 4 is to 6", written $2 : 3 = 4 : 6$ at elementary school. This evidently involves *four terms.* The pattern was extended outside of mathematics already during Greek antiquity. "Medicine is to cuisine as justice is to rhetoric" is an analogy used in Plato's Gorgias.

In the era of AI manipulating numerical representations in combination with machine learning techniques, analogy has been used in two ways. The first way considers Boolean-valued features. Formalization and techniques for the use of analogy on binary representations have been proposed (Klein, 1982; Prade and Richard, 2017) and demonstrated on standard datasets for classification, e.g. (Bounhas and Prade, 2024). To inquire into the foundation of Boolean analogy, connections with logic have also been explored (Antić, 2022, 2024). But numerical representations in modern AI go beyond Boolean representations. This is the case for the concept of four-term analogy "A is to B as C is to D" which has triggered much interest with numerical representations of words derived from distributional semantics in natural language processing (Mikolov et al., 2013). In this case, and this constitutes the second way, analogy is understood as *arithmetic analogy.* That is, if the vector associated with an object X is noted by $\vec{X}$, the analogy "$A : B :: C : D$" is interpreted as $\vec{A} - \vec{B} = \vec{C} - \vec{D}$.

Numerical analogy, as introduced in (Lepage and Couceiro, 2024) and Cunha et al. (2026), generalizes the two previous ways of using analogy. In a numerical analogy, the terms are real numbers. The above-cited articles showed how numerical analogy covers the cases of Boolean analogy with the same semantics. More importantly, a numerical analogy is characterized by a so-called *analogical power*, creating an infinity of different types of analogy. Arithmetic analogy is just one type among this infinity of types, namely the case where the analogical power is 1. The use of numerical analogy has already been exemplified in preliminary works in image reconstruction or in the exploration of number sequences (Pilimon et al., 2025; Qi et al., 2025). Its application to classification is also in progress. As the analogical power characterizes the type of an analogy, its determination for a given quadruple of real numbers is crucial. The implementation of algorithms for that purpose in the C programming language provides relatively fast computation. A current implementation using the bisection method (Lepage and Couceiro, 2025) computes one million analogical powers in less than 2 seconds. Although such speed is reasonable, it might not be enough for the massive computation of analogical powers in processing large images, or, if one envisages the use of numerical analogy in new types of neural networks with intense demand for computing several billions of analogical powers.

The purpose of this article is to explore theoretical foundations in the computation of analogical powers and explore the possibilities of speed-up by proposing approximations in computation.

The main contributions of this paper are threefold.

- Firstly, the existing bisection algorithm for the computation of analogical powers requires initial bounds, and this article proposes suitable ones and provides proofs for their determination.
- Secondly, this article explores approximations of analogical powers. Several approximations are proposed and are experimentally validated by measuring computation time and quality. One aspect of quality is to ensure that approximated values lie between the theoretical bounds.
- Thirdly, we empirically show the feasibility and applicability of analogical reasoning for real-world applications and complex high-dimensional downstream tasks, such as image processing and reconstruction.

The paper is structured as follows. In Section 2, we recall some historical background and justify the terms we use. We give the necessary definitions and reminders for the following developments. In particular, we stress the notion of analogical power because the topic of this article is to give bounds and approximations for it. Because the range of values for analogical powers is different depending on the sign of the terms in an analogy, we recall that all cases can be reduced to only two cases to inquire. Sections 3 and 4 give bounds in these two respective cases. The case where some terms are negative is decomposed into two further sub-cases for which different bounds are given. Section 5 deals with the approximation of the analogical power in the case of positive numbers. In particular, it examines whether the proposed approximations are framed in the bounds obtained in Section 3. Section 6 is a demonstrative application of approximated computation of analogical powers in the task of reconstruction of images.

2. Notation and definitions

When the objects in an analogy are not numbers, the ratios and the equality might not be so well-defined. In all generality, the symbol :: is thus preferably used to =. Whatever the kind of ratio, the symbol : is traditionally used. An analogy is thus noted $a : b :: c : d$ in a standard manner. Over time, specific and standard terms have been established by tradition to name the components in an analogy. a and d are called the *extremes* and b and c the *means*. Nonetheless, we will use non-classical more explicit wordings to reduce the reader's cognitive load: *outer terms* and *inner terms* in place of extremes and means, and sometimes *average* instead of mean.

The formalization of analogy in all generality, on any kind of objects, led to the identification of several fundamental properties. In particular, eight equivalent forms for the same analogy can be obtained by the inversion of ratios, the symmetry of ::, or by permutation of the extremes or the means.

$$\begin{array}{cccc} a : b :: c : d & b : a :: d : c & c : a :: d : b & d : b :: c : a \\ a : c :: b : d & b : d :: a : c & c : d :: a : b & d : c :: b : a \end{array} \tag{1}$$

As mentioned above, the term analogy was first reserved for numbers and the equality of ratios, that is, geometrical analogy. But it was rapidly expanded to other kinds of relations between four terms. Arithmetic analogy, which states that $a - b = c - d$, is a first example. A second example is harmonic analogy emerging from the study of harp chords by the Pythagorean school. As a curiosity, the *musical proportion*[1] is the geometric analogy between two numbers taken as its outer terms (extremes) and their arithmetic

[1] Which might have been known already in Ancient Babylon.

and harmonic means taken as its inner terms (means).[2]

$$a : \frac{a+d}{2} :: \frac{1}{1/2(1/a+1/d)} : d \text{ (geometric)} \;\Leftrightarrow\; a \times d = \frac{a+d}{2} \times \frac{2\,(a \times d)}{a+d} \tag{2}$$

This shows that the link of analogy with the notion of average has been established early in History. This link states that, in an analogy, the average of the outer terms (extremes) is equal to the average of the inner terms (means). For instance, in an arithmetic analogy, the arithmetic mean of a and d, $1/2(a+d)$, is equal to the arithmetic mean of b and c, $1/2(b+c)$.

$$a : b :: c : d \text{ (arithmetic)} \;\Leftrightarrow\; a - b = c - d \;\Leftrightarrow\; \frac{1}{2}(a+d) = \frac{1}{2}(b+c) \tag{3}$$

Although several kinds of averages, harmonic, arithmetic, geometric, or quadratic, had been identified, a generalization of the notion of all these different averages was not offered until the XIX$^{\text{th}}$ century (Hölder, 1889).

Definition 1. *The* generalized mean *(average) of two numbers a and d is defined as*

$$m_p(a,d) = \left(\frac{1}{2}\left(a^p + d^p\right)\right)^{1/p},$$

for $p \in \mathbb{R}^*$, and, if the limits exist, $m_p(a,d) = \lim_{r \to p} m_r(a,d)$, for $p \in \{-\infty, 0, +\infty\}$. For negative numbers, exponentiation is considered using the principal branch of the complex logarithm, which is defined through the principal argument.

This naturally leads to the generalization of analogy by simply positing the equality of the same kind of means for the outer terms (extremes) and inner terms (means).

Definition 2. *Given four numbers a, b, c and d in $\mathbb{R}$, they make a* numerical analogy *if there exists a $p \in \mathbb{R} \cup \{-\infty, +\infty\}$, such that $m_p(a,d) = m_p(b,c)$. We denote this analogy in analogical power p by writing $a : b ::^p c : d$.*

[2] Again, caution with the two meanings of *mean* in this article. In plural, and by contrast with the outer terms which are called the *extremes*, it refers to the inner terms, i.e., b and c, in an analogy $a : b :: c : d$. And in singular or plural, it may refer to the average of two numbers, for instance, the arithmetic, geometric, or harmonic means.

Table 1: Examples of analogies.

Terms a, b, c, d	Power p	Equality of gen. means $m_p(a,d) = m_p(b,c)$	Analogy $a : b ::^p c : d$
2, 2, 5, 67	$-\infty$	$\min(2,67) = \min(2,5)$	$2 : 2 ::^{-\infty} 5 : 67$ minimum
8/5, 2, 4, 8	-1	$\dfrac{1}{\frac{1}{2}\left(\frac{1}{8/5}+\frac{1}{8}\right)} = \dfrac{1}{\frac{1}{2}\left(\frac{1}{2}+\frac{1}{4}\right)}$	$\frac{8}{5} : 2 ::^{-1} 4 : 8$ harmonic
2, 4, 8, 16	0	$\sqrt{2 \times 16} = \sqrt{4 \times 8}$	$2 : 4 ::^{0} 8 : 16$ geometric
1, 64, 343, 512	2/3	$\left(\frac{1}{2}(1^{2/3} + 512^{2/3})\right)^{\frac{3}{2}} = \left(\frac{1}{2}(64^{2/3} + 343^{2/3})\right)^{\frac{3}{2}}$	$1 : 64 ::^{2/3} 343 : 512$ no specific name
5, 7, 10, 12	1	$\frac{1}{2}(5+12) = \frac{1}{2}(7+10)$	$5 : 7 ::^{1} 10 : 12$ arithmetic
1, 81, 100, 144	3/2	$\left(\frac{1}{2}(1^{3/2} + 144^{3/2})\right)^{\frac{2}{3}} = \left(\frac{1}{2}(81^{3/2} + 100^{3/2})\right)^{\frac{2}{3}}$	$1 : 81 ::^{3/2} 100 : 144$ no specific name

Table 1 illustrates the notion with different powers.

Below, we rely on several results established in (Malaca et al., 2026) (under review) that we recall here. The analogy respects the eight equivalent forms in Formula 1 and is invariant to transforming all terms in their additive inverse. Therefore, if the terms are non-null, there are four cases to consider: all terms are positive ($a : b :: c : d$ – Theorem 1), exactly one term is negative ($-a : b :: c : d$ – Theorem 4), only the first two terms are negative ($-a : -b :: c : d$ – Theorem 2), and only the extremes are negative ($-a : b :: c : -d$ – Theorem 3).

We do not consider the special cases where p is not unique or infinite, because these cases can be easily identified and do not require searching for the analogical powers, bounding them, or approximating them.

Secondly, apart from these special cases, Theorems 2 and 3 show that, in fact, only the two cases described in Theorems 1 and 4 have to be considered. As the analogical power is unique, it makes sense to look for its bounds. In Theorems 1, 2, 3, and 4, it is enough to assume that a, b, c, and d are pairwise distinct in absolute value.

Theorem 1 addresses the case where all the terms are positive. It states that there always exists a numerical analogy, that is, there always exists a power p such that $a : b ::^p c : d$ if the terms are positive real numbers and sorted in increasing order.

Theorem 1 (Lepage and Couceiro (2024))**.** *Any four positive real numbers in increasing order make an analogy in a unique power.*
$\forall (a, b, c, d) \in (\mathbb{R}_+)^4, \;\; 0 \leq a < b \leq c < d,$

$$\exists!\, p \in \mathbb{R} \quad s.t. \quad a : b ::^p : c : d.$$

We need to address the cases where some terms are negative. Classically let $2\mathbb{Z}^*$ denote the set of even nonzero integers, $\mathbb{Z}\backslash 2\mathbb{Z}$ the set of odd integers, $\mathbb{Z}_+\backslash 2\mathbb{Z}$ the set of positive odd integers and $\mathbb{Z}_-\backslash 2\mathbb{Z}$ the set of negative odd integers.

Theorem 2 establishes that in the case where the first two terms are negative, the power can be obtained from an equivalent analogy with all terms positive, with the restriction that the power should be an integer. Therefore, the bounds for the analogical power of an analogy with the first two terms negative are given by the bounds for an equivalent analogy where all terms are positive.

Theorem 2. *Given $a, b, c, d \in \mathbb{R}_+^*$, with $-a < -b < c, d$, and $p \in \mathbb{R}$, an analogy $-a : -b ::^p c : d$ exists if and only if either $a : b ::^p c : d$ for some $p \in 2\mathbb{Z}$ or $a : b ::^p d : c$ for some $p \in \mathbb{Z} \backslash 2\mathbb{Z}$.*

Sketch proof. The result is shown by raising the two sides of the equality to p, noticing that $m_p(-a, d)^p = \dfrac{1}{2}\left((-a)^p + d^p\right)$ and $m_p(-b, c)^p = \dfrac{1}{2}\left((-b)^p + c^p\right)$. $\square$

Moreover, Theorem 3 shows that if the extremes are negative, the analogy can be written with positive terms, restricting to even values of p. Again, the bounds for the analogical power can be obtained from bounds for the positive terms.

Theorem 3. *Given $a, b, c, d \in \mathbb{R}_+^*$, with $-a < -d < b \leq c$, and $p \in \mathbb{R}$, an analogy $-a : b ::^p c : -d$ exists if and only if $a : b ::^p c : d$ for some $p \in 2\mathbb{Z}$.*

Sketch proof. As $m_p(b, c)$ is positive, $m_p(-a, -d)$ must be positive too, which only occurs for even p. $\square$

When one of the terms is negative, and the analogical power is a non-null even number, the analogy can be equivalently written with the absolute value of the terms, and the bounds for the case in Theorem 1 will apply. Otherwise, the following theorem states that the equality of the generalized means must hold for an odd p.

Theorem 4. *Given $a, b, c, d \in \mathbb{R}_+^*$, with $b \leq c$, and $p \in \mathbb{R} \setminus 2\mathbb{Z}^*$, an analogy $-a : b ::^p c : d$ exists if and only if $-a^p + d^p = b^p + c^p$ and $p \in \mathbb{Z} \backslash 2\mathbb{Z}$.*

Sketch proof. As $m_p(b, c)$ is positive, $m_p(-a, d)$ must be positive too, which can only occur for $p \in \mathbb{Z}$, but as we excluded $p \in 2\mathbb{Z}^*$, p must be an odd number. $\square$

In summary of all the remarks above, the determination of the analogical powers of all cases of analogies where the terms are non-null, can be deduced by the inspection of only two cases, $a : b ::^p c : d$ and $-a : b ::^p c : d$, where a, b, c and d are positive. Therefore, the sequel of the present paper examines the bounds of analogical powers for these two cases.

