COLLOQUIUM MATHEMATICUM

VOL. 144 2016 NO. 2

ITERATED QUASI-ARITHMETIC MEAN-TYPE MAPPINGS

BY

PAWEŁ PASTECZKA (Kraków)

Abstract. We work with a fixed N-tuple of quasi-arithmetic means $M_1, \ldots, M_N$ generated by an N-tuple of continuous monotone functions $f_1, \ldots, f_N : I \to \mathbb{R}$ (I an interval) satisfying certain regularity conditions. It is known [initially Gauss, later Gustin, Borwein, Toader, Lehmer, Schoenberg, Foster, Philips et al.] that the iterations of the mapping $I^N \ni b \mapsto (M_1(b), \ldots, M_N(b))$ tend pointwise to a mapping having values on the diagonal of I^N. Each of [all equal] coordinates of the limit is a new mean, called the *Gaussian product* of the means $M_1, \ldots, M_N$ taken on b. We effectively measure the speed of convergence to that Gaussian product by producing an effective—doubly exponential with fractional base—majorization of the error.

1. Introduction. In 1799 (1) Gauss introduced the arithmetic-geometric mean. It is obtained as the limit in the following two-term recursion:

$$x_{n+1} = \frac{x_n + y_n}{2}, \qquad y_{n+1} = \sqrt{x_n y_n}, \tag{1.1}$$

where $x_0 = x$ and $y_0 = y$ are positive parameters. Gauss [9, p. 370] proved that both $(x_n)_{n=1}^{\infty}$ and $(y_n)_{n=1}^{\infty}$ converge to a common limit, which is called the *arithmetic-geometric mean* ($\mathcal{AGM}$ for short) of x and y. In fact, this limit is a nonelementary transcendental function of x and y elegantly expressible in terms of complete elliptic integrals:

$$\frac{1}{\mathcal{AGM}(x,y)} = \frac{2}{\pi} \int_0^{\pi/2} \frac{d\theta}{\sqrt{x^2 \cos^2\theta + y^2 \sin^2\theta}}, \qquad x, y \in (0, \infty). \tag{1.2}$$

2010 *Mathematics Subject Classification*: Primary 26E60; Secondary 26A18, 26B15, 39B12.
Key words and phrases: Gaussian product, invariant means, quasi-arithmetic means, iteration, mean, mean-type mapping.
Received 16 December 2014; revised 28 August 2015.
Published online 16 March 2016.

(1) According to [11, p. 23] on May 30th, 1799 Gauss observed that (1.2) holds for $x = 1$ and $y = \sqrt{2}$ up to 11 decimal places. On December 23rd he proved (1.2) in the general case. Most probably he had worked with this definition before, but the author of the present note found no prior information about the arithmetic-geometric mean.

It is also known [2, p. 354] that

$$|x_{n+1}^2 - y_{n+1}^2| < \left(\frac{x_n^2 - y_n^2}{4\mathcal{AGM}(x,y)}\right)^2, \qquad n \in \mathbb{N}. \tag{1.3}$$

In particular, $|x_n - y_n| = |x_n^2 - y_n^2|/(x_n + y_n)$ tends to zero quadratically.

The iteration process described in (1.1) can be carried over to a much more general family of means. For a given interval I, a *mean* defined on I is any function $M\colon \bigcup_{k=1}^\infty I^k \to I$ such that

$$\min(a) \le M(a) \le \max(a) \quad \text{ for every } a \in \bigcup_{k=1}^\infty I^k.$$

The mean is *strict* if

$$\min(a) < M(a) < \max(a) \quad \text{ for every nonconstant vector } a \in \bigcup_{k=1}^\infty I^k.$$

For any twice continuously differentiable, strict means M, N and sequences

$$x_{n+1} = M(x_n, y_n), \qquad y_{n+1} = N(x_n, y_n), \qquad n \in \mathbb{N}_+ \cup \{0\},$$

the difference $|x_n - y_n|$ tends to zero quadratically [1, Theorem 8.2].

Following [1, Section 8.7], we will consider the iteration of multidimensional means. Given $N \in \mathbb{N}$ and a vector of means $M = (M_1, \dots, M_N)$ defined on a common interval I, for $a \in \bigcup_{k=1}^\infty I^k$ we consider the sequence of vectors

$$\begin{aligned} \vec{M}^0(a) &:= a, \\ \vec{M}^{n+1}(a) &:= (M_i(\vec{M}^n(a)))_{i=1}^N, \qquad n \in \mathbb{N}_+ \cup \{0\}. \end{aligned}$$

Whenever for every $i \in \{1, \dots, N\}$ the limit $\lim_{n\to\infty}[\vec{M}^n(a)]_i$ exists, and it does not depend on i, we call it the *Gaussian product* of (M_i) and denote it by $(\bigotimes_{i=1}^N M_i)(a)$. We also introduce a compact notation

$$\delta(a) := \max(a) - \min(a)$$

for every vector a of reals.