In addition, all cases where some of the terms are null lead to an infinite number of analogical powers (e.g., $0 : 0 ::^p d : d \Rightarrow p \in \mathbb{R}_+ \cup \{+\infty\}$), or to an infinite power (e.g., for $b < d$, $0 : b ::^p d : d \Rightarrow p = +\infty$), except for one case. This one case is $0 : \pm b ::^p \pm c : \pm d$, which can be equivalently written with one null term and three strictly positive terms, given some restrictions on p, which Theorem 1 covers.

3. Bounds for analogies with positive terms only

In a previous article (Lepage and Couceiro, 2024), it was briefly mentioned how, given a quadruple of positive numbers, one can determine the power p of the analogy in p by means of a bisection method starting from $-\infty$ and $+\infty$. Below, we exhibit finer bounds for p. These bounds have been presented in (Lepage and Couceiro, 2025), but here we give a shorter and more direct version of the proofs.

Theorem 5. *Bounds with positive terms.*
$\forall p \in \mathbb{R},\ \ \forall (a, b, c, d) \in \left(\mathbb{R}_+^*\right)^4,\ \ 0 < a < b \leq c < d,$

$$a : b ::^p c : d \ \ \Rightarrow \ \ \frac{1}{\log_2 a/b} \leq p \leq \frac{1}{\log_2 d/c}$$

$$0 : b ::^p c : d \ \ \Rightarrow \ \ \frac{1}{\log_2 d/b} \leq p \leq \frac{1}{\log_2 d/c}$$

Figure 1 illustrates for a particular analogy how the analogical power is located between the bounds calculated by Theorem 5.

The remainder of this section consists of the proof of Theorem 5.

3.1. Upper bound

Firstly, we consider a positive p and two real numbers a and d such that $0 \leq a < d$.

$$\begin{aligned}
&d^p \leq a^p + d^p \leq 2d^p \\
&\quad \Leftrightarrow \ \ d^p \leq 2 \times \frac{1}{2}\left(a^p + d^p\right) \leq 2d^p \\
&\quad \Leftrightarrow \ \ \left(d^p\right)^{1/p} \leq 2^{1/p} \times \left(\frac{1}{2}\left(a^p + d^p\right)\right)^{1/p} \leq 2^{1/p}\left(d^p\right)^{1/p} \\
&\quad \Leftrightarrow \ \ d \leq 2^{1/p} m_p(a, d) \leq 2^{1/p} d \\
&\quad \Leftrightarrow \ \ \frac{d}{2^{1/p}} \leq m_p(a, d) \leq d
\end{aligned}$$

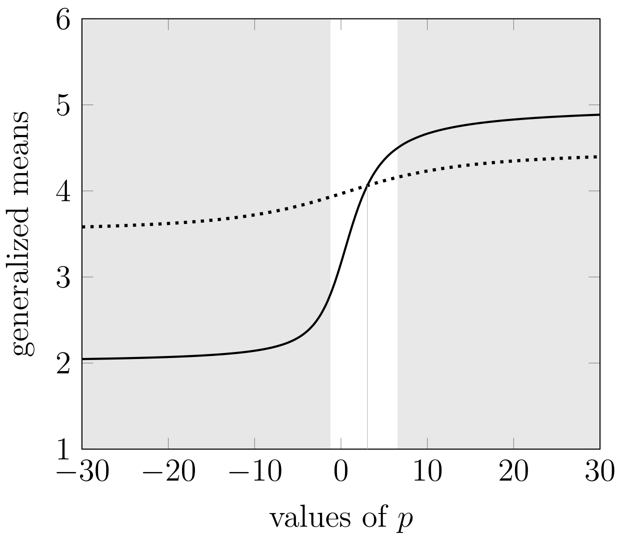


Figure 1: The solid curve shows the generalized means for $a = 2$ and $d = 5$, with p in abscissae. The dotted curve shows the same for $b = 3.5$ and $c = 4.5$. The power p of the analogy $a : b ::^p c : d$ is the argument of the intersection point of the solid and dotted curves. In this case, $p \simeq 3.06$. The clear area indicates the interval to explore to find this power. It is delimited by the bounds obtained using Theorem 5, here -1.24 and 6.58.

The first sequence of inequalities is true because p is positive, a is non-negative, and d is greater than a. Raising to the power $1/p$ does not change the order of the inequalities because p is positive.[3] We apply this result to two other real numbers b and c such that $0 < b \leq c$.

$$\frac{b}{2^{1/p}} \leq m_p(b,c) \leq c$$

Taking the additive inverses reverses the direction of the inequalities.

$$-c \leq -m_p(b,c) \leq -\frac{b}{2^{1/p}}$$

We combine the two previous inequalities under the assumption of an analogy in power p, $a : b ::^p c : d$. That is, we suppose the equality of the generalized means in p.

$\forall p \in \mathbb{R}_+,\ \forall (a,b,c,d) \in \left(\mathbb{R}_+^*\right)^4,\ 0 \leq a < b \leq c < d,$

$$\begin{aligned} a : b ::^p c : d \quad & \Rightarrow \frac{d}{2^{1/p}} - c \leq \underbrace{m_p(a,d) - m_p(b,c)}_{=0} \\ & \Rightarrow \frac{d}{2^{1/p}} \leq c \\ & \Rightarrow \frac{d}{c} \leq 2^{1/p} \\ & \Rightarrow \log_2\left(\frac{d}{c}\right) \leq \frac{1}{p} \\ & \Rightarrow p \leq \frac{1}{\log_2(d/c)} \end{aligned}$$

Since the term $1/(\log_2 d/c)$ is positive, the inequality holds for non-positive p too.

[3] There is a similar well-known result for the L_p norm:

$$\forall p \in \mathbb{R}_+,\ \forall \vec{x} \in \mathbb{R}^n,\ \ \|\vec{x}\|_{+\infty} \leq \|\vec{x}\|_p \leq n^{1/p}\,\|\vec{x}\|_{+\infty}.$$

3.2. Lower bound for $a \neq 0$

Similarly, we consider a negative p and two positive non-zero real numbers a and d such that $0 < a < d$.

$$\begin{aligned}
a^p \leq a^p + d^p \leq 2a^p \quad &\Leftrightarrow \quad a^p \leq 2 \times \frac{1}{2}\left(a^p + d^p\right) \leq 2a^p \\
&\Leftrightarrow \quad 2^{1/p} \ \left(a^p\right)^{1/p} \leq 2^{1/p} \times \left(\frac{1}{2}\left(a^p + d^p\right)\right)^{1/p} \leq \left(a^p\right)^{1/p} \\
&\Leftrightarrow \quad 2^{1/p} \ a \leq 2^{1/p} m_p(a,d) \leq a \\
&\Leftrightarrow \quad a \leq m_p(a,d) \leq \frac{a}{2^{1/p}}
\end{aligned}$$

From the first to the second lines and from the third to the fourth lines, the inequalities are reversed because p and hence $1/p$ are negative.

With b and c such that $0 < b \leq c$, we have

$$b \leq m_p(b,c) \leq \frac{b}{2^{1/p}}.$$

Therefore,

$\forall p \in \mathbb{R}_-, \ \forall (a,b,c,d) \in \left(\mathbb{R}_+^*\right)^4, \ 0 < a < b \leq c < d,$

$$\begin{aligned}
a : b ::^p c : d \quad &\Rightarrow \quad \underbrace{m_p(a,d) - m_p(b,c)}_{=0} \leq \frac{a}{2^{1/p}} - b \\
&\Rightarrow \quad 0 \leq \frac{a}{2^{1/p}} - b \\
&\Rightarrow \quad b \leq \frac{a}{2^{1/p}} \\
&\Rightarrow \quad 2^{1/p} \leq \frac{a}{b} \\
&\Rightarrow \quad \frac{1}{p} \leq \log_2\left(\frac{a}{b}\right) \\
&\Rightarrow \quad \frac{1}{\log_2\left(a/b\right)} \leq p
\end{aligned}$$

Once more, as the term $1/(\log_2 a/b)$ is negative, the inequality holds for non-negative p too.

3.3. Lower bound for $a = 0$

If $a = 0$, then $m_p(a, d) = m_p(0, d) = d/(2^{1/p})$, and so $m_p(0, d) \leq d/(2^{1/p})$. Thus $\forall p \in \mathbb{R}_+,\ \forall (b, c, d) \in \left(\mathbb{R}_+^*\right)^3,\ 0 < b \leq c < d,$

$$\begin{aligned}
0 : b ::^p c : d \quad & \Rightarrow \underbrace{m_p(0, d) - m_p(b, c)}_{=0} \leq \frac{d}{2^{1/p}} - b \\
& \Rightarrow b \leq \frac{d}{2^{1/p}} \\
& \Rightarrow 2^{1/p} \leq \frac{d}{b} \\
& \Rightarrow \frac{1}{p} \leq \log_2 \left(\frac{d}{b}\right) \\
& \Rightarrow \frac{1}{\log_2 (d/b)} \leq p
\end{aligned}$$

3.4. Summarization

To summarize this section, the combination of the inequalities above gives the bounds stated in Theorem 5.

4. Bounds for analogies with one negative term

It has been seen above that the only relevant case with negative terms is the case with one negative term only. The following theorem gives the bounds in that case.

Theorem 6. *Bounds with one negative term.*
$\forall p \in \mathbb{Z}_+ \backslash 2\mathbb{Z}\ \ \forall (a, b, c, d) \in \left(\mathbb{R}_+^*\right)^4,\ \ a < d\ \wedge\ b \leq c < d,$

$$\begin{aligned}
& -a : b ::^p c : d\ \Rightarrow \\
& \max\left\{\frac{1}{\log_3 d/\min\{a, b\}}, \frac{1}{\log_2 d/\max\{b, \min\{a, c\}\}}\right\} \leq p \leq \frac{1}{\log_3 d/\max\{a, c\}}
\end{aligned}$$

Note that the values given in the lower bound can happen to be less than 1. In this case, of course, as we supposed $1 \leq p$, 1 will be the actual lower bound.

4.1. Upper bound

Proof. Remind that we only consider the case where p is a positive odd integer. We thus exploit $(-a)^p = -a^p$.

$$\begin{aligned}
-a : b ::^p c : d &\Leftrightarrow (-a)^p + d^p = b^p + c^p \\
&\Leftrightarrow -a^p + d^p = b^p + c^p \\
&\Leftrightarrow d^p = a^p + b^p + c^p \\
&\Leftrightarrow 1^p = \left(\frac{a}{d}\right)^p + \left(\frac{b}{d}\right)^p + \left(\frac{c}{d}\right)^p \\
&\Leftrightarrow 1 = \left(\frac{a}{d}\right)^p + \left(\frac{b}{d}\right)^p + \left(\frac{c}{d}\right)^p
\end{aligned}$$

The function x^p is a non-decreasing function for positive p. Consequently, $\max\{(a/d)^p, (b/d)^p, (c/d)^p\} = (\max\{a/d, b/d, c/d\})^p$.

$$\begin{aligned}
1 = \left(\frac{a}{d}\right)^p + \left(\frac{b}{d}\right)^p + \left(\frac{c}{d}\right)^p \quad &\Rightarrow \quad \frac{1}{3} \leq \max\left\{\left(\frac{a}{d}\right)^p, \left(\frac{b}{d}\right)^p, \left(\frac{c}{d}\right)^p\right\} \\
&\Rightarrow \quad \frac{1}{3} \leq \left(\max\left\{\frac{a}{d}, \frac{b}{d}, \frac{c}{d}\right\}\right)^p \\
&\Rightarrow \quad \frac{1}{3} \leq \left(\frac{1}{d}\max\{a, b, c\}\right)^p \\
&\Rightarrow \quad -\ln 3 \leq p \ln\left(\frac{1}{d}\max\{a, b, c\}\right) \\
&\Rightarrow \quad p \ln\left(\frac{d}{\max\{a, b, c\}}\right) \leq \ln 3 \\
&\Rightarrow \quad p \leq \ln 3 \,/\, \ln\left(\frac{d}{\max\{a, b, c\}}\right) \\
&\Rightarrow \quad p \leq 1 \,/\, \log_3\left(\frac{d}{\max\{a, b, c\}}\right)
\end{aligned}$$

As we have supposed that $b \leq c$, we can write:

$$p \leq 1 \,/\, \log_3\left(d/\max\{a, c\}\right).$$

□

4.2. Lower bounds

Proof. We compute several lower bounds and their maximum:

$$\max\left\{\frac{\ln(1/3)}{\ln(\min\{a,b\}/d)}, \frac{\ln(1/2)}{\ln(\max\{b,\min\{a,c\}\}/d)}\right\} \;\leq\; p.$$

The first lower bound is similar to the upper bound. By exchanging max with min and changing the direction of the inequalities in the proof above, one gets:

$$1\,/\log_3\,(d/\min\{a,b\}) \leq p. \tag{4}$$

Notice that, instead of $\max\{a,c\}$ in the upper bound, here we have $\min\{a,b\}$ since we suppose $b \leq c$.

To compute the second lower bound, we make an observation about the second greatest term in the sum $(a/d)^p + (b/d)^p + (c/d)^p$. Since all terms in this sum are nonzero and positive, and since the sum is equal to 1, the second greatest term cannot be greater than $1/2$.[4] We can give a formula for this second-greatest term. Recall that we supposed that $b \leq c$, but we do not know how a is positioned relative to b and c. Under these conditions, the second greatest term is $((\max\{b,\min\{a,c\}\})/d)^p$. From the constraint of being less than $1/2$, we derive:

$$\begin{aligned}(\max\{b,\min\{a,c\}\}/d)^p \leq \frac{1}{2} \;&\Leftrightarrow\; -\frac{1}{\log_2(\max\{b,\min\{a,c\}\}/d)} \leq p\\ &\Leftrightarrow\; \frac{1}{\log_2(d/\max\{b,\min\{a,c\}\})} \leq p.\end{aligned}$$

Note that, because d is greater than a, b, and c, the argument of the log in the denominator is greater than 1, hence the logarithm is positive. □

5. Approximation of analogical powers

After studying how we can frame the analogical power of an analogy between a lower and an upper bound, we turn to the problem of directly computing the analogical power.

[4]Either the greatest term is larger than $1/2$, and then the second term is necessarily less than 1 minus this greatest term, i.e., less than $1/2$, or the greatest term is less than $1/2$, and by definition, the second greatest term being less than the greatest term, it is less than $1/2$.

When negative terms are considered, the generalized means assume complex values, and the existence of an analogical power is very sensitive to the input values. For example, according to Theorem 3, an analogy of the form $-a : b ::^p c : -d$ only holds for even p, in which case $|a|^p + |d|^p = |b|^p + |c|^p$. A slight variation in one of the inputs changes the solution, p, to the equation involving the absolute value of the terms. Thus, p is no longer an even number, and an analogical power no longer exists.

In the following, we restrict ourselves to the case of analogies with positive terms. Under this condition, the generalized means, as functions of p, are infinitely differentiable.

In the general case, solving the equation $m_p(a, d) = m_p(b, c)$ for p is not an easy problem. As mentioned earlier, an algorithm based on the bisection method has been proposed and implemented. By construction of this algorithm, multiple recursive calls are executed until reaching a condition of convergence. The program is reasonably fast, but any speed-up is always welcome. A direct computation using some analytical approximation can thus be examined in place of this procedure if the computation time and the quality of the approximation are satisfactory.

We note that the problem of determining the analogical power, given four terms, is ill-conditioned for $a \approx b$, $c \approx d$ and $ad \approx cd$. This is shown in Appendix B. If $a = b$ and $c = d$, any p is an analogical power, and if $ad = bc$, the terms are in geometric analogy and $p = 0$.

In Subsection 5.1, we first study an approximation of the generalized means function via a generalized sigmoid. As the slope of the sigmoid can be controlled by a parameter, which we call B, we first consider constant values of B that can be derived from natural considerations. We then inspect different proposals for B as a function of the two arguments of the generalized means. These proposals rely on an attempt to make the derivatives of the generalized means and the approximating sigmoid coincide at 1. We graphically evaluate which proposal gives the best approximation.

In Subsection 5.2, we then approximate the analogical power using the previously proposed approximations of the generalized means using the different values of B proposed. In addition, we also propose formulae for B that depend on the four terms in the analogy and even discuss the efficiency of some proposals that were shown to be inefficient in terms of approximation or computation time.

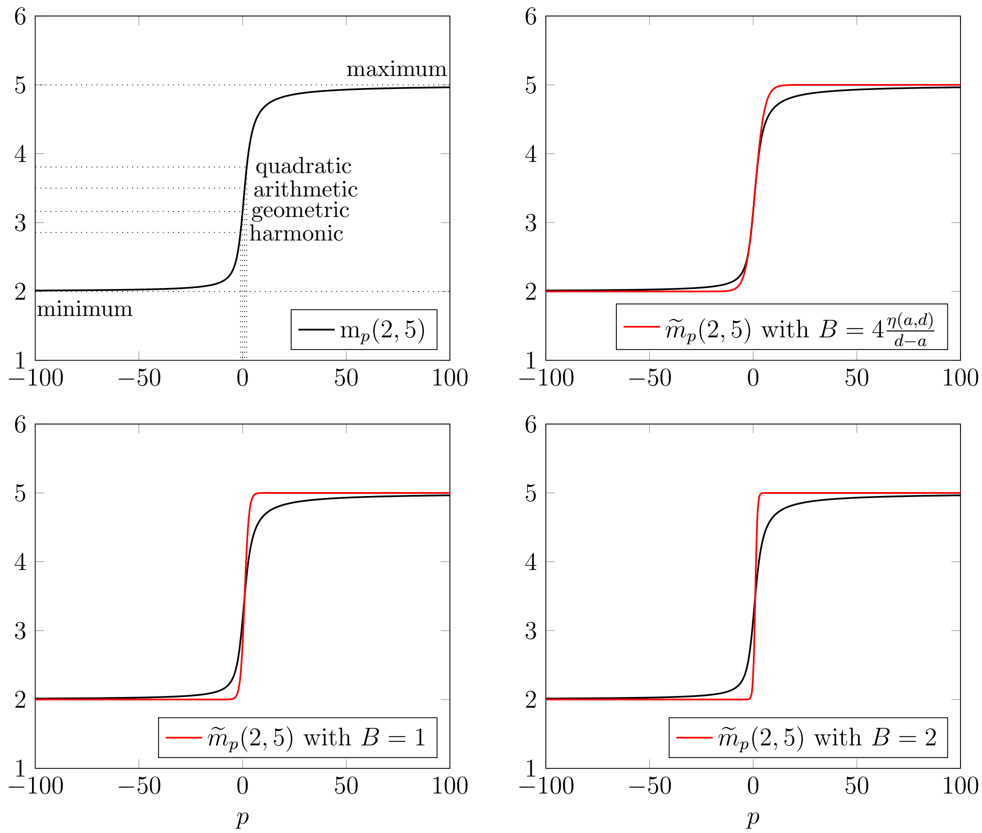


Figure 2: *Top left*: graph of the generalized means of $a = 2$ and $d = 5$, with p on the horizontal axis. In particular, the values of the harmonic (2.86), geometric ($\sqrt{2 \times 5} = \sqrt{10.0} \simeq 3.16$), arithmetic (3.5), and quadratic (3.81) means are indicated. Similarly, the limits at $-\infty$ and $+\infty$ are the minimum and maximum, i.e., 2 and 5, respectively. The three other graphs superimpose approximations in red on the graph of the generalized means.

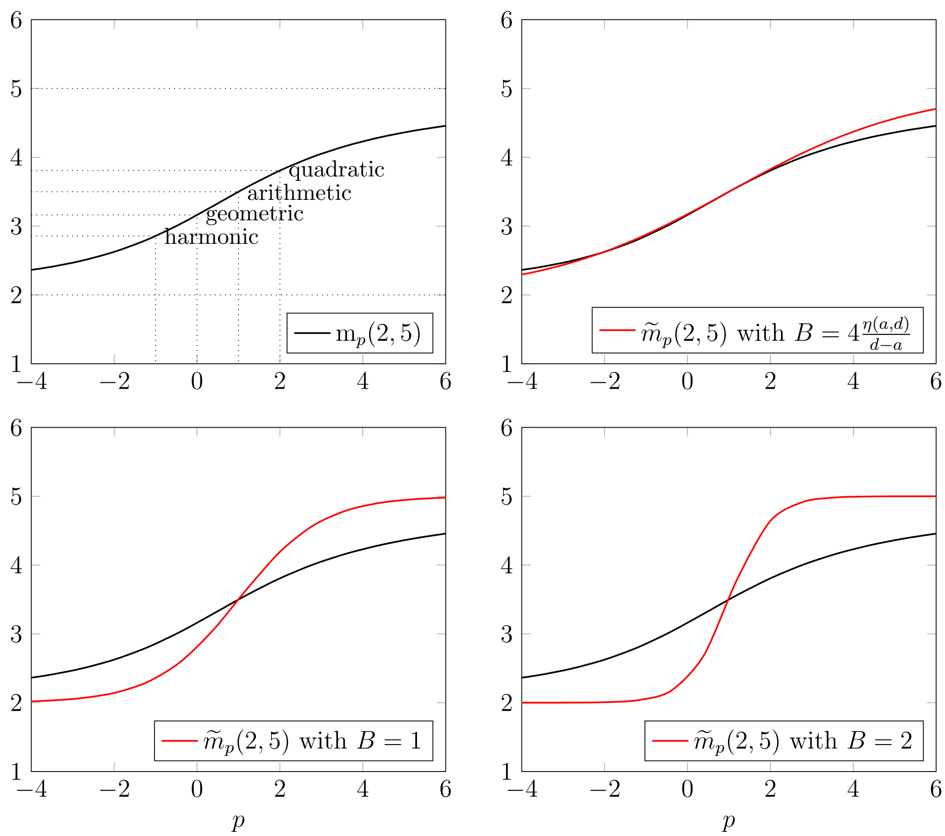


Figure 3: Same as Figure 2, but zooming on the abscissa of the interval $[-4, 6]$.

5.1. Approximation of the generalized means

Sigmoid. The graph of the generalized means between two positive real numbers looks like a sigmoid. However, one needs to observe that there is no central symmetry around a particular point. Figure 2 shows that the behavior before 0 is not symmetrical to the behavior right after 0. The reader should observe that the shape is more round after 0. Appendix A provides a proof that there is no central symmetry in the general case.

We approximate the curve of generalized means between a and d using a sigmoid function. The general formula is as follows:

$$\frac{1}{1+e^{-p}} \tag{5}$$

It supposes that there is a central symmetry relative to 0 ($-p = 0 - p$). For our case, we consider that a better supposed central point, that does not exist as we have shown above, would be 1. Hence, we choose the function below:

$$\frac{1}{1+e^{(1-p)}} \tag{6}$$

As this function should be framed between a and d (with $a < d$), we linearly shift it by a and scale it using a factor of $d - a$. An approximation of the generalized mean is the following one:

$$a + \frac{d-a}{1+e^{(1-p)}}. \tag{7}$$

It verifies the following properties:

- when p tends towards $-\infty$, it tends towards a;
- when p is equal to 1, it is equal to the arithmetic mean $\frac{1}{2}(a+d)$;
- when p tends towards $+\infty$, it tends towards d.

Introduction of factor B. To control the slope of the function, we introduce a factor $B > 0$ in the exponential of the denominator.

Definition 3. *Approximation of m_p with sigmoid parametrized by B.*

$$\widetilde{m}_p(a,d) = a + \frac{d-a}{1+e^{B(1-p)}}. \tag{8}$$

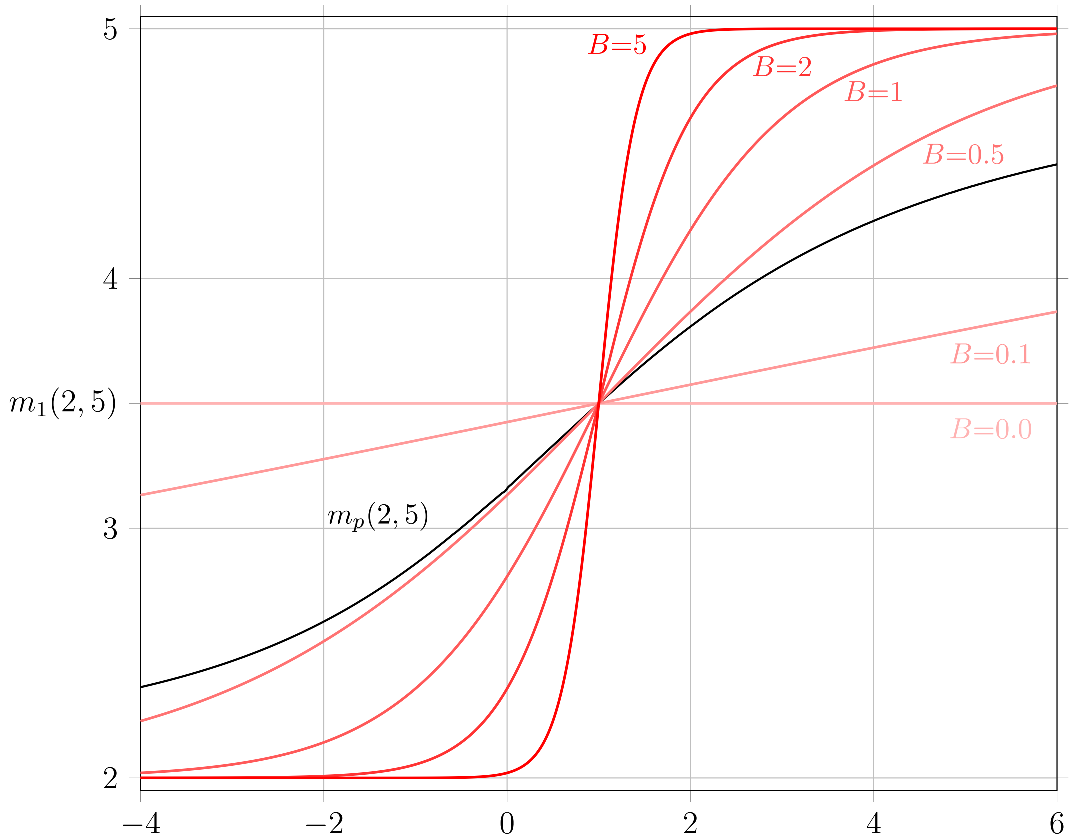


Figure 4: Sigmoid approximation of $m_p(2, 5)$ according to Definition 3. Different values of B show that B controls the slope of the sigmoid.

If B is positive and non-null, its introduction does not influence the values taken for $p = -\infty, 1$ and $+\infty$. Figure 4 illustrates the effect of varying B. When B is larger, the slope of the sigmoid is sharper, and when B tends toward 0, the sigmoid tends toward the horizontal line $y = m_1(a, d)$. When compared with the actual curve of the generalized means of the two numbers a and d, the sigmoid functions parametrized by B are located above or below it, and there should be some value of B that best approximates the generalized means. We shall explore several possible positive values for B.