Gustin [10] proved that the Gaussian product exists for every vector M of power means. Later, J. M. Borwein and P. B. Borwein, extending some ideas of Lehmer [15], Schoenberg [21, 22] and Foster and Philips [8], proved the following (much more general) result

Proposition 1.1 ([1, Theorem 8.8]). *Let $N \in \mathbb{N}$, and $M = (M_1, \dots, M_N)$ be a vector of strict, continuous means.*

(a) *Then $\bigotimes_{i=1}^N M_i$ exists and is a strict, continuous mean.*

(b) *Suppose that the means are symmetric and continuously differentiable. Then, for $a \in \bigcup_{n=1}^{\infty} I^n$ such that $\delta(\vec{M}^n(a)) \neq 0$ for every $n \in \mathbb{N}$,*

$$\lim_{n\to\infty} \delta(\vec{M}^{n+1}(a))/\delta(\vec{M}^n(a)) = 0$$

and the convergence in the Gaussian iteration is superlinear.

(c) *If, in fact, the means are twice continuously differentiable, then the convergence of the Gaussian iteration is quadratic* (*uniformly on compact subsets*).

If the means are twice continuously differentiable, part (c) implies that $\delta(\vec{M}^n(a))$ converges to zero quadratically. Matkowski [16] proved that part (a) remains valid if one of the means is not strict (the Gaussian product does not have to be strict in this case).

We are going to present some effective estimation of $\delta(\vec{M}^n(a))$ in the case when the vector M consists of quasi-arithmetic means satisfying a certain smoothness condition.

2. Quasi-arithmetic means. We will be interested in the special case of strict, continuous means, called quasi-arithmetic means. The idea of these means only glimpsed as a natural generalization of power means in the pioneering paper by Knopp [12]. Shortly thereafter it was materialized in three nearly simultaneous papers in the early 1930s [13, 18, 6]. These means have been extensively dealt with ever since their introduction (cf. e.g. [3, Chap. 4]). Many results concerning power means have counterparts for this family (frequently under some additional assumption).

For an interval I and a continuous, strictly monotone function $f\colon I \to \mathbb{R}$ we define a mean $A_{[f]}\colon \bigcup_{k=1}^{\infty} I^k \to I$ by the equality

$$A_{[f]}(a) := f^{-1}\bigg(\frac{f(a_1) + \cdots + f(a_k)}{k}\bigg), \quad a \in I^k,\ k \in \mathbb{N}.$$

All these *quasi-arithmetic means* are strict and continuous, whence their Gaussian products always exist.

In our setting we will fix $N \in \mathbb{N}$, an interval I, and a vector $\mathbf{f} = (f_1, \ldots, f_N)$ of continuous, strictly monotone functions $f_j\colon I \to \mathbb{R}$, $j \in \{1, \ldots, N\}$. It will lead us to a vector of means $(A_{[f_j]})_{j=1}^N$ and their Gaussian product. The notion of quasi-arithmetic means could also be understood vectorially as a function $\vec{A}_{[\mathbf{f}]}\colon \bigcup_{k=1}^{\infty} I^k \to I^N$, as in the introduction.

Our main Theorem 3.3 is a strengthening of Proposition 1.1(c). Unlike Proposition 1.1 this theorem presents some effective estimation of $\delta(\vec{A}_{[\mathbf{f}]}^n(a))$ when the pertinent means are quasi-arithmetic means generated by functions satisfying certain smoothness conditions (belonging to the family $\mathcal{S}(I)$ defined later).

Lemmas 4.3 and 4.4 that appear in the course of estimating the difference between quasi-arithmetic means continues the 1960s papers [4, 5] and our previous results [19, 20].

2.1. Operator P. Now we turn to the result of Mikusiński [17] (and independently Łojasiewicz [17, footnote 2]), who found a handy tool to compare quasi-arithmetic means in terms of the operator $P_f := f''/f'$. More precisely, their result reads

PROPOSITION 2.1 (Basic comparison). *Let I be an interval and $f, g \in \mathcal{C}^2(I)$ with $f' \cdot g' \neq 0$ on I. Then the following conditions are equivalent:*

(i) $A_{[f]}(a) \geq A_{[g]}(a)$ *for all* $a \in \bigcup_{k=1}^{\infty} I^k$, *with equality only when a is a constant vector,*
(ii) $P_f > P_g$ *on a dense subset of I,*
(iii) $(\operatorname{sgn} f') \cdot (f \circ g^{-1})$ *is strictly convex.*

The operator P is central to our discussion; we will assume that the functions considered are smooth enough for it to be applied.

Moreover, we assume the second derivative of relevant functions is of *almost bounded variation* (i.e. of finite variation when restricted to every compact interval; cf. [14, p. 135]). Using this notion we introduce the class

$$\mathcal{S}(I) := \{f \in \mathcal{C}^2(I) \colon f' \neq 0 \text{ and } f'' \text{ is of almost bounded variation}\}.$$

Very often we will use a global estimation of f''/f', so for $K > 0$ we denote

$$\mathcal{S}_K(I) := \{f \in \mathcal{S}(I) \colon \|f''/f'\|_\infty \leq K\}.$$

This is closely connected with the family of log-exp means (cf. [3, p. 269]) defined for every $a \in \mathbb{R}^k$, $k \in \mathbb{N}_+$, as

$$\mathcal{E}_p(a) := \begin{cases} \dfrac{1}{p} \ln\left(\dfrac{e^{p \cdot a_1} + \cdots + e^{p \cdot a_k}}{k}\right) & \text{if } p \neq 0, \\ \dfrac{1}{k}(a_1 + \cdots + a_k) & \text{if } p = 0. \end{cases}$$

Indeed, a slightly weaker version of Proposition 2.1 is that

$$\mathcal{S}_K(I) = \{f \in \mathcal{S}(I) \colon \mathcal{E}_{-K} \leq A_{[f]} \leq \mathcal{E}_K\}. \tag{2.1}$$

3. Main result. We are going to present a certain majorization of the speed of convergence. The theorem below depends on a free parameter l. There is no universal (optimal) value of l that could be plugged into this theorem; the most natural restriction will be given after the statement.