$B = 1$. The simple form of the approximation of the generalized mean function using the sigmoid function in (7) is just the formula in Definition (3) with a value of 1 for B.

$B = 2$. The sigmoid function recast between a and d given in (7) looks somewhat unbalanced or not so symmetric relative to a and d. If we observe that $a + (d-a)/(1+e^{(1-p)}) = (a-d)/(e^{-(1-p)}+1) + d$, the use of $e^{(1-p)}$ and $e^{-(1-p)}$ naturally triggers an attempt at using the hyperbolic cosine of $1-p$: $\cosh(1-p) = (e^{(1-p)} + e^{-(1-p)})/2$. To take a and d into account, we introduce them in the hyperbolic cosine function as weights and consider the ratio between the hyperbolic cosine weighted by a and d and the true, unweighted, hyperbolic cosine.

$$\begin{aligned}\frac{\dfrac{ae^{(1-p)} + de^{-(1-p)}}{a+d}}{\dfrac{e^{(1-p)} + e^{-(1-p)}}{2}} &= \frac{ae^{(1-p)} + de^{-(1-p)}}{a+d} \times \frac{2}{e^{(1-p)} + e^{-(1-p)}} \\ &= \frac{2}{a+d} \times \frac{ae^{(1-p)} + de^{-(1-p)}}{e^{(1-p)} + e^{-(1-p)}} \\ &= \frac{1}{m_1(a,d)} \times \frac{ae^{(1-p)} + de^{-(1-p)}}{e^{(1-p)} + e^{-(1-p)}} \qquad (9)\end{aligned}$$

With this, the values at 1 and the infinities are respectively 1, $a/m_1(a,d)$, and $d/m_1(a,d)$. To recast this function to the generalized means, and ensure that the values at 1 and the infinities are $m_1(a,d)$, a, and d respectively, we eliminate the factor $m_1(a,d)$ in the denominator. We write:

$$\frac{ae^{-(1-p)} + de^{(1-p)}}{e^{-(1-p)} + e^{(1-p)}}. \qquad (10)$$

Figure 2 gives the graph of this function superimposed onto the graph of the generalized means for particular values of a and b. It shows that this function can be used as an approximation of the generalized means. Now, by observing that

$$\frac{ae^{-(1-p)} + de^{(1-p)}}{e^{-(1-p)} + e^{(1-p)}} = a + \frac{d-a}{1+e^{2(1-p)}}, \qquad (11)$$

we see that this is just Definition 3 with a value of 2 for B.

B for equal derivatives of m_p and $\widetilde{m}_p$ at 1. Alternatively, we consider a B depending on a and d for which the derivative at $p = 1$ is equal to the derivative of the generalized mean, besides meeting the expected values at 1

and the infinities. Appendix C shows how this value can be derived from the calculation of the derivatives. To simplify the expression, we define[5]

$$\mathrm{h}(x) = -x \log x \tag{12}$$

and

$$\begin{aligned}\eta(x, y) &= \mathrm{h}(\frac{1}{2}(x+y)) - \frac{1}{2}(\mathrm{h}(x) + \mathrm{h}(y)) \\ &= \mathrm{h}(m_1(x, y)) - m_1(\mathrm{h}(x), \mathrm{h}(y)).\end{aligned} \tag{13}$$

With this,

$$B = 4 \times \frac{\eta(a, d)}{d-a}. \tag{14}$$

Figure 3 provides a zoom around $[-4, 6]$ showing that the approximation around 1 using (14) is of better quality than the ones obtained with constant values of B.

5.2. Approximation of analogical powers

The intersection of the curves for a and d and b and c obtained using Definition 3 is the point $\tilde{p}$ such that $\widetilde{m}_{\tilde{p}}(a, c) = \widetilde{m}_{\tilde{p}}(b, c)$. It is an approximation of the analogical power of the analogy $a : b ::^p c : d$. Its computation is as follows.

$$\begin{aligned} a + \frac{d-a}{1+e^{B(1-\tilde{p})}} = b + \frac{c-b}{1+e^{B(1-\tilde{p})}} &\Leftrightarrow \frac{(d-a)-(c-b)}{1+e^{B(1-\tilde{p})}} = (b-a) \\ &\Leftrightarrow \frac{(d-c)+(b-a)}{b-a} = 1 + e^{B(1-\tilde{p})} \\ &\Leftrightarrow \frac{d-c}{b-a} + 1 = 1 + e^{B(1-\tilde{p})} \\ &\Leftrightarrow \frac{d-c}{b-a} = e^{B(1-\tilde{p})} \end{aligned}$$

[5] Let us note that h is an extension of a point-wise entropy (extension because x and y may be greater than 1) and η is a measure of the concavity (or convexity) of h.

$$a + \frac{d-a}{1+e^{B(1-\tilde{p})}} = b + \frac{c-b}{1+e^{B(1-\tilde{p})}} \Leftrightarrow \tilde{p} = 1 - \frac{1}{B}\ln\frac{d-c}{b-a}$$
$$\Leftrightarrow \tilde{p} = 1 + \frac{1}{B}\ln\frac{a-b}{c-d} \qquad (15)$$

Remark. It should be observed that, in the case where $a-b = c-d$, i.e., the case of arithmetic analogy, this approximation is exact because

$$\tilde{p} = 1 + \log 1 = 1 = p.$$

We inspect below the consequences of using the different values of B proposed above in (15).

$B = 1$. In the case where $B = 1$, the approximation of the analogical power for the analogy $a : b ::^p c : d$ is thus $\tilde{p} = 1 + \ln((a-b)/(c-d))$.

$B = 2$. Through Equation (11), Formula (10) has been shown to be equivalent to Definition 3 with a value of 2 for B. The approximation of the analogical power is thus $\tilde{p} = 1 + 1/2\ln((a-b)/(c-d))$ in this case. Appendix E shows how this value can also be computed directly from the function defined in (10).

B for equal derivatives of m_p and $\widetilde{m}_p$ at 1. Formula (14) depends solely on two terms, a and d, and not on the four terms in the analogy. This has to do with solving the equality of generalized means in the first line leading to (15). Choosing $4 \times (\eta(a,d))/(d-a)$ implies that there are no guarantees that the derivatives for $m_p(b,c)$ and $\widetilde{m}_p(b,c)$ will be equal at $p = 1$.

In the experiments reported below, we set

$$B = B(a,d) = 4 \times \frac{\eta(a,d)}{d-a}.$$

The general expression of the approximate analogical power is:

$$\tilde{p} = 1 + \frac{1}{4}\frac{d-a}{\eta(a,d)}\ln\frac{a-b}{c-d}.$$

In experiments not reported here, we also tried $B = B(b,c)$, and the arithmetic mean of $B(a,d)$ and $B(b,c)$. The results showed that taking $B = B(a,d)$ delivers a reasonably small difference on average with the exact analogical power, while taking $B = B(b,c)$ is disastrous, with an average absolute difference of 15 to 16 between the approximation and the actual value of the analogical power. Now, taking the arithmetic mean of $B(a,d)$ and $B(b,c)$ was slightly worse than the approximation with $B = 1$.

B such that the derivatives of the differences of the generalized means of the outer terms and the inner terms of the analogy are the same for the exact and approximate generalized means at 1. The equality of these derivatives of differences is:

$$\frac{\partial(m_p(a,d) - m_p(b,c))}{\partial p} = \frac{\partial(\widetilde{m}_p(a,d) - \widetilde{m}_p(b,c))}{\partial p}.$$

Appendix D shows how B can be computed from both derivatives of m_p and $\widetilde{m}_p$ with respect to p. This computation leads to the value:

$$B = 4 \times \frac{\eta(a,d) - \eta(b,c)}{(d-a) - (c-b)} \tag{16}$$

B as twice the ratio of the sum of the differences of the outer terms and of the inner terms and the sum of the terms. Finally, we also consider a value of B that is reasonably simple and fast to compute:

$$B = 2 \times \frac{(d-a) + (c-b)}{a+b+c+d}. \tag{17}$$

Notice that this form is balanced in the use of the outer terms and the inner terms. It ensures that B is positive as we require it to be. It is invariant to multiplication by any positive factor, as is the analogical power itself (for any $\lambda > 0$, $a : b ::^p c : d \implies \lambda a : \lambda b ::^p \lambda c : \lambda d$). Moreover, it increases with c and d, but decreases with a and b, following the slope of the generalized means.

Experimental results reported below show that, on average, this formula delivers a better approximation, with the least standard deviation among all values for B.

5.3. Experiments

We now report experiments to inspect under which conditions the approximated p is framed between the bounds exhibited in the previous sections. When the conditions are met, this can help accelerate programs. Because of the property of invariance of numerical analogy under multiplication, the

Table 2: Computation times for bounds and analogical powers over one million analogies using different values for B

Computation	Method	Time (s)
Bounds		
Lower bound	Theorem 5	0.13
Upper bound		0.13
Powers		
Bisection method	(Lepage and Couceiro, 2025)	1.73
Approximation	$B = 1$	0.13
	$B = 2$	0.14
	$B = 4 \times \dfrac{\eta(a,d)}{d-a}$	0.41
	$B = 4 \times \dfrac{\eta(a,d) - \eta(b,c)}{(d-a)-(c-b)}$	0.66
	$B = 2 \times \dfrac{(d-a)+(c-b)}{a+b+c+d}$	0.20

experiments were restricted to numerical values in the interval $[0; 100]$. The sampling consists of 1 million analogies. The results are listed in Table 2 for times and Tables 3 and 4 for different statistics obtained using different values of B. The best B which balances speed and quality of approximation is given by Formula (17), i.e., $B = 2 \times ((d-a)+(c-b))/(a+b+c+d)$.

Speed (Table 2). When using a constant value of B, the computation of the approximated value of the analogical power is more than 13 times as fast (0.13 s for one million analogies) as the exact computation of the value using the bisection algorithm described in (Lepage and Couceiro, 2025) (1.73 s). Formula (14) involves some computational cost. It is only 3 times as fast (0.41) as the exact computation. Formula (16) requires even more time because its formula is heavier. For Formula (17), the computational cost is less: it is almost 9 times as fast as the exact computation.

Quality of approximation (Table 3). The value computed by the approximations was equal to the value computed by the bisection algorithm in only 2% of the cases. The same percentage is observed for all approximations.

Table 3: Percentage of cases where the approximated analogical power is equal to the exact power and percentage of cases where it falls within the theoretical bounds of Section 5, obtained using different values for B

Computation of powers	Formula	Percentage of cases (%)	
		Approx. value = exact value	Approx. value within bounds
Bisection method		100	100
$B = 1$		2	96
$B = 2$		2	99
$B = 4 \times \dfrac{\eta(a,d)}{d-a}$	(14)	2	97
$B = 4 \times \dfrac{\eta(a,d) - \eta(b,c)}{(d-a)-(c-b)}$	(16)	2	98
$B = 2 \times \dfrac{(d-a)+(c-b)}{a+b+c+d}$	(17)	2	99

Table 4: Absolute difference between the approximated and exact analogical powers using different values for B

Computation of powers	Formula	Difference in absolute value	
		Average $\pm$ Std. dev.	Max. diff.
Bisection method		0.00 ± 0.00	0.00
$B = 1$		1.15 ± 2.55	44.17
$B = 2$		1.49 ± 2.86	45.35
$B = 4 \times \dfrac{\eta(a,d)}{d-a}$	(14)	0.77 ± 1.63	42.93
$B = 4 \times \dfrac{\eta(a,d) - \eta(b,c)}{(d-a)-(c-b)}$	(16)	0.57 ± 1.66	42.97
$B = 2 \times \dfrac{(d-a)+(c-b)}{a+b+c+d}$	(17)	0.37 ± 1.43	39.78

Closer inspection reveals that these cases correspond to arithmetic analogies for which, indifferently of the values of B, $\log(a - b)/(c - d) = \log 1 = 0$, which makes the approximated power correctly equal to 1.

We inquire whether the powers computed by the various approximations are framed between the bounds given in Section 5. All approximations fall massively within the bounds, but some variations are observed across the approximations. The constant value of $B = 2$ and Formulae (17) yield approximations which are almost always framed between the bounds, in 99% of the cases. For the other values of B, slightly lower percentages are obtained (96%, 97% and 98%).

Difference with the exact value (Table 4). We compute the absolute difference between the approximated value of the analogical power and its exact value computed by the bisection algorithm. When the basic approximation of the generalized mean with a sigmoid without B is used, or, equivalently, for $B = 1$, the absolute difference is 1.15 on average with a standard deviation of 2.55. Although $B = 2$ yields an increase in average and standard deviation (1.49 ± 2.86), the average is successively reduced in comparison with $B = 1$ using the values for B in Formulas (14) and (16) until the least value obtained with Formula (17): 0.77, then 0.57, down to 0.37. The standard deviation fluctuates for all formulae, with the highest values for $B = 1$ and $B = 2$ (2.55 and 2.86). The other formulae yield smaller and similar standard deviations, with the best one observed for (17): 1.43. In total, Formula (17) yields the best results for both average and standard deviation. Figure 5 shows the distributions of these absolute differences for the different values of B. The last graph clearly exhibits a higher concentration towards 0.

Use in actual implementation. To summarize the study of different values of B for approximated computation of the analogical powers, Formula (17) looks like the best compromise between execution time and quality of approximation.

6. Reconstruction of images

To confirm the reliability and the usefulness of the approximations proposed above, we run an experiment in the reconstruction of images using analogical pooling. The main purpose of this experiment is to compare the exact computation of powers with their approximate computation in a practical setting. Still, we provide a baseline to demonstrate the potential of a

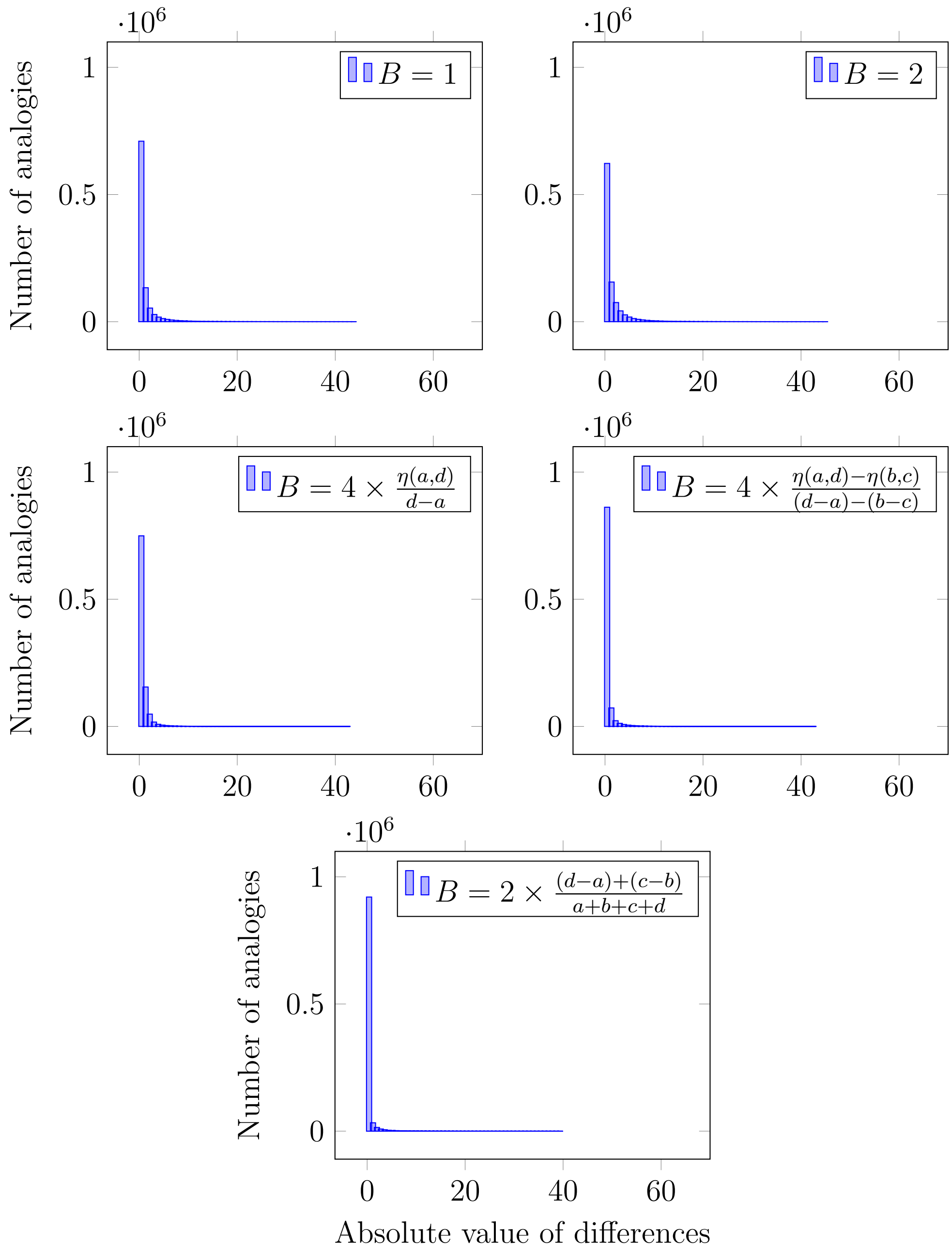


Figure 5: Distribution of the absolute values of the differences between the values computed by approximation and by the bisection algorithm, over a sample of one million analogies. The values for the factor B are indicated in the respective legends.

method based on numerical analogy in this setting. The method considers each patch of 2×2 pixels in an image as an analogy. Each patch can thus be represented by its analogical power, plus some information on the order the terms of the analogy are to be read from the patch. Consequently, each patch can be replaced by only one pixel, capturing the previous information. A square image is successively reduced from a size of $n \times n$ to $(n-1) \times (n-1)$ down to possibly only one pixel. This is the reduction step. An algorithm that reads in a reduced image of size $(n-1) \times (n-1)$, plus the analogical powers and the orderings, can construct an image of size $n \times n$ by solving analogical equations of the type $a : b ::^p: c : x$ for x along a suitable enumeration of the patches. The exact description of the method can be found in (Pilimon et al., 2025).[6]

We apply the above-described method to 10 images from the MNIST dataset (LeCun et al., 1989), i.e., images of hand-written digits of size 28×28, in two settings:

- using the exact computation of analogical powers according to the algorithm described in (Lepage and Couceiro, 2025) starting from the bounds given in Section 3;
- using the approximation proposed in Section 5.

Time. The time to reconstruct the $4 \times 10 = 40$ images shown in Appendix F was 30.62 seconds in total using exact computation of analogical powers, i.e., without approximation. Now, using approximation, this time was reduced to 29.13 seconds. This is how much time it takes to reduce the images using analogical pooling step-by-step, the number of steps indicated at the top of the images, and to reconstruct them back. In addition to the computation of analogical powers for analogical pooling, this includes scanning all patches of an image, solving analogical equations, storing/reading images, etc. Although the difference in times sounds small, it conforms to expectations[7]. The gains obtained in Section 5 are drowned in the total computation here.

[6]This method does not perform compression. Each reduction step stores more information than the $n \times n$ image contains.

[7]The reduction is $30.62 - 29.13 = 1.49$s. There are $40 \times \left(\sum_{n=1}^{27} n^2\right) = 277,200 > 1/4 \times 10^6$ computations of analogical powers in total. Table 2 suggest a reduction of $(1.73 - 0.20)/4 = 0.38$s. The rest should be attributed to other computations that also use the approximations, like the computation of generalized means.

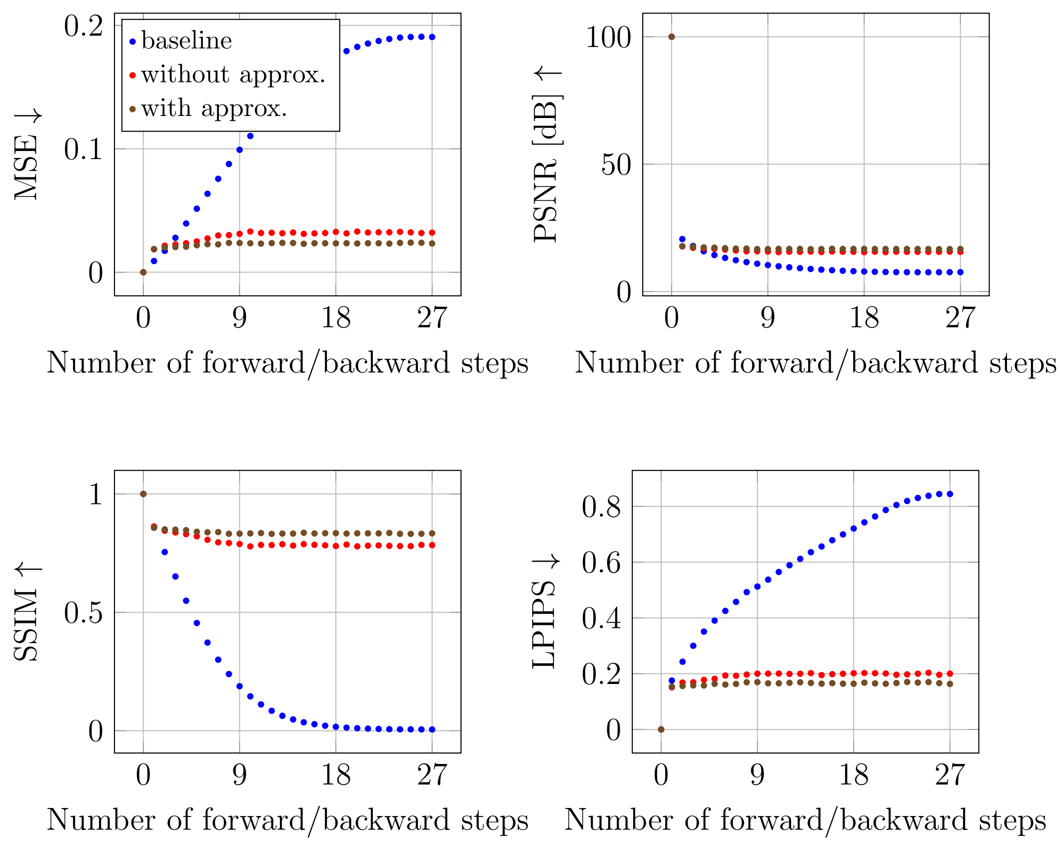


Figure 6: The values of different metrics plotted against the number of reconstruction steps of MNIST images for a method that uses numerical analogy, with or without approximation, and for a baseline that does not use numerical analogy. *Top left:* Mean Squared Error (MSE), the lower the better; *top right:* Peak Signal to Noise Ratio (PSNR), the higher the better; *bottom left:* Structural Similarity Index Measure (SSIM), the higher the better; *bottom right:* Learned Perceptual Image Patch Similarity (LPIPS), the lower the better.

Quality. One would reasonably expect that the quality of the images reconstructed using the approximation of analogical powers deteriorates in comparison with the images reconstructed using the exact computation of powers. Experimental results show the contrary. The quality of the images reconstructed using approximations is improved. This can be observed by comparing the images in Appendix F. This is confirmed using four objective evaluation metrics of standard use in image processing. These metrics compare an image to a ground truth image, in our case, the original image. The graphs in Figure 6 plot the averages for the reconstruction of a sample of 100 images, 10 per digit, from the MNIST dataset. According to all metrics, the images reconstructed using approximations are of better quality than the images reconstructed using exact computation of analogical powers. The exact reason for this phenomenon has still to be inspected, but the images in Appendix F show that the power approximation prevents an effect seen when using exact powers, that vertical bars, e.g., at the top of 5 or at the bottom of 2, tend to fade away when the number of steps increases. Images using the approximation of powers undergo some smoothing that prevents this from happening, and thus compare better with the original images.

Although the main goal of this experiment is to assess the use of approximation in comparison to exact computation, Figure 6 also provides a baseline that does not make use of numerical analogy. This baseline uses average pooling to reduce the image size and bilinear interpolation in reconstruction. Bilinear interpolation estimates the value of a reconstructed pixel by taking a weighted average of the four nearest pixels in the original image. It first interpolates in one direction (e.g., horizontally) between two neighboring pixels, and then interpolates the results in the other direction (vertically). The plots in Figure 6 demonstrate that using numerical analogy is superior to this baseline.

Nonetheless, our experimental observations about the reduction in time and the positive impact of using approximations amply justify our theoretical work.

7. Conclusion

This article presented the notion of numerical analogy again and stated some fundamental results about it. This notion, based on that of generalized means, is centered around the notion of analogical power, which characterizes an analogy between four real numbers, when it exists. A direct computation of analogical powers cannot be realized, as no directly solvable analytical formula exists for it.

However, analogical powers can be framed within bounds. We have exhibited such bounds for the analogical powers of analogies between real numbers when these powers exist and are not special cases that do not require such bounds. These bounds are used in actual program implementations for computing analogical powers. They allow a bisection algorithm to use them as initial bounds and explore the interval between them, leading to an exact computation of analogical powers (up to machine precision).

We have also proposed several practical approximations of analogical powers and determined the percentage of cases in which such approximate values are reasonable, i.e., are framed by the bounds theoretically established. With respect to the acceleration of programs, the best approximation yields a practical speed-up of up to seven times in comparison with the bisection algorithm. An experiment in image reconstruction also confirmed in practice the usefulness of the approximation of analogical powers, in terms of time and quality of results. Although the process was only applied here to MNIST images, it can be used with any real-world images, such as medical images. Investigating the potential for image classification is part of our future research, which is already underway.

The practical justification of the theoretical work presented in this paper is the acceleration of the computation of analogical powers. Speeding up is fundamental to support practical applications of numerical analogy to the different settings that we are currently exploring. One of them is in word embedding models, where we analyze the actual shape of word analogies through the inspection of analogical powers on each dimension of the word representation space. The analysis of one word analogy requires the computation of several hundred numerical analogies in the usual word embedding models. Consequently, the inspection of dozens of thousands of word analogies entails the computation of millions of analogical powers. Fast computation allows us to run experiments and compare across various embedding models using various learning methods on various corpora. Another application also deals with word embedding models. Enforcing word analogies to be faithfully represented as arithmetic analogies in each vector dimension can be achieved by designing loss functions that require frequent computation of analogical powers, since backpropagation is performed on their gradients. Since the computation must be performed for every word in the training corpus, which typically contains several million words, any practical implementation of such loss functions requires the fast computation of analogical powers. Yet another setting where fast computation of analogical powers is required is image processing, the domain we chose solely for illustrative purposes. Our example did not involve learning and used only a small number of images. However, over sets of several tens of thousands of images, with basically n^2 computations for images of size $n \times n$, the use of analogical pooling will require the computation of millions of analogical powers. Testing the usefulness of using images that undergo analogical pooling in tasks like classification using machine learning methods will obviously require fast computation.

8. Acknowledgements

This work was supported in part by a grant for a special research project from Waseda University, number 2025R-034, entitled “Numerical analogy: mathematical definitions and practical applications in machine learning”. This work was also partially supported by the French National Research Agency (ANR)[8] project “Analogies: from theory to tools and applications” (AT2TA), ANR-22-CE23-0023. This work was also supported by Portuguese national funds through FCT under project UIDB/50021/2020[9]. Additionally, the work was partially supported by the project Center for Responsible AI, reference C628696807-00454142.