Throughout, I is an interval, $k \in \mathbb{N}$, and $a \in I^k$ is a nonconstant vector.

Remark 3.1. It is important to note that the result depends *only* on the input vector a (more precisely $\delta(a)$) and the number K. In particular the number of functions as well as the functions themselves are not essential.

Theorem 3.2. *Let $K \in (0,\infty)$ and $\mathbf{f}$ be a vector of functions from $\mathcal{S}_K(I)$. Let $\alpha = (3+7e)/3$ ($\alpha \approx 7.34$). Then*

$$\max \vec{A}^n_{[\mathbf{f}]}(a) - \min \vec{A}^n_{[\mathbf{f}]}(a) < \frac{1}{\alpha K}(\alpha l)^{2^{n-n_0}}$$

for every $l \in (0,1)$ and

$$n \geq \left\lceil \log_2\left(\frac{\exp(K(\max a - \min a)) - 1}{e^l - 1}\right)\right\rceil =: n_0.$$

Taking l minimizing the right hand side of the first inequality is the most natural choice. However, if we wish to decrease n_0, we may change the value of l.

Minimization of the right hand side is realized for $l \approx 0.05$ ($l = \xi$ in the setting below). Thus, we obtain the following

Theorem 3.3. *Let $K \in (0,\infty)$ and $\mathbf{f}$ be a vector of functions from $\mathcal{S}_K(I)$. Let $\alpha = (3+7e)/3$ and let μ be the minimum value of the function*

$$(0,1) \ni l \mapsto (\alpha l)^{(e^l-1)/2}$$

achieved for $l = \xi$ ($\alpha \approx 7.34$, $\mu \approx 0.97$, $\xi \approx 0.05$). Then

$$\max \vec{A}^n_{[\mathbf{f}]}(a) - \min \vec{A}^n_{[\mathbf{f}]}(a) < \frac{1}{\alpha K} \cdot \mu^{\frac{2^n}{\exp(K(\max a - \min a))-1}}$$

for every

$$n \geq \log_2(e) \cdot K \cdot (\max a - \min a) - \log_2(e^{\xi} - 1) + 1 =: n_1$$

($n_1 \approx 1.443 \cdot K \cdot (\max a - \min a) + 5.25$).

The proofs of these theorems are postponed until Sections 4.4 and 4.5, respectively. They will be preceded by a number of technical lemmas (in Sections 4.1–4.3).

3.1. Possible reformulation. In both theorems above we can restrict the interval I to $[\min a, \max a]$ and the relevant functions to $\mathcal{S}([\min a, \max a])$ (by taking K best possible).

More precisely, we can change the order of assumptions in the following way: First, take an interval I, a natural number k, and a k-tuple $\mathbf{f} = (f_1, \dots, f_k)$ with $f_i \in \mathcal{S}(I)$ for every $i \in \{1, \dots, k\}$. Then we *define*

$$K := \sup_{\substack{x \in [\min a, \max a] \\ i \in \{1,\dots,k\}}} |P_{f_i}(x)|.$$

Such a reformulation is natural but (i) we need to calculate K, which could be difficult and (ii) Remark 3.1 is no longer applicable. However we will apply this procedure in Section 5.

4. Proofs of Theorems 3.2 and 3.3

4.1. Assumption $K = 1$. For fixed $K > 0$ and an interval I we define an operator ${}^*\colon \mathcal{S}_K(I) \to \mathcal{S}_1(K \cdot I)$ by $f^*(x) := f(x/K)$. Then

$$P_{f^*}(x) = \frac{1}{K} P_f(x/K), \qquad x \in K \cdot I,\ f \in \mathcal{S}_K(I).$$

Moreover, for every continuous, monotone function $f \in \mathcal{S}_K(I)$,

$$A_{[f]}(a) = \frac{1}{K} A_{[f^*]}(K \cdot a).$$

Hence, if $\mathbf{f} = (f_1, \dots, f_N)$ with $f_i \in \mathcal{S}_K(I)$ for $i = 1, \dots, N$ and $\mathbf{f}^* := (f_1^*, \dots, f_N^*)$, then

$$\vec{A}_{[\mathbf{f}]}(a) = \frac{1}{K} \vec{A}_{[\mathbf{f}^*]}(K \cdot a).$$

Thus, iterating, we get

$$\vec{A}^n_{[\mathbf{f}]}(a) = \frac{1}{K} \vec{A}^n_{[\mathbf{f}^*]}(K \cdot a), \qquad n \in \mathbb{N}.$$

Using these properties we easily obtain

$$\begin{aligned}
\max a - \min a &= \frac{1}{K}(\max Ka - \min Ka),\\
\max A^n_{[\mathbf{f}]}(a) - \min A^n_{[\mathbf{f}]}(a) &= \frac{1}{K}\big(\max \mathbf{A}^n_{[\mathbf{f}^*]}(K \cdot a) - \min \mathbf{A}^n_{[\mathbf{f}^*]}(K \cdot a)\big).
\end{aligned}$$

Hence, from now on, we will assume in many proofs (or even formulations) that $K = 1$.