The experiments reported in Section 6 and Appendix F have been conducted by Jakub Pilimon, a master’s student at Waseda University.

References

Antić, C., 2022. Analogical proportions. Annals of Mathematics and Artificial Intelligence 90, 595–644. URL: https://doi.org/10.1007/s10472-022-09798-y, doi:10.1007/s10472-022-09798-y.

Antić, C., 2024. Boolean proportions. Logical methods in computer science 20, 2–20. URL: https://doi.org/10.46298/lmcs-20(2:2)2024, doi:10.46298/lmcs-20(2:2)2024.

Bounhas, M., Prade, H., 2024. Revisiting analogical proportions and analogical inference. International Journal of Approximate Reasoning 171,

[8]Agence Nationale de la Recherche.

[9]DOI: 10.54499/UIDB/50021/2020.

109202. URL: https://www.sciencedirect.com/science/article/pii/S0888613X24000896, doi:https://doi.org/10.1016/j.ijar.2024.109202. synergies between Machine Learning and Reasoning.

Cunha, F., Lepage, Y., Bouraoui, Z., Couceiro, M., 2026. Generalizing analogical inference from Boolean to contiuous domains, in: Proceedings of the Thirty-Ninenth AAAI Conference on Artificial Intelligence, AAAI 2026 (to appear), pp. ??–??

Hofstadter, D., the Fluid Analogies Research Group, 1994. Fluid Concepts and Creative Analogies. Basic Books, New York.

Hölder, O.L., 1889. Ueber einen Mittelwerthssatz. Nachrichten von der königlichen Gesellschaft der Wissenschaft zu Göttingen und der Georg-Augusts-Universität zu Göttingen 1889, 38–47.

Klein, S., 1982. Culture, mysticism and social structure and the calculation of behavior, in: Proceedings of the European Conference on Artificial Intelligence (ECAI 1982), pp. 141–146.

LeCun, Y., Boser, B., Denker, J.S., Henderson, D., Howard, R.E., Hubbard, W., Jackel, L.D., 1989. Backpropagation applied to handwritten zip code recognition. Neural Computation 1, 541–551. doi:https://doi.org/10.1162/neco.1989.1.4.541.

Lepage, Y., Couceiro, M., 2024. Analogie et moyenne généralisée, in: Mailly, J.G., Schwarzentruber, F., Wilczynski, A. (Eds.), Actes des 18es Journées d'Intelligence Artificielle Fondamentale et des 19es Journées Francophones sur la Planification, la Décision et l'Apprentissage pour la conduite de systèmes (JIAF 2024 et JFPDA 2024), La Rochelle, France. pp. 114–124. URL: https://hal.science/AFIA/hal-04620491v1.

Lepage, Y., Couceiro, M., 2025. L'analogie numérique revisitée et étendue, précisée, rétrécie ou agrandie, in: Mailly, J.G., Schwarzentruber, F., Wilczynski, A. (Eds.), Actes des 19es Journées d'Intelligence Artificielle Fondamentale et des 20es Journées Francophones sur la Planification, la Décision et l'Apprentissage pour la conduite de systèmes (JIAF 2025 et JFPDA 2025), Dijon, France. pp. 133–142.

Malaca, F., Couceiro, M., Lepage, Y., 2026. A unified framework of numerical analogies: Definitions, conditions, and algorithms. Submitted.

Mikolov, T., Yih, W.T., Zweig, G., 2013. Linguistic regularities in continuous space word representations, in: Proceedings of the 2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT 2013), Association for Computational Linguistics, Atlanta, Georgia. pp. 746–751. URL: http://www.aclweb.org/anthology/N13-1090.

Pilimon, J., Couceiro, M., Lepage, Y., 2025. Analogical pooling for image reconstruction, in: Couceiro, M., Schahaf, D. (Eds.), Proceedings of the Workshop Interactions between analogies and machine learning, co-located with the International Joint Conference on Artificial Intelligence (IARML@IJCAI 2025), p. 12 pages. URL: https://hal.science/hal-05315980.

Prade, H., Richard, G., 2017. Analogical proportions and analogical reasoning - an introduction, in: Aha, D.W., Lieber, J. (Eds.), Case-Based Reasoning Research and Development, Springer International Publishing, Cham. pp. 16–32.

Qi, Y., Yang, Y., Lepage, Y., 2025. Towards number sequence completion through the analysis of number sequence patterns using numerical analogy, in: Couceiro, M., Schahaf, D. (Eds.), Proceedings of the Workshop Interactions between analogies and machine learning, co-located with the International Joint Conference on Artificial Intelligence (IARML@IJCAI 2025), p. 12 pages.

Trefethen, L.N., Bau, D., 1997. Numerical Linear Algebra. Society for Industrial and Applied Mathematics.

Appendix A. No central symmetry for the generalized mean function in the general case

Lemma 1. *The curve of the generalized means of two different numbers has no central symmetry.*

Proof. If there were a central point for a certain p, it would necessarily be for a value of $1/2 \times (a+d)$ because the limits of the function $m_p(a,d)$ at the infinities are a and d respectively for $-\infty$ and $+\infty$. This is the arithmetic mean, obtained for $p=1$. If there were central symmetry around this point, the symmetry would hold for the two special values of the geometric mean ($p=0$) and the quadratic mean ($p=2$), equally distant from $p=1$ on its left and on its right. The following equality would thus hold.

$$m_1(a,d) \;=\; \frac{1}{2} \times \left(m_0(a,d) + m_2(a,d)\right). \tag{A.1}$$

The following developments show that this entails the equality of a and d. We first replace the arithmetic, geometric, and quadratic means with their definitions.

$$\begin{aligned}
m_1(a,d) \;&=\; \frac{1}{2} \times \left(m_0(a,d) + m_2(a,d)\right) \\
\Leftrightarrow\; \frac{1}{2} \times (a+d) \;&=\; \frac{1}{2} \times \left(\sqrt{ad} + \sqrt{1/2 \times (a^2+d^2)}\right) \\
\Leftrightarrow\; a+d \;&=\; \sqrt{ad} + \frac{1}{\sqrt{2}}\sqrt{a^2+d^2}
\end{aligned}$$

We raise to the power of 2. This gives the following implication.

$$\begin{aligned}
m_1(a,d) \;&=\; \frac{1}{2} \times \left(m_0(a,d) + m_2(a,d)\right) \\
\Rightarrow\; (a+d)^2 \;&=\; \left(\sqrt{ad} + \frac{1}{\sqrt{2}}\sqrt{a^2+d^2}\right)^2 \\
\Rightarrow\; (a+d)^2 \;&=\; ad + \frac{1}{2}(a^2+d^2) + \frac{2}{\sqrt{2}}\sqrt{ad}\sqrt{a^2+d^2} \\
\Rightarrow\; (a+d)^2 \;&=\; \frac{1}{2}(a+d)^2 + \sqrt{2}\sqrt{ad}\sqrt{a^2+d^2} \\
\Rightarrow\; (a+d)^2 \;&=\; 2\sqrt{2}\sqrt{ad}\sqrt{a^2+d^2}
\end{aligned}$$

We raise to the power of 2 again.

$$\begin{aligned}
m_1(a,d) &= \frac{1}{2} \times (m_0(a,d) + m_2(a,d)) \\
&\Rightarrow (a+d)^4 = 8ad(a^2+d^2) \\
&\Rightarrow a^4 + 4a^3d + 6a^2d^2 + 4ad^3 + d^4 = 8a^3d + 8ad^3 \\
&\Rightarrow a^4 - 4a^3d + 6a^2d^2 - 4ad^3 + d^4 = 0 \\
&\Rightarrow (a-d)^4 = 0
\end{aligned}$$

To summarize, central symmetry implies $a = d$. In this case, all generalized means are equal to $a = d$, i.e., the curve of the function of generalized means degenerates into a horizontal straight line. Any point on the line defines a central symmetry for the curve.

By contrapositive, $a \neq d$ implies that the equality does not hold, and, hence, that there is no central symmetry for the curve of the generalized means. □

Appendix B. Stability and condition number

The purpose of this appendix is to provide a study of the numerical stability of the proposed approximation of analogical powers. For that, we use the notion of *condition number*. The condition number of a mathematical problem measures how sensitive the solution is to small changes in the input data. In our case, it measures how much p varies when the input terms, i.e., the terms of the analogy a, b, c, and d, change even slightly. For example, let $a = 1 - \epsilon$, $b = 1$, $c = 2 - \epsilon$, and $d = 2$, with ϵ a small number greater than 0. In this case, $p = 1$. If we now change c to 2, which corresponds to a very slight modification, p becomes infinite, because a and b are still different, but c and d become equal. Therefore, the generalized mean that makes an analogy is the maximum, which corresponds to an analogy in $p = +\infty$. This suggests that the problem is ill-conditioned at $(1, 1, 2, 2)$.

Let $(a, b, c, d) \in (\mathbb{R}_+^*)^4$. An analogical power p is implicitly defined by $F(a, b, c, d, p) = 0$, where

$$F(a,b,c,d,p) = \begin{cases} a^p + d^p - b^p - c^p = 0, \text{ if } p \neq 0 \\ ad - bc, \text{ otherwise,} \end{cases}$$

provided that the four terms belong to

$$D = \{(a,b,c,d) \in (\mathbb{R}_+^*)^4 : a < b \leq c < d\}.$$

Under this condition, the analogical power exists and it is unique, by Theorem 1. As F is differentiable on $D \times \mathbb{R}^*$, the Implicit Function Theorem implies that if $(a_0, b_0, c_0, d_0, p_0)$ is such that

$$F(a_0,b_0,c_0,d_0,p_0) = 0, \text{ wherever } \frac{\partial F}{\partial p}(a_0,b_0,c_0,d_0,p_0) \neq 0,$$

then p is the unique differentiable function such that

$$p(a_0,b_0,c_0,d_0) = p_0 \text{ and } F(a,b,c,d,p(a,b,c,d)) = 0,$$

in a neighborhood of (a_0, b_0, c_0, d_0). Therefore, following (Trefethen and Bau, 1997, p. 90) and using the infinity norm, the relative conditioned number is given by

$$\kappa(a,b,c,d) = \frac{||\nabla p(a,b,c,d)||_{+\infty} \; ||(a,b,c,d)||_{+\infty}}{|p(a,b,c,d)|}.$$

To facilitate the computation of the gradient, note that, for $p \neq 0$,

$$\frac{\partial F(a,b,c,d,p)}{\partial p} = a^p \ln a + d^p \ln d - b^p \ln b - c^p \ln c. \tag{B.1}$$

Also, observe that

$$\frac{\partial F(a,b,c,d,p)}{\partial a} = pa^{p-1}, \tag{B.2}$$

and similarly for the other three arguments b, c and d. Therefore, where the Implicit Function Theorem can be applied, to define the unique differentiable function $p(a,b,c,d)$,

$$\begin{aligned}\frac{\partial p(a,b,c,d)}{\partial a} &= -\frac{\dfrac{\partial F}{\partial a}(a,b,c,d,p)}{\dfrac{\partial F}{\partial p}(a,b,c,d,p)} \\ &= -\frac{pa^{p-1}}{a^p \ln a + d^p \ln d - b^p \ln b - c^p \ln c},\end{aligned} \tag{B.3}$$

and similarly for b, c and d. Let

$$L(a,b,c,d) = \frac{\partial F}{\partial p}(a,b,c,d,p(a,b,c,d)) = a^p \ln a + d^p \ln d - b^p \ln b - c^p \ln c.$$

The condition number is thus given by

$$\begin{aligned}\kappa(a,b,c,d) &= \frac{\left|\dfrac{\max\{pa^{p-1}, pb^{p-1}, c^{p-1}, pd^{p-1}\}}{L(a,b,c,d)}\right| \times \max\{a,b,c,d\}}{|p|} \\ &= \frac{\max\{a^{p-1}, b^{p-1}, c^{p-1}, d^{p-1}\}}{|L(a,b,c,d)|} \times \max\{a,b,c,d\}.\end{aligned}$$

Lemma 2 indicates the points where we cannot apply the Implicit Function Theorem.

Lemma 2. *Given* $(a,b,c,d,p) \in \left(\mathbb{R}_+^*\right)^4 \times \mathbb{R}^*$*, with* $a \leq b \leq c \leq d$*, if*

$$F(a,b,c,d,p) = 0 \quad \textit{and} \quad \frac{\partial F}{\partial p}(a,b,c,d,p) = 0$$

then $a = b$ and $c = d$.

Proof. Assume $p \neq 0$. Let $\alpha = a^p$, $\beta = b^p$, $\gamma = c^p$ and $\delta = d^p$. Then

$$\alpha + \delta = \beta + \gamma \ \wedge \ \alpha \ln \alpha + \delta \ln \delta = \beta \ln \beta + \gamma \ln \gamma.$$

Let $s = \alpha + \delta = \beta + \gamma$, $\phi(x) = x \ln x$,[10] and $g(x) = \phi(x) + \phi(s - x)$. As ϕ is a strictly convex function in its domain, ϕ' is a strictly increasing function. Therefore,

$$\begin{aligned} g'(x) > 0 &\Leftrightarrow \phi'(x) - \phi'(s-x) > 0 \\ &\Leftrightarrow \phi'(x) > \phi'(s-x) \\ &\Leftrightarrow x > s - x \\ &\Leftrightarrow x > \frac{s}{2}. \end{aligned}$$

Similarly, $g'(x) < 0$ iff $x < s/2$, and $g'(x) = 0$ iff $x = s/2$. Moreover, note that $g(x) = g(s - x)$. Therefore, if

$$g(\alpha) = g(\beta),$$

i.e., $\alpha \ln \alpha + \delta \ln \delta = \beta \ln \beta + \gamma \ln \gamma$, then either $\alpha = \beta$ or $\alpha = s - \beta = \gamma$. Hence, $a = b$ (and $d = c$) or $a = c$ (and $d = b$). As we assume $a \leq b \leq c \leq d$, the only solution (for $p \neq 0$) is $(a, c) = (b, d)$. $\square$

Following Lemma 2, we cannot apply the Implicit Function Theorem at (b, b, c, c, p). However, we know that if a is approximately b, but different, and c is approximately d, but different, the problem is ill-conditioned. For example, for $\epsilon > 0$, $p(b - \epsilon, b, c, c + \epsilon) = 1$, but $p(b - \epsilon, b, c, c) = +\infty$ and $p(b, b, c, c + \epsilon) = -\infty$.