4.2. Approximate value of quasi-arithmetic means. The aim of the present section is to establish an approximate value of the quasi-arithmetic mean generated by a function belonging to the family $\mathcal{S}_1(I)$. Let us introduce a compact notation for the arithmetic mean and variation of a vector:

$$\begin{aligned}
\bar{a} &:= \frac{1}{k}(a_1 + \dots + a_k), && a \in \mathbb{R}^k,\ k \in \mathbb{N},\\
\operatorname{Var}(a) &:= \frac{1}{k}\sum_{i=1}^{k}(a_i - \bar{a})^2 = \frac{1}{k}\sum_{i=1}^{k} a_i^2 - \bar{a}^2, && a \in \mathbb{R}^k,\ k \in \mathbb{N}.
\end{aligned}$$

Note that functions belonging to $\mathcal{S}(I)$ are (only) twice differentiable. However in the lemmas below it would be handy to use third derivatives. To avoid this drawback, we turn to Riemann–Stieltjes integrals (see Lemma 4.1 below).

Remark. This is just one of possible solutions—otherwise we may consider functions from $\mathcal{C}^\infty(I) \cap \mathcal{S}(I)$ only, and use a density-type argument to extend Theorems 3.2 and 3.3 to the whole space $\mathcal{S}(I)$.

Lemma 4.1. *For every* $f \in \mathcal{S}(I)$,

$$A_{[f]}(a) = \bar{a} + \frac{1}{2}\operatorname{Var}(a)\cdot P_f(\bar{a}) + \frac{1}{2k\cdot f'(\bar{a})}\sum_{i=1}^{k}\int_{\bar{a}}^{a_i}(a_i - t)^2\,df''(t) + \int_{\bar{a}}^{A_{[f]}(a)} \frac{\big(f(u) - f(A_{[f]}(a))\big)f''(u)}{f'(u)^2}\,du.$$

Proof. By Taylor's theorem applied to f at $\bar{a}$ and to f^{-1} at $f(\bar{a})$ (with integral remainder in both cases; cf. [7, equation 2.4]) we obtain

$$f(x) = f(\bar{a}) + (x-\bar{a})f'(\bar{a}) + (x-\bar{a})^2\frac{f''(\bar{a})}{2} + \int_{\bar{a}}^{x}\tfrac{1}{2}(x-t)^2\,df''(t), \tag{4.1}$$

$$f^{-1}(f(\bar{a})+\delta) = \bar{a} + \frac{\delta}{f'(\bar{a})} + \int_{f(\bar{a})}^{f(\bar{a})+\delta}\frac{\big(t-(f(\bar{a})+\delta)\big)f''(f^{-1}(t))}{f'(f^{-1}(t))^3}\,dt. \tag{4.2}$$

By (4.1), for every $i \in \{1,\dots,k\}$,

$$f(a_i) = f(\bar{a}) + (a_i-\bar{a})f'(\bar{a}) + (a_i-\bar{a})^2\frac{f''(\bar{a})}{2} + \int_{\bar{a}}^{a_i}\tfrac{1}{2}(a_i-t)^2\,df''(t). \tag{4.3}$$

But it can be easily verified that $\sum_{i=1}^{k}(a_i-\bar{a}) = 0$. So, by (4.3),

$$\begin{aligned}\frac{1}{k}\sum_{i=1}^{k}f(a_i) &= f(\bar{a}) + \frac{1}{k}\sum_{i=1}^{k}\left((a_i-\bar{a})^2\frac{f''(\bar{a})}{2} + \int_{\bar{a}}^{a_i}\tfrac{1}{2}(a_i-t)^2\,df''(t)\right)\\ &= f(\bar{a}) + \operatorname{Var}(a)\cdot\frac{f''(\bar{a})}{2} + \frac{1}{k}\sum_{i=1}^{k}\int_{\bar{a}}^{a_i}\tfrac{1}{2}(a_i-t)^2\,df''(t).\end{aligned}$$

Thus

$$\begin{aligned}\delta &:= \frac{1}{k}\sum_{i=1}^{k}f(a_i) - f(\bar{a})\\ &= \operatorname{Var}(a)\cdot\frac{f''(\bar{a})}{2} + \frac{1}{k}\sum_{i=1}^{k}\int_{\bar{a}}^{a_i}\tfrac{1}{2}(a_i-t)^2\,df''(t).\end{aligned} \tag{4.4}$$

Combining (4.2) and (4.4) one gets

$$\begin{aligned}
A_{[f]}(a) &= f^{-1}\Big(\frac{1}{k}\sum_{i=1}^{k} f(a_i)\Big)\\
&= f^{-1}(f(\bar a)+\delta)\\
&= \bar a + \mathrm{Var}(a)\cdot\frac{f''(\bar a)}{2f'(\bar a)} + \frac{1}{k\cdot f'(\bar a)}\sum_{i=1}^{k}\int_{\bar a}^{a_i}\tfrac{1}{2}(a_i-t)^2\,df''(t)\\
&\quad + \int_{f(\bar a)}^{\frac{1}{k}\sum_{i=1}^{k} f(a_i)} \frac{(t-\frac{1}{k}\sum_{i=1}^{k} f(a_i))f''(f^{-1}(t))}{f'(f^{-1}(t))^3}\,dt.
\end{aligned}$$

Lastly, upon setting $t=f(u)$ in the last integral, one obtains $dt=f'(u)du$ and

$$\begin{aligned}
A_{[f]}(a) &= \bar a + \mathrm{Var}(a)\cdot\frac{f''(\bar a)}{2f'(\bar a)} + \frac{1}{2k\cdot f'(\bar a)}\sum_{i=1}^{k}\int_{\bar a}^{a_i}(a_i-t)^2\,df''(t)\\
&\quad + \int_{\bar a}^{A_{[f]}(a)} \frac{(f(u)-f(A_{[f]}(a)))f''(u)}{f'(u)^2}\,du. \ \blacksquare
\end{aligned}$$

Now, we majorize the last two terms in Lemma 4.1.

LEMMA 4.2. *For every $f\in\mathcal{S}_1(I)$,*

(i) $$\Big|\int_{\bar a}^{A_{[f]}(a)} \frac{\big(f(u)-f(A_{[f]}(a))\big)f''(u)}{f'(u)^2}\,du\Big| < (A_{[f]}(a)-\bar a)^2\cdot\exp(\|P_f\|_*),$$

(ii) $$\Big|\frac{1}{2k\cdot f'(\bar a)}\sum_{i=1}^{k}\int_{\bar a}^{a_i}(a_i-t)^2\,df''(t)\Big| \le \frac{1}{6k}\cdot\exp(\|P_f\|_*)\cdot\sum_{i=1}^{k}|a_i-\bar a|^3,$$

where $\|P_f\|_* := \sup_{a,b\in I}|\int_a^b P_f(t)\,dt|$.

Proof. Note that

$$\frac{f'(\Omega)}{f'(\Theta)} = \exp\Big(\int_{\Theta}^{\Omega} P_f(u)\,du\Big) \le \exp(\|P_f\|_*) \quad \text{for all } \Omega,\Theta\in I. \tag{4.5}$$

(i) Applying the mean value theorem we obtain

$$\begin{aligned}
&\Big|\int_{\bar a}^{A_{[f]}(a)} \frac{\big(f(u)-f(A_{[f]}(a))\big)f''(u)}{f'(u)^2}\,du\Big|\\
&\quad = |A_{[f]}(a)-\bar a|\cdot\Big|\frac{f(\Theta)-f(A_{[f]}(a))}{f'(\Theta)}\Big|\cdot\Big|\frac{f''(\Theta)}{f'(\Theta)}\Big| \text{ for some } \Theta\in(\bar a, A_{[f]}(a))
\end{aligned}$$

$$= |A_{[f]}(a) - \bar{a}| \cdot |\Theta - A_{[f]}(a)| \cdot \left|\frac{f'(\Omega)}{f'(\Theta)}\right| \cdot |P_f(\Theta)| \text{ for some } \Omega \in (\bar{a}, A_{[f]}(a))$$
$$\le (A_{[f]}(a) - \bar{a})^2 \left|\frac{f'(\Omega)}{f'(\Theta)}\right|$$
$$\le (A_{[f]}(a) - \bar{a})^2 \exp(\|P_f\|_*) \quad \text{by (4.5).}$$

(ii) By the mean value theorem, for every $i \in \{1, \ldots, k\}$ there exists $\beta_i \in (\min(\bar{a}, a_i), \max(\bar{a}, a_i))$ satisfying

$$\int_{\bar{a}}^{a_i} (a_i - t)^2 \, df''(t) = f''(\beta_i) \int_{\bar{a}}^{a_i} (a_i - t)^2 \, dx = \frac{f''(\beta_i)}{3} (a_i - \bar{a})^3.$$

Applying the mean value theorem again, we find that there exists a universal value $\beta \in (\min a, \max a)$ satisfying

$$\sum_{i=1}^{k} \int_{\bar{a}}^{a_i} (a_i - t)^2 \, df''(t) = \frac{f''(\beta)}{3} \sum_{i=1}^{k} (a_i - \bar{a})^3.$$

Combining this equality above with (4.5), we get

$$\begin{aligned}\left|\frac{1}{2k \cdot f'(\bar{a})} \sum_{i=1}^{k} \int_{\bar{a}}^{a_i} (a_i - t)^2 \, df''(t)\right| &= \left|\frac{f''(\beta)}{6k \cdot f'(\bar{a})} \sum_{i=1}^{k} (a_i - \bar{a})^3\right| \\ &\le \frac{1}{6k} \left|\frac{f''(\beta)}{f'(\bar{a})}\right| \sum_{i=1}^{k} |a_i - \bar{a}|^3 \\ &= \frac{1}{6k} \left|\frac{f''(\beta)}{f'(\beta)}\right| \left|\frac{f'(\beta)}{f'(\bar{a})}\right| \sum_{i=1}^{k} |a_i - \bar{a}|^3 \\ &\le \frac{1}{6k} \cdot \exp(\|P_f\|_*) \cdot \sum_{i=1}^{k} |a_i - \bar{a}|^3. \quad \blacksquare\end{aligned}$$

Combining Lemmas 4.1 and 4.2, we obtain

Corollary 4.1. *For every $f \in \mathcal{S}_1(I)$,*

$$|A_{[f]}(a) - \bar{a} - \tfrac{1}{2} \operatorname{Var}(a) P_f(\bar{a})|$$
$$\le \exp(\|P_f\|_*) \cdot \left((A_{[f]}(a) - \bar{a})^2 + \frac{1}{6k} \sum_{i=1}^{k} |a_i - \bar{a}|^3\right).$$

Therefore, the value of a quasi-arithmetic mean can be approximated in the following informal way:

$$A_{[f]}(a) \approx \bar{a} + \tfrac{1}{2} \operatorname{Var}(a) P_f(\bar{a}).$$

Such an expression could be predicted much earlier, after Proposition 2.1. The only things to calculate were $\frac{1}{2} \operatorname{Var}(a)$ and the majorization of error, which was the most difficult part.