It remains to check the points where $p \approx 0$. Consider the situation when the terms approximate the geometric analogy, i.e., d approaches bc/a and, consequently, p approaches 0.

[10] Note that $\phi(x) = x \ln x = -\eta(x)$. η has been introduced in Definition (13) in Section 5.1 and is used in Appendix C and Appendix D. It appears through differentiation with respect to the analogical power p.

$$
\begin{aligned}
\lim_{d\to bc/a} \kappa(a,b,c,d) &= \lim_{d\to bc/a} \frac{\max\{a^{p-1}, d^{p-1}, b^{p-1}, c^{p-1}\}}{|L(a,b,c,d)|} \times \max\{a,b,c,d\} \\
&= \frac{\max\{1/a, a/(bc), 1/b, 1/c\}}{\lim_{d\to bc/a} |L(a,b,c,d)|} \times \max\{a,b,c,bc/a\}
\end{aligned}
$$

Since the numerator is non-null and

$$
\begin{aligned}
\lim_{d\to bc/a} L(a,b,c,d) &= \lim_{d\to bc/a} a^p \ln a + d^p \ln d - b^p \ln b - c^p \ln c \\
&= \ln a + \lim_{d\to bc/a} \ln d - \ln b - \ln c \\
&= \ln a + \ln\left(\frac{bc}{a}\right) - \ln b - \ln c = 0,
\end{aligned}
$$

the problem is also ill-conditioned if the terms are near the geometric analogy.

Appendix C. Approximation function of generalized means with the same derivative at $p = 1$

Theorem 7. *The function*

$$\begin{aligned} B: \quad \mathbb{R}_+ \times \mathbb{R}_+ &\rightarrow \mathbb{R} \\ (a, d) &\mapsto 4 \times \frac{\eta(a, d)}{d - a} \end{aligned}$$

is a function that, when used in the formula of Definition 3, allows for approximating the generalized mean function of a and d. This approximation has the same value of derivative as the generalized mean function at $p = 1$. Note: $B(a, d)$ is not defined for $a = d$.

Proof. For the same value of the derivative at $p = 1$.
We first compute the derivative of $m_p(a, d)$ with respect to p.

$$\begin{aligned} \frac{\partial}{\partial p} m_p(a, d) &= \frac{\partial}{\partial p}\left(\left(\frac{1}{2} \times (a^p + d^p)\right)^{1/p}\right) \\ &= \frac{\partial}{\partial p}\left(\left(\frac{1}{2}\right)^{1/p} \times (a^p + d^p)^{1/p}\right) \\ &= \left(\frac{1}{2}\right)^{1/p} \frac{\partial}{\partial p}\left((a^p + d^p)^{1/p}\right) + \frac{\partial}{\partial p}\left(\left(\frac{1}{2}\right)^{1/p}\right)(a^p + d^p)^{1/p} \\ &= \left(\frac{1}{2}\right)^{1/p} \frac{\partial}{\partial p}\left((a^p + d^p)^{1/p}\right) - \frac{1}{p^2}\left(\frac{1}{2}\right)^{1/p} \times \log\frac{1}{2} \times (a^p + d^p)^{1/p} \\ &= \left(\frac{1}{2}\right)^{1/p} \times (a^p + d^p)^{1/p} \times \left(\frac{\partial}{\partial p}\left(\frac{\log(a^p + d^p)}{p}\right) - \frac{1}{p^2}\log\frac{1}{2}\right) \end{aligned}$$

Now,

$$
\begin{aligned}
\frac{\partial}{\partial p}\left(\frac{\log\left(a^p+d^p\right)}{p}\right) &= -\frac{1}{p^2}\log\left(a^p+d^p\right)+\frac{1}{p}\frac{\partial}{\partial p}\left(\log(a^p+d^p)\right)\\
&= \frac{1}{p^2}\left(-\log\left(a^p+d^p\right)+p\frac{\partial}{\partial p}\left(\log(a^p+d^p)\right)\right)\\
&= \frac{1}{p^2}\left(-\log\left(a^p+d^p\right)+p\frac{\frac{\partial}{\partial p}\left(a^p+d^p\right)}{a^p+d^p}\right)\\
&= \frac{1}{p^2}\left(-\log\left(a^p+d^p\right)+p\frac{a^p\log(a)+d^p\log(d)}{a^p+d^p}\right)\\
&= \frac{1}{p^2}\left(-\log\left(a^p+d^p\right)+\frac{a^p\log a^p+d^p\log d^p}{a^p+d^p}\right)\\
&= \frac{1}{p^2}\frac{1}{a^p+d^p}\left(a^p\log a^p+d^p\log d^p-\left(a^p+d^p\right)\log\left(a^p+d^p\right)\right)
\end{aligned}
$$

Therefore, at $p=1$,

$$
\begin{aligned}
\left.\frac{\partial}{\partial p}m_p(a,d)\right|_{p=1} &= \frac{1}{2}\times(a+d)\times\left(-\log(a+d)+\frac{a\log a+d\log d}{a+d}-\log\frac{1}{2}\right)\\
&= \frac{1}{2}\times(a+d)\times\left(-\log(a+d)+\frac{a\log a+d\log d}{a+d}-\log\frac{1}{2}\right)\\
&= \frac{1}{2}\times(a+d)\times\left(\frac{a\log a+d\log d}{a+d}-\log\left(\frac{1}{2}(a+d)\right)\right)\\
&= \frac{1}{2}\left(a\log a+d\log d\right)-\frac{1}{2}(a+d)\log\left(\frac{1}{2}(a+d)\right)\\
&= -\frac{1}{2}(a+d)\log\left(\frac{1}{2}(a+d)\right)-\frac{1}{2}\Big(-a\log a-d\log d\Big).
\end{aligned}
$$

Recall that we defined $\eta(x,y)=\mathrm{h}(m_1(x,y))-m_1(\mathrm{h}(x),\mathrm{h}(y))$ where $\mathrm{h}(x)=-x\log x$ and m_1 is the arithmetic mean. With this, we have:

$$
\left.\frac{\partial}{\partial p}m_p(a,d)\right|_{p=1}=\eta(a,d).
$$

We then compute the derivative of $\widetilde{m}_p(a,d)$ of Definition 3, with respect to p.

$$\begin{aligned}
\frac{\partial}{\partial p}\widetilde{m}_p(a,d) &= \frac{\partial}{\partial p}\left(a + \frac{d-a}{1+\exp(B(1-p))}\right) \\
&= \frac{\partial}{\partial p}\left(\frac{d-a}{1+\exp(B(1-p))}\right) \\
&= (d-a)\frac{\partial}{\partial p}\left(\frac{1}{1+\exp(B(1-p))}\right) \\
&= (d-a)\left(-\frac{\frac{\partial}{\partial p}\left(1+\exp(B(1-p))\right)}{\left(1+\exp(B(1-p))\right)^2}\right) \\
&= (d-a)\left(-\frac{\frac{\partial}{\partial p}\left(\exp(B(1-p))\right)}{\left(1+\exp(B(1-p))\right)^2}\right) \\
&= (d-a)\left(\frac{\left(\exp(B(1-p))\right)B}{\left(1+\exp(B(1-p))\right)^2}\right) \\
&= \frac{(d-a)\exp(B(1-p))B}{\left(1+\exp(B(1-p))\right)^2}
\end{aligned}$$

Therefore, at $p = 1$,

$$\begin{aligned}
\left.\frac{\partial}{\partial p}\widetilde{m}_p(a,d)\right|_{p=1} &= \frac{(d-a)B}{(1+1)^2} \\
&= \frac{(d-a)B}{4}
\end{aligned}$$

For the two derivatives above to be equal, i.e.,

$$\frac{(d-a)B}{4} = \eta(a,d),$$

B must take the value

$$4 \times \frac{\eta(a,d)}{d-a}.$$

$\square$

Appendix D. Computation of B such that the derivatives of the differences of the generalized means of the extremes and the means are the same for the exact and approximate generalized means at 1

In Appendix C it was proved that

$$\frac{\partial}{\partial p} m_p(a,d)\bigg|_{p=1} = \eta(a,d)$$

and

$$\frac{\partial}{\partial p} \widetilde{m}_p(a,d)\bigg|_{p=1} = \frac{(d-a)B}{4}.$$

Similarly, we have

$$\frac{\partial}{\partial p} m_p(b,c)\bigg|_{p=1} = \eta(b,c)$$

and

$$\frac{\partial}{\partial p} \widetilde{m}_p(b,c)\bigg|_{p=1} = \frac{(c-b)B}{4}.$$

Therefore,

$$\begin{aligned} \frac{\partial}{\partial p}(m_p(a,d) - m_p(b,c))\bigg|_{p=1} &= \frac{\partial}{\partial p}(\widetilde{m}_p(a,d) - \widetilde{m}_p(b,c))\bigg|_{p=1} \\ &\Leftrightarrow \eta(a,d) - \eta(b,c) = \frac{(d-a)B}{4} - \frac{(c-b)B}{4} \\ &\Leftrightarrow B = 4 \times \frac{\eta(a,d) - \eta(b,c)}{(d-a) - (c-b)}. \end{aligned}$$

which is Formula (16).

Appendix E. Direct derivation of $\tilde{p}$ from the balanced approximation of the generalized means

Here, we use Formula (10) to give an approximation of the analogical power of $a : b ::^p c : d$. We approximate the analogy by writing the equality of the approximation of the generalized means in a and d on the one hand and b and c on the other hand.

$$\begin{aligned}
\frac{ae^{(1-\tilde{p})} + de^{-(1-\tilde{p})}}{e^{(1-\tilde{p})} + e^{-(1-\tilde{p})}} = \frac{be^{(1-\tilde{p})} + ce^{-(1-\tilde{p})}}{e^{(1-\tilde{p})} + e^{-(1-\tilde{p})}}& \\
&\Leftrightarrow\ ae^{(1-\tilde{p})} + de^{-(1-\tilde{p})} = be^{(1-\tilde{p})} + ce^{-(1-\tilde{p})} \\
&\Leftrightarrow\ (a-b)e^{(1-\tilde{p})} = (c-d)e^{-(1-\tilde{p})} \\
&\Leftrightarrow\ \frac{a-b}{c-d} = \frac{e^{-(1-\tilde{p})}}{e^{(1-\tilde{p})}} \\
&\Leftrightarrow\ \frac{a-b}{c-d} = e^{-2(1-\tilde{p})} \\
&\Leftrightarrow\ \ln\frac{a-b}{c-d} = 2(\tilde{p}-1) \\
&\Leftrightarrow\ \tilde{p} = 1 + \frac{1}{2}\ln\frac{a-b}{c-d} \qquad \text{(E.1)}
\end{aligned}$$

Appendix F. Reconstruction of images

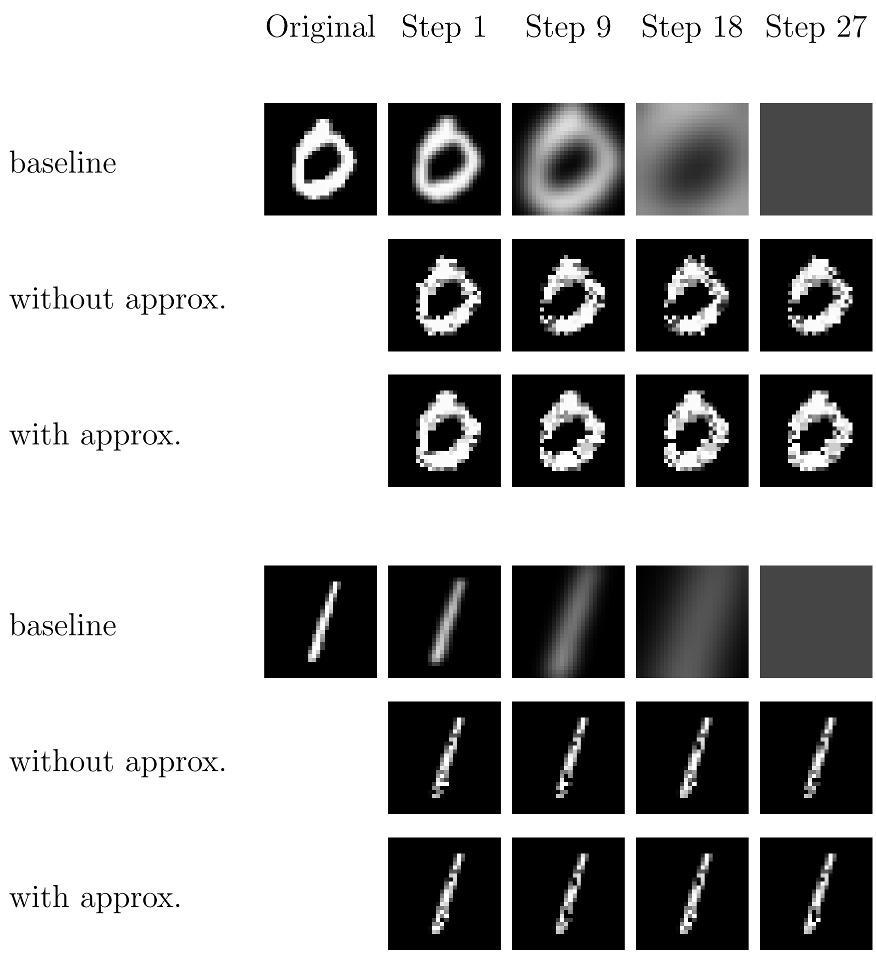


Figure F.7: Reconstruction of MNIST images from different analogical pooling steps using a baseline method, average pooling followed by bilinear interpolation on the first row, and analogical pooling followed by our proposed reconstruction method, using exact computation of analogical powers on the second row, and approximate computation of analogical powers on the third row. Top left is the original image. Digits 0 and 1.