4.3. Estimation of the difference between quasi-arithmetic means. Recall that our aim is to reduce the results to the family $\mathcal{S}_1(I)$. Moreover, in the setting of Gaussian product, we deal with the limit

$$\lim_{n\to\infty} \delta(\mathbf{A}^n_{[\mathbf{f}]}(a)) = 0.$$

If $\delta(\mathbf{A}^n_{[\mathbf{f}]}(a))$ is small enough, we will majorize the difference between $A_{[f_i]}(\mathbf{A}^n_{[\mathbf{f}]}(a))$ and $\bar{a}$ for $i = 1, \ldots, N$ (Lemma 4.3). Otherwise we will use property (2.1) to decrease $\delta(\mathbf{A}^n_{[\mathbf{f}]}(a))$ (Lemma 4.4). The following lemma specifies Proposition 1.1(c) in the case of quasi-arithmetic means.

LEMMA 4.3. *For every $f \in \mathcal{S}_1(I)$,*

$$|A_{[f]}(a) - \bar{a}| < \frac{\alpha}{2} \cdot \delta(a)^2, \quad \textit{where} \quad \alpha = \frac{3+7e}{3}.$$

Proof. If $\delta(a) \geq 1$ we simply get

$$|A_{[f]}(a) - \bar{a}| \leq \delta(a) \leq \alpha/2 \cdot \delta(a)^2.$$

From now on we will assume $\delta(a) < 1$. Let $J = [\min a, \max a]$. We restrict f to the interval J, set $g := f|_J$ and consider the mean $A_{[g]}$. Then $g \in \mathcal{S}_1(J)$ (we consider one-sided derivatives at the endpoints of J), and since $A_{[g]}(a) = A_{[f]}(a)$ we have

$$|A_{[f]}(a) - \bar{a}| = |A_{[g]}(a) - \bar{a}|.$$

Therefore we can apply Corollary 4.1 to g instead of f. By definition,

$$\operatorname{Var}(a) = \frac{1}{k}\sum_{i=1}^{k}(a_i - \bar{a})^2 < \delta(a)^2.$$

Similarly

$$\begin{aligned}
&(A_{[g]}(a) - \bar{a})^2 < \delta(a)^2,\\
&\frac{1}{6k}\sum_{i=1}^{k}|a_i - \bar{a}|^3 < \tfrac{1}{6}\delta(a)^3,\\
&\|P_g\|_* \leq \delta(a) \cdot \|P_g\|_\infty \leq \delta(a).
\end{aligned}$$

Thus, by Corollary 4.1, we obtain

$$\begin{aligned}
|A_{[f]}(a) - \bar{a}| &= |A_{[g]}(a) - \bar{a}|\\
&< \tfrac{1}{2}\operatorname{Var}(a)|P_g(\bar{a})| + \exp(\|P_g\|_*) \cdot \bigg((A_{[g]}(a) - \bar{a})^2 + \frac{1}{6k}\sum_{i=1}^{k}|a_i - \bar{a}|^3\bigg)\\
&\leq \tfrac{1}{2}\delta(a)^2 + e^{\delta(a)}\big(\delta(a)^2 + \tfrac{1}{6}\delta(a)^3\big) = \big(\tfrac{1}{2} + e^{\delta(a)}\big(1 + \tfrac{1}{6}\delta(a)\big)\big)\delta(a)^2.
\end{aligned}$$

Lastly, as $\delta(a) < 1$,

$$|A_{[f]}(a) - \bar{a}| < \frac{3+7e}{6} \cdot \delta(a)^2 = \frac{\alpha}{2} \cdot \delta(a)^2. \blacksquare$$

To majorize the difference in the more general case, in view of (2.1), let us prove the following (seemingly isolated) result.

LEMMA 4.4.
$$\exp(\mathcal{E}_1(a) - \mathcal{E}_{-1}(a)) - 1 \le \tfrac{1}{2}\big(\exp(\delta(a)) - 1\big).$$

Proof. Assume without loss of generality that $a_1 \le \cdots \le a_k$. Then, by simple transformations,

$$\begin{aligned}
\mathcal{E}_1(a) - \mathcal{E}_{-1}(a) &= \ln\left(\frac{\sum_{i=1}^k \exp(a_i)}{k}\,\frac{\sum_{j=1}^k \exp(-a_j)}{k}\right),\\
e^{\mathcal{E}_1(a)-\mathcal{E}_{-1}(a)} &= \frac{1}{k^2}\sum_{i=1}^k e^{a_i}\sum_{j=1}^k e^{-a_j} = \frac{1}{k^2}\sum_{i=1}^k\sum_{j=1}^k e^{a_i-a_j},\\
e^{\mathcal{E}_1(a)-\mathcal{E}_{-1}(a))} - 1 &= \frac{1}{k^2}\sum_{i=1}^k\sum_{j=1}^k (e^{a_i-a_j}-1).
\end{aligned}$$