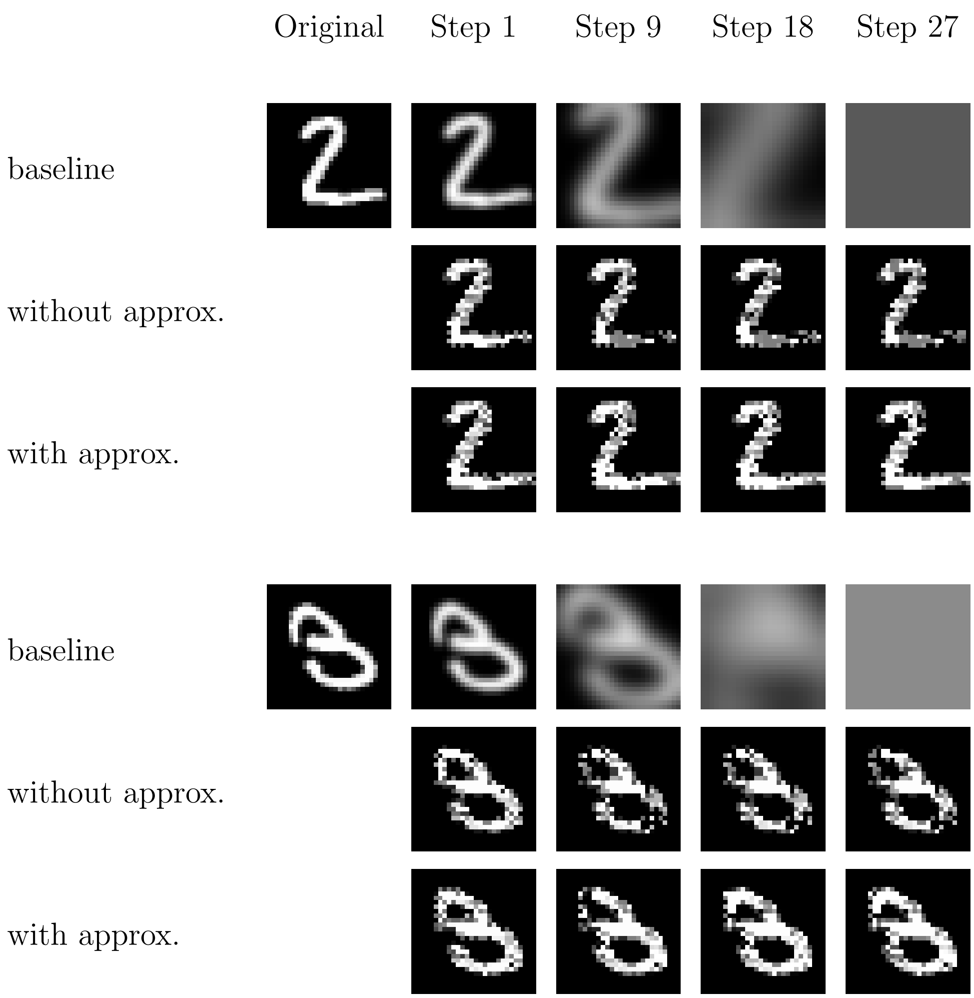


Figure F.8: Same as Figure F.7, but for digits 2 and 3.

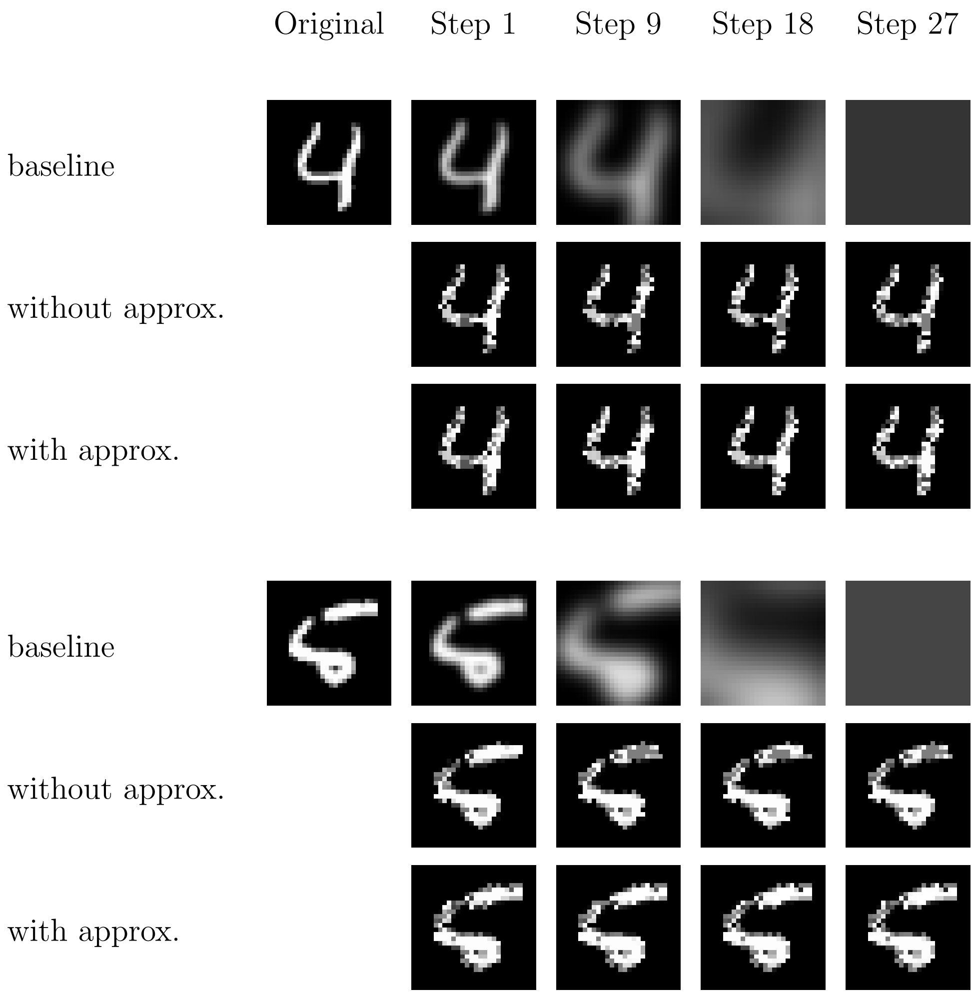


Figure F.9: Same as Figure F.7, but for digits 4 and 5.

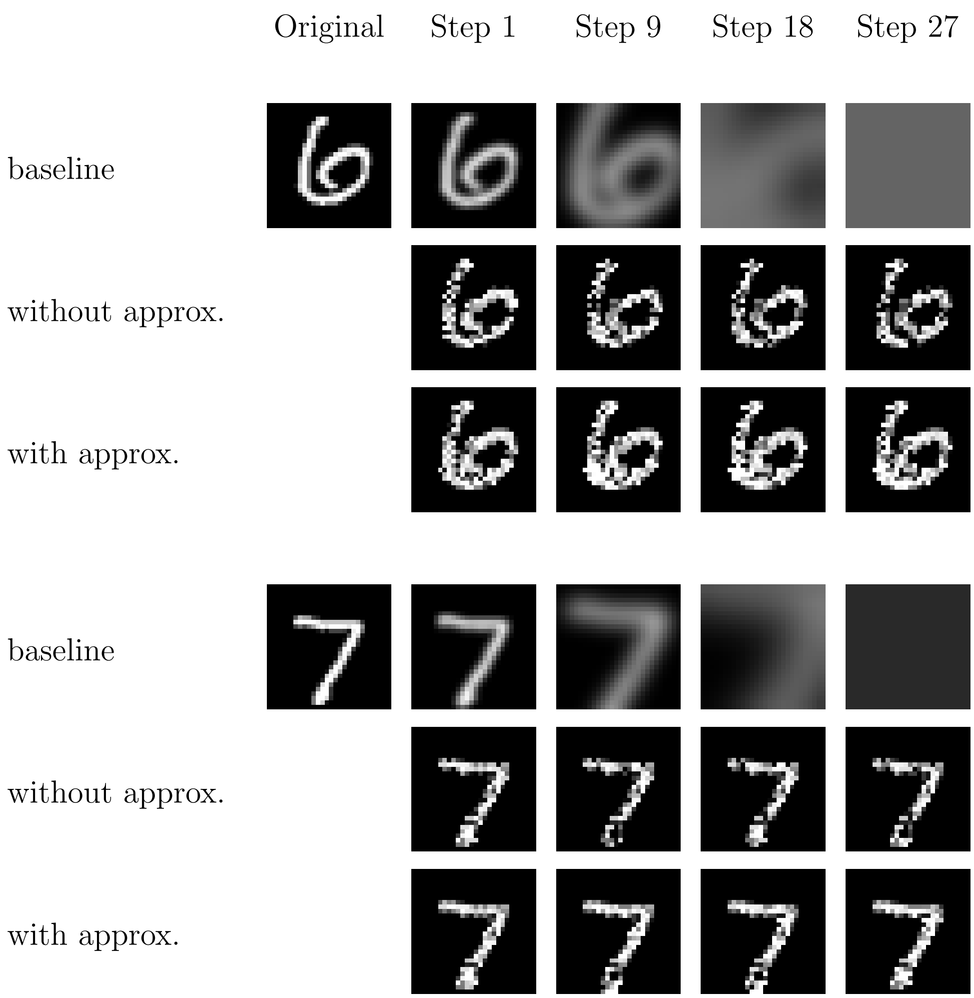


Figure F.10: Same as Figure F.7, but for digits 6 and 7.

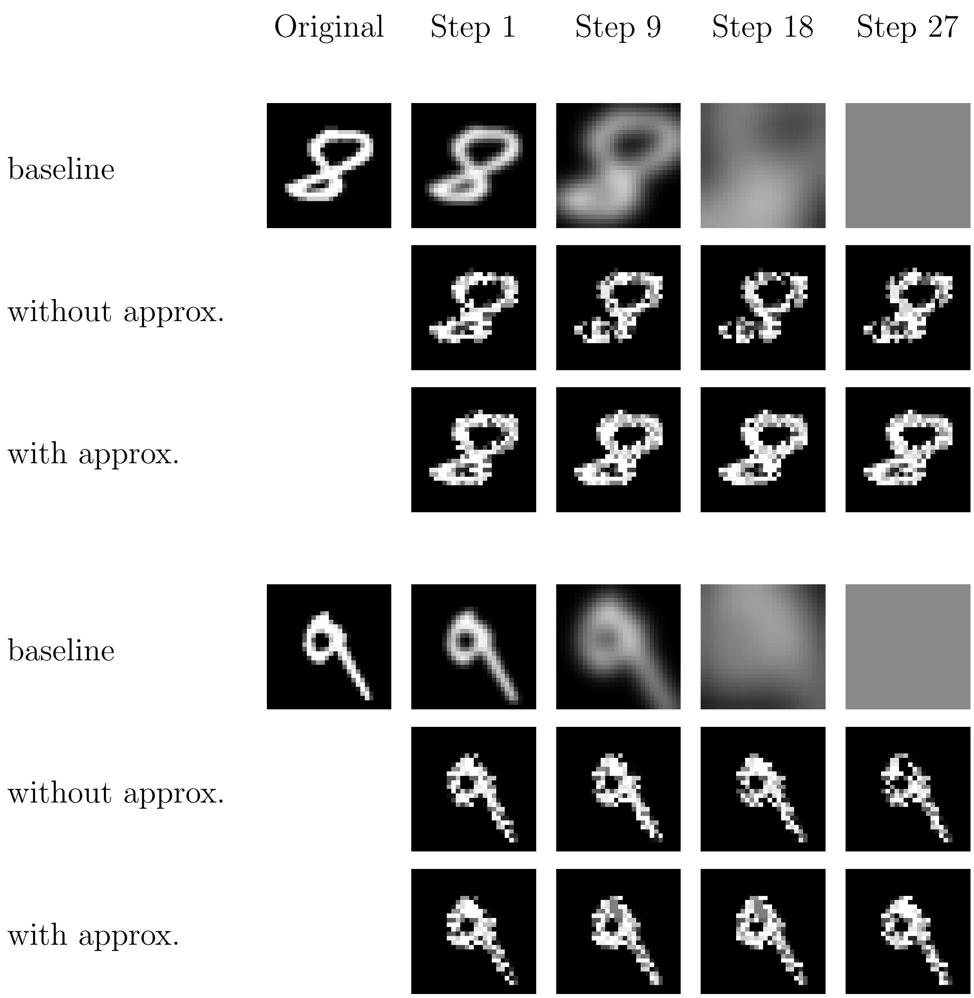


Figure F.11: Same as Figure F.7, but for digits 8 and 9.