The right hand side above will not decrease if we omit the terms $e^{a_i-a_j}-1$ for $i \le j$ (they are nonpositive). For $i > j$ we can use a trivial estimation $a_i - a_j \le \delta(a)$. Thus we obtain

$$\begin{aligned}
e^{\mathcal{E}_1(a)-\mathcal{E}_{-1}(a)} - 1 &\le \frac{1}{k^2}\sum_{i>j}(e^{a_i-a_j}-1) \le \frac{k(k-1)}{2k^2}(e^{\delta(a)}-1)\\
&\le \tfrac{1}{2}(e^{\delta(a)}-1). \blacksquare
\end{aligned}$$

4.4. Proof of Theorem 3.2. By the argument announced in Section 4.1, let $K = 1$ and $f_i \in \mathcal{S}_1(I)$ for every $i \in \{1, \dots, k\}$. By (2.1) and Lemma 4.4,

$$\exp\big(\delta(\vec{A}_{[\mathbf{f}]}(v))\big) - 1 \le \exp(\mathcal{E}_1(v) - \mathcal{E}_{-1}(v)) - 1 \le \tfrac{1}{2}(e^{\delta(v)} - 1), \quad v \in \bigcup_{k=1}^{\infty} I^k.$$

So, by simple induction, using definition of n_0, one has

$$\begin{aligned}
\exp\big(\delta(\vec{A}^{n_0}_{[\mathbf{f}]}(a))\big) - 1 &\le \frac{1}{2^{n_0}}(e^{\delta(a)} - 1)\\
&\le \frac{1}{2^{\log_2\left(\frac{\exp(\max a - \min a)-1}{e^l-1}\right)}}(e^{\delta(a)} - 1)\\
&= \frac{1}{\frac{\exp(\max a - \min a)-1}{e^l-1}}(e^{\delta(a)} - 1) = e^l - 1.
\end{aligned}$$

Hence

$$\delta(\vec{A}^{n_0}_{[\mathbf{f}]}(a)) < l. \tag{4.6}$$

Therefore the assertion of the theorem is satisfied for $n = n_0$. Moreover, the mapping $n \mapsto \delta(\vec{A}^n_{[\mathbf{f}]}(a))$ is decreasing. Thus for $n \ge n_0$, by Lemma 4.3,

$$\begin{aligned}\delta(\vec{A}^{n+1}_{[\mathbf{f}]}(a)) &\le |\max \vec{A}^{n+1}_{[\mathbf{f}]}(a) - \bar{a}| + |\bar{a} - \min \vec{A}^{n+1}_{[\mathbf{f}]}(a)| \\ &\le \alpha \cdot \delta(\vec{A}^n_{[\mathbf{f}]}(a))^2.\end{aligned} \tag{4.7}$$

By simple induction, inequalities (4.6) and (4.7) imply

$$\delta(\vec{A}^n_{[\mathbf{f}]}(a)) < \frac{1}{\alpha}(\alpha l)^{2^{n-n_0}} \qquad \text{for every } n \ge n_0.$$

4.5. Proof of Theorem 3.3. Assume $K = 1$. Moreover, let $l = \xi$ and

$$n^* := \log_2\biggl(\frac{\exp(\max a - \min a) - 1}{e^\xi - 1}\biggr).$$

Applying Theorem 3.2, we have $n_0 = \lceil n^* \rceil$. Thus

$$\begin{aligned} n_0 &< n^* + 1 \\ &= \log_2\bigl(\exp(\max a - \min a) - 1\bigr) - \log_2(e^\xi - 1) + 1 \\ &< \log_2 e \cdot (\max a - \min a) - \log_2(e^\xi - 1) + 1 =: n_1. \end{aligned}$$

Therefore Theorem 3.2 is applicable for $n \ge n_1$ (note that in this case $n \ge n_0$ too). Since $\alpha\xi \in (0, 1)$, this theorem implies

$$\begin{aligned}\delta(\vec{A}^n_{[\mathbf{f}]}(a)) &< \frac{1}{\alpha}(\alpha\xi)^{2^{n-n_0}} < \frac{1}{\alpha}(\alpha\xi)^{2^{n-n^*-1}} = \frac{1}{\alpha} \cdot \bigl((\alpha\xi)^{(e^\xi-1)/2}\bigr)^{\frac{2^n}{\exp(\max a - \min a)-1}} \\ &= \frac{1}{\alpha} \cdot \mu^{\frac{2^n}{\exp(\max a - \min a)-1}}.\end{aligned}$$

5. Arithmetic-geometric mean revisited. In this section we keep the notation (1.1). To obtain the arithmetic-geometric mean we take $I = (0, \infty)$,

$$f_1 \colon I \ni x \mapsto x, \quad f_2 \colon I \ni x \mapsto \ln(x),$$

and their product $\mathbf{f} = (f_1, f_2)$. Then $\vec{A}_{[\mathbf{f}]}(x, y) = \bigl(\frac{1}{2}(x+y), \sqrt{xy}\bigr)$. Using this self-mapping we simply identify $\mathcal{AGM}$.

Fix $x_0, y_0 \in \mathbb{R}_+$, $x_0 > y_0$, $a_0 = (x_0, y_0)$. We will be interested in estimating $\delta(\vec{A}^n_{[\mathbf{f}]}(a))$. We have already mentioned inequality (1.3). To visualise Theorem 3.3, we will apply it in the spirit of Section 3.1.

We have $P_{f_1}(x) = 0$ and $P_{f_2}(x) = -1/x$. Therefore

$$K := \sup_{\substack{x \in [y_0, x_0] \\ i \in \{1,2\}}} |P_{f_i}(x)| = \frac{1}{y_0}.$$

In order to apply our Theorem 3.3, we calculate

$$\begin{aligned} n_1 &= \log_2 e \cdot K \cdot (\max a - \min a) - \log_2(e^{\xi} - 1) + 1 \\ &= \log_2 e \cdot \frac{1}{y_0} \cdot (x_0 - y_0) - \log_2(e^{\xi} - 1) + 1 \\ &= \log_2 e \cdot \frac{x_0}{y_0} - \log_2 e - \log_2(e^{\xi} - 1) + 1 \\ &\approx 1.44 \cdot \frac{x_0}{y_0} + 3.80. \end{aligned}$$

Thus, for $n > n_1$, one has

$$\delta(\vec{A}^n_{[\mathbf{f}]}(a)) < \frac{y_0}{\alpha} \mu^{\overline{\frac{2^n}{\exp(\frac{1}{y_0}(x_0 - y_0)) - 1}}} = \frac{y_0}{\alpha} \mu^{\overline{\frac{2^n}{\exp(\frac{x_0}{y_0} - 1) - 1}}}.$$

Remark. The inequality above remains valid (with the same value of n_1) if $\mathbf{f}$ is any family of power means with parameters between 0 and 2. In particular it holds for a number of classical means: arithmetic-quadratic, quadratic-geometric, arithmetic-geometric-quadratic etc.

Acknowledgements. I am grateful to Professor Janusz Matkowski and Professor Piotr Mormul for their valuable remarks and help in the process of writing down this paper.

REFERENCES

[1] J. M. Borwein and P. B. Borwein, *Pi and the AGM: A Study in Analytic Number Theory and Computational Complexity*, Wiley, New York, 1987.

[2] J. M. Borwein and P. B. Borwein, *The arithmetic-geometric mean and fast computation of elementary functions*, SIAM Rev. 26 (1984), 351–366.

[3] P. S. Bullen, *Handbook of Means and Their Inequalities*, Math. Appl. 560, Kluwer, Dordrecht, 2003.

[4] G. T. Cargo and O. Shisha, *On comparable means*, Pacific J. Math. 4 (1964), 1053–1058.

[5] G. T. Cargo and O. Shisha, *A metric space connected with generalized means*, J. Approx. Theory 2 (1969), 207–222.

[6] B. de Finetti, *Sul concetto di media*, Giornale dell'Istituto Italiano degli Attuari 2 (1931), 369–396.

[7] S. S. Dragomir and S. Abelman, *Approximating the Riemann–Stieltjes integral of smooth integrands and of bounded variation integrators*, J. Inequal. Appl. 154 (2013), 16 pp.

[8] D. M. E. Foster and G. M. Phillips, *The arithmetic-harmonic mean*, Math. Comput. 42 (1984), 183–191.

[9] C. F. Gauss, *Nachlass (Aritmetisch-geometrisches Mittel)*, 1818, in: *Werke* 3, Königlichen Gesell. Wiss., Göttingen, 1876, 357–402.

[10] W. Gustin, *Gaussian means*, Amer. Math. Monthly 54 (1947), 332–335.

[11] F. Klein, *Gauß' wissenschaftliches Tagebuch 1796–1814*, Math. Ann. 57 (1903), 1–34.

[12] K. Knopp, *Über Reihen mit positiven Gliedern*, J. London Math. Soc. 3 (1928), 205–211.
[13] A. Kolmogoroff, *Sur la notion de la moyenne*, Rend. Accad. Lincei 12 (1930), 388–391.
[14] M. Kuczma, B. Choczewski and R. Ger, *Iterative Functional Equations*, Encyclopedia Math. Appl. 32, Cambridge Univ. Press, Cambridge, 1990.
[15] D. H. Lehmer, *On the compounding of certain means*, J. Math. Anal. Appl. 36 (1971), 183–200.
[16] J. Matkowski, *Iterations of mean-type mappings and invariant means*, Ann. Math. Sil. 13 (1999) 211–226.
[17] J. G. Mikusiński, *Sur les moyennes de la forme* $\psi^{-1}[\sum q\psi(x)]$, Studia Math. 10 (1948), 90–96.
[18] M. Nagumo, *Über eine Klasse der Mittelwerte*, Japan. J. Math. 7 (1930), 71–79.
[19] P. Pasteczka, *When is a family of generalized means a scale?*, Real Anal. Exchange 38 (2013), 193–210.
[20] P. Pasteczka, *A new estimate of the difference among quasi-arithmetic means*, Math. Inequal. Appl. 18 (2015), 1321–1327.
[21] I. J. Schoenberg, *On the arithmetic-geometric mean*, Delta 7 (1977), 49–65.
[22] I. J. Schoenberg, *Mathematical Time Exposures*, MAA, Washington, DC, 1982.

Paweł Pasteczka
Institute of Mathematics
Pedagogical University of Cracow
Podchorążych 2
30-084 Kraków, Poland
E-mail: ppasteczka@up.krakow.pl